DECISION-MAKING MODEL FOR ENERGY EFFICIENT TECHNOLOGIES IN GREEN BUILDINGS

M P Wasudha Prabodhani Abeyrathna

198136T

Doctor of Philosophy

Department of Civil Engineering

Faculty of Engineering

University of Moratuwa

Sri Lanka

July 2023

DECISION-MAKING MODEL FOR ENERGY EFFICIENT TECHNOLOGIES IN GREEN BUILDINGS

M P Wasudha Prabodhani Abeyrathna

198136T

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree

Doctor of Philosophy

Department of Civil Engineering

Faculty of Engineering

University of Moratuwa

Sri Lanka

July 2023

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: *UOM Verified Signature* Date: 08/09/2023

The above candidate has carried out research for the PhD/MPhil/Masters thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. Rangika Umesh Halwatura

Signature of the Supervisor: Date: 08. 09 2023

Name of Supervisor: Prof. Arturas Kaklauskas

Signature of the Supervisor: *UOM Verified Signature* Date: 08/09/2023

Name of Supervisor: Dr. Amal Shehan Perera

Signature of the Supervisor: *UOM Verified Signature* Date: 08/09/23

Name of Supervisor: Dr. Fatima Rizna Arooz

Signature of the Supervisor: *UOM Verified Signature* Date: 08/09/2023

DEDICATION

This thesis represents not only the culmination of academic achievements but also the culmination of the love, support, and belief that surrounded me throughout this transformative journey.

This dissertation is dedicated to the love and support given by my loving parents, my sister and her husband, and my loving mother-in-law.

To my loving husband, you deserve equal credit for this accomplishment, and I dedicate this thesis to you with heartfelt gratitude and love. This achievement is not solely mine, but a shared victory that we celebrate together.

With heartfelt appreciation, this dissertation is dedicated to Mother Sri Lanka - my source of knowledge, inspiration, and boundless opportunities that generously granted me free education. As I present this dissertation, I do so with profound gratitude for the doors Sri Lanka's education system has opened for me. This dedication is a tribute to the countless educators, administrators, and policymakers who work tirelessly to make education accessible to all.

To the relentless researchers, the tenacious souls who refuse to give up, the researchers who keep pushing forward, undeterred by failure and setbacks, I dedicate this thesis to you as a tribute to your incredible spirit and celebration of your remarkable accomplishments and the champions of perseverance.

ACKNOWLEDGEMENT

I am deeply grateful for the support and guidance of numerous individuals whose unwavering contributions made this research journey possible, ultimately leading to the successful completion of my PhD thesis.

Foremost, I extend my heartfelt appreciation to Prof. R.U. Halwatura, my research supervisor, whose mentorship and unwavering support were the cornerstone of my research success. Additionally, I wish to express my sincere gratitude to Prof. Arturas Kaklauskas, Dr. Amal Shehan Perera and Dr. Rizna Arooz, my esteemed co-supervisors, whose invaluable guidance and encouragement played a crucial role in achieving my research objectives. As a team, we surmounted obstacles and accomplished our research goals.

I am also indebted to Dr. Surangika Ranatunga and Dr. Nimal Wijerathna of the progress review committee and fellow researchers of the Pro Green Laboratory, University of Moratuwa. Their constructive feedback, unwavering support, and encouragement were instrumental in yielding fruitful research outcomes.

My heartfelt gratitude goes to the support from the European Union Erasmus+ program for Capacity Building in the field of Higher Education for financing the project activities and for the great exposure they have offered me throughout the years. Also, I would like to acknowledge the Department of Construction Management and Real Estate academic staff of Vilnius Gediminas Technical University, Lithuania, for hosting and sponsoring me to study as one of their research students.

I would like to extend my sincere appreciation to the management and staff of the factories and office buildings that graciously granted me permission to conduct surveys and collect crucial building information, especially during the challenging COVID-19 pandemic. Their willingness to cooperate and support my research, despite the unprecedented circumstances, was truly commendable. My sincere appreciation also goes to the participants of the surveys, whose enthusiastic engagement and valuable input were integral to the success of this research. None of these research outcomes would have been possible without their active involvement.

Furthermore, I extend my special thanks to the Construction Management Division of the Department of Civil Engineering, University of Moratuwa, Ms. Priyantha, Ms. Rukma, and Mr. Naveen for their constant and unfaltering support throughout this journey.

I would like to express my deepest gratitude to my parents, husband, sister and her husband, and my mother-in-law for their unwavering belief in me and their endless encouragement. Their steadfast support at every step has been an essential source of strength and motivation in making this journey a resounding success. I want to especially acknowledge Ms. Amavi Silva, who was a beacon of light when I was in a research block and also Mr. Udayantha Giragamage, for going above and beyond in his support and assistance with my research.

To everyone who played a role in this research endeavor, your contributions have not gone unnoticed and have left an indelible mark on my academic and personal growth. Thank you all for being an integral part of this remarkable achievement.

ABSTRACT

Employee satisfaction is paramount as it directly impacts their productivity and health, particularly in the office environment, where thermal comfort plays a crucial role. Existing quantitative methods for evaluating thermal comfort satisfaction solely focus on building structural elements. To bridge this gap, a study was conducted, surveying 1091 staff members across 14 green office buildings to assess their satisfaction with indoor environmental quality (IEQ) comfort. The analysis introduced a proposed network of IEQ comfort features to aid in designing the questionnaire and measuring the environment. To address the issue of an imbalanced dataset, the study implemented various resampling methods along with feature selection techniques that integrated statistical analysis methods and machine learning algorithms. Developing predictive models using the Random Forest algorithm allowed for a comparison with Decision Tree, Lasso Regression and Support Vector Regression models. Three predictive models were created to assess thermal comfort, visual comfort and indoor air quality comfort separately, and one predictive model was created to assess the overall IEQ comfort. The study identified significant factors influencing IEQ comfort satisfaction, the share of the area served by AC, total window area, the thickness of the wall insulation, area served by lighting, and smart controlling. The predictive models achieved more than 75% accuracy, and interpretability supports their practical application in office design. By utilising this predictive model, building designers and managers can make informed decisions, uncovering situations where green building certifications may not meet employees' expected level of thermal comfort. Ultimately, optimising employee thermal comfort can lead to enhanced productivity.

Keywords: employee satisfaction evaluation, green office buildings, IEQ comfort, predictive modelling, random forest regression

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
AC	Air Condition
ASHRAE	American Society of Heating, Refrigerating, And Air-Conditioning Engineers
BCA	Building and Construction Authority
BD & FM	Building Design and Facilities Management
BRE	Building Research Establishment
BREAM	Building Research Establishment Environmental Assessment Method
CDF	Cumulative Distribution Function
CFL	Compact Fluorescent Lamps
CO_2	Carbon Dioxide
DGNB	Deutsche Gesellschaft Für Nachhaltiges Bauen (German Sustainable Building Council)
DT	Decision Tree
EDGE	Excellence In Design for Greater Efficiencies
Emp	Employee
ESE	Employee Satisfaction Survey
GBCA	Green Building Council of Australia
GBCSL	Green Building Council of Sri Lanka
GBDM	Green Building Decision Making Model
G-SEED	Green Standard for Energy and Environmental Design
H_0	Null Hypothesis
H_1	Alternative Hypothesis
HVAC	Heating, Ventilation, and Air Conditioning
IAQ	Indoor Air Quality
IEQ	Indoor Environmental Quality
IFC	International Finance Corporation
IQR	Interquartile Range
ISO	International Organization for Standardization
IWBI	International Well Building Institute
K-S	Kolmogorov-Smirnov
KW test	Kruskal-Wallis Test
LED	Light-Emitting Diodes
LEED	Leadership In Energy and Environmental Design
MAE	Mean Absolute Error
NDA	Non Disclosure Agreement
NGBS	National Green Building Standard

NZEB	Net-Zero Energy Building
PM	Particulate Matter
PPM	Parts Per Million
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QQ	Quantile-Quantile
r	Correlation Coefficient
RF	Random Forest
RMSE	Root Mean Squared Error
SHGC	Solar Heat Gain Coefficient
SVM	Support Vector Mechanism
SVR	Support Vector Regression
S-W	Shapiro-Wilk
U test	Mann-Whitney U Test
USA	United States of America
USGBC	United State Green Building Council
USGBC	U.S. Green Building Council
VAV	Variable Air Volume
VIF	Variance Inflation Factor
VOCs	Volatile Organic Compounds
VT	Visible Transmittance
WWR	Window-to-Wall Ratio

LIST OF APPENDICES

1 INTRODUCTION

Fast-paced and technology-driven world today, office buildings serve as the central hub for numerous activities and interactions, accommodating a significant portion of people's daily lives. These work environments significantly impact individuals, as they spend much time in office buildings (Verbeke & Audenaert, 2018). Therefore, ensuring comfortable and healthy indoor environments in these office spaces has become increasingly important (McArthur & Powell, 2020). One crucial aspect of occupant comfort within office buildings is indoor environmental quality (IEQ) satisfaction, particularly in tropical regions where maintaining optimal parameters presents unique challenges (Ahmed et al., 2021).

Tropical green office buildings, characterised by sustainable design and environmentally friendly practices, have gained prominence recently. These buildings incorporate efficient energy systems, passive cooling techniques, and natural ventilation to create comfortable and sustainable indoor environments (Awadh, 2017). In this study, the focus is specifically on tropical green office buildings due to the presence of proactive measures taken to improve indoor environmental quality comfort. Green buildings adhere to sustainability standards and incorporate design elements prioritising occupant well-being (Grzegorzewska & Kirschke, 2021). It is focused on leveraging the existing measures and standards to reduce the impact of unknown variables that may influence employee satisfaction in conventional office buildings by selecting green buildings as the study context. The choice to focus on green buildings is driven by the growing importance of sustainability and energy efficiency in the built environment. However, despite their ecological focus, the indoor environmental quality comfort in tropical climates often falls below desirable levels compared to other climatic conditions (Kiki et al., 2020), and it could lead to reduced productivity, increased absenteeism, and overall dissatisfaction. Optimising the indoor environmental quality of office buildings is crucial to ensuring the well-being and productivity of employees (Abeyrathna et al., 2023). The level of indoor environmental quality comfort that employees experience at work directly impacts their well-being and productivity (Bueno et al., 2021). Green office buildings, designed to be energy-efficient a sustainable, are increasingly popular due to their environmentally-friendly design. However, while these buildings offer numerous benefits, their indoor air quality comfort can be challenging to maintain (Ahmed et al., 2021). This study aims to capture the various applications used in green buildings to improve indoor environmental quality comfort and analyse how they affect employee satisfaction.

Comprehending the elements that affect workers' indoor environmental quality comfort in building structures is crucial to optimise their construction and use (Bueno et al., 2021). This article thoroughly investigates employee indoor environmental quality comfort in green office buildings, identifying building factors that significantly impact satisfaction. In order to precisely measure and forecast employee indoor environmental quality comfort based on these parameters, the study investigated different machine learning modelling aspects. Understanding the factors influencing employees' indoor environmental quality satisfaction in green buildings will enable the refinement and enhance existing strategies, making them more effective in creating comfortable indoor environments. Moreover, the findings of this research can serve as valuable insights for both green building professionals and practitioners in the broader field of office building design and management.

The main objective of this research is to evaluate the accuracy of the designed predictive model in predicting employee indoor environmental quality comfort and to identify the most crucial building envelope parameters for predicting employee indoor environmental quality comfort. The predictive model can optimise future or existing building envelop based on occupancy patterns and green applications to maintain a comfortable indoor environment while reducing energy consumption. Building managers can contribute to the sustainability of the building and enhance the well-being and productivity of the building occupants by promoting a comfortable indoor environment (Dinmohammadi et al., 2023).

One promising approach is to use machine learning algorithms to develop a predictive model for employee indoor environmental quality comfort (Sibyan et al., 2022). A data-driven prediction model is proposed using machine learning techniques to address the research objectives. Specifically, four widely recognised regression models are employed: Random Forest (RF), Support Vector Machine (SVM), Lasso regression (LR) and Decision Tree (DT). These models have been chosen based on their proven success in various predictive modelling tasks, ability to handle complex and nonlinear relationships, and capability to capture the interactions between different variables. The predictive model was developed using RF and compared with LR, SVM and DT to analyse the accuracy. Survey data from employees in 14 green office buildings in tropical climates were collected to determine their satisfaction levels with the indoor environmental quality comfort of their workplace. The building parameters were measured and compiled using the standard equipment of these buildings.

Random Forest (RF) is an ensemble learning algorithm that combines multiple decision trees to create a robust and accurate prediction model (Dou et al., 2019). RF's capability to handle large datasets, high dimensionality, and non-linear relationships (Predictive modelling for solar indoor environmental quality energy systems, Overview and comparative study of dimensionality reduction techniques for high dimensional data) makes it an ideal choice. RF can identify the most influential variables, providing insights into the key factors that affect employees' indoor environmental quality satisfaction by leveraging the existing measures in green buildings. Lasso regression (LR), on the other hand, is a regularization technique that helps in feature selection and preventing overfitting (Patil & Kim, 2020). It is particularly useful when dealing with high-dimensional datasets. Lasso regression penalizes the absolute values of regression coefficients, effectively shrinking some coefficients to zero and providing a sparse model with only the most relevant features.

In the context of indoor environmental quality satisfaction (Miller et al., 2022), Lasso regression can help identify the most critical factors while excluding less influential ones, thus simplifying the model and enhancing interpretability (Dumitrescu et al., 2022). Support Vector Machine (SVM) is another robust algorithm that constructs a hyperplane to predict the continuous value of indoor environmental quality satisfaction (Peng & Hsieh, 2017). Its effectiveness in handling high-dimensional data and complex patterns makes it a valuable addition to comparing with the RF modelling approach. By mapping input features into a higher-dimensional space, SVM identifies decision boundaries that accurately separate the data points, enabling precise predictions of indoor environmental quality satisfaction (Aryal & Gerber, 2020) Decision Tree (DT), a simple yet effective algorithm, learns decision rules from the data to create a hierarchical structure of if-else conditions (Song & Lu, 2015). Decision trees are easily interpretable and offer insights into the factors affecting indoor environmental quality satisfaction. Although prone to overfitting, ensemble methods like RF can mitigate this limitation (Tang et al., 2022) . Applying DT alongside RF and SVM allowed a comprehensive understanding of different modelling techniques' predictive power and interpretability.

The study depicts the leverage of the proactive measures and standards already in place to reduce the impact of unknown variables. The focus on green buildings aligns with the growing importance of sustainability and energy efficiency, offering an opportunity to refine and optimise existing strategies. This research aims to advance green building design and the broader field of office building management, ultimately fostering comfortable and sustainable workspaces by employing RF predictive model.

1.1 Aims and objectives

This investigation aims to build a decision-making model to evaluate employee Indoor Environmental Quality satisfaction with the energy-efficient applications of green buildings.

The objectives of this research:

1. Conduct a comprehensive literature review to investigate the decision-making-related tools and techniques for the green building industry
2. To identify the green technology awareness, reputation and popularity over time and to track professional’s perception of energy-efficient technology applications
3. To evaluate the opinions over the satisfaction of applying energy efficient tools to green buildings
4. To develop a green building decision-making (GBDM) model

Investigating the factors that influence energy efficiency improves the quality of life for individuals and communities. Energy-efficient buildings offer better indoor air quality, thermal comfort, and lighting conditions, resulting in healthier and more productive living and working environments. Understanding these factors helps policymakers and building professionals design and implement energy.

1.2 Research gap

Several challenges have been identified in the existing approaches in the context of assessing occupant satisfaction with indoor environmental quality (IEQ) parameters. One major concern is the lack of statistical validation of current measurement tools used to evaluate occupant satisfaction (Tekce et al., 2020). The reliability and accuracy of data collected through these tools are consequently questioned, hindering the ability to draw meaningful conclusions.

Another significant issue lies in the absence of consensus among researchers and professionals regarding the specific parameters to consider when assessing occupant satisfaction with IEQ (Cheung et al., 2021). This lack of standardization leads to inconsistency in research findings and makes it difficult to compare results across different studies.

Additionally, occupant comfort requirements can vary significantly, posing a challenge in establishing a universally applicable approach to assessing satisfaction. Different individuals may have diverse comfort preferences and needs, making it crucial to understand these variations to create tailored and occupant-centric building designs (Rohde et al., 2019).

Current studies often focus on evaluating individual physical factors, such as lighting, thermal conditions, or indoor air quality, in isolation. This approach fails to capture the complex interactions between these parameters that significantly influence occupant comfort. A more comprehensive investigation that considers multiple physical factors together is needed to provide a holistic understanding of how the indoor environment affects occupants (Zhang et al., 2022).

The focus on energy efficiency in green buildings has garnered significant attention in the pursuit of sustainable construction and operation. While energy efficiency is undoubtedly a critical aspect, this research seeks to shed light on the potential consequences of overemphasizing energy-related applications. Specifically, this study aims to demonstrate that energy-efficient measures not only impact the energy aspects of green buildings but also exert considerable influence on Indoor Environmental Quality (IEQ) parameters. These, in turn, play a vital role in the comfort and well-being of employees within these buildings.

(a) Energy Efficiency as a Key Element:

Energy efficiency is a cornerstone of green building design and operation, often cited for its potential to reduce the environmental footprint. It is regarded as crucial due to its economic and ecological implications associated with energy consumption (Paramati et al., 2022). Energy-efficient technologies and practices hold the promise of substantial energy savings, making buildings more sustainable and cost-effective in the long term. However, this emphasis on energy-related applications has raised concerns about potential imbalances in green building priorities (Uddin et al., 2021).

(b) The Predominance of Energy:

One of the identified problems is the overemphasis on energy-related applications to the detriment of other sustainability aspects. In some instances, energy efficiency is mistakenly treated as the predominant, if not the sole, element to be considered in green building design and operation (Uddin et al., 2021). This narrow focus has led to a neglect of other critical factors, including Indoor Environmental Quality, which encompasses aspects such as air quality, thermal comfort, lighting, and acoustics.

(c) The Economic and Environmental Significance:

The strong emphasis on energy is not without merit, given its economic and environmental significance. The economic costs and environmental impacts associated with energy use in buildings are substantial (Dräger & Letmathe, 2022). Energy-efficient measures hold the promise of reducing these costs and mitigating environmental harm. Furthermore, the potential energy savings resulting from widespread deployment of energy efficiency measures in buildings are projected to be considerably higher than in other sectors, such as transportation and industry (Neofytou et al., 2020).

Furthermore, building performance metrics that specifically prioritize occupant comfort with multiple environmental variables are limited (Parkinson et al., 2023). While energy efficiency and sustainability are widely addressed, a comprehensive approach that equally considers occupant well-being is essential for balanced and sustainable building design (Zhang & Tu, 2021).

An important aspect often overlooked is the real-world experiences and perspectives of building occupants themselves. Assessing occupant satisfaction should not rely solely on the views of professionals involved in building design and construction. Incorporating the insights of those who directly experience the indoor environment is vital to gaining practical and user-centric evaluation (Khoury, 2019).

Green buildings in Sri Lanka have shown better performance in terms of Indoor Environmental Quality (IEQ) parameters when compared to conventional buildings (Elnaklah et al., 2021). As green buildings become more prevalent in the country, expectation of more empirical evidence to emerge in the future (Khoshbakht, Gou, Lu, et al., 2018). This expectation is driven by the increasing popularity of green buildings in Sri Lanka and the growing interest in post-occupancy studies. More in-depth studies to compare the actual performance of green buildings and conventional buildings in terms of indoor environmental quality is a timely need (Deegahawature & Rupasinghe, 2019).

The adoption of sustainable building practices, including green buildings, can have a significant positive impact on the environment and public health in developing countries like Sri Lanka. As awareness and resources increase in these regions, more attention is likely to be directed towards understanding the benefits of green buildings in contrast to conventional ones.

Addressing these challenges requires collaborative efforts from researchers, architects, engineers, and building occupants. Rigorous statistical validation of assessment tools, consensus-building on key satisfaction parameters, and comprehensive studies encompassing multiple physical factors will contribute to a more robust understanding of occupant satisfaction with IEQ.

Prioritizing occupants' experiences and developing performance metrics that holistically consider both energy efficiency and occupant comfort, building designers and managers can create healthier, more comfortable, and more productive workspaces. This, in turn, fosters enhanced employee well-being, satisfaction, and productivity, ultimately benefiting both occupants and building stakeholders alike.

1.3 Research philosophy

Embracing Interpretivism, this research philosophy guides the approach to exploring the intricate relationship between energy efficiency and indoor environmental quality (IEQ) in green buildings. Interpretivism empowers an investigation into the subjective experiences, perceptions, and narratives of occupants and professionals within the context of IEQ. This philosophy enables the uncovering of human dimensions in IEQ, providing insights that complement quantitative and qualitative data and enrich the understanding of how energy-efficient applications impact the indoor environment

1.4 Chapter summary

Chapter 1 introduces the research topic of green office buildings and the importance of indoor environmental quality (IEQ) in creating sustainable and comfortable work environments. It begins by highlighting the rapid urbanisation and the increasing need for energy-efficient and environmentally friendly buildings.

The chapter emphasises the significance of office buildings as crucial contributors to energy consumption and greenhouse gas emissions. It discusses conventional office buildings' challenges regarding energy inefficiency, poor indoor air quality, and inadequate thermal comfort. The research objectives include evaluating the employees' indoor environmental quality satisfaction in office buildings and developing a predictive model to assess employee comfort based on building structural parameters.

The significance of the research is discussed, emphasising the potential benefits of green office buildings in terms of energy efficiency, occupant health and well-being, and environmental sustainability. The chapter highlights the need for research on the challenges and opportunities presented by green office buildings in tropical climates, where energy-efficient design strategies and passive cooling techniques are crucial. The chapter concludes by providing an outline of the subsequent chapters, which will include a literature review on green buildings and energy efficiency measures, an analysis of the survey data and development of the predictive model, and a discussion of the findings and recommendations for the design and operation of green office buildings.

The paste research studies of occupant satisfaction with Indoor Environmental Quality (IEQ) parameters, have been identified several gaps in the context which are;

(a) Lack of Statistical Validation: Current measurement tools used to evaluate occupant satisfaction lack adequate statistical validation, raising concerns about data reliability and accuracy.
(b) Absence of Consensus: There is no consensus among researchers and professionals regarding the specific parameters to consider when assessing occupant satisfaction with IEQ. This lack of standardization leads to inconsistent research findings and hinders result comparison.
(c) Variability in Comfort Requirements: Occupant comfort requirements can vary significantly among individuals, making it challenging to establish a universally applicable approach to assessing satisfaction.
(d) Limited Comprehensive Investigations: Current studies often focus on evaluating individual physical factors in isolation, disregarding the complex interactions between parameters that influence occupant comfort. A more holistic approach is needed.
(e) Lack of Occupant-Centric Metrics: Building performance metrics often prioritize energy efficiency and sustainability over occupant well-being, neglecting the importance of a balanced and user-centric building design.
(f) Neglecting Occupant Perspectives: Real-world experiences and perspectives of building occupants are often overlooked in assessing satisfaction. It is essential to incorporate their insights for practical and user-centric evaluations.

The green buildings have shown promising results in terms of indoor environmental quality in Sri Lanka and that more comprehensive evidence is expected to emerge as green building practices gain momentum and are studied more rigorously in the future. These findings could have implications for shaping future building policies and standards to prioritize sustainability and occupant well-being.While energy efficiency remains a cornerstone of green building design, this research endeavors to demonstrate that its overemphasis may lead to unintended consequences, particularly in the realm of Indoor Environmental Quality (IEQ). This study aims to underscore that green building design and operation should strike a balance between energy efficiency and the well-being of building occupants. It is able to work towards more holistic and sustainable green building practices that prioritize not only energy conservation but also the comfort and health of employees by recognizing the multifaceted impact of energy-related applications.

Overall, Chapter 1 sets the stage for the research by introducing the topic, outlining the research objectives and methodology, and emphasising the significance of green office buildings and indoor environmental quality in creating sustainable and comfortable work environments.

2 LITERATURE REVIEW

The drive towards sustainable and environmentally friendly practices has led to the emergence of green buildings as a prominent solution. Green or sustainable buildings are designed to minimise environmental impact while maximising resource use efficiency and occupant well-being(Attaianese, 2012). Energy efficiency is a crucial aspect of green buildings, aiming to optimise energy and reduce greenhouse gas emissions. Implementing various strategies and technologies, green buildings can significantly reduce energy consumption, minimise water usage, promote renewable energy, improve indoor air quality, and utilise sustainable materials (Ragheb et al., 2016). Identifying the factors that influence the energy efficiency of green buildings is essential for designing, constructing, and operating these structures in an environmentally responsible manner. Factors such as building design, HVAC systems, insulation and sealing, lighting systems, renewable energy integration, and occupant behaviour all play critical roles in determining the overall energy efficiency of green buildings (Chen et al., 2020). Understanding and addressing these factors, green buildings can substantially contribute to a more sustainable and energy-efficient future. This section clarifies the scope of the literature review, focusing on the specific types of green buildings and energy efficiency measures to be discussed. It outlines the parameters and objectives of the study, providing a clear understanding of the topics that will be covered.

2.1 Defining “green buildings” and “energy-efficiency parameters of a building structure”

Green buildings, also known as sustainable or environmentally friendly, are designed, constructed, operated, and demolished to minimise environmental impact while maximising resource efficiency and occupant well-being. These buildings incorporate various strategies and technologies to reduce energy consumption, minimise water usage, promote renewable energy, improve indoor air quality, and utilise sustainable materials.

One of the fundamental principles of green buildings is energy efficiency, which is achieved by optimising the building's design and systems to reduce energy consumption and minimise greenhouse gas emissions (Brambilla et al., 2018). Green buildings employ energy-efficient lighting, heating, ventilation, air conditioning (HVAC), and insulation materials to minimise energy losses and enhance thermal comfort. They often incorporate renewable energy technologies such as solar panels, wind turbines, or geothermal systems to generate clean energy on-site, reducing reliance on fossil fuel-based electricity (Hertwich et al., 2016).

Water efficiency is another important aspect of green buildings. These structures utilise various measures to conserve water, such as efficient plumbing fixtures, rainwater harvesting systems, and greywater recycling (Sheth, 2017). Green buildings help preserve local water resources by minimising water wastage and reducing the strain on municipal water supply systems.

Green buildings also prioritise the use of sustainable materials and construction practices. They employ recycled, reclaimed, or locally sourced materials to reduce the environmental impact of extraction, manufacturing, and transportation (Hossain et al., 2016). Additionally, these buildings aim to minimise construction waste and implement strategies for recycling and reusing materials.

Green buildings ' indoor environmental quality is crucial as they strive to provide healthy and comfortable living or working spaces (Al horr et al., 2016). They employ proper ventilation systems to ensure adequate fresh air circulation and reduce the accumulation of pollutants. Additionally, green buildings often use low-emission materials and adopt design principles that maximise natural daylight and minimise glare, creating a pleasant indoor environment (Balaban & Puppim, 2017).

Beyond the individual building scale, green buildings can contribute to the sustainability of communities and cities. They may incorporate urban design strategies that promote walkability, encourage the use of public transportation, and foster the creation of green spaces. These structures contribute to a more sustainable and resilient urban fabric by minimising the environmental impact of buildings and their surroundings (Moreno et al., 2021).

Energy efficiency parameters consider insulation, fenestration, HVAC, lighting, and renewable energy integration. These parameters ensure that the building is well-insulated, reducing heat transfer and the need for excessive heating or cooling (Bevilacqua et al., 2019). High-performance windows and doors and adequately sealed building envelopes help prevent energy losses due to air leakage and improve thermal comfort (Tanyer et al., 2018).

Efficient HVAC systems optimise energy use by employing advanced technologies, smart controls, and proper sizing to provide heating, ventilation, and air conditioning while minimising energy waste (Taheri et al., 2022). Lighting systems that utilise energy-efficient fixtures, daylight harvesting techniques, and occupancy sensors can significantly reduce electricity consumption (Zolfaghari & Jones, 2022).

Furthermore, integrating renewable energy technologies, such as solar panels or wind turbines, allows buildings to generate clean energy on-site, reducing dependence on traditional energy sources and enhancing energy efficiency (Khalil et al., 2021).

By considering and implementing these energy efficiency parameters, buildings can achieve significant energy savings, lower greenhouse gas emissions, and create healthier and more sustainable environments for occupants.

2.2 Types of green buildings and energy-efficient buildings

As the world embraces sustainable practices, specific types of green buildings and energy efficiency measures have emerged as crucial solutions (Ng, 2018). These innovative approaches aim to reduce environmental impact, optimise resource utilisation, and create healthier and more energy-efficient structures. This study will explore some prominent types of green buildings and the energy efficiency measures associated with them.

(a) Passive House: One notable type of green building is the Passive House, also known as Passivhaus. The Passive House standard focuses on achieving exceptional energy efficiency and comfort without relying heavily on mechanical heating or cooling systems (Colclough et al., 2018). Key energy efficiency measures employed in Passive House construction include meticulous insulation, airtight building envelopes, high-performance windows, and heat recovery ventilation (Hachem-Vermette, 2020). These measures create a highly insulated and virtually airtight structure that minimises heat loss and reduces the need for active heating or cooling.
(b) Net-Zero Energy Building: Net-Zero Energy Buildings (NZEBs) are designed to produce as much energy as they consume over a given period (Ohene et al., 2022). These buildings typically integrate renewable energy systems, such as solar panels, wind turbines, or geothermal systems, to generate clean energy on-site. Simultaneously, they incorporate energy efficiency measures like advanced insulation, efficient lighting, HVAC, and smart energy management systems (W. Wei & Skye, 2021). By optimising energy use and utilising renewable energy sources, NZEBs strive to achieve a net-zero energy balance, significantly reducing reliance on fossil fuels and minimising carbon emissions (Kazmi et al., 2022).
(c) LEED-Certified Buildings: The Leadership in Energy and Environmental Design (LEED) certification is a globally recognised sustainable building design and construction standard (US Green Building Council, 2023). LEED-certified buildings prioritise energy efficiency through a comprehensive approach. They incorporate various measures, including efficient lighting systems, water-saving fixtures, green roofs, and advanced HVAC systems 8. LEED-certified buildings also promote sustainable materials and construction practices, such as using recycled or locally sourced materials and implementing strategies to minimise construction waste (Tomasella et al., 2022).
(d) Green Roofs and Living Walls: Green roofs and living walls are innovative energy efficiency measures that enhance the sustainability of buildings (Tokazhanov et al., 2022). Green roofs cover the rooftop with vegetation, which provides insulation, reduces the heat island effect, and mitigates stormwater runoff (Pragati et al., 2023). These roofs improve energy efficiency by reducing the need for heating and cooling and contribute to the overall well-being of occupants and the environment. Living walls, or vertical gardens, similarly incorporate vegetation on building facades, offering insulation, air purification, and aesthetic benefits (Manso et al., 2021). These measures improve thermal performance, enhance indoor air quality, and create a more sustainable urban environment.
(e) Daylighting and Efficient Lighting: Daylighting is a design strategy that maximises the use of natural daylight within a building to minimise the reliance on artificial lighting (Radić et al., 2019). Energy-efficient lighting technologies contribute to energy savings, such as light-emitting diodes (LEDs) and compact

fluorescent lamps (CFLs). These technologies consume less energy and have longer lifespans than traditional incandescent bulbs (Ngarambe et al., 2022). Additionally, advanced lighting control systems, such as occupancy and daylight sensors, optimise lighting usage and reduce unnecessary energy consumption.

(f) Smart Building Automation Systems: Smart building automation systems utilise advanced technologies to optimise building energy efficiency (Pode, 2020). These systems integrate various components, such as HVAC, lighting, and energy management systems, to control and monitor energy usage intelligently. Smart sensors and controls adjust temperature, lighting, and ventilation based on occupancy and environmental conditions, ensuring energy is used efficiently and only when needed (Minoli et al., 2017). Real-time data analysis and feedback mechanisms further enable continuous monitoring and optimisation of energy consumption in the building (Han & Zhang, 2020).

The research mainly focused on buildings which are certified as green buildings.

2.3 Green building certification types in the world

As the world increasingly recognizes the importance of sustainable development, green building certification systems have gained prominence. These systems provide guidelines and standards for designing, constructing, and operating environmentally friendly and energy-efficient buildings. This section will explore some of the most widely recognized green building certification systems globally.

(a) LEED (Leadership in Energy and Environmental Design): The Leadership in Energy and Environmental Design (LEED) certification system, developed by the United State Green Building Council (USGBC), is one of the most widely known and adopted green building certification programs worldwide. LEED provides a comprehensive framework for evaluating and recognizing sustainable building practices. It considers various aspects, including site selection, energy and water efficiency, indoor environmental quality, materials selection, and innovation. Buildings can achieve different levels of LEED certification, such as Certified, Silver, Gold, or Platinum, based on their performance and adherence to the prescribed criteria (US Green Building Council, 2023.).

(b) BREEAM (Building Research Establishment Environmental Assessment Method): BREEAM, developed by the Building Research Establishment (BRE) in the United Kingdom, is another prominent green building certification system. It evaluates buildings based on various environmental and sustainability criteria, including energy and water efficiency, ecological impact, indoor environmental quality, and management processes. BREEAM provides ratings ranging from Pass, Good, Very Good, Excellent, and Outstanding, enabling buildings to showcase their environmental performance and commitment to sustainability (BREEAM - BRE Group, 2023.).

(c) Green Star: Green Star, developed by the Green Building Council of Australia (GBCA), is a comprehensive green building rating system widely used in

Australia and New Zealand. It assesses buildings based on energy efficiency, water conservation, materials selection, indoor environment quality, and innovation. Green Star certification offers levels, including 4-Star, 5-Star, 6-Star, and above, reflecting the building's environmental performance and sustainability features (Green Star Rating System | Green Building Council of Australia, 2023.).

(d) DGNB (Deutsche Gesellschaft für Nachhaltiges Bauen): DGNB, the German Sustainable Building Council, provides a holistic certification system that assesses and certifies the sustainability performance of buildings in Germany and other European countries. The DGNB certification considers various aspects, including ecological quality, economic viability, socio-cultural factors, technical quality, and processes involved in planning, constructing, and operating buildings. It provides ratings ranging from Bronze to Platinum, reflecting the sustainability level achieved by the facility (German Sustainable Building Council | DGNB GmbH, 2023.).

(e) Green Mark: Green Mark, developed by Singapore's Building and Construction Authority (BCA), is Asia's leading green building certification system. It evaluates buildings based on energy efficiency, water conservation, environmental protection, indoor environmental quality, and other sustainability parameters. Green Mark certification offers ratings from Certified, Gold, GoldPlus, and Platinum, encouraging buildings to strive for higher sustainability and energy efficiency (Green Mark Certification Scheme | Building and Construction Authority (BCA), 2023).

(f) WELL Building Standard: The WELL Building Standard focuses on promoting health and well-being within buildings. Developed by the International WELL Building Institute (IWBI), this certification system evaluates air quality, water quality, lighting, acoustics, comfort, fitness, and mental well-being. The WELL certification recognizes buildings prioritising occupant health and comfort, creating spaces that enhance well-being and productivity (WELL - International WELL Building Institute | IWBI, 2023).

(g) Estidama: Estidama, meaning "sustainability" in Arabic, is a green building rating system developed by the Abu Dhabi Urban Planning Council in the United Arab Emirates. It assesses buildings based on energy and water efficiency, materials selection, indoor environmental quality, and overall sustainability. Estidama provides a Pearl Rating ranging from one to five, with a higher rating indicating a higher level of sustainability achieved (Estidama Program, 2023.).

(h) Green Building Council of Sri Lanka: A set of performance criteria called the GREENSL® Rating System for New and Existing Buildings is employed to certify the operations and upkeep of commercial or institutional buildings of various sizes, both public and private. The goal is to promote high-performing, wholesome, long-lasting, and reasonably priced ecologically sound facilities (Green Building Council, 2023).

2.4 Green building certification types in Sri Lanka

The selection criteria for green buildings in any study hold significant importance as they serve as the foundation for evaluating the impact of sustainable design on occupant comfort and well-being. These criteria ensure that the chosen buildings meet established standards and guidelines for green and sustainable construction, setting a benchmark for environmental performance. By considering factors such as building certifications, climatic variations, floor plan design, window features, and clothing requirements, the selection criteria help control for variables, enhance representativeness, and eliminate potential confounding factors. Such measures enable researchers to conduct focused investigations and provide valuable insights into the role of green buildings in promoting employee environmental quality comfort.

Some prominent green building certification systems in Sri Lanka:

(a) Green Building Council of Sri Lanka (GBCSL): The Green Building Council of Sri Lanka is a non-profit organization that promotes sustainable building practices in the country. It offers the "Green Mark" certification, which assesses the environmental performance of buildings based on various criteria such as energy efficiency, water conservation, waste management, indoor environmental quality, and sustainable materials (Green Building Council, 2023).

(b) Leadership in Energy and Environmental Design (LEED): The LEED certification, developed by the U.S. Green Building Council (USGBC), is internationally recognized and widely adopted in Sri Lanka. LEED evaluates buildings across several categories, including sustainable site development, water efficiency, energy and atmosphere, materials and resources, indoor environmental quality, and innovation in design (US Green Building Council, 2023).

(c) EDGE (Excellence in Design for Greater Efficiencies): EDGE is a certification system developed by the International Finance Corporation (IFC), a member of the World Bank Group. EDGE focuses on resource-efficient building design and offers a simplified and cost-effective process for assessing and certifying residential and commercial buildings. It evaluates energy, water, and materials usage to determine the building's environmental impact (EDGE - Excellence in Design for Greater Efficiencies (EN) - EDGE Buildings, 2023).

(d)National Green Building Standard (NGBS): The National Green Building Standard is a rating system developed by the Green Building Certification Institute (GBCI) in the United States. It provides guidelines for sustainable construction practices, emphasizing energy efficiency, resource conservation, indoor air quality, and occupant comfort. The NGBS certification can be pursued by projects in Sri Lanka as well (National Green Building Standard (NGBS), 2023).

Building requirements to be "certified" under LEED or GBCSL criteria are fundamental in evaluating green buildings. It is a reliable indicator that the facilities meet established green and sustainable construction standards. This criterion ensures that the structures under consideration have comprehensively assessed their environmental performance and implemented sustainable practices throughout their design, construction, and operation.

LEED (Leadership in Energy and Environmental Design) and GBCSL (Green Building Council Sri Lanka) are widely recognized and respected certification systems benchmarking green building practices. Both approaches have rigorous criteria and certification processes to assess buildings' sustainability and environmental performance. By requiring facilities to be certified under these criteria, this case study ensures that the selected buildings have met predetermined standards, ensuring their credibility as green buildings.

One of the critical advantages of requiring LEED or GBCSL certification is that it provides a standardized framework for evaluating the sustainability of buildings (Weerasinghe et al., 2021). These certification systems consider various aspects of building performance, including energy efficiency, water conservation, indoor environmental quality, materials selection, and site sustainability. By adhering to these criteria, certified buildings demonstrate their commitment to minimizing environmental impact and maximising resource efficiency.

Furthermore, LEED and GBCSL certification requires a thorough documentation and verification process, which adds an extra layer of credibility to the buildings. The certification process involves submitting detailed information about the building's design, construction materials, energy systems, water usage, and indoor environmental quality. Independent assessors then verify this documentation to ensure compliance with the specified standards.

This will ensure that the buildings selected for evaluation have undergone a comprehensive sustainability assessment. This guarantees they have implemented sustainable features and practices and facilitates meaningful comparisons between green buildings. It allows for a more accurate analysis of the factors contributing to employee environmental quality comfort and helps identify best practices that can be replicated in future green building projects.

Moreover, the LEED or GBCSL certification requirement aligns with the growing global emphasis on sustainable development and environmental stewardship (Samaraweera et al., 2019). As societies recognize the need to mitigate climate change and reduce resource consumption, green building certification systems have become increasingly important in promoting sustainable practices in the construction industry. Focusing on certified buildings contributes to this larger goal of creating a more sustainable built environment.

2.5 Significance of understanding the factors influencing energy efficiency

Achieving and maintaining high levels of energy efficiency in various sectors requires a deep understanding of the factors that influence it. Understanding the factors affecting energy efficiency is pivotal for driving sustainable development. By identifying and analyzing these factors, policymakers, researchers, and stakeholders can develop effective strategies and initiatives to promote energy-efficient practices (Iqbal et al., 2022). This knowledge facilitates the formulation of targeted policies that align with sustainability goals, reducing energy consumption, mitigating environmental impact, and promoting resource conservation (Shaikh et al., 2017).

Comprehending the factors influencing energy efficiency allows for developing and implementing effective measures to reduce greenhouse gas emissions, combat climate change, and protect the environment (Riahi et al., 2017). Through detailed research and analysis, policymakers can identify critical areas where energy efficiency measures can have the most significant environmental impact. By focusing on these factors, such as building design, transportation systems, and industrial processes, energy consumption can be significantly reduced, decreasing carbon dioxide emissions and overall ecological footprint.

Understanding the factors influencing energy efficiency is crucial for realizing substantial economic benefits. Energy-efficient practices reduce energy consumption, resulting in significant cost savings for individuals, businesses, and governments, by identifying the drivers and barriers to energy -efficient technologies and practices that prioritize occupant comfort, health, and well-being (Gulbinas & Taylor, 2014).

Comprehensive knowledge of the factors influencing energy efficiency enables policymakers to make informed decisions and develop effective policies. Policymakers can create targeted and impactful policies that incentivise energy efficiency measures by researching, analyzing data, and understanding the complex interplay of various factors (Rosenow et al., 2017). This understanding helps overcome barriers and challenges, fosters stakeholder collaboration, and ensures the successful implementation of energy-efficient practices across sectors.

The literature review presents four main building categories when discussing the energy efficiency influencing factors: Residential, Hotel, office and General. Office-specific buildings were chosen for this study because there are so many factors than the building structure influencing occupant satisfaction in those categories. Also, most rated green buildings worldwide are commercial buildings, and occupants evaluation is most effective when considering commercial buildings. Office buildings can again be divided into manufacturing/factory offices and general offices.

2.6 Indoor environmental quality satisfaction of employees

In response to the global environment protection movement, carbon neutrality and energy efficiency have become the core of the building industry’s sustainability agenda. These targets have, thus, reinforced the importance of green building policy (Altomonte et al., 2016; Ravindu et al., 2015)

Green rating tools have been initiated to accelerate the transformation of the building sector towards a more environmentally friendly model (Roderick et al., 2009). All green building rating tools share the same concept of maximising energy and resource efficiency and improving occupant health and well-being (Gou & Xie, 2017). However, the role of green rating tools in improving indoor environmental quality and occupant experience in green buildings is uncertain (Altomonte et al., 2016; Gou et al., 2014).

Occupant satisfaction is of great importance for many organizations, particularly commercial organizations and institutions, as it helps to elevate employee work performance and productivity (Veitch et al., 2007). Occupant satisfaction can also be correlated with turnover intentions and retaining a talented and skilled workforce (Van Dick et al., 2004). Therefore, the success of a sustainable building relies on its indoor environmental quality (IEQ), which directly affects the quality of the occupant's life. It is, thus, essential to assess whether green-certified buildings are genuinely successful as sustainable buildings by evaluating the satisfaction of their occupants.

Numerous post-occupancy studies have investigated the relationship between green certifications and occupant satisfaction. Most green building occupant satisfaction studies come from the U.S. and the U.K.; recent studies are emerging from Asia (Liu et al., 2018; Ravindu et al., 2015). The global evidence review depicts a contradictory body of knowledge regarding the impact of green buildings on occupant satisfaction.

Factors influencing occupant experience and environmental satisfaction in buildings are Indoor Air Quality (IAQ) and Building Design and Facilities Management (BD&FM) factors. When gauging occupant satisfaction in green buildings, the IAQ factors and BD&FM factors considered in most of the research are summarized in Table 1Indoor Environmental Quality (IEQ) satisfaction of employees is a crucial aspect of building design and operation. IEQ refers to the overall quality of the indoor environment, including factors such as air quality, thermal comfort, lighting, acoustics, and spatial layout. Employees' satisfaction with their indoor environment significantly impacts their productivity, well-being, and overall job satisfaction.

Air quality is a critical factor in IEQ satisfaction. Adequate ventilation and controlling pollutants, such as volatile organic compounds (VOCs) and allergens, contribute to a healthier and more comfortable indoor environment. Employees prefer spaces with fresh and clean air, free from odours and harmful substances.

Thermal comfort plays a vital role in employee satisfaction. Maintaining optimal temperature and humidity levels and providing individual control over thermal conditions promotes comfort and productivity. Employees should feel neither hot nor cold, avoiding discomfort and distractions.

Appropriate lighting is essential for both visual comfort and psychological well-being. Sufficient natural daylight, coupled with well-designed artificial lighting, helps prevent eye strain, supports circadian rhythms, and creates a pleasant and inviting atmosphere.

Acoustics greatly influence the comfort and productivity of employees. Controlling noise levels and ensuring proper sound insulation enhance concentration and reduce stress. Providing quiet workspaces and incorporating acoustic design elements like sound-absorbing materials can significantly improve IEQ satisfaction.

Spatial layout and ergonomics are also crucial considerations. A well-designed layout with adequate space, functional furniture, and ergonomic considerations fosters workplace comfort, movement, and efficiency.

To ensure high IEQ satisfaction, building professionals must consider these factors during the design, construction, and operation phases. Regular monitoring, maintenance, and periodic evaluation of IEQ parameters are essential to address issues and improve employee satisfaction. By prioritizing IEQ, employers can create healthier and more satisfying work environments that promote employee well-being and productivity.

As reflected in the majority of the literature, Indoor air quality (IAQ) in LEED buildings is perceived to be higher when compared with non-green buildings (Abbaszadeh et al., 2006; Huizenga et al., 2005; Issa et al., 2011; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; Turner, 2006). Studies conducted on green buildings in China (B. Lin et al., 2016; Pei et al., 2015) and Taiwan (Liang et al., 2014) have similarly reported a higher perceived IAQ in green buildings.

However, post-occupant evaluations conducted on BREEAM buildings have reported lower satisfactory IAQ than their conventional counterparts (Altomonte et al., 2016; Leaman & Bordass, 2007). Results reported from Australia (Paul & Taylor, 2008), South Korea (Sediso & Lee, 2016) and Sri Lanka (Ravindu et al., 2015) have indicated no significant difference in IAQ of green buildings compared to non-green buildings. Furthermore, some studies on LEED buildings (Altomonte & Schiavon, 2013) have also indicated that the IAQ of green buildings is comparable with conventional buildings. Incidentally, Gou, Lau, & Zhang (2012) reported that green buildings perform better in summer but worse in winter. This study is supported by (Gou, Lau, & Shen, 2012), who reported that the summer performance of LEED buildings in Hong Kong concerning IAQ was much better than in winter.

Regarding lighting performance, most research has detected no significant difference in LEED buildings (Abbaszadeh et al., 2006; Altomonte & Schiavon, 2013; Huizenga et al., 2005). However, some studies have indicated a higher satisfaction score (Issa et al., 2011; S.-K. Kim et al., 2015; Turner, 2006), whereas others have reported a lower satisfaction score in LEED buildings (Brown et al., 2010; Y. S. Lee & Kim, 2008). Similarly, in BREEAM buildings, some studies have perceived a higher satisfaction score (Baird et al., 2012; Y. Zhang & Altan, 2011), while others have reported no significant differences (Altomonte et al., 2016; Leaman & Bordass, 2007) in satisfaction between green and non-green groups.

In the Chinese context, two studies (Gou, Lau, & Shen, 2012; Gou, Lau, & Zhang, 2012) have reported no significant differences in lighting performance, whereas another study (Pei et al., 2015) indicated higher perceived lighting scores in green buildings compared to their conventional counterparts. Post-occupancy evaluation surveys conducted on Green Star buildings in Australia (Khoshbakht, Gou, Xie, et al., 2018; Paul & Taylor, 2008) have reported no significant differences in lighting performance between green and non-green buildings. However, studies conducted in South Korea (Sediso & Lee, 2016) on G-SEED buildings (Green Standard for Energy and Environmental Design) have reported higher perceived satisfaction with lighting performance in green buildings.

Concerning literature, green buildings are the least successful in noise performance. The majority of papers have reported either no significant differences (Abbaszadeh et al., 2006; Altomonte & Schiavon, 2013; Huizenga et al., 2005) or lower satisfaction scores (Brown et al., 2010; Issa et al., 2011; Y. S. Lee & Kim, 2008; Turner, 2006) in LEED buildings in comparison with non-LEED buildings. A study on higher education buildings in Australia (Khoshbakht, Gou, Xie, et al., 2018) has also reported lower satisfaction levels with noise in Green Star-certified buildings. Similarly, post-occupant evaluations conducted on BREEAM buildings (Altomonte et al., 2016; Leaman & Bordass, 2007; Y. Zhang & Altan, 2011), Green Star buildings (Paul & Taylor, 2008), G-SEED facilities in South Korea (Sediso & Lee, 2016) and LEED buildings in Sri Lanka (Ravindu et al., 2015) have reported no significant differences in the noise performance of green and non-green buildings. Contrarily, some studies (Liang et al., 2014; Newsham et al., 2013) have indicated a higher perceived satisfaction score in the noise performance of green buildings.

Regarding thermal comfort, most studies have detected a higher performance in green buildings than in conventional buildings. Post-occupant evaluations conducted on LEED buildings (Brown et al., 2010; Huizenga et al., 2005; Issa et al., 2011; S.-K. Kim et al., 2015; Newsham et al., 2013; Y. Zhang & Altan, 2011) have indicated greater thermal comfort compared to their conventional counterparts. In the Chinese context, most studies (Gou et al., 2014; B. Lin et al., 2016; Pei et al., 2015) have reported higher thermal satisfaction in green buildings. Similarly, studies conducted in Taiwan (Liang et al., 2014) and South Korea (Sediso & Lee, 2016) reported that green buildings significantly outperform non-green buildings concerning thermal comfort.

However, few studies conducted on LEED buildings (Altomonte & Schiavon, 2013), BREEAM buildings (Altomonte et al., 2016), China Three Star buildings (Gou, Lau, & Zhang, 2012)and Green Star buildings (Menadue et al., 2014; Paul & Taylor, 2008), have reported comparable thermal satisfaction in green and non-green buildings, with no significant differences. Table 2.1 summarizes the above description.

Table 2.1: Satisfaction levels reported in different case studies

	IAQ	Lighting	Acoustics	Thermal Comfort	Layout	Furnishings	Privacy	Cleanliness	Operation & Maintenance	Workspace	Colours and textures
(Abbaszadeh et al., 2006)	✓	◦	◦	✓		✓		✓	✓	✓	
(Issa et al., 2011)	✓	✓	×	✓							
(S.-K. Kim et al., 2015)	✓	✓	✓	✓	✓						
(Y. S. Lee & Kim, 2008)	✓	×	×	✓	×	✓		✓	✓		
(B. Lin et al., 2016)	✓	✓	✓	✓		✓	✓	◦	✓		
(Pei et al., 2015)	✓	✓	✓	✓							
(Turner, 2006)	✓	✓	×				×				
(Altomonte et al., 2016)	×	◦	◦	◦			◦	◦			
(Sediso & Lee, 2016)	◦	✓	◦	✓	✓	✓		✓	✓		
(Altomonte & Schiavon, 2013)	◦	◦	◦	◦		◦	◦	◦	◦		
(Huizenga et al., 2005)	✓	◦	◦	✓							
(Brown et al., 2010)		×	×	✓							
(Paul & Taylor, 2008)	◦	◦	◦	◦							
(Khoshbakht, Gou, Xie, et al., 2018)	✓	◦	×	✓							
(Liang et al., 2014)	✓	✓	✓	✓							
(Leaman & Bordass, 2007)	×	◦	◦	×							
(Ravindu et al., 2015)	◦	✓	◦	×	✓	✓	✓	✓			
(Baird et al., 2012)	✓	✓	◦	×		✓		✓	✓		
(Y. Zhang & Altan, 2011)		✓	◦	✓							
(Newsham et al., 2013)			✓	✓						✓	✓
(Liang et al., 2014)	✓	✓	✓	✓							

**Higher satisfaction than conventional buildings (✓), equal satisfaction with conventional buildings (◦) and lower satisfaction than conventional buildings (×)*

Contradictory results on the perceived thermal performance of green buildings have also been reported. Baird et al. (2012) reported lower satisfaction scores in BREEAM buildings than non-BREEAM buildings. Another Australian-based study (Leaman et al., 2007) reported that Green Star buildings underperformed their conventional counterparts regarding thermal comfort satisfaction. Similarly, a survey in Sri Lanka (Ravindu et al., 2015) reported a lower perceived thermal comfort in green buildings.

As reflected in the literature, Building Design and Facilities Management (BD&FM) of green buildings are perceived to be better than conventional buildings. Studies conducted on LEED buildings (Brown et al., 2010; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; Newsham et al., 2013) and BREEAM buildings (Baird et al., 2012) have recorded a more satisfactory performance in green buildings. In the Chinese context, higher perceived satisfaction scores were achieved in green buildings in terms of operation and maintenance (B. Lin et al., 2016), health, and productivity (Gou et al., 2014). Furthermore, green buildings in Sri Lanka (Ravindu et al., 2015), South Korea (Sediso & Lee, 2016) and Green Star buildings in Australia (Khoshbakht, Gou, Xie, et al., 2018) have shown satisfactory performance in terms of BD&FM parameters when compared with non-green facilities. Issa et al. (2011) also reported that student and staff absenteeism in green schools was 2-7.5 % lower, and students' performance was 8-19% higher when compared with conventional schools. However, one study conducted in China (Gou, Lau, & Zhang, 2012) indicated that occupant satisfaction in green and non-green buildings was comparable. The factors affecting IEQ comfort in a building and impacting employee satisfaction in an office building will be further analysed in Chapter 4, the systematic literature review results.

2.7 Using machine learning predictive models to predict employee comfort over building structural parameters

Recently, a growing interest has been in leveraging machine learning predictive models to enhance our understanding of the complex relationship between building structural parameters and employee comfort in indoor environments.

Researchers and building professionals can now develop predictive models that can accurately estimate employee comfort levels based on building structural parameters by harnessing the power of machine learning algorithms (W. P. Abeyrathna et al., 2023). These models consider data collected from sensors, surveys, and other sources to understand the indoor environment better.

Machine learning models can analyze vast amounts of data, identifying patterns and relationships that might go unnoticed (Daniel, 2015). They can learn from historical data to predict future employee comfort levels under different scenarios. This allows building managers and designers to optimise structural parameters to create environments promoting optimal comfort and well-being.

The use of machine learning in predicting employee comfort has the potential to revolutionize building design and operation (Seyedzadeh et al., 2020). Optimising building structural parameters based on predictive models, energy consumption can be minimised, and occupant satisfaction can be maximised. This approach promotes sustainability, productivity, and occupant health in the built environment (Azar et al., 2020).

However, it is essential to note that developing accurate and reliable predictive models requires high-quality data, including comprehensive information on building structural parameters and corresponding employee comfort assessments (García Kerdan & Morillón Gálvez, 2020). Careful model validation and ongoing data collection and analysis are necessary to ensure the reliability and effectiveness of the predictive models (D. Yan et al., 2015).

2.7.1 The concept of regression models and their application in predicting and optimising IEQ factors

Achieving optimal IEQ comfort requires a comprehensive understanding of the factors influencing employee comfort and effective decision-making strategies (Roumi et al., 2023). In this context, machine learning regression models have emerged as valuable tools for analyzing and predicting IEQ conditions (Salamone et al., 2020). This research aims to evaluate the effectiveness of different regression models in enhancing employee IEQ comfort through decision-making processes.

Machine learning regression models play a pivotal role in enhancing IEQ comfort through their ability to analyze and predict complex relationships between various factors (S. Li et al., 2023). These models can effectively process large volumes of sensor data, including temperature, humidity, air quality, lighting, and noise levels (Balogun et al., 2021). Leveraging historical data and patterns, machine learning regression models provide insights into how these factors impact employee comfort. This knowledge allows organizations to make data-driven decisions to create a more comfortable and productive work environment.

The primary objective of this research is to evaluate the effectiveness of different regression models in enhancing employee IEQ comfort through decision-making processes. Specifically, the research aims to compare and analyze the performance of the random forest model, lasso regression model, decision tree regression model, and support vector regression model in predicting and optimising IEQ factors. The study evaluates these models to identify the most suitable approach for IEQ comfort enhancement.

Comprehensively assessing the strengths and limitations of each model can determine which regression model offers the highest predictive accuracy, interpretability, and scalability. Furthermore, we will explore each model's unique advantages and applications in the context of IEQ comfort optimisation. This analysis will provide valuable insights into the potential of machine learning regression models for decision-making processes related to employee IEQ comfort.

This research evaluates and compares different regression models to identify the most effective approach for enhancing employee IEQ comfort. By achieving this objective, organizations can make informed decisions prioritising employee well-being, resulting in improved productivity and job satisfaction.

Regression models can be applied to predict and optimise IEQ factors. One common approach is to develop predictive models that estimate the comfort level based on historical data (Crosby & Rysanek, 2022). By training a regression model on a dataset that includes employee feedback on their comfort level and corresponding IEQ variables, organizations can create a model that accurately predicts comfort based on the current IEQ conditions. This enables proactive adjustments to the environment to maintain optimal comfort levels.

Furthermore, regression models can optimise IEQ factors by identifying the ideal values or ranges that maximise employee comfort (Roskams & Haynes, 2019). Regression models can determine the optimal conditions that lead to the highest level of employee comfort by analyzing the relationships between IEQ variables and comfort levels (Kaushik et al., 2020). Organizations can then use this information to adjust environmental parameters or implement control strategies to achieve the desired comfort level.

2.7.2 Overview of the random forest regression model and its key characteristics

The random forest regression model is a powerful ensemble learning technique that combines multiple decision trees to make predictions (Sagi & Rokach, 2018). Each decision tree in the random forest is built on a different subset of the training data using a random selection of features (Jaiswal & Samikannu, 2017). The final prediction is obtained by aggregating the predictions of individual trees.

Key characteristics of the random forest model include:

(a) Ensemble Approach: The random forest model leverages the power of multiple decision trees to improve prediction accuracy and reduce overfitting (Sagi & Rokach, 2018).
(b) Feature Randomness: Each decision tree is trained on a random subset of features, reducing the correlation among trees and enhancing model robustness (Cui et al., 2018).
(c) Bagging: The random forest employs a bagging technique, where each tree is trained on a bootstrapped sample of the training data, allowing for better generalization (Altman & Krzywinski, 2017).
(d) Variable Importance: The model measures feature importance, allowing analysts to identify the most influential factors in the prediction process (Gregorutti et al., 2017).
(e) Nonlinear Relationships: The random forest model can effectively capture nonlinear relationships between input and target variables, making it suitable for complex data patterns (M. Chen et al., 2021).

The random forest model offers several advantages in handling IEQ data and capturing complex relationships:

(a) Robustness to Noise and Outliers: The random forest model is robust to noise and outliers in the data, making it suitable for handling real-world IEQ datasets that may contain measurement errors or anomalies (Afanador et al., 2016).
(b) Handling High-Dimensional Data: IEQ optimisation often involves many variables. The random forest model can effectively handle high-dimensional data, making it suitable for analyzing multiple IEQ factors simultaneously (Venkatesh & Anuradha, 2019).

(c) Nonlinear Relationship Capture: IEQ comfort is influenced by various nonlinear relationships among factors, such as the interaction between temperature, humidity, and air quality. The random forest model captures such complex relationships, enabling accurate predictions (Shao et al., 2020).
(d) Feature Importance Assessment: The random forest model measures feature importance, allowing analysts to identify the most influential IEQ factors affecting comfort. This information helps prioritize and allocate resources effectively for IEQ improvement (Gregorutti et al., 2017).

Several case studies and examples demonstrate the effectiveness of the random forest model in IEQ comfort prediction and optimisation:

(a) According to the research studies conducted by (Jin et al., 2021; Q. Y. Li et al., 2021; H. Zhang et al., 2022) utilised the random forest model to predict thermal comfort in an office building based on temperature, humidity, air velocity, and clothing insulation data. The models accurately predicted comfort levels and identified the key factors contributing to comfort variations, leading to targeted interventions to improve thermal conditions.
(b) The case study conducted by (Sun et al., 2020) proved that in an innovative building project, the random forest model optimised lighting conditions for employee comfort. The model analyzed data on lighting levels, natural light availability, and employee feedback on perceived comfort. Adjusting lighting parameters based on the model's recommendations significantly improved employee comfort.
(c) The model's effectiveness has been demonstrated through various case studies, highlighting its accuracy in predicting IEQ comfort levels and facilitating targeted interventions. Overall, the random forest model is valuable for organizations seeking to enhance employee comfort and well-being in indoor environments.

2.7.3 Overview of the lasso regression model and its key characteristics

The Lasso (Least Absolute Shrinkage and Selection Operator) regression model is a linear regression technique incorporating feature selection and regularization (Januaviani et al., 2019). It is designed to overcome the limitations of traditional linear regression models by introducing a penalty term that encourages sparse and interpretable models (De Bock & De Caigny, 2021).

The Lasso regression model's primary purpose is feature selection and regularization. Feature selection refers to identifying the most influential variables that contribute to the prediction of the target variable (Otchere et al., 2022). On the other hand, regularisation aims to prevent overfitting by introducing a penalty term that discourages the model from relying too heavily on many features (Tredennick et al., 2021).

In Lasso regression, the penalty term is based on the sum of the absolute values of the regression coefficients (Dudek, 2016). The model can shrink the coefficients of less influential features to zero, eliminating them from the model by tuning the regularisation strength parameter. This characteristic makes Lasso regression particularly useful in situations with many features, some of which may be irrelevant or redundant.

In the context of IEQ comfort, various factors such as temperature, humidity, air quality, lighting levels, and noise levels may influence the overall comfort experienced by employees. The Lasso regression model can analyze historical data that includes these factors and the corresponding comfort rating employees provide (Assaf & Srour, 2021). Identifying the most influential features, the model helps organizations pinpoint the key factors significantly impacting IEQ comfort.

Moreover, the Lasso regression model offers interpretability by assigning non-zero coefficients to the selected features (Jain & Xu, 2021). This allows organizations to understand the direction and magnitude of the relationship between each feature and IEQ comfort. This information can guide decision-making processes, helping organizations prioritize interventions and allocate resources effectively to improve specific IEQ factors.

Several case studies and examples demonstrate the effectiveness of the Lasso model in IEQ comfort prediction and optimisation:

(a) Study on Office Environment: (Alsaleem et al., 2020; Y. Kim et al., 2021; Sikram et al., 2020) research studies analysed that, the Lasso regression model was used to predict thermal comfort in office environments based on temperature, humidity, air velocity, and clothing insulation. The model identified the most influential factors and helped prioritize interventions to optimise thermal conditions and enhance employee comfort.

(b) Smart Building Case Study: The Lasso regression model was applied to optimise lighting conditions for IEQ comfort in an innovative building project. According to (Aguilar et al., 2022; Kapoor et al., 2021; Nag, 2019), the model identified the key lighting parameters that significantly impacted comfort by analyzing data on lighting levels, natural light availability, and employee feedback. This information guided the adjustment of lighting systems to create a more comfortable working environment.

(c) Indoor Air Quality Optimisation: Another studies (Sethi & Mittal, 2021 and Shi et al., 2023) used the Lasso regression model to optimise indoor air quality for IEQ comfort. The model analyzed factors such as CO2 levels, VOCs, temperature, and humidity to identify the most influential variables affecting air quality. This information helped organizations implement targeted strategies to improve air quality and enhance employee comfort.

These examples highlight the importance of the Lasso regression model in IEQ comfort optimisation. Effectively selecting and prioritizing influential factors, the model helps organizations make informed decisions and implement targeted interventions to create a more comfortable and productive work environment. The interpretability of the Lasso regression model further enhances its value in understanding the relationship between IEQ factors and comfort, facilitating data-driven decision-making processes.

2.7.4 Overview of the decision tree regression model and its key characteristics

The decision tree (DT) regression model is a powerful and intuitive machine learning algorithm for predictive modelling and decision-making (Elhazmi et al., 2022). A non-parametric supervised learning method creates a tree-like model of decisions and their possible consequences (Nugroho et al., 2022). Each internal node of the tree represents a feature or attribute, while the leaf nodes represent the predicted target value.

Characteristics of the decision tree regression model include:

(g) Interpretability: Decision trees are easy to interpret and understand. The model's structure resembles a flowchart, making it intuitive for users to follow the decision-making process and comprehend the factors influencing the prediction (Wu et al., 2023).
(h) Nonlinear Relationship Capture: Decision trees can capture nonlinear relationships between input and target variables. They can identify complex interactions and patterns among the IEQ factors, allowing for accurate predictions (Pastoriza et al., 2022).
(i) Handling Categorical and Numerical Data: Decision trees can take categorical and numerical data, making them versatile for various IEQ factors. The model can split nodes based on categorical attributes and use numerical thresholds to partition the data (Tan et al., 2017).
(j) Feature Importance Assessment: Decision trees measure feature importance, indicating the relative contribution of each variable in the prediction process. This information helps identify the most influential IEQ factors affecting comfort (Fritz et al., 2022).

Decision trees excel at capturing complex interactions between IEQ factors due to their hierarchical structure and ability to make multiple splits (Himeur et al., 2023). The model can automatically identify and exploit interactions among variables, considering both main effects and interactions among features.

For example, in IEQ comfort, decision trees can capture interactions between temperature, humidity, and air quality. By recursively splitting the data based on different IEQ thresholds, the model can identify specific temperature, humidity, or air quality ranges where comfort levels significantly change (Gupta et al., 2017). This enables a comprehensive understanding of how different factors interact to influence comfort.

Additionally, decision trees can handle interactions between categorical and numerical variables (Nag, 2019). They can create branches that separate data based on categorical attributes and further split the data based on numerical thresholds, allowing for the analysis of interactions between different IEQ factors.

Several case studies and examples demonstrate the effectiveness of the DT model in IEQ comfort prediction and optimisation:

(a) A study conducted by (Aguilar et al., 2022; Bavaresco et al., 2021; Fassio et al., 2014) focused on optimising IEQ comfort in a workplace, decision tree regression models were used to predict comfort levels based on temperature, humidity, air quality, and lighting conditions. The models identified the key factors and their interactions that significantly impacted comfort. This information guided organizations in implementing targeted interventions to improve specific IEQ factors and enhance overall comfort.

(k) Decision tree regression models have been used in the studies published by (Bourhnane et al., 2020; Farzaneh et al., 2021; Ridwana et al., 2020)context of energy-efficient buildings to optimise IEQ factors while minimizing energy consumption. By considering the interactions between temperature, humidity, and lighting levels, the models provided insights into the optimal settings that balance comfort and energy efficiency.

(l) Decision tree regression models have been applied by (Aguilar et al., 2022; Kapoor et al., 2021; Y. Zhang et al., 2019) to optimise indoor air quality for improved IEQ comfort. The models identified the critical variables and their interactions affecting air quality by analyzing factors such as CO_2 levels, VOCs, temperature, and humidity. This knowledge helped organizations implement targeted strategies to enhance air quality and employee comfort.

2.7.5 Overview of the support vector mechanism regression model and its key characteristics

The support vector regression (SVR) model is a supervised machine learning algorithm for regression tasks (Kavitha et al., 2017). It is based on the principles of support vector machines (SVM). It aims to find the optimal hyperplane that maximises the margin between the predicted values and the actual data points (Ramaraj, 2013).

The SVR model is particularly applicable to IEQ comfort optimisation because it can handle both numerical and categorical IEQ factors and effectively capture the nonlinear relationships between these factors and comfort levels (Himeur et al., 2023). By incorporating a tolerance margin, SVR can balance the trade-off between accurately predicting comfort levels and allowing some flexibility within a specific range.

The model works by mapping the input variables into a higher-dimensional feature space and finding the optimal hyperplane that separates the data points (Rienow & Goetzke, 2015). Unlike traditional regression models, SVR does not aim to minimise the error for all data points but focuses on achieving a maximum margin between the predicted values and a tolerance region around the target variable.

Support Vector Regression offers several advantages in handling nonlinear relationships and outliers:

(a) Nonlinear Relationship Capture: SVR can capture nonlinear relationships between IEQ factors and comfort levels. By utilizing kernel functions, such as the radial basis function (RBF), the model can map the data into a higher-dimensional space where nonlinear relationships become linear. This enables SVR to capture complex interactions among IEQ factors effectively (Mao et al., 2019).

(m) Robustness to Outliers: SVR is robust to outliers, common in real-world datasets. The model focuses on finding a hyperplane that maximises the margin between the predicted values and the data points within the tolerance region. Outliers have a minimal impact on the position of the hyperplane, making SVR suitable for handling data with noisy or irregular values (Balasundaram & Meena, 2019).

(n) The margin of Tolerance: SVR incorporates a margin of tolerance that allows some flexibility in the predictions. This feature is valuable in IEQ comfort optimisation as it accounts for individual preferences and variations. The tolerance region enables the model to consider comfort levels within a certain range rather than strictly adhering to a specific target value (García-Floriano et al., 2018).

Several case studies and examples demonstrate the effectiveness of the SVR model in IEQ comfort prediction and optimisation:

(a) According to (Carli et al., 2019; T. M. S. Kumar & Kurian, 2022; X. Yan, 2023)thermal comfort optimisation in buildings, SVR was used to predict comfort levels based on temperature, humidity, air quality, and clothing insulation data. The model accurately captured the nonlinear relationships and provided insights into the optimal ranges of IEQ factors for improved thermal comfort.

(o) SVR was applied to analyze a large dataset in several studies (Altomonte et al., 2019; Y. K. Kim et al., 2022; Zagreus et al., 2004)containing IEQ factors and occupant satisfaction ratings. The model effectively identified the key factors influencing satisfaction and their nonlinear relationships. This knowledge guided the implementation of targeted interventions to enhance IEQ comfort and overall occupant satisfaction.

(p) SVR has been used to develop personalized comfort models in the studies conducted by (Arakawa Martins et al., 2022; Feng et al., 2020; J. Kim et al., 2018) considering individual preferences and variations. By incorporating data on personal factors, such as age, gender, and activity levels, the model can adapt comfort predictions to specific individuals. This approach allows for a more tailored and personalized indoor environment that maximises individual comfort.

These examples illustrate the effectiveness of the support vector regression model in enhancing IEQ comfort. SVR's ability to capture nonlinear relationships, handle outliers, and consider a margin of tolerance makes it a valuable tool for optimising IEQ factors. Table 2.2 summarises the advantages and disadvantages of each machine learning regression model.

Machine learning regression models are crucial in optimising IEQ comfort by providing accurate predictions and insights into the factors that affect comfort levels. These models enable organizations to make data-driven decisions and implement targeted interventions, improving employee well-being, productivity, and satisfaction. By analyzing historical data on IEQ factors and comfort ratings, regression models can identify the most influential variables and their relationships with comfort. This information helps organizations understand the complex interactions among different IEQ factors and prioritize interventions for maximum impact. Machine learning regression models also allow for exploring nonlinear relationships, capturing the intricacies of IEQ comfort optimisation.

Table 2.2: Advantages and Disadvantages of using RF, DT, Lasso and SVM regression models

Regression Model	**Advantages**	**Disadvantages**
Random Forest	Ability to handle high-dimensional data with numerous IEQ factors. Robustness to noise and outliers. Capturing complex relationships and interactions among IEQ factors. Feature importance assessment for prioritizing interventions. Suitable for both numerical and categorical variables.	Less interpretable compared to simpler models. Potential overfitting if hyperparameters are not carefully tuned. Computationally expensive, especially for large datasets. Difficulty in visualizing the entire model due to its ensemble nature.
Lasso Regression	Feature selection and regularization to identify influential factors. Improves model interpretability. Reduces the impact of irrelevant or redundant variables.	The selection of the regularization parameter (alpha) may require careful tuning. May shrink coefficients to zero, eliminating potentially valuable variables.

	Handles high-dimensional data effectively. Ability to handle both numerical and categorical variables.	May face challenges when dealing with multicollinearity among predictor variables. Assumes a linear relationship between predictors and the target variable, limiting its ability to capture nonlinear relationships.
Decision Tree Regression	Capturing complex interactions and patterns between IEQ factors. Can handle both numerical and categorical variables. An intuitive interpretation of the model structure. Can handle nonlinear relationships. Identifies feature importance.	Prone to overfitting if not appropriately pruned. Sensitive to small changes in the training data. Difficulty in capturing interactions beyond the depth of the tree. It may generate overly complex models, resulting in decreased interpretability.
Support Vector Mechanism	Effective for capturing complex nonlinear relationships. Requires careful selection of the kernel function and hyperparameters. Can handle both numerical and categorical variables through appropriate kernel functions. Tends to generalize well with small to medium-sized datasets. Robust to outliers due to the use of a loss function that penalizes errors. Can handle high-dimensional data	Computationally expensive, especially for large datasets. Difficult to interpret the learned parameters and understand the model's inner workings. May not perform well when the number of features is much larger than the number of samples. SVM can be sensitive to noise in the training data. May suffer from slow training and inference times for large datasets. Can be challenging to handle missing data and imbalanced datasets.

	effectively through the kernel trick. Provides flexibility in choosing different kernel functions to capture various data patterns. Performs well in the presence of multicollinearity among predictor variables.	

2.8 Chapter summary

Chapter 2 provides a comprehensive literature review on green buildings and energy efficiency measures. It begins by highlighting the importance of ensuring energy-efficient and sustainable buildings in the fast-paced and technology-driven world. The chapter emphasizes the significance of indoor environmental quality (IEQ) satisfaction in office buildings, particularly in tropical regions.

The focus of the study is on tropical green office buildings, which are characterized by sustainable design and environmentally friendly practices. These buildings incorporate energy-efficient systems, passive cooling techniques, and natural ventilation to create comfortable and sustainable indoor environments. The chapter explains that despite their ecological focus, indoor environmental quality comfort in tropical climates often falls below desirable levels compared to other climatic conditions.

The research objectives are outlined, which include investigating decision-making tools and techniques for the green building industry, evaluating the awareness and perception of green technology applications, and developing a predictive model for employee indoor environmental quality comfort.

Most research shows that green buildings provide improved IAQ and thermal comfort compared to conventional buildings. A higher inconsistency was observed in lighting performance, with over 50% of the research indicating either no difference or poorer performance in green buildings when compared with conventional buildings. Based on the findings, the acoustic performance in green buildings was comparable to non-green facilities in most reported work. Regarding the BD&FM parameters, most papers indicated better performance in green buildings, particularly concerning furnishing, cleanliness, operation, and maintenance.

A literature review into occupant satisfaction in green-certified buildings discloses a contrary body of research. As depicted in Table 2, there is no consistent global evidence to prove that green buildings are more satisfactory than non-green buildings. However, based on the literature, the following findings can be surmised.

A literature review into occupant satisfaction in green-certified buildings discloses a contrary body of research. As depicted in Table 2, there is no consistent global evidence to prove that green buildings are more satisfactory than non-green buildings. However, based on the literature, the following findings can be surmised.

The contradicting results in the literature review can be attributed to various factors influencing occupant evaluations.

