THE UNIVERSITY OF
BUCKINGHAM

Data-Driven Framework for Enhancing Student Applications, Acceptances, and Registrations at the University of Buckingham

Mr. Chidera Kenechukwu Onwumbiko

Supervised by: Prof. Harin Sellahewa

ABSTRACT

The goal of this research is to develop a comprehensive data-driven framework to enhance the student application, acceptance, and registration processes at the University of Buckingham. By leveraging modern data science tools such as Python for data preprocessing and Tableau for visualisation and dashboard building, the framework seeks to provide insights that help better understand the university's student enrolment. The project aims to address challenges faced by the university, including increased competition and fluctuating student numbers, by utilising historical data to predict and optimise enrolment outcomes.

Data collected includes attributes such as school/department, academic level, number of applications, acceptances, and registrations, etc., which were thoroughly preprocessed and analysed using advanced techniques like unpivoting, merging, and feature engineering. A predictive model was built to estimate the number of students that need to be accepted to get a target number of student registrations, using a statistical method where the predicted student acceptances are calculated based on the target student registrations divided by the probability of student registrations. This model was implemented in Tableau using parameters and calculated fields, allowing for dynamic adjustment and scenario testing.

The model's performance was evaluated using root mean square error (RMSE), ensuring that it was both accurate and reliable for decision-making purposes. RMSE provided a measure of the model's predictive accuracy. Furthermore, the interactive visualisations created with Tableau help university administrators and stakeholders easily interpret complex trends in student enrolment. This dashboard provides insights into student applications, acceptances, and registrations.

By integrating data-driven decision support with predictive modelling, this framework serves as a crucial tool for the University of Buckingham to navigate the increasingly competitive landscape of higher education enrolment. The insights gained can help develop targeted recruitment strategies and improve the overall efficiency of the admission processes.

Key contributions of this research include the development of a predictive model, a deeper understanding of enrolment dynamics through exploratory data analysis (EDA), and the creation of intuitive visualisations that support strategic decisions. The proposed data-driven approach ultimately aims to contribute to the university's ability to attract, engage, and retain students effectively, ensuring sustainable growth in an increasingly challenging environment.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION

1.1 Background

The rapid growth of data availability in educational institutions has created opportunities to improve decision-making processes, particularly for student enrolment, retention, and engagement. Universities in the United Kingdom, like the University of Buckingham, face increasing competition in attracting and retaining students. The ability to leverage data to understand trends, predict outcomes, and effectively strategise can make a substantial difference. This project uses a data-driven approach to enhance the application, acceptance, and registration processes, focusing on optimising decision-making based on empirical evidence.

In recent years, universities have faced significant challenges due to fluctuating student numbers, both domestic and international. According to recent reports, top-grade universities in the UK have continued to attract a large share of students, leading to an unequal distribution of student populations across institutions. Data from the Universities and Colleges Admissions Service (UCAS) reveals that prestigious universities are absorbing a greater number of students compared to others, which presents challenges for the Universities of Buckingham to find effective ways to boost their application and acceptance rates (BBC News, 2024).

Further complicating the landscape, there has been a fall in the overall number of students accepted into universities in the UK. This trend indicates increased competition among institutions and emphasises the need for effective recruitment strategies based on accurate data insights to maintain or grow enrolment figures (BBC News, 2024). Additionally, a record number of UK students have been heading for university, driven by increased government initiatives and a cultural shift toward higher education as a necessary step for career development (BBC News, 2024). However, this influx has not been evenly distributed, and universities that do not adapt quickly risk falling behind.

International admissions have also been a critical aspect of the higher education landscape. The UK government has been reviewing international admissions policies, which could affect the number of overseas students and impact institutions that rely heavily on this demographic (BBC News, 2024). Although there has been a recent rise in overseas student applications, universities must be prepared to navigate these uncertain conditions and ensure that they maintain a steady flow of international enrolments (BBC News, 2024). On the other hand, recent reports have also highlighted a drop in foreign student visa applications, pointing to potential challenges for UK universities in maintaining their international student numbers (BBC News, 2024).

These dynamics highlight the importance of adopting data-driven approaches for enhancing student applications, acceptances, and registrations. With modern data science tools, universities can identify key trends, target prospective students more effectively, and optimise their marketing and engagement strategies. This project aims to address these challenges by building a robust framework that utilises predictive

modelling and data visualisation to support the University of Buckingham in navigating the increasingly complex landscape of higher education enrolment.

1.2 Problem Statement

The University of Buckingham, recognised for its excellent teaching quality and outstanding student satisfaction, faces a critical need to optimise its student application, acceptance, and registration processes to maintain its competitive edge in the higher education sector. Despite the university's reputation, existing methods for evaluating student data and managing enrolment processes lack the analytical depth needed for strategic decision-making.