(a) The occupancy period – If the survey was conducted on a newly constructed green building, a short period after occupancy, it could manifest artificially higher satisfaction scores. Singh et al., (2010) reported this as the “honeymoon” effect. Therefore, the occupancy period can bias the evaluation and influence occupant satisfaction scores.
(q) Socio-economic background: The occupant satisfaction studies considered in Table 2 have been conducted in various countries of different socio-economic backgrounds. In developed countries like U.S. and U.K., where standards are very high and stringent, improvements brought by green building designs may be marginal (Darko, Chan, et al., 2017; Darko, Zhang, et al., 2017). Therefore, modifications to occupant satisfaction might be minimal. However, building design and service standards are relatively low in developing countries like China and Sri Lanka. In these countries, the improvements brought by green building concepts can significantly improve the building design and operation (He et al., 2014; D. X. Zhao et al., 2015). This will lead to considerable improvements in occupant satisfaction. Therefore, the socio-economic background of the countries must also be considered when evaluating the occupant satisfaction scores.
(r) Green building features vary from building to building. This could contribute to the inconsistent results in occupant satisfaction observed in the literature.
(s) The size and characteristics of the sample will also affect the findings. If the number of respondents in green and non-green buildings is disproportionate, this asymmetry might lead to biases when comparing their responses. Therefore, the effect of sample size must be given due consideration.

The chapter discusses the use of machine learning algorithms, specifically Random Forest (RF), Support Vector Machine (SVM), and Decision Tree (DT), to develop a predictive model for employee comfort. These models are chosen based on their success in predictive modelling tasks and their ability to capture complex relationships between variables. Survey data from employees in fourteen green office buildings in tropical climates were collected to determine their satisfaction levels with indoor environmental quality comfort.

The literature review section explores various types of green buildings, including commercial, residential, educational, healthcare, and government buildings. It also delves into different energy efficiency measures implemented in green facilities, such as building design, HVAC systems, lighting systems, renewable energy integration, and occupant behaviour.

The chapter discusses the importance of understanding the factors influencing energy efficiency in green buildings. It highlights the environmental conservation and climate change mitigation benefits, resource efficiency and energy security advantages, economic benefits and cost savings, health, comfort, productivity enhancements, regulatory compliance, and market demand.

Additionally, the chapter provides an overview of the specific types of green buildings and energy efficiency measures. It mentions Passive House, Net-Zero Energy Building, LEED-certified buildings, green roofs, living walls, daylighting, efficient lighting, and smart building automation systems as examples of these types and measures.

The significance of green building certifications is also addressed, focusing on LEED, BREEAM, Green Star, DGNB, Green Mark, WELL Building Standard, Estidama, and the Green Building Council of Sri Lanka's GREENSL® Rating System.

Furthermore, the chapter discusses the structural parameters that influence the energy efficiency of green office buildings. It highlights the importance of building envelopes, orientation and layout, energy-efficient materials, and renewable energy technologies in achieving energy efficiency goals.

The concept of indoor environmental quality satisfaction of employees is explored, emphasizing the significance of factors such as air quality, thermal comfort, lighting, acoustics, and spatial layout. The chapter explains that understanding these factors is crucial for creating healthier and more comfortable work environments.

Finally, the chapter discusses using machine learning predictive models to predict employee comfort based on building structural parameters. It highlights the ability of RF, DT, Lasso and SVR machine learning algorithms to analyze data, identify patterns, and make accurate predictions. The chapter emphasizes the potential of this approach to revolutionize building design and operation by optimising building structural parameters to enhance occupant satisfaction and energy efficiency.

3 METHODOLOGY

The study is driven by a systematic literature review and bibliographic analysis, selecting suitable green buildings, employee satisfaction surveys, developing predictive models and creating a user interface to get employee satisfaction for indoor environmental quality.

The systematic literature review was conducted to perform three main approaches (i) Identify main factors related to the indoor environmental quality of the buildings and (ii) Identify the employee satisfaction-related elements in a building,. An extensive literature review was conducted to identify the factors related to employee comfort and machine learning regression models, which have been demonstrated to have many theoretical and practical applications in decision-making in green buildings and employee satisfaction. A quantitative analysis was carried out to understand what decision-making models have been applied more in building-related applications and the topics related to occupant satisfaction and in which they have been used.

The data collection process consisted of two primary steps. Firstly, an Employee Satisfaction Evaluation (ESE) questionnaire survey was conducted to evaluate employee satisfaction with the thermal comfort of their workplace. The survey was distributed to 14 LEED or GBCSL-certified office building employees. To assess the necessary information for the survey process, we initially evaluated the satisfaction opinion criteria, thermal comfort metrics, and fundamental demographic and geographic data based on the summary of the thermal comfort feature kinds. The ESE (Employee Satisfaction Evaluation) procedure consisted of three main steps: pre-collection of information, dissemination of questionnaires, and environmental measurement. Subsequently, physical measurements were taken to obtain data on the thermal comfort-related parameters of the buildings. The collected information was then utilised for data processing and developing predictive machine learning models. This approach allowed identifying the most crucial building parameters for predicting employee thermal comfort in green office buildings and accurately analysing the collected data. The user interface was created to examine the employee satisfaction response under different IEQ factors and overall IEQ satisfaction. Figure 3.1 illustrates the overall view of the research study.

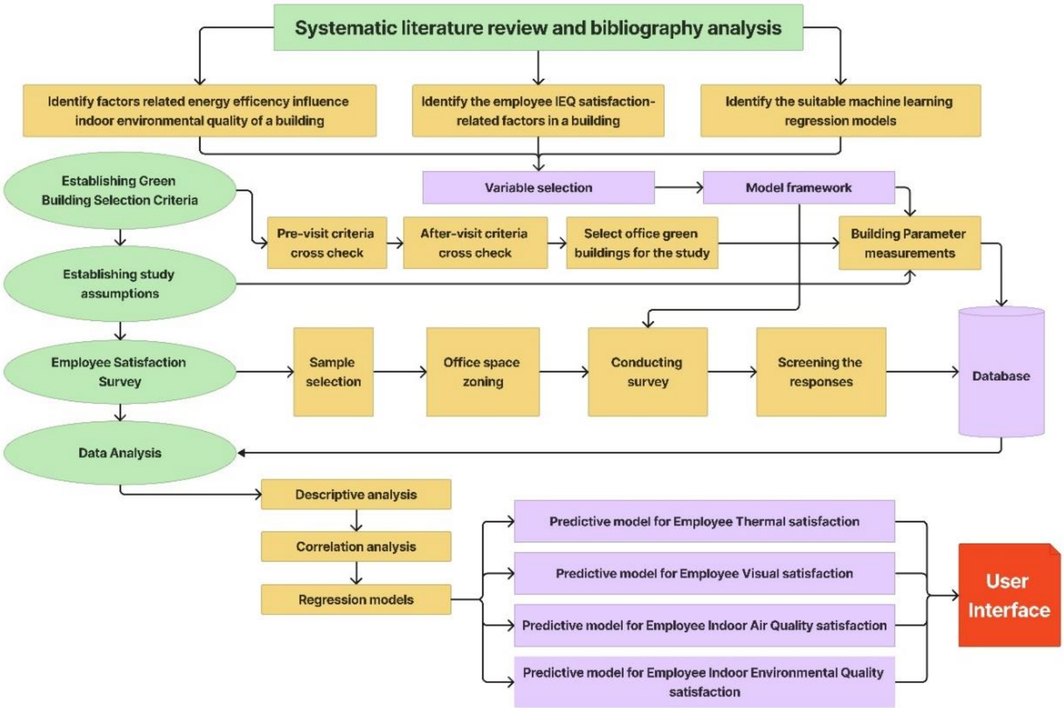


Figure 3.1: Methodology of the research

3.1 Keyword search and review procedure

Indoor environmental quality (IEQ) plays a vital role in ensuring the health and well-being of building occupants. Various factors, including building structures, influence the quality of the indoor environment. This literature review aims to identify the building structural factors that impact indoor environmental quality, factors affecting employees' IEQ satisfaction and the machine learning regression models used for similar research. By examining the existing research, it was allowed to gain insights into the key factors that contribute to a healthier and more sustainable indoor environment.

3.1.1 Research Questions:

The central research questions guiding this literature review are:

1. "What green building structural factors influence the energy efficiency of the building and also impact indoor environmental quality?"
2. "What are the employee satisfaction factors in an office building?"

3.1.2 Selection criteria

The selection criteria for the research study are listed below.

(a) Inclusion and Exclusion Criteria- Studies published between 2010 and 2020 were considered to ensure a comprehensive review. This time frame allows for the incorporation of recent advancements in the field. Additionally, English-language publications were included to maintain consistency and accessibility.

(b) Databases and Search Strategy- The search used multiple databases, including Scopus, Web of Science, and Science Direct. A carefully designed search strategy was employed using relevant keywords and Boolean operators. The selected keywords included "building structures," "indoor environmental quality," "green building," "sustainable buildings," and "structural factors." Combining these keywords using "+" operators, we aimed to retrieve studies encompassing all the specified terms. An example search query is "building structures + indoor environmental quality + green building."
(c) Study Selection- The study selection process followed the PRISMA 2019 and PRISMA 2020 guidelines(PRISMA, 2023). Initially, a database search was conducted, and duplicate studies were removed. Subsequently, a two-step screening process was implemented. In Step 1, titles and abstracts were screened to identify potentially relevant studies. In Step 2, a full-text assessment was performed to determine eligibility based on the inclusion criteria. The number of studies at each stage and the reasons for exclusion were documented. A flow diagram illustrating the study selection process was developed as PRISMA 2019 and PRISMA 2020 recommended.
(d) Data Extraction and Synthesis- Relevant variables and factors related to research questions were identified for data extraction. A structured data extraction form/template was developed to capture the necessary information from the selected studies. The extracted data were then synthesized using a thematic analysis approach. This involved categorizing the data based on common themes or factors and identifying patterns and trends. Appropriate visualization techniques, such as tables and graphs, were utilised to present the synthesized data. Limitations encountered during the data extraction and synthesis process were also discussed.

3.2 Green building selection criteria

Selecting green buildings from different climatic zones in Sri Lanka is a valid and essential approach to this case study as it enables a more comprehensive understanding of how these buildings perform in various environmental conditions. Sri Lanka is known for its diverse climate, ranging from tropical rainforests to dry, arid regions. Including green buildings from different climatic zones can capture the influence of the environment on the performance of these sustainable structures (Su et al., 2021).

One of the primary advantages of this approach is the recognition that green building design and strategies need to adapt to specific climatic conditions. Sri Lanka's climatic zones exhibit temperature, humidity, rainfall patterns, and solar radiation levels variations (Warnasekara et al., 2021). Selecting green buildings from different zones can analyze how these buildings respond to and mitigate the challenges of specific climatic factors.

Green buildings designed for a particular climatic zone may employ different strategies and technologies to enhance energy efficiency, thermal comfort, and overall environmental performance. For instance, facilities in hot and humid regions may focus on effective ventilation systems, shading devices, and passive cooling techniques to reduce the reliance on mechanical cooling. On the other hand, buildings in cooler regions might prioritize insulation, energy-efficient heating systems, and measures to maximise solar gain.

Case studies can identify region-specific best practices and solutions by including green buildings from different climatic zones. It allows for evaluating the effectiveness of various design and technological interventions in specific climatic contexts. This knowledge can then be used to develop targeted strategies and guidelines for green building design and operation in different climatic regions of Sri Lanka.

Furthermore, studying green buildings across diverse climatic zones helps address the issue of generalizability. The performance of sustainable building strategies and technologies can vary significantly depending on the climate in which they are implemented. By selecting buildings from different climatic zones, this case study increases the robustness of the findings. It enhances the ability to draw broader conclusions applicable to various climatic conditions.

In addition to the technical aspects, buildings from different climatic zones also account for the occupant experience. Comfort preferences and satisfaction levels can differ across different climatic regions. Considering structures from diverse climatic zones can assess the impact of green building features on occupant comfort and well-being under other environmental conditions. This knowledge can inform future green building designs, highlighting the importance of localized solutions and occupant-centred design approaches.

It allows for analysing how green buildings perform in varying environmental conditions, provides insights into region-specific design strategies, and contributes to a more nuanced understanding of the relationship between green building practices and occupant comfort. The study advances sustainable building practices tailored to specific climatic contexts by capturing the diversity of Sri Lanka's climate.

Selecting only low-rise buildings with occupied desks/workstations/compartments for this case study is a prudent approach that helps mitigate the potential climatic factor fluctuations associated with different altitudes in high-rise buildings (Mirrahimi et al., 2016). This criterion ensures a more consistent analysis and provides valuable insights into the impact of green building features on employee environmental quality comfort.

High-rise buildings, by their very nature, exhibit variations in temperature, wind patterns, and air pressure as one moves vertically. These variations can significantly affect the indoor environmental conditions experienced by occupants on different floors. Factors such as temperature gradients, airflows, and exposure to sunlight can vary substantially between the ground floor and upper levels of high-rise buildings.

Focusing solely on low-rise buildings will eliminate the confounding factor of altitude-related climatic variations (Verma et al., 2023). This allows for a more controlled analysis of other variables, such as the performance of green building features and their influence on employee comfort. Low-rise buildings can better isolate the effects of specific design and operational strategies implemented within the building envelope and mechanical systems.

Additionally, selecting low-rise buildings with occupied desks/workstations/compartments ensures that the analysis reflects the actual experiences of occupants in their day-to-day work environment. Occupants' comfort and perception of environmental quality vary depending on proximity to windows, natural light availability, and direct exposure to external elements.

Focusing on occupied spaces can evaluate the real-life impact of green building features on the well-being and productivity of employees.

Moreover, low-rise buildings tend to have a more straightforward floor plan and layout than high-rise buildings. This simplicity helps minimise the potential influences from building shape and configuration, reducing confounding effects on the study's outcomes. Selecting low-rise structures with a consistent and rectangular floor plan can better isolate and analyze the impact of other variables, such as indoor air quality, thermal comfort, and lighting conditions.

It's important to acknowledge that high-rise buildings have unique challenges and benefits regarding sustainable design and occupant comfort. However, for the specific goals of this case study, focusing on low-rise buildings with occupied work areas allows for a more targeted investigation into the impact of green building features on employee environmental quality comfort.

Employing this criterion can provide valuable insights into the specific design strategies and operational practices that contribute to improved comfort and well-being in low-rise green buildings. The findings can serve as a basis for informing future building designs, retrofitting projects, and sustainability guidelines for low-rise commercial spaces.

The decision to select buildings with working areas with externally openable windows for testing windows-related hypotheses in this case study is a strategic approach that allows for a focused investigation into the impact of windows on employee environmental quality comfort. This criterion provides an opportunity to evaluate the influence of natural ventilation, daylighting, and outdoor views on occupant satisfaction and well-being.

Externally openable windows offer the advantage of facilitating natural ventilation within a building (H. Zhang et al., 2021). Openable windows can help improve indoor air quality, regulate temperature, and reduce reliance on mechanical ventilation systems by allowing fresh air to enter and circulate through the workspace. This criterion enables the study to assess the effectiveness of natural ventilation strategies in creating a comfortable and healthy work environment.

In addition to ventilation benefits, openable windows also play a crucial role in providing access to daylight and views of the outdoor environment. Daylighting has positively impacted occupant well-being, productivity, and satisfaction. It contributes to a more visually comfortable workspace, reduces reliance on artificial lighting, and offers a connection to the natural environment. Selecting buildings with externally openable windows can examine the influence of daylighting on employee comfort and the potential correlations between access to natural light and environmental quality.

Furthermore, the criterion of externally openable windows aligns with the principles of sustainable design and energy efficiency. By utilizing natural ventilation and maximising daylighting, buildings can reduce energy consumption, particularly for cooling, heating, and lighting. This criterion allows the study to explore the energy-saving potential associated with well-designed windows that promote passive cooling and minimise the need for artificial lighting during daylight hours.

Focusing on buildings with working areas with externally openable windows can test specific hypotheses about windows' impact on occupant comfort and environmental quality. For example, it can evaluate the correlation between window size, placement, and orientation regarding indoor thermal and visual comfort. It can also examine the relationship between natural ventilation provided by openable windows and perceived indoor air quality.

The findings from such an analysis can inform building design strategies and guidelines for incorporating windows that enhance employee comfort and well-being. The results can provide evidence-based insights into the benefits of natural ventilation and daylighting in green building design and contribute to developing more sustainable and occupant-centric work environments.

However, it is essential to note that not all buildings may have externally openable windows due to various factors such as architectural design, building codes, and climate considerations. Considering this criterion in the context of the availability and feasibility of openable windows in the selected buildings is essential.

In this case, study, selecting office spaces with a rectangular floor plan is a strategic choice to avoid potential influences due to building shape. This criterion allows for a more focused analysis of other variables related to employee environmental quality comfort and minimises confounding factors associated with irregular or unconventional building layouts.

A rectangular floor plan offers certain advantages in terms of space utilization and flexibility in interior design. Selecting office spaces with this shape creates a more standardized and comparable setting for evaluating the impact of other factors on occupant comfort (Meena et al., 2022). This allows for a more controlled analysis and facilitates meaningful spatial comparisons.

One advantage of a rectangular floor plan is its inherent simplicity. It provides a regular and uniform layout that minimises variations in room size, configuration, and circulation patterns. This consistency reduces potential biases and allows for a more accurate assessment of the effects of other design and operational variables on employee comfort and well-being.

Furthermore, a rectangular floor plan helps avoid the potential influences of unusual building shapes on airflow patterns, thermal distribution, and lighting conditions. Irregularly shaped buildings can introduce complexities regarding air circulation, creating areas of stagnant air or poor ventilation. Similarly, non-rectangular spaces may have limited daylight penetration, resulting in variations in lighting levels across the workspace. Focusing on rectangular floor plans mitigates these potential confounding factors and ensures a more reliable analysis of other variables.

Additionally, a rectangular floor plan offers practical benefits regarding furniture arrangement, space organization, and workflow efficiency. The regular shape facilitates the configuration of workstations, desks, and partitions logically and functionally. It allows for more straightforward circulation paths and minimises wasted space, enhancing the overall functionality and productivity of the office environment.

By selecting office spaces with a rectangular floor plan, this case study can better focus on variables directly related to employee environmental quality comfort. For example, it can evaluate the impact of factors such as indoor air quality, thermal comfort, acoustic performance, and lighting conditions without the potential interference of irregular building layouts. This targeted approach allows a more accurate understanding of how these variables contribute to occupant comfort and well-being in green buildings.

However, it is essential to acknowledge that real-world office spaces come in various shapes and configurations. While selecting rooms with a rectangular floor plan helps streamline the analysis, it is crucial to consider the limitations of this criterion. The findings should be interpreted within the context of the specific building shape and generalizability to other building types.

The decision to consider companies that made wearing company uniforms mandatory for employees in this study is a thoughtful approach aimed at neutralizing the impact of clothing choice on employee environmental quality comfort. This criterion helps control the variability in clothing rate and allows for a more focused analysis of other factors influencing occupant comfort in green buildings.

Clothing choice can significantly impact an individual's thermal comfort perception. Different clothing materials, styles, and layers can affect heat transfer, insulation, and moisture management (Ullah et al., 2021). By selecting companies with mandatory uniforms, minimise the influence of individual clothing choices on the study's outcomes and create a more consistent baseline for evaluating the effects of other variables.

Mandatory company uniforms ensure a standardized clothing rate among employees within the selected companies. This means that the thermal properties of the clothing employees wear are relatively uniform and consistent. Controlling for clothing variability can better isolate the effects of other factors, such as indoor air temperature, humidity, and ventilation, on occupant comfort.

Additionally, considering companies with mandatory uniforms helps reduce the potential bias introduced by individual preferences and clothing habits. Employees may have varying comfort preferences and clothing practices, which can influence their perception of thermal comfort. Focusing on mandatory uniforms minimises these subjective influences and creates a more controlled environment for evaluating the impact of other factors on occupant comfort.

Moreover, this criterion enables examining the impact of other design and operational strategies on occupant comfort without interfering with clothing-related variables. For example, it can evaluate the effectiveness of thermal comfort strategies such as passive heating or cooling systems, insulation, and air conditioning without the confounding effects of varying clothing types and layers.

By neutralizing the impact of clothing rates by selecting companies with mandatory uniforms, the study can provide valuable insights into the design and operation of green buildings for occupant comfort. The findings can help inform strategies to enhance thermal comfort, optimise HVAC systems, and create more comfortable indoor environments for employees.

It is essential to acknowledge that the criterion of mandatory uniforms may not apply to all types of companies or work environments. Some industries or job roles may have specific clothing requirements essential for safety, hygiene, or professional appearance. In such cases, it is necessary to consider alternative approaches to controlling clothing variability, such as providing standardized clothing options for study participants.

Based on the above factors, the building selection criteria can be summarized as follows,

(a) The Buildings should be at least "certified" and awarded under LEED or GBCSL criteria.
(b) The green buildings were selected from every climatic zone to represent all the green office buildings in Sri Lanka.
(c) Only Low-rise buildings were selected with occupied desks/workstations/compartments to avoid climatic factor fluctuations due to different altitudes in high-rise buildings.
(d) Buildings with working areas with externally openable windows were selected to test the windows-related hypothesis.
(e) Office spaces with a rectangular floor plan were selected to avoid possible influences due to the building's shape.
(f) Companies that made wearing company uniforms mandatory for employees were considered for the study to neutralise the impact of clothing rate.

According to the LEED USA and GBCSL 2020 database, there were 72 "Certified' or above-rated office green buildings in Sri Lanka. Out of 72 buildings, 08 office spaces and 06 factory spaces were selected for the study (Figure 3.2). Figure 3.3 illustrates the distribution of green office buildings in Sri Lanka, and the map in the right corner depicts the buildings selected for the research.

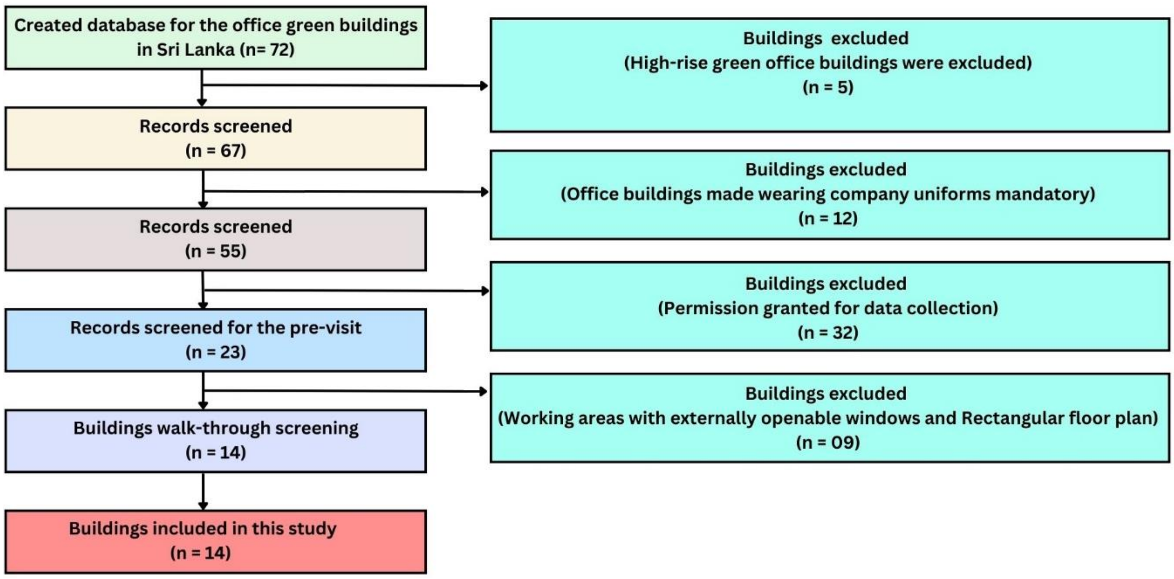


Figure 3.2: Green office buildings selection process for the study

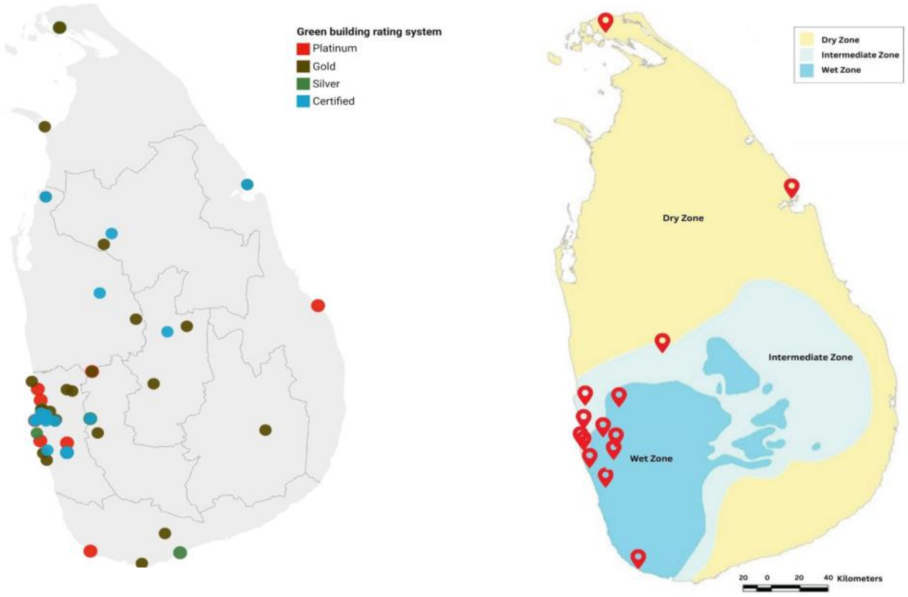


Figure 3.3: Green office buildings in Sri Lanka (right) and the selected green office buildings (left)

3.3 Employee satisfaction evaluation: employee sample selection

The sample selection process in the study employed a combination of stratified random and random sampling techniques to ensure a representative and diverse set of office spaces and factories were included. The objective was to capture a range of environmental conditions and factors that could potentially impact employee comfort and indoor environmental quality in green buildings. In the case of office spaces, a stratified sampling approach was adopted based on the distance from the windows. This zoning strategy was implemented to account for the potential variation in environmental conditions across different zones within the office spaces (Figure 3.4). The rationale behind this approach is that proximity to windows can influence

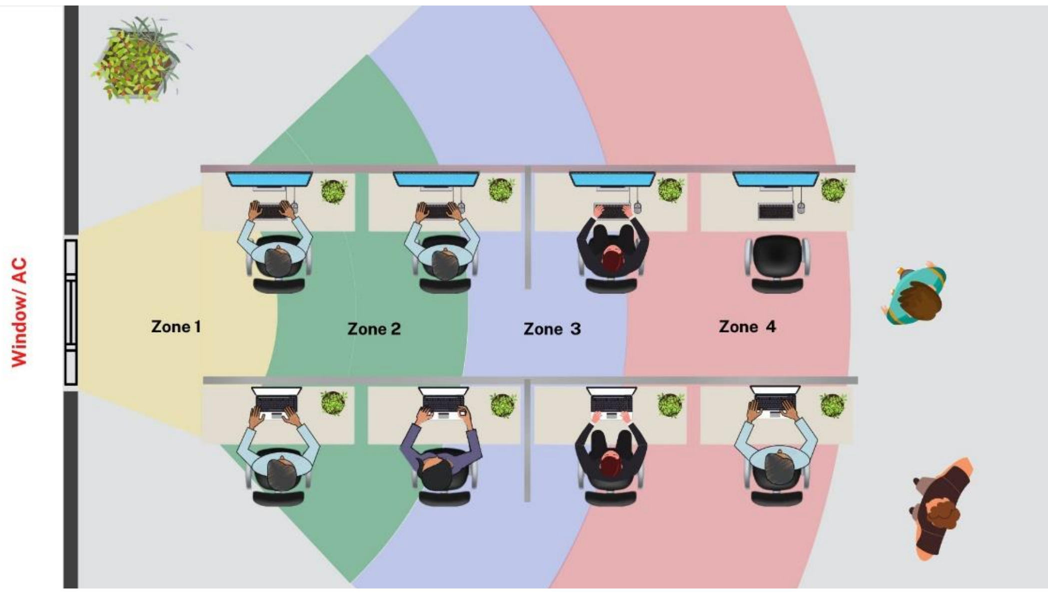


Figure 3.4: The zoning layout of the general office buildings

factors such as natural daylighting, views, and ventilation, which in turn can impact the employee's perception of environmental comfort. The study aimed to capture a range of conditions and assess their potential effects on employee comfort by stratifying the office spaces based on window distance.

On the other hand, factories were divided into four equal zones (Figure 3.5), and each central zone was again divided into four zones to cover all corners of the premises, stratified based on centralized HVAC vents. Unlike office spaces, factories often have different architectural layouts and operational characteristics that may result in limited access to operable windows or reduced attention to window-related factors. Therefore, dividing factories into equal zones ensured a fair representation of different areas within the factory spaces. This approach aimed to capture the variability in environmental conditions across the factory floor and assess their impact on employee comfort.

Stratified random sampling is a valuable technique in research as it allows for selecting samples from different strata or subgroups within a population. The study obtained a comprehensive understanding of the impact of other environmental conditions on employee comfort by including samples from various zones based on window distance in office spaces. This approach also helps control for potential confounding factors and improves the generalizability of the findings to the broader population of green office spaces. Similarly, random sampling was used to select samples within each stratum or zone. Random selection ensures that each unit within the stratum has an equal chance of being selected, reducing bias and increasing the sample's representativeness. By employing both stratified random and random sampling techniques, the study minimised the potential for selection bias and enhanced the statistical validity of the findings.

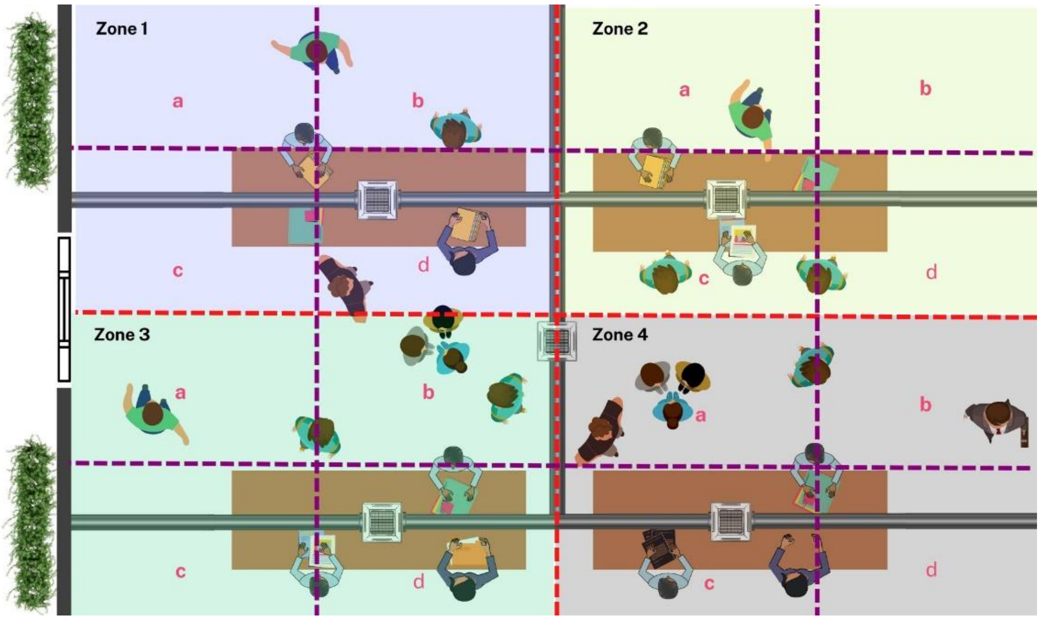


Figure 3.5: The zoning layout of the factory buildings

Overall, the sample selection methodology employed in the study demonstrates a rigorous and systematic approach to ensure a diverse representation of office spaces and factories. The stratified random sampling based on window distance in office spaces and equal zone division in factories allows for investigating various environmental conditions and their impact on employee comfort. Using these sampling techniques, the study aimed to provide valuable insights into the relationship between green buildings, employee comfort, and indoor environmental quality, contributing to advancing knowledge in this field.

The present study employed different sampling strategies to ensure the representativeness and generalizability of the findings. The aim was to minimise sampling error while considering the total number of employees in factory and office spaces. The summary of the sampling process is depicted in Figure 3.6.

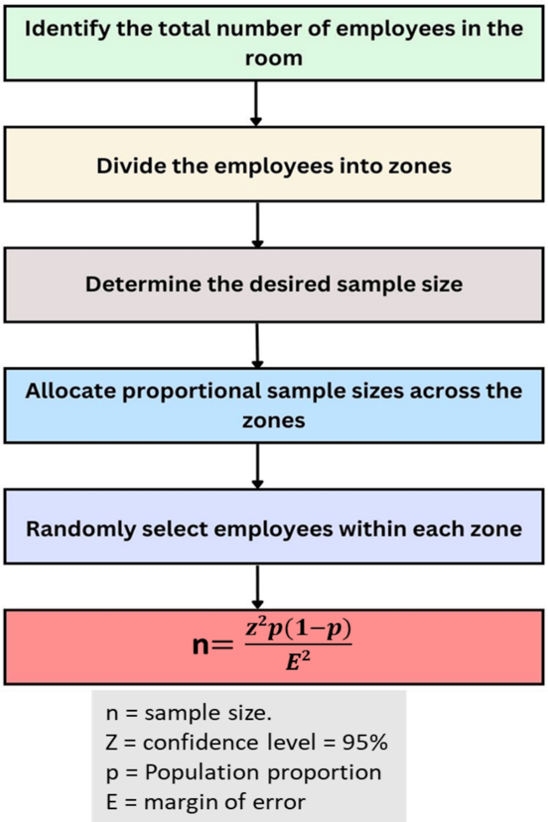


Figure 3.6: The employee sampling process

3.4 Study assumptions

Assumptions play a crucial role in any research study as they provide a foundation for the methodology, data analysis, and interpretation of results. In the context of this study on green buildings, several key assumptions were made to guide the research process and ensure the validity and reliability of the findings. The following assumptions were considered:

(a) The difference in outside weather parameters of the selected locations of the green buildings was negligible:
One of the assumptions made in this study is that the outside weather parameters, such as temperature, humidity, and solar radiation, among the selected locations of the green buildings were relatively similar. This assumption is crucial because variations in weather conditions could significantly impact the performance and perception of green buildings. By assuming negligible differences, the study aims to isolate the effects of building design and features on occupant satisfaction and energy efficiency.

(b) The average metabolic rates of all female and male employees were assumed to be equivalent:
To simplify the analysis and interpretation of the data, it was assumed that the average metabolic rates of all female and male employees were equivalent. Although there may be slight variations in metabolic rates between genders, this assumption allows for a more straightforward data comparison and eliminates potential confounding factors related to different energy requirements. However, it is essential to acknowledge that individual variations in metabolic rates may exist within the sample population, and this assumption should be considered when interpreting the results.

(c) Slight bias in employee response data would not significantly influence the results and analysis:
Every survey or self-report study is susceptible to response bias, where participants may provide answers influenced by personal preferences or subjective interpretations. In this study, a slight bias in employee response data was acknowledged. However, this bias was assumed to not significantly influence the overall results and analysis. While efforts were made to minimise bias through careful survey design and data collection methods, the assumption helps ensure that any potential bias does not compromise the validity of the findings.

(d) New employees with less than one year of experience in the building were not included in the sample:
New employees who had been in the building for less than one year were excluded from the sample of respondents. This decision was based on the assumption that the initial period after occupancy can lead to a "Honeymoon effect," where employees may perceive the building more positively due to novelty or unfamiliarity. By excluding new employees, the study focuses on long-term experiences and avoids potential biases associated with the early occupancy stages.

(e) The uniform material of the employees is the same:
Another assumption made in this study is that the uniform material worn by the employees is consistent across the sample population. This assumption ensures that factors related to clothing material and its potential influence on thermal comfort and perception are standardized. By assuming uniform material, the study aims to isolate the effects of building design and features on occupant satisfaction without confounding clothing-related variables.

(f) Increasing the window head height will increase the depth of helpful daylight penetration:
In this study, it was assumed that increasing the window head height would result in deeper penetration of helpful daylight into the indoor space. This assumption is based on

established principles of daylighting design, where higher window openings allow for a more significant influx of natural light, potentially enhancing occupant well-being and reducing reliance on artificial lighting. By assuming this relationship, the study explores the potential benefits of increased window head heights in green buildings.

(g) The employees are aligned with the ISO2004 definition as standard occupants:
The study assumes that the employees in the sample population align with the ISO2004 definition of standard occupants. This definition considers various anthropometric characteristics such as height, weight, body surface area, and basal metabolic rate. By assuming these standard values for male and female employees, the study aims to provide a consistent framework for analyzing and interpreting the data, ensuring comparability across different individuals within the sample population.

(h) Occupant behaviour and usage patterns remain consistent:
Assuming that occupants' behaviour and usage patterns in the selected green buildings remain consistent throughout the study helps minimise the confounding effects of varying occupant activities on the study outcomes. It allows for a more accurate evaluation of the impact of the building design and features on occupant satisfaction and energy efficiency. However, it is essential to acknowledge that individual behaviours may still vary to some extent, and this assumption should be considered in light of any potential changes in occupant behaviour over time.

(i) Building management practices are consistent:
Assuming consistent building management practices across the selected green buildings helps ensure that any differences observed in occupant experiences and energy performance can be primarily attributed to the building design. Consistency in temperature control, ventilation strategies, and maintenance procedures reduces the potential influence of management practices on occupant satisfaction and energy efficiency ratings. However, it is essential to acknowledge that minor variations in management practices may still exist, and their potential impact should be considered.

(j) Building certifications reflect accurate performance:
Assuming that the green building certifications, such as LEED or BREEAM, accurately reflect the performance and sustainability features of the selected buildings is crucial for validating the study's findings. Green building certifications provide a standardized framework for evaluating and recognizing sustainable building design. By assuming the accuracy of these certifications, the analysis can align its evaluation with widely recognized industry standards and ensure that the findings align with the intended goals and benchmarks of green building certification programs.

(k) Occupants are aware of the green building features:
Assuming that occupants in the selected buildings have sufficient awareness and knowledge about the green building features and their potential benefits acknowledges that occupant perceptions and responses may be influenced by their understanding of the building's sustainable design elements. It is essential to consider that occupant awareness levels may vary, and this assumption can be evaluated through survey responses or interviews to ensure that participants have a reasonable understanding of the green features being assessed.

(l) Sample participants are representative of the overall population:
Assuming that the selected sample participants in offices and factories represent the broader population within those building types enhances the generalizability of the study findings. Ensuring that the sample is diverse and adequately represents the occupants' demographic range, job roles, and work activities is vital. This assumption allows for broader implications and applicability of the study results to similar office and factory settings.

(m) Data collection instruments are reliable and valid:
Assuming that the survey questionnaires, interviews, or other data collection instruments used in the study are reliable and valid is essential for obtaining accurate and meaningful data. Reliability ensures that the instruments consistently measure the intended constructs, while validity ensures they measure what they are supposed to measure. Pilot testing and validation of the data collection instruments should be conducted to establish their reliability and validity before implementing them in the actual study.

(n) Green building features are effectively implemented and operational:
Assuming that the green building features and systems, such as energy-efficient HVAC, lighting controls, and sustainable materials, are effectively implemented and operational in the selected buildings is crucial for evaluating their impact. It assumes that these features' intended benefits and effects can be observed and measured accurately. However, it is essential to consider the potential variability in implementation quality and operational performance, which may influence the study's outcomes. This assumption can be supported by verifying the building systems' performance through maintenance records or commissioning reports.

These assumptions provide a basis for the design and execution of the study on green buildings and occupant experiences. They help guide the selection of variables, data collection methods, and data analysis techniques. While these assumptions are reasonable and informed by existing literature and best practices, it is essential to acknowledge their limitations and potential impact on the study's findings.

3.5 Developing machine learning predictive model

This section presents a detailed overview of developing a predictive model for IEQ satisfaction using Python programming. The model is constructed based on the code provided, encompassing various data preprocessing techniques, model training, evaluation, and performance assessment. This research outlines the steps taken to build the model, explores the implemented coding architecture, and discusses the significance of the chosen approaches in the context of the research objective. The predictive models were created separately for Thermal comfort, Visual comfort and IEQ satisfaction. A separate model was created for the overall IEQ satisfaction, considering the overall satisfaction of employees towards the IEQ of their workplace.

The development of the predictive model involves the utilization of multiple Python libraries and modules. These libraries are crucial for data manipulation, analysis, visualization, and implementation of machine learning. Notably, the code imports libraries such as `numpy`, `pandas`, `seaborn`, `matplotlib`, `scipy`, and `sklearn` to facilitate the functionalities required for model construction and evaluation.

The data used for the model is imported from an Excel file using the `pandas` library, and preliminary exploratory analysis is conducted to gain insights into the dataset's structure and characteristics. Following this, data preprocessing techniques are employed to handle non-numerical columns, remove duplicate columns, and identify descriptive columns of interest. These steps ensure that the data is appropriately prepared for subsequent modelling tasks.

The model development process employs various regression algorithms, including Support Vector Regression (SVR), Lasso Regression, Decision Tree Regression, and Random Forest Regression. Each algorithm is evaluated using appropriate performance metrics, such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), which provide insights into the model's accuracy and predictive power.

Furthermore, cross-validation techniques assess the models' generalizability and robustness. The mean MAE scores obtained from cross-validation aid in comparing and selecting the most suitable regression model for the given task.

Hyperparameter tuning is performed to optimise the Random Forest Regression model using a grid search approach. This step aims to find the best combination of hyperparameters that maximise the model's performance. The best model, along with the optimised parameters, is reported.

The Variance Inflation Factor (VIF) is calculated to assess potential multicollinearity among the features used in the model. This analysis helps identify if any variables in the dataset are highly correlated, which may affect the model's interpretability and stability.

Additionally, the developed predictive model is serialized using the pickle module in Python. The serialized model is stored in a "Thermal.pickle" file to design the user interface/form.

Following this comprehensive approach, the developed predictive model demonstrates its ability to IEQ based on the provided dataset. The subsequent sections will delve into each step of the model development process, presenting detailed analyses, discussing the findings, and providing insights into the model's performance and applicability in the context of the research objectives.The pickle. dump() function is utilised to serialize the best_model object and save it in the specified file. Serialization allows the model to be stored persistently, enabling its later retrieval and usage without retraining.

The serialized model can be loaded for predictions in other Python scripts or applications. Incorporating this code, the predictive model is developed, evaluated, and made readily available for future use, ensuring its practicality and scalability in real-world scenarios. The model infrastructure is shown in Figure 3.7.

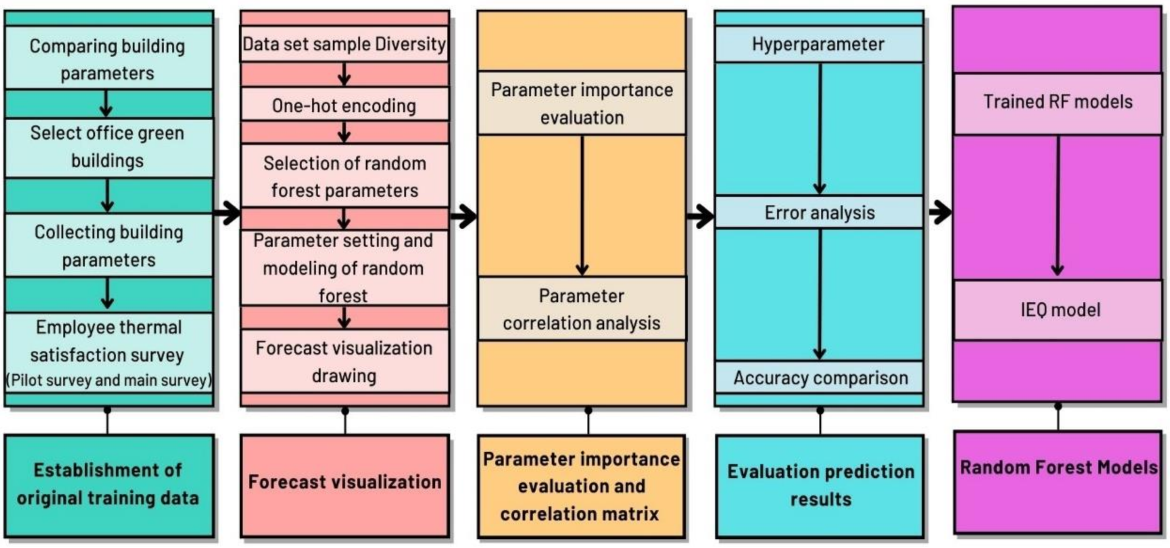


Figure 3.7: The model infrastructure

The following libraries were imported to support the development and analysis:

(a) \`numpy\` (version 1.23.0): This library supports large, multi-dimensional arrays and matrices and a collection of mathematical functions to operate on these arrays efficiently.
(b) \`pandas\` (version 1.4.3): It is a powerful data manipulation and analysis library. It provides data structures and functions for efficiently handling structured data, including various operations such as filtering, grouping, and merging.
(c) \`seaborn\` (version 0.12.2): This library is built on top of \`matplotlib\` and offers a high-level interface for statistical data visualization. It provides a variety of aesthetically pleasing and informative visualizations to analyze relationships and patterns in the data.
(d) \`matplotlib\` (version 3.5.2): This library is widely used for visualizations in Python. It provides a flexible framework for generating plots, charts, histograms, and other graphical representations.
(e) \`scipy.stats\` (version 1.8.1): This module within the SciPy library provides various statistical functions and distributions. In this particular case, this module's \`norm\` function is imported to work with normal distribution statistics.

For data preprocessing and transformation, the following modules were imported:

(a) \`LabelEncoder\` from \`sklearn.preprocessing\` (version 1.2.2): This class converts non-numerical columns into numerical ones. It assigns a unique numeric label to each category, allowing machine learning algorithms to handle categorical data effectively.
(b) \`MinMaxScaler\` from \`sklearn.preprocessing\` (version 1.2.2): This class is utilised for feature scaling, ensuring all features are on a similar scale. Scaling is important for specific machine learning algorithms to prevent any particular feature from dominating the learning process due to its magnitude.

For model evaluation and analysis, the following modules and classes were imported:

(a) \`accuracy_score\`, \`confusion_matrix\`, and \`r2_score\` from \`sklearn.metrics\` (version 1.2.2): These functions provide metrics to assess the performance of the predictive model.

`accuracy_score` calculates the accuracy of classification models, `confusion_matrix` creates a table to evaluate classification results, and `r2_score` computes the coefficient of determination for regression models.

(b) `DecisionTreeClassifier` from `sklearn.tree` (version 1.2.2): This class implements decision tree algorithms for classification tasks. Decision trees are widely used due to their interpretability and ability to handle numerical and categorical data.

(c) `export_graphviz` from `sklearn.tree` (version 1.2.2): This function exports decision tree models in Graphviz format. It allows visualizing and analyzing the decision tree structure.

(d) `train_test_split` from `sklearn.model_selection` (version 1.2.2): This function facilitates splitting the dataset into training and testing sets. It is essential to have separate datasets for model training and evaluation.

Lastly, the `warnings` module was imported to suppress warning messages during the execution of the code.

In addition to these libraries, the `%matplotlib inline` command was used to ensure that plots are displayed directly within the Jupyter Notebook or IPython environment for seamless visualization. These imported libraries provide the necessary tools and functions to develop and evaluate the predictive model.

3.6 Chapter summary

The study aims to analyze the impact of green building factors on employee satisfaction with indoor environmental quality (IEQ). It involves a systematic literature review to identify key factors related to IEQ, employee satisfaction factors, and machine learning regression models. Data collection includes an employee satisfaction survey and physical measurements of building parameters. A user interface is developed to assess employee satisfaction with different IEQ factors. The study focuses on selecting green buildings from diverse climatic zones in Sri Lanka to understand their performance under different environmental conditions. Low-rise buildings with occupied work areas and externally openable windows are selected to ensure consistent analysis and evaluate specific hypotheses. The research analyses variables related to employee environmental quality comfort in green buildings by selecting office spaces with rectangular floor plans. This approach allows for a targeted evaluation of factors such as indoor air quality, thermal comfort, and lighting conditions without the interference of irregular building layouts. The study also considers companies that enforce mandatory uniforms for employees to neutralize the impact of clothing choices on occupant comfort. The building selection criteria include choosing "certified" green buildings from different climatic zones in Sri Lanka, low-rise buildings with occupied workstations, spaces with externally openable windows, and those with a rectangular floor plan. The aim is to represent a diverse range of green office spaces.

The sample selection process involves stratified random and random sampling techniques. Office spaces are stratified based on window distance to capture various environmental conditions, while factories are divided into zones based on centralized HVAC vents. These sampling techniques ensure a representative and diverse sample of office spaces and factories.

Assumptions are made in the study to guide the research process. These assumptions include negligible differences in outside weather parameters, equivalent average metabolic rates for male and female employees, minimal bias influence on data, exclusion of new employees in the sample, consistent, uniform material, increased window head height leading to deeper daylight penetration, standard occupants as per ISO2004 definition, consistent occupant behaviour and building management practices, accurate building certifications, occupant awareness of green features, representative sample participants, reliable data collection instruments, effective implementation of green building features, and consistent occupant behaviour and usage patterns.

The study also includes the development of a machine learning predictive model for employee satisfaction with indoor environmental quality (IEQ). The model uses Python programming and involves data preprocessing, model training, evaluation, and performance assessment. Various regression algorithms are utilised, and performance metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are employed for model evaluation. Cross-validation techniques and hyperparameter tuning are used to enhance model generalizability and performance.

4 RESULTS AND DISCUSSION

The results and discussion include the systematic reviews, the parameters of selecting buildings, the employee survey process and results, the descriptive analysis, the machine learning predictive model and the user interface.

4.1 The systematic literature review and bibliographic analysis

The identified factors and the behaviour of the identified factors will be widely discussed in this chapter.

4.1.1 Green building structural factors influencing indoor environmental quality

The literature proved that IEQ refers to the overall quality of the indoor environment, including factors such as air quality, thermal comfort, lighting, acoustics, and spatial layout. Among these factors, the spatial layout mainly depends on the interior design of the building. It was decided to omit the spatial layout and continue the keyword search. The factors were primarily divided into three main categories which are, Thermal comfort, Lighting and visual comfort, Air quality.

According to the identified main IEQ comfort categories, the keyword search criteria were as follows,

(a) "thermal comfort", "building structure"
(b) "visual comfort", "building structure"
(c) "air quality", "building structure"

The search results were narrowed down using the following procedure; Initial Search Keywords are "Building", "IEQ", and "Occupants" Search Years are 2011-2021, and the search sources are Google Scholar, Web of Science, Scopus and ScienceDirect. The Bibliographic analysis steps is summarised in Figure 4.1 The following factors mentioned in Table 4.1 were identified as the building structural parameters influencing indoor environmental quality.

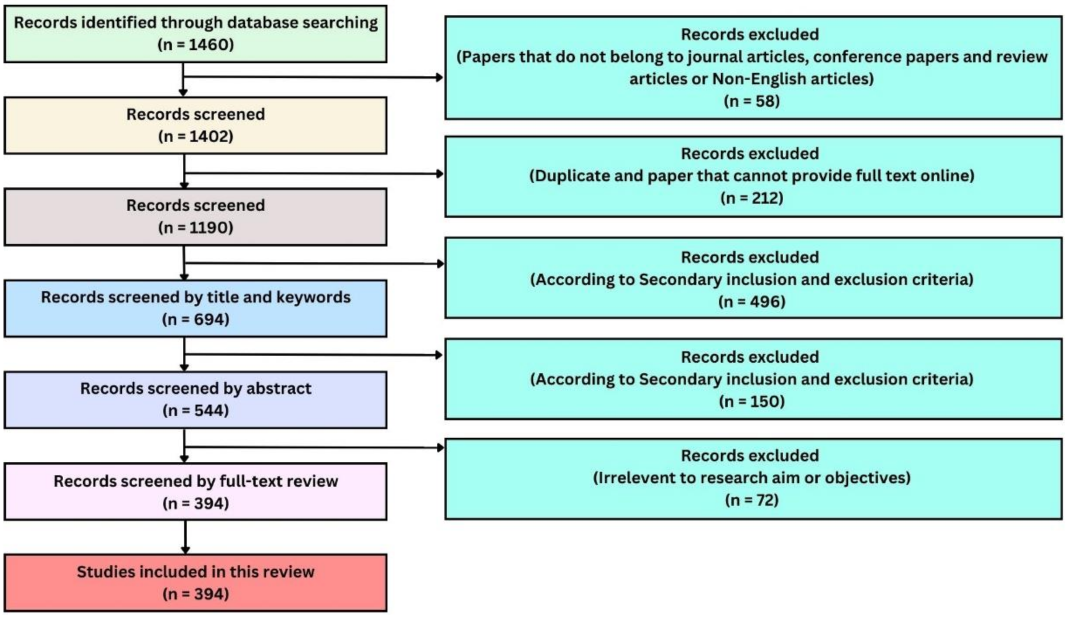


Figure 4.1: The steps conducted for bibliographic analysis to identify building factors

Table 4.1: Identified factors affecting employee comfort in office buildings

Employee comfort factor	Literature
Thermal Comfort	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Brown et al., 2010; Huizenga et al., 2005; Issa et al., 2011; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Leaman & Bordass, 2007; Y. S. Lee & Kim, 2008; Liang et al., 2014; B. Lin et al., 2016; Newsham et al., 2013; Paul & Taylor, 2008; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016; Y. Zhang & Altan, 2011)
Visual Comfort	(Altomon te et al., 2016; Altomonte & Schiavon, 2013; Lee & Kim, 2008; Liang et al., 2014; Lin et al., 2016; Newsham et al., 2013; Paul & Taylor, 2008; Pei et al., 2015; Zhang & Altan, 2011)
Indoor Air Quality	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Brown et al., 2010; Huizenga et al., 2005; Issa et al., 2011; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; Liang et al., 2014; B. Lin et al., 2016; Paul & Taylor, 2008; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016)
Acoustic comfort	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Brown et al., 2010;

	Huizenga et al., 2005; Issa et al., 2011; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Leaman & Bordass, 2007; Y. S. Lee & Kim, 2008; Liang et al., 2014; B. Lin et al., 2016; Newsham et al., 2013; Paul & Taylor, 2008; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016; Y. Zhang & Altan, 2011)
Illumination	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Brown et al., 2010; Huizenga et al., 2005; Issa et al., 2011; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Leaman & Bordass, 2007; Y. S. Lee & Kim, 2008; Liang et al., 2014; B. Lin et al., 2016; Paul & Taylor, 2008; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016; Y. Zhang & Altan, 2011)
Office Furnishings	(Abbaszadeh et al., 2006; Baird et al., 2012; Issa et al., 2011; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008)
Furnishing adjustability	(Altomonte & Schiavon, 2013; Baird et al., 2012; Y. S. Lee & Kim, 2008; Ravindu et al., 2015)
Ease of interaction	(Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Y. S. Lee & Kim, 2008; B. Lin et al., 2016; Pei et al., 2015)
Building cleanliness	(Abbaszadeh et al., 2006; Altomonte & Schiavon, 2013; Baird et al., 2012; Issa et al., 2011; Y. S. Lee & Kim, 2008)
General Workspace	(Altomonte et al., 2016; Baird et al., 2012; Huizenga et al., 2005; Issa et al., 2011)
Visual privacy	(Altomonte et al., 2016; Altomonte & Schiavon, 2013; Y. S. Lee & Kim, 2008; B. Lin et al., 2016; Pei et al., 2015; Ravindu et al., 2015)
Office Layout	(Abbaszadeh et al., 2006; Huizenga et al., 2005; Issa et al., 2011; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; Ravindu et al., 2015; Sediso & Lee, 2016)
Workstation’s distance from a window	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Khoshbakht, Gou, Xie, et al., 2018; Y. S. Lee & Kim, 2008; Newsham et al., 2013; Paul & Taylor, 2008; Sediso & Lee, 2016)
Ventilation	(Baird et al., 2012; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Leaman & Bordass, 2007; Paul & Taylor, 2008; Ravindu et al., 2015)
Gender	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Huizenga et al., 2005; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; Liang et al., 2014; Newsham et al.,

	2013; Paul & Taylor, 2008; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016)
Age	(Abbaszadeh et al., 2006; Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Huizenga et al., 2005; Khoshbakht, Gou, Xie, et al., 2018; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; Liang et al., 2014; Newsham et al., 2013; Paul & Taylor, 2008; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016)
Duration of working in the building	(Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Khoshbakht, Gou, Xie, et al., 2018; Y. S. Lee & Kim, 2008; Ravindu et al., 2015; Sediso & Lee, 2016)
Mood	(Brown et al., 2010; Khoshbakht, Gou, Xie, et al., 2018)
Cleanliness	(Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Huizenga et al., 2005; Khoshbakht, Gou, Xie, et al., 2018; B. Lin et al., 2016; Pei et al., 2015; Ravindu et al., 2015; Sediso & Lee, 2016)
Workplace colours and textures	(Altomonte et al., 2016; Altomonte & Schiavon, 2013; Baird et al., 2012; Issa et al., 2011; S.-K. Kim et al., 2015; Y. S. Lee & Kim, 2008; B. Lin et al., 2016; Paul & Taylor, 2008; Pei et al., 2015)

Thermal comfort is critical to indoor environments, influencing building occupants' well-being, productivity, and satisfaction. Achieving optimal thermal comfort requires understanding the various factors that contribute to it. This article comprehensively analyses the building factors influencing thermal comfort, including temperature, humidity, air velocity, radiant temperature, insulation and building envelope, HVAC systems, solar heat gain, occupant density, thermal zoning, and building orientation and location.

(a) Temperature : Temperature is perhaps the most apparent and influential factor affecting thermal comfort. The ambient temperature should be maintained within a range that suits the majority of occupants. The recommended range typically falls between 20-26 degrees Celsius (68-79 degrees Fahrenheit), although individual preferences may vary (Franco et al., 2021). Deviations from this range can lead to discomfort, affecting productivity and overall satisfaction.

- Radiant Temperature: Radiant heat exchange between people and surfaces, such as walls, windows, and furniture, affects thermal comfort (C. Zhang et al., 2020). Cold surfaces can lead to discomfort, while warm surfaces provide a sense of cosiness. Optimising the design and selection of materials for building elements can help achieve a balanced, radiant temperature that contributes to overall comfort.

(b) Humidity: Humidity refers to the level of moisture present in the air. It plays a significant role in thermal comfort, as high humidity can make the environment feel warmer, sticky, and uncomfortable, while low humidity can lead to dryness and discomfort.

Maintaining an optimal humidity level, typically around 40-60% (Hsu et al., 2021), contributes to a more pleasant and comfortable indoor environment.

(c) Air Velocity : Air movement, whether through natural ventilation or mechanical systems, influences thermal comfort. A gentle breeze can enhance comfort by promoting evaporative cooling and aiding in moisture evaporation from the skin (Fang et al., 2021). However, excessive air velocity or drafts can create discomfort and should be avoided. Proper air distribution and control are crucial in maintaining a comfortable environment.

(d) Insulation and Building Envelope: The insulation properties and quality of the building envelope significantly impact thermal comfort. Adequate insulation helps minimise heat transfer through walls, roofs, and floors, reducing temperature fluctuations and heat loss/gain (Danaci & Akin, 2022). Well-designed windows with appropriate glazing also play a crucial role in preventing thermal discomfort. Minimizing heat transfer, insulation, and the building envelope contributes to maintaining a stable and comfortable indoor environment.

- Wall insulation U value: The Wall insulation U value is an essential parameter for evaluating building insulation's thermal performance and energy efficiency. It measures the thermal conductivity of the insulation material, indicating its ability to resist heat transfer through the walls.

  The measurement formula for the Wall insulation U value is as follows: The measurement formula for the Glazing U value is as follows, R = The gas constant at the T temperature W = Watts per square meter, K= Temperature in Kelvin, m = meter.

$$U = \frac{1}{R_T} = \frac{W}{m^2 * K} \tag{1}$$

The unit of measurement for the Wall insulation U value is watts per square meter per Kelvin (W/m^2K). It represents the amount of heat energy that can transfer through one square meter of the insulation material per degree of temperature difference.

Building maintenance information and literature are commonly used as measurement sources to determine the Wall insulation U value. This information provides data on the insulation materials used in the walls, including their thermal properties and conductivity. By referring to building maintenance records and relevant literature, professionals can obtain the necessary information to assess the U value of the wall insulation.

The Wall insulation U value is a critical factor in evaluating the effectiveness of insulation materials in reducing heat transfer. A lower U value indicates better insulation properties, meaning the insulation material provides higher resistance to heat flow (Akkurt et al., 2020). This helps to minimise heat loss during colder months and heat gain during hotter months, contributing to improved energy efficiency and occupant comfort.

By selecting insulation materials with lower U values, architects and engineers can enhance the overall thermal performance of the building envelope. This reduces the reliance on heating and cooling systems, leading to energy savings and a more comfortable indoor environment.

Considering the Wall insulation U value in conjunction with other factors such as the insulation thickness and building envelope design is essential. Combining proper insulation materials, adequate thickness, and effective installation techniques helps achieve optimal thermal performance.

Furthermore, building codes and energy efficiency standards often specify maximum U values for wall insulation to ensure compliance with energy performance requirements. Buildings can demonstrate their commitment to energy efficiency and sustainability by meeting or exceeding these standards.

Some common types of wall insulation materials and their corresponding U values are summarized in Table 4.2

Table 4.2: Wall insulation materials that are commonly used

Insulation Type	**U Value (W/m^2K)**	**Special Characteristics**	**References**
Fibreglass Insulation	0.25 - 0.5	Cost-effective and widely available insulation material	(Dong et al., 2023; Roque & Santos, 2017; Cabeza et al., 2010 ; D. Kumar et al., 2020; Schiavoni et al., 2016; Aditya et al., 2017;Wang et al., 2018; Deshmukh et al., 2017)
Cellulose Insulation	0.035 - 0.06	Made from recycled materials, eco-friendly	
Mineral Wool Insulation	0.035 - 0.045	Fire-resistant and provides good sound insulation	
Polyurethane (PUR) Insulation	0.02 - 0.035	High thermal resistance and good moisture resistance	
Polyisocyanurate (PIR) Insulation	0.02 - 0.035	Excellent thermal performance and fire resistance	
Expanded Polystyrene (EPS) Insulation	0.03 - 0.05	Lightweight, rigid, and resistant to moisture	
Extruded Polystyrene (XPS) Insulation	0.025 - 0.035	High compressive strength and moisture resistance	
Vacuum Insulation Panels (VIP)	As low as 0.005	Exceptional thermal performance with minimal thickness	
Aerogel Insulation	As low as 0.015	Superb thermal insulator, lightweight and thin	

- The thickness of wall insulation: The Thickness of wall insulation is a crucial parameter that measures the average thickness of the insulation material applied to the walls of a building. It is essential in determining the wall assembly's insulation properties and thermal performance.

  The Thickness of wall insulation is typically measured through physical measurements, where professionals directly measure the average width of the insulation material. The unit of measurement for this parameter is millimetres (mm). It directly affects the insulation properties of the material (Akkurt et al., 2020). A thicker insulation layer provides higher thermal resistance, reducing heat transfer through the walls. It helps to create a more stable and comfortable indoor environment by minimizing heat loss during colder months and heat gain during hotter months.

  The selection of the appropriate insulation thickness depends on several factors, including climate conditions, building design, and energy efficiency goals. Climate conditions play a significant role as regions with more extreme temperatures may require thicker insulation to ensure optimal thermal performance. Building design considerations, such as the wall construction type and available space, influence the maximum achievable insulation thickness.

  Building codes and energy efficiency standards often provide guidelines or minimum insulation thickness requirements to ensure compliance with energy performance regulations. Compliance with these standards helps to achieve energy-efficient buildings and contributes to sustainability goals.

(e) Roof insulation materials and U value: Roof insulation is crucial in enhancing buildings' energy efficiency and thermal comfort (Rumiantcev et al., 2016). Various insulation materials are available worldwide, each with different properties that impact their thermal performance, including their U values. Popular wall insulation materials and their properties are summarized in Table 4.3

Table 4.3: Roof insulation materials that are commonly used

Insulation Type	**U Value (W/m^2K)**	**Special Characteristics**	**References**
Fibreglass Insulation	0.2 - 0.4	Affordable, effective insulation material	(Bozsaky, 2010; Aditya et al., 2017; A. Kumar & Suman, 2013; Tariku et al., 2023; Zakaria Salem et al., 2018; Abdou & Budaiwi, 2013; Kalhor &
Cellulose Insulation	0.2 - 0.5	Eco-friendly, made from recycled materials	
Mineral Wool Insulation	0.035 - 0.045	Excellent thermal performance reduces heat transfer	
Polyurethane (PUR) Insulation	0.02 - 0.035	High thermal resistance, commonly used in buildings	

Polyisocyanurate (PIR) Insulation	0.02 - 0.035	Energy-efficient, suitable for residential and commercial buildings	Emaminejad, 2020; Kunič, 2017; Cabeza et al., 2010; Gesa et al., 2014)
Expanded Polystyrene (EPS) Insulation	0.03 - 0.05	Lightweight, good insulation properties	
Extruded Polystyrene (XPS) Insulation	0.025 - 0.035	Excellent thermal resistance, commonly used in residential and commercial settings.	
Vacuum Insulation Panels (VIP)	As low as 0.005	Exceptional thermal performance, thin insulation with minimal thickness	

When selecting roof insulation materials, it is essential to consider factors such as local climate conditions, building regulations, and energy efficiency requirements. Consulting with building professionals and referring to relevant standards and certifications can help ensure the appropriate selection and installation of insulation materials to achieve optimal thermal performance and energy savings.

(f) Heating, Ventilation, and Air Conditioning (HVAC) Systems: HVAC systems are essential components of maintaining thermal comfort in buildings. Proper design, sizing, and maintenance of HVAC systems ensure efficient temperature and humidity control (Zhuang & Wang, 2020). Advanced systems may incorporate variable air volume (VAV), demand-controlled ventilation, and smart thermostats, providing precise control over indoor conditions and improving occupant comfort.

- **Area served by AC (Air Conditioning):** The parameter "Area served by AC" is an essential factor in evaluating the cooling coverage and energy efficiency of air conditioning systems in a building. It measures the percentage of the building's floor area served by air conditioning, providing insights into the extent of cooling coverage within the space.

  The measurement formula for the "Area served by AC" parameter is as follows:

  $$\text{Area AC\%} = \frac{\Sigma Area\ served\ by\ AC(m^2)}{\Sigma Gross\ floor\ area\ (m^2)} \qquad (2)$$

  The unit of measurement for this parameter is a percentage (%), representing the proportion of the total floor area serviced by air conditioning systems. To determine the "Area served by AC," both the building floor plan and physical measurements are utilised. The building floor plan provides information about the layout and designated areas for installing air conditioning systems. Physical measuring may also be necessary to obtain accurate measurements of the regions served by air conditioning in complex or non-standard spaces.

The "Area served by AC" parameter is valuable for evaluating the energy efficiency and effectiveness of the air conditioning systems. A higher percentage indicates that a more significant portion of the building's floor area is being cooled, suggesting a greater demand for cooling and potentially higher energy consumption. Moreover, this parameter is also relevant for assessing occupant comfort. Adequate cooling coverage ensures occupants have a comfortable indoor environment, particularly in regions with high temperatures or during hot seasons. The "Area served by AC" parameter helps determine the overall cooling capacity and distribution throughout the building.

(g) Occupant Density: The number of people occupying a space influences thermal comfort. High occupant density can increase heat generation, leading to a warmer environment (Ali et al., 2020). On the other hand, low occupancy may lead to more excellent conditions. Proper ventilation and air distribution systems should consider occupant density to maintain comfortable conditions for all occupants.

- Number of employees sharing the same workstation: The number of employees sharing the same workstation is a parameter that provides insights into the occupancy density and utilization of workspaces within a building (Mantesi et al., 2022). While it is not measured in a specific unit, it is an essential consideration in understanding the dynamics of the workplace environment.

  This parameter is typically sourced from employee support and engagement (ESE) data, which captures information about the organisation's seating arrangements and allocation of workstations. It helps to identify how many employees are assigned to a particular workstation or desk. The measurement of this parameter does not involve a specific quantitative measurement but instead relies on collecting and analyzing data on workstation assignments and employee seating arrangements. It can vary based on the organization's policies, spatial constraints, and workforce requirements.

  The number of employees sharing the same workstation has implications for occupant comfort, collaboration, and productivity. High workstation-sharing ratios can lead to overcrowding and reduced personal space, potentially impacting employee well-being and satisfaction. On the other hand, efficient space utilization and flexible seating arrangements can promote collaboration and interaction among employees.

(h) Thermal Zoning: Different areas within a building may have varying thermal comfort requirements. Thermal zoning allows for customized temperature control to meet the needs of other spaces and occupants (J. Kim et al., 2019). By dividing a building into zones, temperature adjustments can be made based on occupancy, activity levels, and individual preferences, optimising thermal comfort and energy efficiency.

(i) Building Orientation and Location :The orientation of a building and its geographic location can significantly influence thermal comfort. Proper orientation maximises or minimises solar heat gain, depending on the climate. By considering the position of windows, shading devices, and the layout of the building, architects and designers can optimise thermal comfort and reduce energy consumption (J. Zhao & Du, 2020).

Thermal comfort in buildings is influenced by a combination of factors, both personal and related to the building itself. Temperature, humidity, air velocity, radiant temperature, insulation, HVAC systems, solar heat gain, occupant density, thermal zoning, and building orientation and location all contribute to the overall comfort experienced by occupants. By considering these factors during the design, construction, and operation phases of a building, it is possible to create comfortable and energy-efficient indoor environments that promote well-being and productivity. Achieving optimal thermal comfort requires balancing the building's design, systems, and individual adjustments to meet the occupants' needs.

(a) Lighting Design: Effective lighting design plays a vital role in creating visual comfort. It involves strategically placing and arranging light sources to ensure adequate illumination throughout the space. A well-designed lighting system should provide uniform lighting levels without creating stark contrasts or dark spots, enhancing visual clarity and reducing eye strain.

(b) Daylighting: Daylighting refers to the controlled utilization of natural light to illuminate interior spaces. Incorporating ample natural light enhances visual comfort and offers employees numerous health and productivity benefits. Well-placed windows, skylights, and light wells can optimise daylight penetration, reducing reliance on artificial lighting and fostering a connection with the outdoors (Walker & Kimberly, 2020).

(c) Artificial Lighting: Artificial lighting becomes crucial in areas where natural light is limited or during nighttime. The selection of appropriate artificial lighting fixtures, such as LED or fluorescent lights, should consider factors like colour temperature, intensity, and distribution (Saraswati et al., 2021). Adequate and well-balanced artificial lighting helps maintain visual comfort and prevents visual fatigue.

(d) Glare Control: Glare occurs when there is excessive contrast between bright and dark areas, leading to discomfort and visual impairment. Proper glare control measures, such as window treatments, light diffusers, and anti-glare coatings on screens and surfaces, ensure visual comfort. Minimizing glare helps employees maintain focus, reduces eye strain, and improves overall visual well-being.

(e) Area Served by Lighting: The parameter "Area Served by Lighting" is crucial in assessing the lighting coverage and effectiveness within a building. It quantifies the percentage of the total floor area covered by lighting fixtures, providing insights into the extent of lighting provision throughout the space.

The measurement formula for the "Area served by Lighting" parameter is as follows:

$$\text{Area Lighting\%} = \frac{\Sigma Area\ served\ by\ Lighting(m^2)}{\Sigma Gross\ floor\ area\ (m^2)} \qquad (3)$$

The unit of measurement for this parameter is a percentage (%), representing the proportion of the total floor area served by lighting fixtures.

To determine the "Area served by Lighting," both the building floor plan and physical measuring are utilised. The building floor plan provides information about the layout and designated areas for installing lighting fixtures. Physical measuring may also be necessary to accurately measure the areas covered by lighting fixtures, especially in complex or non-standard spaces.

Efficient lighting design aims to provide sufficient illumination precisely where needed, avoiding over- or under-lighting certain areas. By analyzing the "Area served by Lighting" parameter, architects and lighting designers can identify areas where lighting improvements can be made, such as adjusting fixture placement, optimising light levels, or implementing lighting controls. Moreover, this parameter is also relevant for energy efficiency considerations. Lighting typically accounts for a significant portion of a building's energy consumption. By assessing the "Area served by Lighting" and optimising lighting design, it is possible to minimise energy waste, reduce electricity costs, and lower the environmental impact of lighting usage.

(a) Smart Control of Lighting: The parameter "Smart control of Lighting" plays a significant role in assessing the level of sophistication and energy efficiency of the lighting systems within a building. It provides information on whether smart control systems have been implemented to regulate and optimise lighting operations.

    Measuring the "Smart control of Lighting" parameter is straightforward, employing a simple binary classification. It is determined by a "Yes" or "No" response, indicating whether or not smart control systems are in place. The unit of measurement for this parameter is not applicable, as it represents a categorical measurement rather than a numerical value.

    Utilising occupancy sensors, smart control systems can detect the presence or absence of occupants within a space and automatically adjust lighting accordingly. This feature ensures that lighting is only activated when needed, reducing energy consumption during unoccupied periods (Chew et al., 2017). Daylight sensors, another component of smart lighting control, measure the amount of natural light available in a space and adjust artificial lighting levels accordingly. Using natural light can reduce energy usage while maintaining appropriate lighting levels.

    Time scheduling features enable lighting to be programmed to turn on and off at specific times, aligning with occupancy patterns and operational requirements. This prevents unnecessary energy consumption when the building is not in use. Dimming capabilities allow for flexible control of lighting levels, adjusting brightness based on specific needs or preferences. Dimming also contributes to energy savings by reducing lighting fixtures' energy consumption.

(b) Lux level: The parameter "Lux level" is a crucial factor in assessing the lighting quality within a building. It measures the level of illuminance, which refers to the amount of light falling on a surface and is expressed in Lux (lx). The Lux level provides valuable information about the intensity and adequacy of lighting in a particular area.

    The "Lux level" parameter is measured using a Lux meter, designed to measure illuminance. The Lux meter measures the amount of light reaching a surface in Lux units, considering natural and artificial lighting sources. The unit of measurement for this parameter is Lux (lx), which represents one lumen per square meter. Lux is a standard unit used in lighting design and is widely recognized for accurately quantifying the level of illuminance.

The Lux level measurement is essential for several reasons. Firstly, it ensures that lighting conditions meet recommended standards and provides adequate visibility for occupants 10. Different activities and tasks require varying levels of illuminance, and the Lux level helps determine if the lighting is appropriate for the intended use of the space. Furthermore, the Lux level measurement allows for the evaluation of lighting uniformity. Uniform lighting distribution is crucial to avoid excessive brightness or darkness, resulting in discomfort, eye strain and decreased productivity. By measuring the Lux level at various points, lighting professionals can assess the lighting uniformity across the space and make necessary adjustments.

The Lux level measurement is also relevant for compliance with lighting regulations and standards. Many building codes and guidelines specify minimum Lux-level requirements for different spaces, such as offices, classrooms, and industrial areas. Lighting design professionals, facility managers, and building owners should consult the relevant ASHRAE standards, such as ASHRAE/IESNA 90.1 (ASHRAE Standard 90.1, 2019) (Energy Standard for Buildings Except Low-Rise Residential Buildings), for detailed guidelines on lighting levels, energy efficiency, and other lighting-related considerations. By measuring and documenting the Lux levels, building owners and operators can ensure compliance with these standards and provide occupants with a safe and comfortable environment. ASHRAE does not explicitly define Lux levels for different spaces; it offers recommendations and best practices for achieving appropriate lighting levels in office and factory spaces. The recommended Lux levels may vary based on the specific tasks performed in each area. The general guidelines for office and factory spaces are summarized in Table 4.4.

Table 4.4: The general guidelines for comfortable lux levels in office and factory spaces

Space Type	**Recommended Lux Levels**	**Description**	**References**
		Office Spaces	
General Office Areas	300-500 Lux	Open-plan workstations, corridors, and circulation areas with non-demanding visual tasks.	(Huang et al., 2012; Mui & Wong, 2011; Chua et al., 2016; Yun et al., 2012; Dangol et al., 2013; Suk, 2019; Taleb & Antony, 2020)
Conference Rooms	500-750 Lux	Rooms for meetings, conferences, and collaborative activities.	
Task-Oriented Areas	500-1000 Lux	Workstations or areas requiring detailed tasks such as reading, writing, and computer work.	
Reception Areas	200-500 Lux	Entrance areas and lobbies where visitors wait ensure adequate lighting for orientation and visual comfort.	

Factory Spaces			
Assembly and Packaging Areas	300-500 Lux	Areas involving manual assembly, packaging, and inspection activities.	(Chowdhury & Alam, 2011; Wijewardane et al., 2018; Manju & Jacob P, 2022; Mui & Wong, 2011)
Machine Work and Precision Tasks	500-1000 Lux	Workstations require enhanced visibility and accuracy for intricate or detailed tasks like machining or electronics assembly.	
Storage and Warehousing	100-200 Lux	Areas dedicated to storage, inventory management, and warehousing with lower lighting requirements for material handling.	

It's important to note that these Lux-level recommendations are general guidelines and may vary based on specific industry standards, local regulations, and individual project requirements. Additionally, task-specific lighting considerations, such as glare control and uniformity, should be considered when designing lighting systems.

By adhering to recommended Lux levels, organizations can provide appropriate lighting conditions in their office and factory spaces, promoting visual comfort, safety, and productivity for occupants while considering energy efficiency and sustainability.

In general, Offices are required to have a Lux level of 500. Regarding screen-based devices such as Computers/laptops, Lux levels can vary from 300-500. But measuring the Lux level is a bit complicated, especially in the garment factories. Usually, the garment factories maintain specific Lux levels in the sewing machines according to their customer's requirements. The study has only considered the Lux level in the general workspace, not their particular working Lux level.

(c) Window Design: Windows play a significant role in visual comfort by providing views, natural light, and a connection to the external environment. The design of windows should consider factors like size, orientation, and placement to optimise daylight penetration while minimizing glare and thermal discomfort (Mirrahimi et al., 2016). Well-designed windows offer employees a pleasant visual experience and contribute to their well-being.

(d) WWR (Window-to-Wall Ratio): The Window-to-Wall Ratio (WWR) is a crucial parameter in assessing the design and performance of green buildings. It measures the glazing area (windows) in proportion to the total exterior wall area. The WWR is calculated by dividing the sum of the glazing area in square meters by the sum of the gross outer wall area in square meters and expressing it as a percentage.

The measurement formula for WWR is as follows:

$$\text{WWR\%} = \frac{\Sigma Glazing\ area\ (m^2)}{\Sigma Gross\ exterior\ wall\ area\ (m^2)} \tag{4}$$

The unit of measurement for WWR is the percentage (%), which represents the ratio of the glazing area to the total wall area.

The building floor plan is the primary source for obtaining the necessary measurements to calculate the WWR. Analyzing the floor plan, the glazing area and gross exterior wall area can be determined, enabling the calculation of the WWR.

The WWR parameter provides valuable insights into the fenestration (window) characteristics of a building. It indicates the extent of window coverage in the overall wall area. A higher WWR demonstrates greater glazing, allowing more natural light to enter the building and potentially reducing the need for artificial lighting during daylight hours (Albatayneh et al., 2021). However, if improperly designed and insulated, excessive glazing may also lead to increased solar heat gain and potential energy inefficiencies. A low WWR, on the other hand, suggests a building with a smaller window-to-wall ratio. This could be due to energy efficiency considerations, insulation requirements, or architectural design preferences. A lower WWR may result in reduced daylight penetration and reliance on artificial lighting systems.

Quantifying the WWR, researchers and architects can assess the balance between natural light, energy efficiency, and visual comfort within a building. It allows for comparisons between different buildings or areas within a single building to determine the impact of window coverage on energy consumption, thermal comfort, and occupant satisfaction.

(e) Total window area: The Total window area is a significant parameter in assessing the design and characteristics of windows in a building. It measures the combined area covered by all windows within the building. The total window area is calculated by summing up the products of the length and height of each window.

The measurement formula for the Total window area is as follows:

$$\text{Total window area} = \Sigma\, Length_{window}\ X\ Height_{window} \quad (5)$$

The unit of measurement for the Total window area is square meters (m^2), which represents the total surface area covered by all windows in the building. Information from the building floor plan and physical measurements are required to determine the total window area. The floor plan provides the layout and dimensions of each window, including the length and height. Physical measuring may also be necessary to obtain accurate measurements for irregularly shaped or non-standard windows.

A larger Total window area indicates more excellent window coverage, allowing more natural light to enter the space. The Total window area is a crucial parameter in green building design and evaluation. It affects various aspects of building performance, including daylighting, energy efficiency, and thermal comfort. A larger window area can enhance the availability of natural light, potentially reducing the need for artificial lighting during daylight hours and improving occupant well-being.

(f) Glazing U value: The Glazing U value is a crucial parameter used to assess the thermal performance and energy efficiency of glazing systems in buildings. It represents the overall heat transfer coefficient of the glazing system, which is influenced by factors such as the window material, thickness, and thermal properties.

The unit of measurement for the Glazing U value is watts per square meter per Kelvin (W/m^2K). It indicates the amount of heat energy that transfers through the glazing per unit area and the degree of temperature difference.

Building maintenance information and literature are commonly used as measurement sources to determine the Glazing U value. This information provides data on the properties of the glazing system, including the materials used, thickness, and thermal characteristics. The U value can be calculated using established formulas or reference tables by analysing these parameters.

The Glazing U value is a key factor in evaluating the thermal performance of windows. A lower U value indicates better insulation properties and reduced heat transfer through the glazing. This helps to minimise heat loss during colder months and heat gain during hotter months, contributing to energy efficiency and occupant comfort. The selection of glazing materials and technologies with lower U values can significantly impact the overall energy consumption of a building. By choosing windows with lower U values, architects and engineers can improve the building's energy efficiency, reduce reliance on heating and cooling systems, and create a more comfortable indoor environment for occupants.

It is important to note that the Glazing U value should be considered in conjunction with other factors, such as solar heat gain coefficient (SHGC) and visible transmittance (VT), to evaluate the overall performance of the glazing system (Berardi, 2019). These parameters provide insights into the window's ability to control solar heat gain, allow natural light to enter the building, and balance energy efficiency with daylighting requirements. The window material, thickness, and thermal properties determine it.

Some common types of window glazing and their corresponding U and SHGC values are described in Table 4.5.

Table 4.5: Common types of windows glazing types

Glazing Type	**U Value**	**SHGC**
Single Pane Glazing	High	High
Double Pane Glazing	Lower than single pane glazing	Variable (depending on glass type, coatings, and gas fillings)
Triple Pane Glazing	Lowest	Variable (depending on specific requirements)
Low-E Coated Glazing	Lower than untreated glazing (improves insulation)	Variable (can be designed for selective solar heat control)
Tinted Glazing	Moderate	Low (reflects a portion of solar radiation)

The summary of the popular window glazing system types of the world is summarized in Table 4.6.

Table 4.6: Common types of the window glazing system types (References: Gasparella et al., 2011; Pereira et al., 2022; Mahmoud, 2022; Hee et al., 2015; Saadatian et al., 2021; Es-Sakali et al., 2022)

Simple Glazing System Descriptor	Glazing System Type	Outer Glass Type	U- Value	SHGC
Single-glazed, clear	Single-glazing	clear	.67	0.57
Single-glazed, tint	Single-glazing	tint	.66	0.41
Single-glazed, high solar gain low E	Single-glazing	high solar gain, low E	.54	0.49
Single-glazed, low solar gain low E	Single-glazing	low solar gain, low E	.56	0.36
Double-glazed, clear/clear, air-fill	Double-glazing with air fill	clear	.48	0.51
Double-glazed, tint / clear, air-fill	Double-glazing with air fill	tint	.52	0.35
Double-glazed, high solar gain, low E / clear, air-fill	Double-glazing with air fill	high solar gain, low E	.43	0.47
Double-glazed, low solar gain, low E / clear, air-fill	Double-glazing with air fill	low solar gain, low E	.49	0.33
Double-glazed, clear/clear, argon fill	Double-glazing with argon fill	clear	.45	0.3
Double-glazed, tint / clear, argon fill	Double-glazing with argon fill	tint	.51	0.32
Double-glazed, high solar gain, low E / clear, argon fill	Double-glazing with argon fill	high solar gain, low E	.41	0.47

(g) Solar Heat Gain

Solar heat gain refers to the amount of sunlight entering a building. Depending on the climate and building orientation, excessive solar radiation can lead to overheating, causing discomfort and increasing the reliance on cooling systems (Kisilewicz, 2019). Proper shading techniques, such as overhangs, blinds, or tinted glazing, can help regulate solar heat gain, contributing to thermal comfort.

The SHGC is not measured in specific units but is expressed as a dimensionless value between 0 and 1. A higher SHGC indicates that more solar heat is transmitted through the glazing, while a lower SHGC implies that less solar heat can pass through.

To determine the SHGC, building maintenance information and literature serve as important measurement sources. This information provides data on the glazing material used in the building, including its specific properties related to solar heat gain. Any tinting or coating applied to the glazing can also influence the SHGC. By referencing building maintenance records and relevant literature, professionals can obtain the necessary information to assess the SHGC.

(h) Colour and Material Selection
The colours and materials used in interior spaces influence visual comfort. Light-coloured walls, ceilings, and furniture help reflect light, creating a brighter and more visually comfortable environment. A careful selection of materials minimising reflections and reducing contrast can contribute to a harmonious and visually pleasing workspace (Kingma, 2018).

(i) Spatial Layout
The spatial layout of a building affects visual comfort by considering factors like room dimensions, furniture arrangement, and workstation design. Adequate spacing between workstations, proper placement of furniture, and consideration of sightlines contribute to an open and visually comfortable environment. A well-designed layout ensures that employees have sufficient visual access, reducing feelings of confinement and enhancing their overall comfort.

(j) Gross floor area: Gross floor area is a fundamental parameter that measures the total area of a building, encompassing all floors and spaces within its boundaries. It is a key metric used in architectural and construction design, providing an overall assessment of the size and scale of the structure.

The measurement of gross floor area is obtained from the building floor plan, which outlines the layout and dimensions of each floor and space. The total gross floor area can be determined by summing up the individual sites of all floors, including corridors, rooms, common areas, and service areas. The unit of measurement for gross floor area is square meters (m^2). Gross floor area is critical for various purposes, including space planning, construction cost estimation, zoning regulations, and energy efficiency calculations. It serves as a baseline for determining building requirements, such as the number of occupants, ventilation needs, fire safety measures, and electrical load capacity.

In sustainable design practices, a common goal is to minimise the gross floor area while maximising the usable space. This approach promotes compact and efficient building designs, reducing material usage, construction costs, and environmental impact. It also encourages effective space planning and the integration of sustainable features such as daylighting, natural ventilation, and efficient use of resources.

(k) Area of the Green Roof: The parameter "Area of the Green Roof" is an essential factor in assessing the sustainability and environmental impact of a building. It measures the area covered by green roofs as a percentage of the total floor area. Green roofs refer to implementing vegetation and plant life on the roof surface, providing numerous benefits regarding energy efficiency, stormwater management, and biodiversity promotion.

Measuring the "Area of the Green Roof" parameter involves determining the total area covered by green roofs within the building. This can be obtained by analyzing the building's floor plan and physically measuring the green roof area. The unit of measurement for this parameter is square meters (m^2), which represents the area covered by the green roof.

This parameter's measurement source includes the building floor plan and physical measuring. The floor plan estimates the green roof area based on the designated space for vegetation on the roof. However, physical measuring is also crucial to ensure accurate calculations and account for any variations or modifications made during the construction or implementation of the green roof.

The "Area of the Green Roof" parameter holds significant importance due to the numerous benefits associated with green roofs. Green roofs help mitigate the urban heat island effect by reducing surface temperatures and minimizing heat absorption (Bevilacqua et al., 2017). The vegetation acts as a natural insulator, reducing the energy demand for cooling and improving the thermal performance of the building. Moreover, green roofs contribute to stormwater management by absorbing and retaining rainwater, reducing the burden on drainage systems and mitigating the risk of flooding. The plants and soil on the green roof act as a natural filtration system, removing pollutants and improving runoff water quality. Additionally, green roofs have aesthetic value, improving the visual appeal of the building and creating a greener and more pleasant environment for occupants. They can also provide recreational spaces for employees or residents, offering opportunities for relaxation, socialization, and outdoor activities.

(l) Indoor air quality (IAQ)

The IAQ is critical in creating a healthy and comfortable work environment for employees. The structural design of a building plays a significant role in determining IAQ. This article analyses the building factors influencing indoor air quality, including ventilation systems, building materials, pollutant sources, air filtration, moisture control, and maintenance practices. Understanding these factors is crucial for creating a conducive and healthy indoor environment for employees.

(m) PM 2.5 Concentration: PM 2.5, which stands for Particulate Matter 2.5, is a crucial parameter in assessing air quality and understanding the concentration of fine particles in the atmosphere. It refers to airborne particles with a diameter of 2.5 micrometres or smaller, which are small enough to be easily inhaled into the respiratory system.

The measurement source for PM 2.5 is physical measuring using an air quality meter. This device employs advanced sensing technology to detect and quantify the concentration of fine particles present in the air. Analyzing the particulate matter levels provides valuable information about the air quality and potential health risks associated with exposure to high concentrations of PM 2.5. PM 2.5 particles are of significant concern due to their small size and ability to penetrate deep into the respiratory system. These particles can originate from various sources, including combustion processes, vehicle emissions, industrial activities, and natural sources such as dust and pollen. Long-term exposure to high levels of PM 2.5 has been associated with adverse health effects, including respiratory problems, cardiovascular diseases, and even premature death. The unit of measurement for PM 2.5 is micrograms per cubic meter (µg/m³), which represents the mass of particles per unit volume of air.

(n) PM10: also known as Particulate Matter 10, is a parameter used to measure the concentration of airborne particles with a diameter of 10 micrometres or smaller. These particles can include dust, pollen, mould spores, and other solid or liquid particles suspended in the air.

PM_{10} is typically measured using an air quality meter, which provides a meter reading indicating the concentration of these particles in the surrounding environment. The unit of measurement for PM_{10} is micrograms per cubic meter (µg/m³), which represents the mass of particles per unit volume of air. The measurement source for PM_{10} is physical measuring by an air quality meter. This specialized device utilises sensors and detectors to collect air samples and analyze the concentration of particulate matter present. Measuring the levels of PM_{10} provides valuable information about the air quality and potential health risks associated with exposure to these particles.

PM_{10} particles are larger than $PM_{2.5}$ but small enough to inhale into the respiratory system. They can originate from various sources, including natural processes like wind-blown dust, industrial activities, construction sites, and vehicle emissions. Prolonged exposure to high concentrations of PM_{10} can harm human health, particularly affecting the respiratory system and potentially exacerbating existing respiratory conditions.PM_{10} is a parameter used to measure the concentration of particulate matter with a diameter of 10 micrometres or smaller in the air. It is calculated using an air quality meter and expressed in micrograms per cubic meter (µg/m³).

ASHRAE recommends monitoring and controlling the levels of PM, specifically PM2.5 and PM10, which are fine particles suspended in the air. High concentrations of PM can lead to respiratory problems and other health issues.

(o) CO_2 Concentration: CO_2, or carbon dioxide, is a parameter used to measure the concentration of this gas in the indoor environment. It is an essential parameter to consider when assessing occupant comfort, indoor air quality, and the effectiveness of ventilation systems.

The measurement of CO_2 is typically conducted using an air quality meter, which provides a meter reading indicating the concentration of carbon dioxide in parts per million (ppm). This measurement source involves physical measuring, where air samples are collected and analyzed to determine the CO_2 levels present. Carbon dioxide is a naturally occurring gas released through various human activities, such as breathing, combustion processes, and the use of fossil fuels. In indoor environments, the concentration of CO_2 can increase due to inadequate ventilation, high occupancy levels, and poor air circulation. Elevated CO_2 levels can lead to feelings of stuffiness, decreased concentration, and potential health effects, particularly in enclosed spaces.

In green building studies, monitoring CO_2 levels is especially important as it provides valuable data on the performance of energy-efficient ventilation systems and sustainable building practices. By correlating CO_2 measurements with other parameters such as occupancy patterns, outdoor air quality, and thermal comfort, researchers can evaluate the impact of green building features on maintaining healthy indoor environments and reducing the carbon footprint associated with ventilation. ASHRAE (ASHRAE Standard 90.1, 2019) recommends maintaining appropriate levels of CO_2 in indoor spaces. Elevated CO_2 levels can indicate poor ventilation and may lead to drowsiness, poor concentration, and discomfort.

Following ASHRAE guidelines and standards, building professionals can ensure that indoor air quality parameters are monitored, controlled, and maintained within acceptable limits. This promotes a healthier and more comfortable indoor environment for occupants, enhancing their well-being and productivity. According to ASHRAE guidelines, the accepted levels are summarized in Table 4.7.

Table 4.7: Acceptable Air Quality parameters according to ASHRAE 90.1 (ASHRAE Standard 90.1, 2019)

PM 2.5 level ($\mu g/m^3$)	**PM 10 level ($\mu g/m^3$)**	**CO_2 (ppm)**	**Level of health concern**
0.0-12.0	0-54	0-700	Good
12.1-35.4	55-154	701-1000	Moderate
35.5-55.4	155-254	1001-1500	Unhealthy for sensitive groups
55.5-150.4	255-354	1501-2500	Unhealthy
150.5-250.4	355-424	2501-5000	Very unhealthy
≥250.4	≥425	≥5000	Hazardous

After analysing the factors from the literature review, they were categorized to make it easy to study. The factors are mainly divided into External factors and Internal factors. External factors again were subdivided into Socio-economic and legal-related characteristics and Climate-related characteristics. And internal factors are subdivided into, building services system-related characteristics, Building related characteristics and occupant-related characteristics. The factors are shown in Figure 4.2

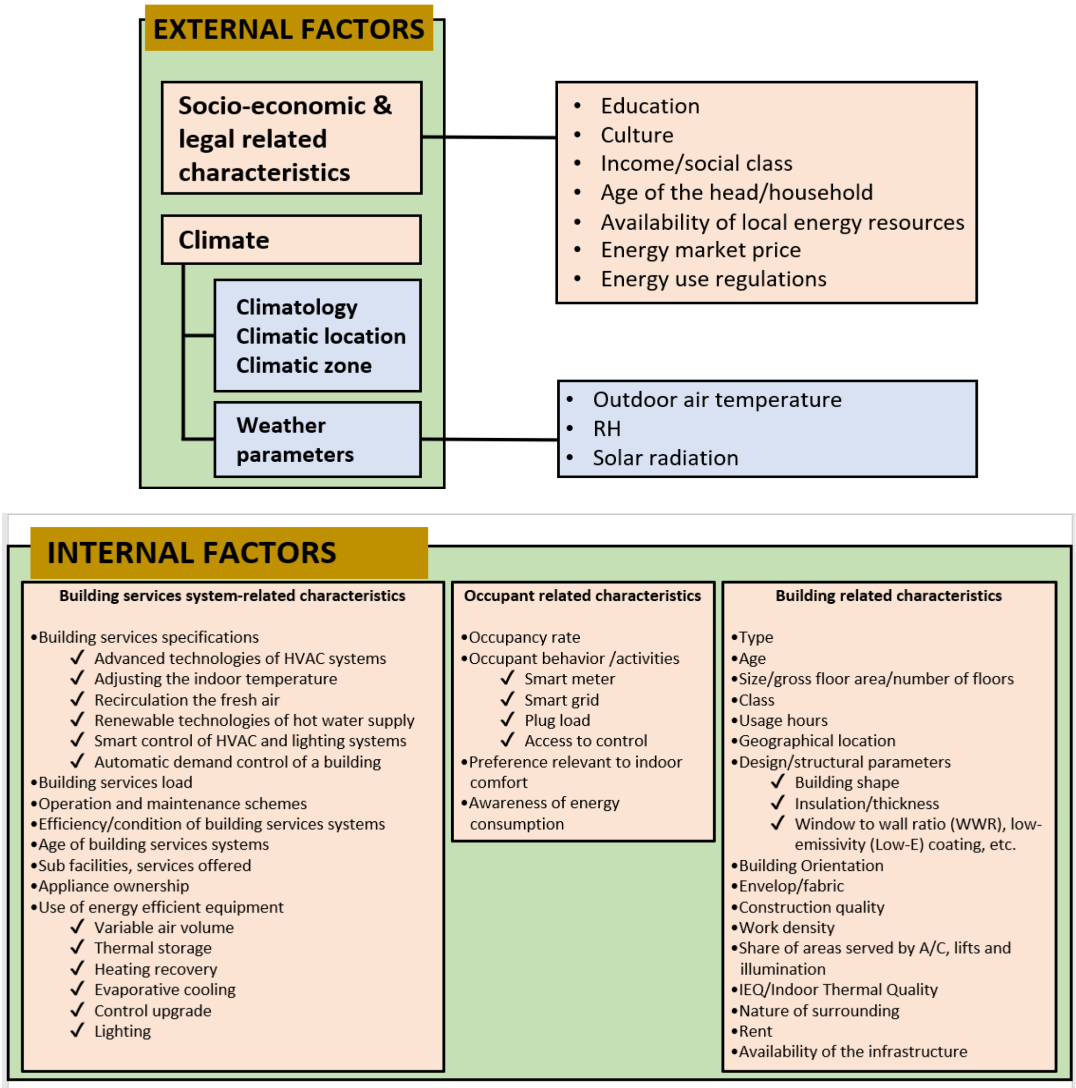


Figure 4.2: The building-related factors influencing IEQ

4.1.2 Factors affecting employee satisfaction in an office building

According to the systematic literature review, the keyword search criteria were as follows,

1. "thermal comfort", "Occupant satisfaction"
2. "visual comfort", " Occupant satisfaction "
3. "air quality", " Occupant satisfaction"
4. “IEQ/Indoor Environmental Quality”, “Occupant satisfaction”

The search results were narrowed down using the following procedure; Initial Search Keywords are “Satisfaction”, “IEQ”, and “Occupants” Search Years are 2011-2021, and the search sources are Google Scholar, Web of Science, Scopus and ScienceDirect. The systematic literature review process is summarised in Figure 4.3

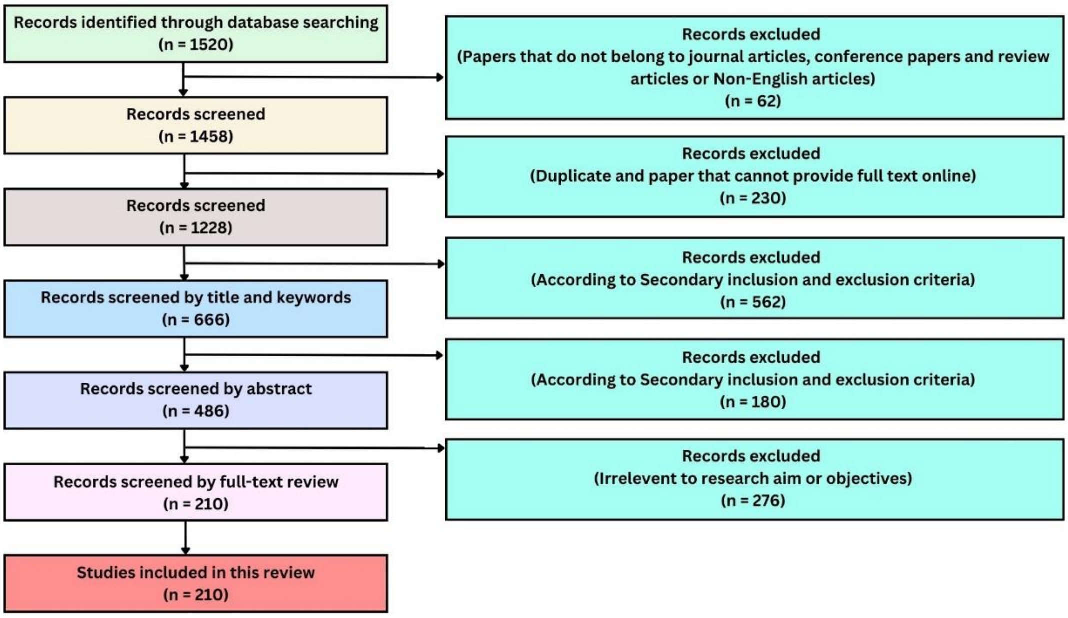


Figure 4.3: The steps conducted for bibliographic analysis to identify occupant factors

Physical comfort plays a crucial role in employee satisfaction. Comfortable furniture, ergonomic workstations, appropriate lighting, proper temperature and ventilation control, and noise reduction measures contribute to a comfortable work environment. Employees feeling physically comfortable in their workspace enhances their overall satisfaction and productivity. Physical comfort in an office building can directly impact employee satisfaction.

(a) Productivity: Employees who are physically comfortable in their workspace can better focus on their tasks and perform their work efficiently (Al-Omari & Okasheh, 2017). Comfortable furniture, ergonomic workstations, and appropriate lighting reduce physical strain and fatigue, allowing employees to work comfortably for extended periods without discomfort. As a result, employees can be more productive, leading to higher job satisfaction.

(b) Health and Well-being: Physical comfort is vital in maintaining employee health and well-being. Uncomfortable work environments, such as poorly designed chairs or workstations, can lead to musculoskeletal issues like back pain, neck strain, and repetitive injuries. On the other hand, ergonomic furniture and proper workstation setup can help alleviate these issues and promote better posture (Woo et al., 2015), reducing the risk of discomfort and long-term health problems. When employees feel physically well and comfortable, their overall satisfaction with their work environment improves.

(c) Focus and Concentration: An office building that provides proper temperature control, adequate ventilation, and noise reduction measures can help create an environment conducive to focus and concentration. Excessive heat or cold, poor air quality, and high noise levels can distract and disrupt employees' ability to concentrate (Rahman et al., 2022). Employees can focus better by ensuring physical comfort through optimal temperature and air quality regulation and minimizing noise disturbances, increasing job satisfaction.

(d) Stress Reduction: Physical discomfort can increase employee stress levels. Being in an environment with uncomfortable temperatures, poor ventilation, or excessive noise can create a sense of unease and frustration (Sutherland & Cooper, 2022). On the contrary, a physically comfortable office space can help alleviate stress and promote a sense of calm and well-being. When employees feel more at ease physically, their overall satisfaction with their work environment improves, and they can better manage workplace stress.

(e) Employee Retention and Engagement: A physically comfortable office environment demonstrates an organization's commitment to employee well-being. Employees who feel valued and cared for are more likely to be satisfied with their work environment and more inclined to stay with the company. Physical comfort, therefore, plays a role in employee retention. Moreover, physically comfortable employees are more likely to be engaged in their work, as they can focus on their tasks without distractions or discomfort.

(f) Perception of Organizational Support: Providing physical comfort in the office building contributes to the perception of organizational support. When employees have access to comfortable workstations, appropriate lighting, and other physical amenities that enhance their comfort, they perceive that their employer values their well-being. This perception of support positively influences employee satisfaction and fosters a positive employer-employee relationship.

Workspace design and layout significantly impact employee satisfaction in an office building. Here are some ways in which workspace design and layout can influence employee satisfaction:

(a) Collaboration and Communication: The design and layout of the workspace can either facilitate or hinder cooperation and communication among employees. Open, flexible arrangements with shared spaces and collaborative areas encourage interaction, teamwork, and idea exchange.
(b) Privacy and Focus: While collaboration is essential, employees also require privacy and focused work time. Well-designed workspace balances open collaborative spaces and private areas for individual work or confidential discussions. Providing designated quiet zones, private meeting rooms, and personal workstations helps employees concentrate on tasks that require focus and minimises distractions, leading to increased satisfaction with the workspace.
(c) Comfort and Ergonomics: A well-designed workspace prioritizes employee comfort and ergonomic principles (Voordt & Jensen, 2023). Ergonomic furniture, adjustable workstations, and proper lighting create a comfortable work environment. When employees can adjust their chairs and desks to suit their needs, it reduces the risk of discomfort and musculoskeletal issues, enhancing their overall satisfaction and well-being.
(d) Natural Light and Views: Access to natural light and views can positively impact employee satisfaction (Mourato et al., 2020). Ample windows, skylights, and open spaces that allow

natural light to reach work areas create a more pleasant and stimulating environment. Natural light has been linked to improved mood, productivity, and overall well-being. Incorporating views of nature or outdoor areas can also contribute to a sense of connection and provide mental respite for employees.

(e) Spatial Layout and Traffic Flow: A well-thought-out spatial layout ensures efficient traffic flow within the workspace. Properly organized workstations, clear pathways, and logical placement of common areas (e.g., meeting rooms, break areas, restrooms) contribute to a smooth workflow and minimise disruptions (Lin, 2016). When employees can navigate the workspace easily and move between zones without obstacles, it enhances efficiency and satisfaction.

(f) Flexibility and Adaptability: A workspace that allows flexibility and adaptability can positively impact employee satisfaction. Design features such as movable furniture, modular workstations, and reconfigurable spaces enable employees to customize their work environment based on their preferences and the nature of their tasks. Flexibility in workspace design empowers employees, promotes autonomy, and accommodates changing work needs, leading to higher job satisfaction.

(g) Aesthetics and Visual Appeal: The aesthetic quality of the workspace can influence employee satisfaction. A visually appealing environment, with attention to colour schemes, artwork, and overall aesthetics, can create a positive and enjoyable atmosphere (Lin, 2016). A well-designed and visually pleasing workspace can contribute to employee pride in their workplace and positively impact their satisfaction and engagement.

(h) Well-being and Amenities: Incorporating elements that support employee well-being, such as breakout areas, relaxation spaces, or wellness rooms, can enhance employee satisfaction. Providing amenities like comfortable seating, plants, natural elements, and refreshment access promotes employee well-being and provides opportunities for relaxation and recharge during the workday (Andrews, 2022).

Considering and implementing thoughtful workspace design and layout principles, employers can create an environment that promotes employee satisfaction, collaboration, productivity, and well-being. A well-designed workspace enhances the employee experience and contributes to a positive work culture within the office building. The summarized influencing factors for employees' physical satisfaction in an office building are in Figure 4.4. The factors can be divided into two main categories: IEQ-related and non-IEQ factors.

Both IEQ-related and non-IEQ-related factors contribute to employee comfort in an office building. Ensuring a high standard of Indoor Environmental Quality, including proper air quality, lighting, and acoustics, is crucial for creating a comfortable and productive workspace. Additionally, factors like office layout, amenities, and company culture also play essential roles in promoting employee comfort and well-being. A holistic approach that considers both the physical environment and other aspects of the workplace is essential for optimizing employee comfort in an office setting.

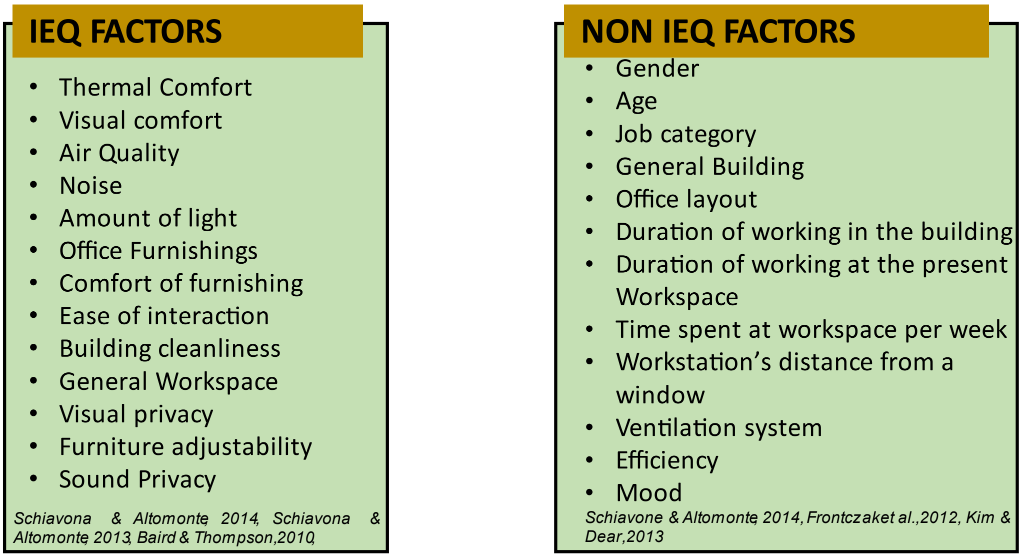


Figure 4.4: The occupant-related factors for IEQ comfort

4.2 Conceptual framework

The conceptual framework was designed to categorize the dependent and independent variables to understand the relationship between different factors.

1. Dependent variable: The dependent variable is the outcome or the response variable that researchers are trying to measure or predict. In this study, the dependent variable might be related to the performance, effectiveness, or any other measurable metric of the building structure or system under investigation. For example, it could be the energy efficiency, structural stability, maintenance cost, or any other relevant parameter related to the building's structural characteristics.

2. Independent variable: The independent variable, also known as the predictor variable, is what researchers manipulate or control in the study. It is expected to have an effect on the dependent variable. In this case, the focus is on structural parameters, which means variables related to the physical attributes and design of the building. These could include things like the building's materials, layout, architectural features, and other design elements that are directly related to the structural performance.
3. Narrowing the scope: External factors, such as location, climate, and surrounding environment, are omitted from the independent variable. The reason is that these factors may vary from one building to another and are not directly related to the building's structural parameters. Similarly, building system-related characteristics, like HVAC (Heating, Ventilation, and Air Conditioning) systems or lighting systems, might depend on the specific applications or equipment installed and are therefore left out from the study.
4. Excluding psychological factors: Psychological factors, like employee efficiency and mood, are highly complex and subjective aspects. It can be challenging to quantify and objectively measure emotions and human behavior solely through a questionnaire survey. As a result,

these factors are not considered part of the independent variable in this study. Instead, the research is focused on objective and measurable structural parameters to maintain a clear and more straightforward analysis.

The study is honing in on the impact of specific structural parameters on the dependent variable (e.g., building performance metrics) by excluding external factors, building system-related characteristics, and psychological factors. This narrowing of the scope allows the study to isolate and understand the direct influence of the building's physical attributes on the measured outcomes without the complexities introduced by the omitted variables. The conceptual framework is implied in Figure 4.5.

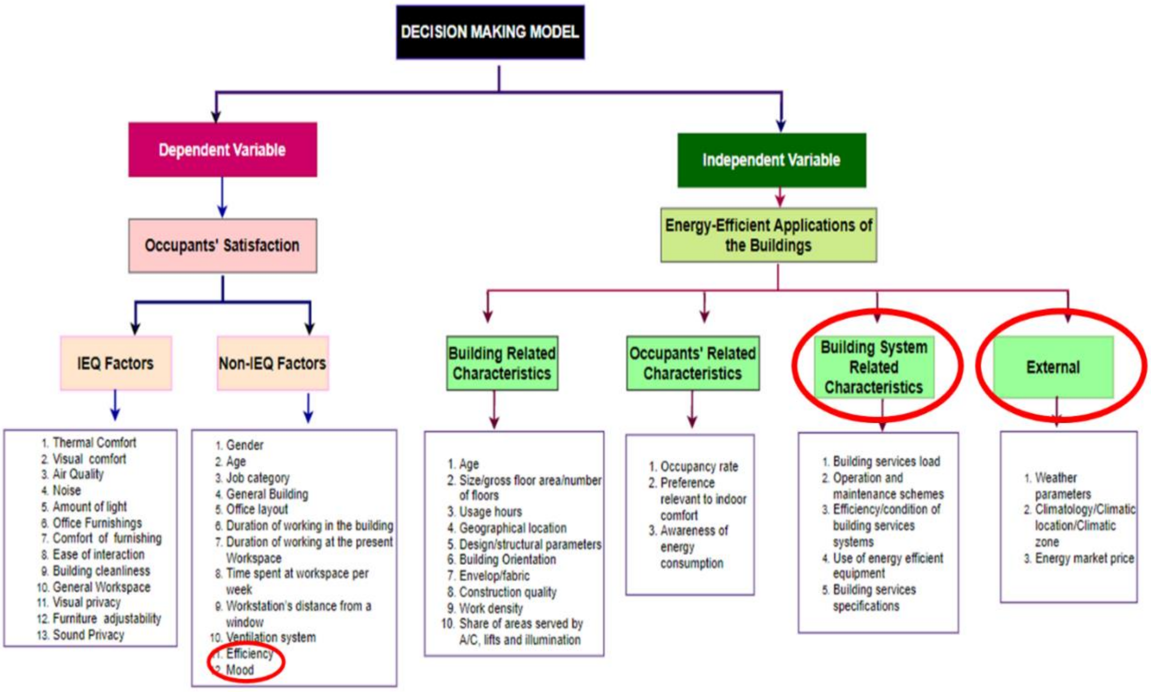


Figure 4.5: Conceptual Framework of the Study

4.3 Parameter Measurements Criteria

After identifying the relevant parameters for the study, the study proceeded to establish a clear and consistent measurement criterion to quantify these parameters effectively. This was achieved through the creation of Table 4.8, which likely outlines the specific details of how each parameter will be measured and where the data will be obtained. The measurement criteria provide a standardized approach to collect and analyze data, ensuring the research is conducted in a systematic and reliable manner.

The different methods used for parameter measurement can be categorized into four main approaches:

1. Physical measurements: Some parameters may require direct physical measurements of certain building characteristics. For example, structural parameters such as the dimensions

of the building, wall thickness, or floor area may be measured using tools like tape measures, laser distance meters, or other specialized equipment.

2. Specific parameter meters: In some cases, specialized meters or instruments are used to measure certain parameters accurately. For instance, devices like air quality monitors can measure indoor air quality parameters like particulate matter, carbon dioxide levels, and humidity. Similarly, thermal comfort parameters may be measured using temperature and humidity sensors.

3. Building plan analysis: Building plans and architectural drawings are valuable resources for obtaining theoretical values for certain parameters. These plans may contain information on building dimensions, layout, and design elements that contribute to the study. These theoretical values can serve as a baseline or reference for the actual measurements.

4. Employee survey: To gather data on certain subjective parameters, such as employee comfort or satisfaction, a survey might be conducted. The survey could include questions related to perceived indoor environmental quality, thermal comfort, lighting satisfaction, and other relevant aspects. The data from the employee survey can provide insights into how the building environment affects the occupants and can complement the objective measurements.

To ensure the accuracy and reliability of the data, the obtained measurements with theoretical values were cross-referenced from the literature and relevant departments involved in the building's design and maintenance. This helps validate the measurements and provides a broader perspective on the parameter values.

By combining physical measurements, specific parameter meters, building plan analysis, and employee surveys, the researchers can obtain a comprehensive dataset to assess the building's performance in terms of the identified parameters. This data-driven approach allows for a more robust analysis, leading to meaningful conclusions and recommendations for improving indoor environmental quality and employee comfort in the office building.

Table 4.8: Parameter measuring criteria

Parameter	**Parameter Measurement**	**Unit**	**Measurement source**
WWR	WWR% = $\frac{\Sigma Glazing\ area\ (m^2)}{\Sigma Gross\ exterior\ wall\ area\ (m^2)}$	%	Building floor plan
Total window area	Total window area= $\Sigma\ Length_{window} * Heightwindo_{w}$	m^2	Building floor plan and physical measuring
Glazing U value	U = $\frac{1}{R_T} = \frac{W}{M^2 * K}$	W/m^2K	Building maintenance

			information, Literature
SHGC	Glazing material, Glazing tinted material	-	Building maintenance information, Literature
Area served by AC (%)	Area AC% = $\frac{\Sigma Area\ served\ by\ AC(m^2)}{\Sigma Gross\ floor\ area\ (m^2)}$	%	Building floor plan and physical measuring
Smart control of HVAC	Yes/ No	-	ESE
Wall insulation U value	U = $\frac{1}{R_T} = \frac{W}{M^2 * K}$	W/m^2K	Building maintenance information, Literature
The thickness of wall insulation	Insulation thickness = $Average\ \Sigma\ Width\ insulation\ material$	mm	Physical measuring
Area served by Lighting (%)	Area Lighting% = $\frac{\Sigma Area\ served\ by\ lights(m^2)}{\Sigma Gross\ floor\ area\ (m^2)}$	%	Building floor plan and physical measuring
Roof insulation U value	U = $1R_T$=WM^2 *K	W/m^2K	Building maintenance information, Literature
Smart Control of Lighting	Yes/ No	-	ESE
Lux level	Meter reading	lx	Physical measuring by a Lux meter
Area of the Green roof	Green roof Area = $\frac{\Sigma Green\ roof\ area(m^2)}{\Sigma Gross\ floor\ area\ (m^2)}$	m^2	Building floor plan and

			physical measuring
$PM_{2.5}$	Meter reading	µg/m³	Physical measuring by an air quality meter
PM_{10}	Meter reading	µg/m³	Physical measuring by an air quality meter
CO_2	Meter reading	ppm	Physical measuring by an air quality meter
Number of employees sharing the same workstation	-	m	ESE
Gross floor area	$\Sigma Gross\ floor\ area\ (m^2)$	m^2	Building floor plan

4.4 Main features of selected office green buildings

Fourteen buildings were selected for the study island-wide. Most of the office green buildings are situated in the Colombo area. All the buildings are Gold or Platinum rated buildings. Eight general office buildings and six factory spaces were studied. The non-disclosure agreement (NDA) was signed between each selected building and the research team. According to the NDA, the names of the office spaces cannot be disclosed for any reason, and the data can be published after getting approvals from them and without using their brand name. The factories are notated by F, and O notates the offices for identification. The building features are listed in Table 4.9 and Table 4.10

Factories

F1: SL's first LEAD gold-certified factory was a leading apparel manufacturing factory.

Figure 4.6: Factory 1

F2: SL's first GBCSL Platinum-certified factory was a toothbrush and razor manufacturing factory.

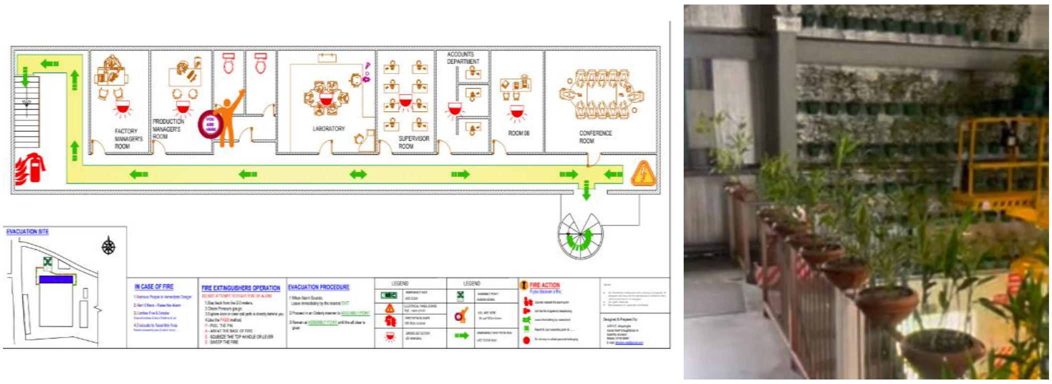


Figure 4.7 : Factory 2

F3: The first LEED platinum-awarded green building in SL. This building has the best green features, including argon-filled double-layer glazing and imported high-grade insulation materials. It was a leading apparel manufacturing factory.

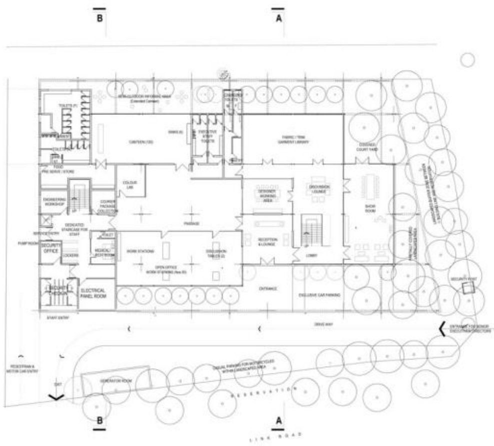


Figure 4.8: Factory 3

F4: LEED Gold certified apparel manufacturing factory

Figure 4.9: Factory 4

F5: GBCSL gold-certified cement factory

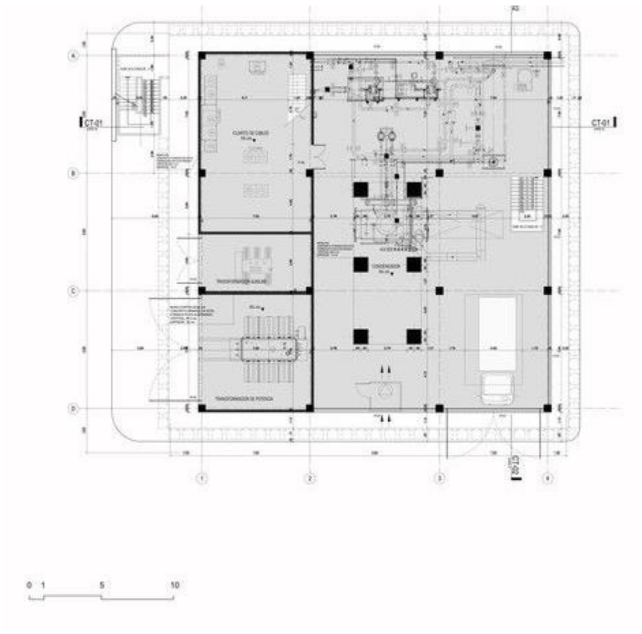


Figure 4.10: Factory 5

F6: LEED Gold - certified apparel manufacturing factory

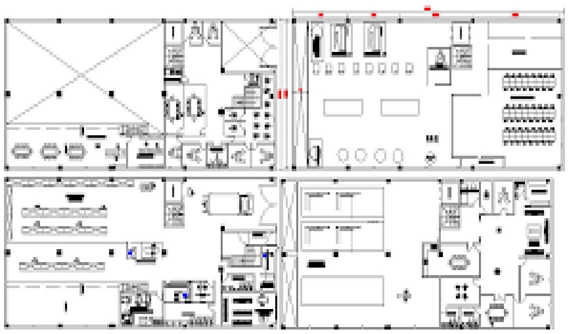


Figure 4.11: Factory 6

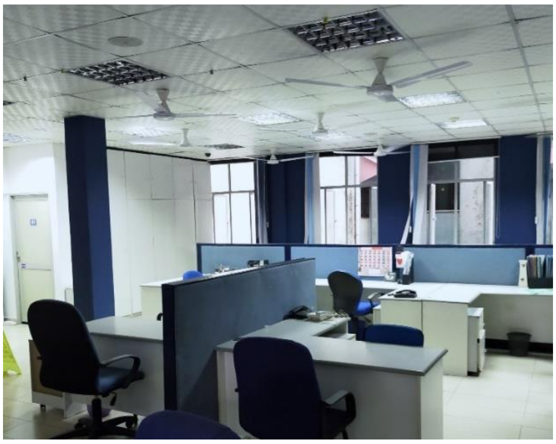
Table 4.9: Building features summary - Factories

Features	F1	F2	F3	F4	F5	F6
WWR %	33	16	35	16	25	33
Total window area (m^2)	31.5	64.5	38	64.5	33	32
Glazing U value (W/m^2K)	1	0.48	0.2	0.48	0.48	1
SHGC	0.79	0.76	0.76	0.7	0.7	0.76
Area served by AC (%)	90	12	34	50	82	92
Smart control of HVAC	Yes	No	Yes	No	Yes	No
Wall insulation U value (W/m^2K)	0.17	0.35	0.12	0.35	0.32	0.3
The thickness of wall insulation (mm)	3	24	40	48	32	30
Area served by Lighting (%)	80	50	80	85	30	78
Smart Control of Lighting	No	No	Yes	No	Yes	No
Lux level (lx)	800	900	700	750	200	750
Area of the Green roof (m^2)	1500	1897	0	1897	0	0
$PM_{2.5}$ ($\mu g/m^3$)	40	38	30	35	52	37
PM_{10} ($\mu g/m^3$)	158	150	120	130	170	147

CO_2 ppm	6000	5000	4000	4500	5200	4000
Number of employees sharing the same workstation	18	05	8	04	04	8
Gross floor area (m^2)	9600	8256	3675	8630	15000	12000

General office buildings

O1: This is an office situated in Colombo, which is GBCSL gold-certified space



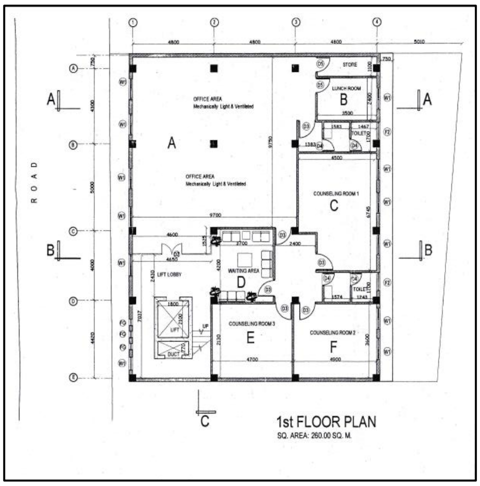


Figure 4.12: Office 1

O2: This is an office situated in Colombo, which is GBCSL gold-certified space

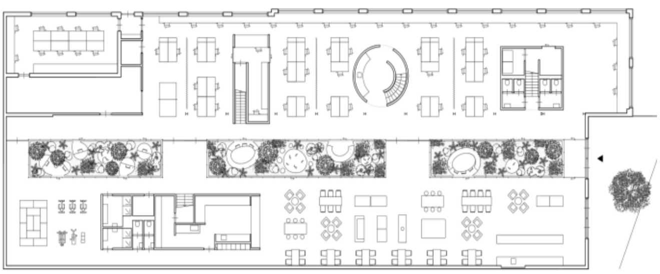


Figure 4.13: Office 2

O3: This is a GBCSL gold-certified building. This is a historical building which is later converted into an office space.

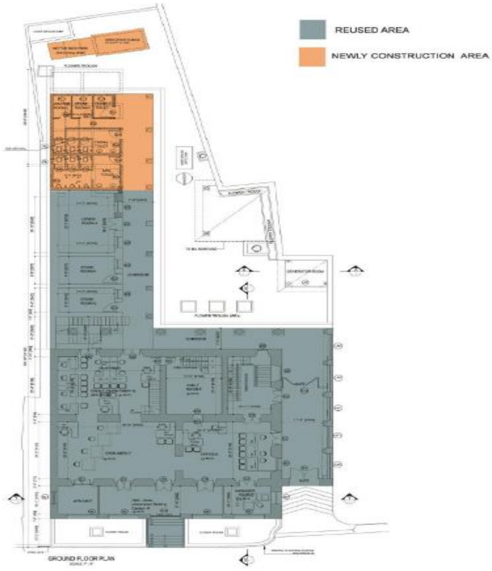


Figure 4.14: Office 3

O4: This is a GBCSL gold-certified building. This is a vehicle showroom office space.

Figure 4.15: Office 4

O5: This is a GBCSL gold-certified building. This is a bank office space.

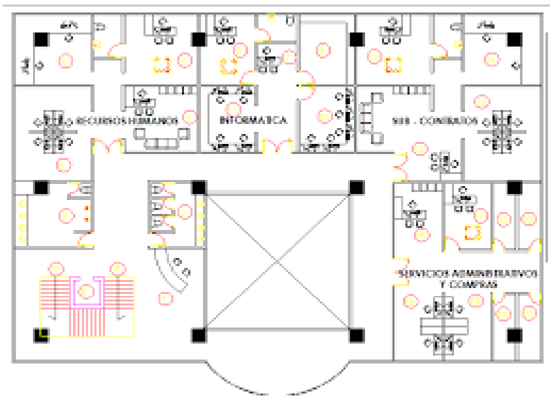


Figure 4.16: Office 5

O6: This is a GBCSL gold-certified bank.

Figure 4.17: Office 6

O7: This is a GBCSL gold-certified building.

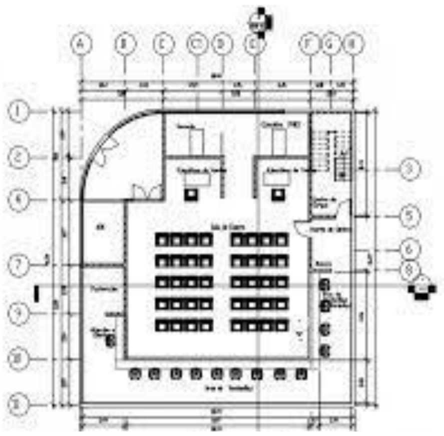


Figure 4.18: Office 7

O8: This is a GBCSL gold-certified low-rise office building.

Figure 4.19: Office 8

Table 4.10: Building features summary - General office spaces

Office	O1	O2	O3	O4	O5	O6	O7	O8
WWR %	25	55	28	60	60	45	50	28
Total window area (m^2)	56.34	42.3	45.68	57.5	50	35	45	58
Glazing U value (W/m^2K)	0.48	0.48	1	0.48	0.48	0.5	0.48	1.5
SHGC	0.76	0.76	0.79	0.76	0.76	0.7	0.79	0.7
Area served by AC (%)	10	79.6	29	70	70	90	78	85
Smart control of HVAC	Yes	Yes	No	No	No	no	no	yes
Wall insulation U value (W/m^2K)	0.17	0.12	0.4	0.12	0.4	0.26	0.26	0.2
The thickness of wall	32	40	33	22	20	28	32	28

insulation (mm)								
Area served by Lighting (%)	50	60	40	55	45	40	40	62
Smart Control of Lighting	Yes	Yes	No	No	No	no	no	yes
Lux level (lx)	300	500	200	1000	250	300	450	400
Area of the Green roof (m^2)	1423	1200	0	2500	0	0	0	0
$PM_{2.5}$ (µg/m³)	15	12	10	16	12	20	18	12
PM_{10} (µg/m³)	55	35	25	20	35	40	37	18
CO_2 ppm	1500	1200	1500	1200	1200	2000	2500	1200
Number of employees sharing the same workstation	02	03	1	1	2	1	3	2
Gross floor area (m^2)	5689	8600	682	5263	7632	3062	7632	3200

4.5 Identified building parameter ranges

The observed parameter ranges are summarized in Table 4.11 to Table 4.16. These observed variable ranges were then used as input data for developing a predictive model, enabling the study to gain valuable insights into building performance and occupant comfort. The analysis and conclusions drawn from this model can inform strategies to enhance the indoor environment and ultimately improve the well-being and productivity of employees in the office building.

Table 4.11: The observed window glazing types of the selected green buildings

Simple Glazing System Descriptor	**Glazing System Type**	**Outer Glass Type**	**U- Value (W/m^2K)**	**SHGC**
Single-glazed, clear	Single-glazing	clear	.67	0.57
Single-glazed, tint	Single-glazing	tint	.66	0.41
Single-glazed, high solar gain low E	Single-glazing	high solar gain, low E	.54	0.49
Double-glazed, clear/clear, air-fill	Double-glazing with air fill	clear	.48	0.51
Double-glazed, clear/clear, argon fill	Double-glazing with argon fill	clear	.45	0.3

Table 4.12: The observed roof insulation material types of the selected green buildings

Roof insulation materials	**U- Value (W/m^2K)**
Al foil	0.3
Rigid insulation board	0.3
Structural insulated panel	0.2-0.5

Table 4.13: The observed wall insulation material types of the selected green buildings

Wall insulation materials	Thickness (mm)	U- Value (W/m^2K)
Stud wall	80	0.48
Al foil	24	0.3
Rigid insulation board	24	0.3
Cavity wall	40	0.5

Table 4.14: The observed IAQ parameters of the selected green buildings

PM 2.5 level ($\mu g/m^3$)	PM 10 level ($\mu g/m^3$)	CO_2 (ppm)	Level of health concern
0.0-12.0	0-54	0-700	Good
12.1-35.4	55-154	701-1000	Moderate
35.5-55.4	155-254	1001-1500	Unhealthy for sensitive groups
		1501-2500	Unhealthy
		2501-5000	Very Unhealthy
		≥5000	Hazardous

Table 4.15: The observed Lux levels of the selected green buildings

Working Space Type	Standard LUX Level (lx)	Identified Range (lx)
Office Space	300-500	300-1000
Factory Space	300-700	300-1000

Table 4.16: The other observed parameters and their ranges

Parameter	Range
Gross Floor Area (m^2)	682-78000
WWR (%)	16-60
Total Window Area (m^2)	31-575
Share of the area served by AC(%)	10-90
Area of the green roof (m^2)	0-2500
Smart control of HVAC	Yes/No
Smart control of the lighting system	Yes/No
Area Served by Lighting (m^2)	30-80
Distance between the seat and window (m)	1-6

4.6 Pilot survey

The pilot survey is a critical phase in the research process, allowing researchers to gather valuable data and insights before conducting the substantive study on employee satisfaction with green buildings. This article provides a comprehensive overview of the objectives, methodology, conclusions, and adjustments based on the pilot survey results.

4.6.1 Objectives of the Pilot Survey:

(a) To gather data to guide a substantive study adapted to employee satisfaction with green buildings:

(b) One of the primary objectives of the pilot survey is to collect preliminary data that can guide the design and implementation of the main study. This involves determining the appropriate research methods, identifying potential challenges, and refining the research objectives based on initial findings.

(c) To test the response feasibility of the given questionnaire:

(d) Another crucial objective is to assess the feasibility and practicality of the questionnaire used in the pilot survey. By administering the questionnaire to a small sample, researchers can evaluate its clarity, ease of understanding, and suitability for the target population. This assessment helps identify any issues or improvements needed for the main survey.

(e) To develop the theoretical model as a tool for decision-making for employee satisfaction with energy-efficiency applications in green buildings:

(f) The pilot survey provides an opportunity to validate and refine the theoretical model employed in the substantive study. Researchers can identify key factors influencing employee satisfaction with energy-efficiency applications in green buildings by analysing the pilot survey responses. This information allows for developing a robust theoretical framework to guide decision-making in the field.

4.6.2 Methodology of the pilot survey

The pilot survey employed a sample of participants from office spaces and factory settings. Stratified random sampling techniques ensured representative samples from different zones within each space. Office spaces were zoned based on the distance from windows, while factories were divided into four equal zones.

The questionnaire was distributed to the selected participants, and data collection was conducted online and offline. Participants were given clear instructions on completing the questionnaire, and any queries or concerns were addressed promptly. The data collected included responses to Likert-scale questions, open-ended questions, and demographic information.

4.6.3 Conclusions of the Pilot Survey

The pilot survey yielded valuable insights into employee satisfaction with green buildings. The following conclusions were drawn from the analysis of the pilot survey data:

(a) Simplification of Questions: Feedback from participants highlighted the need for simplification of the questionnaire. Scientific terminology was identified as a barrier to understanding, leading to confusion among respondents. To address this, the research team simplified the language used in the main survey, making it more accessible and understandable to all participants.

(b) Respondents' Familiarity with Terminology: Approximately 60% of the pilot survey respondents reported unfamiliarity with specific terms related to green technology, humidity, and HVAC. This finding underscored the importance of using language that aligns with the respondents' familiarity and expertise. Consequently, the research team adapted the questionnaire by providing explicit definitions and explanations for key terms to ensure a common understanding among participants.

(c) Length of Questionnaire: Participants expressed concerns about the length of the questionnaire in the pilot survey. It was observed that a lengthy questionnaire can lead to respondent fatigue and potentially compromise data quality. To address this issue, the research team reduced the number of questions in the main survey, focusing on the most relevant and impactful aspects of employee satisfaction with green buildings.

(d) Likert Scale Distribution: The pilot survey indicated that participants perceived the 7-point Likert scale distribution used in the questionnaire as complex. Feedback suggested that respondents found differentiating between the various response options challenging. The Likert scale was reduced to 5 points in the main survey to simplify the response process, allowing for a more straightforward assessment of employee satisfaction levels.

(e) Open-ended Questions: Approximately 80% of the open-ended questions in the pilot survey were left unanswered. This indicated respondents' potential reluctance or difficulty in providing detailed written responses. To mitigate this issue and enhance participation, the research team removed open-ended questions from the main survey and focused solely on structured response formats.

(f) Variability in Satisfaction among Employees Sharing the Same Workstation:

(g) An intriguing finding from the pilot survey was the significant variability in overall satisfaction levels among employees sharing the same workstation. Despite working nearby, the mean satisfaction score of 3.562 and a standard deviation of 1.968 highlighted the diversity of experiences and the need for further investigation into the factors influencing satisfaction within shared workspaces.

(h) Reliability Assessment: The Cronbach's alpha coefficient was calculated to evaluate the internal consistency and reliability of the questionnaire. The pilot survey yielded a Cronbach's alpha of 0.939, indicating high reliability. This provided confidence in the consistency of the survey instrument and its ability to measure the intended constructs.

4.7 The main survey

Based on the conclusions drawn from the pilot survey, several adjustments were made to the main survey. The questionnaire is attached in APPENDIX A.

(a) Simplification and Translation of Questions: The questions were simplified further to ensure clarity and understanding among participants. Additionally, the questionnaire was translated into local languages such as Sinhalese and Tamil to accommodate participants with diverse linguistic backgrounds.

(b) Removal of Scientific Terminology: Scientific words and technical terminology were eliminated or replaced with simpler language to enhance participant comprehension and engagement.

(c) Likert Scale Reduction: The Likert scale was reduced from 7 points to 5 points to address participant feedback in the main survey. This modification aimed to facilitate more accessible response selection and minimise confusion.

(d) Reduction in Number of Questions: The number of questions was reduced to avoid respondent fatigue and improve overall response rates. Only the most relevant and impactful questions were included in the main survey.

(e) Focus on Micro-Climate: Additional questions were added to explore the influence of micro-climate factors on employee experiences within shared workstations. This adjustment aimed to capture a more nuanced understanding of the variance in satisfaction levels among employees working in close proximity.

The pilot survey played a crucial role in refining the methodology and questionnaire design for the main study of employee satisfaction with green buildings. The research team optimised the survey instrument to ensure maximum reliability, validity, and participant engagement by addressing the identified limitations and incorporating participant feedback. The adjustments made based on the pilot survey conclusions will contribute to the collection of robust and meaningful data in the main study, ultimately enhancing our understanding of employee satisfaction with green buildings and informing decision-making for sustainable workplace design and management.

4.7.1 Sampling Error calculations

In the case of factory spaces, where a relatively large number of employees were present, the sampling error was targeted to be between 5% and 6% (Table 4.17). This approach was adopted to achieve higher precision in the estimates and minimise the variability between the sample and the population. By maintaining a narrow margin of error, the researchers aimed to enhance the reliability of the findings and provide accurate insights into the factory spaces.

Table 4.17: Sampling error for factories

Building Category	Total employees in the selected layout	Selected sample	Sampling error (for F)
F1	3842	200	5%
F2	1104	180	5.6%
F3	1700	100	5.8%
F4	2450	160	5%
F5	1262	90	5.8%
F6	1537	100	5%
Total	11895	830	

Conversely, a minimum sampling percentage of 25% of the total population was deemed appropriate in office spaces with comparatively fewer employees (Table 4.18). This decision balanced the need for accuracy with practical limitations such as resource constraints and accessibility. While a larger sample size would have offered greater precision, it was necessary to consider the feasibility of surveying a significantly higher percentage of the population. Sampling at least 25% of the people ensured a reasonable representation of the office spaces while considering practical considerations.

Table 4.18: Percentage of population representation for general office buildings

Building Category	Total employees in the selected layout	Selected sample	Population representation (for O)
O1	176	60	34%
O2	400	150	37.5%
O3	150	60	40%
O4	180	70	38%
O5	100	25	25%
O6	100	25	25%
O7	220	120 (89)	40%
O8	150	60 (43)	40%
Total	1476	570	

Random sampling techniques were implemented to minimise biases and ensure that each individual in the population had an equal chance of being selected. Furthermore, factors such as demographic diversity and geographic distribution were considered during the sample selection process to reflect the population's characteristics accurately.

To further strengthen the validity and reliability of the findings, statistical analyses were conducted to estimate the sampling error and establish confidence intervals around the sample estimates. These analyses provided valuable insights into the precision and variability of the results, allowing for a comprehensive understanding of the study's robustness. The statistical findings were essential for interpreting the results in the context of the population and assessing the potential impact of sampling error on the conclusions drawn from the study.

The sampling strategies employed in this study aimed to minimise sampling error and enhance the representativeness and generalizability of the findings. By targeting a specific margin of error in factory spaces and sampling at least 25% of the population in office spaces, the researchers sought to balance accuracy and practical constraints. Using random sampling techniques, consideration of relevant variables, and statistical analyses further contributed to the reliability and validity of the research findings

Based on the objectives of the research study and the nature of the questions given in the questionnaire, it mainly focused on Thermal comfort, Visual comfort and Indoor Air Quality Satisfaction. The questions are shown in the Table 4.19.

Table 4.19: Nature of the questions used to get the average IEQ satisfaction of the employees

Independent Variable	Question Descriptive
Thermal Comfort	Distance between the seat and the window
	Are you feeling too cold in this room someday?
	Are you feeling too hot in this room someday?
	Satisfaction level of Room temperature
	Satisfaction level of freedom to adjust the room temperature
Visual Comfort	Satisfaction level of the attractiveness of the room
	Satisfaction level of Sunlight glare of the room
	Satisfaction level of Distance between the seat and the window
	Satisfaction level of Lighting level
	Satisfaction level of Light Glare level in the room
IAQ Satisfaction	Satisfaction level of Ventilation in the room
	Satisfaction level of Distance between the seat and the window
	Satisfaction level of Dryness of the room
	Satisfaction level of Fresh air inside the room
	Satisfaction level of Dust level in the room

4.8 Pre-Screening Survey Results: Ensuring Data Integrity and Quality

The pre-screening phase of a survey is a crucial step in the data collection process. It involves screening the collected responses to ensure data integrity and quality. This article analyses the pre-screening survey results, focusing on the number of responses collected, the exclusion criteria applied, and the final sample size available for analysis.

Total Responses Collected:

A total of 1,369 survey questionnaires were distributed to the target population. These questionnaires aimed to gather valuable insights into various aspects of the research topic. The response rate was encouraging, indicating a high participant engagement and interest in the study.

Responses Excluded:

However, not all the distributed questionnaires were suitable for further analysis. Several exclusion criteria were applied to ensure the quality and validity of the data. The following sections detail the responses excluded at each stage of the pre-screening process.

(a) Exclusion of Non-Returned Questionnaires: Out of the total distributed questionnaires, 38 were never returned. These non-returned questionnaires were not considered for further analysis. The reasons for non-return could vary, such as participants' forgetfulness or lack of interest. It is essential to acknowledge that non-returned questionnaires may introduce a potential bias if those who did not respond differ systematically from those who did.
(b) Exclusion of Partially Filled Questionnaires: From the remaining responses, 215 questionnaires were partially filled. These questionnaires lacked essential information or contained incomplete answers, rendering them unsuitable for analysis. Excluding these partially filled questionnaires ensures data integrity and maintains the dataset's quality.
(c) Exclusion of Damaged Questionnaire Sheets: Out of the screened responses, a small number of five questionnaire sheets were found to be damaged. These damaged sheets could include torn or illegible answers, making extracting meaningful data impossible. As a result, these five questionnaire sheets were excluded from the dataset.

Final Sample Size:

After applying the exclusion criteria, the final sample size for analysis was 1,091 responses. These responses met the required completeness, legibility, and suitability standards for the research objectives. The final sample size of 1,091 provides a substantial dataset for analysis, allowing for robust statistical inference and generalizability of the findings. Figure 4.20 describes the pre-screening process.

(a) Importance of Pre-Screening: The pre-screening phase is crucial to ensure data quality and reliability. By applying strict criteria for response inclusion, researchers can minimise the risk of bias and improve the study's overall validity. It helps filter out incomplete or unreliable responses, ensuring that the analyzed data accurately represent the target population.
(b) Limitations: While pre-screening enhances data integrity, it is essential to acknowledge its limitations. Exclusion criteria may introduce potential biases, as participants who did not return the questionnaire or partially filled it may have different characteristics from those included in the final sample. Also, damaged questionnaire sheets may lead to data loss and reduce the sample size. Researchers should be aware of these limitations and consider them when interpreting the results.
(c) The pre-screening survey results provide valuable insights into the data collection process and the dataset's quality. Through the application of exclusion criteria, non-returned questionnaires, partially filled questionnaires, and damaged questionnaire sheets were identified and excluded from the analysis. The final sample size of 1,091 responses ensures a robust dataset for analysis, maintaining data integrity and enhancing the validity of the

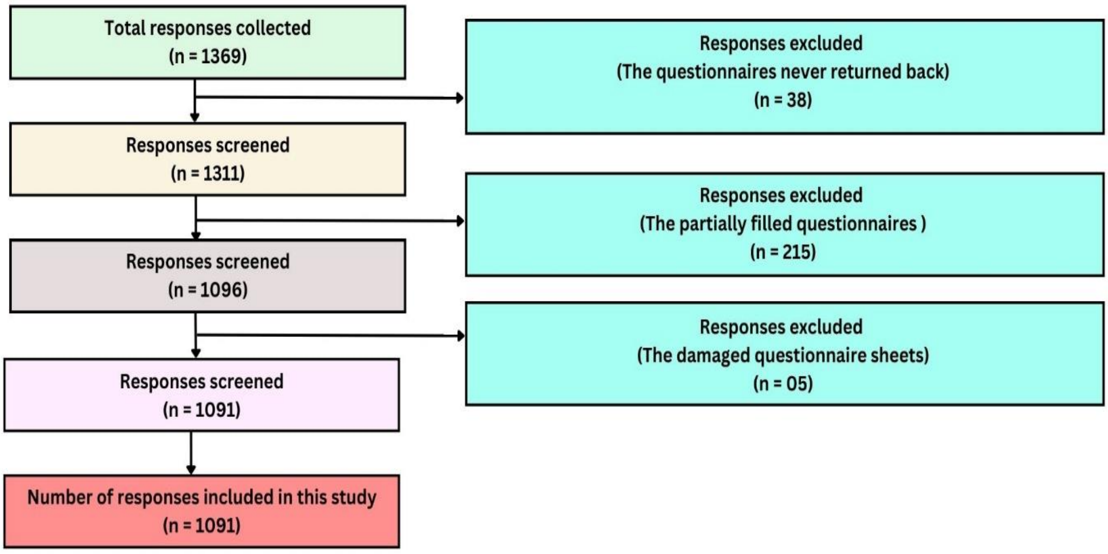


Figure 4.20: Pre-screening steps of the survey results

research findings. Pre-screening plays a vital role in maintaining the quality of survey data, allowing researchers to draw accurate conclusions and make informed decisions based on the analyzed dataset.

4.9 Descriptive analysis

The descriptive analysis was conducted to identify the survey's demographic distribution, response behaviour and hypothesis testing.

(a) Workplace Distribution: The distribution of respondents across different workplace settings provides a comprehensive understanding of employee experiences in green buildings. With 434 respondents from office spaces and 657 from factories (Figure 4.21), the data allows for a comparison between these two distinct work environments. This comparison can show potential differences in employee satisfaction, comfort, and environmental conditions between office spaces and factory settings. It enables researchers to examine factors such as indoor air quality, thermal comfort, lighting, and overall workspace design that may vary between these contexts.

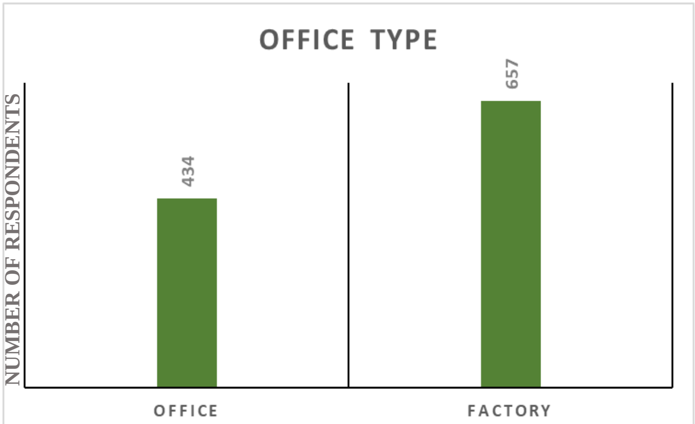
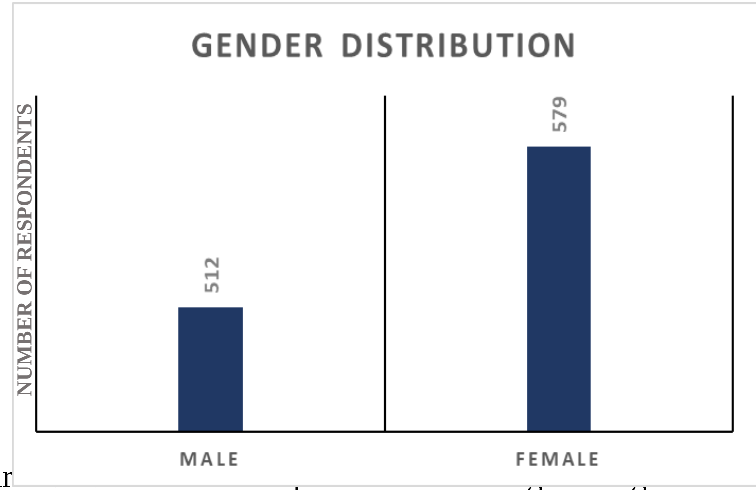


Figure 4.21: Number of respondents according to the office type

(b) Gender Distribution: The gender distribution of the respondents, with 512 males and 579 females (Figure 4.22), offers valuable insights into potential gender-related differences in the perception and experience of green buildings. By examining the data through a gender lens, researchers can explore whether male and female employees have different preferences, needs, and satisfaction levels regarding indoor environmental quality comfort. This gender analysis can create more inclusive and gender-responsive green building design and management strategies.



Figur

(c) Hometown Climate Zones: Categorizing respondents' hometowns into three main climatic zones in Sri Lanka (dry, wet, and intermediate) allows for examining the influence of climatic conditions on employee satisfaction and comfort. The more significant number of respondents from the wet climate zone (850) indicates the significance of studying the impact of high humidity, rainfall, and other weather factors on occupant experiences. Additionally, the distribution of respondents across the dry (61) and intermediate (180) zones allows for comparisons between different climatic regions (Figure 4.23), providing insights into the unique challenges and opportunities posed by each climate type in green buildings.