Furthermore, the fluctuating number of student applications, combined with increased competition from top-grade institutions, has underscored the importance of utilising data science and data analysis to make informed decisions. There is a need for a system that not only captures and analyses data but also provides actionable insights that help in boosting application and acceptance rates. Such a system can also ensure that marketing and recruitment efforts are focused where they will have the greatest impact.

This project aims to address these challenges by developing a robust data-driven framework that utilises data science and data analysis techniques to provide deeper insights to improve student enrolment outcomes. By leveraging predictive modelling and data visualisation, the University of Buckingham can better understand their student enrolment processes and enhance its ability to meet enrolment targets effectively.

1.3 Objectives

The main objectives of this research are:

- **Collect and Preprocessing Enrolment Data**: To extract and preprocess historical data on student enrolment data from university records.
- **Analyse Enrolment Data**: To conduct a thorough analysis of historical student application, acceptance, and registration data to identify key trends and patterns. This includes identifying peak periods for applications, acceptances or registrations and evaluating departmental differences in acceptance and registration rates.
- **Visualise Data Trends**: To create interactive data visualisations using Tableau that provide actionable insights into the student enrolments. These visualisations are designed to help the university administrators and stakeholders quickly understand complex data and identify areas that require strategic attention. Visualising key metrics such as application trends, acceptance rates, and registration growth will help in effectively communicating insights to administrators and stakeholders.
- **Build a Predictive Model**: To develop a predictive model that can accurately forecast student acceptance and registration rates. This model will be used to simulate different scenarios and predict how much student applications that need to be accepted to improve target student registration numbers. The goal

is to provide the university with a tool that supports data-driven decision-making.

1.4 Research Questions

The following research questions guide the study:

- **How can data science be utilised to improve student application, acceptance, and registration rates effectively?** This question aims to understand how different data science methodologies, such as machine learning or statistical modelling, can be applied to predict enrolment metrics and improve forecasting accuracy for the University of Buckingham.
- **How can visualisations help administrators and stakeholders understand complex enrolment trends and make informed decisions?** This question focuses on exploring the role of interactive data visualisations in simplifying and conveying insights to university administrators and stakeholders, enabling them to grasp complex data patterns and use those insights to drive decisions related to student enrolment.
- **In what ways can data-driven decision-making improve the effectiveness of student recruitment strategies at the University of Buckingham?** This question aims to understand how leveraging data-driven strategies can help the university enhance its student recruitment processes, leading to increased application, acceptance, and registration rates.

1.5 Scope

This research focuses on developing a data-driven framework to improve student applications, acceptances, and registrations at the University of Buckingham. The scope includes several key components:

- **Data Collection**: This research uses historical records of student applications, acceptances, and registrations, which include various attributes such as school/department, academic level, number of applications, acceptances, and registrations, etc. The study focuses on both domestic UK-based and international students.
- **Data Preprocessing**: Data preprocessing activities include handling missing values, data cleaning, unpivoting the datasets, merging the datasets, and feature engineering. These steps are essential to ensure the quality and consistency of the datasets, making it suitable for analysis and modelling.
- **Exploratory Data Analysis (EDA)**: The scope includes conducting EDA to uncover trends, relationships, and patterns within the datasets. This involves analysing application trends over time and department-specific variations.
- **Predictive Modelling**: The research develops a predictive model to estimate the number of student acceptances needed to reach a target number of registrations. The predictive model is built using statistical methods, with parameters and calculated fields in Tableau to simulate different scenarios. The goal is to accurately forecast enrolment outcomes to support strategic decision-making.

- **Data Visualisation**: Tableau is used to create an interactive dashboard that present insights into application, acceptance, and registration trends. Tableau is also used to display the predictive model using parameters and calculated fields. This visualisation helps university administrators and stakeholders understand complex data intuitively and make informed decisions regarding enrolment strategies.
- **Stakeholder Engagement**: The scope also includes providing actionable insights to university administrators and stakeholders including admissions teams, marketing departments, and senior management. The insights generated from the predictive model and visualisation are intended to guide recruitment efforts, resource allocation, and policymaking.
- **Evaluation Metrics**: The model developed in this research is evaluated using root mean square error (RMSE). This metric is used to assess the accuracy and reliability of the predictive model in forecasting student enrolment outcomes.