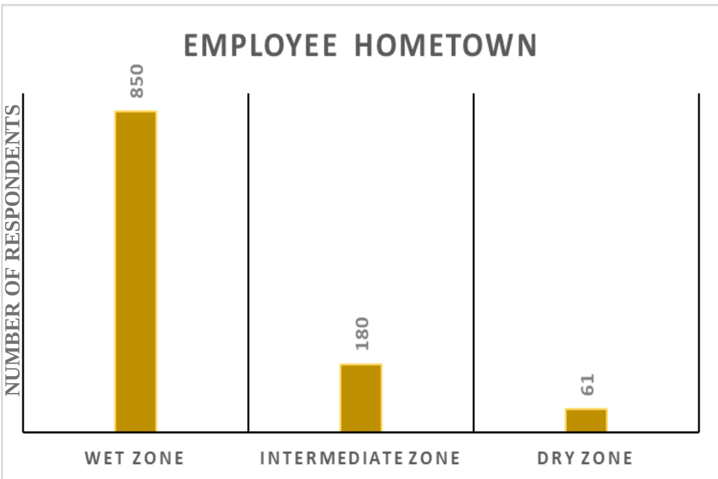


Figure 4.23: Number of respondents according to the home town climatic zone

(d) Age Groups: Classifying respondents into different age groups enables researchers to analyze potential variations in IEQ comfort levels among different generations. Most respondents fall into the 25-29 and 40-44 age groups (Fiure 4.24).

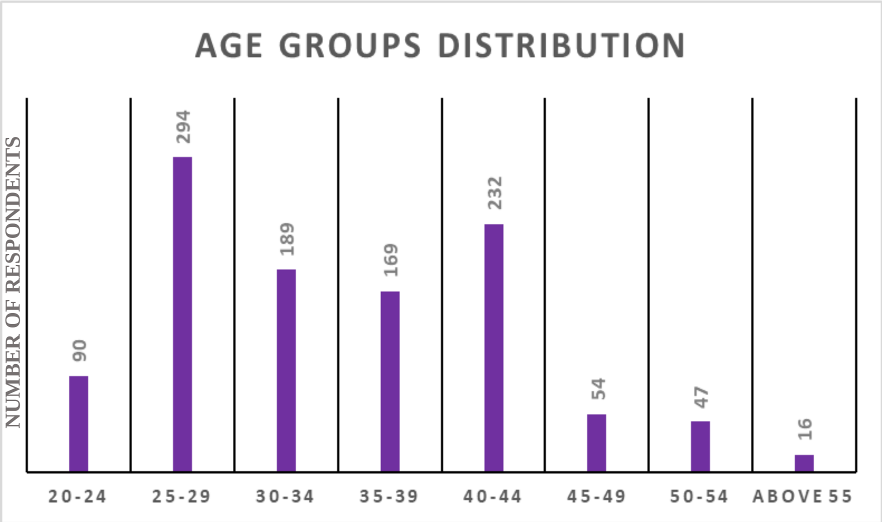


Figure 4.24: Number of respondents according to the age groups

(e) Working Hours: Understanding the distribution of working hours among respondents is crucial for assessing the impact of time spent in the workplace on employee satisfaction and comfort. The prevalence of 8-hour workdays among most respondents indicates a standard work schedule. However, the presence of respondents working for 10 and 12 hours (Figure 4.25), especially those engaged in double shifts, warrants attention. Longer working hours can have implications for employee IEQ comfort. Analyzing the satisfaction levels and comfort of employees working extended hours can provide insights into the effectiveness of green building features in supporting their needs during prolonged periods of occupancy.

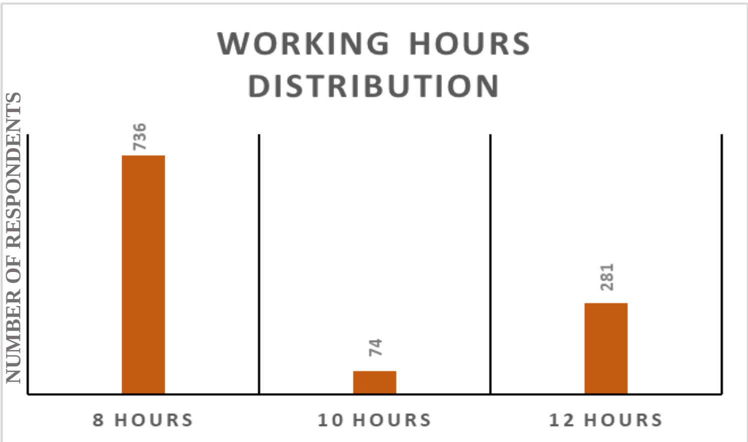


Figure 4.25: Number of respondents according to the working hours

(f) Work Experience: Analyzing the work experience of respondents within the same building offers insights into the potential variations in satisfaction and comfort over time. The majority of respondents reported work experience of 2-5 years. This finding highlights the importance of understanding the impact of work experience on employee perceptions of the indoor environment.

Figure 4.26: Number of respondents according to the work experience

4.10 Accuracy and reliability analysis

Cronbach's alpha is a statistical measure to assess a scale or questionnaire's reliability or internal consistency. It measures the extent to which items within a scale or questionnaire are interrelated and measures the same underlying construct. In other words, Cronbach's alpha helps to determine the extent to which the items in a survey instrument are reliable and consistent in measuring a specific variable or concept.

Cronbach's alpha is based on the correlation coefficients between the items within a scale. It ranges from 0 to 1, with higher values indicating greater internal consistency and reliability (Bujang et al., 2018). A Cronbach's alpha value of 1 represents perfect reliability, while 0 shows no reliability.

The calculation of Cronbach's alpha considers the number of items in the scale, the inter-item correlations, and the total score variance. It is important to note that Cronbach's alpha is influenced by the number of items in the scale: scales with more things tend to have higher Cronbach's alpha values.

Interpreting Cronbach's alpha values:

a. Values above 0.7 are generally considered acceptable and indicate good internal consistency.
b. Values above 0.8 are considered very good, indicating high internal consistency.
c. Values below 0.7 may suggest low internal consistency and a need to examine the scale or questionnaire further.

It is important to note that the specific field of study and the nature of the measured construct can influence the acceptable range of Cronbach's alpha values. Some research fields may require higher levels of internal consistency than others.

In the context of survey research, Cronbach's alpha is commonly used to assess the reliability of multi-item scales that measure constructs such as attitudes, perceptions, or behaviours. By calculating Cronbach's alpha, researchers can determine whether the items in their survey instrument reliably measure the intended construct. If Cronbach's alpha value is low, it may indicate the need to revise or remove certain items from the scale to improve reliability.

The limitations of Cronbach's alpha assume that all items in the scale measure the same construct, and it may not account for other sources of error or variation. Additionally, Cronbach's alpha is sensitive to the number of items in the scale, and small-scale sizes may result in less reliable estimates.

Table 4.20: Cronbach's alpha values for the models

Variable	**Cronbach's α**
Thermal Comfort	0.924
Visual Comfort	0.854
IAQ Satisfaction	0.822

`

4.11 Exploring the normality of the data

A normality test was performed using different methods and obtained several results, including the Kolmogorov-Smirnov (K-S) test and Shapiro-Wilk test, as well as graphical representations such as QQ plots and box plots.

The Kolmogorov-Smirnov (K-S) test is a nonparametric test that compares the observed cumulative distribution function (CDF) of the data to the expected CDF of a specified distribution, typically the normal distribution in this context. The K-S test provides a statistical test statistic and a p-value that assesses the goodness of fit between the observed data and the expected normal distribution. A significant p-value suggests that the data significantly deviates from normality (Significance level is 0.05) (Whitnall et al., 2011).

The Shapiro-Wilk (S-W) test is another commonly used test for normality. It calculates a test statistic based on the correlation between the observed data and the expected average values. The Shapiro-Wilk test also provides a p-value that indicates whether the data significantly departs from a normal distribution (J. Wei, 2022).

In addition to these numerical tests, graphical methods were used to assess normality. The QQ (quantile-quantile plot) compares the observed data's quantiles to a normal distribution. The data is usually distributed if the data points fall along a straight line. Deviations from the straight line suggest departures from normality.

The observed values versus the expected regular QQ plot help visualize the departure from normality. If the points in the plot deviate significantly from the straight line, it suggests non-normality. Additionally, the observed values versus deviation from the normal plot illustrate the variations of the experimental values from the expected average values, providing further insight into the distribution characteristics.

Furthermore, the box plot is a graphical representation that displays the data distribution, including measures such as the median, quartiles, and outliers. It allows for a visual assessment of the skewness and symmetry of the data, indicating a departure from normality.

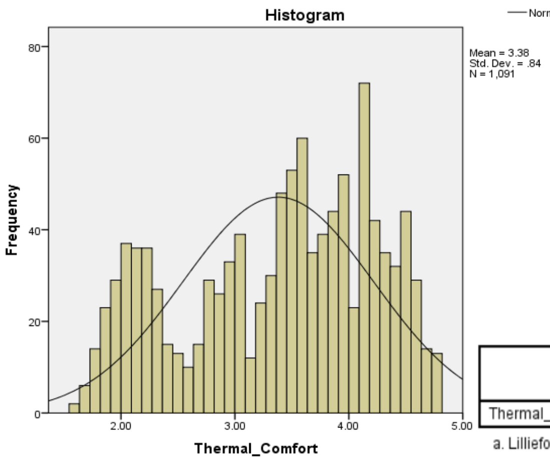
When interpreting the results of these normality tests and graphical representations, it is essential to consider the significance level (alpha) used and the sample size. A non-significant p-value suggests that the data can be assumed to follow a normal distribution, while a significant p-value indicates departures from normality.

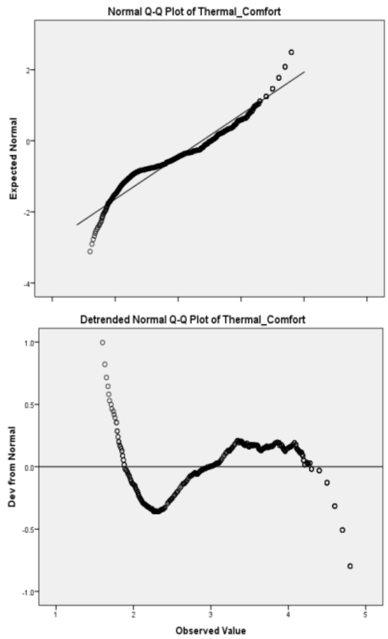
4.11.1 Normality Test for Thermal Comfort

The analysis rejected the null hypothesis and reflected that the data are not normally distributed. The K-S and S-W p values are 0.000. Figure 4.27 illustrates the normality test results.

Descriptives

			Statistic	Std. Error
Thermal_Comfort	Mean		3.3801	.02543
	95% Confidence Interval for Mean	Lower Bound	3.3302	
		Upper Bound	3.4300	
	5% Trimmed Mean		3.3946	
	Median		3.5300	
	Variance		.705	
	Std. Deviation		.83981	
	Minimum		1.60	
	Maximum		4.80	
	Range		3.20	
	Interquartile Range		1.34	
	Skewness		-.354	.074
	Kurtosis		-1.002	.148





Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Thermal_Comfort	.087	1091	.000	.949	1091	.000

a. Lilliefors Significance Correction

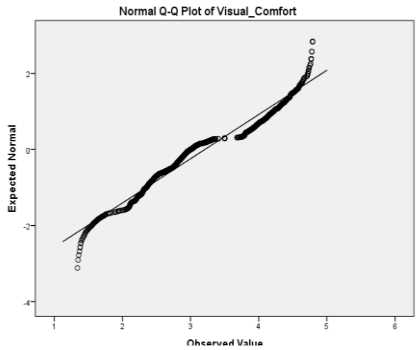
Figure 4.27: The normality test illustration of thermal comfort

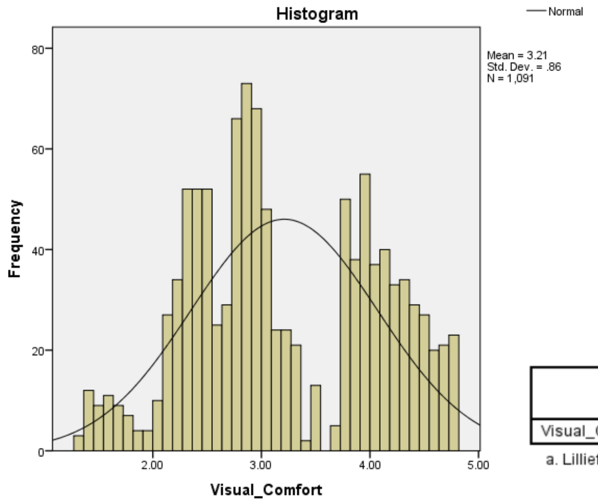
4.11.2 Normality Test for Visual Comfort

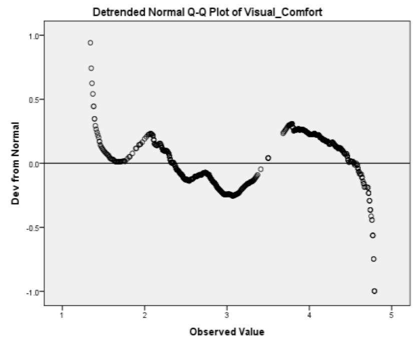
The analysis rejected the null hypothesis and reflected that the data are not normally distributed. The K-S and S-W p values are 0.000. Figure 4.28 illustrates the normality test results.

Descriptives

			Statistic	Std. Error
Visual_Comfort	Mean		3.2100	.02603
	95% Confidence Interval for Mean	Lower Bound	3.1589	
		Upper Bound	3.2610	
	5% Trimmed Mean		3.2181	
	Median		3.0000	
	Variance		.739	
	Std. Deviation		.85976	
	Minimum		1.34	
	Maximum		4.79	
	Range		3.45	
	Interquartile Range		1.47	
	Skewness		.075	.074
	Kurtosis		-1.009	.148







Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Visual_Comfort	.109	1091	.000	.958	1091	.000

a. Lilliefors Significance Correction

Figure 4.28: The normality test illustration of visual comfort

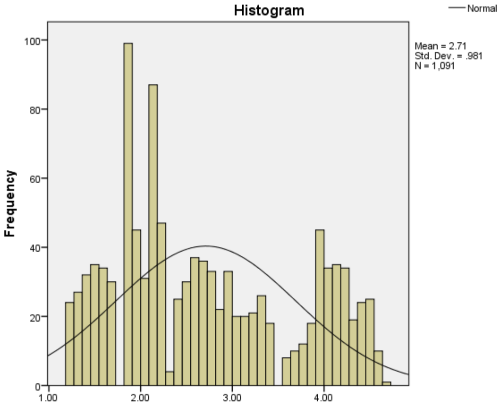
4.11.3 Normality Test for IAQ Satisfaction

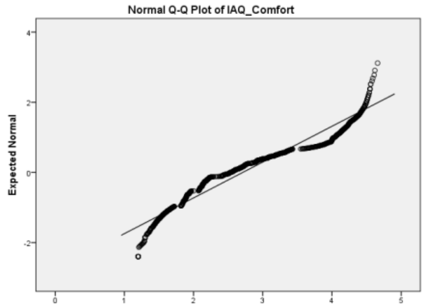
The analysis rejected the null hypothesis and reflected that the data are not normally distributed. The K-S and S-W p values are 0.000. Figure 4.29 illustrates the normality test results.

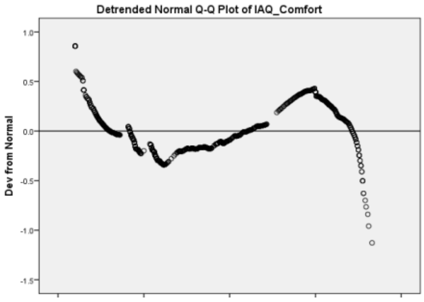
According to the analysis of the main three independent variables, the data were not normally distributed, and further research will be conducted using non-parametric statistical tests.

Descriptives

			Statistic	Std. Error
IAQ_Comfort	Mean		2.7107	.02969
	95% Confidence Interval for Mean	Lower Bound	2.6525	
		Upper Bound	2.7690	
	5% Trimmed Mean		2.6918	
	Median		2.5300	
	Variance		.962	
	Std. Deviation		.98059	
	Minimum		1.20	
	Maximum		4.66	
	Range		3.46	
	Interquartile Range		1.68	
	Skewness		.392	.074
	Kurtosis		-1.103	.148







Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IAQ_Comfort	.131	1091	.000	.931	1091	.000

a. Lilliefors Significance Correction

Figure 4.29: The normality test illustration of IAQ comfort

4.12 Hypothesis testing

A series of hypothesis analyses were conducted to understand the nature of employee responses to gain a deeper understanding. These analyses aimed to explore the relationships, patterns, and associations between various variables related to employee responses in the context of IEQ comfort. This study's hypotheses formulated and tested provide valuable insights into the underlying factors influencing employee responses.

One of the primary hypotheses examined in this research focused on the relationship between employee satisfaction and factors such as building features, indoor environmental quality, and workplace design. It was hypothesized that there would be a positive correlation between these variables and employee satisfaction.

Another set of hypotheses explored the differences in employee responses based on demographic characteristics such as gender, age, and work experience. These hypotheses aimed to investigate whether these demographic factors influenced employee perceptions, satisfaction, or preference. Non-parametric statistical tests, such as the Mann-Whitney U test or the Kruskal-Wallis test, were employed to examine the significance of these differences. The findings shed light on the variations in employee responses across different demographic groups, allowing for a more comprehensive understanding of the factors that shape employee experiences. When the test results indicate a significant difference among the groups, post-hoc pairwise comparisons are often conducted to determine which groups differ. The pairwise comparison in Kruskal-Wallis results involves comparing each group against every other group to determine if significant differences exist in the distribution of the variable being tested (S. Lee & Lee, 2018). The results of pairwise comparisons are typically presented in the form of p-values or statistical significance levels. Interpreting pairwise comparisons in Kruskal-Wallis results involves considering the p-values or significance levels associated with each comparison. A significant p-value indicates evidence of a significant difference between the two groups being compared.

The hypothesis analyses were conducted using appropriate statistical techniques, ensuring the validity and reliability of the findings. The significance levels and effect sizes were considered to evaluate the practical implications of the results. In cases where the data supported the hypotheses, the findings provided evidence to support the theoretical frameworks or conceptual models underpinning the study.

Hypothesis 01

$H1_0$ = There is no significant difference between the respondents' **gender group** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

$H1_1$ = There is a significant difference between the respondents' **gender group** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	.667	Retain the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	.885	Retain the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	.835	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

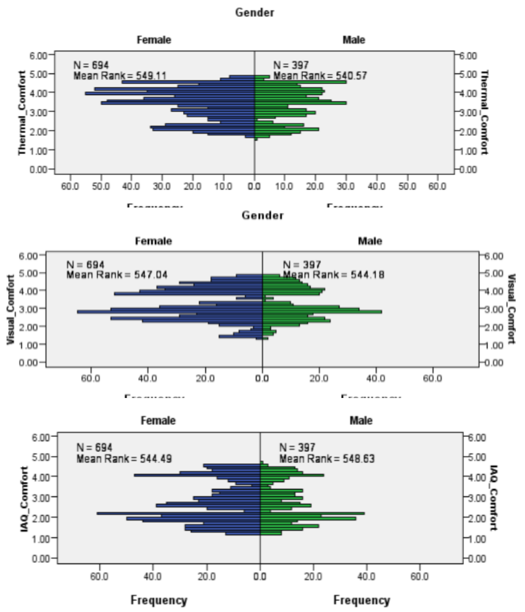


Figure 4.30: Independent samples Mann-Whitney U test results illustration of gender group and test satisfaction variables

The p-value is more significant than 0.05 for all three variables, and because of the H_0 is accepted. Therefore, it is evident that there is no significant difference between the responses and the gender of the respondents (Figure 4.30). Since there is no impact on the gender of the respondents, the Gender variable is omitted from the model framework.

Hypothesis 02

Office

$H2i_0$ = There is no significant difference between the respondents' **age group** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

$H2i_1$ = There is a significant difference between the respondents' **age group** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	.350	Retain the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	.179	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

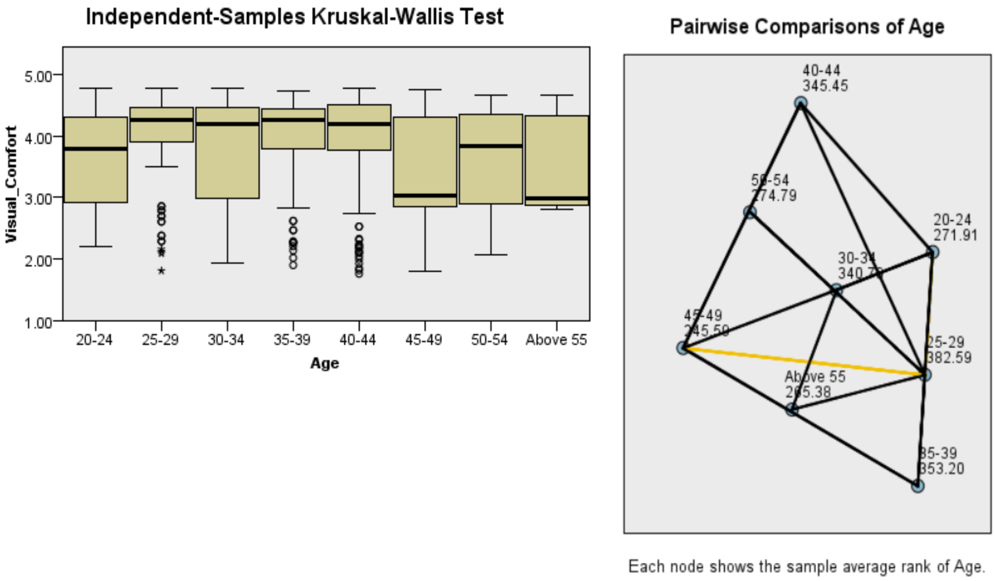


Figure 4.31: Independent samples Kruskal-Wallis test results illustration of age group and test satisfaction variables in office buildings

The p-value is greater than 0.05 for thermal comfort and IAQ satisfaction, and H_0 is accepted for those variables. The p-value for visual comfort is less than 0.05 (Figure 4.31), rejecting the H_0. This means that the particular age group in general office spaces can significantly impact visual comfort responses.

The pairwise comparison reflected a significant difference between (p = 0.007) Age groups 45-49 and 25-29.

Factory

H2ii$_0$ = There is no significant difference between the respondents' **age group** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

H2ii$_1$ = There is a significant difference between the respondents' **age group** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	.003	Reject the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	.142	Retain the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Age.	Independent-Samples Kruskal-Wallis Test	.240	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

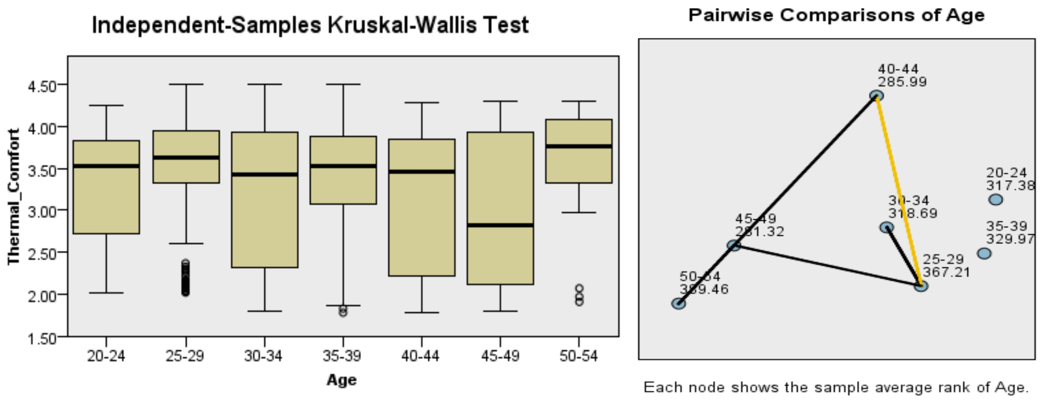


Figure 4.32: Independent samples Kruskal-Wallis test results illustration of age group and test satisfaction variables in factory buildings

The p-value is greater than 0.05 for visual comfort and IAQ satisfaction, and H$_0$ is accepted for those variables. The p-value for thermal comfort is less than 0.05 (Figure 4.32), resulting in rejecting the H$_0$. This means that the particular age group in factory spaces can significantly impact the response to visual comfort.

The pairwise comparison reflected a significant difference between (p = 0.02) Age groups 40-44 and 25-29.

Hypothesis 03

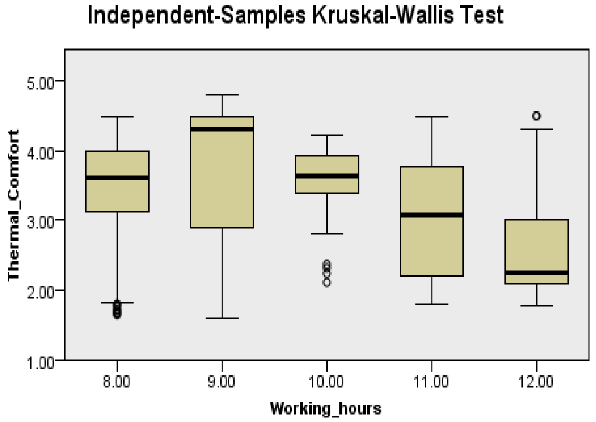
$H3_0$ = There is no significant difference between the respondents' **working hours** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

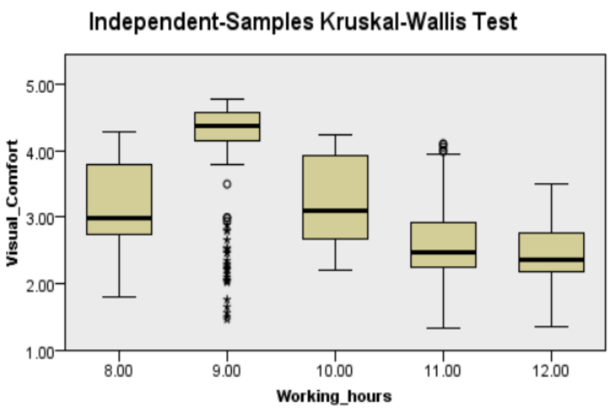
$H3_1$ = There is a significant difference between the respondents' **working hours** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Working_hours.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Working_hours.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Working_hours.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.





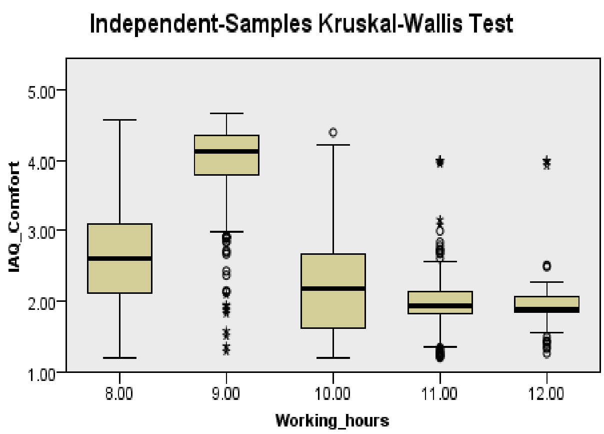


Figure 4.33: Independent samples Kruskal-Wallis test results illustration of working and test satisfaction variables

The p-value is less than 0.05 for all thermal comfort, visual comfort and IAQ satisfaction (Figure 4.33), and H_0 is rejected for those variables. This means that the working hours in factory and general office spaces can significantly impact the response to IEQ comfort.

Hypothesis 04

$H4_0$ = There is no significant difference between the respondents' **home town** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

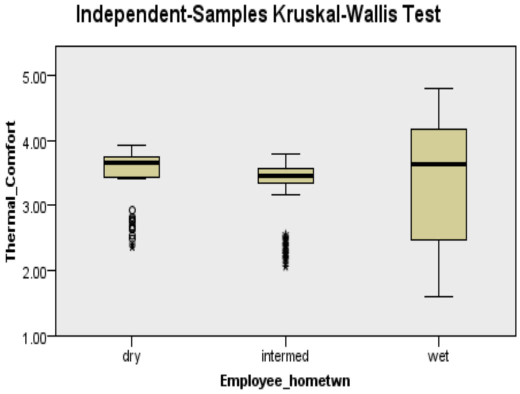
$H4_1$ = There is a significant difference between the respondents' **home town** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

The p-value is less than 0.05 for all thermal comfort, visual comfort and IAQ satisfaction, and H_0 is rejected for those variables (Figure 4.34).

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Employee_hometwn.	Independent-Samples Kruskal-Wallis Test	.021	Reject the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Employee_hometwn.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Employee_hometwn.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.



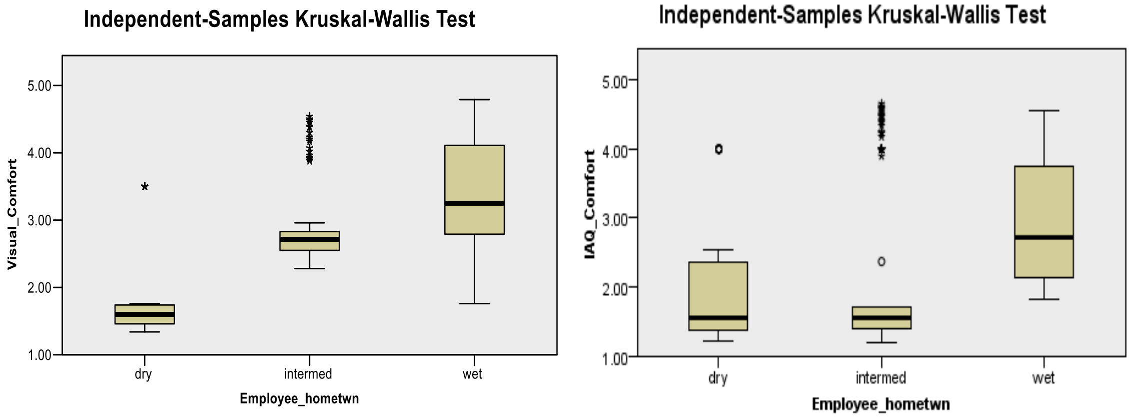


Figure 4.34: Independent samples Kruskal-Wallis test results illustration of home town and test satisfaction variables

The employee hometowns were categorized into three main climatic zones in Sri Lanka and analysed accordingly. This means that the particular climatic zone of the hometown of the employees in the factory and general office spaces can significantly impact the response to IEQ comfort.

The PowerBI output reflected a visible satisfaction level difference in the thermal comfort variable (Figure 4.35). It visualized that if the employee's hometown is in a dry zone or intermediate zone and the office/factory is in a wet zone, their overall thermal satisfaction is higher than the employees' hometown in a wet zone and the office/factory in a dry zone. It also visualized that there is no significant visible difference in the location if the employee's hometown and the office/factory locations are in the same climatic zone. The diameter of the circles on the map represent the magnitude of the satisfaction and they are directly proportional.

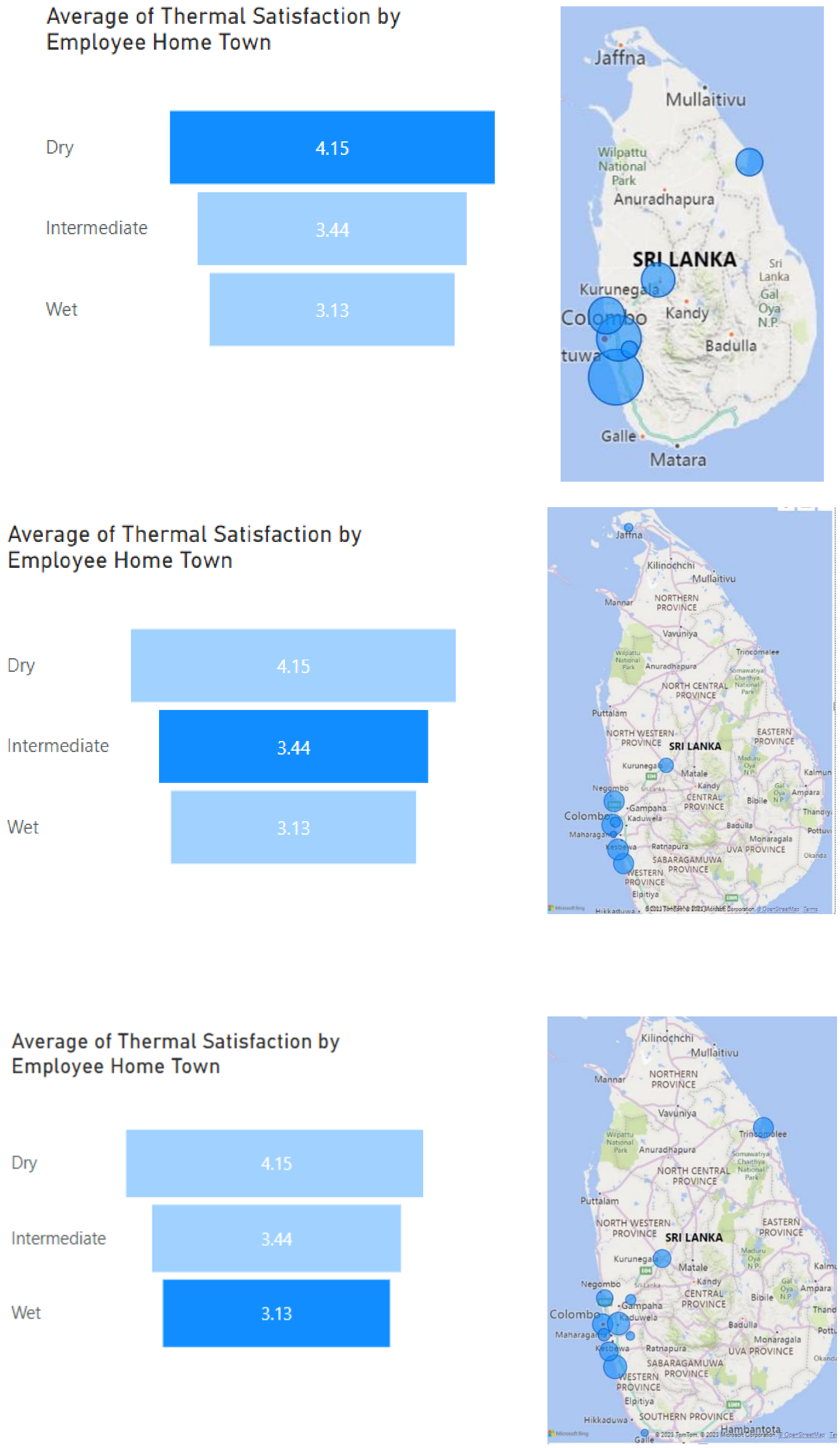


Figure 4.35: The satisfaction level Vs climatic zone of the hometown of the employees

Generated by: PowerBI

Hypothesis 05

Office

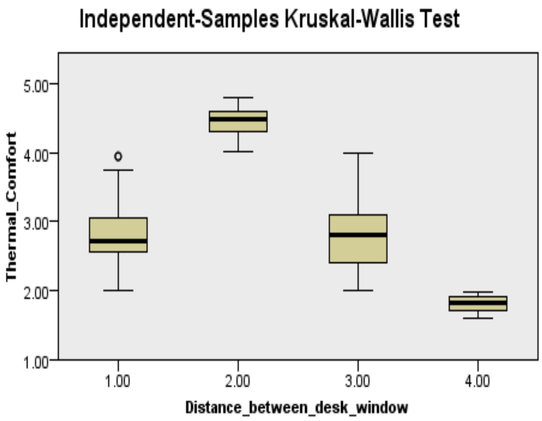
$H5i_0$ = There is no significant difference between the respondents' **distance between the window** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

$H5i_1$ = There is a significant difference between the respondents' **distance between the window** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Distance_between_desk_window.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Distance_between_desk_window.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Distance_between_desk_window.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.



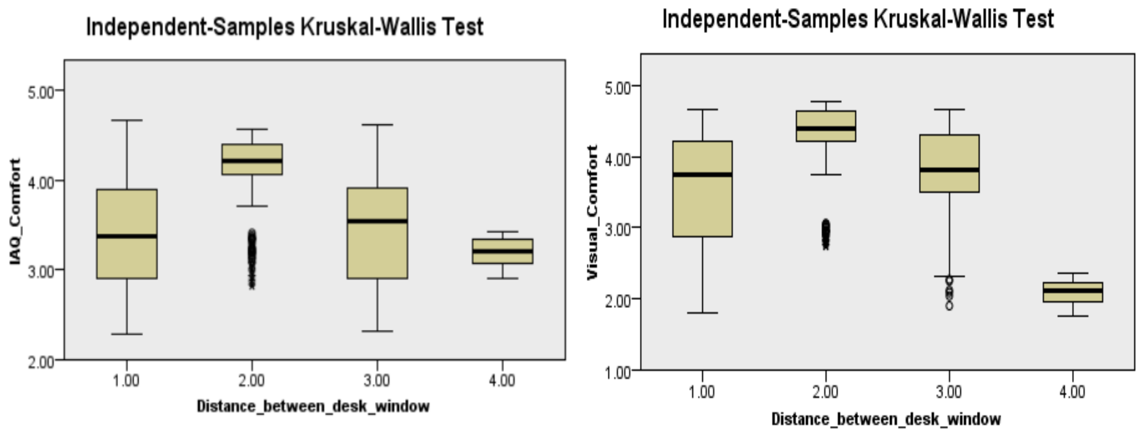


Figure 4.36: Independent samples Kruskal-Wallis test results in illustration of the distance between the window and test satisfaction variables in office buildings

The p-value is less than 0.05 for all thermal comfort, visual comfort and IAQ satisfaction (Figure 4.36), and H_0 is rejected for those variables. This means the distance from windows to the workstation in office spaces can significantly impact the response to IEQ comfort.

4.13 The IEQ comfort and distance between the window in general office spaces

The distance to the window can influence several aspects of IEQ, including natural light exposure, ventilation, and thermal comfort. Numerous studies have indicated that employees generally express higher satisfaction when close to windows. This finding can be attributed to several factors related to the benefits associated with window proximity.

(a) Natural Light: Being closer to windows gives individuals greater access to natural light. Natural light is known to affect mood, well-being, and productivity positively. Exposure to natural light has been linked to reduced stress, improved circadian rhythm regulation, and increased vitamin D synthesis. The presence of daylight can also enhance visual comfort, reduce eyestrain, and create a more pleasant and visually stimulating work environment.

(b) Views and Connection to the Outdoors: Proximity to windows allows employees to enjoy views of the external environment, which can have a positive psychological impact. Views of nature, green spaces, or cityscapes can contribute to feelings of relaxation, restoration, and connection to the outside world. Viewing natural elements has improved cognitive function, reduced stress, and enhanced creativity.

(c) Sense of Space and Openness: Being close to windows can create a perception of spaciousness and openness. The ability to gaze beyond the confines of the office and observe the surrounding environment can enhance the perceived spaciousness of the workspace, thereby fostering a sense of freedom and comfort.

(d) Visual and Environmental Variability: Window proximity exposes individuals to a greater variety of visual stimuli, including changes in natural light patterns, weather conditions, and external activities. This variability can break the monotony of the work environment and contribute to a more engaging and dynamic setting. Additionally, being closer to windows can allow individuals to regulate their immediate environment by allowing for better control of views, privacy, and exposure to natural ventilation.

Not all employees may prefer or benefit equally from proximity to windows. Personal preferences, sensitivity to light, thermal comfort requirements, and task demands may influence individual responses. Additionally, specific workplace characteristics, building design, and climate conditions can control how window proximity affects employee satisfaction.

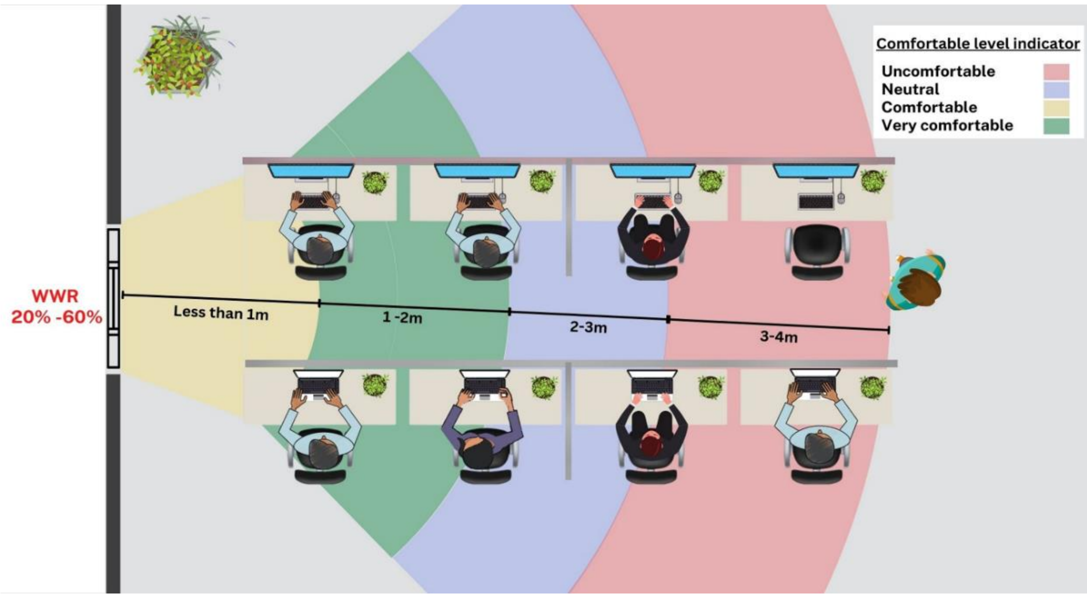


Figure 4.37: The comfortable levels of the employees with respect to the distance from the window- General office buildings

The results were further analysed to identify the comfort zones of the employees. It was visible that the employees within a 1-2m distance from the window were more comfortable than those in 0-1m (Figure 4.37).

An analysis was conducted to determine why the employees weren't right next to the window delighted as employees in 1-2 m distance from the window. No shading devices were used in most buildings observed during the study. A hypothesis was created to understand whether there is a correlation between the sunlight glare and the satisfaction of the distance between the window and the work desk.

Correlations

			Distance_between_desk_window	Sun_light_glare
Spearman's rho	Distance_between_desk_window	Correlation Coefficient	1.000	.926**
		Sig. (2-tailed)	.	.000
		N	434	434
	Sun_light_glare	Correlation Coefficient	.926**	1.000
		Sig. (2-tailed)	.000	.
		N	434	434

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4.38: Correlation between satisfaction level of the sunlight glare and the distance between work desk and window

And it revealed a high correlation (r=0.928) between the considered variables.

Factory

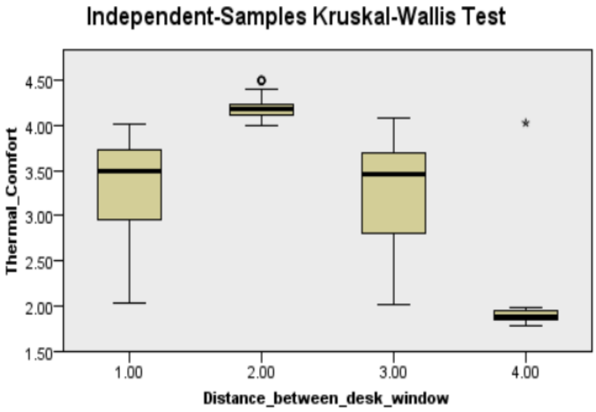
H5ii$_0$ = There is no significant difference between the respondents’ **distance between the window** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

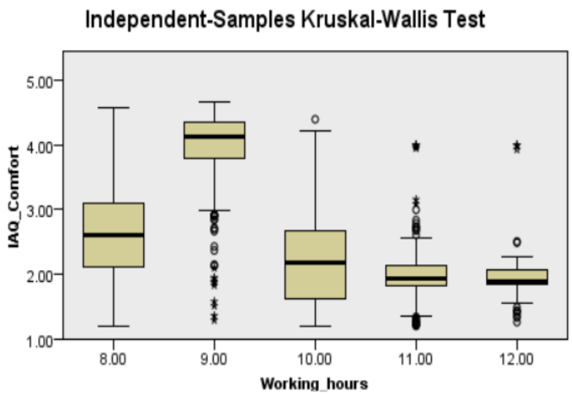
H5ii$_1$ = There is a significant difference between the respondents' **distance between the window** and the Thermal comfort/Visual comfort/ IAQ satisfaction.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Thermal_Comfort is the same across categories of Working_hours.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The distribution of Visual_Comfort is the same across categories of Working_hours.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of IAQ_Comfort is the same across categories of Working_hours.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.





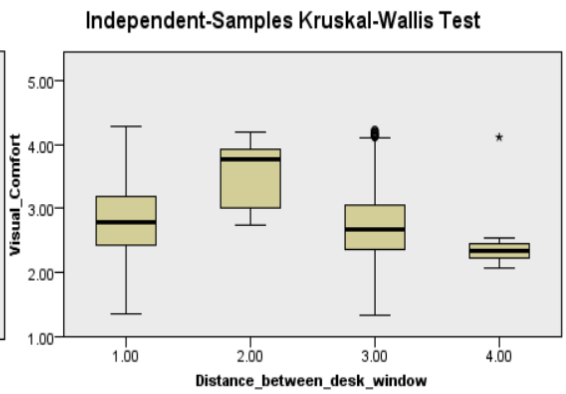


Figure 4.39: Independent samples Kruskal-Wallis test results illustration of distance between the window and test satisfaction variables in factory buildings

The p-value is less than 0.05 for all thermal comfort, visual comfort and IAQ satisfaction, and H0 is rejected for those variables (Figure 4.39) This means that the distance from windows to the workstation in factory spaces can significantly impact the response to IEQ comfort.

The results were further analysed to identify the comfort zones of the employees in the factory. Although there is a statistically significant difference between the distance of the window and the work desk, the results reflected that the comfort zones are very different from the general office spaces (Figure 4.40).

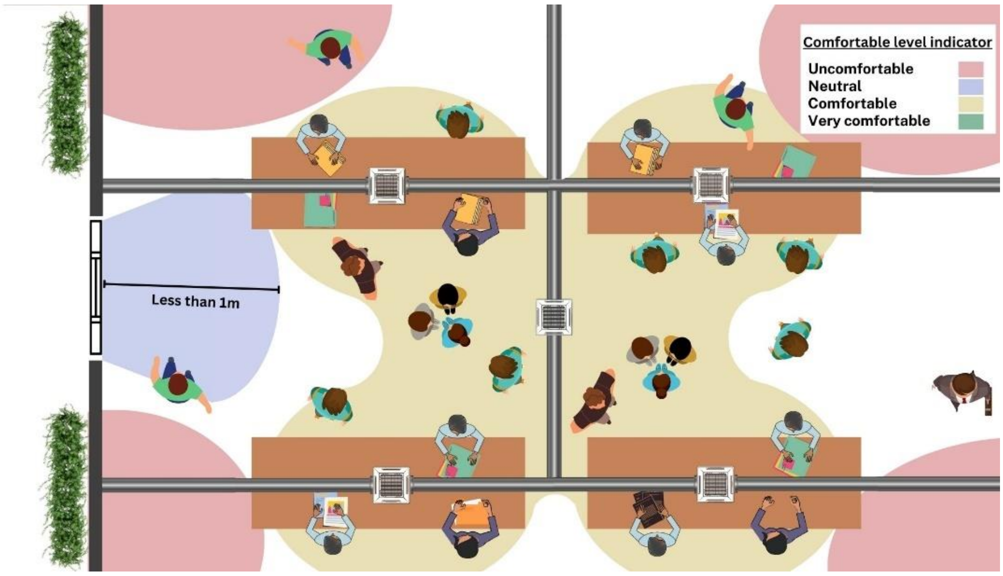


Figure 4.40: The comfortable levels of the employees with respect to the distance from the window- Factories

According to the illustration, the people in the corners of the office space are comparatively more uncomfortable or not satisfied with their IEQ than the other areas. Employees in the middle of the space are satisfied with their IEQ with respect to the different areas of the room.

4.14 Developing predictive model for Thermal comfort – Model 01

Model 1 utilises a dataset containing 1,091 rows and 28 columns. The dataset consists of various features such as 'EmpID', 'Building Type', 'Building', 'Building Location', 'Employee Home Town', and several other attributes related to thermal satisfaction. The target variable of interest is 'Thermal Satisfaction'.

Prior to modelling, specific columns that are not relevant to the analysis are dropped from the dataset. These include 'Age', 'Working Hours', 'Is there a blind wall', 'Distance between this work desk and the nearest window?', 'Smart control of lighting system', 'PM 2.5 level', 'PM 10 level', 'CO_2 PPM', 'Area served by lighting', 'Lux level', 'Visual Satisfaction', 'Indoor Air Quality', and 'Overall Satisfaction'.

A duplicate of the original dataset, denoted as 'data1_copy', is created for further analysis. The duplicate dataset has dimensions of 1,091 rows and 15 columns, excluding the dropped columns.

Further exploration of the dataset reveals information about the remaining columns. The dataset contains information related to 'EmpID', 'Building Type', 'Building', 'Building Location', 'Employee Home Town', 'Gross Floor Area', 'Wall Insulation U value', 'Roof Insulation U value', 'Thickness of the Wall Insulation', 'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area', 'Share of the area served by AC(%)', 'Smart control of HVAC', and 'Thermal Satisfaction'.

The dataset's information is then displayed using the `data1.info()` function, providing details about the column names, non-null counts, and data types. This information helps understand the dataset's structure and prepare it for further analysis.

Table 4.21 summarises the columns in the dataset used for Model 1. It includes information such as the column name, the count of non-null values, and the data type for each column. The dataset contains a mix of integer, float, and object data types, representing different variables related to the thermal satisfaction analysis. The 'Building Location' column has a count of 0 non-null values, indicating that it may be an empty or missing column in the dataset.

Table 4.21: Features and data types used in Model 1

Column	**Dtype**
EmpID	int64
Building Type	object
Building	object
Building Location	float64
Employee Home Town	object
Gross Floor Area	int64
Wall Insulation U value	float64
Roof Insulation U value	float64
The thickness of the Wall Insulation	int64
Window to Wall Ratio (WWR)	int64
Glazing U value	float64
Total Window Area	float64
Share of the area served by AC(%)	float64
Smart control of HVAC	int64
Thermal Satisfaction	float6

Following the previous steps, the dataset is further analyzed to explore the correlation between the remaining variables and the 'Thermal Satisfaction' column (Table 4.22). The correlation coefficients are computed using the `df4.corr()['Thermal Satisfaction']` code.

Table 4.22: Correlation values between variables and thermal comfort

Variable	**Correlation**
Gross Floor Area	0.350823
Wall Insulation U value	-0.021777
Roof Insulation U value	-0.519799
The thickness of the Wall Insulation	0.486843
Window to Wall Ratio (WWR)	0.348407
Glazing U value	-0.001262
Total Window Area	-0.322638
Share of the area served by AC(%)	0.654350
Smart control of HVAC	0.313969

Building Type_Factory	-0.061776
Building Type_Office	0.061776
Employee Home Town_dry	0.024022
Employee Home Town_intermediate	-0.012228
Employee Home Town_wet	-0.002363

Some key findings from the correlation analysis:

(a) 'Gross Floor Area' positively correlates 0.350823 with 'Thermal Satisfaction', indicating that larger floor areas may contribute to higher thermal satisfaction.

(b) 'Wall Insulation U value' and 'Roof Insulation U value' have negative correlations of -0.021777 and -0.519799, respectively, implying that better insulation values are associated with higher thermal satisfaction.

(c) 'Thickness of the Wall Insulation' and 'Window to Wall Ratio (WWR)' exhibit positive correlations of 0.486843 and 0.348407, respectively, suggesting that thicker wall insulation and higher window-to-wall ratios may positively impact thermal satisfaction.

(d) 'Glazing U value', 'Total Window Area', 'Share of the area served by AC(%)', and 'Smart control of HVAC' show relatively weak correlations with 'Thermal Satisfaction'.

Additionally, the correlation analysis indicates a weak positive correlation between 'Thermal Satisfaction' and the 'Building Type' variable, represented by the 'Building Type_Factory' and 'Building Type_Office' columns. Similarly, the 'Employee Home Town' variable, categorized as 'Employee Home Town_dry', 'Employee Home Town_intermediate', and 'Employee Home Town_wet', has a minor impact on 'Thermal Satisfaction'.

These correlations provide insights into the relationships between the variables and the target variable (Figure 4.41), helping identify potential predictors of thermal satisfaction. Further analysis and modelling can be performed to investigate these relationships further.

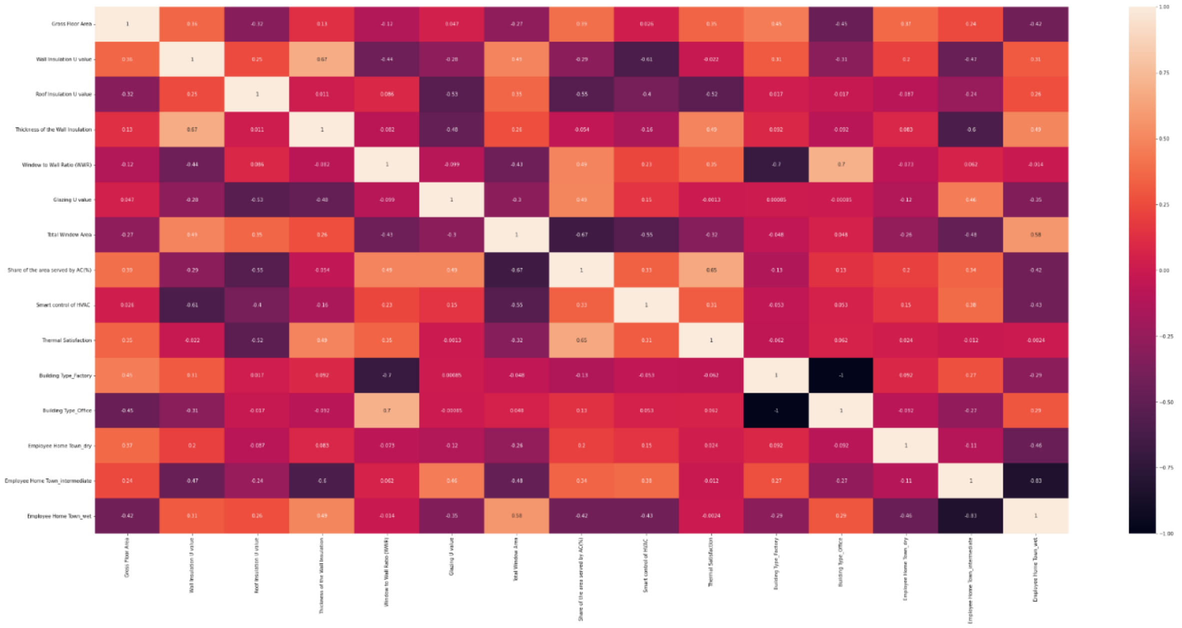


Figure 4.41: Correlation plot of Model 1
Generated by: Python

After preparing the dataset by separating the independent variables (`X`) and the dependent variable (`y`), the dataset is split into training and testing sets using the `train_test_split()` function from sci-kit-learn. The testing set size is specified as 25% of the total dataset. The shapes of `X_train`, `X_test`, `y_train`, and `y_test` are displayed, indicating the number of samples and features in each set.

A function called `model_acc()` is defined to assess the performance of various regression models. This function takes a model as an input, fits the model on the training data, and evaluates its accuracy on the testing data using the `score()` function. The accuracy score is then printed.

Several regression models are evaluated in Model 1:

1. Support Vector Regression (SVR): The SVR model with a radial basis function (RBF) kernel is created (`SVR(kernel='rbf')`) and passed to the `model_acc()` function.

2. Lasso Regression: The Lasso regression model is created (`Lasso()`) and evaluated using `model_acc()`.

3. Decision Tree Regressor: The Decision Tree regressor model is created (`DecisionTreeRegressor()`) and assessed using `model_acc()`.

4. Random Forest Regressor: The Random Forest regressor model is created (`RandomForestRegressor()`) and evaluated using `model_acc()`.

The accuracy values for each model are printed, indicating how well each model performs in predicting the 'Thermal Satisfaction' variable.

Next, the code calculates the root mean squared error (RMSE) for each model using the `mean_squared_error()` function from scikit-learn. The RMSE measures the average deviation between the predicted and actual values of the target variable. The RMSE values for the Random Forest, Lasso Regression, Decision Tree, and SVR models are printed.

The mean absolute error (MAE) is also calculated for each model using the `mean_absolute_error()` function. The MAE represents the average absolute difference between the predicted and actual values. The MAE values for the Random Forest, Lasso Regression, Decision Tree, and SVR models are displayed.

Lastly, cross-validation is performed to evaluate the performance of the models further. The models, including Lasso Regression, Decision Tree Regressor, SVR, and Random Forest Regressor, are defined. Cross-validation scores are computed using the `cross_val_score()` function with five-fold cross-validation. The mean MAE scores for each model are printed, representing the average MAE across all folds.

The results indicate the accuracy, RMSE, MAE, and cross-validated MAE for each model, providing insights into their performance in predicting the 'Thermal Satisfaction' variable.

The results from Model 1 provide valuable insights into the performance of different regression models in predicting the 'Thermal Satisfaction' variable. The findings based on the accuracy, RMSE, MAE, and cross-validated MAE values for each model:

1. Support Vector Regression (SVR):

- Accuracy: 0.119

- RMSE: 0.789

- MAE: 0.599

- Cross-validated MAE: 0.617

The SVR model performs poorly with low accuracy and relatively high RMSE and MAE values. It may not be the most suitable model for accurately predicting 'Thermal Satisfaction' based on the given independent variables.

2. Lasso Regression:

- Accuracy: 0.664

- RMSE: 0.488

- MAE: 0.396

- Cross-validated MAE: 0.401

The Lasso Regression model shows significantly improved performance compared to SVR, with higher accuracy and lower RMSE and MAE values. It demonstrates a better fit to the data and offers relatively accurate 'Thermal Satisfaction' predictions.

3. Decision Tree Regressor:

- Accuracy: 0.835

- RMSE: 0.342

- MAE: 0.231

- Cross-validated MAE: 0.238

The Decision Tree Regressor performs even better than the Lasso Regression model, with higher accuracy and lower RMSE and MAE values. It exhibits a strong ability to capture the relationships between the independent variables and 'Thermal Satisfaction', making it a promising model for prediction.

4. Random Forest Regressor:

- Accuracy: 0.834

- RMSE: 0.342

- MAE: 0.230

- Cross-validated MAE: 0.238

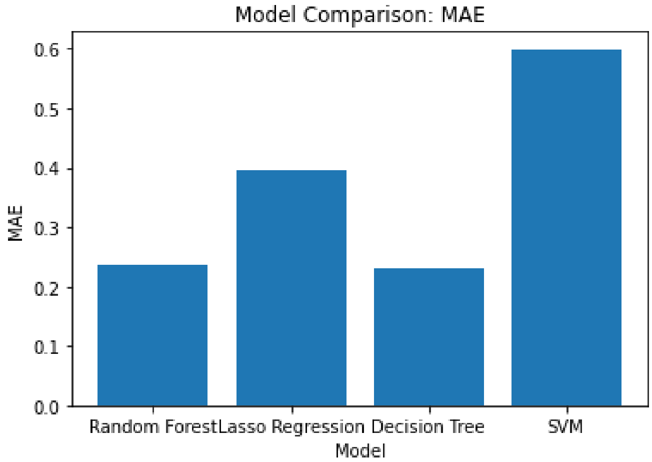


Figure 4.42: Model 1 comparison MAE values

Generated by: Python

The models were compared and summarized in Table 4.23.

Table 4.23: The model comparison for Thermal comfort

Model	R^2	RMSE	MAE	Cross-Validated MAE
SVR	0.119	0.789	0.599	0.617
Lasso Regression	0.664	0.488	0.396	0.401
Decision Tree Regressor	0.834	0.342	0.231	0.238
Random Forest Regressor	0.835	0.342	0.230	0.238

The Random Forest Regressor achieves results similar to the Decision Tree Regressor, demonstrating high accuracy and low RMSE and MAE values. Combining multiple decision trees to improve prediction accuracy and generalization benefits the ensemble approach.

Based on these results, the Decision Tree Regressor and Random Forest Regressor outperform the SVR and Lasso Regression models. They exhibit higher accuracy and lower error metrics, indicating their superior predictive capability for the 'Thermal Satisfaction' variable.

A grid search approach is employed to find the best model for this scenario. The \`GridSearchCV\` class from scikit-learn is utilised to perform an exhaustive search over specified parameter values for the Random Forest Regressor. The following parameters are considered:

a. `n_estimators`: Number of trees in the random forest (10, 50, 100)
b. `criterion`: The function to measure the quality of a split ('squared_error', 'absolute_error', 'Poisson')

The grid search uses the training data (`X_train` and `y_train`). The best model is determined based on the highest score obtained during the grid search.

After fitting the grid search object to the training data, the best model is obtained using `grid_fit.best_estimator_`. In this case, the best model is a Random Forest Regressor with the criterion set to 'poisson' and 50 estimators.

To assess the accuracy of the best model, its score on the testing data (`X_test` and `y_test`) is calculated using `best_model.score(X_test, y_test)`. The accuracy score for the best model is 0.836.

Based on the grid search results, the Random Forest Regressor with the specified parameters is the best model for this scenario. It achieves a high accuracy score, indicating its predictive solid performance for the 'Thermal Satisfaction' variable.

The Variance Inflation Factor (VIF) is calculated for each feature to assess multicollinearity among the independent variables in the dataset. The VIF measures the extent to which the variance of the estimated regression coefficient is increased due to multicollinearity.

The provided code selects a subset of the dataset (`dfdata`) containing the relevant independent variables. The variables included are 'Gross Floor Area', 'Wall Insulation U value', 'Roof Insulation U value', 'Thickness of the Wall Insulation', 'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area', 'Share of the area served by AC(%)', and 'Smart control of HVAC'.

To calculate the VIF, a constant term is added to the dataset using `add_constant()` from the `statsmodels—tools` module. Then, the VIF values are computed for each variable using the `variance_inflation_factor()` function from the `statsmodels.stats.outliers_influence` module. The VIF values are stored in a Pandas Series, where the index corresponds to the feature names. The VIF values for each feature are mentioned in Table

Table 4.24: VIF values for Model 1

Feature	**VIF**
Gross Floor Area	5.623841
Wall Insulation U value	6.811833
Roof Insulation U value	3.976865
The thickness of the Wall Insulation	10.678259
Window to Wall Ratio (WWR)	3.020962
Glazing U value	6.451798
Total Window Area	2.718279
Share of the area served by AC(%)	7.413206
Smart control of HVAC	3.273872

These VIF values indicate the level of multicollinearity present in the dataset. Generally, VIF values above 10 are considered problematic, suggesting a high degree of correlation among the independent variables. In this case, the 'Wall Insulation U value' variable has relatively high VIF values, indicating the presence of multicollinearity.

A code snippet is provided to utilise the `plot_tree` function from scikit-learn's `tree` module to visualize a decision tree from the best model. The tree being visualized is the 21st estimator in the ensemble (`best_model.estimators_[20]`). To plot the decision tree, a more significant figure size of 100x100 is set using `plt. figure(figsize=(100,100))`. Then, the `plot_tree` function is called with the specified parameters:

(a) `best_model.estimators_[20]`: The decision tree estimator from the best model.
(b) `feature_names`: The names of the features in the dataset (`df4.columns`).
(c) `class_names`: The class names for the target variable ('Thermal Satisfaction').
(d) `filled=True`: This parameter fills the tree nodes with colour based on the majority class.

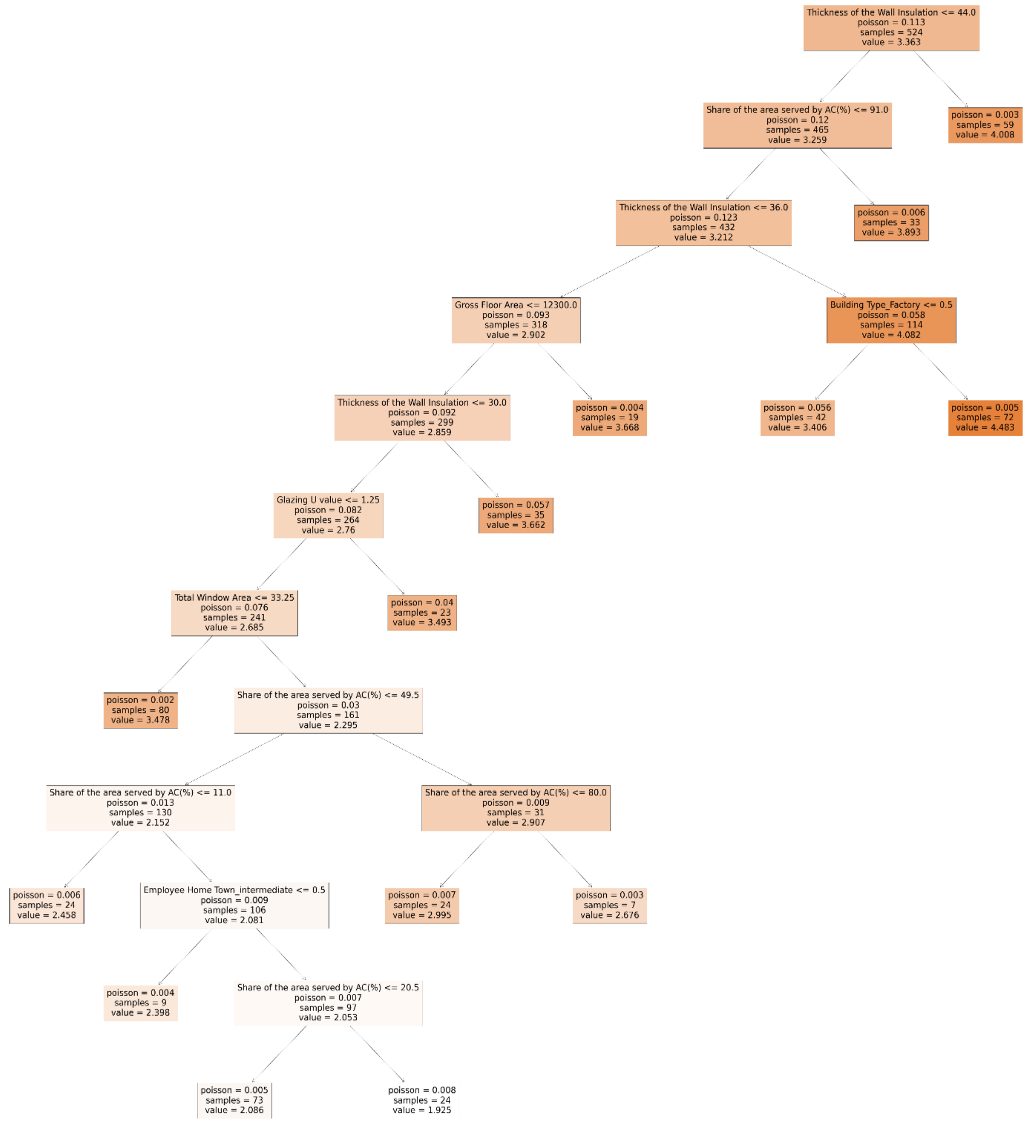


Figure 4.43: Decision tree of the Model 1
Generated by: Python

The resulting decision tree visualization (Figure 4.43) provides insights into the structure and decision-making process of the model, showcasing the splits and conditions used to predict the 'Thermal Satisfaction' class.

The splits in the tree represent the relationships between the building's structural factors and thermal satisfaction. Each split defines a condition based on a specific feature, and the tree navigates through the conditions to make predictions. The Poisson values indicate the expected thermal satisfaction values within each box, and the number of samples and values provide additional information about the data distribution.

The Poisson values in the Random Forest regression tree context represent the predicted thermal satisfaction values within each box or leaf node of the tree. The Poisson distribution is a discrete probability distribution that models the number of events occurring in a fixed interval of time or space. In this case, it is used to model the thermal satisfaction values.

When interpreting the Poisson values within a box or leaf node, they indicate the expected or average thermal satisfaction value for the samples within that node. The Poisson values can be considered the predicted mean or central tendency of the thermal satisfaction in that particular data subgroup.

It's important to note that the Poisson values are specific to each box or leaf node and are generated based on the training data and the regression algorithm used. They provide insight into the expected thermal satisfaction level within that particular subgroup, aiding in understanding the relationship between the selected features and the target variable.

The root node, the first split in the tree, is based on the "Thickness of the wall insulation" feature. If the thickness of the wall insulation is less than or equal to 44.0, the tree proceeds to the left side of the split, otherwise to the right side.

The fact that the tree ends with the "Share of the area served by AC" feature and that it is one of the highest-ranked boxes indicates the importance of this feature in predicting thermal satisfaction.

When the tree reaches the end node based on the " Share of the area served by AC " feature, it means that this feature alone provides sufficient information to predict thermal satisfaction. The tree has determined that the share of the area served by air conditioning is a significant factor in determining the level of thermal satisfaction experienced by individuals.

The high ranking of this box suggests that the " Share of the area served by AC " feature strongly influences thermal satisfaction outcomes. It implies that the proportion or extent of the area covered by air conditioning in a building notably impacts people's thermal comfort and satisfaction levels. A higher share of the area served by AC might indicate better temperature control and enhanced thermal comfort, leading to higher satisfaction levels.

Therefore, in the context of the provided random forest regression tree, the " Share of the area served by AC " feature plays a crucial role in predicting and understanding the variation in thermal satisfaction, and it is one of the key factors to consider when assessing the impact of building structural factors on thermal comfort.

A code performed cross-validation for the Random Forest model using 5-fold validation and calculated the Mean Absolute Error (MAE) as the evaluation metric.

The output indicates that the average MAE across the 5 cross-validation folds is 0.237. The lower the MAE value, the better the model's predictive performance, as it represents the average absolute difference between the predicted values and the actual values of the target variable (thermal satisfaction) across the test data.

The boxplot visualizes (Figure 4.44) the distribution of MAE values obtained from cross-validation. Each box represents the interquartile range (IQR), with the median indicated by the horizontal line within the box. The whiskers extend to the minimum and maximum values, excluding outliers. The boxplot helps assess the consistency and spread of the MAE values across the folds. A smaller spread and lower median indicate better model performance.

The relatively low MAE value and the compact boxplot suggest that the Random Forest model has achieved good performance and consistency in predicting thermal satisfaction. The model's average absolute error in predicting thermal satisfaction is approximately 0.237, indicating a relatively small deviation from the actual values.

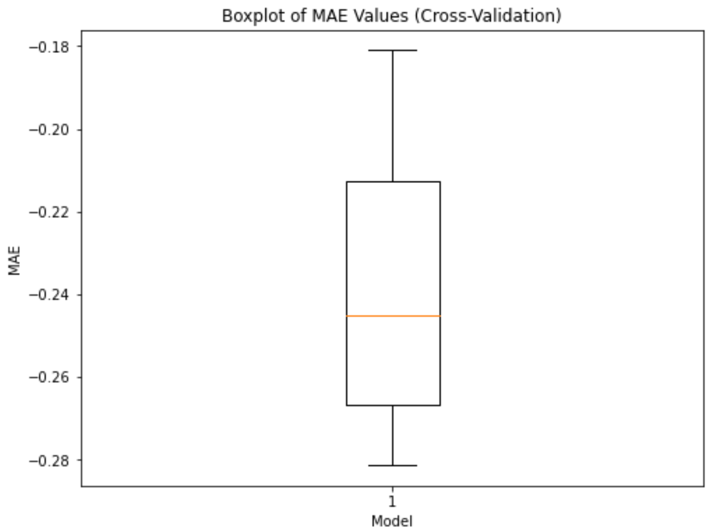


Figure 4.44: Box plot MAE values (cross-validation) for Model 1

Generated by: Python

A scatter plot was developed comparing the actual values of the target variable (thermal satisfaction) with the predicted values from the Random Forest model.

In the scatter plot, each point represents a data instance. The x-axis represents the actual values of the target variable, while the y-axis represents the predicted values by the Random Forest model. The plot visualizes how well the model's predictions align with the actual values.

The diagonal red dashed line in the plot represents the line of perfect prediction, where the predicted values perfectly match the actual values. Ideally, the points on the scatter plot should be close to this line, indicating a strong correlation between the predicted and actual values.

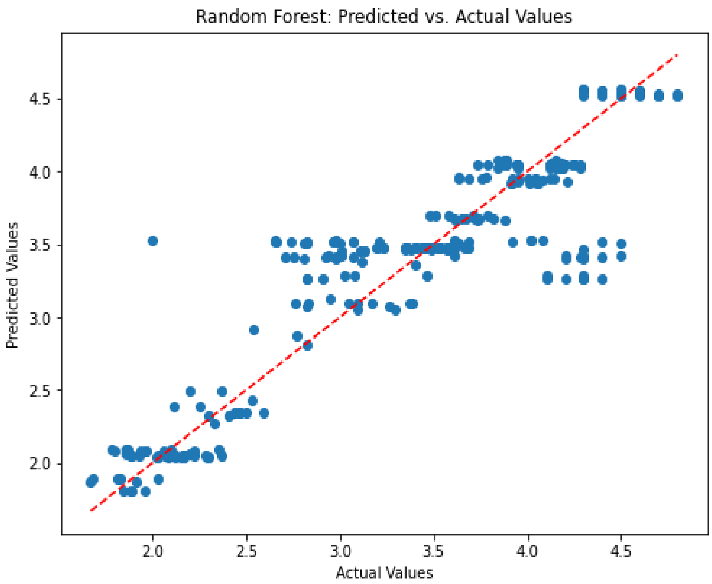


Figure 4.45: Predicted Vs Actual value of RF for Model 1

Generated by: Python

By observing the scatter plot, assess the model's predictive performance. The model's predictions are accurate and reliable if the points are closely scattered around the red dashed line. On the other hand, if the facts are spread out with a more significant deviation from the red dashed line, it suggests that the model's predictions have more substantial errors or inconsistencies (Figure 4.45).

According to the figure, the points are closely scattered around the red dashed line, which indicates accurate and reliable predictions.

The code was developed to train a Random Forest model on the training data and calculates the feature importance using the trained model's `feature_importances_` attribute. The feature importance represents the relative significance of each feature in the model's decision-making process.

The resulting feature importance values are then stored in a DataFrame and sorted in descending order. Finally, a bar plot is created to visualize the feature importance, with the y-axis representing the features and the x-axis representing their corresponding importance scores.

Interpreting the feature importance plot allows (Figure 4.46) us to identify which features significantly impact the Random Forest model's predictions for thermal satisfaction. Features with higher importance values contribute more to the model's decision-making process, indicating their more substantial influence on the target variable.

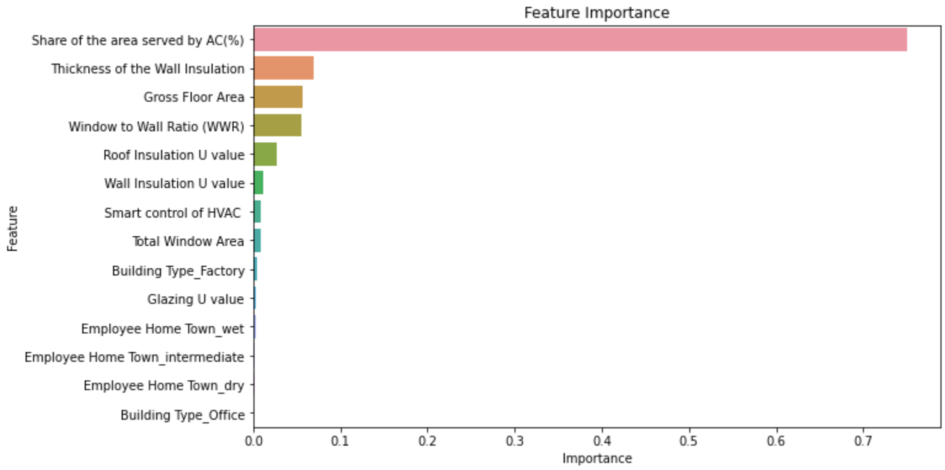


Figure 4.46: Feature importance of the thermal comfort model variables.

Generated by: Python

The figure highlights the feature index according to the importance of thermal comfort. The share of the area served by AC is the most important factor. The second most important factor is the thickness of the wall insulation. Another visual representation was created to visualize the importance of other variables when we omit AC's share of the area served (Figure 4.47).

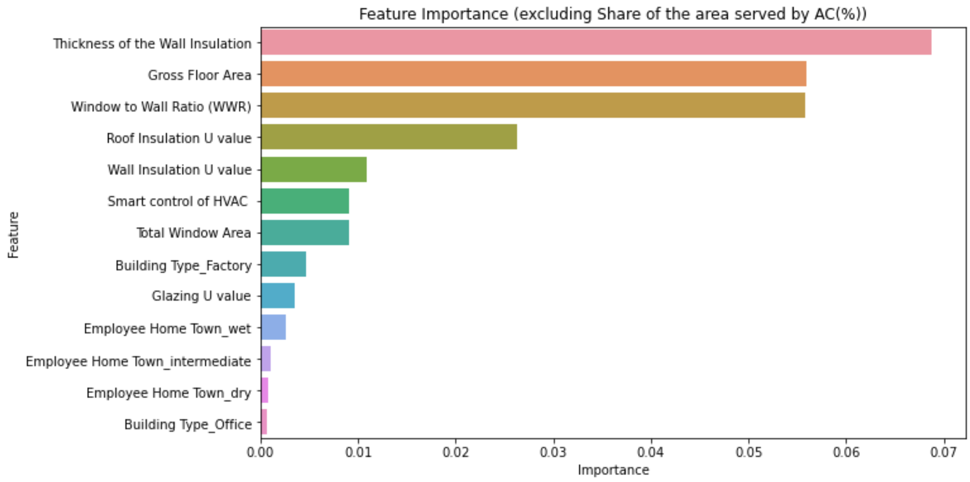


Figure 4.47: Feature importance of the thermal comfort model variables (Excluding area shared by AC)

The least important factor is the building type, whether the building is a general office or a factory and then the employee home town.

4.15 Developing predictive model for Visual Comfort – Model 02

The model infrastructure, the methodology and the data were all the same as what was used in Model 01. Prior to modelling, specific columns that are not relevant to the analysis are dropped from the dataset. This table provides a summary of the columns in the dataset used for the Model 2 is outlined in the Table 4.25.

Table 4.25: Features and data types used in Model 2

Column	**Dtype**
EmpID	int64
Building Type	object
Building	object
Building Location	float64
Employee Home Town	object
Gross Floor Area	int64
Window to Wall Ratio (WWR)	int64
Total Window Area	float64
Smart control of the lighting system	int64
PM 10 level	int64
Area served by lighting	int64
Lux level	int64
Visual Satisfaction	float64

Following the previous steps, the dataset is further analyzed to explore the correlation between the remaining variables and the 'Visual Satisfaction' column (Table 4.26).

Table 4.26: Correlation values between variables and visual comfort

Variable	**Correlation**
Gross Floor Area	-0.179
Window to Wall Ratio (WWR)	0.469
Total Window Area	0.224
Smart control of the lighting system	0.241
PM 10 level	-0.580
Area served by lighting	0.303
Lux level	0.021
Visual Satisfaction	1.000
Building Type_Factory	-0.466
Building Type_Office	0.466
Employee Home Town_dry	-0.354
Employee Home Town_intermediate	-0.205
Employee Home Town_wet	0.379

The analysis of the correlation between the selected factors and Visual Satisfaction revealed the following key findings:

1. Building Factors:

- Window Wall Ratio (WWR) and Total Window Area showed positive correlations with Visual Satisfaction, indicating that higher WWR and larger window areas are associated with increased satisfaction.

- Building Type strongly correlated with Visual Satisfaction, with offices having higher satisfaction and factories having lower satisfaction.

2. Lighting Factors:

- Smart control of lighting systems demonstrated a positive correlation with Visual Satisfaction, suggesting that better control of lighting systems leads to higher satisfaction.

- Area served by lighting positively correlated with Visual Satisfaction, indicating that a more extensive area served by lighting contributes to higher satisfaction.

3. Environmental Factors:

- Gross Floor Area exhibited a weak negative correlation with Visual Satisfaction, implying that larger floor areas may lead to slightly lower satisfaction.

- PM 10 level, a measure of particulate matter pollution, had a strong negative correlation with Visual Satisfaction, highlighting the negative impact of higher pollution levels on satisfaction.

- LUX level, which measures illumination intensity, showed a very weak positive correlation with Visual Satisfaction, suggesting a minimal relationship.

4. Employee Home Town:

- Employee Home Town had correlations with Visual Satisfaction, with employees from wet home towns showing higher satisfaction and those from dry and intermediate home towns having lower satisfaction.

These findings emphasize the importance of window design, lighting control, building type, environmental quality, and employee background in influencing Visual Satisfaction. The correlation plot for Model 2 is shown in Figure 4.48.

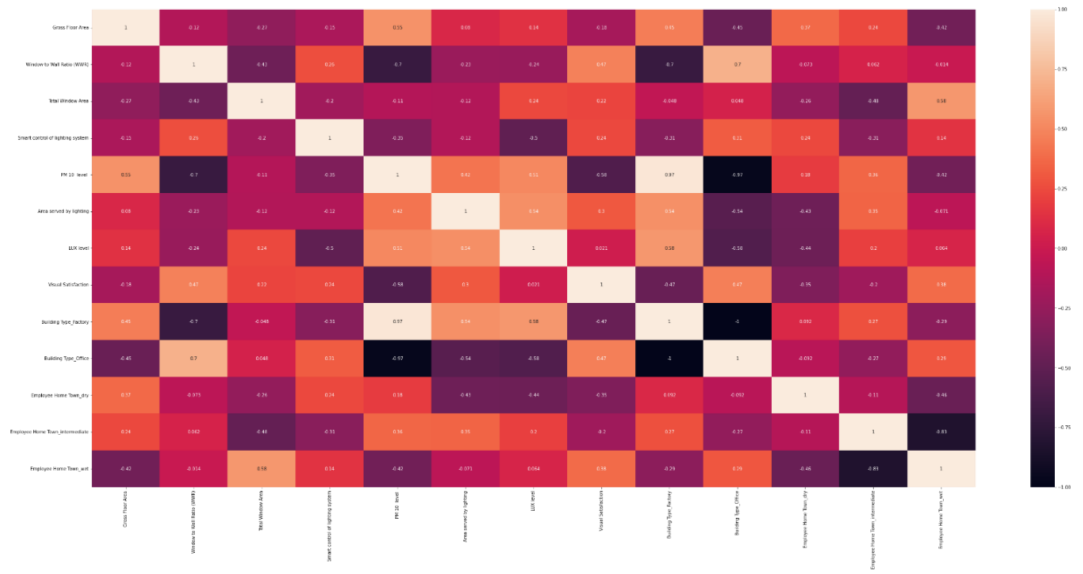


Figure 4.48: Correlation plot of Model 2

Generated by: Python

The models were evaluated using accuracy, RMSE, MAE (Figure 4.49) and cross-validation metrics based on the results.

1. Support Vector Regression (SVR):

- Accuracy: 0.342

- RMSE: 0.701

- MAE: 0.626

- Cross-validated MAE: 0.659

The SVR model shows relatively low accuracy and high RMSE, indicating that it may not be the best model for accurately predicting the target variable. The cross-validated MAE also suggests that the model's performance may vary across different cross-validation folds.

2. Lasso Regression:

- Accuracy: 0.871

- RMSE: 0.311

- MAE:0.244

- Cross-validated MAE: 0.251

The Lasso Regression model demonstrates higher accuracy and lower RMSE than SVR, indicating better performance. The cross-validated MAE suggests that the model's average absolute error is relatively low, meaning good prediction accuracy.

3. Decision Tree Regressor:

- Accuracy: 0.952

- RMSE: 0.189

- MAE: 0.145

- Cross-validated MAE: 0.150

The Decision Tree Regressor shows high accuracy and low RMSE, indicating excellent performance. The cross-validated MAE suggests that the model performs well across different cross-validation folds.

4. Random Forest Regressor:

- Accuracy: 0.952

- RMSE: 0.188

- MAE: 0.144

- Cross-validated MAE: 0.150

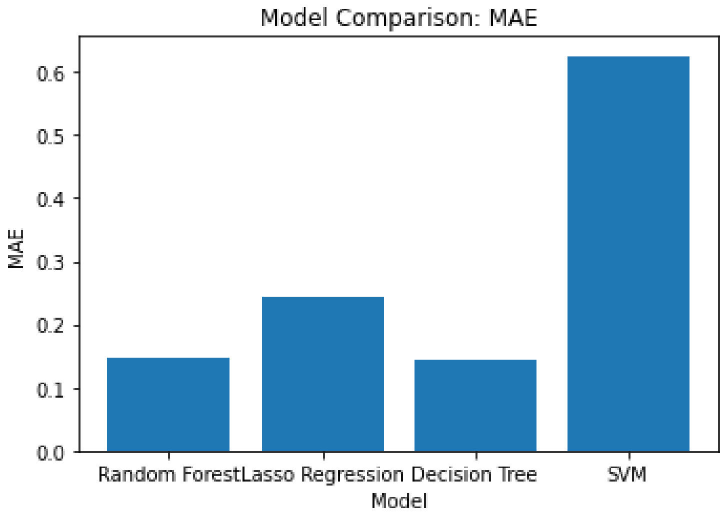


Figure 4.49: Model 2 comparison MAE values

Generated by: Python

The Random Forest Regressor performs similarly to the Decision Tree Regressor, with high accuracy and low RMSE. The cross-validated MAE suggests consistent performance across different folds.

Overall, the Decision Tree Regressor and Random Forest Regressor models perform exceptionally well, outperforming the SVR and Lasso Regression models. They show high accuracy and low RMSE, indicating their suitability for predicting the target variable. The cross-validated MAE scores suggest that these models consistently deliver accurate predictions across different cross-validation folds. The model comparison is summarized in Table 4.27.

Table 4.27: The model comparison for Visual comfort

Model	**R^2**	**RMSE**	**MAE**	**Cross-Validated MAE**
Support Vector Regression (SVR)	0.342	0.701	0.626	0.659
Lasso Regression	0.871	0.311	0.244	0.251
Decision Tree Regressor	0.952	0.189	0.145	0.150
Random Forest Regressor	0.952	0.188	0.144	0.150

After the hyperparameter, the accuracy score for the best model is 0.953.

The Variance Inflation Factor (VIF) is calculated for each feature to assess multicollinearity among the independent variables in the dataset (Table 4.28).

Table 4.28: VIF values for Model 2

Feature	**VIF**
Gross Floor Area	2.572
Window to Wall Ratio (WWR)	10.089
Total Window Area	6.300
Smart control of the lighting system	1.462
PM 10 level	10.316
Area served by lighting	1.940
LUX level	5.868

Figure 4.50 : Decision tree of the Model 2

The root node, the first split in the tree, is based on the "Lux level" feature if the thickness of the wall insulation is less than or equal to 950 lx, the tree proceeds to the left side of the split, otherwise to the right side.

The high ranking of this box suggests that the "window-to-wall ratio" feature strongly influences visual satisfaction outcomes. Figure 4.51 represents the model accuracy visually.

The output indicates that the average MAE across the 5 cross-validation folds is 0.151.

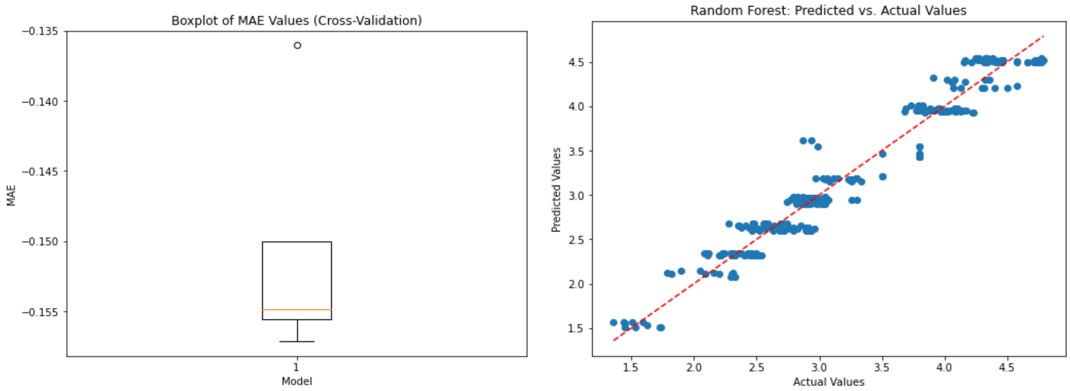


Figure 4.51: Predicted Vs Actual value of RF for Model 2 (Left), Box plot MAE values (cross-validation) for Model 2 (Right)

The model's average absolute error in predicting thermal satisfaction is approximately 0.151, indicating a relatively small deviation from the actual values.

The scatter plot points are highly scattered around the red line, which means the accuracy and reliability of the model are in a good position.

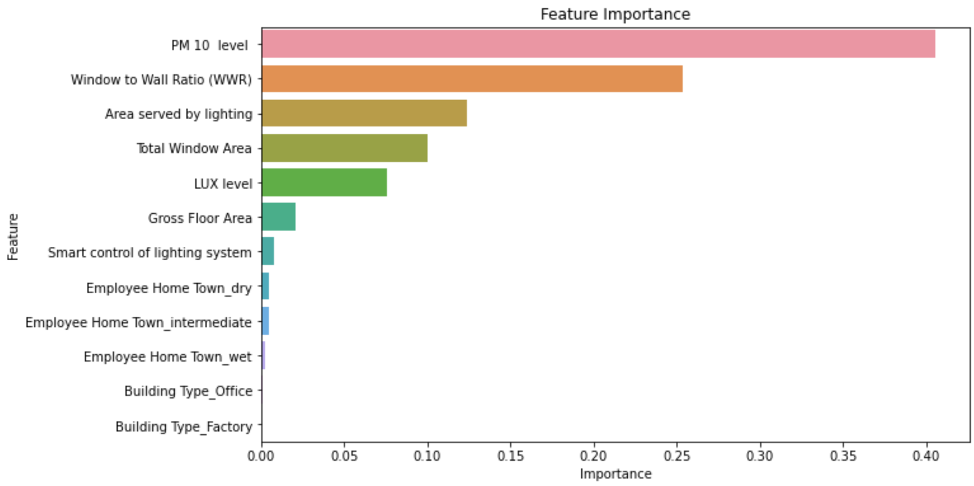


Figure 4.52: Feature importance of the visual comfort model variables.

The figure highlights the feature index according to the importance of visual comfort. The PM 10 level is the most important factor (Figure 4.52) The most important building structural factors are window-to-wall ratio and area served by lighting.

4.16 Developing a predictive model for Indoor Air Quality – Model 03

Prior to modelling, specific columns that are not relevant to the analysis are dropped from the dataset. This table provides a summary of the columns in the dataset used for Model 3 is outlined in the Table 4.29.

Table 4.29: Features and data types used in Model 3

Column	Dtype
EmpID	int64
Building Type	object
Building	object
Building Location	float64
Employee Home Town	object
Gross Floor Area	int64
Window to Wall Ratio (WWR)	int64
Total Window Area	float64
PM 2.5 level	int64
PM 10 level	int64
CO2 PPM	int64
Indoor Air Quality	float64

Following the previous steps, the dataset is further analyzed to explore the correlation between the remaining variables and the 'IAQ Satisfaction' column (Table 4.30).

Table 4.30: Correlation values between variables and IAQ comfort

Variable	Correlation
Gross Floor Area	-0.312798
Window to Wall Ratio (WWR)	0.621136
Total Window Area	0.144883
PM 2.5 level	-0.822760
PM 10 level	-0.848444
CO2 PPM	-0.842865
Indoor Air Quality	1.000000
Building Type_Factory	-0.798600
Building Type_Office	0.798600
Employee Home Town_dry	-0.197425
Employee Home Town_intermediate	-0.303220
Employee Home Town_wet	0.380628

Key Findings:

Window Wall Ratio (WWR), Building Type_Office, and Employee Home Town_wet have a relatively higher positive correlation with Indoor Air Quality. This suggests that these factors may have a positive influence on the quality of indoor air.

PM $_{2.5}$ level, PM $_{10}$ level, CO_2 PPM, and Building Type_Factory have a strong negative correlation with Indoor Air Quality. This indicates that higher levels of these factors are associated with lower indoor air quality.

Gross Floor Area, Total Window Area, Employee Home Town_dry, and Employee Home Town_intermediate have a weaker correlation with Indoor Air Quality but still show some influence on the air quality.

The strongest correlation is observed between Indoor Air Quality and itself, which is expected. These findings suggest that factors such as window-to-wall ratio, building type, and external pollutant levels (PM $_{2.5}$, PM $_{10}$, CO_2) significantly determine indoor air quality. The correlation plot for Model 3 is shown in Figure 4.53.

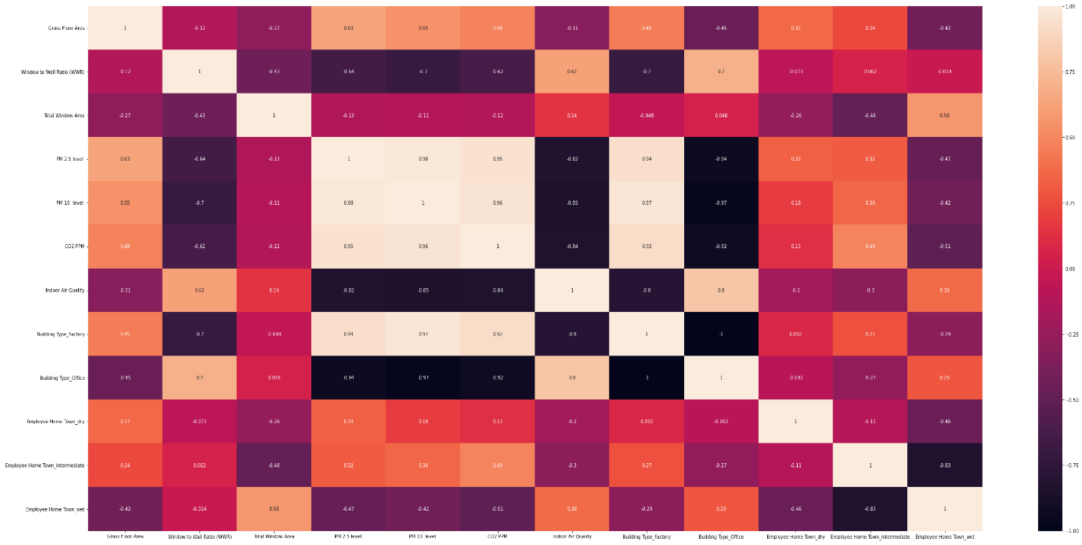


Figure 4.53: The correlation plot for Model 3

The models were evaluated using accuracy, RMSE, MAE and cross-validation metrics based on the results. Based on the performance metrics,

1. Support Vector Regression (SVR):

- Accuracy: 0.743

- RMSE: 0.489

- MAE: 0.265

- Cross-validated MAE: 0.263

The SVR model performs reasonably well with decent accuracy and relatively low RMSE and MAE values. It shows a moderate ability to predict 'Indoor Air Quality' based on the given independent variables.

2. Lasso Regression:

- Accuracy: 0.761

- RMSE: 0.472

- MAE: 0.307

- Cross-validated MAE: 0.294

The Lasso Regression model performs slightly better than SVR, with higher accuracy and lower RMSE and MAE values. It demonstrates a better fit to the data and offers relatively accurate 'Indoor Air Quality' predictions.

3. Decision Tree Regressor:

- Accuracy: 0.825

- RMSE: 0.404

- MAE: 0.186

- Cross-validated MAE: 0.188

The Decision Tree Regressor performs even better than the Lasso Regression model, with higher accuracy and lower RMSE and MAE values. It exhibits a strong ability to capture the relationships between the independent variables and 'Indoor Air Quality', making it a promising model for prediction.

4. Random Forest Regressor:

- Accuracy: 0.825

- RMSE: 0.404

- MAE: 0.187

- Cross-validated MAE: 0.188

The Random Forest Regressor performs similarly to the Decision Tree model, with high accuracy and low RMSE and MAE (Figure 4.54) values. It leverages the ensemble of multiple decision trees to make more accurate predictions.

Overall, the Decision Tree and Random Forest models demonstrate the best performance among the four models, with the lowest RMSE and MAE values. Based on the given dataset, these models are recommended for predicting 'Indoor Air Quality'. The model comparison is summarized in Table 4.31

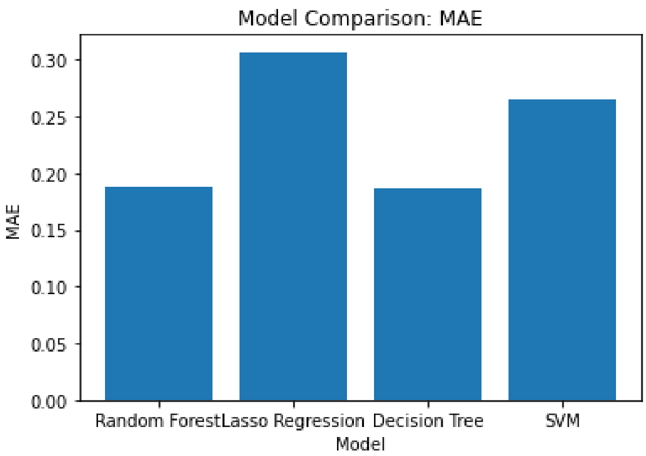


Figure 4.54: Model 3 comparison MAE values

.

Table 4.31: The model comparison for IAQ comfort

Model	R^2	RMSE	MAE	Cross-Validated MAE
Support Vector Regression	0.743	0.489	0.265	0.263
Lasso Regression	0.761	0.472	0.307	0.294
Decision Tree Regressor	0.825	0.404	0.186	0.188
Random Forest Regressor	0.825	0.404	0.187	0.188

The Variance Inflation Factor (VIF) is calculated for each feature to assess multicollinearity among the independent variables in the dataset (Table 4.32).

Table 4.32: VIF values for Model 3

Feature	VIF
Gross Floor Area	3.228
Window to Wall Ratio (WWR)	5.719
Total Window Area	2.215
$PM_{2.5}$ level	8.046
PM_{10} level	8.915
CO_2 PPM	10.487

The features "$PM_{2.5}$ level," "PM_{10} level," and "CO_2 PPM" have relatively high VIF values, indicating a higher degree of correlation with other variables in the model. This suggests that these variables may have a higher influence on the target variable and may need further examination to avoid multicollinearity issues in the analysis.

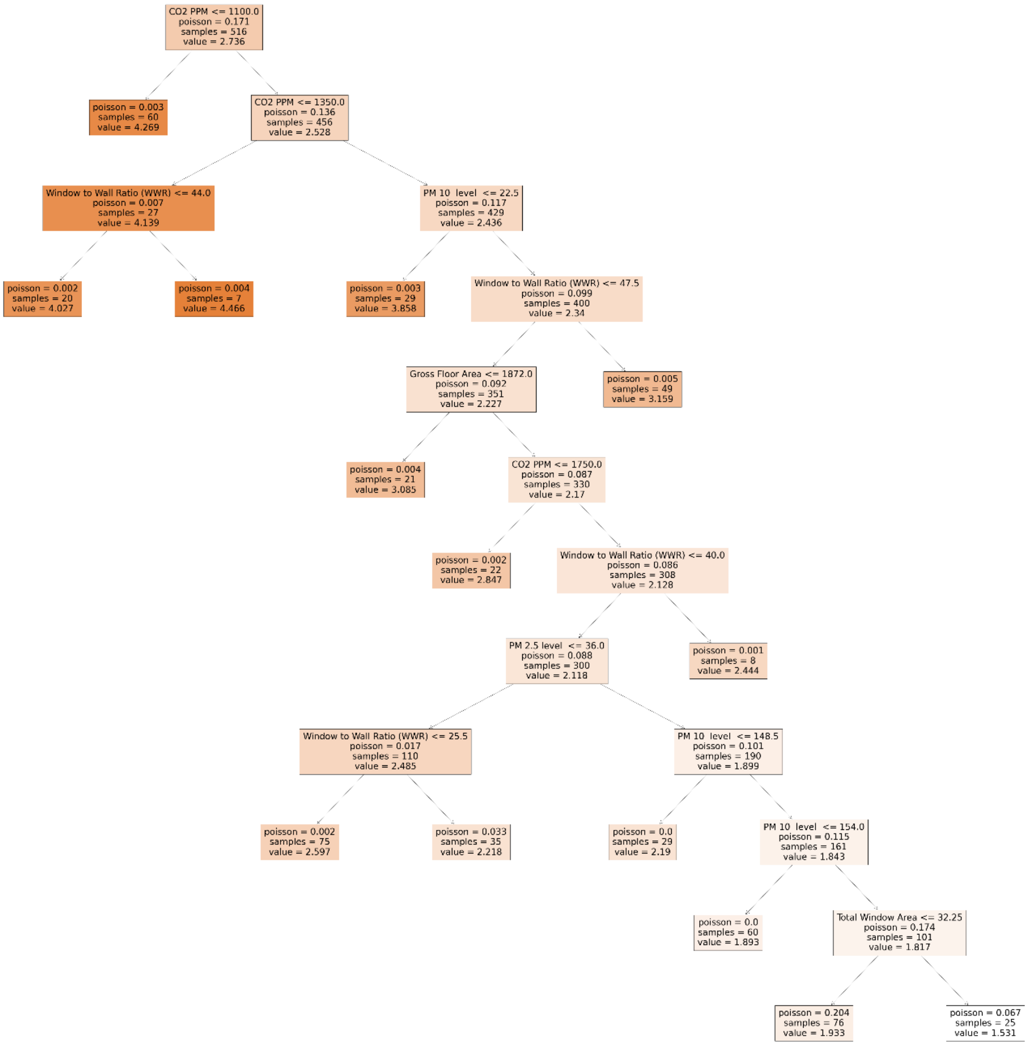


Figure 4.55: Decision tree of the Model 3

The root node, the first split in the tree, is based on the "CO_2 ppm" feature. If the window-to-wall ratio is less than or equal to 47.5, the tree proceeds to the left side of the split, otherwise to the right side (Figure 4.55).

The high ranking of this box suggests that the "window-to-wall ratio" feature strongly influences visual satisfaction outcomes.

The output indicates that the average MAE across the 5 cross-validation folds is 0.188. Figure 4.56 represents the model accuracy visually.

The model's average absolute error in predicting thermal satisfaction is approximately 0.151, indicating a relatively small deviation from the actual values.

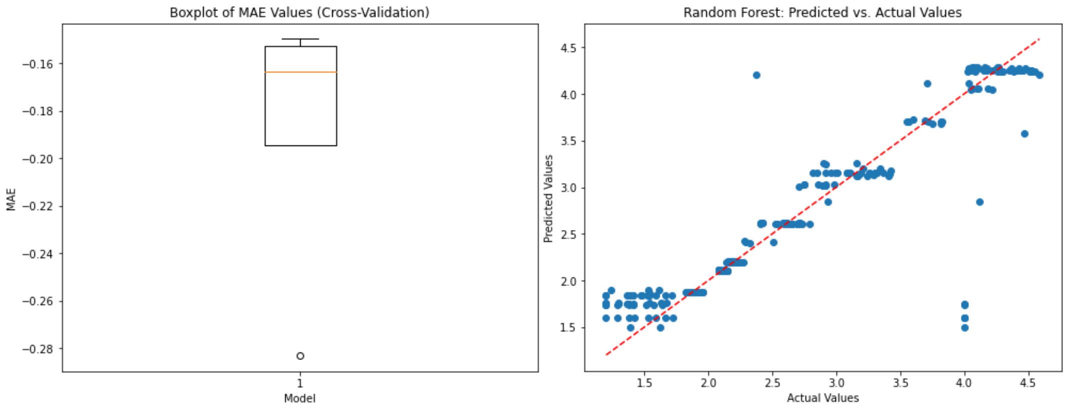


Figure 4.56: Predicted Vs Actual value of RF for Model 3 (Left), Box plot MAE values (cross-validation) for Model 3 (Right)

The scatter plot points are highly scattered around the red line, which means the accuracy and reliability of the model are in a good position. There are a few outliers are also visible in the illustration.

The Figure 4.57 highlights the feature index according to the importance of visual comfort.

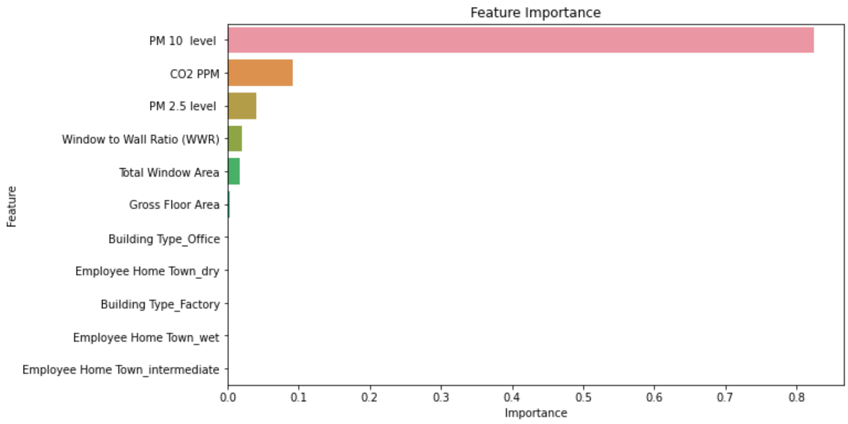


Figure 4.57: Feature importance of the IAQ comfort model variables

The PM $_{10}$ level is the most important factor. The second most important factor is the CO_2 level. The most important building structural factors are window-to-wall ratio and total window area.

4.17 Developing predictive model for Indoor Environmental Quality (IEQ) for overall performance- Model 04

This table provides a summary of the columns in the dataset used for Model 4 is outlined in the Table 4.33. All the variables considered in the Model 01, Model 02 and Model 03 were included.

Table 4.33: Features and data types used in Model 4

Column	Dtype
EmpID	int64
Building Type	object
Building	object
Building Location	float64
Employee Home Town	object
Gross Floor Area	int64
Wall Insulation U value	float64
Roof Insulation U value	float64
The thickness of the Wall Insulation	int64
Window to Wall Ratio (WWR)	int64
Glazing U value	float64
Total Window Area	float64
Share of the area served by AC(%)	float64
Smart control of HVAC	int64
Smart control of the lighting system	int64
PM 2.5 level	int64
PM 10 level	int64
CO2 PPM	int64
Area served by lighting	int64
LUX level	int64
Overall Satisfaction	float64

Following the previous steps, the dataset is further analyzed to explore the correlation between the remaining variables and the 'IEQ Satisfaction' column (Table 4.34).

Table 4.34: Correlation values between variables and IEQ comfort

Variable	Correlation
Gross Floor Area	0.046
Wall Insulation U value	-0.039
Roof Insulation U value	-0.332
The thickness of the Wall Insulation	0.436
Window to Wall Ratio (WWR)	0.458
Glazing U value	-0.044
Total Window Area	-0.029
Share of the area served by AC(%)	0.447
Smart control of HVAC	0.151
Smart control of the lighting system	0.287
PM 2.5 level	-0.411
PM 10 level	-0.445

CO2 PPM	-0.452
Area served by lighting	0.292
LUX level	-0.100
Building Type_Factory	-0.353
Building Type_Office	0.353
Employee Home Town_dry	-0.170
Employee Home Town_intermediate	-0.153
Employee Home Town_wet	0.231

Key Findings:

Features with a positive correlation: Window to Wall Ratio (WWR), Thickness of the Wall Insulation, Share of the area served by AC(%), Smart control of HVAC, Smart control of lighting system, Area served by lighting, and Employee Home Town_wet. These features have a positive influence on overall IEQ satisfaction.

Features with a negative correlation: Roof Insulation U value, PM $_{2.5}$ level, PM $_{10}$ level, CO_2 PPM, and Building Type_Factory. These features have a negative influence on overall IEQ satisfaction.

It's important to note that correlation does not imply causation, and other factors not included in the dataset may also influence overall IEQ satisfaction. The correlation plot for Model 3 is shown in Figure 4.58.

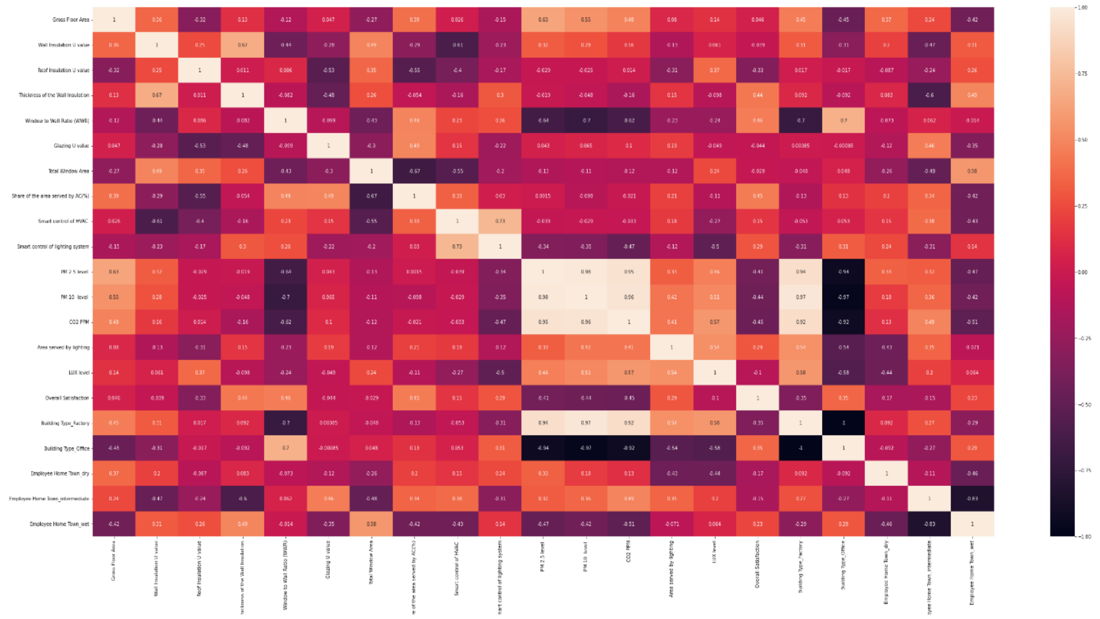


Figure 4.58: The correlation plot for Model 3

The models were evaluated using accuracy, RMSE, MAE and cross-validation metrics based on the results. Based on the performance metrics,

The performance metrics for the other regression models are as follows:

1. Support Vector Regression (SVR):

- Accuracy: 0.370

- RMSE: 0.580

- MAE: 0.431

- Cross-validated MAE: 0.460

The SVR model shows relatively lower accuracy and higher RMSE, MAE, and cross-validated MAE values than the other models. It may not be the best choice for accurately predicting the target variable.

2. Lasso Regression:

- Accuracy: 0.747

- RMSE: 0.367

- MAE: 0.263

- Cross-validated MAE: 0.262

The Lasso Regression model performs better than SVR but still has higher RMSE, MAE, and cross-validated MAE values than the Decision Tree and Random Forest models.

3. Decision Tree Regressor:

- Accuracy: 0.802

- RMSE: 0.325

- MAE: 0.179

- Cross-validated MAE: 0.179

The Decision Tree Regressor model shows higher accuracy and lower RMSE, MAE (Figure 4.59), and cross-validated MAE values than SVR and Lasso Regression. It demonstrates good predictive performance and effectively captures the underlying patterns in the data.

4. Random Forest Regressor:

- Accuracy: 0.802

- RMSE: 0.325

- MAE: 0.179

-Cross-validated MAE: 0.180

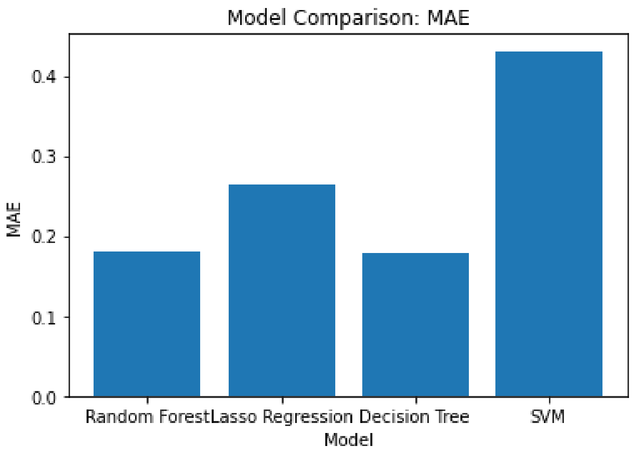


Figure 4.59: Model 4 comparison MAE values

The Random Forest Regressor model performs similarly to the Decision Tree model with slightly higher cross-validated MAE. It also shows good predictive performance and can be reliable for predicting the target variable.

In summary, both the Decision Tree and Random Forest models outperform SVR and Lasso Regression regarding accuracy, RMSE, MAE, and cross-validated MAE. Therefore, they are recommended for predicting the target variable in this scenario.

Table 4.35: The model comparison for IEQ comfort

Model	**R^2**	**RMSE**	**MAE**	**Cross-Validated MAE**
SVR	0.370	0.580	0.431	0.460
Lasso Regression	0.747	0.367	0.263	0.262
Decision Tree Regressor	0.802	0.325	0.179	0.180
Random Forest Regressor	0.802	0.325	0.179	0.179

The Variance Inflation Factor (VIF) is calculated for each feature to assess multicollinearity among the independent variables in the dataset (Table 4.36).

Table 4.36: VIF values for Model 4

Feature	**VIF**
Gross Floor Area	3.100
Wall Insulation U value	5.429
Roof Insulation U value	2.476
The thickness of the Wall Insulation	10.816
Window to Wall Ratio (WWR)	2.066
Glazing U value	2.174

Total Window Area	2.484
Share of the area served by AC(%)	8.102
Smart control of HVAC	10.016
Smart control of the lighting system	6.704

The features "PM $_{2.5}$ level," "PM $_{10}$ level," and "CO_2 PPM" have relatively high VIF values, indicating a higher degree of correlation with other variables in the model. This suggests that these variables may have a higher influence on the target variable and may need further examination to avoid multicollinearity issues in the analysis.

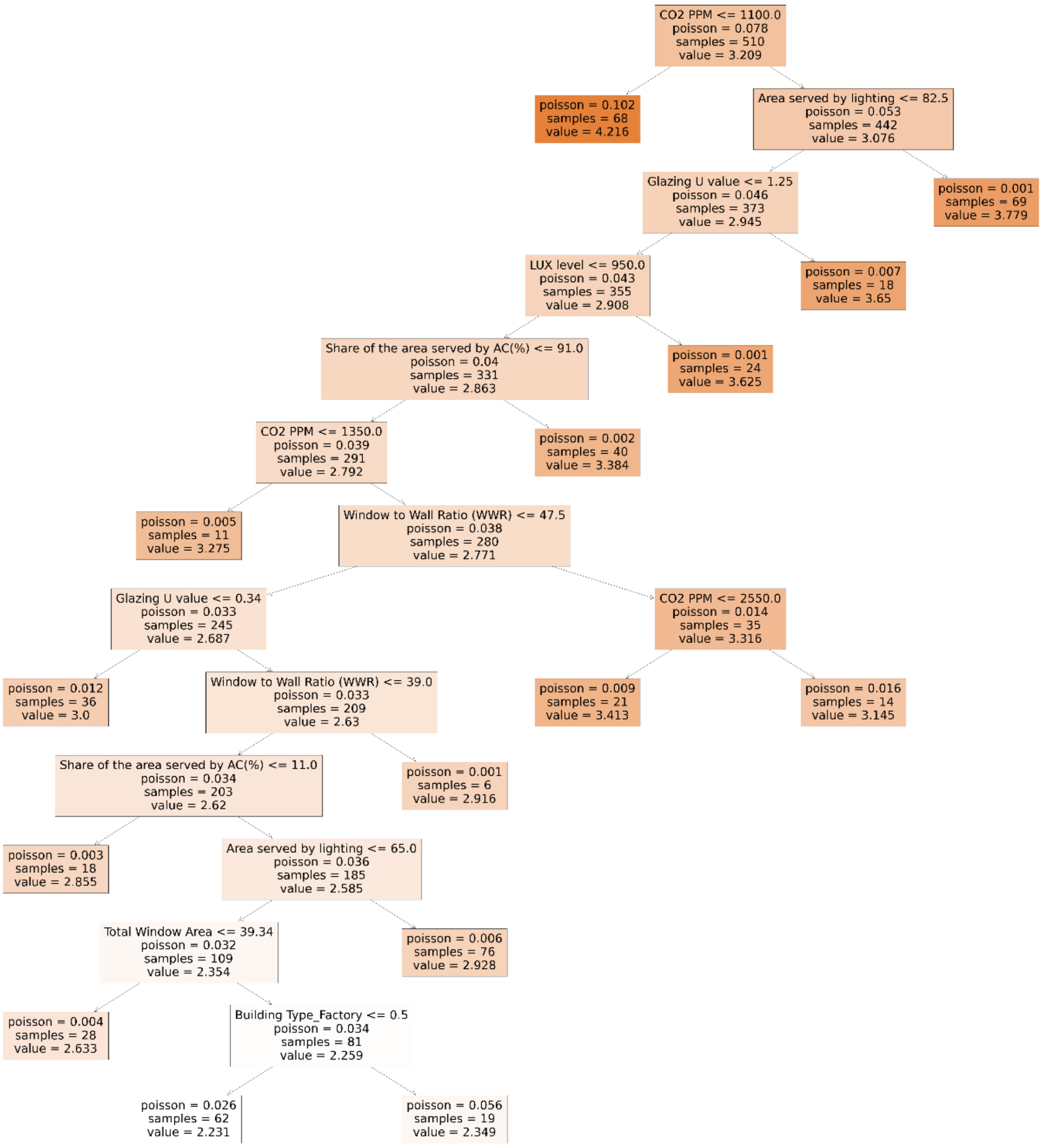


Figure 4.60: Decision tree of the Model 4

The root node, the first split in the tree, is based on the "CO_2 ppm" feature. If the window-to-wall ratio is less than or equal to 47.5, the tree proceeds to the left side of the split, otherwise to the right side (Figure 4.60).

The high ranking of this box suggests that the "window-to-wall ratio" feature strongly influences visual satisfaction outcomes. Figure 4.61 represents the model accuracy visually.

The output indicates that the average MAE across the 5 cross-validation folds is 0.188.

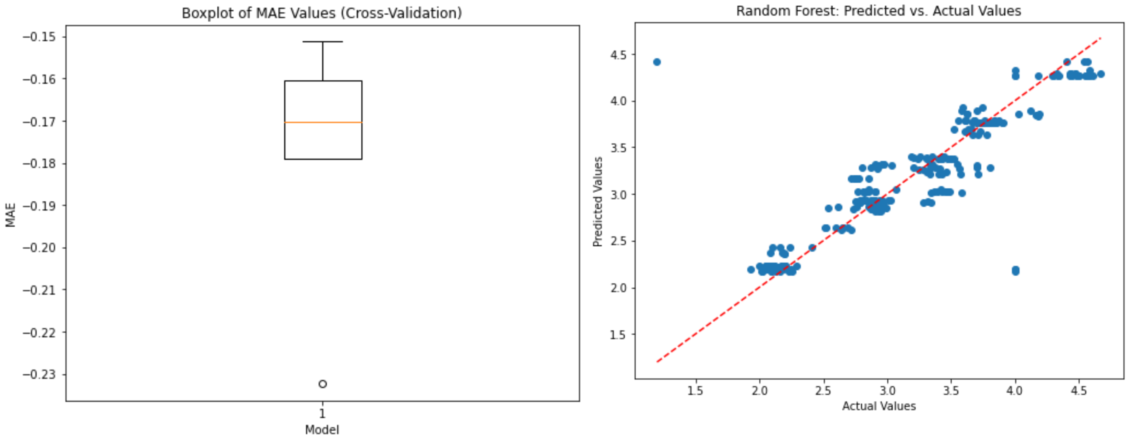


Figure 4.61: Predicted Vs Actual value of RF for Model 4 (Left), Box plot MAE values (cross-validation) for Model 4 (Right)

The model's average absolute error in predicting thermal satisfaction is approximately 0.179, indicating a relatively small deviation from the actual values.

The scatter plot points are highly scattered around the red line, which means the accuracy and reliability of the model are in a good position. There are a few outliers are also visible in the illustration.

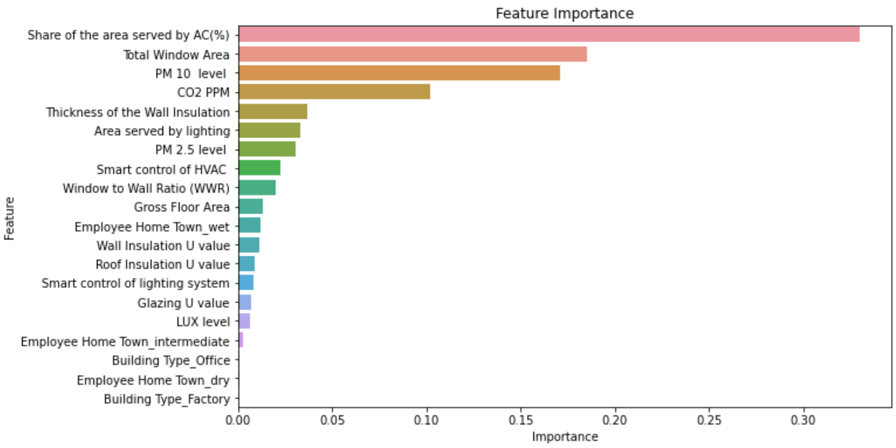


Figure 4.62: Feature importance of the IEQ comfort model variables

The figure highlights the feature index according to the importance of the overall IEQ comfort. The share of the area served by the AC is the most important factor. The second most important factor is the window area (Figure 4.62).

The codes files written for the predictive models are attached in APPENDIX B. While both Random Forest and Decision Tree models show similar results in terms of accuracy, RMSE, and MAE, there are several reasons why this study chose RF instead of DT for the further analysis.

(a) Bias-Variance Tradeoff: Decision Trees tend to have high variance and low bias, which means they can easily overfit the training data and may not generalize well to unseen data. Random Forest, on the other hand, is an ensemble model that combines multiple decision trees, reducing the variance and improving generalization performance.
(b) Model Robustness: Random Forest is more robust to outliers and noisy data compared to Decision Trees. The ensemble nature of Random Forest helps in reducing the impact of individual noisy data points or outliers.
(c) Reduced Overfitting: Random Forest introduces randomness in feature selection during the tree-building process, which helps to decorrelate the trees and reduces the risk of overfitting. Decision Trees can easily overfit the training data, leading to poor performance on the test data.
(d) Feature Importance: Random Forest provides a feature importance measure, which can help to identify the most significant variables that contribute to the target variable. This information can be valuable for understanding the underlying relationships and making informed decisions.
(e) Higher Accuracy: While both models may have similar accuracy in this specific study, Random Forest generally tends to perform better on average for complex and high-dimensional datasets. It is known to be one of the most powerful and versatile machine learning algorithms.
(f) Reduced Variance: Random Forest averages predictions from multiple trees, which reduces the variance of the predictions and provides more stable results.
(g) Cross-validation performance: While Decision Trees may show good performance on the training data, they can sometimes perform poorly on unseen data. Random Forest typically provides more consistent performance across different cross-validation folds.

4.18 Developing a user interface

All the models were saved as pickle files to create forms. The code was provided as a Flask web application that serves as a user interface for making predictions using machine learning models. The Visual Studio code file written for the application is attached in APPENDIX C. The summary of the application is as follows:

(a) The application is built using the Flask framework, a lightweight and extensible web framework for Python.
(b) It imports the necessary libraries, including Flask, pickle, numpy, and Sklearn.
(c) The Flask application is created using the `Flask(__name__)` constructor.
(d) Four prediction functions (`prediction`, `prediction_visual`, `prediction_IAQ`, `prediction_OS`) are defined, each of which loads a trained machine learning model from a saved pickle file and uses it to make predictions on input data.

(e) The Flask routes are defined using the `@app.route()` decorator. There are four routes: `'/'`, `'/Visual'`, `'/IAQ'`, and `'/OS'`.
(f) Each route corresponds to a different prediction task (Thermal, Visual, Indoor Air Quality, Overall Satisfaction).
(g) The form data submitted by the user is extracted using the `request. Form` dictionary and the input features are collected in a list.
(h) The categorical features are one-hot encoded by traversing predefined lists (`Building_Type_list`, `Employee_Home_Town_list`), and the resulting features are added to the feature list.
(i) The `prediction` functions are called with the feature list as input to obtain the predicted values.
(j) The predicted values are rendered in the `index.html` template and displayed to the user.

Overall, this Flask application provides a user-friendly interface for users to input data and obtain predictions using pre-trained machine learning models.

Designers, engineers or other relevant professionals can feed the data and assess the IEQ comfort of their employees according to their building factors. The interface enables professionals to assess Thermal, visual and IAQ comfort separately or asses the overall IEQ comfort of their built environment. The created user interface is shown in Figure 4.63. The form is published on http://127.0.0.1:5000 .

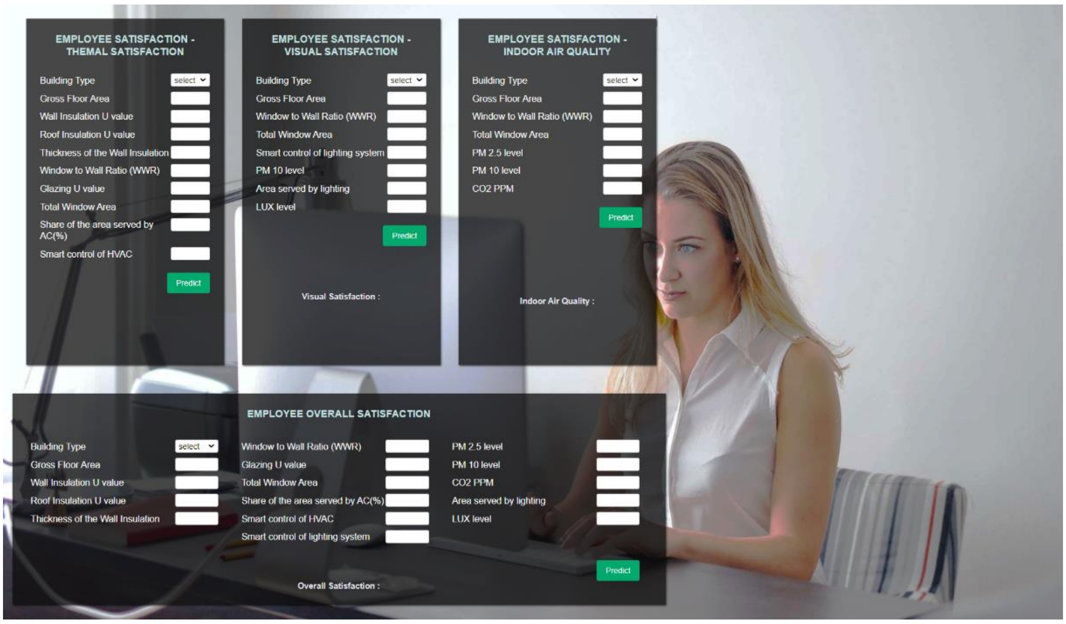


Figure 4.63: User interface to evaluate the employee IEQ satisfaction

4.19 Chapter summary

The pilot survey in the research on employee satisfaction with green buildings had several objectives, including gathering data for the main study, testing the feasibility of the questionnaire, and developing a theoretical model for decision-making. The survey employed stratified random sampling and collected data online and offline. The conclusions drawn from the pilot survey led to adjustments in the primary study, such as simplifying questions, removing scientific terminology, reducing the Likert scale, and focusing on micro-climate factors. The pre-screening survey collected 1,369 responses, with exclusions for non-returned, partially filled, and damaged questionnaires. The final sample size for analysis was 1,091 responses, ensuring data integrity and quality. Pre-screening is important for maintaining data quality, but it has limitations and potential biases to consider.

The descriptive analysis of the research findings provides insights into various aspects of the study. The distribution of respondents across workplace settings (office spaces and factories), gender, hometown climate zones, age groups, working hours, and work experience is examined. The variables measured in the study include thermal comfort, visual comfort, and indoor air quality (IAQ) satisfaction. Cronbach's alpha values indicate high internal consistency and reliability for these variables. The data deviate from a normal distribution, warranting non-parametric statistical tests.

The analysis of hypotheses explores the relationships between variables and employee responses. Gender does not significantly impact satisfaction levels. However, age groups, working hours, hometown climate zones, and distance from windows significantly influence satisfaction and comfort. Employees in office spaces 1-2 meters from windows are more comfortable. Shading devices are identified as a potential factor contributing to this difference. In factory spaces, the comfort zones differ from those in office spaces.

The descriptive analysis provides a comprehensive understanding of employee experiences in green buildings. It highlights the distribution of respondents and significant relationships between satisfaction, comfort, and various factors. These findings have implications for future research and design considerations.

A predictive model for Thermal comfort is developed using Python programming. The model development process involves various steps, including data preprocessing, model training and evaluation, performance assessment, hyperparameter tuning, and feature analysis.

The dataset used for the model contains information related to various features, such as 'EmpID', 'Building Type', 'Building', 'Building Location', 'Employee Home Town', and several attributes related to thermal satisfaction. The target variable of interest is 'Thermal Satisfaction'. Data preprocessing techniques are applied to handle non-numerical columns, remove irrelevant columns, and prepare the data for modelling tasks. Exploratory data analysis is conducted to gain insights into the dataset's structure and characteristics.

Multiple regression algorithms, including Support Vector Regression (SVR), Lasso Regression, Decision Tree Regression, and Random Forest Regression, are employed to develop predictive models for thermal comfort. These models are evaluated using performance metrics such as accuracy, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). Cross-validation techniques are also applied to assess the models' generalizability and robustness. The Random Forest Regressor is the best model for this scenario based on its high accuracy and low error metrics. A grid search approach optimises the model's hyperparameters and determines the best model.

The Variance Inflation Factor (VIF) is calculated to assess potential multicollinearity among the independent variables. The analysis identifies variables with a high degree of correlation, which can affect the model's interpretability and stability.The decision tree from the best Random Forest model is visualized using the `plot_tree` function, providing insights into the model's decision-making process and feature importance.

The model's performance is further assessed through cross-validation and visualized through a scatter plot comparing actual and predicted values of thermal satisfaction. The model demonstrates exemplary performance and consistency in its predictions.The feature importance of the Random Forest model is calculated and visualized to identify the most significant features contributing to thermal satisfaction prediction. The share of the area served by air conditioning is the most important factor, followed by the thickness of the wall insulation.

The developed predictive models for thermal comfort, visual comfort, IAQ satisfaction and IEQ comfort demonstrate their ability to predict thermal satisfaction based on the provided dataset accurately. The model selection, evaluation, and feature analysis provide valuable insights into the factors influencing thermal comfort and can be used to inform decision-making in improving indoor environmental quality. Finally, a successful user interface was created.

5 CONCLUSION

The literature review highlights the significance of energy efficiency in green buildings and its impact on environmental conservation, resource efficiency, economic benefits, occupant health and comfort, and regulatory compliance. Green buildings are crucial in mitigating climate change by reducing energy consumption and greenhouse gas emissions. They contribute to resource efficiency and security by minimizing energy waste and incorporating renewable energy sources. Energy-efficient practices in green buildings lead to substantial cost savings and improved financial performance for building owners and occupants. Moreover, they prioritize occupants' health, comfort, and productivity by providing a sustainable and conducive indoor environment.

The review emphasizes the factors influencing the energy efficiency of green buildings, including building design, HVAC systems, insulation and sealing, lighting systems, renewable energy integration, and occupant behaviour. When properly understood and addressed, these factors contribute to achieving and sustaining energy efficiency in green buildings. Additionally, the review acknowledges the geographic and time considerations by encompassing studies conducted in different locations and incorporating recent advancements and trends in energy efficiency measures.

Based on the systematic literature review and bibliographic analysis conducted in this study, the research has achieved its objectives of identifying main factors and existing decision-making models, performing bibliographic analysis and thematic mapping, and presenting the most appropriate methods for a hybrid decision-making model framework. By utilizing tools such as journal search engines, ArcGIS, and NodeXL, the study successfully analyzed various publications and correlations related to employee satisfaction, thermal comfort, and building-related decision models.

Overall, the literature review provides a comprehensive understanding of the importance of energy efficiency in green buildings, the factors influencing energy efficiency, and the specific types of green buildings and energy efficiency measures employed in different sectors. This knowledge is invaluable for designing and constructing environmentally responsible buildings that contribute to a more sustainable and energy-efficient future. By implementing energy-efficient practices, green buildings can significantly address global sustainability challenges and create healthier and more comfortable indoor environments for occupants.

The data collection process comprised an Employee Satisfaction Evaluation (ESE) questionnaire survey and physical measurements of thermal comfort-related parameters in 14 LEED or GBCSL-certified office buildings. This comprehensive approach allowed for identifying key building parameters and developing predictive machine learning models to analyze the collected data accurately.

The review also focused on the impact of building structural factors on indoor environmental quality (IEQ) and employee satisfaction. The study gained insights into the crucial factors contributing to a healthier and more sustainable indoor environment by examining existing research. The research questions guiding the literature review were centred around identifying green building structural factors that impact IEQ, employee satisfaction factors related to IEQ, and the types of machine learning regression models used in similar research.

The study's results revealed various building structural factors influencing indoor environmental quality, including building materials, ventilation systems, thermal insulation, lighting design, and acoustics. These findings contribute to a comprehensive understanding of the current state of knowledge and provide insights for architects, engineers, and policymakers to improve building design and construction practices for healthier and more sustainable indoor environments.

This study has successfully identified the factors and decision-making models relevant to employee satisfaction and indoor environmental quality in green buildings. The study has provided valuable insights into the key parameters and factors impacting occupant comfort using machine learning regression models and thorough data collection processes. This research's findings can inform future decision-making processes and guide the design and operation of green buildings, ultimately leading to improved occupant satisfaction, well-*being*, and overall environmental sustainability.

The descriptive analysis of the research findings provides valuable insights into various aspects of the study. The distribution of respondents across different workplace settings, including office spaces and factories, allows for a comparison between these two distinct work environments. The gender distribution of the respondents provides insights into potential gender-related differences in the perception and experience of green buildings. Categorizing respondents' hometowns into different climatic zones allows for examining the influence of climatic conditions on employee satisfaction and comfort. Analyzing the age groups of the respondents helps understand variations in satisfaction levels among different generations. The distribution of working hours and work experience provides insights into the impact of time spent in the workplace on employee satisfaction and comfort.

The descriptive analysis also highlights the variables and constructs measured in the study. The variables include thermal comfort, visual comfort, and indoor air quality (IAQ) satisfaction. The Cronbach's alpha values for these variables indicate high internal consistency and reliability. The normality tests, such as the Kolmogorov-Smirnov and Shapiro-Wilk tests and graphical representations, show that the data deviate from a normal distribution. Therefore, non-parametric statistical tests are appropriate for further analysis.

The analysis of hypotheses helps to understand the nature of employee responses and explore relationships between variables. The hypotheses examine the influence of gender, age, working hours, hometown climate zones, and distance from windows on employee satisfaction and comfort. The statistical tests, including Mann-Whitney U, Kruskal-Wallis, and pairwise comparisons, reveal significant differences in employee responses based on these factors.

The results suggest no significant difference in satisfaction levels based on gender. However, age groups and working hours significantly impact visual comfort and IAQ satisfaction in office spaces and factories. The employees' hometown climate zones also significantly influence thermal comfort, visual comfort, and IAQ satisfaction. The distance between the window and the work desk shows a significant relationship between employee satisfaction and comfort, with employees closer to windows generally expressing higher satisfaction levels.

Further analysis reveals that employees in office spaces within 1-2 meters from the window are more comfortable than those in closer proximity. The presence of shading devices is identified as a potential factor contributing to this difference. In factory spaces, the comfort zones differ from those in office spaces, with employees in the middle expressing higher satisfaction levels.

The descriptive analysis provides a comprehensive understanding of the research findings. It highlights the distribution of respondents across different workplace settings, demographic characteristics, climate zones, and work-related factors. Analyzing variables and constructs and exploring hypotheses reveal significant relationships between employee satisfaction, comfort, and various factors. These findings contribute to the overall understanding of employee experiences in green buildings and provide valuable insights for future research and design considerations.

Correlation analysis revealed relationships between the remaining variables and thermal satisfaction in Model 1. Key findings included: Gross floor area, thickness of wall insulation, and window-to-wall ratio positively correlated with thermal satisfaction, Roof insulation U value and glazing U value negatively correlated with thermal satisfaction and Share of the area served by AC showed a strong positive correlation with thermal satisfaction. Cross-validation demonstrated the model's good predictive performance, with a low Mean Absolute Error (MAE) of approximately 0.237 in the Model 1, thermal comfort.

Model 2 focuses on developing a predictive model for Visual Comfort. The correlation analysis revealed Building Factors such as Window Wall Ratio (WWR) and Total Window Area positively correlate with Visual Satisfaction, while Building Type is also influential, with offices having higher satisfaction and factories having lower satisfaction. Lighting Factors such as Smart control of lighting systems and Area served by lighting show positive correlations with Visual Satisfaction. Environmental Factors such as Gross Floor Area exhibits a weak negative correlation, while PM 10 level negatively impacts Visual Satisfaction.The Decision Tree Regressor and Random Forest Regressor show high accuracy and low RMSE, outperforming the Support Vector Regression and Lasso Regression models. The random forest model achieves an accuracy score of 0.953. The scatter plot and model accuracy visually demonstrate the reliability and accuracy of the predictive model for Visual Comfort. The most influential factor in predicting visual satisfaction is the PM 10 level, followed by the window-to-wall ratio.

Model 3 aims to develop a predictive model for Indoor Air Quality (IAQ) Satisfaction. The correlation analysis revealed, Factors positively influencing IAQ Satisfaction: Window Wall Ratio (WWR), Building Type_Office, and Employee Home Town_wet, factors negatively impacting IAQ Satisfaction: PM 2.5 level, PM 10 level, CO2 PPM, and Building Type_Factory and factors with weaker correlations: Gross Floor Area, Total Window Area, Employee Home Town_dry, and Employee Home Town_intermediate. The Variance Inflation Factor (VIF) analysis is conducted, and features like "PM 2.5 level," "PM 10 level," and "CO2 PPM" show relatively high VIF values, indicating a higher correlation with other variables and the need for further examination to avoid multicollinearity issues.The most influential factors in predicting IAQ Satisfaction are PM 10 level and CO2 PPM, indicating their importance in determining indoor air quality. The most important building structural factors are window-to-wall ratio and total window are

Model 4 focuses on developing a predictive model for overall Indoor Environmental Quality (IEQ) Satisfaction. The dataset includes all the variables considered in Models 1, 2, and 3. The correlation analysis revealed, features positively influencing overall IEQ Satisfaction: Window to Wall Ratio (WWR), Thickness of the Wall Insulation, Share of the area served by AC(%), Smart control of HVAC, Smart control of lighting system, Area served by lighting, and Employee Home Town_wet, and features negatively impacting overall IEQ Satisfaction: Roof Insulation U value, PM 2.5 level, PM 10 level, CO2 PPM, and Building Type_Factory. The Variance Inflation Factor (VIF) analysis is conducted, and features like "PM 2.5 level," "PM 10 level," and "CO2 PPM" show relatively high VIF values, indicating a higher correlation with other variables and the need for further examination to avoid multicollinearity issues. The most influential factors in predicting overall IEQ Satisfaction are the "Share of the area served by AC" and "Total Window Area," indicating their importance in determining indoor environmental quality. Model 4 provides valuable insights into the factors affecting overall IEQ Satisfaction and presents reliable predictive models to help optimize indoor environmental conditions for occupants' comfort and well-being.

The web application provides a user-friendly interface, enabling professionals such as designers and engineers to easily assess the Thermal, Visual, and IAQ comfort separately or the overall IEQ comfort of their building environment.

6 RECOMMENDATIONS FOR FUTURE WORKS

The context of green office buildings might introduce certain biases that need to be considered when extrapolating the results. While this performance is notable, it indicates room for improvement. The model's effectiveness may vary across different datasets or settings, and employing additional evaluation metrics could provide valuable insights into its strengths and weaknesses.

The study primarily focused on the factors influencing indoor air quality comfort satisfaction within the surveyed buildings. However, it is crucial to consider the potential influence of external factors, such as seasonal variations, climate change, or individual preferences, on long-term indoor air quality comfort. These factors can significantly impact occupants' satisfaction and well-being, and incorporating them into the study would offer a more comprehensive understanding of the indoor environment's impact.

Furthermore, it's important to acknowledge that indoor air quality comfort is a dynamic concept that can change over time. The study might have overlooked the variability in occupants' preferences and perceptions of comfort during different hours of the day, seasons, or with changes in building occupancy. Future research could explore real-time sensor data integration with advanced machine learning techniques to create dynamic and adaptable models. By doing so, we could anticipate and enhance indoor air quality comfort in real-time, ensuring a consistently pleasant environment for occupants.

This research highlights the significance of considering occupant satisfaction and well-being in green offices by bridging the gap between building design, human comfort, and machine learning. By expanding the scope of the study to include a broader range of influencing factors and adopting dynamic modeling approaches, we can create healthier, more comfortable, and sustainable indoor environments for office occupants.

REFERENCES & BIBLIOGRAPHY

Abbaszadeh, S., Zagreus, L., Lehrer, D., & Huizenga, C. (2006). Occupant satisfaction with indoor environmental quality in green buildings. *Healthy Buildings Conference*.

Abdou, A., & Budaiwi, I. (2013). The variation of thermal conductivity of fibrous insulation materials under different levels of moisture content. *Construction and Building Materials*, *43*, 533–544. https://doi.org/10.1016/J.CONBUILDMAT.2013.02.058

Abeyrathna, W., John, G. K. P., Jayasinghe, R., Ariyarathna, R. I. S., Hendawitharana, M. P., Halwaturaa, R. U., Kaklauskas, A., Arooz, F. R., & Perera, A. S. (2023). [Review] Exploring an integrated decision-making model to enhance the employee-oriented built environment in urban green buildings: A Review. *Qeios*. https://doi.org/10.32388/70OGM8.5

Abeyrathna, W. P., Ariyarathna, I. S., Halwatura, R. U., Arooz, F. R., Perera, A. S., & Kaklauskas, A. (2023). ANN prediction model to improve employees' thermal satisfaction in tropical green office buildings. *Asian Journal of Civil Engineering*, 1–16. https://doi.org/10.1007/S42107-023-00779-Y/METRICS

Aditya, L., Mahlia, T. M. I., Rismanchi, B., Ng, H. M., Hasan, M. H., Metselaar, H. S. C., Muraza, O., & Aditiya, H. B. (2017). A review on insulation materials for energy conservation in buildings. *Renewable and Sustainable Energy Reviews*, *73*, 1352–1365. https://doi.org/10.1016/J.RSER.2017.02.034

Afanador, N. L., Smolinska, A., Tran, T. N., & Blanchet, L. (2016). Unsupervised random forest: a tutorial with case studies. *Journal of Chemometrics*, *30*(5), 232–241. https://doi.org/10.1002/CEM.2790

Aguilar, A. J., de la Hoz-Torres, M. L., Ruiz, D. P., & Martínez-Aires, M. D. (2022). Monitoring and Assessment of Indoor Environmental Conditions in Educational Building Using Building Information Modelling Methodology. *International Journal of Environmental Research and Public Health 2022, Vol. 19, Page 13756*, *19*(21), 13756. https://doi.org/10.3390/IJERPH192113756

Ahmed, T., Kumar, P., & Mottet, L. (2021). Natural ventilation in warm climates: The challenges of thermal comfort, heatwave resilience and indoor air quality. *Renewable and Sustainable Energy Reviews*, *138*, 110669. https://doi.org/10.1016/J.RSER.2020.110669

Akkurt, G. G., Aste, N., Borderon, J., Buda, A., Calzolari, M., Chung, D., Costanzo, V., Del Pero, C., Evola, G., Huerto-Cardenas, H. E., Leonforte, F., Lo Faro, A., Lucchi, E., Marletta, L., Nocera, F., Pracchi, V., & Turhan, C. (2020). Dynamic thermal and hygrometric simulation of historical buildings: Critical factors and possible solutions. *Renewable and Sustainable Energy Reviews*, *118*, 109509. https://doi.org/10.1016/J.RSER.2019.109509

Al-Omari, K., & Okasheh, H. (2017). The Influence of Work Environment on Job Performance: A Case Study of Engineering Company in Jordan. *International Journal of Applied Engineering Research*, *12*, 15544–15550. http://www.ripublication.com

Al horr, Y., Arif, M., Katafygiotou, M., Mazroei, A., Kaushik, A., & Elsarrag, E. (2016). Impact of indoor environmental quality on occupant well-being and comfort: A review of the literature. *International Journal of Sustainable Built Environment*, *5*(1), 1–11.

https://doi.org/10.1016/J.IJSBE.2016.03.006

Albatayneh, A., Juaidi, A., Abdallah, R., & Manzano-Agugliaro, F. (2021). Influence of the Advancement in the LED Lighting Technologies on the Optimum Windows-to-Wall Ratio of Jordanians Residential Buildings. *Energies 2021, Vol. 14, Page 5446*, *14*(17), 5446. https://doi.org/10.3390/EN14175446

Ali, K. A., Ahmad, M. I., & Yusup, Y. (2020). Issues, Impacts, and Mitigations of Carbon Dioxide Emissions in the Building Sector. *Sustainability 2020, Vol. 12, Page 7427*, *12*(18), 7427. https://doi.org/10.3390/SU12187427

Alsaleem, F., Tesfay, M. K., Rafaie, M., Sinkar, K., Besarla, D., & Arunasalam, P. (2020). An IoT Framework for Modeling and Controlling Thermal Comfort in Buildings. *Frontiers in Built Environment*, *6*, 482045. https://doi.org/10.3389/FBUIL.2020.00087/BIBTEX

Altman, N., & Krzywinski, M. (2017). Ensemble methods: bagging and random forests. *Nature Methods*, *14*(10), 933–935. https://go.gale.com/ps/i.do?p=HRCA&sw=w&issn=15487091&v=2.1&it=r&id=GALE%7CA625179388&sid=googleScholar&linkaccess=fulltext

Altomonte, S. ., Saadouni, S. ., & Schiavon, S. (2016). Occupant satisfaction in LEED and BREEAM-certified office buildings. *N Proceedings of the PLEA 2016-36th International Conference on Passive and Low Energy Architecture: Cities, Buildings, People: Towards Regenerative Environments*.

Altomonte, S., & Schiavon, S. (2013). Occupant satisfaction in LEED and non-LEED certified buildings. *Building and Environment*, *68*, 66–76. https://doi.org/10.1016/j.buildenv.2013.06.008

Altomonte, S., Schiavon, S., Kent, M. G., & Brager, G. (2019). Indoor environmental quality and occupant satisfaction in green-certified buildings. *Building Research and Information*, *47*(3), 255–274. https://doi.org/10.1080/09613218.2018.1383715

Andrews, D. C. (2022). Commentary: A Space for Place in Business Communication Research Updated. *Https://Doi.Org/10.1177/23294884221126491*. https://doi.org/10.1177/23294884221126491

Arakawa Martins, L., Soebarto, V., Williamson, T., & Pisaniello, D. (2022). Personal thermal comfort models: a deep learning approach for predicting older people's thermal preference. *Smart and Sustainable Built Environment*, *11*(2), 245–270. https://doi.org/10.1108/SASBE-08-2021-0144/FULL/XML

Aryal, A., & Becerik-Gerber, B. (2020). Thermal comfort modeling when personalized comfort systems are in use: Comparison of sensing and learning methods. *Building and Environment*, *185*, 107316. https://doi.org/10.1016/J.BUILDENV.2020.107316

ASHRAE Standard 90.1. (2019). *ANSI/ASHRAE/IESNA Addenda a, b, c, g, h, i, j, k, l, m, n, p, q, s, t, u, w, y, ad, and aw to ANSI/ASHRAE/IESNA Standard 90.1-2007*. *4723*.