1.6 Structure of the Report

The report is organised into five chapters, each providing a detailed examination of various aspects of the project:

- **Introduction**: This chapter introduces the context of the research, detailing the challenges in student enrolment at the University of Buckingham and the need for a data-driven solution. It presents the problem statement, research objectives, research questions, and scope of the study, setting the foundation for the subsequent chapters.
- **Literature Review**: The literature review discusses existing research and practices in data science applications within higher education. It covers predictive analytics for student enrolment, strategies to enhance student recruitment, and the role of data visualisation. The chapter also identifies key challenges and gaps in current research, which this study aims to address.
- **Methodology**: This chapter outlines the research approach, and the tools used to conduct the study. It provides a detailed description of the data collection process, data preprocessing techniques, and the predictive modelling approach. It also explains the use of Tableau for data visualisation and describes the metrics used to evaluate model performance.
- **Experiment, Results, and Discussion**: This chapter presents the findings from the analysis, including the results of predictive modelling and insights gained from data visualisations. It discusses the implementations of these results for improving student enrolment strategies at the University of Buckingham.
- **Conclusion and Future Work**: The final chapter summarises the key findings of the research and its contributions to improving student enrolment processes. It discusses the limitations of the current study and suggests areas for future research, such as incorporating additional data sources and expanding the predictive models to enhance their accuracy and applicability.

CHAPTER TWO: LITERATURE REVIEW

2.1 Overview

The application of data science in higher education has been gaining momentum, reshaping how universities manage recruitment, admissions, and student success initiatives. This growing reliance on data is largely driven by the need for more efficient and informed decision-making processes that improve student outcomes and institutional performance. The increasing availability of educational data, coupled with advancements in big data analytics and machine learning, has enabled universities to adopt more sophisticated methods for understanding student behaviour and optimising enrolment strategies.

This literature review explores existing approaches to data-driven enrolment strategies, predictive modelling, and the use of data visualisation tools in higher education. It also highlights the challenges and opportunities that come with implementing data-driven practices. By reviewing studies conducted by various educational institutions, this chapter sheds light on how predictive modelling, data visualisation, and business intelligence systems are transforming the educational landscape. Furthermore, it identifies existing research gaps and discusses how the University of Buckingham can leverage these techniques to enhance student applications, acceptances, and registrations.

The growing body of literature review reveals an agreement on the benefits of using data-driven approaches to improve educational processes, such as recruitment, student engagement, and retention. However, several challenges remain, particularly regarding the ethical use of data, data quality, and the potential biases introduced by machine learning models. Despite these challenges, the potential for data science to enhance institutional efficiency, personalise student experiences, and improve overall academic success remains important. This literature review aims to establish a foundation for developing a comprehensive framework that the University of Buckingham can use to optimise their enrolment strategies and better navigate the complexities of higher education.

2.2 Data-Driven Strategies in Higher Education

The use of data to improve quality and efficiency in higher education is not a new concept, but recent advancements in data science have significantly enhanced the ability to gather, process, and interpret large volumes of data. The integration of data science into higher education allows institutions to make informed decisions, optimise operations, and ultimately improve student outcomes. According to McFarland et al. (2021), the field of education data science has evolved considerably over the past decades, particularly with the emergence of big data analytics that enables educational institutions to gain valuable insights into student behaviours and performance (McFarland et al., 2021).

Data-driven strategies in higher education typically involve the use of business intelligence systems, predictive analytics, and strategic interventions to enhance both institutional efficiency and the student experience. For instance, the University of Nicosia implemented a business intelligence system aimed at improving decision-

making processes by leveraging historical student data (Daniel, 2015). This system incorporated descriptive, diagnostic, and predictive analytics to support recruitment and resource allocation, leading to more targeted and efficient strategies for improving enrolment and academic success.

Furthermore, Cope and Kalantzis (2016) discussed the implications of big data for learning, assessment, and research, emphasising the potential of data analytics to reshape the educational landscape through personalised learning and improved assessment methods (Cope and Kalantzis, 2016). Personalised learning, driven by data insights, allows institutions to provide strategic educational experiences that meet the unique needs of individual students, thereby significantly improving student satisfaction and success rates.

A study by Daniel (2015) also points out the transformative impact that business intelligence and data analytics have on higher education institutions. They underline how leveraging predictive and prescriptive analytics can help universities manage recruitment strategies more effectively, improve student retention, and optimise resource allocation. By using these data-driven strategies, institutions can ensure that they focus their efforts on high-impact areas that yield better outcomes for both students and the university.

Moreover, educational big data enables universities to collect and analyse diverse datasets to make informed decisions on student recruitment, admissions, and resource allocation. For example, by examining historical data on student performance, universities can identify patterns that indicate students at risk of dropping out and provide targeted support to mitigate these risks (Baker and Yacef, 2009). The use of data-driven decision-making frameworks has been proven effective in enhancing institutional efficiency, as demonstrated by various universities that have successfully integrated these strategies into their operational processes.

The adoption of data-driven strategies also involves predictive modelling to forecast student behaviour. Predictive analytics can help determine which students are most likely to accept an offer, enrol, or face academic challenges. The University of Nicosia's case highlights how business intelligence systems can enhance enrolment predictions and refine recruitment strategies, ultimately improving the effectiveness of institutional practices (Daniel, 2015). Business intelligence systems integrate different datasets and apply advanced analytics to identify patterns, predict outcomes, and make informed decisions.

In addition to improving operational efficiency, data-driven strategies can foster a culture of continuous improvement within educational institutions. For instance, by using analytics to monitor student engagement, universities can identify the most effective teaching methods and adapt their practices accordingly. Cope and Kalantzis (2016) also mention that big data enables universities to identify at-risk students early and provide strategic interventions, which can significantly improve student retention and success rates. When integrated into an institutional strategy, predictive modelling can support improved outcomes for both students and the institution (Cope and Kalantzis, 2016).

While the benefits of adopting data-driven strategies in higher education are evident, there are also challenges that need to be addressed. These challenges include ensuring data quality, managing privacy concerns, and addressing potential biases in predictive models. Universities must invest in data governance frameworks that ensure data accuracy, protect student privacy, and promote fairness in decision-making processes. By addressing these challenges, institutions can fully realize the potential of data-driven strategies to transform higher education and enhance student outcomes.

Data-Driven Strategies for Higher Education: Data-driven strategies have become an integral part of modern higher education, providing institutions with the tools they need to improve student recruitment, engagement, and retention. By leveraging business intelligence systems, predictive analytics, and personalised learning approaches, universities can enhance both institutional efficiency and student success. The University of Buckingham can benefit from the insights gained from this research, which continues the work of using data-driven strategies to develop a comprehensive data-driven framework that optimises enrolment processes.

2.3 Predictive Modelling for Enrolment

Predictive modelling has become a powerful tool for universities seeking to forecast student behaviours and optimise enrolment strategies. It enables institutions to make data-driven decisions regarding student recruitment, engagement, and retention, ultimately improving institutional efficiency and student outcomes. This section explores various approaches to predictive modelling in higher education and their impact on enrolment management.

The growing complexity of student behaviour and the competitive nature of higher education necessitate the use of advanced predictive modelling techniques. According to McFarland et al. (2021), predictive modelling methods such as logistic regression, decision trees, and ensemble methods like random forests have been successfully applied in higher education to predict student enrolment outcomes (McFarland et al., 2021). These models help identify significant factors that influence student decisions, including socio-economic background, academic performance, and engagement metrics. By leveraging these insights, institutions can design more effective recruitment campaigns and allocate resources where they are most needed.

Business intelligence systems are important to predictive modelling efforts. For example, the University of Nicosia used a business intelligence system to enhance decision-making processes related to student enrolment. This system incorporated predictive analytics to forecast enrolment numbers and refine recruitment strategies, ultimately resulting in more targeted marketing efforts and resource optimisation (Daniel, 2015). The predictive model, in this research, will be invaluable for estimating the number of acceptances required to achieve specific registration targets and for simulating different enrolment scenarios.

Another example of predictive modelling in higher education is the use of machine learning algorithms to improve the accuracy of predictions related to student success and retention. Baker and Yacef (2009) highlighted the use of educational data mining

to identify students at risk of dropping out, allowing institutions to intervene early and provide support (Baker and Yacef, 2009). Machine learning models, such as support vector machines and neural networks, have been particularly effective in detecting non-linear relationships between student characteristics and their likelihood of success, which are often missed by traditional statistical methods.

Early warning systems represent a crucial application of predictive modelling in higher education. Cope and Kalantzis (2016) discussed the role of predictive analytics in monitoring student engagement and academic performance to identify those who may require additional support (Cope and Kalantzis, 2016). By implementing early warning systems, universities can proactively address issues before they lead to student dropout. These systems leverage predictive models to detect warning signs, such as declining grades or poor attendance, and generate alerts that prompt intervention. This proactive approach has been shown to significantly improve student retention and success rates.

Predictive modelling is also used extensively in optimising the admissions process. By analysing historical application data, universities can determine which prospective students are most likely to accept an offer and ultimately enrol. This information is used to personalise communication strategies, ensuring that recruitment messages resonate with the intended audience. A study conducted by Green et al. (2009) demonstrated that predictive models could accurately forecast enrolment based on applicant characteristics, allowing institutions to develop targeted outreach efforts and enhance recruitment efficiency (Green et al., 2009). Personalising recruitment efforts in this manner not only improves conversion rates but also enhances the overall student experience by creating a more individualised interaction.