Assaf, S., & Srour, I. (2021). Using a data driven neural network approach to forecast building occupant complaints. *Building and Environment*, *200*, 107972. https://doi.org/10.1016/J.BUILDENV.2021.107972

Attaianese, E. (2012). A broader consideration of human factor to enhance sustainable building design. *IOS Press*, *41*(Supplement 1), 2155–2159. https://doi.org/10.3233/WOR-2012-1020-2155

Awadh, O. (2017). Sustainability and green building rating systems: LEED, BREEAM, GSAS and Estidama critical analysis. *Journal of Building Engineering*, *11*, 25–29. https://doi.org/10.1016/J.JOBE.2017.03.010

Azar, E., O'Brien, W., Carlucci, S., Hong, T., Sonta, A., Kim, J., Andargie, M. S., Abuimara, T., El Asmar, M., Jain, R. K., Ouf, M. M., Tahmasebi, F., & Zhou, J. (2020). Simulation-aided occupant-centric building design: A critical review of tools, methods, and applications. *Energy and Buildings*, *224*, 110292. https://doi.org/10.1016/J.ENBUILD.2020.110292

Baird, G., Leaman, A., & Thompson, J. (2012). A comparison of the performance of sustainable buildings with conventional buildings from the point of view of the users. *Architectural Science Review*, *55*(2), 135–144. https://doi.org/10.1080/00038628.2012.670699

Balaban, O., & Puppim de Oliveira, J. A. (2017). Sustainable buildings for healthier cities: assessing the co-benefits of green buildings in Japan. *Journal of Cleaner Production*, *163*, S68–S78. https://doi.org/10.1016/J.JCLEPRO.2016.01.086

Balasundaram, S., & Meena, Y. (2019). Robust Support Vector Regression in Primal with Asymmetric Huber Loss. *Neural Processing Letters*, *49*(3), 1399–1431. https://doi.org/10.1007/S11063-018-9875-8/METRICS

Balogun, A. L., Tella, A., Baloo, L., & Adebisi, N. (2021). A review of the inter-correlation of climate change, air pollution and urban sustainability using novel machine learning algorithms and spatial information science. *Urban Climate*, *40*, 100989. https://doi.org/10.1016/J.UCLIM.2021.100989

Bavaresco, M. V., Ghisi, E., D'Oca, S., & Pisello, A. L. (2021). Triggering occupant behaviour for energy sustainability: Exploring subjective and comfort-related drivers in Brazilian offices. *Energy Research & Social Science*, *74*, 101959. https://doi.org/10.1016/J.ERSS.2021.101959

Berardi, U. (2019). Light transmittance characterization and energy-saving analysis of a new selective coating for in situ window retrofit. *Https://Doi.Org/10.1080/23744731.2019.1620546*, *25*(9), 1152–1163. https://doi.org/10.1080/23744731.2019.1620546

Bevilacqua, P., Benevento, F., Bruno, R., & Arcuri, N. (2019). Are Trombe walls suitable passive systems for the reduction of the yearly building energy requirements? *Energy*, *185*, 554–566. https://doi.org/10.1016/J.ENERGY.2019.07.003

Bevilacqua, P., Mazzeo, D., Bruno, R., & Arcuri, N. (2017). Surface temperature analysis of an extensive green roof for the mitigation of urban heat island in southern mediterranean climate. *Energy and Buildings*, *150*, 318–327. https://doi.org/10.1016/J.ENBUILD.2017.05.081

Bourhnane, S., Abid, M. R., Lghoul, R., Zine-Dine, K., Elkamoun, N., & Benhaddou, D. (2020). Machine learning for energy consumption prediction and scheduling in smart buildings. *SN Applied Sciences*, *2*(2), 1–10. https://doi.org/10.1007/S42452-020-2024-9/FIGURES/11

Bozsaky, D. (2010). The historical development of thermal insulation materials. *Periodica Polytechnica Architecture*, *41*(2), 49–56. https://doi.org/10.3311/PP.AR.2010-2.02

Brambilla, A., Salvalai, G., Imperadori, M., & Sesana, M. M. (2018). Nearly zero energy building renovation: From energy efficiency to environmental efficiency, a pilot case study. *Energy and Buildings*, *166*, 271–283. https://doi.org/10.1016/J.ENBUILD.2018.02.002

BREEAM - BRE Group. (n.d.). Retrieved July 16, 2023, from https://bregroup.com/products/breeam/

Brown, Z., Cole, R. J., Robinson, J., & Dowlatabadi, H. (2010). Evaluating user experience in green buildings in relation to workplace culture and context. *Facilities*, *28*(3/4), 225–238. https://doi.org/10.1108/02632771011023168

Bueno, A. M., de Paula Xavier, A. A., & Broday, E. E. (2021). Evaluating the Connection between Thermal Comfort and Productivity in Buildings: A Systematic Literature Review. *Buildings 2021, Vol. 11, Page 244*, *11*(6), 244. https://doi.org/10.3390/BUILDINGS11060244

Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A Review on Sample Size Determination for Cronbach's Alpha Test: A Simple Guide for Researchers. *The Malaysian Journal of Medical Sciences : MJMS*, *25*(6), 85. https://doi.org/10.21315/MJMS2018.25.6.9

Cabeza, L. F., Castell, A., Medrano, M., Martorell, I., Pérez, G., & Fernández, I. (2010). Experimental study on the performance of insulation materials in Mediterranean construction. *Energy and Buildings*, *42*(5), 630–636. https://doi.org/10.1016/J.ENBUILD.2009.10.033

Carli, R., Cavone, G., Dotoli, M., Epicoco, N., & Scarabaggio, P. (2019). Model predictive control for thermal comfort optimization in building energy management systems. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, *2019-October*, 2608–2613. https://doi.org/10.1109/SMC.2019.8914489

Chen, M., Bai, J., Zhu, S., Yang, B., & Dai, F. (2021). The influence of neighborhood-level urban morphology on PM2.5 variation based on random forest regression. *Atmospheric Pollution Research*, *12*(8), 101147. https://doi.org/10.1016/J.APR.2021.101147

Chen, S., Zhang, G., Xia, X., Setunge, S., & Shi, L. (2020). A review of internal and external influencing factors on energy efficiency design of buildings. *Energy and Buildings*, *216*, 109944. https://doi.org/10.1016/j.enbuild.2020.109944

Cheung, T., Schiavon, S., Graham, L. T., & Tham, K. W. (2021). Occupant satisfaction with the indoor environment in seven commercial buildings in Singapore. *Building and Environment*, *188*(November). https://doi.org/10.1016/j.buildenv.2020.107443

Chew, I., Karunatilaka, D., Tan, C. P., & Kalavally, V. (2017). Smart lighting: The way forward? Reviewing the past to shape the future. *Energy and Buildings*, *149*, 180–191. https://doi.org/10.1016/J.ENBUILD.2017.04.083

Chowdhury, S., & Alam, M. R. (2011). *DESIGN APPROACH OF ENERGY EFFICIENT READY MADE GARMENTS FACTORY IN VIEW OF THERMAL COMFORT*. https://doi.org/10.13140/RG.2.1.2858.2002

Chua, S. J. L., Ali, A. S., & Lim, M. E. L. (2016). Physical Environment Comfort Impacts on Office Employee's Performance. *MATEC Web of Conferences*, *66*, 00124. https://doi.org/10.1051/MATECCONF/20166600124

Colclough, S., Kinnane, O., Hewitt, N., & Griffiths, P. (2018). Investigation of nZEB social housing built to the Passive House standard. *Energy and Buildings*, *179*, 344–359. https://doi.org/10.1016/J.ENBUILD.2018.06.069

Crosby, S., & Rysanek, A. (2022). Predicting thermal satisfaction as a function of indoor CO2 levels: Bayesian modelling of new field data. *Building and Environment*, *209*, 108569. https://doi.org/10.1016/J.BUILDENV.2021.108569

Cui, H., Huang, D., Fang, Y., Liu, L., & Huang, C. (2018). Webshell detection based on random forest-gradient boosting decision tree algorithm. *Proceedings - 2018 IEEE 3rd International Conference on Data Science in Cyberspace, DSC 2018*, 153–160. https://doi.org/10.1109/DSC.2018.00030

Danaci, H. M., & Akin, N. (2022). Thermal insulation materials in architecture: a comparative test study with aerogel and rock wool. *Environmental Science and Pollution Research*, *29*(48), 72979–72990. https://doi.org/10.1007/S11356-022-20927-2/METRICS

Dangol, R., Islam, M. S., Hyvärinen, M., Bhushal, P., Puolakka, M., & Halonen, L. (2013). User acceptance studies for LED office lighting: Preference, naturalness and colourfulness. *Http://Dx.Doi.Org/10.1177/1477153513514424*, *47*(1), 36–53. https://doi.org/10.1177/1477153513514424

Daniel, B. (2015). Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, *46*(5), 904–920. https://doi.org/10.1111/BJET.12230

Darko, A., Chan, A. P. C., Ameyaw, E. E., He, B.-J., & Olanipekun, A. O. (2017). Examining issues influencing green building technologies adoption: The United States green building experts' perspectives. *Energy and Buildings*, *144*, 320–332. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.03.060

Darko, A., Zhang, C., & Chan, A. P. C. (2017). Drivers for green building: A review of empirical studies. *Habitat International*, *60*, 34–49. https://doi.org/10.1016/J.HABITATINT.2016.12.007

De Bock, K. W., & De Caigny, A. (2021). Spline-rule ensemble classifiers with structured sparsity regularization for interpretable customer churn modeling. *Decision Support Systems*, *150*, 113523. https://doi.org/10.1016/J.DSS.2021.113523

Deegahawature, M. M. D. R., & Rupasinghe, O. H. S. N. (2019). Indoor environment quality in green buildings: A case of apparel firm in Sri Lanka. *Journal of Management Matters*, *6*(2), 35–44. www.rjt.ac.lk/mgt

Deshmukh, G., Birwal, P., Datir, R., & Patel, S. (2017). Thermal Insulation Materials: A Tool for Energy Conservation. *J Food Process Technol*, *8*, 670. https://doi.org/10.4172/2157-7110.1000670

Dinmohammadi, F., Han, Y., & Shafiee, M. (2023). Predicting Energy Consumption in Residential Buildings Using Advanced Machine Learning Algorithms. *Energies 2023, Vol. 16, Page 3748*, *16*(9), 3748. https://doi.org/10.3390/EN16093748

Dong, Y., Kong, J., Mousavi, S., Rismanchi, B., & Yap, P.-S. (2023). Wall Insulation Materials in Different Climate Zones: A Review on Challenges and Opportunities of Available Alternatives. *Thermo 2023, Vol. 3, Pages 38-65*, *3*(1), 38–65. https://doi.org/10.3390/THERMO3010003

Dou, J., Yunus, A. P., Tien Bui, D., Merghadi, A., Sahana, M., Zhu, Z., Chen, C. W., Khosravi, K., Yang, Y., & Pham, B. T. (2019). Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Science of The Total Environment*, *662*, 332–346. https://doi.org/10.1016/J.SCITOTENV.2019.01.221

Dräger, P., & Letmathe, P. (2022). Value losses and environmental impacts in the construction industry – Tradeoffs or correlates? *Journal of Cleaner Production*, *336*, 130435. https://doi.org/10.1016/J.JCLEPRO.2022.130435

Dudek, G. (2016). Pattern-based local linear regression models for short-term load forecasting. *Electric Power Systems Research*, *130*, 139–147. https://doi.org/10.1016/J.EPSR.2015.09.001

Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, *297*(3), 1178–1192. https://doi.org/10.1016/J.EJOR.2021.06.053

E.G.Hertwich, J. A. de L., & A. Arvesen, P. Bayer, J. Bergesen, E. Bouman, T. Gibon, G. Heath, C. Peña, P. Purohit, A. Ramirez, S. S. (2016). Green energy choices: the benefits, Risks Electricity, and trade-offs of low-carbon technologies for production — Summary for policy makers. In *United Nations Environment Programme*. www.unep.org/resourcepanel.

EDGE - Excellence in Design for Greater Efficiencies (EN) - EDGE Buildings. (2023). https://edgebuildings.com/edge-excellence-in-design-for-greater-efficiencies/

Elhazmi, A., Al-Omari, A., Sallam, H., Mufti, H. N., Rabie, A. A., Alshahrani, M., Mady, A., Alghamdi, A., Altalaq, A., Azzam, M. H., Sindi, A., Kharaba, A., Al-Aseri, Z. A., Almekhlafi, G. A., Tashkandi, W., Alajmi, S. A., Faqihi, F., Alharthy, A., Al-Tawfiq, J. A., … Arabi, Y. M. (2022). Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU. *Journal of Infection and Public Health*, *15*(7), 826–834. https://doi.org/10.1016/J.JIPH.2022.06.008

Elnaklah, R., Walker, I., & Natarajan, S. (2021). Moving to a green building: Indoor environment quality, thermal comfort and health. *Building and Environment*, *191*, 107592. https://doi.org/10.1016/J.BUILDENV.2021.107592

Es-Sakali, N., Kaitouni, S. I., Laasri, I. A., Mghazli, M. O., Cherkaoui, M., & Pfafferott, J. (2022). Assessment of the energy efficiency for a building energy model using different glazing windows in a semi-arid climate. *2022 13th International Renewable Energy Congress, IREC 2022*. https://doi.org/10.1109/IREC56325.2022.10001934

Estidama Program. (n.d.). Retrieved July 16, 2023, from https://pages.dmt.gov.ae/en/Urban-Planning/Estidama-Program

Fang, Y., Chen, G., Bick, M., & Chen, J. (2021). Smart textiles for personalized thermoregulation. *Chemical Society Reviews*, *50*(17), 9357–9374. https://doi.org/10.1039/D1CS00003A

Farzaneh, H., Malehmirchegini, L., Bejan, A., Afolabi, T., Mulumba, A., & Daka, P. P. (2021). Artificial Intelligence Evolution in Smart Buildings for Energy Efficiency. *Applied Sciences 2021, Vol. 11, Page 763*, *11*(2), 763. https://doi.org/10.3390/APP11020763

Fassio, F., Fanchiotti, A., & de Lieto Vollaro, R. (2014). Linear, Non-Linear and Alternative

Algorithms in the Correlation of IEQ Factors with Global Comfort: A Case Study. *Sustainability 2014, Vol. 6, Pages 8113-8127*, *6*(11), 8113–8127. https://doi.org/10.3390/SU6118113

Feng, Y., Wang, N., & Wang, J. (2020). Design of Real-Time Individualized Comfort Monitor System Used in Healthcare Facilities. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12429 LNCS*, 261–270. https://doi.org/10.1007/978-3-030-59987-4_19/COVER

Franco, A., Miserocchi, L., & Testi, D. (2021). A method for optimal operation of HVAC with heat pumps for reducing the energy demand of large-scale non residential buildings. *Journal of Building Engineering*, *43*, 103175. https://doi.org/10.1016/J.JOBE.2021.103175

Fritz, H., Tang, M., Kinney, K., & Nagy, Z. (2022). Evaluating machine learning models to classify occupants' perceptions of their indoor environment and sleep quality from indoor air quality. *Https://Doi.Org/10.1080/10962247.2022.2105439*, *72*(12), 1381–1397. https://doi.org/10.1080/10962247.2022.2105439

García-Floriano, A., López-Martín, C., Yáñez-Márquez, C., & Abran, A. (2018). Support vector regression for predicting software enhancement effort. *Information and Software Technology*, *97*, 99–109. https://doi.org/10.1016/J.INFSOF.2018.01.003

García Kerdan, I., & Morillón Gálvez, D. (2020). Artificial neural network structure optimisation for accurately prediction of exergy, comfort and life cycle cost performance of a low energy building. *Applied Energy*, *280*, 115862. https://doi.org/10.1016/J.APENERGY.2020.115862

Gasparella, A., Pernigotto, G., Cappelletti, F., Romagnoni, P., & Baggio, P. (2011). Analysis and modelling of window and glazing systems energy performance for a well insulated residential building. *Energy and Buildings*, *43*(4), 1030–1037. https://doi.org/10.1016/J.ENBUILD.2010.12.032

German Sustainable Building Council | DGNB GmbH. (n.d.). Retrieved July 16, 2023, from https://www.dgnb.de/en

Gesa, N., Roy Joseph Sarwuan, A., Newton, F., Roy, A. A., & Solomon, I. (2014). Investigation of the Thermal Insulation Properties of Selected Ceiling Materials used in Makurdi Metropolis (Benue State-Nigeria). *American Journal of Engineering Research (AJER)*, *03*(11), 245–250. https://doi.org/10.13140/RG.2.1.3701.3364

Gou, Z., Lau, S. S.-Y., & Shen, J. (2012). Indoor Environmental Satisfaction in Two LEED Offices and its Implications in Green Interior Design. *Indoor and Built Environment*, *21*(4), 503–514. https://doi.org/10.1177/1420326X11418700

Gou, Z., Lau, S. S.-Y., & Zhang, Z. (2012). A COMPARISON OF INDOOR ENVIRONMENTAL SATISFACTION BETWEEN TWO GREEN BUILDINGS AND A CONVENTIONAL BUILDING IN CHINA. *Journal of Green Building*, *7*(2), 89–104. https://doi.org/10.3992/jgb.7.2.89

Gou, Z., Prasad, D., & Lau, S. S.-Y. (2014). Impacts of green certifications, ventilation and office types on occupant satisfaction with indoor environmental quality. *Architectural Science Review*, *57*(3), 196–206. https://doi.org/10.1080/00038628.2014.908113

Gou, Z., & Xie, X. (2017). Evolving green building: triple bottom line or regenerative design?

Journal of Cleaner Production, *153*, 600–607. https://doi.org/10.1016/j.jclepro.2016.02.077

Green Building Council | Build a Brighter Future. (n.d.). Retrieved July 16, 2023, from https://www.srilankagbc.org/

Green Mark Certification Scheme | Building and Construction Authority (BCA). (n.d.). Retrieved July 16, 2023, from https://www1.bca.gov.sg/buildsg/sustainability/green-mark-certification-scheme

Green Star Rating System | Green Building Council of Australia. (n.d.). Retrieved July 16, 2023, from https://new.gbca.org.au/green-star/rating-system/

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, *27*(3), 659–678. https://doi.org/10.1007/S11222-016-9646-1/METRICS

Grzegorzewska, M., & Kirschke, P. (2021). The Impact of Certification Systems for Architectural Solutions in Green Office Buildings in the Perspective of Occupant Well-Being. *Buildings 2021, Vol. 11, Page 659*, *11*(12), 659. https://doi.org/10.3390/BUILDINGS11120659

Gulbinas, R., & Taylor, J. E. (2014). Effects of real-time eco-feedback and organizational network dynamics on energy efficient behavior in commercial buildings. *Energy and Buildings*, *84*, 493–500. https://doi.org/10.1016/J.ENBUILD.2014.08.017

Gupta, B., Uttarakhand, P., & Rawat, I. A. (2017). Analysis of Various Decision Tree Algorithms for Classification in Data Mining. *International Journal of Computer Applications*, *163*(8), 975–8887.

Hachem-Vermette, C. (2020). Selected high-performance building envelopes. *Green Energy and Technology*, 67–100. https://doi.org/10.1007/978-3-030-47016-6_3/COVER

Han, K. H., & Zhang, J. (2020). Energy-saving building system integration with a smart and low-cost sensing/control network for sustainable and healthy living environments: Demonstration case study. *Energy and Buildings*, *214*, 109861. https://doi.org/10.1016/J.ENBUILD.2020.109861

He, B. jie, Yang, L., Ye, M., Mou, B., & Zhou, Y. (2014). Overview of rural building energy efficiency in China. *Energy Policy*, *69*, 385–396. https://doi.org/10.1016/J.ENPOL.2014.03.018

Hee, W. J., Alghoul, M. A., Bakhtyar, B., Elayeb, O., Shameri, M. A., Alrubaih, M. S., & Sopian, K. (2015). The role of window glazing on daylighting and energy saving in buildings. *Renewable and Sustainable Energy Reviews*, *42*, 323–343. https://doi.org/10.1016/J.RSER.2014.09.020

Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2023). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial Intelligence Review*, *56*(6), 4929–5021. https://doi.org/10.1007/S10462-022-10286-2/FIGURES/12

Hossain, M. U., Poon, C. S., Lo, I. M. C., & Cheng, J. C. P. (2016). Comparative environmental evaluation of aggregate production from recycled waste materials and virgin sources by LCA. *Resources, Conservation and Recycling*, *109*, 67–77. https://doi.org/10.1016/J.RESCONREC.2016.02.009

Hsu, W.-L., Tsai, H.-H., Yang, M.-L., Lai, S.-C., Ho, M.-C., & Shiau, Y.-C. (2021). Development of Smart Residential Environment Control System. *Sensors and Materials*, *33*(9), 3361–3377. https://doi.org/10.18494/SAM.2021.3420

Huang, L., Zhu, Y., Ouyang, Q., & Cao, B. (2012). A study on the effects of thermal, luminous, and acoustic environments on indoor environmental comfort in offices. *Building and Environment*, *49*(1), 304–309. https://doi.org/10.1016/J.BUILDENV.2011.07.022

Huizenga, C., Zagreus, L., Abbaszadeh, S., Lehrer, D., Goins, J., & Hoe, L. (2005). LEED post-occupancy evaluation: Taking responsibility for the occupants. *Proceedings of GreenBuild*.

Iqbal, M., Ma, J., Ahmad, N., Ullah, Z., & Hassan, A. (2022). Energy-Efficient supply chains in construction industry: An analysis of critical success factors using ISM-MICMAC approach. *Https://Doi.Org/10.1080/15435075.2022.2038609*, *20*(3), 265–283. https://doi.org/10.1080/15435075.2022.2038609

Issa, M. H., Rankin, J. H., Attalla, M., & Christian, A. J. (2011). Absenteeism, Performance and Occupant Satisfaction with the Indoor Environment of Green Toronto Schools. *Indoor and Built Environment*, *20*(5), 511–523. https://doi.org/10.1177/1420326X11409114

Jain, R., & Xu, W. (2021). HDSI: High dimensional selection with interactions algorithm on feature selection and testing. *PLOS ONE*, *16*(2), e0246159. https://doi.org/10.1371/JOURNAL.PONE.0246159

Jaiswal, J. K., & Samikannu, R. (2017). Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression. *Proceedings - 2nd World Congress on Computing and Communication Technologies, WCCCT 2017*, 65–68. https://doi.org/10.1109/WCCCT.2016.25

Jin, L., Liu, T., & Ma, J. (2021). Modeling Thermal Sensation Prediction Using Random Forest Classifier. *Communications in Computer and Information Science*, *1469 CCIS*, 552–561. https://doi.org/10.1007/978-981-16-7213-2_53/COVER

Kalhor, K., & Emaminejad, N. (2020). Qualitative and quantitative optimization of thermal insulation materials: Insights from the market and energy codes. *Journal of Building Engineering*, *30*, 101275. https://doi.org/10.1016/J.JOBE.2020.101275

Kapoor, N. R., Kumar, A., Meena, C. S., Kumar, A., Alam, T., Balam, N. B., & Ghosh, A. (2021). A Systematic Review on Indoor Environmental Quality in Naturally Ventilated School Classrooms: A Way Forward. *Advances in Civil Engineering*, *2021*. https://doi.org/10.1155/2021/8851685

Kaushik, A., Arif, M., Tumula, P., & Ebohon, O. J. (2020). Effect of thermal comfort on occupant productivity in office buildings: Response surface analysis. *Building and Environment*, *180*, 107021. https://doi.org/10.1016/J.BUILDENV.2020.107021

Kavitha, S., Varuna, S., & Ramya, R. (2017). A comparative analysis on linear regression and support vector regression. *Proceedings of 2016 Online International Conference on Green Engineering and Technologies, IC-GET 2016*. https://doi.org/10.1109/GET.2016.7916627

Kazmi, H., Keijsers, M., Mehmood, F., & Miller, C. (2022). Energy balances, thermal performance, and heat stress: Disentangling occupant behaviour and weather influences in a Dutch net-zero energy neighborhood. *Energy and Buildings*, *263*, 112020.

https://doi.org/10.1016/J.ENBUILD.2022.112020

Khalil, M. I., Jhanjhi, N. Z., Humayun, M., Sivanesan, S. K., Masud, M., & Hossain, M. S. (2021). Hybrid smart grid with sustainable energy efficient resources for smart cities. *Sustainable Energy Technologies and Assessments*, *46*, 101211. https://doi.org/10.1016/J.SETA.2021.101211

Khoshbakht, M., Gou, Z., Lu, Y., Xie, X., & Zhang, J. (2018). Are green buildings more satisfactory? A review of global evidence. *Habitat International*, *74*, 57–65. https://doi.org/10.1016/J.HABITATINT.2018.02.005

Khoshbakht, M., Gou, Z., Xie, X., He, B., & Darko, A. (2018). Green building occupant satisfaction: Evidence from the Australian higher education sector. *Sustainability (Switzerland)*, *10*(8), 1–21. https://doi.org/10.3390/su10082890

Khoury, K. B. (2019). Effective Communication Processes for Building Design, Construction, and Management. *Buildings 2019, Vol. 9, Page 112*, *9*(5), 112. https://doi.org/10.3390/BUILDINGS9050112

Kiki, G., Kouchadé, C., Houngan, A., Zannou-Tchoko, S. J., & André, P. (2020). Evaluation of thermal comfort in an office building in the humid tropical climate of Benin. *Building and Environment*, *185*, 107277. https://doi.org/10.1016/J.BUILDENV.2020.107277

Kim, J., Bauman, F., Raftery, P., Arens, E., Zhang, H., Fierro, G., Andersen, M., & Culler, D. (2019). Occupant comfort and behavior: High-resolution data from a 6-month field study of personal comfort systems with 37 real office workers. *Building and Environment*, *148*, 348–360. https://doi.org/10.1016/J.BUILDENV.2018.11.012

Kim, J., Zhou, Y., Schiavon, S., Raftery, P., & Brager, G. (2018). *Personal comfort models: predicting individuals' thermal preference using occupant heating and cooling behavior and machine learning*. *129*, 96–106. https://doi.org/10.1016/j.buildenv.2017.12.011

Kim, S.-K., Hwang, Y., Lee, Y. S., & Corser, W. (2015). Occupant Comfort and Satisfaction in Green Healthcare Environments: A Survey Study Focusing on Healthcare Staff. *Journal of Sustainable Development*, *8*(1). https://doi.org/10.5539/jsd.v8n1p156

Kim, Y. K., Abdou, Y., Abdou, A., & Altan, H. (2022). Indoor Environmental Quality Assessment and Occupant Satisfaction: A Post-Occupancy Evaluation of a UAE University Office Building. *Buildings 2022, Vol. 12, Page 986*, *12*(7), 986. https://doi.org/10.3390/BUILDINGS12070986

Kim, Y., Shin, Y., & Cho, H. (2021). Influencing factors on thermal comfort and biosignals of occupant-a review. *Journal of Mechanical Science and Technology 2021 35:9*, *35*(9), 4201–4224. https://doi.org/10.1007/S12206-021-0832-5

Kingma, S. (2018). New ways of working (NWW): work space and cultural change in virtualizing organizations. *Https://Doi.Org/10.1080/14759551.2018.1427747*, *25*(5), 383–406. https://doi.org/10.1080/14759551.2018.1427747

Kisilewicz, T. (2019). On the Role of External Walls in the Reduction of Energy Demand and the Mitigation of Human Thermal Discomfort. *Sustainability 2019, Vol. 11, Page 1061*, *11*(4), 1061. https://doi.org/10.3390/SU11041061

Kumar, A., & Suman, B. M. (2013). Experimental evaluation of insulation materials for walls and roofs and their impact on indoor thermal comfort under composite climate. *Building and Environment*, *59*, 635–643. https://doi.org/10.1016/J.BUILDENV.2012.09.023

Kumar, D., Alam, M., Zou, P. X. W., Sanjayan, J. G., & Memon, R. A. (2020). Comparative analysis of building insulation material properties and performance. *Renewable and Sustainable Energy Reviews*, *131*, 110038. https://doi.org/10.1016/J.RSER.2020.110038

Kumar, T. M. S., & Kurian, C. P. (2022). Real-time data based thermal comfort prediction leading to temperature setpoint control. *Journal of Ambient Intelligence and Humanized Computing*, *1*, 1–12. https://doi.org/10.1007/S12652-022-03754-8/TABLES/11

Kunič, R. (2017). Carbon footprint of thermal insulation materials in building envelopes. *Energy Efficiency*, *10*(6), 1511–1528. https://doi.org/10.1007/S12053-017-9536-1/METRICS

Leaman, A., & Bordass, B. (2007). Are users more tolerant of 'green' buildings? *Building Research & Information*, *35*(6), 662–673. https://doi.org/10.1080/09613210701529518

Leaman, A., Thomas, L., & Vandenberg, M. (2007). "Green" buildings: What Australian users are saying. *EcoLibrium(R)*.

Lee, S., & Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean Journal of Anesthesiology*, *71*(5), 353–360. https://doi.org/10.4097/KJA.D.18.00242

Lee, Y. S., & Kim, S.-K. (2008). Indoor Environmental Quality in LEED-Certified Buildings in the U.S. *Journal of Asian Architecture and Building Engineering*, *7*(2), 293–300. https://doi.org/10.3130/jaabe.7.293

Li, Q. Y., Han, J., & Lu, L. (2021). A Random Forest Classification Algorithm Based Personal Thermal Sensation Model for Personalized Conditioning System in Office Buildings. *The Computer Journal*, *64*(3), 500–508. https://doi.org/10.1093/COMJNL/BXAA165

Li, S., Zhang, X., Li, Y., Gao, W., Xiao, F., & Xu, Y. (2023). A comprehensive review of impact assessment of indoor thermal environment on work and cognitive performance - Combined physiological measurements and machine learning. *Journal of Building Engineering*, *71*, 106417. https://doi.org/10.1016/J.JOBE.2023.106417

Liang, H.-H., Chen, C.-P., Hwang, R.-L., Shih, W.-M., Lo, S.-C., & Liao, H.-Y. (2014). Satisfaction of occupants toward indoor environment quality of certified green office buildings in Taiwan. *Building and Environment*, *72*, 232–242. https://doi.org/10.1016/j.buildenv.2013.11.007

Lin, B., Liu, Y., Wang, Z., Pei, Z., & Davies, M. (2016). Measured energy use and indoor environment quality in green office buildings in China. *Energy and Buildings*, *129*, 9–18. https://doi.org/10.1016/j.enbuild.2016.07.057

Lin, I. Y. (2016). Effects of visual servicescape aesthetics comprehension and appreciation on consumer experience. *Journal of Services Marketing*, *30*(7), 692–712. https://doi.org/10.1108/JSM-08-2015-0258/FULL/XML

Liu, Y., Wang, Z., Lin, B., Hong, J., & Zhu, Y. (2018). Occupant satisfaction in Three-Star-certified office buildings based on comparative study using LEED and BREEAM. *Building and Environment*, *132*, 1–10. https://doi.org/10.1016/j.buildenv.2018.01.011

Magdalena Adelheid Januaviani, T., Gusriani, N., Joebaedi, K., & Talib Bon, A. (2019). *The LASSO (Least Absolute Shrinkage and Selection Operator) Method to Predict Indonesian Foreign Exchange Deposit Data*.

Mahmoud, A. R. (2022). Investigating the Impact of Different Glazing Types on the Energy Performance in Hot Arid Climate. *Journal of Advanced Engineering Trends*, *42*(1), 69–84. https://doi.org/10.21608/JAET.2021.96026.1121

Manju M, & Jacob P. (2022). A STUDY ON INDUSTRIAL DAYLIGHTING IN WARM HUMID CLIMATE. *JOURNAL OF ALGEBRAIC STATISTICS*, *13*(2), 1265–1273. https://doi.org/10.52783/JAS.V13I2.285

Manso, M., Teotónio, I., Silva, C. M., & Cruz, C. O. (2021). Green roof and green wall benefits and costs: A review of the quantitative evidence. *Renewable and Sustainable Energy Reviews*, *135*, 110111. https://doi.org/10.1016/J.RSER.2020.110111

Mantesi, E., Chmutina, K., & Goodier, C. (2022). The office of the future: Operational energy consumption in the post-pandemic era. *Energy Research & Social Science*, *87*, 102472. https://doi.org/10.1016/J.ERSS.2021.102472

Mao, F., Zhou, X., & Song, Y. (2019). Environmental and Human Data-Driven Model Based on Machine Learning for Prediction of Human Comfort. *IEEE Access*, *7*, 132909–132922. https://doi.org/10.1109/ACCESS.2019.2940910

McArthur, J. J., & Powell, C. (2020). Health and wellness in commercial buildings: Systematic review of sustainable building rating systems and alignment with contemporary research. *Building and Environment*, *171*, 106635. https://doi.org/10.1016/J.BUILDENV.2019.106635

Meena, R. K., Raj, R., & Anbukumar, S. (2022). Effect of wind load on irregular shape tall buildings having different corner configuration. *Sadhana - Academy Proceedings in Engineering Sciences*, *47*(3), 1–17. https://doi.org/10.1007/S12046-022-01895-2/METRICS

Menadue, V., Soebarto, V., & Williamson, T. (2014). Perceived and actual thermal conditions: case studies of green-rated and conventional office buildings in the City of Adelaide. *Architectural Science Review*, *57*(4), 303–319. https://doi.org/10.1080/00038628.2014.986433

Miller, A., Panneerselvam, J., & Liu, L. (2022). A review of regression and classification techniques for analysis of common and rare variants and gene-environmental factors. *Neurocomputing*, *489*, 466–485. https://doi.org/10.1016/J.NEUCOM.2021.08.150

Minoli, D., Sohraby, K., & Occhiogrosso, B. (2017). IoT Considerations, Requirements, and Architectures for Smart Buildings-Energy Optimization and Next-Generation Building Management Systems. *IEEE Internet of Things Journal*, *4*(1), 269–283. https://doi.org/10.1109/JIOT.2017.2647881

Mirrahimi, S., Mohamed, M. F., Haw, L. C., Ibrahim, N. L. N., Yusoff, W. F. M., & Aflaki, A. (2016). The effect of building envelope on the thermal comfort and energy saving for high-rise buildings in hot–humid climate. *Renewable and Sustainable Energy Reviews*, *53*, 1508–1519. https://doi.org/10.1016/J.RSER.2015.09.055

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-Minute City": Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities 2021, Vol. 4, Pages 93-111*, *4*(1), 93–111. https://doi.org/10.3390/SMARTCITIES4010006

Mourato, J., Pinto Ferreira, L., Sá, J. C., Silva, F. J. G., Dieguez, T., & Tjahjono, B. (2020).

Improving internal logistics of a bus manufacturing using the lean techniques. *International Journal of Productivity and Performance Management*, *70*(7), 1930–1951. https://doi.org/10.1108/IJPPM-06-2020-0327/FULL/XML

Mui, K. W., & Wong, L. T. (2011). Acceptable Illumination Levels for Office Occupants. *Http://Dx.Doi.Org/10.3763/Asre.2006.4915*, *49*(2), 116–119. https://doi.org/10.3763/ASRE.2006.4915

Nag, P. K. (2019). *Assessing IEQ Performance in Buildings*. 311–340. https://doi.org/10.1007/978-981-13-2577-9_11

Neelamegam, S., & Ramraj, E. (2013). Classification algorithm in Data mining: An Overview. *International Journal of P2P Network Trends and Technology*, *4*. http://www.ijpttjournal.org

Neofytou, H., Sarafidis, Y., Gkonis, N., Mirasgedis, S., & Askounis, D. (2020). Energy Efficiency contribution to sustainable development: A multi-criteria approach in Greece. *Https://Doi.Org/10.1080/15567249.2020.1849449*, *15*(10–12), 572–604. https://doi.org/10.1080/15567249.2020.1849449

Newsham, G. R., Birt, B. J., Arsenault, C., Thompson, A. J. L., Veitch, J. A., Mancini, S., Galasiu, A. D., Gover, B. N., Macdonald, I. A., & Burns, G. J. (2013). Do 'green' buildings have better indoor environments? New evidence. *Building Research & Information*, *41*(4), 415–434. https://doi.org/10.1080/09613218.2013.789951

Ng, A. W. (2018). From sustainability accounting to a green financing system: Institutional legitimacy and market heterogeneity in a global financial centre. *Journal of Cleaner Production*, *195*, 585–592. https://doi.org/10.1016/J.JCLEPRO.2018.05.250

Ngarambe, J., Adilkhanova, I., Uwiragiye, B., & Yun, G. Y. (2022). A review on the current usage of machine learning tools for daylighting design and control. *Building and Environment*, *223*, 109507. https://doi.org/10.1016/J.BUILDENV.2022.109507

Nugroho, B. C. P., Chan, J. F., Vananda, V., Lucky, H., & Suhartono, D. (2022). Detecting Phishing Websites with Non-Parametric Machine Learning. *2022 International Conference on Informatics Electrical and Electronics, ICIEE 2022 - Proceedings*. https://doi.org/10.1109/ICIEE55596.2022.10009939

Ohene, E., Hsu, S. C., & Chan, A. P. C. (2022). Feasibility and retrofit guidelines towards net-zero energy buildings in tropical climates: A case of Ghana. *Energy and Buildings*, *269*, 112252. https://doi.org/10.1016/J.ENBUILD.2022.112252

Otchere, D. A., Ganat, T. O. A., Ojero, J. O., Tackie-Otoo, B. N., & Taki, M. Y. (2022). Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. *Journal of Petroleum Science and Engineering*, *208*, 109244. https://doi.org/10.1016/J.PETROL.2021.109244

Paramati, S. R., Shahzad, U., & Doğan, B. (2022). The role of environmental technology for energy demand and energy efficiency: Evidence from OECD countries. *Renewable and Sustainable Energy Reviews*, *153*, 111735. https://doi.org/10.1016/J.RSER.2021.111735

Parkinson, T., Schiavon, S., Kim, J., & Betti, G. (2023). Common sources of occupant dissatisfaction with workspace environments in 600 office buildings. *Buildings and Cities*, *4*(1), 17–35. https://doi.org/10.5334/BC.274

Patil, A. R., & Kim, S. (2020). Combination of Ensembles of Regularized Regression Models with Resampling-Based Lasso Feature Selection in High Dimensional Data. *Mathematics 2020, Vol. 8, Page 110*, *8*(1), 110. https://doi.org/10.3390/MATH8010110

Paul, W. L., & Taylor, P. A. (2008). A comparison of occupant comfort and satisfaction between a green building and a conventional building. *Building and Environment*, *43*(11), 1858–1870. https://doi.org/10.1016/j.buildenv.2007.11.006

Pei, Z., Lin, B., Liu, Y., & Zhu, Y. (2015). Comparative study on the indoor environment quality of green office buildings in China with a long-term field measurement and investigation. *Building and Environment*, *84*, 80–88. https://doi.org/10.1016/j.buildenv.2014.10.015

Peng, B., & Hsieh, S. J. (2017). Data-Driven Thermal Comfort Prediction With Support Vector Machine. *ASME 2017 12th International Manufacturing Science and Engineering Conference, MSEC 2017 Collocated with the JSME/ASME 2017 6th International Conference on Materials and Processing*, *3*. https://doi.org/10.1115/MSEC2017-3003

Pereira, J., Teixeira, H., Gomes, M. da G., & Rodrigues, A. M. (2022). Performance of Solar Control Films on Building Glazing: A Literature Review. *Applied Sciences 2022, Vol. 12, Page 5923*, *12*(12), 5923. https://doi.org/10.3390/APP12125923

Pode, R. (2020). Organic light emitting diode devices: An energy efficient solid state lighting for applications. *Renewable and Sustainable Energy Reviews*, *133*, 110043. https://doi.org/10.1016/J.RSER.2020.110043

Pragati, S., Shanthi Priya, R., Pradeepa, C., & Senthil, R. (2023). Simulation of the Energy Performance of a Building with Green Roofs and Green Walls in a Tropical Climate. *Sustainability 2023, Vol. 15, Page 2006*, *15*(3), 2006. https://doi.org/10.3390/SU15032006

PRISMA. (2023). http://www.prisma-statement.org/

Radić, M., Dodig, M. B., & Auer, T. (2019). Green Facades and Living Walls—A Review Establishing the Classification of Construction Types and Mapping the Benefits. *Sustainability 2019, Vol. 11, Page 4579*, *11*(17), 4579. https://doi.org/10.3390/SU11174579

Ragheb, A., El-Shimy, H., & Ragheb, G. (2016). Green Architecture: A Concept of Sustainability. *Procedia - Social and Behavioral Sciences*, *216*, 778–787. https://doi.org/10.1016/J.SBSPRO.2015.12.075

Rahman, A., Farrok, O., & Haque, M. M. (2022). Environmental impact of renewable energy source based electrical power plants: Solar, wind, hydroelectric, biomass, geothermal, tidal, ocean, and osmotic. *Renewable and Sustainable Energy Reviews*, *161*, 112279. https://doi.org/10.1016/J.RSER.2022.112279

Ravindu, S., Rameezdeen, R., Zuo, J., Zhou, Z., & Chandratilake, R. (2015). Indoor environment quality of green buildings: Case study of an LEED platinum certified factory in a warm humid tropical climate. *Building and Environment*, *84*, 105–113. https://doi.org/10.1016/j.buildenv.2014.11.001

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., … Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/J.GLOENVCHA.2016.05.009

Ridwana, I., Nassif, N., & Choi, W. (2020). Modeling of Building Energy Consumption by Integrating Regression Analysis and Artificial Neural Network with Data Classification. *Buildings 2020, Vol. 10, Page 198*, *10*(11), 198. https://doi.org/10.3390/BUILDINGS10110198

Rienow, A., & Goetzke, R. (2015). Supporting SLEUTH – Enhancing a cellular automaton with support vector machines for urban growth modeling. *Computers, Environment and Urban Systems*, *49*, 66–81. https://doi.org/10.1016/J.COMPENVURBSYS.2014.05.001

Roderick, Y. ., McEwan, D. ., Wheatley, C. ., & Alonso, C. (2009). Comparison of energy performance assessment between LEED, BREEAM and Green Star. *In Proceedings of the Eleventh International IBPSA Conference*, 27–30.

Rohde, L., Larsen, T. S., Jensen, R. L., & Larsen, O. K. (2019). Framing holistic indoor environment: Definitions of comfort, health and well-being. *Https://Doi.Org/10.1177/1420326X19875795*, *29*(8), 1118–1136. https://doi.org/10.1177/1420326X19875795

Roque, E., & Santos, P. (2017). The Effectiveness of Thermal Insulation in Lightweight Steel-Framed Walls with Respect to Its Position. *Buildings 2017, Vol. 7, Page 13*, *7*(1), 13. https://doi.org/10.3390/BUILDINGS7010013

Rosenow, J., Kern, F., & Rogge, K. (2017). The need for comprehensive and well targeted instrument mixes to stimulate energy transitions: The case of energy efficiency policy. *Energy Research & Social Science*, *33*, 95–104. https://doi.org/10.1016/J.ERSS.2017.09.013

Roskams, M., & Haynes, B. (2019). Predictive analytics in facilities management: A pilot study for predicting environmental comfort using wireless sensors. *Journal of Facilities Management*, *17*(4), 356–370. https://doi.org/10.1108/JFM-03-2019-0008/FULL/XML

Roumi, S., Zhang, F., Stewart, R. A., & Santamouris, M. (2023). Weighting of indoor environment quality parameters for occupant satisfaction and energy efficiency. *Building and Environment*, *228*, 109898. https://doi.org/10.1016/J.BUILDENV.2022.109898

Rumiantcev, B., Zhukov, A., Zelenshikov, D., Chkunin, A., Ivanov, K., & Sazonova, Y. (2016). Insulation systems of the building construtions. *MATEC Web of Conferences*, *86*, 04027. https://doi.org/10.1051/MATECCONF/20168604027

Saadatian, S., Freire, F., & Simões, N. (2021). Embodied impacts of window systems: A comparative assessment of framing and glazing alternatives. *Journal of Building Engineering*, *35*, 102042. https://doi.org/10.1016/J.JOBE.2020.102042

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), e1249. https://doi.org/10.1002/WIDM.1249

Salamone, F., Bellazzi, A., Belussi, L., Damato, G., Danza, L., Dell'aquila, F., Ghellere, M., Megale, V., Meroni, I., & Vitaletti, W. (2020). Evaluation of the Visual Stimuli on Personal Thermal Comfort Perception in Real and Virtual Environments Using Machine Learning Approaches. *Sensors 2020, Vol. 20, Page 1627*, *20*(6), 1627. https://doi.org/10.3390/S20061627

Samaraweera, B. G. T. N., Gunawardhana, W. H. T., & Lankan Journal, S. (2019). The authenticity of Sustainable Practices in Sri Lankan Construction Industry. *Sri Lanka Journal of Real Estate*, *15*. http://journals.sjp.ac.lk/index.php/SLJRE/article/view/4717

Schiavoni, S., D'Alessandro, F., Bianchi, F., & Asdrubali, F. (2016). Insulation materials for the building sector: A review and comparative analysis. *Renewable and Sustainable Energy Reviews*, *62*, 988–1011. https://doi.org/10.1016/J.RSER.2016.05.045

Sediso, B. G., & Lee, M. S. (2016). Indoor environmental quality in Korean green building certification criteria—certified office buildings—occupant satisfaction and performance. *Science and Technology for the Built Environment*, *22*(5), 606–618. https://doi.org/10.1080/23744731.2016.1176849

Septina Saraswati, R., Arie Wibawa, B., Eko Saputra, B., Prihatiningrum, A., Anom Ramawangsa, P., Seftyarizki -, D., Kusumawardani, L., Ramadhan, T., & Maknun -, J. (2021). *IOP Conference Series: Earth and Environmental Science The Role of Artificial Lighting in Architectural Design: A Literature Review You may also like Optimization of Natural and Artificial Lighting System in UPGRIS Lecturer's Workspace using Dialux Evo Im*. https://doi.org/10.1088/1755-1315/665/1/012008

Sethi, J. K., & Mittal, M. (2021). An efficient correlation based adaptive LASSO regression method for air quality index prediction. *Earth Science Informatics*, *14*(4), 1777–1786. https://doi.org/10.1007/S12145-021-00618-1/METRICS

Seyedzadeh, S., Pour Rahimian, F., Oliver, S., Rodriguez, S., & Glesk, I. (2020). Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making. *Applied Energy*, *279*, 115908. https://doi.org/10.1016/J.APENERGY.2020.115908

Shaikh, P. H., Nor, N. B. M., Sahito, A. A., Nallagownden, P., Elamvazuthi, I., & Shaikh, M. S. (2017). Building energy for sustainable development in Malaysia: A review. *Renewable and Sustainable Energy Reviews*, *75*, 1392–1403. https://doi.org/10.1016/J.RSER.2016.11.128

Shao, Y., Ma, Z., Wang, J., & Bi, J. (2020). Estimating daily ground-level PM2.5 in China with random-forest-based spatiotemporal kriging. *Science of The Total Environment*, *740*, 139761. https://doi.org/10.1016/J.SCITOTENV.2020.139761

Sheth, K. N. (2017). *Water efficient technologies in green buildings*. https://www.researchgate.net/publication/316582554

Shi, T., Yang, W., Qi, A., Li, P., & Qiao, J. (2023). LASSO and attention-TCN: a concurrent method for indoor particulate matter prediction. *Applied Intelligence*, 1–15. https://doi.org/10.1007/S10489-023-04507-6/TABLES/5

Sibyan, H., Svajlenka, J., Hermawan, H., Faqih, N., & Arrizqi, A. N. (2022). Thermal Comfort Prediction Accuracy with Machine Learning between Regression Analysis and Naïve Bayes Classifier. *Sustainability 2022, Vol. 14, Page 15663*, *14*(23), 15663. https://doi.org/10.3390/SU142315663

Sikram, T., Ichinose, M., & Sasaki, R. (2020). Assessment of Thermal Comfort and Building-Related Symptoms in Air-Conditioned Offices in Tropical Regions: A Case Study in Singapore and Thailand. *Frontiers in Built Environment*, *6*, 567787. https://doi.org/10.3389/FBUIL.2020.567787/BIBTEX

Singh, A., Syal, M., Grady, S. C., & Korkmaz, S. (2010). Effects of green buildings on employee health and productivity. *American Journal of Public Health*, *100*(9), 1665–1668. https://doi.org/10.2105/AJPH.2009.180687

Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, *27*(2), 130. https://doi.org/10.11919/J.ISSN.1002-0829.215044

Su, H., Han, G., Li, L., & Qin, H. (2021). The impact of macro-scale urban form on land surface temperature: An empirical study based on climate zone, urban size and industrial structure in China. *Sustainable Cities and Society*, *74*, 103217. https://doi.org/10.1016/J.SCS.2021.103217

Suk, J. Y. (2019). Luminance and vertical eye illuminance thresholds for occupants' visual comfort in daylit office environments. *Building and Environment*, *148*, 107–115. https://doi.org/10.1016/J.BUILDENV.2018.10.058

Sun, D., Xu, J., Wen, H., & Wang, Y. (2020). An Optimized Random Forest Model and Its Generalization Ability in Landslide Susceptibility Mapping: Application in Two Areas of Three Gorges Reservoir, China. *Journal of Earth Science*, *31*(6), 1068–1086. https://doi.org/10.1007/S12583-020-1072-9/METRICS

Sutherland, V. J., & Cooper, C. L. (2022). Stress in the work environment. *Human Stress and the Environment*, 131–159. https://doi.org/10.1201/9780367810641-6/STRESS-WORK-ENVIRONMENT-VALERIE-SUTHERLAND-CARY-COOPER

Taheri, S., Hosseini, P., & Razban, A. (2022). Model predictive control of heating, ventilation, and air conditioning (HVAC) systems: A state-of-the-art review. *Journal of Building Engineering*, *60*, 105067. https://doi.org/10.1016/J.JOBE.2022.105067

Taleb, H. M., & Antony, A. G. (2020). Assessing different glazing to achieve better lighting performance of office buildings in the United Arab Emirates (UAE). *Journal of Building Engineering*, *28*, 101034. https://doi.org/10.1016/J.JOBE.2019.101034

Tan, W. C., Zhang, M., Elmeleegy, H., & Srivastava, D. (2017). Reverse engineering aggregation queries. *Proceedings of the VLDB Endowment*, *10*(11), 1394–1405. https://doi.org/10.14778/3137628.3137648

Tang, H., Liu, X., Geng, Y., Lin, B., & Ding, Y. (2022). Assessing the perception of overall indoor environmental quality: Model validation and interpretation. *Energy and Buildings*, *259*, 111870. https://doi.org/10.1016/J.ENBUILD.2022.111870

Tanyer, A. M., Tavukcuoglu, A., & Bekboliev, M. (2018). Assessing the airtightness performance of container houses in relation to its effect on energy efficiency. *Building and Environment*, *134*, 59–73. https://doi.org/10.1016/J.BUILDENV.2018.02.026

Tariku, F., Shang, Y., & Molleti, S. (2023). Thermal performance of flat roof insulation materials: A review of temperature, moisture and aging effects. *Journal of Building Engineering*, *76*, 107142. https://doi.org/10.1016/J.JOBE.2023.107142

Tekce, I., Ergen, E., & Artan, D. (2020). Structural Equation Model of Occupant Satisfaction for Evaluating the Performance of Office Buildings. *Arabian Journal for Science and Engineering*, *45*(10), 8759–8784. https://doi.org/10.1007/s13369-020-04804-z

Tokazhanov, G., Galiyev, O., Lukyanenko, A., Nauyryzbay, A., Ismagulov, R., Durdyev, S., Turkyilmaz, A., & Karaca, F. (2022). Circularity assessment tool development for construction projects in emerging economies. *Journal of Cleaner Production*, *362*, 132293. https://doi.org/10.1016/J.JCLEPRO.2022.132293

Tomasella, M., De Nardi, E., Petruzzellis, F., Andri, S., Castello, M., & Nardini, A. (2022). Green roof irrigation management based on substrate water potential assures water saving without affecting plant physiological performance. *Ecohydrology*, *15*(4), e2428. https://doi.org/10.1002/ECO.2428

Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, *102*(6), e03336. https://doi.org/10.1002/ECY.3336

Troncoso-Pastoriza, F., Martínez-Comesaña, M., Ogando-Martínez, A., López-Gómez, J., Eguía-Oller, P., & Febrero-Garrido, L. (2022). IoT-based platform for automated IEQ spatio-temporal analysis in buildings using machine learning techniques. *Automation in Construction*, *139*, 104261. https://doi.org/10.1016/J.AUTCON.2022.104261

Turner, C. (2006). *LEED building performance in the Cascadia region: a post occupancy evaluation report*.

Uddin, M. N., Wei, H. H., Chi, H. L., Ni, M., & Elumalai, P. (2021). Building information modeling (BIM) incorporated green building analysis: an application of local construction materials and sustainable practice in the built environment. *Journal of Building Pathology and Rehabilitation*, *6*(1), 1–25. https://doi.org/10.1007/S41024-021-00106-5/METRICS

Ullah, H. M. K., Lejeune, J., Cayla, A., Monceaux, M., Campagne, C., & Devaux, É. (2021). A review of noteworthy/major innovations in wearable clothing for thermal and moisture management from material to fabric structure. *Https://Doi.Org/10.1177/00405175211027799*, *92*(17–18), 3351–3386. https://doi.org/10.1177/00405175211027799

US Green Building Council. (2023). *Mission and vision | U.S. Green Building Council*. https://www.usgbc.org/about/mission-vision

Van Dick, R., Christ, O., Stellmacher, J., Wagner, U., Ahlswede, O., Grubba, C., Hauptmeier, M., Hohfeld, C., Moltzen, K., & Tissington, P. A. (2004). Should I Stay or Should I Go? Explaining Turnover Intentions with Organizational Identification and Job Satisfaction*. *British Journal of Management*, *15*(4), 351–360. https://doi.org/10.1111/j.1467-8551.2004.00424.x

Veitch, J. A., Charles, K. E., Farley, K. M. J., & Newsham, G. R. (2007). A model of satisfaction with open-plan office conditions: COPE field findings. *Journal of Environmental Psychology*, *27*(3), 177–189. https://doi.org/10.1016/j.jenvp.2007.04.002

Venkatesh, B., & Anuradha, J. (2019). A Hybrid Feature Selection Approach for Handling a High-Dimensional Data. *Lecture Notes in Networks and Systems*, *74*, 365–373. https://doi.org/10.1007/978-981-13-7082-3_42/COVER

Verbeke, S., & Audenaert, A. (2018). Thermal inertia in buildings: A review of impacts across climate and building use. *Renewable and Sustainable Energy Reviews*, *82*, 2300–2318. https://doi.org/10.1016/J.RSER.2017.08.083

Verma, R., Kumar, S., Rakshit, D., & Premachandran, B. (2023). Design and Optimization of Energy Consumption for a Low-Rise Building With Seasonal Variations Under Composite Climate of India. *Journal of Solar Energy Engineering*, *145*(1).

https://doi.org/10.1115/1.4054831

Voordt, T. van der, & Jensen, P. A. (2023). The impact of healthy workplaces on employee satisfaction, productivity and costs. *Journal of Corporate Real Estate*, *25*(1), 29–49. https://doi.org/10.1108/JCRE-03-2021-0012/FULL/PDF

Walker, W., & Kimberly, V. (2020). *Blueprint for Greening Affordable Housing, Revised Edition - Walker Wells, Kimberly Vermeer - Google Books*. https://books.google.lk/books?hl=en&lr=&id=rDfnDwAAQBAJ&oi=fnd&pg=PP1&dq=Well-placed+windows,+skylights,+and+light+wells+can+optimise+daylight+penetration,+reducing+reliance+on+artificial+lighting+and+fostering+a+connection+with+the+outdoors.&ots=tKez8YvW

Wang, H., Chiang, P. C., Cai, Y., Li, C., Wang, X., Chen, T. L., Wei, S., & Huang, Q. (2018). Application of Wall and Insulation Materials on Green Building: A Review. *Sustainability 2018, Vol. 10, Page 3331*, *10*(9), 3331. https://doi.org/10.3390/SU10093331

Warnasekara, J., Agampodi, S., & Rupika Abeynayake, R. (2021). Time series models for prediction of leptospirosis in different climate zones in Sri Lanka. *PLOS ONE*, *16*(5), e0248032. https://doi.org/10.1371/JOURNAL.PONE.0248032

Weerasinghe, A. S., Ramachandra, T., & Rotimi, J. O. B. (2021). Comparative life-cycle cost (LCC) study of green and traditional industrial buildings in Sri Lanka. *Energy and Buildings*, *234*, 110732. https://doi.org/10.1016/J.ENBUILD.2021.110732

Wei, J. (2022). The adoption of repeated measurement of variance analysis and Shapiro—Wilk test. *Frontiers of Medicine*, *16*(4), 659–660. https://doi.org/10.1007/S11684-021-0908-8/METRICS

Wei, W., & Skye, H. M. (2021). Residential net-zero energy buildings: Review and perspective. *Renewable and Sustainable Energy Reviews*, *142*, 110859. https://doi.org/10.1016/J.RSER.2021.110859

WELL - International WELL Building Institute | IWBI. (n.d.). Retrieved July 16, 2023, from https://www.wellcertified.com/

What's National Green Building Standard (NGBS). (2023). https://www.ngbs.com/the-ngbs-green-promise

Whitnall, C., Oswald, E., & Mather, L. (2011). An exploration of the Kolmogorov-Smirnov test as a competitor to mutual information analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7079 LNCS*, 234–251. https://doi.org/10.1007/978-3-642-27257-8_15/COVER

Wijewardane, M. A., Sudasinghe, S., & ... (2018). Experimental Investigation of Visual Comfort Requirement in Garment Factories and Identify the Cost Saving Opportunities. *... Journal of Urban ...*, 1–12. https://publications.waset.org/10009163/experimental-investigation-of-visual-comfort-requirement-in-garment-factories-and-identify-the-cost-saving-opportunities

Woo, E. H. C., White, P., & Lai, C. W. K. (2015). Ergonomics standards and guidelines for computer workstation design and the impact on users' health – a review. *Http://Dx.Doi.Org/10.1080/00140139.2015.1076528*, *59*(3), 464–475.

https://doi.org/10.1080/00140139.2015.1076528

Wu, C., Zhang, R., Kotagiri, R., & Bouvry, P. (2023). Strategic Decisions: Survey, Taxonomy, and Future Directions from Artificial Intelligence Perspective. *ACM Computing Surveys*, *55*(12). https://doi.org/10.1145/3571807

Yan, D., O'Brien, W., Hong, T., Feng, X., Burak Gunay, H., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, *107*, 264–278. https://doi.org/10.1016/J.ENBUILD.2015.08.032

Yan, X. (2023). Research on Financial Field Integrating Artificial Intelligence: Application Basis, Case Analysis, and SVR Model-Based Overnight. *Https://Doi.Org/10.1080/08839514.2023.2222258*, *37*(1). https://doi.org/10.1080/08839514.2023.2222258

Yun, G. Y., Kong, H. J., Kim, H., & Kim, J. T. (2012). A field survey of visual comfort and lighting energy consumption in open plan offices. *Energy and Buildings*, *46*, 146–151. https://doi.org/10.1016/J.ENBUILD.2011.10.035

Zagreus, L., Huizenga, C., Arens, E., & Lehrer, D. (2004). Listening to the occupants: A Web-based indoor environmental quality survey. *Indoor Air, Supplement*, *14*(8), 65–74. https://doi.org/10.1111/j.1600-0668.2004.00301.x

Zakaria Salem, F., Sadek, M., El-Azab, H., Zakaria, F., Radwan, M. A., Sadek, M. A., & Elazab, H. A. (2018). Insulating material based on shredded used tires and inexpensive polymers for different roofs. *Article in International Journal of Engineering and Technology*, *7*(4), 1983–1988. https://doi.org/10.14419/ijet.v7i4.14081

Zhang, C., Pomianowski, M., Heiselberg, P. K., & Yu, T. (2020). A review of integrated radiant heating/cooling with ventilation systems- Thermal comfort and indoor air quality. *Energy and Buildings*, *223*, 110094. https://doi.org/10.1016/J.ENBUILD.2020.110094

Zhang, D., & Tu, Y. (2021). Green building, pro-environmental behavior and well-being: Evidence from Singapore. *Cities*, *108*, 102980. https://doi.org/10.1016/J.CITIES.2020.102980

Zhang, F., Liu, S., Hu, W., & Yadav, M. (2022). Editorial: Effects of indoor environmental quality on human performance and productivity. *Frontiers in Built Environment*, *8*, 1095443. https://doi.org/10.3389/FBUIL.2022.1095443/BIBTEX

Zhang, H., Yang, D., Tam, V. W. Y., Tao, Y., Zhang, G., Setunge, S., & Shi, L. (2021). A critical review of combined natural ventilation techniques in sustainable buildings. *Renewable and Sustainable Energy Reviews*, *141*, 110795. https://doi.org/10.1016/J.RSER.2021.110795

Zhang, H., Yang, X., Tu, R., Huang, J., & Li, Y. (2022). Thermal Comfort Modeling of Office Buildings Based on Improved Random Forest Algorithm. *Proceedings of 2022 IEEE 11th Data Driven Control and Learning Systems Conference, DDCLS 2022*, 1369–1376. https://doi.org/10.1109/DDCLS55054.2022.9858536

Zhang, Y., & Altan, H. (2011). A comparison of the occupant comfort in a conventional high-rise office block and a contemporary environmentally-concerned building. *Building and Environment*, *46*(2), 535–545. https://doi.org/10.1016/j.buildenv.2010.09.001

Zhang, Y., Wang, H., Gao, W., Wang, F., Zhou, N., Kammen, D. M., & Ying, X. (2019). A survey of the status and challenges of green building development in various countries. *Sustainability (Switzerland)*, *11*(19), 1–29. https://doi.org/10.3390/su11195385

Zhao, D. X., He, B. J., Johnson, C., & Mou, B. (2015). Social problems of green buildings: From the humanistic needs to social acceptance. *Renewable and Sustainable Energy Reviews*, *51*, 1594–1609. https://doi.org/10.1016/J.RSER.2015.07.072

Zhao, J., & Du, Y. (2020). Multi-objective optimization design for windows and shading configuration considering energy consumption and thermal comfort: A case study for office building in different climatic regions of China. *Solar Energy*, *206*, 997–1017. https://doi.org/10.1016/J.SOLENER.2020.05.090

Zhuang, C., & Wang, S. (2020). Uncertainty-based robust optimal design of cleanroom air-conditioning systems considering life-cycle performance. *Https://Doi.Org/10.1177/1420326X19899442*, *29*(9), 1214–1226. https://doi.org/10.1177/1420326X19899442

Zolfaghari, Z., & Jones, J. (2022). A multi-variable building energy optimization: assessing the role of energy efficient lighting technology in changing the optimal window-to-wall ratio in an office building. *Taylor & Francis Online*, *41*(11), 1819–1835. https://doi.org/10.1080/14786451.2022.2118276

APPENDICES

APPENDIX A

Employers' satisfaction of energy-efficient applications in buildings

The objective of this survey is to analyze "Employers' satisfaction of energy-efficient applications in green buildings/factories" - Ph.D.: Major component of Research

This survey is highly confidential and intended solely for acedamic purposes. All data is stored in a password protected electronic format. To help protect your confidentiality, the survey will not contain information that will personally identify you.

Please provide your honest opinion to make this survey a success.

There are 29 questions in this survey.

A note on privacy
This survey is anonymous.
The record of your survey responses does not contain any identifying information about you, unless a specific survey question explicitly asked for it. If you used an identifying token to access this survey, please rest assured that this token will not be stored together with your responses. It is managed in a separate database and will only be updated to indicate whether you did (or did not) complete this survey. There is no way of matching identification tokens with survey responses.

✱ Where is your organization/ company located?

✱ What is your department?

Please choose...

✱ How would you describe the work you do (designation)?

Please choose...

✱ What is your gender?

♀ Female	♂ Male

✱ What is your age category?

- 18-20
- 21-30
- 31-40
- 41-50
- Over 50

✱ Your hometown located at

Please choose...

✱ Have you changed your residence from your hometown?

Yes	No

✱ Which of the following best describes the area of your hometown?

Please choose...

✱ Is your organization/ factory certified as "Green Building" (GBCSL/LEED)?

- ○ Yes
- ○ No
- ○ Don't know

✱ What is your level of education?

- ◯ Primary education
- ◯ Higher education
- ◯ Bachelor Degree
- ◯ Postgraduate

✱ Number of hours per day you normally spend at your desk/ workstation?

Only numbers may be entered in this field.

✱ How long have you worked in this building (in months)?

Only numbers may be entered in this field.

✱ How many people do you share your workstation with? (not including yourself)?

- ◯ less than 2
- ◯ 3-5

- 6-10
- 10-20
- 20-30
- above 30

✱ Rate the satisfaction level of surrounding of your workstation accordingly

	1	2	3	4	5
Attractiveness of the room	1	2	3	4	5
Ventilation in the room	1	2	3	4	5
Thermal comfort in the room	1	2	3	4	5
Noise level of the room	1	2	3	4	5
Dryness of the room	1	2	3	4	5
Sunlight glare of the room	1	2	3	4	5
Distance between your seat and the window	1	2	3	4	5
Fresh air inside the room	1	2	3	4	5
Space you have to interact with the co-workers	1	2	3	4	5

Please rate your answer 1 to 5 (1=very unsatisfied 2=unsatisfied 3=neutral 4=satisfied 5=very satisfied)

✱ Is there a blind wall (wall without an opening) in your room?

Yes	No

✱
Is there a window in your room?

Yes	No

✱ Does sunlight enter to your room?

Yes	No

✱ Do you use A/C at your workspace?

Yes	No

✱ Do you feel too cold or too hot in your room someday?

- ○ Yes,Too hot

- Yes, Too cold
- Yes, somedays too hot, some days too cold
- I feel average temperature everyday

✱ Rate your opinion of the below

	1	2	3	4	5
Dust level in your room	1	2	3	4	5
Comfort of your seating arrangement	1	2	3	4	5
Ease of interaction with the office mates	1	2	3	4	5
Cleanliness of your office space	1	2	3	4	5

(1=very low 2=low =3=neutral 4=high 4=very high)

✱ Rate your comfort in these in the work place?

	1	2	3	4	5
Noise level	1	2	3	4	5
Lighting level	1	2	3	4	5
Glare level in the room	1	2	3	4	5
Distance of your seat from the window	1	2	3	4	5

	1	2	3	4	5
Room temperature	1	2	3	4	5
Ventilation inside the room	1	2	3	4	5
Freshness of your room	1	2	3	4	5
Odour in the room	1	2	3	4	5
State of your health when in the room	1	2	3	4	5
Colors of the room	1	2	3	4	5
The control you have over your local environment (can change the lights or tem-pertature according to your comfort)	1	2	3	4	5
Amount of the working space you have in your room	1	2	3	4	5
Immediate colleagues (their conversation)	1	2	3	4	5
Management	1	2	3	4	5
The outward appearance in your building in general	1	2	3	4	5

(1 = Very unccomfortable, 2= uncomfortable,3=neutral, 4= comfortable, 5= very comfortable)

✱ What is the distance between your work desk and the nearest window?

Please choose...