The use of a predictive model to optimise enrolment strategies is demonstrated in this research for the University of Buckingham. A predictive model will be employed to estimate and improve acceptance and registration rates, enabling the university administrators and stakeholders to make informed decisions about resource allocation and marketing strategies. By using a parameter-driven approach in Tableau, the university administrators and stakeholders will be able to simulate different registration targets and determine the number of acceptances needed to meet those targets. This interactive modelling approach will not only improve decision-making but also allow for greater flexibility in planning recruitment initiatives.

However, there are challenges associated with implementing predictive modelling in higher education. One of the most significant challenges is ensuring the quality and completeness of the data used to train predictive models. Inaccurate or incomplete data can lead to biased predictions, which may have negative consequences for both students and institutions. McFarland et al. (2021) emphasised the importance of data validation and quality assurance processes to ensure that predictive models produce reliable and accurate results (McFarland et al., 2021). Moreover, there are ethical considerations related to privacy, bias, and fairness in predictive modelling. Institutions must develop robust data governance frameworks to address these issues and ensure that predictive models are used responsibly.

Another challenge is the interpretability of complex predictive models. While machine learning algorithms such as neural networks offer high accuracy, they are often considered "black boxes" due to their lack of transparency. This makes it difficult for decision-makers to understand how predictions are made, which can hinder trust in the model's outcomes. To address this issue, universities are increasingly adopting interpretable machine learning techniques and providing visual explanations of model predictions. By improving the transparency of predictive models, institutions can build trust among their stakeholders and ensure that data-driven decisions are well-supported.

Despite these challenges, the benefits of predictive modelling for enrolment management are significant. Predictive models provide universities with actionable insights that can be used to optimise recruitment campaigns, allocate resources more effectively, and improve student success. By shifting from reactive to proactive decision-making, institutions can anticipate challenges, identify opportunities, and implement strategies that enhance the overall student experience.

Predictive modelling is a valuable tool for higher education institutions aiming to optimise enrolment strategies and improve student outcomes. By leveraging historical data and applying machine learning techniques, universities can gain a deeper understanding of the factors that influence student behaviour and use these insights to make informed, data-driven decisions.

2.4 Visualisation as a Communication Tool

Data visualisation has become an indispensable tool for enhancing data-driven decision-making in higher education. Effective visualisation techniques help transform complex datasets into accessible, intuitive insights, allowing stakeholders, ranging from university administrators to faculty members to understand trends, make informed decisions, and take strategic actions without requiring deep technical knowledge. In the context of student enrolment and recruitment, visualisation tools like Tableau have been widely adopted to present enrolment data in an easily understandable format, facilitating the exploration of trends and insights.

Visualisation serves as an essential bridge between data analysis and strategic decision-making. According to Daniel (2015), visualisations can provide actionable insights into enrolment data, enabling university leaders to quickly identify trends, correlations, and outliers that might not be immediately visible in raw data (Daniel, 2015). For instance, visual elements such as tree maps, stacked bar charts, and line graphs can effectively represent enrolment metrics, allowing for a more detailed understanding of factors influencing student decisions and outcomes. In this research for the University of Buckingham, visualisations will be used to demonstrate seasonal trends in applications and department-specific variations in acceptance rates, which will support more strategic resource allocation and recruitment planning.

One of the primary advantages of data visualisation in higher education is its ability to enhance communication across different stakeholder groups. A study conducted by Green et al. (2009) found that visualisation methods significantly improve data comprehension, leading to better decision-making outcomes (Green et al., 2009). By

presenting data visually, complex predictive models and analytical results become easier to interpret, enabling stakeholders who may not possess a strong background in data science to understand the insights and take appropriate action. This accessibility of data insights empowers all stakeholders, from faculty to university executives, to engage actively in the decision-making process.

Interactive dashboards are a particularly powerful form of data visualisation, providing a dynamic and user-friendly way to explore and interpret enrolment data. Tableau, a leading data visualisation tool, allows users to create interactive dashboards that can present data at multiple levels of detail. For example, in this research, decision-makers can filter data by school/department, academic level, or admission year to identify specific trends. This research allows stakeholders to simulate different enrolment scenarios by adjusting the target registration numbers parameter, thereby enhancing the university's ability to plan recruitment initiatives more effectively.

Moreover, using visualisation tools to display relationships can play a significant role in identifying patterns and anomalies that may warrant further investigation. By visually representing these relationships, universities can gain a deeper understanding of how different factors interact and influence student behaviour. This type of visual insight is crucial for optimising recruitment efforts, focusing marketing resources on the most promising segments, and making data-informed adjustments to admissions policies.

The use of visualisation also enhances the presentation of predictive modelling outputs. Predictive models are often complex, relying on multiple variables to forecast enrolment outcomes or identify students at risk of dropping out. By visualising the results of these models, universities can more effectively communicate the underlying insights to stakeholders. In this research, although a machine learning model was not used, visualisation with Tableau is used to show the output of the statistical predictive model. The interactive nature of Tableau allows users to adjust model parameters and immediately visualise the resulting changes, fostering a deeper understanding of the implications of different scenarios.

Another advantage of data visualisation in higher education is its role in fostering transparency and accountability. By making data and analysis results accessible to a wider audience, visualisations can help ensure that decisions are based on evidence rather than intuition. This transparency not only builds trust among stakeholders but also encourages a culture of continuous improvement, where data-driven insights are used to refine strategies and enhance institutional performance.

Despite its many advantages, there are challenges associated with the use of data visualisation in higher education. One of the challenges is ensuring that visualisations are designed in a way that effectively communicates the intended message. Poorly designed visualisations can lead to misinterpretation of data, resulting in misguided decisions. To address this challenge, it is essential for institutions to invest in training staff on best practices for data visualisation, including selecting appropriate chart types, using effective labelling, and avoiding visual clutter. Moreover, institutions must ensure that the data being visualised is accurate and up to date, as errors in the underlying data can undermine the credibility of the visualisations and the decisions made based on them.

Data visualisation is a critical tool for enhancing data-driven decision-making in higher education. By transforming complex data into accessible insights, visualisations empower stakeholders to make informed decisions that enhance student recruitment, engagement, and success. The use of interactive dashboards and other visualisation techniques allow universities to communicate insights effectively, foster transparency, and support strategic planning.

2.5 Application of Big Data in Higher Education

The application of big data in higher education has revolutionised the way universities approach recruitment, retention, and academic success. Big data refers to large volumes of complex data that, when analysed, can reveal patterns, trends, and associations, particularly regarding student behaviours and institutional performance. In the context of higher education, big data enables universities to make informed decisions, enhance student experiences, and optimise resource allocation.

One of the applications of big data in higher education is in analysing student engagement. By collecting data from various sources, including learning management systems (LMS), social media, and institutional databases, universities can gain insights into how students interact with course materials, participate in discussions, and access resources. According to Daniel (2015), big data analytics provides real-time insights into student engagement, enabling educators to identify students who are struggling and intervene before their performance declines further (Daniel, 2015). This proactive approach helps improve student retention and ensures that students receive the support they need to succeed.

Another significant application of big data is in predicting academic success. Predictive analytics, powered by big data, allows institutions to forecast student outcomes based on historical data, such as grades, attendance, and engagement metrics. Cope and Kalantzis (2016) highlights the potential of big data to predict which students are likely to perform well and which may need additional support (Cope and Kalantzis, 2016). By identifying at-risk students early, universities can implement targeted interventions, such as tutoring or mentoring, to improve academic outcomes. This ability to predict academic success and provide personalised support is a key benefit of big data in education.

Big data also plays a critical role in enhancing personalised learning experiences. Personalised learning involves tailoring educational content, pacing, and support to meet the unique needs of each student. By analysing data on student preferences, learning styles, and performance, institutions can create customised learning pathways that maximise student engagement and achievement. According to Chen et al. (2020), big data is instrumental in creating adaptive learning environments that adjust in real-time based on student interactions, making education more responsive and effective (Chen et al., 2020).

In addition to improving learning outcomes, big data analytics can be used to optimise recruitment strategies. By analysing data from past application years, universities can identify the characteristics of applicants who are most likely to accept an offer and enrol. This information enables institutions to develop targeted marketing campaigns

and allocate recruitment resources more efficiently. For example, by understanding which demographics are most responsive to specific marketing messages, universities can focus their efforts on high-impact areas, thereby improving conversion rates and reducing recruitment costs (McFarland et al., 2021).

Resource allocation is another area where big data has had a transformative impact. Universities must allocate resources, such as faculty, facilities, and financial aid, in a way that maximises efficiency and supports student success. Big data analytics provides insights into which programs are most in demand, where student needs are greatest, and how resources can be distributed to meet these needs effectively. By analysing data on enrolment trends, course demand, and student feedback, institutions can make data-driven decisions about hiring faculty, expanding programs, or investing in new technologies (University of Nicosia, 2014).

However, the use of big data in higher education also presents several challenges. One of the concerns is data privacy. The collection and analysis of large volumes of student data raise ethical questions about how that data is used, who has access to it, and how student privacy is protected. Cope and Kalantzis (2016) emphasised the need for institutions to develop robust data governance frameworks that ensure data is used ethically and responsibly (Cope and Kalantzis, 2016). This includes obtaining informed consent from students, anonymising data where possible, and implementing security measures to prevent unauthorised access.