✱ Satisfaction level of your workplace in general

1 2 3 4 5

Please rate 1 for highly unsatisfied and 5 for highly satisfied

Submit

APPENDIX B

iaq-copy1

July 17, 2023

[370]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import norm

#Converting all Non-Numerical Columns to Numerical
from sklearn.preprocessing import LabelEncoder

from sklearn.metrics import accuracy_score, confusion_matrix, r2_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.tree import DecisionTreeClassifier, export_graphviz
from sklearn.model_selection import train_test_split
import warnings
warnings.filterwarnings('ignore')

%matplotlib inline
```

1 Load the Thermal dataset

[371]:
```
data1 = pd.read_excel("Data.xlsx")
```

[372]:
```
data1.shape
```

[372]: (1091, 29)

[373]:
```
# duplicate the dataset

data1_copy = data1.copy()
data1_copy.shape
```

[373]: (1091, 29)

[374]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location', 'Gender',
       'Employee Home Town', 'Age', 'Working Hours', 'Is there a blind wall ',
       'Distance between your work desk and the nearest window?',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Thermal Satisfaction', 'Visual Satisfaction', 'Indoor Air Quality',
       'Overall Satisfaction'],
      dtype='object')>
```

[375]:
```
data1 = data1.drop(columns = ['Gender','Employee Home Town','Age','Working␣
↪Hours', 'Is there a blind wall ','Distance between your work desk and the␣
↪nearest window?','Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation','Glazing U␣
↪value','Share of the area served by AC(%)', 'Smart control of HVAC ','Smart␣
↪control of lighting system', 'Area served by lighting', 'LUX level',
       'Thermal Satisfaction', 'Visual Satisfaction','Overall Satisfaction'])
```

[ ]:

[ ]:

[ ]:

[376]:
```
# duplicate the dataset

data1_copy = data1.copy()
data1_copy.shape
```

[376]: (1091, 11)

[377]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Window to Wall Ratio (WWR)', 'Total Window Area',
       'PM 2.5 level ', 'PM 10  level ', 'CO2 PPM', 'Indoor Air Quality'],
      dtype='object')>
```

[378]:
```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1091 entries, 0 to 1090
```

```
Data columns (total 11 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   EmpID                       1091 non-null   int64
 1   Building Type               1091 non-null   object
 2   Building                    1091 non-null   object
 3   Building Location           0 non-null      float64
 4   Gross Floor Area            1091 non-null   int64
 5   Window to Wall Ratio (WWR)  1091 non-null   int64
 6   Total Window Area           1091 non-null   float64
 7   PM 2.5 level                1091 non-null   int64
 8   PM 10  level                1091 non-null   int64
 9   CO2 PPM                     1091 non-null   int64
 10  Indoor Air Quality          1091 non-null   float64
dtypes: float64(3), int64(6), object(2)
memory usage: 93.9+ KB
```

[ ]:

[379]:
```
print(data1['Building'].unique().tolist())
```

```
['07', 'F4', '03', 'F2', '02', '04', '06', '05', '01', 'F1', 'F6', '08', 'F3',
'F5']
```

[ ]:

[380]:
```
data1 = data1.drop(columns = ['Building'])
```

[ ]:

[381]:
```
#Checking descriptive columns

tex_columns = data1.columns[(data1.dtypes =='object').values].tolist()
tex_columns
```

[381]:
```
['Building Type']
```

[382]:
```
data1.head(2)
```

[382]:
```
   EmpID Building Type  Building Location  Gross Floor Area  \
0      1        Office                NaN              7632
1      2        Office                NaN              7632

   Window to Wall Ratio (WWR)  Total Window Area  PM 2.5 level   \
0                          50               45.0             18
1                          50               45.0             18
```

```
   PM 10  level   CO2 PPM  Indoor Air Quality
0             37     2500                3.38
1             37     2500                3.37
```

[383]:
```
df1=data1
```

[384]:
```
print(df1['Building Type'].unique().tolist())
```

```
['Office', 'Factory']
```

[385]:
```
#print(df1['Building'].unique().tolist())
```

[386]:
```
df1.corr()['Indoor Air Quality']
```

[386]:
```
EmpID                          -0.217991
Building Location                    NaN
Gross Floor Area               -0.312798
Window to Wall Ratio (WWR)      0.621136
Total Window Area               0.144883
PM 2.5 level                   -0.822760
PM 10  level                   -0.848444
CO2 PPM                        -0.842865
Indoor Air Quality              1.000000
Name: Indoor Air Quality, dtype: float64
```

[ ]:
```

```

[387]:
```
df3 = df1.copy()
```

[388]:
```
df3 = df3.drop(columns = ['EmpID','Building Location'])
```

2 one-hot encoding

[389]:
```
df4 = pd.get_dummies(df3)
```

[390]:
```
df4.shape
```

[390]:
```
(1091, 9)
```

[391]:
```
#correlation of the variables to the Thermal satisfaction

df4.corr()['Indoor Air Quality']
```

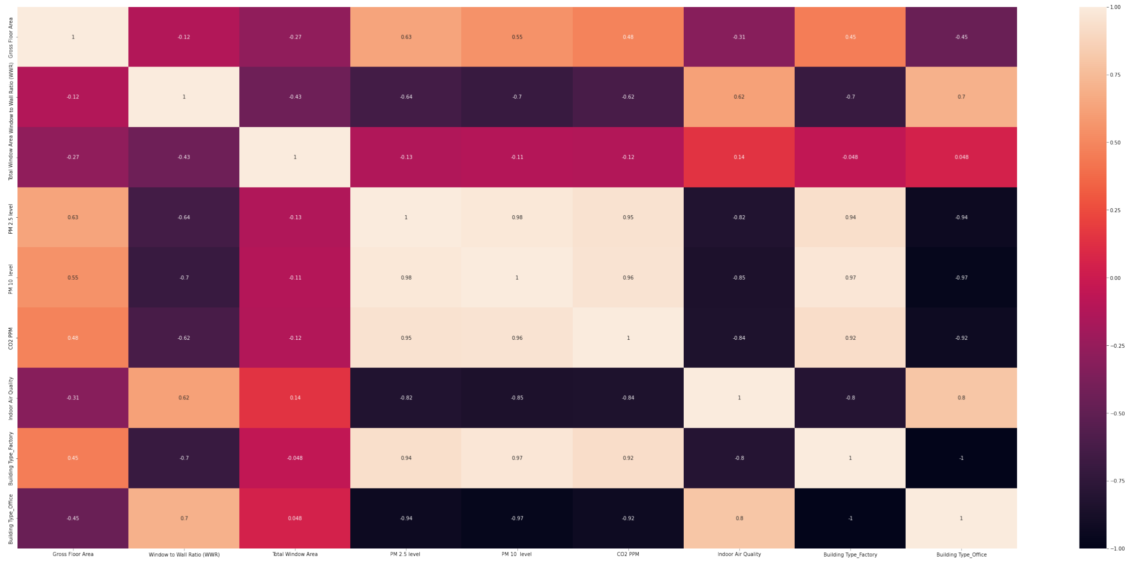
[391]:
```
Gross Floor Area               -0.312798
Window to Wall Ratio (WWR)      0.621136
Total Window Area               0.144883
```

```
PM 2.5 level              -0.822760
PM 10  level              -0.848444
CO2 PPM                   -0.842865
Indoor Air Quality         1.000000
Building Type_Factory     -0.798600
Building Type_Office       0.798600
Name: Indoor Air Quality, dtype: float64
```

[392]:
```python
df4.head(2)
```

[392]:
```
   Gross Floor Area  Window to Wall Ratio (WWR)  Total Window Area  \
0              7632                          50               45.0
1              7632                          50               45.0

   PM 2.5 level   PM 10  level   CO2 PPM  Indoor Air Quality  \
0             18             37     2500                3.38
1             18             37     2500                3.37

   Building Type_Factory  Building Type_Office
0                      0                     1
1                      0                     1
```

[393]:
```python
# correlation metrix

plt.figure(figsize = (45,20))
sns.heatmap(df4.corr(), annot=True)
plt.show()
```



[ ]:

```

```

[394]:
```
#df4 = df4.drop(columns = ['EmpID','Age','Changed your residence','Hometown␣
↪nature (1-3)'])
```

[395]:
```
df4.dtypes
```

[395]:
```
Gross Floor Area                 int64
Window to Wall Ratio (WWR)       int64
Total Window Area              float64
PM 2.5 level                     int64
PM 10  level                     int64
CO2 PPM                          int64
Indoor Air Quality             float64
Building Type_Factory            uint8
Building Type_Office             uint8
dtype: object
```

[396]:
```
df4.shape
```

[396]:
```
(1091, 9)
```

3 Model Training

[397]:
```
#Independent variables and dependent variables

X = df4.drop(['Indoor Air Quality'], axis=1)# Input features (attributes)
y = df4['Indoor Air Quality'] # Target vector
print('X shape: {}'.format(np.shape(X)))
print('y shape: {}'.format(np.shape(y)))


#train and test split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25)
```

```
X shape: (1091, 8)
y shape: (1091,)
```

[398]:
```
X_train.shape, X_test.shape
```

[398]:
```
((818, 8), (273, 8))
```

[399]:
```
#Function to return the model name and the accuracy value

def model_acc(model):
```

```
    model.fit(X_train, y_train)
    acc = model.score(X_test, y_test)
    print(str(model)+ ' --> ' +str(acc))
```

[400]:
```
#Find the best regression model


#Support Vector Regression

from sklearn.svm import SVR
svma = SVR(kernel = 'rbf')
model_acc(svma)


#LassoRegression

from sklearn.linear_model import Lasso
lasso = Lasso()
model_acc(lasso)

#DecisionTreeRegressor

from sklearn.tree import DecisionTreeRegressor
dt = DecisionTreeRegressor()
model_acc(dt)

#RandomForestRegressor

from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
model_acc(rf)
```

```
SVR() --> 0.6709831935235746
Lasso() --> 0.7070587959814456
DecisionTreeRegressor() --> 0.80250825464357
RandomForestRegressor() --> 0.8020524081505687
```

[ ]:
```

```

[401]:
```
# Calculate the RMSE for each model


from sklearn.metrics import classification_report, confusion_matrix,␣
↪accuracy_score
from sklearn.metrics import mean_squared_error
```

```
classifier_rf = rf.fit(X_train,y_train)
y_pred_rf = classifier_rf.predict(X_test)


classifier_dt = dt.fit(X_train,y_train)
y_pred_dt = classifier_dt.predict(X_test)

classifier_ls = lasso.fit(X_train,y_train)
y_pred_ls = classifier_ls.predict(X_test)

classifier_svm = svma.fit(X_train,y_train)
y_pred_svm = classifier_svm.predict(X_test)


# Calculate the RMSE for each model
from sklearn.metrics import mean_squared_error
rf_rmse = mean_squared_error(y_test, y_pred_rf, squared=False)
lasso_rmse = mean_squared_error(y_test, y_pred_ls, squared=False)
dt_rmse = mean_squared_error(y_test, y_pred_dt, squared=False)
svm_rmse = mean_squared_error(y_test, y_pred_svm, squared=False)

print("Random Forest RMSE    :", rf_rmse)
print("Lasso Regression RMSE :", lasso_rmse)
print("Decision Tree RMSE    :", dt_rmse)
print("SVM RMSE              :", svm_rmse)
```

```
Random Forest RMSE    : 0.4176983251111671
Lasso Regression RMSE : 0.5082257297525499
Decision Tree RMSE    : 0.4172927995900807
SVM RMSE              : 0.5386112971998044
```

[ ]:

```

```

[402]:

```
from sklearn.metrics import mean_absolute_error

rf_mae = mean_absolute_error(y_test, y_pred_rf)
lasso_mae = mean_absolute_error(y_test, y_pred_ls)
dt_mae = mean_absolute_error(y_test, y_pred_dt)
svm_mae = mean_absolute_error(y_test, y_pred_svm)

print("Random Forest MAE    :", rf_mae)
print("Lasso Regression MAE :", lasso_mae)
print("Decision Tree MAE    :", dt_mae)
print("SVM MAE              :", svm_mae)
```

```
Random Forest MAE    : 0.19247975725351033
Lasso Regression MAE : 0.3340752016904185
```

```
Decision Tree MAE     : 0.19385421881436954
SVM MAE               : 0.3000416182003782
```

[ ]:

[403]:
```
from sklearn.model_selection import cross_val_score

# Define the models
lasso = Lasso()
dt = DecisionTreeRegressor()
svm = SVR()
rf = RandomForestRegressor()

# Perform cross-validation and print the mean MAE scores
lasso_scores = cross_val_score(lasso, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Lasso Regression MAE (Cross-Validation) :", -np.mean(lasso_scores))

dt_scores = cross_val_score(dt, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Decision Tree MAE (Cross-Validation)    :", -np.mean(dt_scores))

svm_scores = cross_val_score(svm, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("SVM MAE (Cross-Validation)              :", -np.mean(svm_scores))

rf_scores = cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Random Forest MAE (Cross-Validation)    :", -np.mean(rf_scores))
```

```
Lasso Regression MAE (Cross-Validation) : 0.34800662415965816
Decision Tree MAE (Cross-Validation)    : 0.21295652924514097
SVM MAE (Cross-Validation)              : 0.309360362192954
Random Forest MAE (Cross-Validation)    : 0.21310962398897368
```

[ ]:

4 Hyperparameter tunning using Randomforest

[404]:
```
#find the best model for this scenario

from sklearn.model_selection import GridSearchCV

parameters = {'n_estimators':[10, 50, 100],
              'criterion':['squared_error','absolute_error','poisson']}
```

```
grid_obj = GridSearchCV(estimator=rf, param_grid=parameters)

grid_fit = grid_obj.fit(X_train, y_train)

best_model = grid_fit.best_estimator_

best_model
```

[404]: RandomForestRegressor(criterion='poisson')

[405]:
```
#Score accuracy of the best model

best_model.score(X_test, y_test)
```

[405]: 0.8023532906368369

[406]:
```
X_test.columns
```

[406]:
```
Index(['Gross Floor Area', 'Window to Wall Ratio (WWR)', 'Total Window Area',
       'PM 2.5 level ', 'PM 10  level ', 'CO2 PPM', 'Building Type_Factory',
       'Building Type_Office'],
      dtype='object')
```

[ ]:

5 Save the Model as pickle file

[407]:
```
import pickle
with open('IndoorAirQuality.pickle', 'wb') as file:
    pickle.dump(best_model, file)
```

[408]:
```
df4.head(2)
```

[408]:
```
   Gross Floor Area  Window to Wall Ratio (WWR)  Total Window Area  \
0              7632                          50               45.0
1              7632                          50               45.0

   PM 2.5 level   PM 10  level   CO2 PPM  Indoor Air Quality  \
0             18             37     2500                3.38
1             18             37     2500                3.37

   Building Type_Factory  Building Type_Office
0                      0                     1
1                      0                     1
```

6 Test the prediction

[409]:
```
pred_value = best_model.predict([[7000,50,45,20,40,3000,0,1]])
pred_value
```

[409]:
```
array([2.94640391])
```

[ ]:
```

```

[410]:
```
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.model_selection import cross_val_predict
```

[411]:
```
# Make predictions using the model

IAQ_preds = best_model.predict(X_test)
```

[412]:
```
# Evaluate the performance of the model

mse = mean_squared_error(y_test, IAQ_preds)
print("IAQ_preds-model MSE:", mse)
```

```
IAQ_preds-model MSE: 0.17426991612766862
```

[ ]:
```

```

[413]:
```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[ ]:
```

```

[414]:
```
# Evaluate Performance
MSE = mean_squared_error(y_test, cv_predictions)
RMSE = np.sqrt(MSE)
MAE = mean_absolute_error(y_test, cv_predictions)
R2 = r2_score(y_test, cv_predictions)
```

[ ]:
```

```

[415]:
```
# Interpret Results

print("Mean Squared Error (MSE):", MSE)
print("Root Mean Squared Error (RMSE):", RMSE)
print("Mean Absolute Error (MAE):", MAE)
print("R-squared (R2) Score:", R2)
```

```
Mean Squared Error (MSE): 0.19804948255807253
```

```
Root Mean Squared Error (RMSE): 0.4450275076420249
Mean Absolute Error (MAE): 0.2133350802819815
R-squared (R2) Score: 0.7753839022335687
```

[ ]:

[416]:
```python
# calculating VIF for each feature

from statsmodels.stats.outliers_influence import variance_inflation_factor
from statsmodels.tools.tools import add_constant

dfdata = pd.DataFrame(df4[['Gross Floor Area', 'Window to Wall Ratio (WWR)',␣
 ↪'Total Window Area',
       'PM 2.5 level ', 'PM 10  level ', 'CO2 PPM']])

X = add_constant(dfdata)
X = dfdata.assign(const=1)

pd.Series([variance_inflation_factor(X.values, i)
               for i in range(X.shape[1])],
              index=X.columns)
```

[416]:
```
Gross Floor Area                3.227520
Window to Wall Ratio (WWR)      5.718684
Total Window Area               2.214636
PM 2.5 level                    8.045625
PM 10  level                    8.915499
CO2 PPM                        18.486910
const                         150.271725
dtype: float64
```

[417]:
```python
from sklearn.tree import plot_tree

plt.figure(figsize=(100,100))
plot_tree(best_model.estimators_[20], feature_names = df4.
 ↪columns,class_names=['Indoor Air Quality'],filled=True);
```

[ ]:

```

```

[418]:

```python
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[419]:

```python
from sklearn.metrics import mean_absolute_error

rf_mse = mean_squared_error(y_test, cv_predictions)
print("Random Forest MSE     :", rf_mse)

from sklearn.metrics import mean_squared_error
rf_rmse = np.sqrt(mean_squared_error(y_test, cv_predictions))
print("Random Forest RMSE    :", rf_rmse)
```

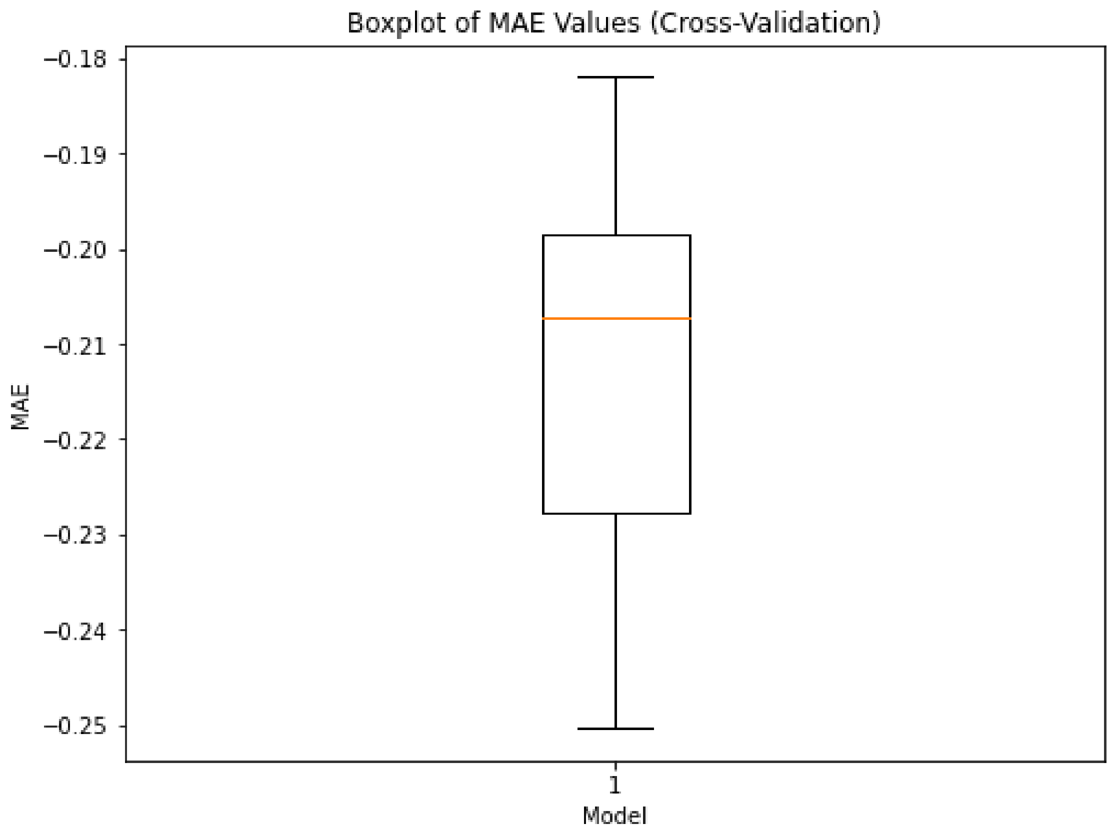
```
rf_mae = mean_absolute_error(y_test, cv_predictions)
print("Random Forest MAE     :", rf_mae)

from sklearn.metrics import r2_score
rf_r2 = r2_score(y_test, cv_predictions)
print("Random Forest R2 Score:", rf_r2)
```

```
Random Forest MSE     : 0.19896050595547832
Random Forest RMSE    : 0.44604989177835064
Random Forest MAE     : 0.21509378599046217
Random Forest R2 Score: 0.7743506729725975
```

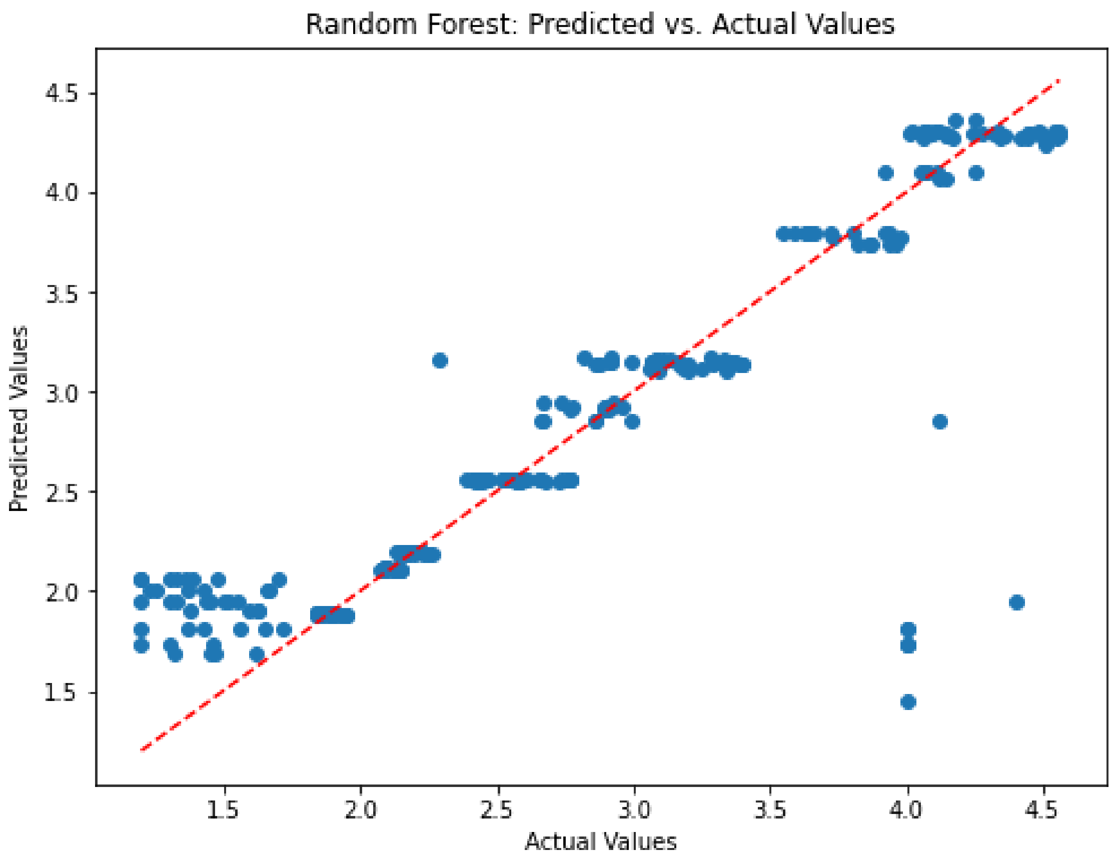
[420]:

```
from sklearn.model_selection import cross_val_score

# Perform cross-validation with 5 folds
rf_cv_mae = -cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')

print("Random Forest MAE (Cross-Validation):", rf_cv_mae.mean())

import matplotlib.pyplot as plt

# MAE values from cross-validation
mae_values = [-rf_cv_mae[fold] for fold in range(5)]

# Boxplot of MAE values
plt.figure(figsize=(8, 6))
plt.boxplot(mae_values)
plt.title('Boxplot of MAE Values (Cross-Validation)')
plt.xlabel('Model')
plt.ylabel('MAE')
plt.show()
```

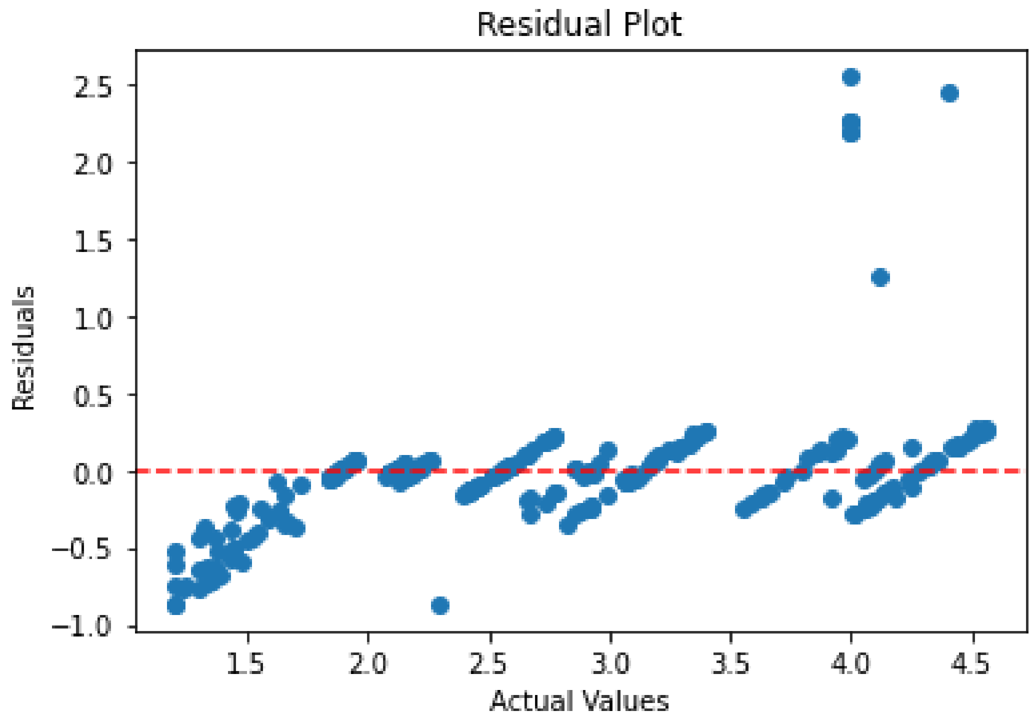
```
Random Forest MAE (Cross-Validation): 0.21318695168856522
```



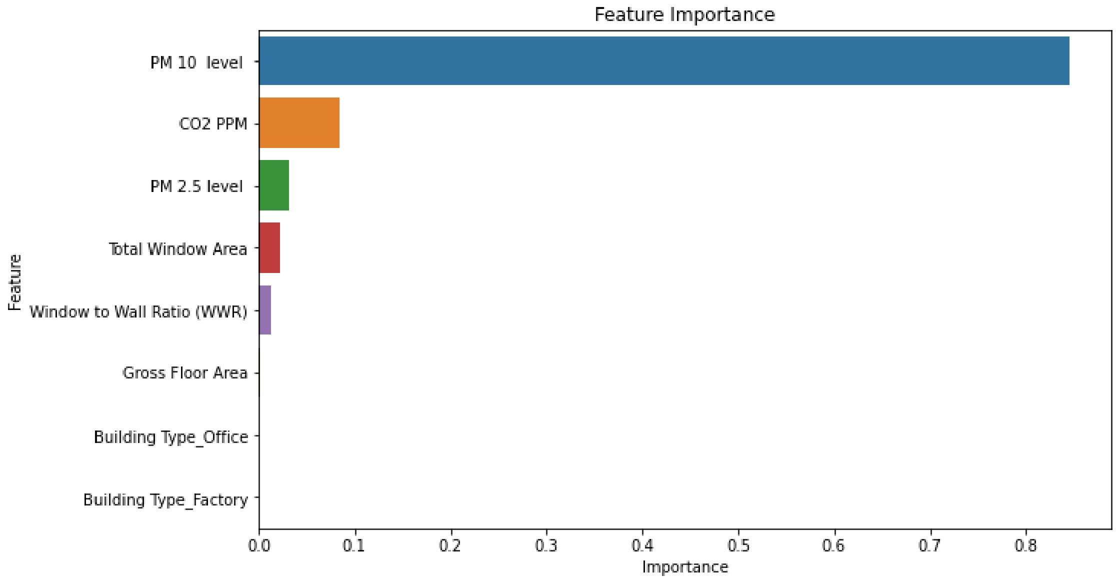
[421]:
```
rf_preds=cv_predictions

plt.figure(figsize=(8, 6))
plt.scatter(y_test, rf_preds)
plt.xlabel("Actual Values")
plt.ylabel("Predicted Values")
plt.title("Random Forest: Predicted vs. Actual Values")
plt.plot([min(y_test), max(y_test)], [min(y_test), max(y_test)], color='red',␣
 ↪linestyle='--')
plt.show()
```



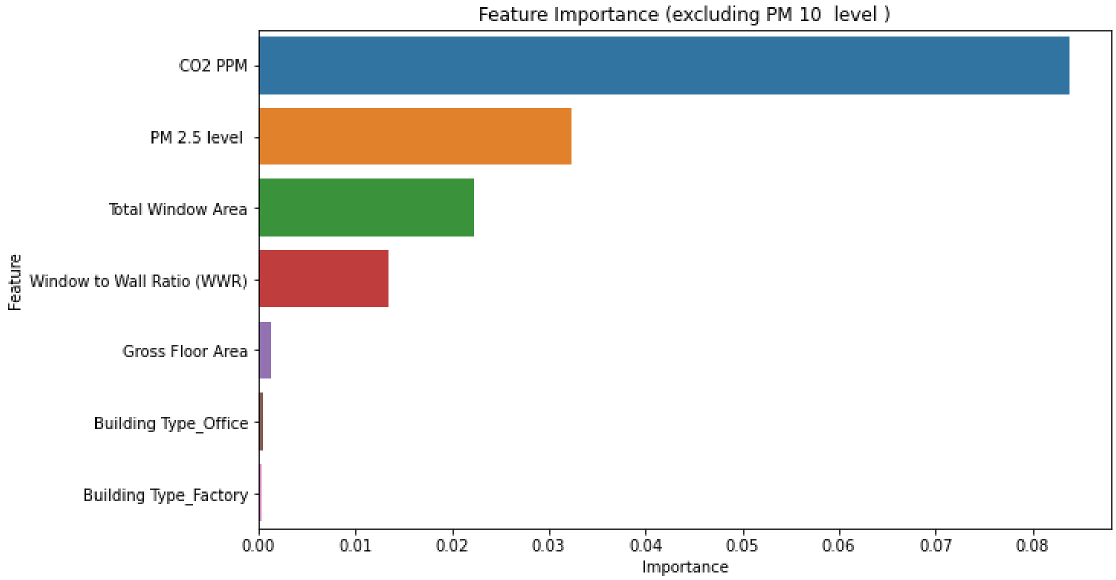
[422]:
```
residuals = y_test - rf_preds
plt.scatter(y_test, residuals)
plt.axhline(y=0, color='r', linestyle='--')
plt.xlabel('Actual Values')
plt.ylabel('Residuals')
plt.title('Residual Plot')
plt.show()
```



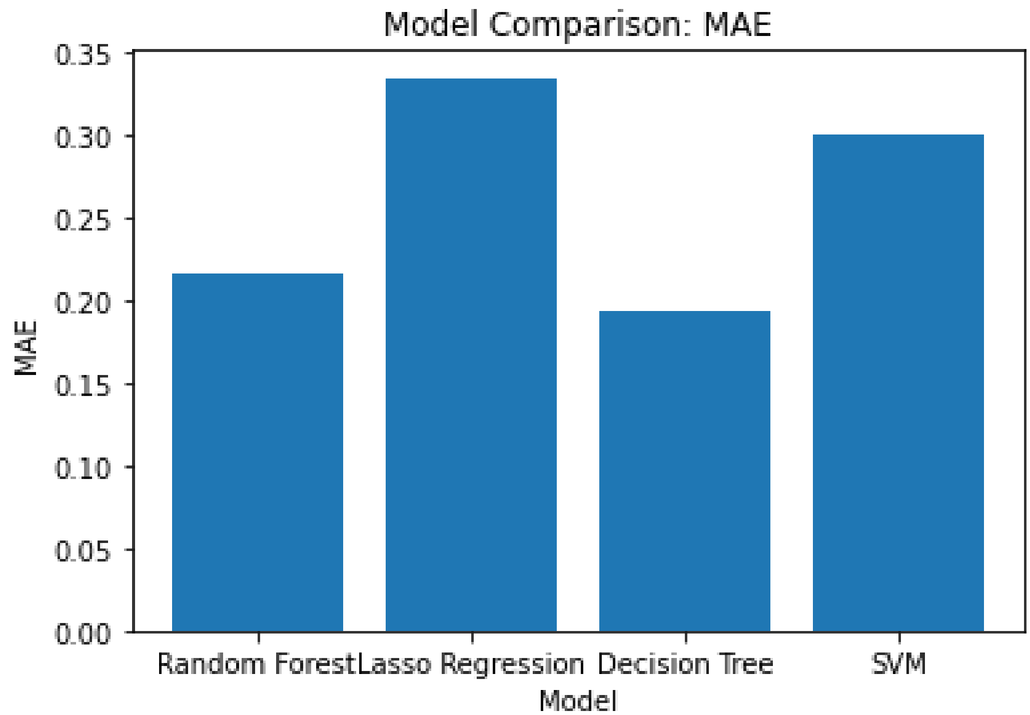
[423]:
```
# Train the random forest model
rf = RandomForestRegressor()
rf.fit(X_train, y_train)

# Calculate feature importance
importance = rf.feature_importances_

# Create a dataframe to store feature importance
feature_importance_df = pd.DataFrame({'Feature': X_train.columns, 'Importance':␣
↪importance})

# Sort the features by importance in descending order
feature_importance_df = feature_importance_df.sort_values('Importance',␣
↪ascending=False)

# Plot the feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df)
plt.title('Feature Importance')
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[424]:
```
# Remove a specific featuShare of PM 10  level from the dataframe
feature_to_remove = 'PM 10  level '
feature_importance_df_filtered =␣
↪feature_importance_df[feature_importance_df['Feature'] != feature_to_remove]

# Plot the updated feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df_filtered)
plt.title('Feature Importance (excluding {})'.format(feature_to_remove))
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[426]:
```
model_names = ['Random Forest', 'Lasso Regression', 'Decision Tree', 'SVM']
mae_values = [rf_mae, lasso_mae, dt_mae, svm_mae]

plt.bar(model_names, mae_values)
plt.xlabel('Model')
plt.ylabel('MAE')
plt.title('Model Comparison: MAE')
plt.show()
```


Model Comparison: MAE
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
MAE
Random Forest
Lasso Regression
Decision Tree
SVM
Model

[ ]:

[ ]:

ts-copy1

July 17, 2023

[18]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import norm

#Converting all Non-Numerical Columns to Numerical
from sklearn.preprocessing import LabelEncoder

from sklearn.metrics import accuracy_score, confusion_matrix, r2_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.tree import DecisionTreeClassifier, export_graphviz
from sklearn.model_selection import train_test_split
import warnings
warnings.filterwarnings('ignore')

%matplotlib inline
```

1 Load the Thermal dataset

[19]:
```
data1 = pd.read_excel("Data.xlsx")
```

[20]:
```
data1.shape
```

[20]:
```
(1091, 29)
```

[21]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location', 'Gender',
       'Employee Home Town', 'Age', 'Working Hours', 'Is there a blind wall ',
       'Distance between your work desk and the nearest window?',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
```

```
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Thermal Satisfaction', 'Visual Satisfaction', 'Indoor Air Quality',
       'Overall Satisfaction'],
      dtype='object')>
```

[22]:
```
data1 = data1.drop(columns = ['Gender','Employee Home Town','Age','Working␣
↪Hours', 'Is there a blind wall ','Distance between your work desk and the␣
↪nearest window?','Smart control of lighting system', 'PM 2.5 level ', 'PM 10␣
↪ level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level', 'Visual␣
↪Satisfaction', 'Indoor Air Quality','Overall Satisfaction'])
```

[23]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Thermal Satisfaction'],
      dtype='object')>
```

[ ]:
```

```

[24]:
```
# duplicate the dataset

data1_copy = data1.copy()
data1_copy.shape
```

[24]:
```
(1091, 14)
```

[25]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Thermal Satisfaction'],
      dtype='object')>
```

[26]:
```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1091 entries, 0 to 1090
Data columns (total 14 columns):
 #   Column                             Non-Null Count  Dtype
---  ------                             --------------  -----
 0   EmpID                              1091 non-null   int64
 1   Building Type                      1091 non-null   object
 2   Building                           1091 non-null   object
 3   Building Location                  0 non-null      float64
 4   Gross Floor Area                   1091 non-null   int64
 5   Wall Insulation U value            1091 non-null   float64
 6   Roof Insulation U value            1091 non-null   float64
 7   Thickness of the Wall Insulation   1091 non-null   int64
 8   Window to Wall Ratio (WWR)         1091 non-null   int64
 9   Glazing U value                    1091 non-null   float64
 10  Total Window Area                  1091 non-null   float64
 11  Share of the area served by AC(%)  1091 non-null   float64
 12  Smart control of HVAC              1091 non-null   int64
 13  Thermal Satisfaction               1091 non-null   float64
dtypes: float64(7), int64(5), object(2)
memory usage: 119.5+ KB
```

[ ]:

[27]:
```
print(data1['Building'].unique().tolist())
```

```
['07', 'F4', '03', 'F2', '02', '04', '06', '05', '01', 'F1', 'F6', '08', 'F3',
'F5']
```

[ ]:

[28]:
```
data1 = data1.drop(columns = ['Building'])
```

[ ]:

[29]:
```
#Checking descriptive columns

tex_columns = data1.columns[(data1.dtypes =='object').values].tolist()
tex_columns
```

[29]:
```
['Building Type']
```

[30]:
```
data1.head(2)
```

[30]:
```
   EmpID Building Type  Building Location  Gross Floor Area  \
0      1        Office                NaN              7632
1      2        Office                NaN              7632
```

```
   Wall Insulation U value  Roof Insulation U value  \
0                     0.26                     0.2
1                     0.26                     0.2

   Thickness of the Wall Insulation  Window to Wall Ratio (WWR)  \
0                                32                          50
1                                32                          50

   Glazing U value  Total Window Area  Share of the area served by AC(%)  \
0             0.48               45.0                               78.0
1             0.48               45.0                               78.0

   Smart control of HVAC   Thermal Satisfaction
0                      1                    2.0
1                      1                    3.0
```

[31]:
```
df1=data1
```

[32]:
```
print(df1['Building Type'].unique().tolist())
```

```
['Office', 'Factory']
```

[33]:
```
#print(df1['Building'].unique().tolist())
```

[34]:
```
df1.corr()['Thermal Satisfaction']
```

[34]:
```
EmpID                                 0.066888
Building Location                          NaN
Gross Floor Area                      0.350823
Wall Insulation U value              -0.021777
Roof Insulation U value              -0.519799
Thickness of the Wall Insulation      0.486843
Window to Wall Ratio (WWR)            0.348407
Glazing U value                      -0.001262
Total Window Area                    -0.322638
Share of the area served by AC(%)     0.654350
Smart control of HVAC                 0.313969
Thermal Satisfaction                  1.000000
Name: Thermal Satisfaction, dtype: float64
```

[ ]:
```

```

[35]:
```
df3 = df1.copy()
```

[36]:
```
df3 = df3.drop(columns = ['EmpID','Building Location'])
```

2 one-hot encoding

[37]:
```
df4 = pd.get_dummies(df3)
```

[38]:
```
df4.shape
```

[38]:
```
(1091, 12)
```

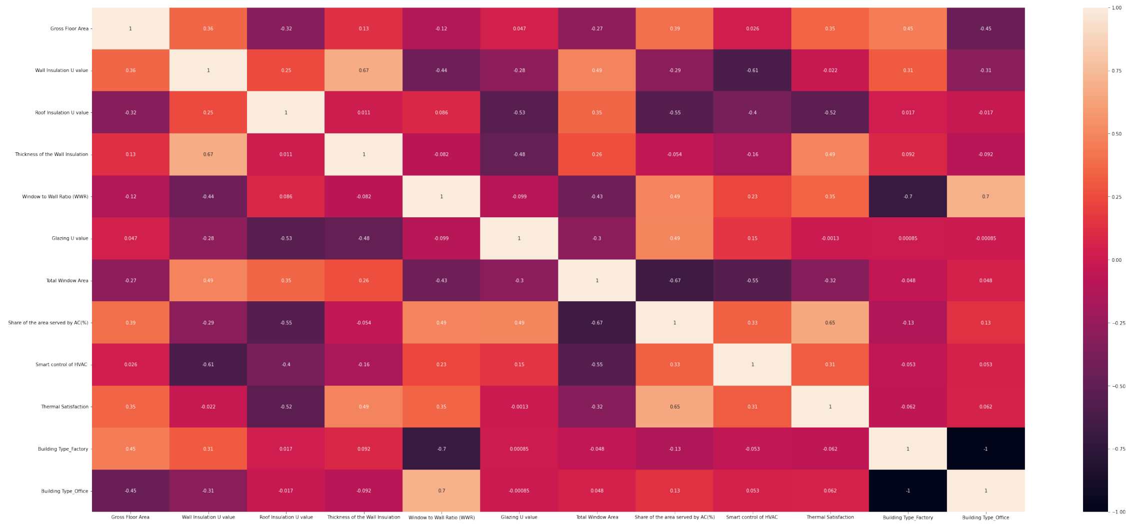
[39]:
```
#correlation of the variables to the Thermal satisfaction

df4.corr()['Thermal Satisfaction']
```

[39]:
```
Gross Floor Area                      0.350823
Wall Insulation U value              -0.021777
Roof Insulation U value              -0.519799
Thickness of the Wall Insulation      0.486843
Window to Wall Ratio (WWR)            0.348407
Glazing U value                      -0.001262
Total Window Area                    -0.322638
Share of the area served by AC(%)     0.654350
Smart control of HVAC                 0.313969
Thermal Satisfaction                  1.000000
Building Type_Factory                -0.061776
Building Type_Office                  0.061776
Name: Thermal Satisfaction, dtype: float64
```

[40]:
```
df4.head(2)
```

[40]:
```
   Gross Floor Area  Wall Insulation U value  Roof Insulation U value  \
0              7632                     0.26                      0.2
1              7632                     0.26                      0.2

   Thickness of the Wall Insulation  Window to Wall Ratio (WWR)  \
0                                32                          50
1                                32                          50

   Glazing U value  Total Window Area  Share of the area served by AC(%)  \
0             0.48               45.0                               78.0
1             0.48               45.0                               78.0

   Smart control of HVAC  Thermal Satisfaction  Building Type_Factory  \
0                      1                   2.0                      0
1                      1                   3.0                      0

   Building Type_Office
0                     1
1                     1
```

[41]:
```
# correlation metrix

plt.figure(figsize = (45,20))
sns.heatmap(df4.corr(), annot=True)
plt.show()
```



[ ]:
```

```

[42]:
```
#df4 = df4.drop(columns = ['EmpID','Age','Changed your residence','Hometown␣
↪nature (1-3)'])
```

[43]:
```
df4.dtypes
```

[43]:
```
Gross Floor Area                       int64
Wall Insulation U value              float64
Roof Insulation U value              float64
Thickness of the Wall Insulation       int64
Window to Wall Ratio (WWR)             int64
Glazing U value                      float64
Total Window Area                    float64
Share of the area served by AC(%)    float64
Smart control of HVAC                  int64
Thermal Satisfaction                 float64
Building Type_Factory                  uint8
Building Type_Office                   uint8
dtype: object
```

[44]:
```
df4.shape
```

[44]: (1091, 12)

3 Model Training

[45]:
```
#Independent variables and dependent variables

X = df4.drop(['Thermal Satisfaction'], axis=1)# Input features (attributes)
y = df4['Thermal Satisfaction'] # Target vector
print('X shape: {}'.format(np.shape(X)))
print('y shape: {}'.format(np.shape(y)))


#train and test split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25)
```

```
X shape: (1091, 11)
y shape: (1091,)
```

[46]:
```
X_train.shape, X_test.shape
```

[46]: ((818, 11), (273, 11))

[47]:
```
#Function to return the model name and the accuracy value

def model_acc(model):
    model.fit(X_train, y_train)
    acc = model.score(X_test, y_test)
    print(str(model)+ ' --> ' +str(acc))
```

[ ]:
```

```

[48]:
```
#Find the best regression model



#Support Vector Regression

from sklearn.svm import SVR
svma = SVR(kernel = 'rbf')
model_acc(svma)


#LassoRegression

from sklearn.linear_model import Lasso
```

```python
lasso = Lasso()
model_acc(lasso)

#DecisionTreeRegressor

from sklearn.tree import DecisionTreeRegressor
dt = DecisionTreeRegressor()
model_acc(dt)

#RandomForestRegressor

from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
model_acc(rf)
```

```
SVR() --> 0.21219691455017586
Lasso() --> 0.6671797248899305
DecisionTreeRegressor() --> 0.857371769080836
RandomForestRegressor() --> 0.8577996612568597
```

[49]:

```python
# Calculate the RMSE for each model


from sklearn.metrics import classification_report, confusion_matrix,␣
↪accuracy_score
from sklearn.metrics import mean_squared_error

classifier_rf = rf.fit(X_train,y_train)
y_pred_rf = classifier_rf.predict(X_test)

classifier_dt = dt.fit(X_train,y_train)
y_pred_dt = classifier_dt.predict(X_test)

classifier_ls = lasso.fit(X_train,y_train)
y_pred_ls = classifier_ls.predict(X_test)

classifier_svm = svma.fit(X_train,y_train)
y_pred_svm = classifier_svm.predict(X_test)


# Calculate the RMSE for each model
from sklearn.metrics import mean_squared_error
rf_rmse = mean_squared_error(y_test, y_pred_rf, squared=False)
lasso_rmse = mean_squared_error(y_test, y_pred_ls, squared=False)
dt_rmse = mean_squared_error(y_test, y_pred_dt, squared=False)
svm_rmse = mean_squared_error(y_test, y_pred_svm, squared=False)
```

```
print("Random Forest RMSE    :", rf_rmse)
print("Lasso Regression RMSE :", lasso_rmse)
print("Decision Tree RMSE    :", dt_rmse)
print("SVM RMSE              :", svm_rmse)
```

```
Random Forest RMSE    : 0.29791673440223704
Lasso Regression RMSE : 0.45543837153146
Decision Tree RMSE    : 0.29814496402366236
SVM RMSE              : 0.7007023419193515
```

[ ]:

[50]:
```
from sklearn.metrics import mean_absolute_error

rf_mae = mean_absolute_error(y_test, y_pred_rf)
lasso_mae = mean_absolute_error(y_test, y_pred_ls)
dt_mae = mean_absolute_error(y_test, y_pred_dt)
svm_mae = mean_absolute_error(y_test, y_pred_svm)

print("Random Forest MAE    :", rf_mae)
print("Lasso Regression MAE :", lasso_mae)
print("Decision Tree MAE    :", dt_mae)
print("SVM MAE              :", svm_mae)
```

```
Random Forest MAE    : 0.20493123004002636
Lasso Regression MAE : 0.35769851936407815
Decision Tree MAE    : 0.20510322660088365
SVM MAE              : 0.5426112771297875
```

[ ]:

[52]:
```
from sklearn.model_selection import cross_val_score

# Define the models
lasso = Lasso()
dt = DecisionTreeRegressor()
svm = SVR()
rf = RandomForestRegressor()

# Perform cross-validation and print the mean MAE scores
lasso_scores = cross_val_score(lasso, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Lasso Regression MAE (Cross-Validation) :", -np.mean(lasso_scores))

dt_scores = cross_val_score(dt, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
```

```
print("Decision Tree MAE (Cross-Validation)    :", -np.mean(dt_scores))

svm_scores = cross_val_score(svm, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("SVM MAE (Cross-Validation)              :", -np.mean(svm_scores))

rf_scores = cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Random Forest MAE (Cross-Validation)    :", -np.mean(rf_scores))
```

```
Lasso Regression MAE (Cross-Validation) : 0.36700371364041623
Decision Tree MAE (Cross-Validation)    : 0.20729891581141882
SVM MAE (Cross-Validation)              : 0.5494714891981947
Random Forest MAE (Cross-Validation)    : 0.20731586172516553
```

[ ]:

[ ]:

4 Hyperparameter tunning using Randomforest

[53]:
```
#find the best model for this scenario

from sklearn.model_selection import GridSearchCV

parameters = {'n_estimators':[10, 50, 100],
              'criterion':['squared_error','absolute_error','poisson']}

grid_obj = GridSearchCV(estimator=rf, param_grid=parameters)

grid_fit = grid_obj.fit(X_train, y_train)

best_model = grid_fit.best_estimator_

best_model
```

[53]:
```
RandomForestRegressor(criterion='poisson', n_estimators=50)
```

[54]:
```
#Score accuracy of the best model

best_model.score(X_test, y_test)
```

[54]:
```
0.8565175302011823
```

[55]:
```
X_test.columns
```

[55]: Index(['Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Building Type_Factory', 'Building Type_Office'],
      dtype='object')

[ ]:

5 Save the Model as pickle file

[56]:
```
import pickle
with open('Thermal.pickle', 'wb') as file:
    pickle.dump(best_model, file)
```

[57]:
```
df4.head(2)
```

[57]:
```
   Gross Floor Area  Wall Insulation U value  Roof Insulation U value  \
0              7632                     0.26                      0.2
1              7632                     0.26                      0.2

   Thickness of the Wall Insulation  Window to Wall Ratio (WWR)  \
0                                32                          50
1                                32                          50

   Glazing U value  Total Window Area  Share of the area served by AC(%)  \
0             0.48               45.0                               78.0
1             0.48               45.0                               78.0

   Smart control of HVAC   Thermal Satisfaction  Building Type_Factory  \
0                      1                    2.0                      0
1                      1                    3.0                      0

   Building Type_Office
0                     1
1                     1
```

6 Test the prediction

[59]:
```
pred_value = best_model.predict([[7000,0.26,0.18,28,50,0.4,50,70,1,0,1]])
pred_value
```

[59]: array([2.99338464])

[ ]:

[60]:
```
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.model_selection import cross_val_predict
```

[61]:
```
# Make predictions using the model

Thermal_preds = best_model.predict(X_test)
```

[62]:
```
# Evaluate the performance of the model

mse = mean_squared_error(y_test, Thermal_preds)
print("Thermal-model MSE:", mse)
```

```
Thermal-model MSE: 0.08942280823050086
```

[ ]:
```

```

[63]:
```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[ ]:
```

```

[64]:
```
# Evaluate Performance
MSE = mean_squared_error(y_test, cv_predictions)
RMSE = np.sqrt(MSE)
MAE = mean_absolute_error(y_test, cv_predictions)
R2 = r2_score(y_test, cv_predictions)
```

[ ]:
```

```

[65]:
```
# Interpret Results

print("Mean Squared Error (MSE)      :", MSE)
print("Root Mean Squared Error (RMSE):", RMSE)
print("Mean Absolute Error (MAE)     :", MAE)
print("R-squared (R2) Score          :", R2)
```

```
Mean Squared Error (MSE)      : 0.09580432237193134
Root Mean Squared Error (RMSE): 0.3095227332069994
Mean Absolute Error (MAE)     : 0.20703448263880467
R-squared (R2) Score          : 0.8462781357090261
```

[ ]:
```

```

[68]:
```
# calculating VIF for each feature

from statsmodels.stats.outliers_influence import variance_inflation_factor
```

```python
from statsmodels.tools.tools import add_constant

dfdata = pd.DataFrame(df4[['Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ']])

X = add_constant(dfdata)
X = dfdata.assign(const=1)

pd.Series([variance_inflation_factor(X.values, i)
               for i in range(X.shape[1])],
              index=X.columns)
```

[68]:
```
Gross Floor Area                       5.623841
Wall Insulation U value                6.811833
Roof Insulation U value                3.976865
Thickness of the Wall Insulation      10.678259
Window to Wall Ratio (WWR)             3.020962
Glazing U value                        6.451798
Total Window Area                      2.718279
Share of the area served by AC(%)      7.413206
Smart control of HVAC                  3.273872
const                                181.210084
dtype: float64
```

[69]:
```python
from sklearn.tree import plot_tree

plt.figure(figsize=(100,100))
plot_tree(best_model.estimators_[20], feature_names = df4.
 ↪columns,class_names=['Thermal Satisfaction'],filled=True);
```

[ ]:

[ ]:

[70]:
```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[71]:
```
from sklearn.metrics import mean_absolute_error

rf_mse = mean_squared_error(y_test, cv_predictions)
print("Random Forest MSE     :", rf_mse)

from sklearn.metrics import mean_squared_error
rf_rmse = np.sqrt(mean_squared_error(y_test, cv_predictions))
```

```
print("Random Forest RMSE    :", rf_rmse)

rf_mae = mean_absolute_error(y_test, cv_predictions)
print("Random Forest MAE     :", rf_mae)

from sklearn.metrics import r2_score
rf_r2 = r2_score(y_test, cv_predictions)
print("Random Forest R2 Score:", rf_r2)
```

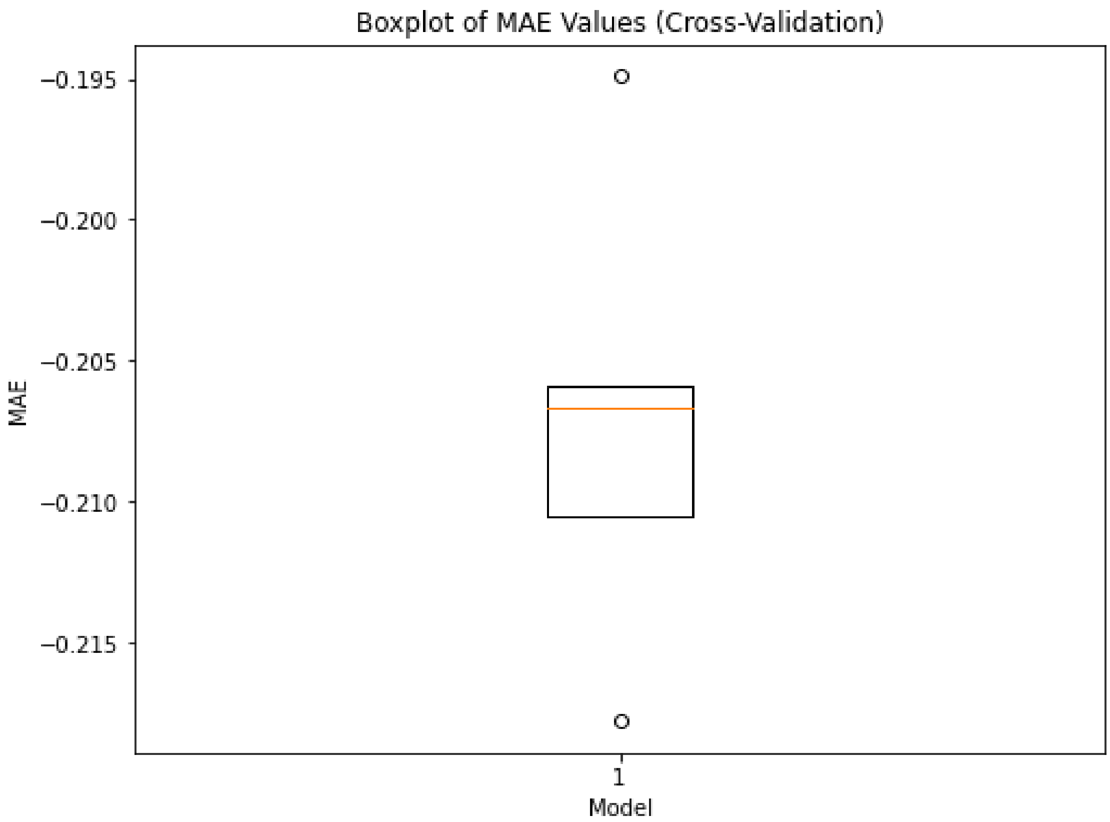
```
Random Forest MSE     : 0.0974550225377857
Random Forest RMSE    : 0.31217787003211117
Random Forest MAE     : 0.20894095245003166
Random Forest R2 Score: 0.8436295213187959
```

[ ]:

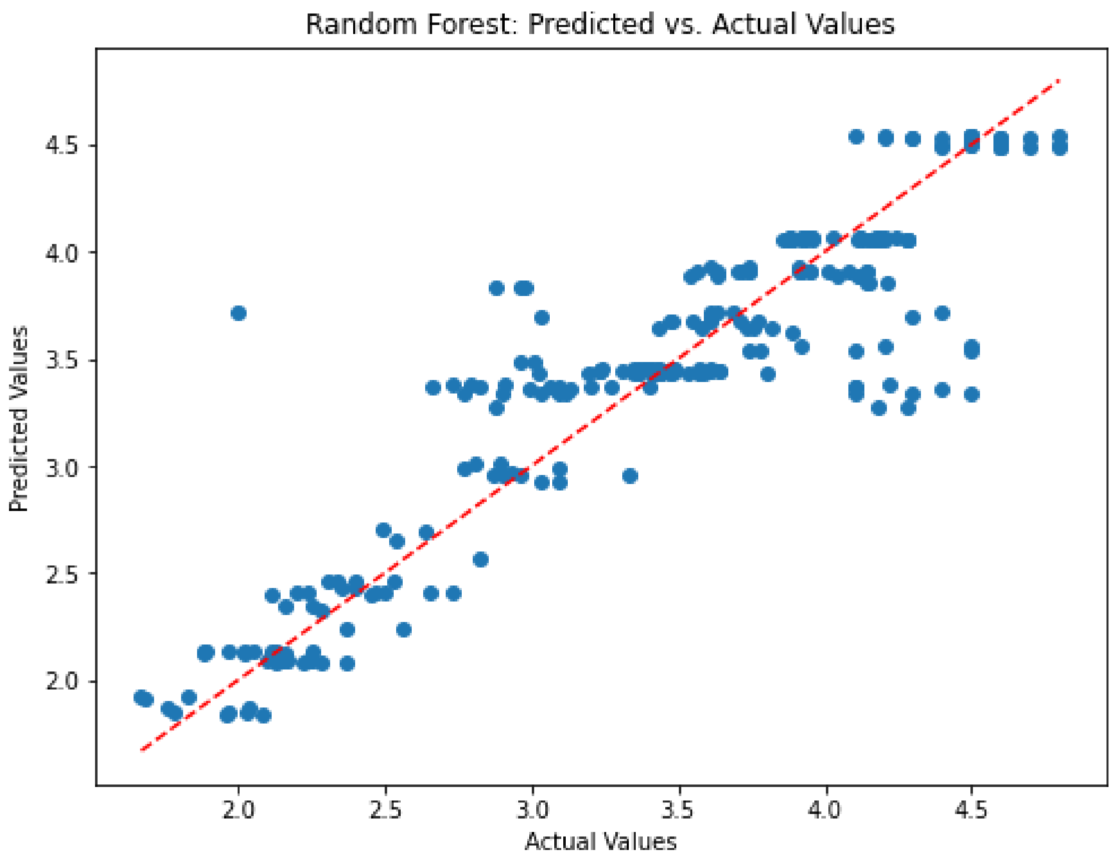
[72]:

```
from sklearn.model_selection import cross_val_score

# Perform cross-validation with 5 folds
rf_cv_mae = -cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')

print("Random Forest MAE (Cross-Validation):", rf_cv_mae.mean())

import matplotlib.pyplot as plt

# MAE values from cross-validation
mae_values = [-rf_cv_mae[fold] for fold in range(5)]

# Boxplot of MAE values
plt.figure(figsize=(8, 6))
plt.boxplot(mae_values)
plt.title('Boxplot of MAE Values (Cross-Validation)')
plt.xlabel('Model')
plt.ylabel('MAE')
plt.show()
```

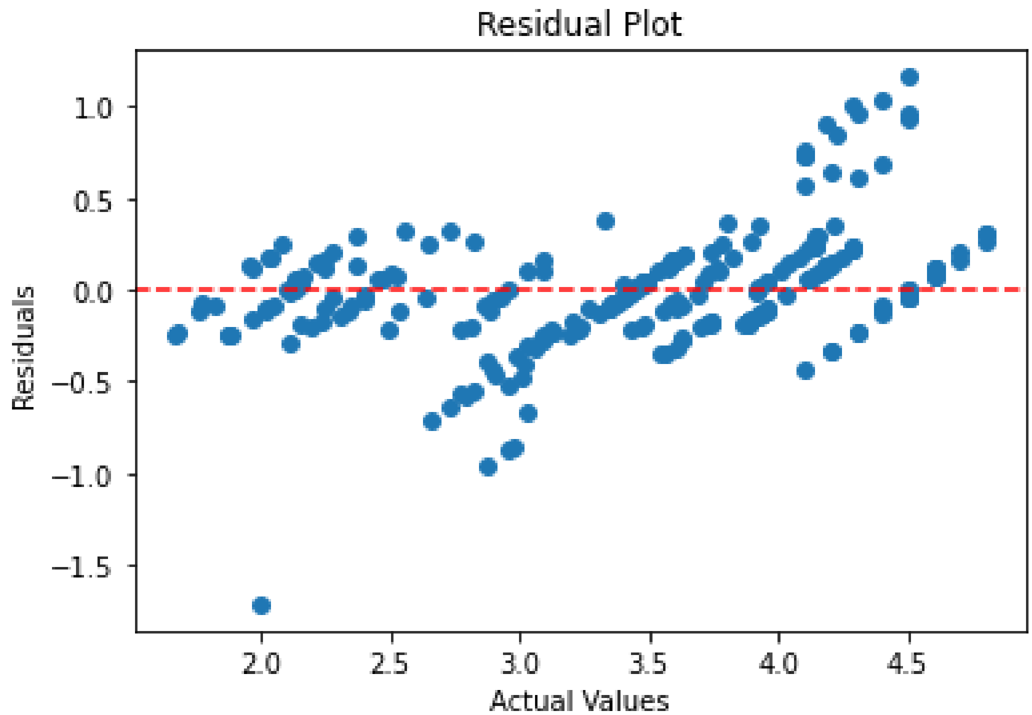
```
Random Forest MAE (Cross-Validation): 0.207162827297539
```



[73]:
```
rf_preds=cv_predictions

plt.figure(figsize=(8, 6))
plt.scatter(y_test, rf_preds)
plt.xlabel("Actual Values")
plt.ylabel("Predicted Values")
plt.title("Random Forest: Predicted vs. Actual Values")
plt.plot([min(y_test), max(y_test)], [min(y_test), max(y_test)], color='red',␣
↪linestyle='--')
plt.show()
```

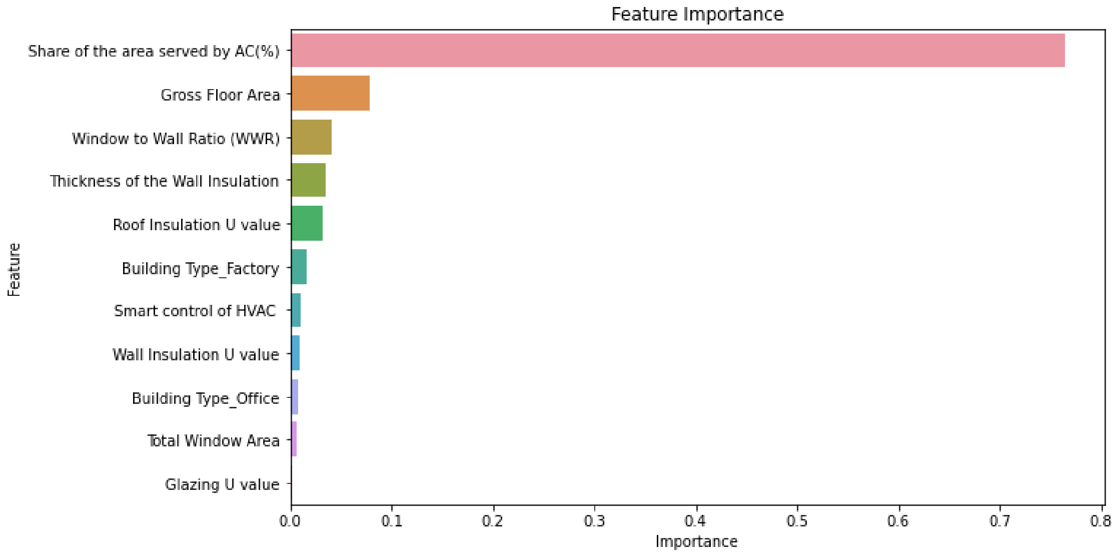


[74]:
```
residuals = y_test - rf_preds
plt.scatter(y_test, residuals)
plt.axhline(y=0, color='r', linestyle='--')
plt.xlabel('Actual Values')
plt.ylabel('Residuals')
plt.title('Residual Plot')
plt.show()
```



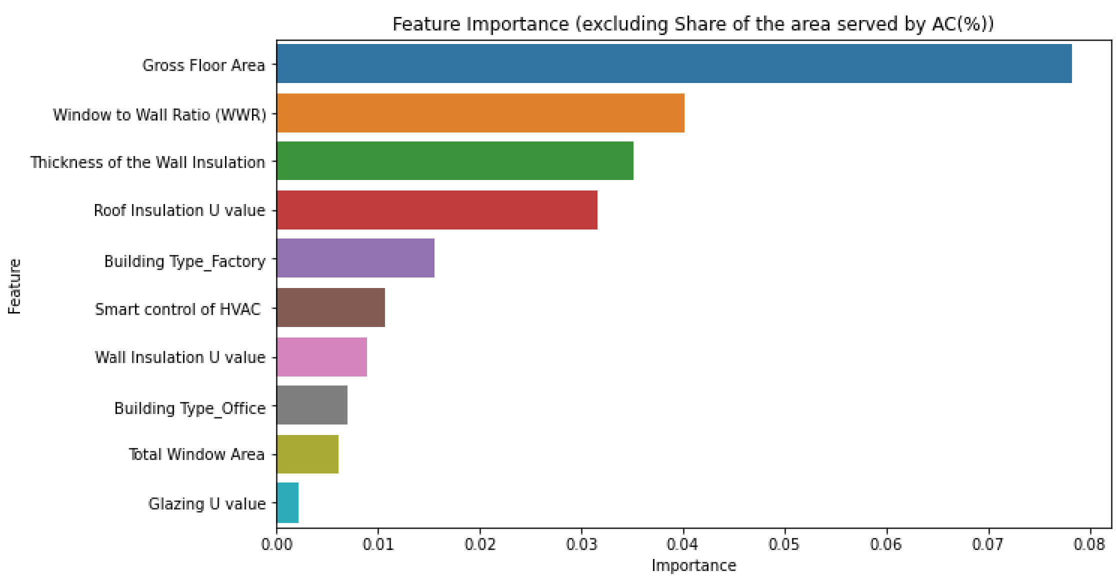
[75]:

```
# Train the random forest model
rf = RandomForestRegressor()
rf.fit(X_train, y_train)

# Calculate feature importance
importance = rf.feature_importances_

# Create a dataframe to store feature importance
feature_importance_df = pd.DataFrame({'Feature': X_train.columns, 'Importance':␣
↪importance})

# Sort the features by importance in descending order
feature_importance_df = feature_importance_df.sort_values('Importance',␣
↪ascending=False)

# Plot the feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df)
plt.title('Feature Importance')
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```

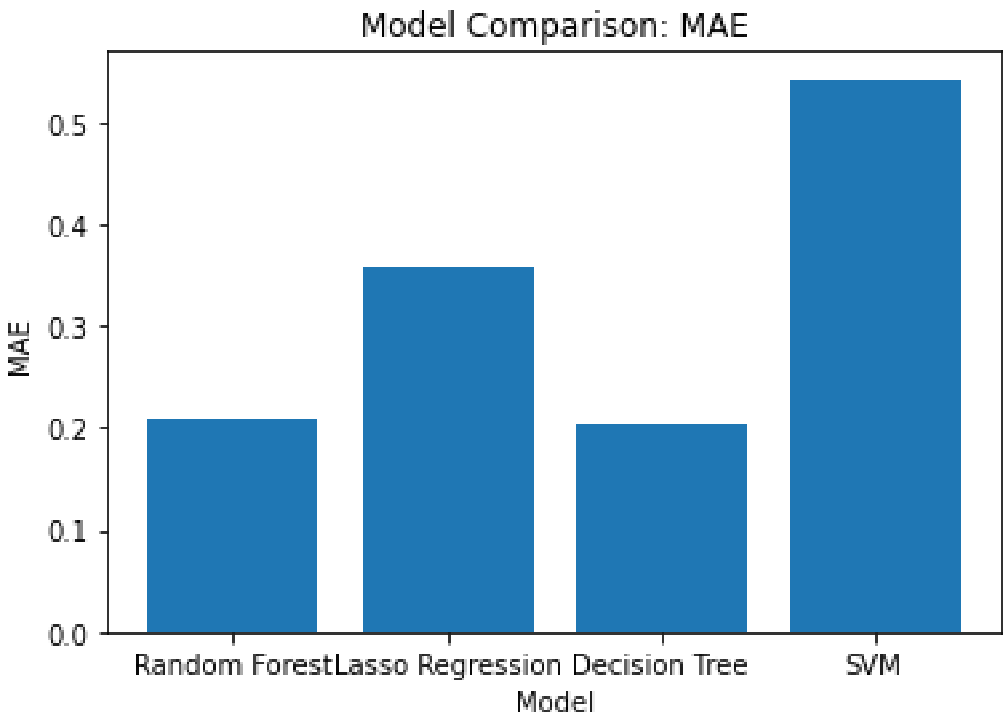


[76]:

```
# Remove a specific featuShare of the area served by AC(%re from the dataframe
feature_to_remove = 'Share of the area served by AC(%)'
feature_importance_df_filtered =␣
↪feature_importance_df[feature_importance_df['Feature'] != feature_to_remove]

# Plot the updated feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df_filtered)
plt.title('Feature Importance (excluding {})'.format(feature_to_remove))
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[78]:
```python
model_names = ['Random Forest', 'Lasso Regression', 'Decision Tree', 'SVM']
mae_values = [rf_mae, lasso_mae, dt_mae, svm_mae]

plt.bar(model_names, mae_values)
plt.xlabel('Model')
plt.ylabel('MAE')
plt.title('Model Comparison: MAE')
plt.show()
```



[ ]:
```python

```

vs-copy1

July 17, 2023

[7]:
```python
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import norm

#Converting all Non-Numerical Columns to Numerical
from sklearn.preprocessing import LabelEncoder

from sklearn.metrics import accuracy_score, confusion_matrix, r2_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.tree import DecisionTreeClassifier, export_graphviz
from sklearn.model_selection import train_test_split
import warnings
warnings.filterwarnings('ignore')

%matplotlib inline
```

1 Load the Visual dataset

[8]:
```python
data1 = pd.read_excel("Data.xlsx")
```

[9]:
```python
data1.shape
```

[9]:
```
(1091, 29)
```

[10]:
```python
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location', 'Gender',
       'Employee Home Town', 'Age', 'Working Hours', 'Is there a blind wall ',
       'Distance between your work desk and the nearest window?',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
```

```
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Thermal Satisfaction', 'Visual Satisfaction', 'Indoor Air Quality',
       'Overall Satisfaction'],
      dtype='object')>
```

[11]:
```
data1 = data1.drop(columns = ['Gender','Employee Home Town','Age','Working␣
↪Hours', 'Is there a blind wall ','Distance between your work desk and the␣
↪nearest window?','Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation','Glazing U␣
↪value','Share of the area served by AC(%)', 'Smart control of HVAC ', 'PM 2.
↪5 level ',
       'CO2 PPM', 'Thermal Satisfaction', 'Indoor Air Quality','Overall␣
↪Satisfaction'])
```

[12]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Window to Wall Ratio (WWR)', 'Total Window Area',
       'Smart control of lighting system', 'PM 10  level ',
       'Area served by lighting', 'LUX level', 'Visual Satisfaction'],
      dtype='object')>
```

[ ]:
```

```

[ ]:
```

```

[13]:
```
# duplicate the dataset

data1_copy = data1.copy()
data1_copy.shape
```

[13]: (1091, 12)

[14]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Window to Wall Ratio (WWR)', 'Total Window Area',
       'Smart control of lighting system', 'PM 10  level ',
       'Area served by lighting', 'LUX level', 'Visual Satisfaction'],
      dtype='object')>
```

[15]:
```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1091 entries, 0 to 1090
Data columns (total 12 columns):
 #   Column                           Non-Null Count  Dtype  
---  ------                           --------------  -----  
 0   EmpID                            1091 non-null   int64  
 1   Building Type                    1091 non-null   object 
 2   Building                         1091 non-null   object 
 3   Building Location                0 non-null      float64
 4   Gross Floor Area                 1091 non-null   int64  
 5   Window to Wall Ratio (WWR)       1091 non-null   int64  
 6   Total Window Area                1091 non-null   float64
 7   Smart control of lighting system 1091 non-null   int64  
 8   PM 10  level                     1091 non-null   int64  
 9   Area served by lighting          1091 non-null   int64  
 10  LUX level                        1091 non-null   int64  
 11  Visual Satisfaction              1091 non-null   float64
dtypes: float64(3), int64(7), object(2)
memory usage: 102.4+ KB
```

[ ]:

[16]:
```python
print(data1['Building'].unique().tolist())
```

```
['07', 'F4', '03', 'F2', '02', '04', '06', '05', '01', 'F1', 'F6', '08', 'F3',
'F5']
```

[ ]:

[17]:
```python
data1 = data1.drop(columns = ['Building'])
```

[ ]:

[18]:
```python
#Checking descriptive columns

tex_columns = data1.columns[(data1.dtypes =='object').values].tolist()
tex_columns
```

[18]: ['Building Type']

[19]:
```python
data1.head(2)
```

[19]:
```
   EmpID Building Type  Building Location  Gross Floor Area  \
0      1        Office                NaN              7632   
1      2        Office                NaN              7632   

   Window to Wall Ratio (WWR)  Total Window Area  \
```

```
0                          50                45.0
1                          50                45.0

   Smart control of lighting system  PM 10  level   Area served by lighting  \
0                                 1             37                       40
1                                 1             37                       40

   LUX level  Visual Satisfaction
0        450                 3.05
1        450                 2.95
```

[20]:
```
df1=data1
```

[21]:
```
print(df1['Building Type'].unique().tolist())
```

```
['Office', 'Factory']
```

[22]:
```
#print(df1['Building'].unique().tolist())
```

[23]:
```
df1.corr()['Visual Satisfaction']
```

[23]:
```
EmpID                              -0.176501
Building Location                        NaN
Gross Floor Area                   -0.178820
Window to Wall Ratio (WWR)          0.469265
Total Window Area                   0.224169
Smart control of lighting system    0.240502
PM 10  level                       -0.580440
Area served by lighting             0.303341
LUX level                           0.020538
Visual Satisfaction                 1.000000
Name: Visual Satisfaction, dtype: float64
```

[ ]:
```

```

[24]:
```
df3 = df1.copy()
```

[25]:
```
df3 = df3.drop(columns = ['EmpID','Building Location'])
```

2 one-hot encoding

[26]:
```
df4 = pd.get_dummies(df3)
```

[27]:
```
df4.shape
```

[27]:
```
(1091, 10)
```

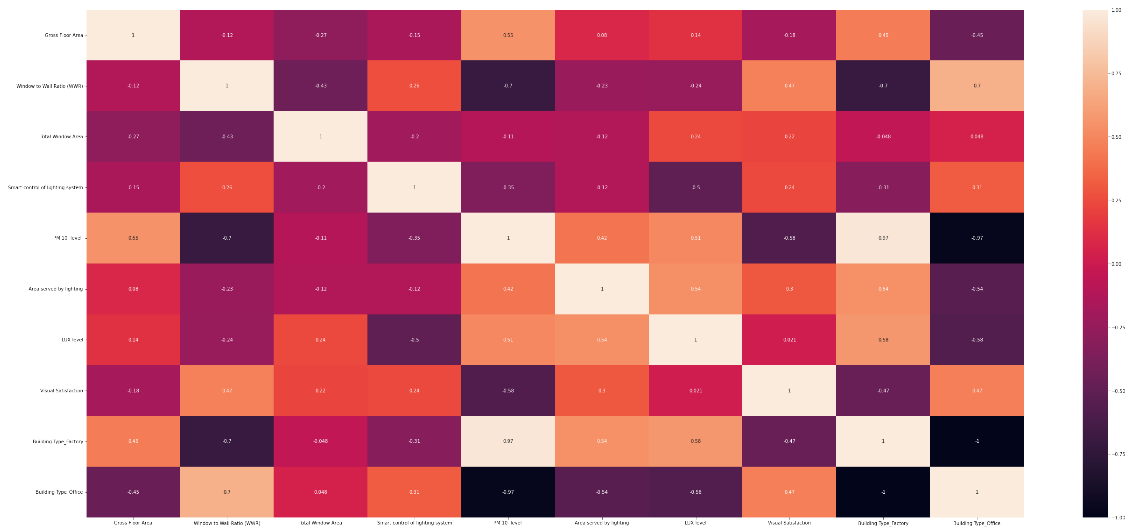
[28]:
```
#correlation of the variables to the Thermal satisfaction

df4.corr()['Visual Satisfaction']
```

[28]:
```
Gross Floor Area                     -0.178820
Window to Wall Ratio (WWR)            0.469265
Total Window Area                     0.224169
Smart control of lighting system      0.240502
PM 10  level                         -0.580440
Area served by lighting               0.303341
LUX level                             0.020538
Visual Satisfaction                   1.000000
Building Type_Factory                -0.466265
Building Type_Office                  0.466265
Name: Visual Satisfaction, dtype: float64
```

[29]:
```
df4.head(2)
```

[29]:
```
   Gross Floor Area  Window to Wall Ratio (WWR)  Total Window Area  \
0              7632                          50               45.0
1              7632                          50               45.0

   Smart control of lighting system  PM 10  level   Area served by lighting  \
0                                 1            37                        40
1                                 1            37                        40

   LUX level  Visual Satisfaction  Building Type_Factory  Building Type_Office
0        450                 3.05                      0                     1
1        450                 2.95                      0                     1
```

[30]:
```
# correlation metrix

plt.figure(figsize = (45,20))
sns.heatmap(df4.corr(), annot=True)
plt.show()
```



[ ]:
```

```

[ ]:
```

```

[31]:
```
df4.dtypes
```

[31]:
```
Gross Floor Area                        int64
Window to Wall Ratio (WWR)              int64
Total Window Area                     float64
Smart control of lighting system        int64
PM 10  level                            int64
Area served by lighting                 int64
LUX level                               int64
Visual Satisfaction                   float64
Building Type_Factory                   uint8
Building Type_Office                    uint8
dtype: object
```

[32]:
```
df4.shape
```

[32]:
```
(1091, 10)
```

3 Model Training

[33]:
```
#Independent variables and dependent variables

X = df4.drop(['Visual Satisfaction'], axis=1)# Input features (attributes)
y = df4['Visual Satisfaction'] # Target vector
```

```
print('X shape: {}'.format(np.shape(X)))
print('y shape: {}'.format(np.shape(y)))


#train and test split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25)
```

```
X shape: (1091, 9)
y shape: (1091,)
```

[34]:
```
X_train.shape, X_test.shape
```

[34]:
```
((818, 9), (273, 9))
```

[35]:
```
#Function to return the model name and the accuracy value

def model_acc(model):
    model.fit(X_train, y_train)
    acc = model.score(X_test, y_test)
    print(str(model)+ ' --> ' +str(acc))
```

[36]:
```
#Find the best regression model


#Support Vector Regression

from sklearn.svm import SVR
svma = SVR(kernel = 'rbf')
model_acc(svma)


#LassoRegression

from sklearn.linear_model import Lasso
lasso = Lasso()
model_acc(lasso)

#DecisionTreeRegressor

from sklearn.tree import DecisionTreeRegressor
dt = DecisionTreeRegressor()
model_acc(dt)

#RandomForestRegressor

from sklearn.ensemble import RandomForestRegressor
```

```
rf = RandomForestRegressor()
model_acc(rf)
```

```
SVR() --> 0.23028769240366043
Lasso() --> 0.8914741185160888
DecisionTreeRegressor() --> 0.9571532720499681
RandomForestRegressor() --> 0.95706482145853
```

[ ]:

```

```

[37]:

```
# Calculate the RMSE for each model


from sklearn.metrics import classification_report, confusion_matrix,␣
↪accuracy_score
from sklearn.metrics import mean_squared_error


classifier_rf = rf.fit(X_train,y_train)
y_pred_rf = classifier_rf.predict(X_test)


classifier_dt = dt.fit(X_train,y_train)
y_pred_dt = classifier_dt.predict(X_test)

classifier_ls = lasso.fit(X_train,y_train)
y_pred_ls = classifier_ls.predict(X_test)

classifier_svm = svma.fit(X_train,y_train)
y_pred_svm = classifier_svm.predict(X_test)


# Calculate the RMSE for each model
from sklearn.metrics import mean_squared_error
rf_rmse = mean_squared_error(y_test, y_pred_rf, squared=False)
lasso_rmse = mean_squared_error(y_test, y_pred_ls, squared=False)
dt_rmse = mean_squared_error(y_test, y_pred_dt, squared=False)
svm_rmse = mean_squared_error(y_test, y_pred_svm, squared=False)


print("Random Forest RMSE    :", rf_rmse)
print("Lasso Regression RMSE :", lasso_rmse)
print("Decision Tree RMSE    :", dt_rmse)
print("SVM RMSE              :", svm_rmse)
```

```
Random Forest RMSE    : 0.18540274583988794
Lasso Regression RMSE : 0.29544277191949625
```

```
Decision Tree RMSE    : 0.18563742657899082
SVM RMSE              : 0.7868119073773935
```

[ ]:

```
[38]:
```

```python
from sklearn.metrics import mean_absolute_error

rf_mae = mean_absolute_error(y_test, y_pred_rf)
lasso_mae = mean_absolute_error(y_test, y_pred_ls)
dt_mae = mean_absolute_error(y_test, y_pred_dt)
svm_mae = mean_absolute_error(y_test, y_pred_svm)

print("Random Forest MAE    :", rf_mae)
print("Lasso Regression MAE :", lasso_mae)
print("Decision Tree MAE    :", dt_mae)
print("SVM MAE              :", svm_mae)
```

```
Random Forest MAE    : 0.14177649292978503
Lasso Regression MAE : 0.23219511553417072
Decision Tree MAE    : 0.1417280881988713
SVM MAE              : 0.6436441343016225
```

[ ]:

```
[40]:
```

```python
from sklearn.model_selection import cross_val_score

# Define the models
lasso = Lasso()
dt = DecisionTreeRegressor()
svm = SVR()
rf = RandomForestRegressor()

# Perform cross-validation and print the mean MAE scores
lasso_scores = cross_val_score(lasso, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Lasso Regression MAE (Cross-Validation) :", -np.mean(lasso_scores))

dt_scores = cross_val_score(dt, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Decision Tree MAE (Cross-Validation)    :", -np.mean(dt_scores))

svm_scores = cross_val_score(svm, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("SVM MAE (Cross-Validation)              :", -np.mean(svm_scores))

rf_scores = cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
```

```
print("Random Forest MAE (Cross-Validation)    :", -np.mean(rf_scores))
```

```
Lasso Regression MAE (Cross-Validation) : 0.2258918598965224
Decision Tree MAE (Cross-Validation)    : 0.14296168006309415
SVM MAE (Cross-Validation)              : 0.6905411489249411
Random Forest MAE (Cross-Validation)    : 0.14357278680679547
```

[ ]:

4 Hyperparameter tunning using Randomforest

[41]:
```
#find the best model for this scenario

from sklearn.model_selection import GridSearchCV

parameters = {'n_estimators':[10, 50, 100],
              'criterion':['squared_error','absolute_error','poisson']}

grid_obj = GridSearchCV(estimator=rf, param_grid=parameters)

grid_fit = grid_obj.fit(X_train, y_train)

best_model = grid_fit.best_estimator_

best_model
```

[41]:
```
RandomForestRegressor(criterion='poisson')
```

[42]:
```
#Score accuracy of the best model

best_model.score(X_test, y_test)
```

[42]:
```
0.9574226767589256
```

[43]:
```
X_test.columns
```

[43]:
```
Index(['Gross Floor Area', 'Window to Wall Ratio (WWR)', 'Total Window Area',
       'Smart control of lighting system', 'PM 10  level ',
       'Area served by lighting', 'LUX level', 'Building Type_Factory',
       'Building Type_Office'],
      dtype='object')
```

[ ]:

5 Save the Model as pickle file

[44]:
```
import pickle
with open('Visual.pickle', 'wb') as file:
    pickle.dump(best_model, file)
```

[45]:
```
df4.head(2)
```

[45]:
```
   Gross Floor Area  Window to Wall Ratio (WWR)  Total Window Area  \
0              7632                          50               45.0
1              7632                          50               45.0

   Smart control of lighting system  PM 10  level   Area served by lighting  \
0                                 1            37                         40
1                                 1            37                         40

   LUX level  Visual Satisfaction  Building Type_Factory  Building Type_Office
0        450                 3.05                      0                     1
1        450                 2.95                      0                     1
```

6 Test the prediction

[47]:
```
pred_value = best_model.predict([[7000,45,56,0,30,40,400,0,1]])
pred_value
```

[47]:
```
array([3.02783488])
```

[ ]:

[48]:
```
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.model_selection import cross_val_predict
```

[67]:
```
# Make predictions using the model

Visual_preds = best_model.predict(X_test)
```

[68]:
```
# Evaluate the performance of the model

mse = mean_squared_error(y_test, Visual_preds)
print("Visual-model MSE:", mse)
```

```
Visual-model MSE: 0.034244574260495396
```

[ ]:

[69]:
```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[ ]:
```