Another challenge is the integration of data from multiple sources. Higher education institutions often collect data from a variety of systems, including LMS, student information systems, and financial databases. Integrating these diverse data sources into a cohesive dataset that can be analysed effectively is a complex task that requires significant technical expertise and infrastructure. According to McFarland et al. (2021), successful integration of big data requires collaboration across departments, as well as investment in data management tools and technologies (McFarland et al., 2021).

Additionally, the sheer volume and complexity of big data can be overwhelming for institutions that lack the necessary expertise to analyse it effectively. Universities need skilled data scientists and analysts who can extract meaningful insights from big data and communicate those insights to decision-makers. This skills gap is a significant barrier to the effective use of big data in higher education, and institutions must invest in training and hiring qualified professionals to bridge this gap.

Despite these challenges, the opportunities presented by big data in higher education are substantial. By leveraging big data, universities can make more informed decisions, provide personalised support to students, and optimise their operations to enhance both student success and institutional efficiency.

2.6 Opportunities and Challenges in Data-Driven Enrolment

The adoption of data-driven strategies in higher education has opened numerous opportunities for improving student recruitment, retention, and overall institutional performance. However, it also presents a variety of challenges that institutions must address to fully leverage the potential of data analytics. This section explores both the

opportunities and challenges associated with data-driven enrolment in higher education.

2.6.1 Opportunities

One of the most significant opportunities presented by data-driven enrolment is the ability to personalise recruitment and retention strategies. By using predictive models, universities can identify the students who are most likely to succeed and tailor their marketing and recruitment campaigns to these individuals. This targeted approach not only increases the likelihood of attracting the right students but also improves the efficiency of recruitment efforts. For example, using predictive analytics to analyse applicant data can help institutions determine which students are most likely to accept offers, allowing them to focus their resources on high-potential candidates (Daniel, 2015).

Data-driven decision-making also enables institutions to proactively address student needs, improving retention and student success. Predictive models can be used to identify at-risk students early, providing an opportunity for timely intervention. This proactive approach has been shown to significantly improve retention rates, as students receive the support they need before their challenges become overwhelming. According to Baker and Yacef (2009), educational data mining can effectively identify students at risk of dropping out, enabling institutions to implement timely support mechanisms (Baker and Yacef, 2009).

Furthermore, data analytics allows universities to optimise resource allocation. By understanding patterns in student applications, enrolment, and retention, institutions can allocate resources more efficiently to areas that require the most attention. Business intelligence systems provide insights that help decision-makers optimise marketing budgets, allocate staff, and ensure that support services are available where they are most needed (McFarland et al., 2021). This level of optimisation is crucial for improving institutional efficiency and ensuring that resources are used effectively to achieve enrolment targets.

The use of interactive data visualisation tools, such as Tableau, also enhances the ability of universities to communicate data insights to stakeholders. Visualisations provide an intuitive way to understand complex data, allowing decision-makers to quickly identify trends, correlations, and areas of concern. This improved communication of data insights fosters a culture of evidence-based decision-making across institutions and encourages collaboration among different departments (Green et al., 2009).

2.6.2 Challenges

While the opportunities presented by data-driven enrolment are substantial, several challenges must be addressed to ensure the successful implementation of these strategies. One of the challenges is ensuring data quality and completeness. Predictive models rely on accurate and comprehensive data, and any gaps or inaccuracies in the data can lead to flawed predictions. According to McFarland et al. (2021), data quality is a critical factor in the success of predictive modelling, and

institutions must invest in data cleaning and validation processes to ensure that the data used is reliable (McFarland et al., 2021).

Privacy and ethical considerations are also major challenges in data-driven enrolment. The use of student data for predictive modelling raises concerns about privacy and the potential misuse of information. Institutions must develop robust data governance frameworks that ensure student data is collected, stored, and used ethically, with appropriate safeguards in place to protect privacy. Cope and Kalantzis (2016) caution against the indiscriminate use of big data in educational contexts without adequate safeguards, highlighting the importance of transparency and fairness in data use (Cope and Kalantzis, 2016). Bias in predictive models is another concern, as biased data can lead to unfair outcomes for certain student groups, particularly those from underrepresented backgrounds.

Another challenge associated with predictive modelling is the interpretability of complex models. Many machine learning algorithms, such as neural networks, offer high accuracy but lack transparency, making it difficult for stakeholders to understand how predictions are made. This "black box" nature of machine learning can hinder trust in the model's outcomes and limit the adoption of data-driven strategies. To address this issue, universities are increasingly adopting interpretable machine learning techniques and providing visual explanations of model predictions to ensure that decision-makers can understand and trust the insights generated by these models (McFarland et al., 2021).

Moreover, the successful implementation of data-driven strategies requires a cultural shift within institutions. Data-driven decision-making must be embraced at all levels of the organisation, from leadership to faculty and staff. Institutions must also invest in the necessary technological infrastructure to support data collection, analysis, and visualisation, which can be a significant financial and logistical challenge.