```

[70]:
```
# Evaluate Performance
MSE = mean_squared_error(y_test, cv_predictions)
RMSE = np.sqrt(MSE)
MAE = mean_absolute_error(y_test, cv_predictions)
R2 = r2_score(y_test, cv_predictions)
```

[ ]:
```

```

[71]:
```
# Interpret Results

print("Mean Squared Error (MSE):", MSE)
print("Root Mean Squared Error (RMSE):", RMSE)
print("Mean Absolute Error (MAE):", MAE)
print("R-squared (R2) Score:", R2)
```

```
Mean Squared Error (MSE): 0.03984939667533331
Root Mean Squared Error (RMSE): 0.19962313662332157
Mean Absolute Error (MAE): 0.14365919885974818
R-squared (R2) Score: 0.9504540301683716
```

[ ]:
```

```

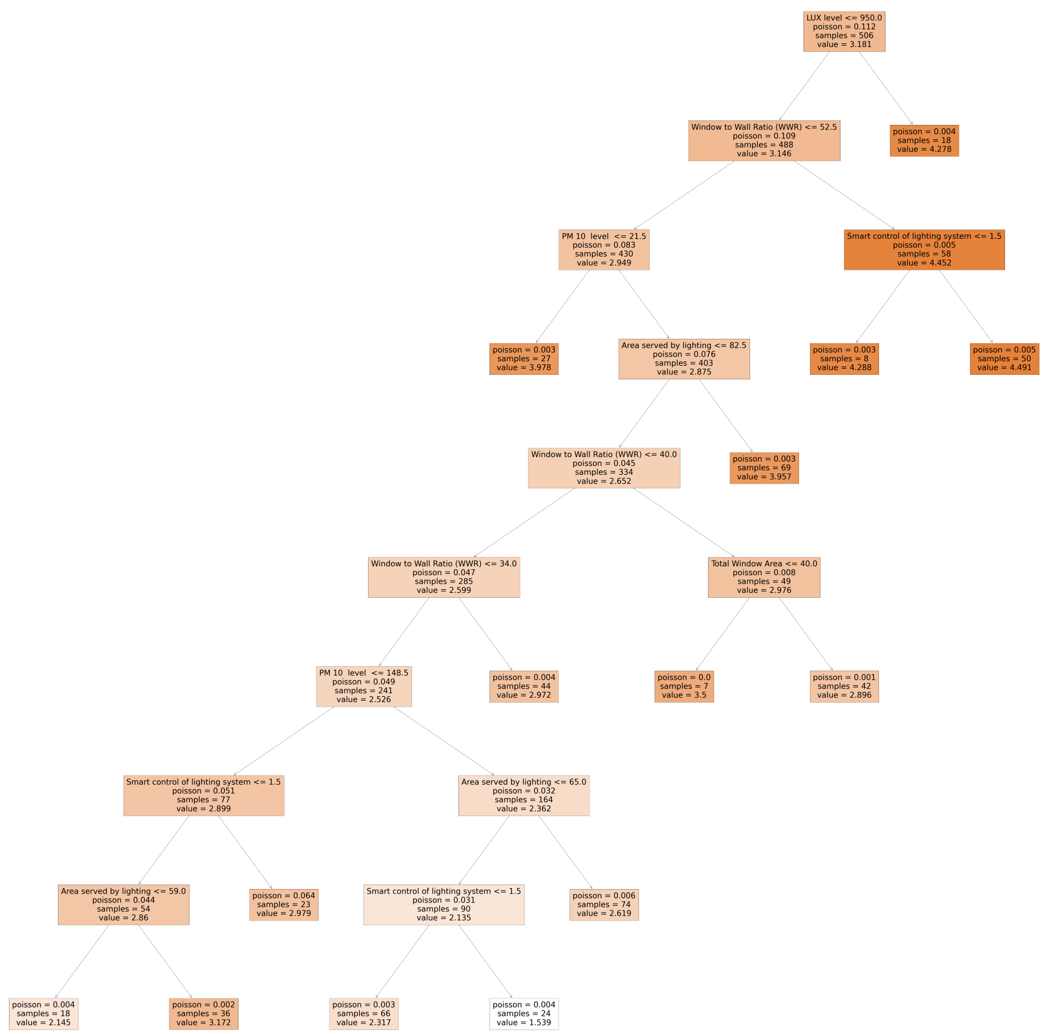
[72]:
```
# calculating VIF for each feature

from statsmodels.stats.outliers_influence import variance_inflation_factor
from statsmodels.tools.tools import add_constant

dfdata = pd.DataFrame(df4[['Gross Floor Area', 'Window to Wall Ratio (WWR)',␣
 ↪'Total Window Area',
       'Smart control of lighting system', 'PM 10  level ',
       'Area served by lighting', 'LUX level']])

X = add_constant(dfdata)
X = dfdata.assign(const=1)

pd.Series([variance_inflation_factor(X.values, i)
               for i in range(X.shape[1])],
              index=X.columns)
```

[72]:
```
Gross Floor Area                      2.571516
Window to Wall Ratio (WWR)           13.089250
Total Window Area                     6.300439
```

```
Smart control of lighting system      1.462450
PM 10  level                          8.315792
Area served by lighting               1.940036
LUX level                             5.868143
const                                61.030025
dtype: float64
```

[73]:

```python
from sklearn.tree import plot_tree

plt.figure(figsize=(100,100))
plot_tree(best_model.estimators_[20], feature_names = df4.
 ↪columns,class_names=['Visual Satisfaction'],filled=True);
```



[ ]:

[56]:
```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

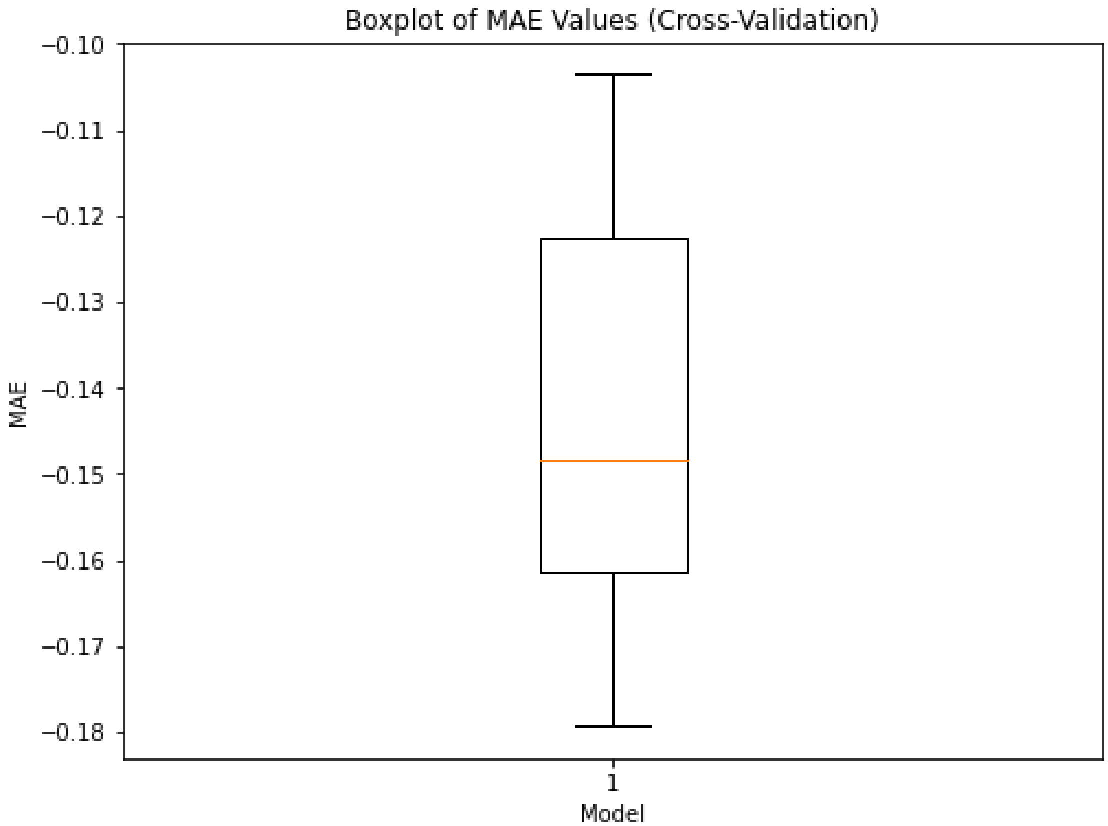
[57]:
```
from sklearn.metrics import mean_absolute_error

rf_mse = mean_squared_error(y_test, cv_predictions)
print("Random Forest MSE     :", rf_mse)

from sklearn.metrics import mean_squared_error
rf_rmse = np.sqrt(mean_squared_error(y_test, cv_predictions))
print("Random Forest RMSE    :", rf_rmse)

rf_mae = mean_absolute_error(y_test, cv_predictions)
print("Random Forest MAE     :", rf_mae)

from sklearn.metrics import r2_score
rf_r2 = r2_score(y_test, cv_predictions)
print("Random Forest R2 Score:", rf_r2)
```

```
Random Forest MSE     : 0.04033858550884948
Random Forest RMSE    : 0.20084468006110962
Random Forest MAE     : 0.14345095644091316
Random Forest R2 Score: 0.9498458067770659
```

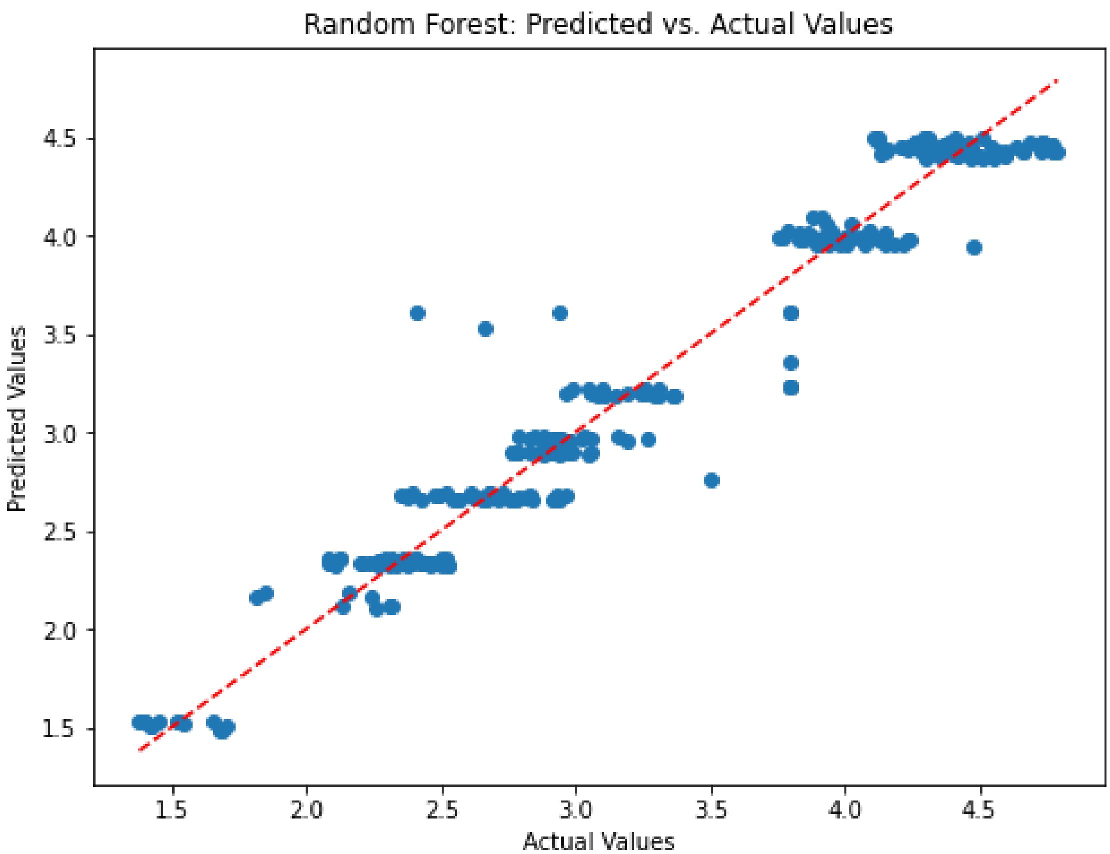
[ ]:

[58]:
```
from sklearn.model_selection import cross_val_score

# Perform cross-validation with 5 folds
rf_cv_mae = -cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')

print("Random Forest MAE (Cross-Validation):", rf_cv_mae.mean())

import matplotlib.pyplot as plt

# MAE values from cross-validation
mae_values = [-rf_cv_mae[fold] for fold in range(5)]

# Boxplot of MAE values
plt.figure(figsize=(8, 6))
plt.boxplot(mae_values)
plt.title('Boxplot of MAE Values (Cross-Validation)')
plt.xlabel('Model')
plt.ylabel('MAE')
plt.show()
```

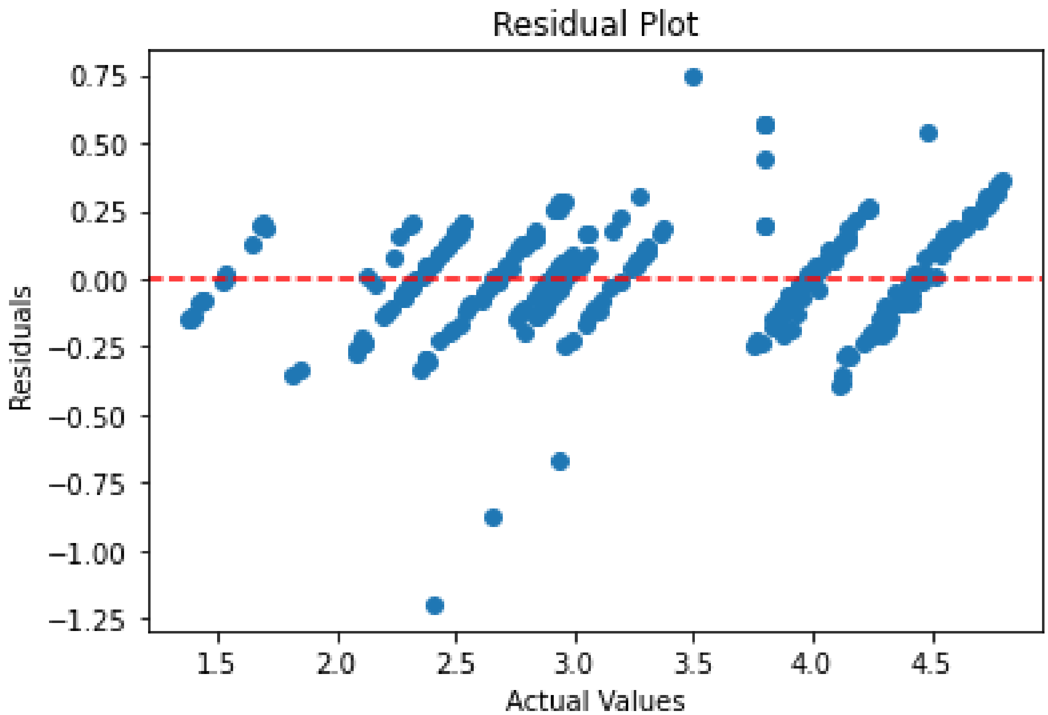
```
Random Forest MAE (Cross-Validation): 0.143047431046207
```



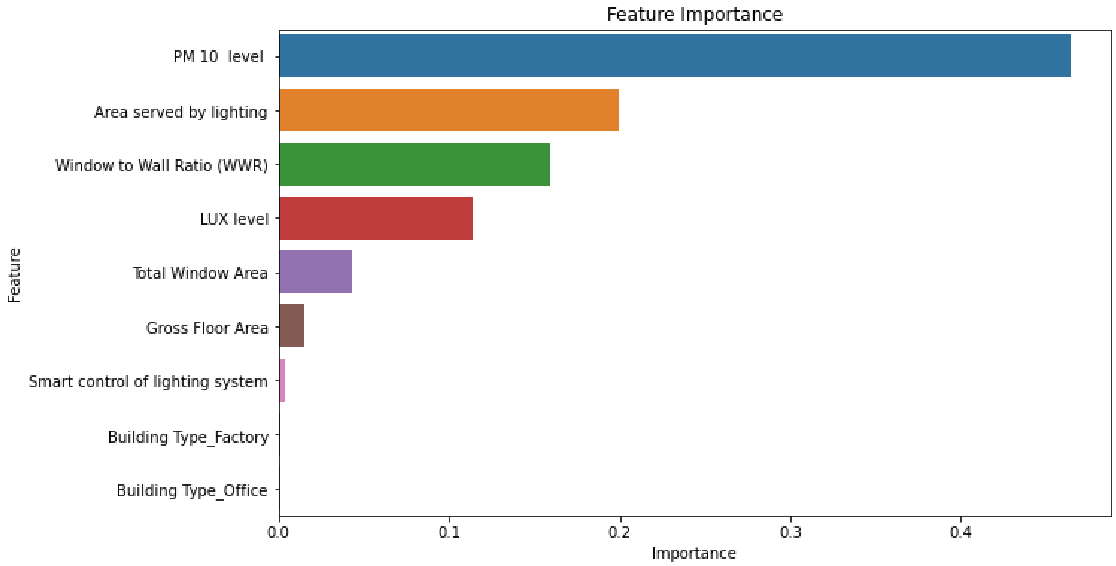
[59]:
```
rf_preds=cv_predictions

plt.figure(figsize=(8, 6))
plt.scatter(y_test, rf_preds)
plt.xlabel("Actual Values")
plt.ylabel("Predicted Values")
plt.title("Random Forest: Predicted vs. Actual Values")
plt.plot([min(y_test), max(y_test)], [min(y_test), max(y_test)], color='red',␣
 ↪linestyle='--')
plt.show()
```



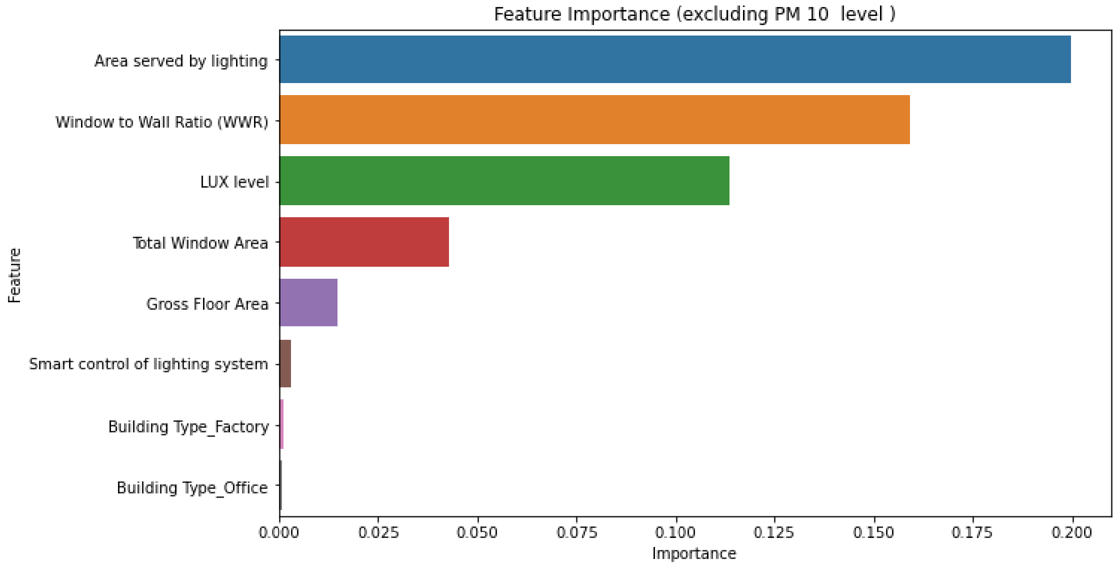
[60]:
```
residuals = y_test - rf_preds
plt.scatter(y_test, residuals)
plt.axhline(y=0, color='r', linestyle='--')
plt.xlabel('Actual Values')
plt.ylabel('Residuals')
plt.title('Residual Plot')
plt.show()
```



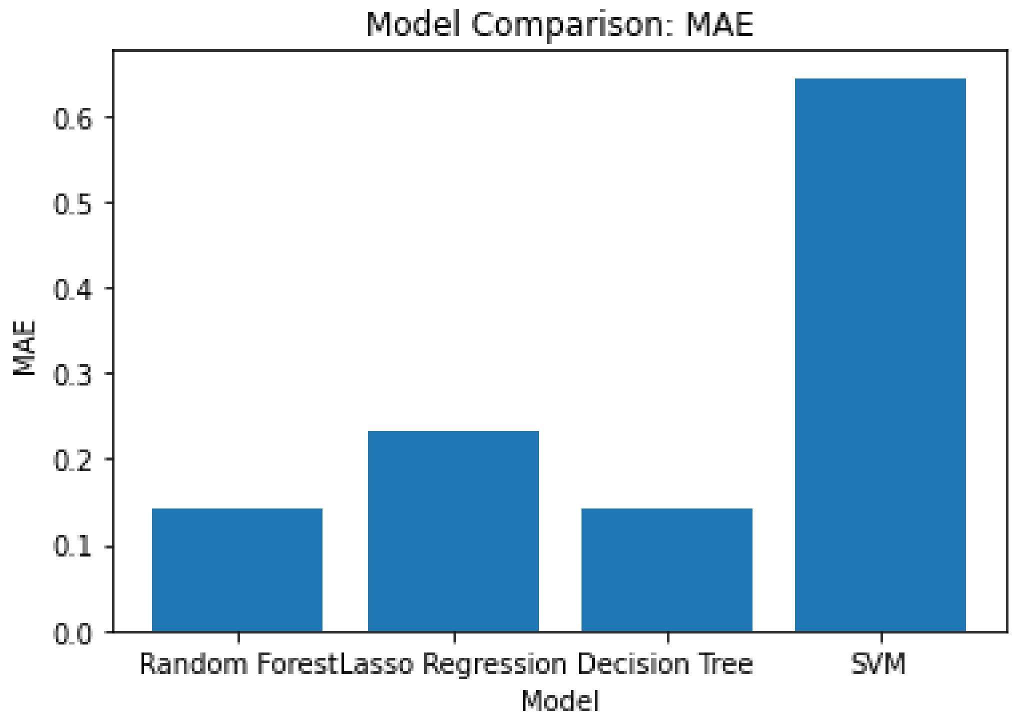
[61]:
```
# Train the random forest model
rf = RandomForestRegressor()
rf.fit(X_train, y_train)

# Calculate feature importance
importance = rf.feature_importances_

# Create a dataframe to store feature importance
feature_importance_df = pd.DataFrame({'Feature': X_train.columns, 'Importance':␣
↪importance})

# Sort the features by importance in descending order
feature_importance_df = feature_importance_df.sort_values('Importance',␣
↪ascending=False)

# Plot the feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df)
plt.title('Feature Importance')
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[62]:
```
# Remove a specific featuShare of PM 10  level from the dataframe
feature_to_remove = 'PM 10  level '
feature_importance_df_filtered =␣
 ↪feature_importance_df[feature_importance_df['Feature'] != feature_to_remove]

# Plot the updated feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df_filtered)
plt.title('Feature Importance (excluding {})'.format(feature_to_remove))
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[64]:
```
model_names = ['Random Forest', 'Lasso Regression', 'Decision Tree', 'SVM']
mae_values = [rf_mae, lasso_mae, dt_mae, svm_mae]

plt.bar(model_names, mae_values)
plt.xlabel('Model')
plt.ylabel('MAE')
plt.title('Model Comparison: MAE')
plt.show()
```


Model Comparison: MAE
MAE
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Random Forest
Lasso Regression
Decision Tree
SVM
Model

[ ]:

[ ]:

os-copy1

July 17, 2023

[2]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import norm

#Converting all Non-Numerical Columns to Numerical
from sklearn.preprocessing import LabelEncoder

from sklearn.metrics import accuracy_score, confusion_matrix, r2_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.tree import DecisionTreeClassifier, export_graphviz
from sklearn.model_selection import train_test_split
import warnings
warnings.filterwarnings('ignore')

%matplotlib inline
```

1 Load the Thermal dataset

[3]:
```
data1 = pd.read_excel("Data.xlsx")
```

[4]:
```
data1.shape
```

[4]: (1091, 29)

[5]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location', 'Gender',
       'Employee Home Town', 'Age', 'Working Hours', 'Is there a blind wall ',
       'Distance between your work desk and the nearest window?',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
```

```
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Thermal Satisfaction', 'Visual Satisfaction', 'Indoor Air Quality',
       'Overall Satisfaction'],
      dtype='object')>
```

[6]:
```
data1 = data1.drop(columns = ['Gender','Employee Home Town','Age','Working␣
↪Hours', 'Is there a blind wall ','Distance between your work desk and the␣
↪nearest window?','Thermal Satisfaction','Visual Satisfaction', 'Indoor Air␣
↪Quality'])
```

[7]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Overall Satisfaction'],
      dtype='object')>
```

[ ]:
```

```

[8]:
```
# duplicate the dataset

data1_copy = data1.copy()
data1_copy.shape
```

[8]: (1091, 20)

[9]:
```
#list(data1.columns.values)
print(data1.columns.tolist)
```

```
<bound method IndexOpsMixin.tolist of Index(['EmpID', 'Building Type',
'Building', 'Building Location',
       'Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Overall Satisfaction'],
      dtype='object')>
```

[10]:
```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1091 entries, 0 to 1090
Data columns (total 20 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   EmpID                             1091 non-null   int64
 1   Building Type                     1091 non-null   object
 2   Building                          1091 non-null   object
 3   Building Location                 0 non-null      float64
 4   Gross Floor Area                  1091 non-null   int64
 5   Wall Insulation U value           1091 non-null   float64
 6   Roof Insulation U value           1091 non-null   float64
 7   Thickness of the Wall Insulation  1091 non-null   int64
 8   Window to Wall Ratio (WWR)        1091 non-null   int64
 9   Glazing U value                   1091 non-null   float64
 10  Total Window Area                 1091 non-null   float64
 11  Share of the area served by AC(%) 1091 non-null   float64
 12  Smart control of HVAC             1091 non-null   int64
 13  Smart control of lighting system  1091 non-null   int64
 14  PM 2.5 level                      1091 non-null   int64
 15  PM 10  level                      1091 non-null   int64
 16  CO2 PPM                           1091 non-null   int64
 17  Area served by lighting           1091 non-null   int64
 18  LUX level                         1091 non-null   int64
 19  Overall Satisfaction              1091 non-null   float64
dtypes: float64(7), int64(11), object(2)
memory usage: 170.6+ KB
```

[ ]:

[11]:
```
print(data1['Building'].unique().tolist())
```

```
['07', 'F4', '03', 'F2', '02', '04', '06', '05', '01', 'F1', 'F6', '08', 'F3',
'F5']
```

[ ]:

[12]:
```
data1 = data1.drop(columns = ['Building'])
```

[ ]:

[13]:
```
#Checking descriptive columns

tex_columns = data1.columns[(data1.dtypes =='object').values].tolist()
tex_columns
```

[13]: ['Building Type']

[14]:
```
data1.head(2)
```

[14]:
```
   EmpID Building Type  Building Location  Gross Floor Area  \
0      1        Office                NaN              7632
1      2        Office                NaN              7632

   Wall Insulation U value  Roof Insulation U value  \
0                     0.26                      0.2
1                     0.26                      0.2

   Thickness of the Wall Insulation  Window to Wall Ratio (WWR)  \
0                                32                          50
1                                32                          50

   Glazing U value  Total Window Area  Share of the area served by AC(%)  \
0             0.48               45.0                               78.0
1             0.48               45.0                               78.0

   Smart control of HVAC   Smart control of lighting system  PM 2.5 level   \
0                      1                                  1             18
1                      1                                  1             18

   PM 10  level   CO2 PPM  Area served by lighting  LUX level  \
0             37      2500                       40        450
1             37      2500                       40        450

   Overall Satisfaction
0                 2.580
1                 3.045
```

[15]:
```
df1=data1
```

[16]:
```
print(df1['Building Type'].unique().tolist())
```

```
['Office', 'Factory']
```

[17]:
```
#print(df1['Building'].unique().tolist())
```

[18]:
```
df1.corr()['Overall Satisfaction']
```

[18]:
```
EmpID                                -0.085151
Building Location                          NaN
Gross Floor Area                      0.045743
Wall Insulation U value              -0.039489
Roof Insulation U value              -0.331803
```

```
Thickness of the Wall Insulation     0.436058
Window to Wall Ratio (WWR)           0.458497
Glazing U value                     -0.043622
Total Window Area                   -0.028734
Share of the area served by AC(%)    0.446629
Smart control of HVAC                0.151362
Smart control of lighting system     0.287155
PM 2.5 level                        -0.411077
PM 10  level                        -0.439911
CO2 PPM                             -0.452158
Area served by lighting              0.292150
LUX level                           -0.099652
Overall Satisfaction                 1.000000
Name: Overall Satisfaction, dtype: float64
```

[ ]:

[19]:
```
df3 = df1.copy()
```

[20]:
```
df3 = df3.drop(columns = ['EmpID','Building Location'])
```

2 one-hot encoding

[21]:
```
df4 = pd.get_dummies(df3)
```

[22]:
```
df4.shape
```

[22]:
```
(1091, 18)
```

[23]:
```
#correlation of the variables to the Thermal satisfaction

df4.corr()['Overall Satisfaction']
```

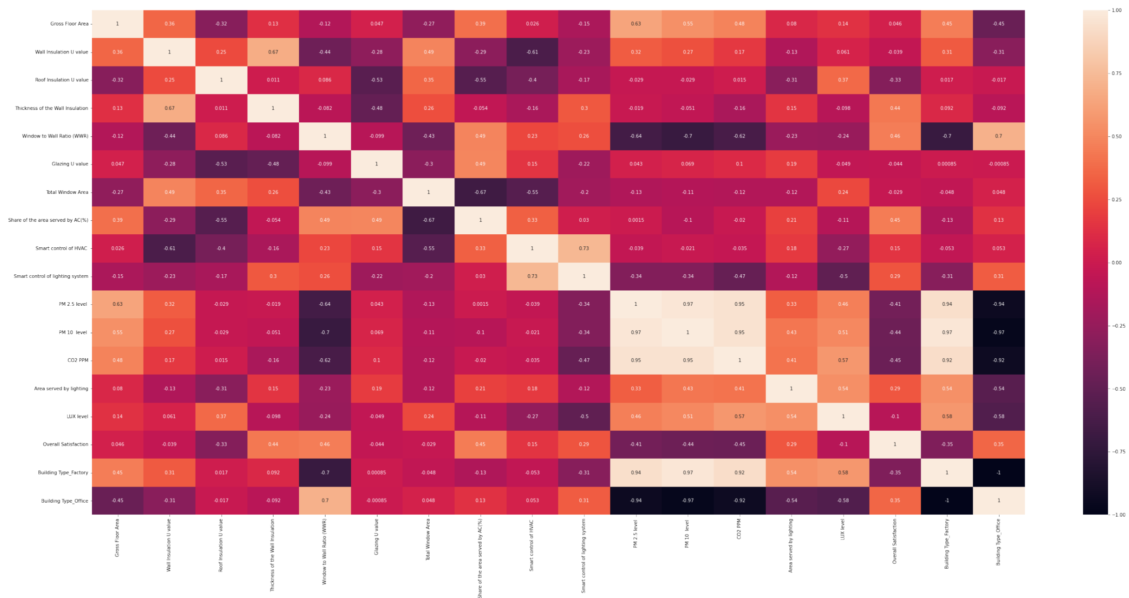
[23]:
```
Gross Floor Area                     0.045743
Wall Insulation U value             -0.039489
Roof Insulation U value             -0.331803
Thickness of the Wall Insulation     0.436058
Window to Wall Ratio (WWR)           0.458497
Glazing U value                     -0.043622
Total Window Area                   -0.028734
Share of the area served by AC(%)    0.446629
Smart control of HVAC                0.151362
Smart control of lighting system     0.287155
PM 2.5 level                        -0.411077
PM 10  level                        -0.439911
CO2 PPM                             -0.452158
```

```
Area served by lighting              0.292150
LUX level                           -0.099652
Overall Satisfaction                 1.000000
Building Type_Factory               -0.353265
Building Type_Office                 0.353265
Name: Overall Satisfaction, dtype: float64
```

[24]:
```
df4.head(2)
```

[24]:
```
   Gross Floor Area  Wall Insulation U value  Roof Insulation U value  \
0              7632                     0.26                      0.2
1              7632                     0.26                      0.2

   Thickness of the Wall Insulation  Window to Wall Ratio (WWR)  \
0                                32                          50
1                                32                          50

   Glazing U value  Total Window Area  Share of the area served by AC(%)  \
0             0.48               45.0                               78.0
1             0.48               45.0                               78.0

   Smart control of HVAC   Smart control of lighting system  PM 2.5 level   \
0                      1                                 1             18
1                      1                                 1             18

   PM 10  level   CO2 PPM  Area served by lighting  LUX level  \
0            37      2500                       40        450
1            37      2500                       40        450

   Overall Satisfaction  Building Type_Factory  Building Type_Office
0                 2.580                      0                     1
1                 3.045                      0                     1
```

[25]:
```
# correlation metrix

plt.figure(figsize = (45,20))
sns.heatmap(df4.corr(), annot=True)
plt.show()
```



[ ]:

[26]:
```
#df4 = df4.drop(columns = ['EmpID','Age','Changed your residence','Hometown␣
↪nature (1-3)'])
```

[27]:
```
df4.dtypes
```

[27]:
```
Gross Floor Area                        int64
Wall Insulation U value               float64
Roof Insulation U value               float64
Thickness of the Wall Insulation        int64
Window to Wall Ratio (WWR)              int64
Glazing U value                       float64
Total Window Area                     float64
Share of the area served by AC(%)     float64
Smart control of HVAC                   int64
Smart control of lighting system        int64
PM 2.5 level                            int64
PM 10  level                            int64
CO2 PPM                                 int64
Area served by lighting                 int64
LUX level                               int64
Overall Satisfaction                  float64
Building Type_Factory                   uint8
Building Type_Office                    uint8
dtype: object
```

[28]:
```python
df4.shape
```

[28]: (1091, 18)

3 Model Training

[29]:
```python
#Independent variables and dependent variables

X = df4.drop(['Overall Satisfaction'], axis=1)# Input features (attributes)
y = df4['Overall Satisfaction'] # Target vector
print('X shape: {}'.format(np.shape(X)))
print('y shape: {}'.format(np.shape(y)))


#train and test split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25)
```

```
X shape: (1091, 17)
y shape: (1091,)
```

[30]:
```python
X_train.shape, X_test.shape
```

[30]: ((818, 17), (273, 17))

[31]:
```python
#Function to return the model name and the accuracy value

def model_acc(model):
    model.fit(X_train, y_train)
    acc = model.score(X_test, y_test)
    print(str(model)+ ' --> ' +str(acc))
```

[ ]:

[ ]:

[ ]:

[32]:
```python
#Find the best regression model


#Support Vector Regression

from sklearn.svm import SVR
svma = SVR(kernel = 'rbf')
model_acc(svma)
```

```
#LassoRegression

from sklearn.linear_model import Lasso
lasso = Lasso()
model_acc(lasso)

#DecisionTreeRegressor

from sklearn.tree import DecisionTreeRegressor
dt = DecisionTreeRegressor()
model_acc(dt)

#RandomForestRegressor

from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
model_acc(rf)
```

```
SVR() --> 0.369444229150685
Lasso() --> 0.7622669520931384
DecisionTreeRegressor() --> 0.8221693983773841
RandomForestRegressor() --> 0.8229534742347487
```

[ ]:

```
```

[33]:

```
# Calculate the RMSE for each model


from sklearn.metrics import classification_report, confusion_matrix,␣
 ↪accuracy_score
from sklearn.metrics import mean_squared_error


classifier_rf = rf.fit(X_train,y_train)
y_pred_rf = classifier_rf.predict(X_test)


classifier_dt = dt.fit(X_train,y_train)
y_pred_dt = classifier_dt.predict(X_test)

classifier_ls = lasso.fit(X_train,y_train)
y_pred_ls = classifier_ls.predict(X_test)

classifier_svm = svma.fit(X_train,y_train)
y_pred_svm = classifier_svm.predict(X_test)
```

```
# Calculate the RMSE for each model
from sklearn.metrics import mean_squared_error
rf_rmse = mean_squared_error(y_test, y_pred_rf, squared=False)
lasso_rmse = mean_squared_error(y_test, y_pred_ls, squared=False)
dt_rmse = mean_squared_error(y_test, y_pred_dt, squared=False)
svm_rmse = mean_squared_error(y_test, y_pred_svm, squared=False)


print("Random Forest RMSE    :", rf_rmse)
print("Lasso Regression RMSE :", lasso_rmse)
print("Decision Tree RMSE    :", dt_rmse)
print("SVM RMSE              :", svm_rmse)
```

```
Random Forest RMSE    : 0.2957631326442059
Lasso Regression RMSE : 0.3439773170365818
Decision Tree RMSE    : 0.297500893529538
SVM RMSE              : 0.5602047384147929
```

[34]:
```
from sklearn.metrics import mean_absolute_error

rf_mae = mean_absolute_error(y_test, y_pred_rf)
lasso_mae = mean_absolute_error(y_test, y_pred_ls)
dt_mae = mean_absolute_error(y_test, y_pred_dt)
svm_mae = mean_absolute_error(y_test, y_pred_svm)
print("Random Forest MAE    :", rf_mae)
print("Lasso Regression MAE :", lasso_mae)
print("Decision Tree MAE    :", dt_mae)
print("SVM MAE              :", svm_mae)
```

```
Random Forest MAE    : 0.17807360349670603
Lasso Regression MAE : 0.2652046663390017
Decision Tree MAE    : 0.17989256215436433
SVM MAE              : 0.43755211255332194
```

[35]:
```
from sklearn.model_selection import cross_val_score

# Define the models
lasso = Lasso()
dt = DecisionTreeRegressor()
svm = SVR()
rf = RandomForestRegressor()
# Perform cross-validation and print the mean MAE scores
lasso_scores = cross_val_score(lasso, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Lasso Regression MAE (Cross-Validation) :", -np.mean(lasso_scores))
```

```
dt_scores = cross_val_score(dt, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Decision Tree MAE (Cross-Validation)    :", -np.mean(dt_scores))

svm_scores = cross_val_score(svm, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("SVM MAE (Cross-Validation)              :", -np.mean(svm_scores))

rf_scores = cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')
print("Random Forest MAE (Cross-Validation)    :", -np.mean(rf_scores))
```

```
Lasso Regression MAE (Cross-Validation) : 0.26321391939656136
Decision Tree MAE (Cross-Validation)    : 0.17764043306215865
SVM MAE (Cross-Validation)              : 0.48740383943117493
Random Forest MAE (Cross-Validation)    : 0.17736623044173644
```

[ ]:

[ ]:

4 Hyperparameter tunning using Randomforest

[36]:
```
#find the best model for this scenario

from sklearn.model_selection import GridSearchCV

parameters = {'n_estimators':[10, 50, 100],
              'criterion':['squared_error','absolute_error','poisson']}

grid_obj = GridSearchCV(estimator=rf, param_grid=parameters)

grid_fit = grid_obj.fit(X_train, y_train)

best_model = grid_fit.best_estimator_

best_model
```

[36]:
```
RandomForestRegressor(criterion='poisson', n_estimators=50)
```

[37]:
```
#Score accuracy of the best model

best_model.score(X_test, y_test)
```

[37]:
```
0.8206133874049724
```

[38]:
```
X_test.columns
```

[38]:
```
Index(['Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Thickness of the Wall Insulation',
       'Window to Wall Ratio (WWR)', 'Glazing U value', 'Total Window Area',
       'Share of the area served by AC(%)', 'Smart control of HVAC ',
       'Smart control of lighting system', 'PM 2.5 level ', 'PM 10  level ',
       'CO2 PPM', 'Area served by lighting', 'LUX level',
       'Building Type_Factory', 'Building Type_Office'],
      dtype='object')
```

[ ]:

5 Save the Model as pickle file

[39]:
```
import pickle
with open('Overall_Satisfaction.pickle', 'wb') as file:
    pickle.dump(best_model, file)
```

[40]:
```
df4.head(2)
```

[40]:
```
   Gross Floor Area  Wall Insulation U value  Roof Insulation U value  \
0              7632                     0.26                      0.2
1              7632                     0.26                      0.2

   Thickness of the Wall Insulation  Window to Wall Ratio (WWR)  \
0                                32                          50
1                                32                          50

   Glazing U value  Total Window Area  Share of the area served by AC(%)  \
0             0.48               45.0                               78.0
1             0.48               45.0                               78.0

   Smart control of HVAC   Smart control of lighting system  PM 2.5 level   \
0                      1                                 1             18
1                      1                                 1             18

   PM 10  level   CO2 PPM  Area served by lighting  LUX level  \
0            37      2500                       40        450
1            37      2500                       40        450

   Overall Satisfaction  Building Type_Factory  Building Type_Office
0                 2.580                      0                     1
1                 3.045                      0                     1
```

6 Test the prediction

[41]:
```
pred_value = best_model.predict([[7000,0.26,0.18,28,50,0.
↪4,50,70,1,0,18,40,3000,50,500,0,1]])
pred_value
```

[41]: array([3.31074854])

[ ]:
```

```

[42]:
```
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.model_selection import cross_val_predict
```

[43]:
```
# Make predictions using the model

OS_preds = best_model.predict(X_test)
```

[44]:
```
# Evaluate the performance of the model

mse = mean_squared_error(y_test, OS_preds)
print("OS-model MSE:", mse)
```

```
OS-model MSE: 0.08928121261002778
```

[ ]:
```

```

[ ]:
```

```

[45]:
```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[ ]:
```

```

[46]:
```
# Evaluate Performance
MSE = mean_squared_error(y_test, cv_predictions)
RMSE = np.sqrt(MSE)
MAE = mean_absolute_error(y_test, cv_predictions)
R2 = r2_score(y_test, cv_predictions)
```

[ ]:
```

```

[47]:
```
# Interpret Results

print("Mean Squared Error (MSE):", MSE)
print("Root Mean Squared Error (RMSE):", RMSE)
print("Mean Absolute Error (MAE):", MAE)
```

```
print("R-squared (R2) Score:", R2)
```

```
Mean Squared Error (MSE): 0.08360001997658392
Root Mean Squared Error (RMSE): 0.28913668044124724
Mean Absolute Error (MAE): 0.17801536004700858
R-squared (R2) Score: 0.8320282178292049
```

[ ]:

```

```

[ ]:

```

```

[48]:

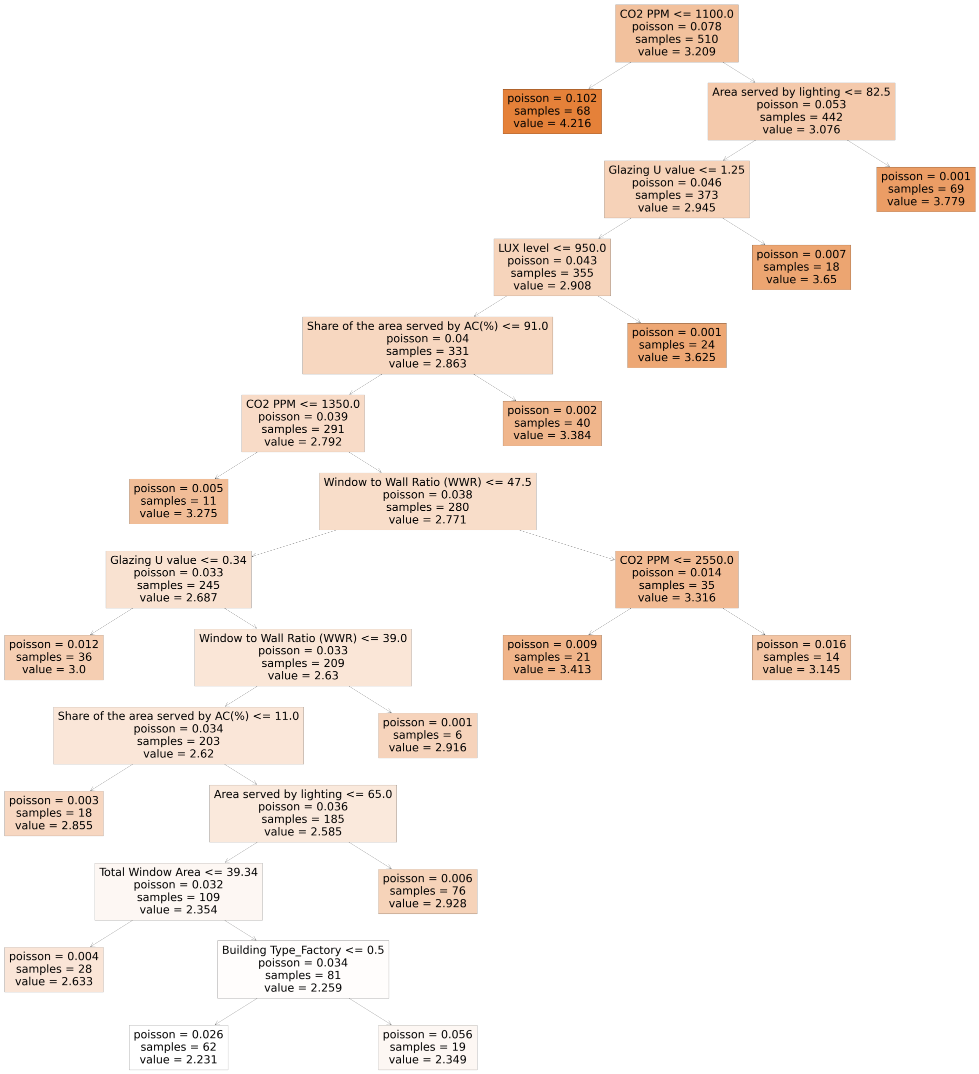
```
# calculating VIF for each feature

from statsmodels.stats.outliers_influence import variance_inflation_factor
from statsmodels.tools.tools import add_constant

dfdata = pd.DataFrame(df4[['Gross Floor Area', 'Wall Insulation U value',
       'Roof Insulation U value', 'Window to Wall Ratio (WWR)', 'Glazing U␣
 ↪value', 'Total Window Area',
        'Smart control of HVAC ', 'Smart control of lighting system', 'Area␣
 ↪served by lighting']])

X = add_constant(dfdata)
X = dfdata.assign(const=1)

pd.Series([variance_inflation_factor(X.values, i)
               for i in range(X.shape[1])],
              index=X.columns)
```

[48]:

```
Gross Floor Area                      3.100576
Wall Insulation U value               5.428763
Roof Insulation U value               2.475935
Window to Wall Ratio (WWR)            2.067447
Glazing U value                       2.173971
Total Window Area                     2.484441
Smart control of HVAC                10.217851
Smart control of lighting system      7.271973
Area served by lighting               1.446202
const                               205.575472
dtype: float64
```

[49]:

```
from sklearn.tree import plot_tree

plt.figure(figsize=(100,100))
plot_tree(best_model.estimators_[20], feature_names = df4.
 ↪columns,class_names=['Overall Satisfaction'],filled=True);
```



[ ]:

```
```

[50]:

```
# Generate predictions using cross-validation
cv_predictions = cross_val_predict(best_model, X_test, y_test, cv=5)
```

[51]:

```
from sklearn.metrics import mean_absolute_error

rf_mse = mean_squared_error(y_test, cv_predictions)
print("Random Forest MSE     :", rf_mse)

from sklearn.metrics import mean_squared_error
rf_rmse = np.sqrt(mean_squared_error(y_test, cv_predictions))
print("Random Forest RMSE    :", rf_rmse)
```

```
rf_mae = mean_absolute_error(y_test, cv_predictions)
print("Random Forest MAE     :", rf_mae)

from sklearn.metrics import r2_score
rf_r2 = r2_score(y_test, cv_predictions)
print("Random Forest R2 Score:", rf_r2)
```

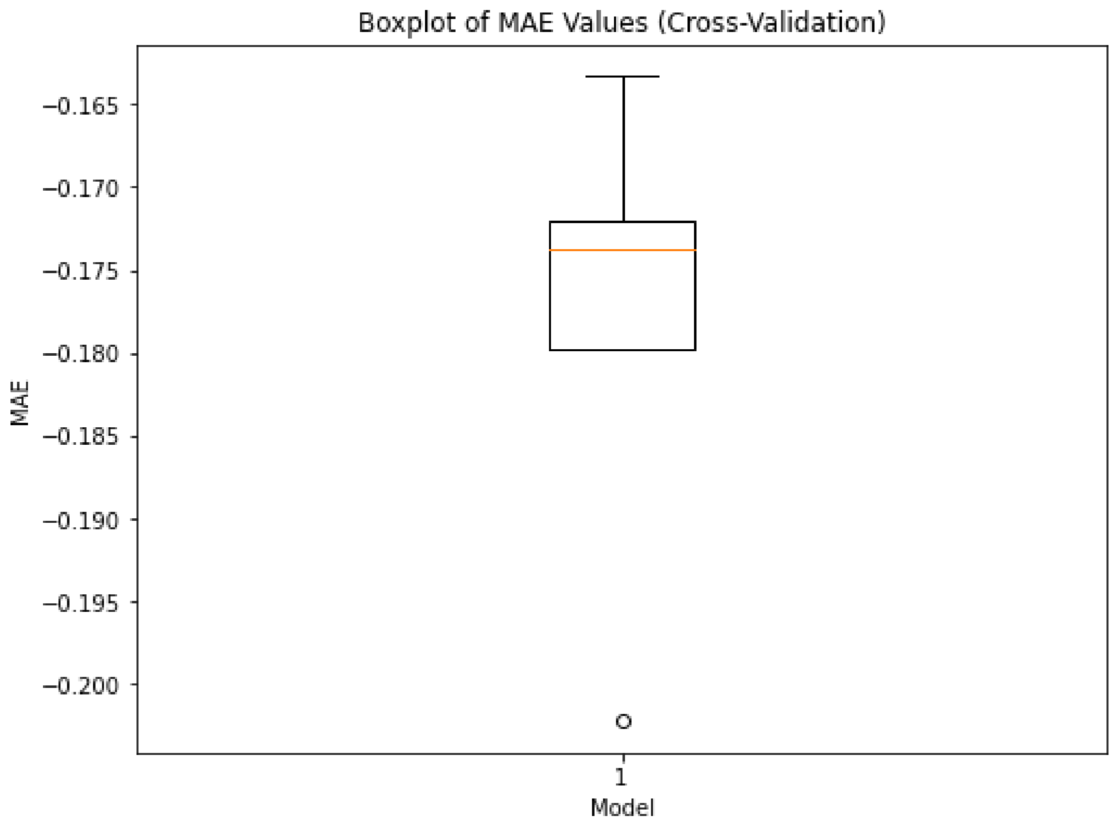
```
Random Forest MSE     : 0.08298220615020027
Random Forest RMSE    : 0.288066322485292
Random Forest MAE     : 0.17902056380094641
Random Forest R2 Score: 0.8332695487463085
```

[ ]:

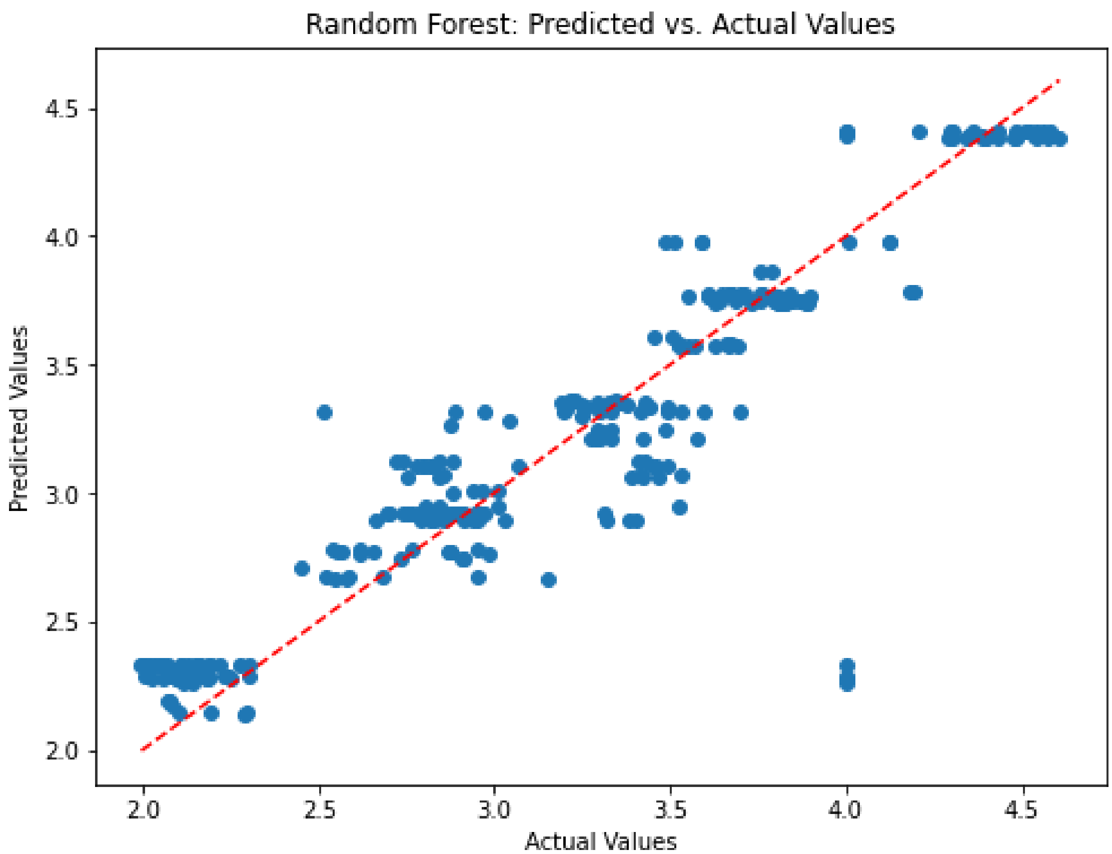
[52]:

```
from sklearn.model_selection import cross_val_score

# Perform cross-validation with 5 folds
rf_cv_mae = -cross_val_score(rf, X_test, y_test, cv=5,␣
 ↪scoring='neg_mean_absolute_error')

print("Random Forest MAE (Cross-Validation):", rf_cv_mae.mean())

import matplotlib.pyplot as plt

# MAE values from cross-validation
mae_values = [-rf_cv_mae[fold] for fold in range(5)]

# Boxplot of MAE values
plt.figure(figsize=(8, 6))
plt.boxplot(mae_values)
plt.title('Boxplot of MAE Values (Cross-Validation)')
plt.xlabel('Model')
plt.ylabel('MAE')
plt.show()
```

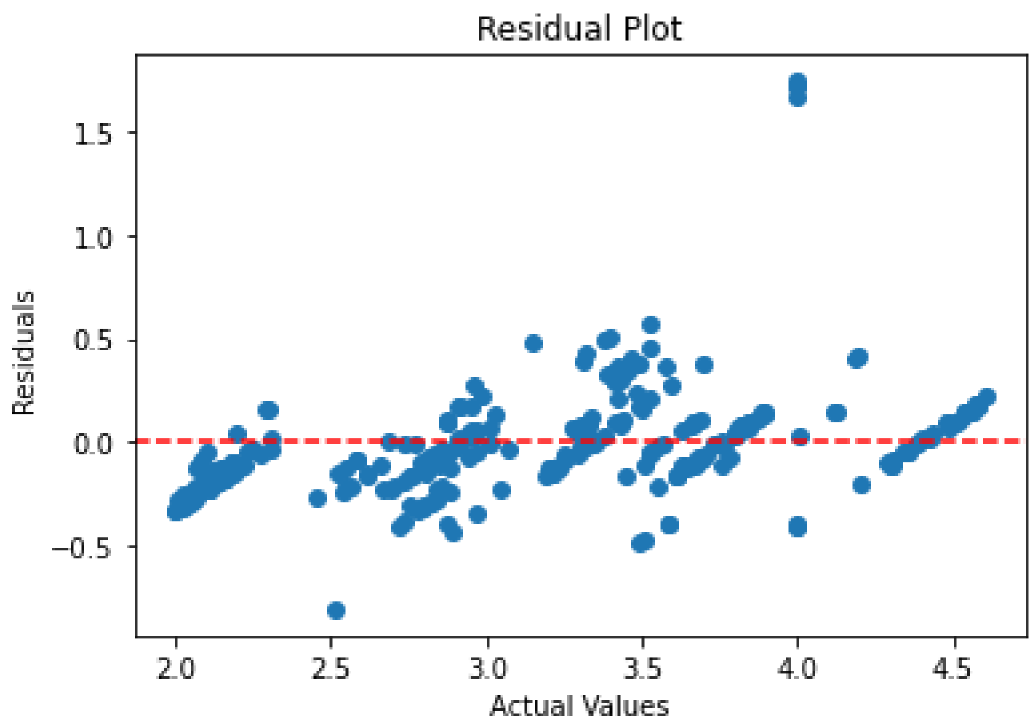
```
Random Forest MAE (Cross-Validation): 0.178211709514956
```



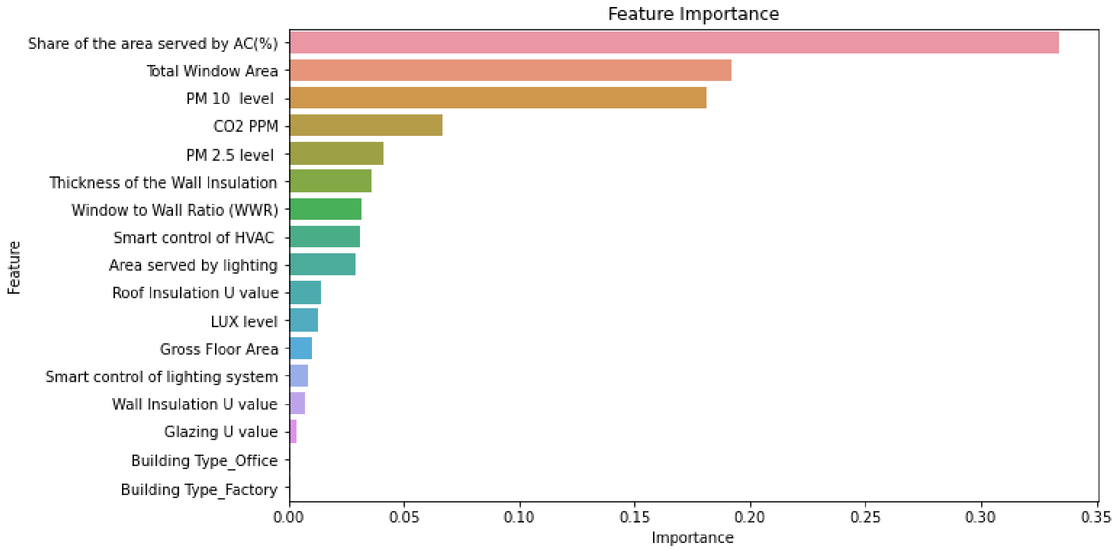
[53]:
```
rf_preds=cv_predictions

plt.figure(figsize=(8, 6))
plt.scatter(y_test, rf_preds)
plt.xlabel("Actual Values")
plt.ylabel("Predicted Values")
plt.title("Random Forest: Predicted vs. Actual Values")
plt.plot([min(y_test), max(y_test)], [min(y_test), max(y_test)], color='red',␣
 ↪linestyle='--')
plt.show()
```



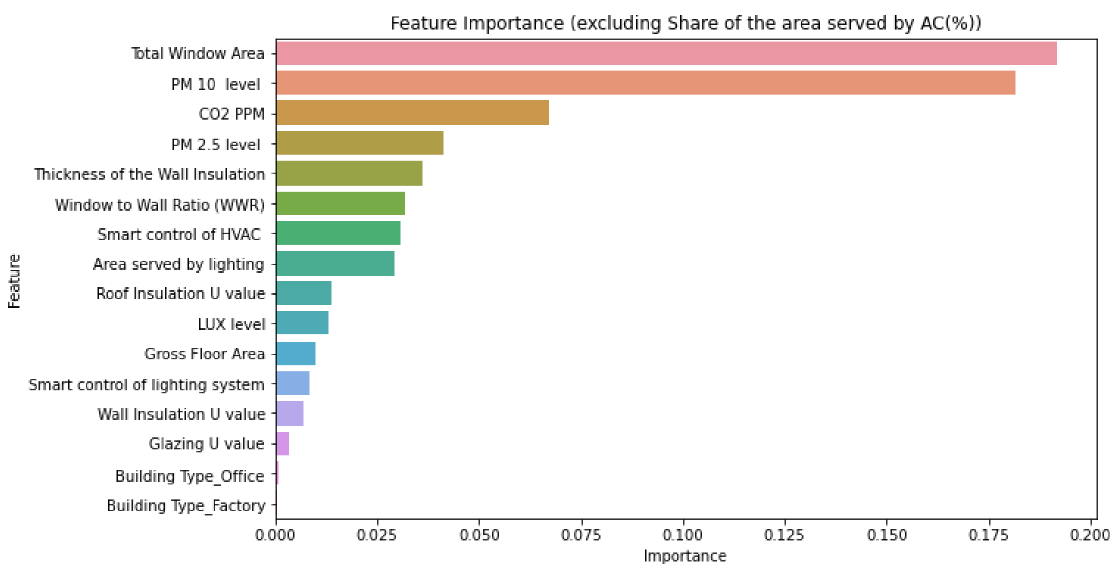
[54]:
```
residuals = y_test - rf_preds
plt.scatter(y_test, residuals)
plt.axhline(y=0, color='r', linestyle='--')
plt.xlabel('Actual Values')
plt.ylabel('Residuals')
plt.title('Residual Plot')
plt.show()
```



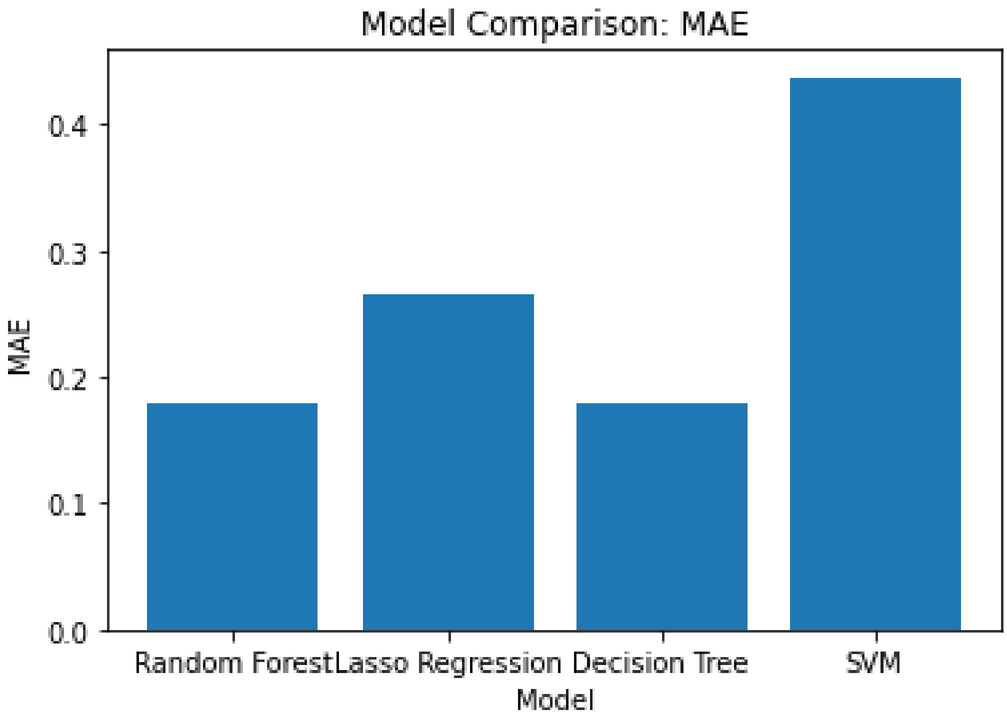
[55]:
```python
# Train the random forest model
rf = RandomForestRegressor()
rf.fit(X_train, y_train)

# Calculate feature importance
importance = rf.feature_importances_

# Create a dataframe to store feature importance
feature_importance_df = pd.DataFrame({'Feature': X_train.columns, 'Importance':␣
↪importance})

# Sort the features by importance in descending order
feature_importance_df = feature_importance_df.sort_values('Importance',␣
↪ascending=False)

# Plot the feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df)
plt.title('Feature Importance')
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[56]:
```python
# Remove a specific featuShare of the area served by AC(%re from the dataframe
feature_to_remove = 'Share of the area served by AC(%)'
feature_importance_df_filtered =␣
↪feature_importance_df[feature_importance_df['Feature'] != feature_to_remove]

# Plot the updated feature importance
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=feature_importance_df_filtered)
plt.title('Feature Importance (excluding {})'.format(feature_to_remove))
plt.xlabel('Importance')
plt.ylabel('Feature')
plt.show()
```



[58]:
```
model_names = ['Random Forest', 'Lasso Regression', 'Decision Tree', 'SVM']
mae_values = [rf_mae, lasso_mae, dt_mae, svm_mae]

plt.bar(model_names, mae_values)
plt.xlabel('Model')
plt.ylabel('MAE')
plt.title('Model Comparison: MAE')
plt.show()
```



[ ]:
```

```

APPENDIX C

```
from flask import Flask, render_template , request
import pickle
import numpy as np
import
sklearn



app = Flask(__name__)


def prediction(lst):
    filename =
'Model/Thermal.pickle'
    with open(filename, 'rb') as file:
        model =
pickle.load(file)
    pred_value = model.predict([lst])
    return pred_value

def
prediction_visual(lst):
    filename = 'Model/Visual.pickle'
    with open(filename, 'rb') as
file:
        model = pickle.load(file)
    pred_value_visual = model.predict([lst])

return pred_value_visual

def prediction_IAQ(lst):
    filename =
'Model/IndoorAirQuality.pickle'
    with open(filename, 'rb') as file:
        model =
pickle.load(file)
    pred_value_IAQ = model.predict([lst])
    return pred_value_IAQ

def
prediction_OS(lst):
    filename = 'Model/Overall_Satisfaction.pickle'
    with
open(filename, 'rb') as file:
        model = pickle.load(file)
    pred_value_OS =
model.predict([lst])
    return pred_value_OS



@app.route('/',
methods=['POST','GET'])

def index():

    pred_value = 0

    if request.method ==
"POST":

        Building_Type = request.form['Building Type']
 #
Employee_Home_Town = request.form['Employee Home Town']
        Gross_Floor_Area =
request.form['Gross Floor Area']
        Wall_Insulation_U_value = request.form['Wall
Insulation U value']
        Roof_Insulation_U_value = request.form['Roof Insulation U
value']
        Thickness_of_the_Wall_Insulation = request.form['Thickness of the Wall
Insulation']
        Window_to_Wall_Ratio_WWR = request.form['Window to Wall Ratio (WWR)']
```

```
     Glazing_U_value = request.form['Glazing U value']
        Total_Window_Area =
request.form['Total Window Area']
        Share_of_the_area_served_by_AC = request.form['Share
of the area served by AC(%)']
        Smart_control_of_HVAC  = request.form['Smart control of
HVAC ']


        feature_list = []


feature_list.append(int(Gross_Floor_Area))

feature_list.append(float(Wall_Insulation_U_value))

feature_list.append(float(Roof_Insulation_U_value))

feature_list.append(int(Thickness_of_the_Wall_Insulation))

feature_list.append(int(Window_to_Wall_Ratio_WWR))

feature_list.append(float(Glazing_U_value))

feature_list.append(float(Total_Window_Area))

feature_list.append(float(Share_of_the_area_served_by_AC))

feature_list.append(int(Smart_control_of_HVAC))



        Building_Type_list =
['Office','Factory']
     #   Employee_Home_Town_list = ['Dry','Intermediate','Wet']




     def traverse_list(lst, value):
            for item in lst:
                if item ==
value:
                    feature_list.append(1)
                else:

feature_list.append(0)

        traverse_list(Building_Type_list, Building_Type)

#    traverse_list(Employee_Home_Town_list, Employee_Home_Town)
    #
traverse_list(Building_list, Building)
    #    traverse_list(Hometown_nature_list,
Hometown_nature)
    #    traverse_list(Satisfaction_level_of_your_workplace_in_general_list,
Satisfaction_level_of_your_workplace_in_general)
    #
traverse_list(Distance_between_your_work_desk_and_the_nearest_window_list,
Distance_between_your_work_desk_and_the_nearest_window)

        pred_value =
prediction(feature_list)

    return render_template('index.html',
pred_value=pred_value)






@app.route('/Visual', methods=['POST','GET'])
def
index1():
```

```
    pred_value_visual = 0

    if request.method == "POST":


Building_Type = request.form['Building Type']
 #       Employee_Home_Town =
request.form['Employee Home Town']
        Gross_Floor_Area = request.form['Gross Floor
Area']
        Window_to_Wall_Ratio_WWR = request.form['Window to Wall Ratio (WWR)']

Total_Window_Area = request.form['Total Window Area']
        Smart_control_of_lighting_system
= request.form['Smart control of lighting system']
        PM_10_level  = request.form['PM 10
level']
        Area_served_by_lighting  = request.form['Area served by lighting']

LUX_level  = request.form['LUX level']


        feature_list = []


feature_list.append(int(Gross_Floor_Area))

feature_list.append(int(Window_to_Wall_Ratio_WWR))

feature_list.append(float(Total_Window_Area))

feature_list.append(int(Smart_control_of_lighting_system))

feature_list.append(int(PM_10_level))

feature_list.append(int(Area_served_by_lighting))

feature_list.append(int(LUX_level))



        Building_Type_list =
['Office','Factory']
#        Employee_Home_Town_list = ['Dry','Intermediate','Wet']




     def traverse_list(lst, value):
            for item in lst:
                if item ==
value:
                    feature_list.append(1)
                else:

feature_list.append(0)

        traverse_list(Building_Type_list, Building_Type)

#    traverse_list(Employee_Home_Town_list, Employee_Home_Town)
    #
traverse_list(Building_list, Building)
    #    traverse_list(Hometown_nature_list,
Hometown_nature)
    #    traverse_list(Satisfaction_level_of_your_workplace_in_general_list,
Satisfaction_level_of_your_workplace_in_general)
    #
traverse_list(Distance_between_your_work_desk_and_the_nearest_window_list,
Distance_between_your_work_desk_and_the_nearest_window)

        pred_value_visual =
prediction_visual(feature_list)

    return render_template('index.html',
```

```
pred_value_visual=pred_value_visual)







@app.route('/IAQ',
methods=['POST','GET'])
def index2():

    pred_value_IAQ = 0

    if request.method ==
"POST":

        Building_Type = request.form['Building Type']
    #
Employee_Home_Town = request.form['Employee Home Town']
        Gross_Floor_Area =
request.form['Gross Floor Area']
        Window_to_Wall_Ratio_WWR = request.form['Window to
Wall Ratio (WWR)']
        Total_Window_Area = request.form['Total Window Area']

PM_2_5_level  = request.form['PM 2.5 level']
        PM_10_level  = request.form['PM 10
level']
        CO2_PPM  = request.form['CO2 PPM']


        feature_list = []


      feature_list.append(int(Gross_Floor_Area))

feature_list.append(int(Window_to_Wall_Ratio_WWR))

feature_list.append(float(Total_Window_Area))
        feature_list.append(int(PM_2_5_level))

       feature_list.append(int(PM_10_level))
        feature_list.append(int(CO2_PPM))




        Building_Type_list = ['Office','Factory']
    #    Employee_Home_Town_list =
['Dry','Intermediate','Wet']



        def traverse_list(lst, value):
            for
item in lst:
                if item == value:
                    feature_list.append(1)

             else:
                    feature_list.append(0)


traverse_list(Building_Type_list, Building_Type)
    #
traverse_list(Employee_Home_Town_list, Employee_Home_Town)
    #
traverse_list(Building_list, Building)
    #    traverse_list(Hometown_nature_list,
Hometown_nature)
    #    traverse_list(Satisfaction_level_of_your_workplace_in_general_list,
Satisfaction_level_of_your_workplace_in_general)
    #
traverse_list(Distance_between_your_work_desk_and_the_nearest_window_list,
Distance_between_your_work_desk_and_the_nearest_window)
```

```
        pred_value_IAQ =
prediction_IAQ(feature_list)

    return render_template('index.html',
pred_value_IAQ=pred_value_IAQ)




@app.route('/OS', methods=['POST','GET'])

def
index3():

    pred_value_OS = 0

    if request.method == "POST":


Building_Type = request.form['Building Type']
    #    Employee_Home_Town =
request.form['Employee Home Town']
        Gross_Floor_Area = request.form['Gross Floor
Area']
        Wall_Insulation_U_value = request.form['Wall Insulation U value']

Roof_Insulation_U_value = request.form['Roof Insulation U value']

Thickness_of_the_Wall_Insulation = request.form['Thickness of the Wall Insulation']

Window_to_Wall_Ratio_WWR = request.form['Window to Wall Ratio (WWR)']
        Glazing_U_value
= request.form['Glazing U value']
        Total_Window_Area = request.form['Total Window
Area']
        Share_of_the_area_served_by_AC = request.form['Share of the area served by
AC(%)']
        Smart_control_of_HVAC  = request.form['Smart control of HVAC ']


Smart_control_of_lighting_system = request.form['Smart control of lighting system']


PM_2_5_level  = request.form['PM 2.5 level']
        PM_10_level  = request.form['PM 10
level']
        CO2_PPM  = request.form['CO2 PPM']

        Area_served_by_lighting  =
request.form['Area served by lighting']
        LUX_level  = request.form['LUX level']





        feature_list = []


feature_list.append(int(Gross_Floor_Area))

feature_list.append(float(Wall_Insulation_U_value))

feature_list.append(float(Roof_Insulation_U_value))

feature_list.append(int(Thickness_of_the_Wall_Insulation))

feature_list.append(int(Window_to_Wall_Ratio_WWR))

feature_list.append(float(Glazing_U_value))

feature_list.append(float(Total_Window_Area))

feature_list.append(float(Share_of_the_area_served_by_AC))
```

```
feature_list.append(int(Smart_control_of_HVAC))


feature_list.append(int(Smart_control_of_lighting_system))


feature_list.append(int(PM_2_5_level))
        feature_list.append(int(PM_10_level))

feature_list.append(int(CO2_PPM))
        feature_list.append(int(Area_served_by_lighting))

      feature_list.append(int(LUX_level))




        Building_Type_list =
['Office','Factory']
    #    Employee_Home_Town_list = ['Dry','Intermediate','Wet']




     def traverse_list(lst, value):
            for item in lst:
                if item ==
value:
                    feature_list.append(1)
                else:

feature_list.append(0)

        traverse_list(Building_Type_list, Building_Type)

#    traverse_list(Employee_Home_Town_list, Employee_Home_Town)
    #
traverse_list(Building_list, Building)
    #    traverse_list(Hometown_nature_list,
Hometown_nature)
    #    traverse_list(Satisfaction_level_of_your_workplace_in_general_list,
Satisfaction_level_of_your_workplace_in_general)
    #
traverse_list(Distance_between_your_work_desk_and_the_nearest_window_list,
Distance_between_your_work_desk_and_the_nearest_window)

        pred_value_OS =
prediction_OS(feature_list)

    return render_template('index.html',
pred_value_OS=pred_value_OS)






if __name__ == '__main__':

app.run(debug=True)
```