CHAPTER THREE: METHODOLOGY

3.1 Overview

The methodology chapter provides a comprehensive description of the research approach used to develop the data-driven framework for enhancing student applications, acceptances, and registrations at the University of Buckingham. The chapter is structured to cover the end-to-end process, starting from data collection, preprocessing, modelling, and visualisation, ensuring that every step contributes meaningfully to achieving the research objectives. The methodology includes the following components:

- **Data Collection**: This section outlines the methodology employed to gather historical records from the University of Buckingham concerning student applications, acceptances, and registrations. The original datasets were sourced directly from the university's official records and encompassed key attributes such as school/department, academic level, number of applications, acceptances, and registrations. Additionally, the datasets were acquired in a pivoted format.

- **Data Preprocessing**: Data preprocessing involved cleaning and preparing the raw datasets to ensure quality and consistency. This included handling missing values, unpivoting the datasets, merging the datasets, and feature engineering. The use of the pandas Python library was critical in implementing these preprocessing steps, ensuring the datasets were structured and ready for analysis.

- **Exploratory Data Analysis (EDA)**: EDA was conducted to understand the underlying patterns, distributions, and relationships within the data. By visualising and summarising the data, the research identified key trends in student enrolment.

- **Predictive Modelling**: This methodology focused on developing a predictive model to estimate the number of student applications that should be accepted to optimise registration numbers. The model was designed using statistical methods, where the predicted number of acceptances was calculated by dividing the target registration numbers by the probability of registration. To enable dynamic predictions and scenario analysis, the model was implemented using Tableau's parameter functionality and calculated fields.

- **Evaluation Metrics**: To assess the accuracy and reliability of the predictive model, the evaluation metric root mean square error (RMSE) was used. This metric provided insights into the model's performance, helping to validate the quality of the predictions and guide any necessary improvements.

- **Data Visualisation**: The final step in the methodology involved visualising the results using Tableau. An interactive dashboard was created to represent the findings, allowing university administrators and stakeholders to explore trends and relationships in the data. Visualisations included stacked bar charts, line charts, and tree maps, each designed to provide an intuitive understanding of the enrolment metrics and help the university make informed decisions.

These methodologies are better described below as individual sections in this chapter.

3.2 Data Collection

The data collection process for this research involved gathering comprehensive historical records from the University of Buckingham's student enrolment system. These records include information related to student applications, acceptances, and registrations over multiple academic years. The collected datasets were extensive, encompassing a range of eight datasets in total for each January and September admission year and several attributes. The attributes collected include:

- **Campus**: The specific campus location where students made applications, got accepted, and registered, which helps analyse differences in enrolment trends across the locations. There were two campus locations, which are Buckingham and Crewe.
- **Group**: Group data refers to the classification of students based on program type, which is either non foundation or foundation, enabling targeted analysis of each group.
- **School/Department**: Data collected at the school or departmental level allowed the analysis to assess enrolment trends within specific academic disciplines, identifying which departments attract more students and which require more focused recruitment efforts.
- **Level**: The academic level of the students was gathered to better understand the progression patterns and where challenges in transitioning to the next academic level may exist. There were several academic levels, such as undergraduate, postgraduate research, postgraduate taught, visiting non-degree, etc.
- **Academic Years**: Specific year information was collected for applications, acceptances, and registrations, enabling analysis of trends over time, including seasonal fluctuations and peak enrolment periods. Applications and acceptances had years covering from 2017 – 2023 while registrations had years covering from 2019 – 2022.
- **Number of Applications, Acceptances, and Registrations**: These key metrics formed the basis of the analysis, providing quantitative data on the number of students applying, accepting offers, and registering for courses. These metrics are crucial for understanding the overall efficiency of the enrolment process.
- **Categories**: The categories home, overseas, and unknown were included in the datasets to ensure understanding of the university's abilities to attract UK based students and international students.

3.3 Data Preprocessing

Data preprocessing is a crucial step in the development of any data-driven model, as it ensures that the data is clean, consistent, and suitable for analysis. In this research, data preprocessing was performed using Python's Pandas library, which provided a comprehensive toolkit for handling various data manipulation tasks. These tasks were carried out on copies of the original datasets, as the original datasets were left untouched in accordance with the cookie-cutter data science framework, ensuring that

the data source remains a reliable reference point. The preprocessing steps undertaken included the following:

- **Data Cleaning (Before Merging the Datasets)**: The data cleaning process involved restructuring the original datasets to ensure compatibility with Python for analysis. This included removing aggregated cells and columns, percentage columns, and data specific to the Crewe campus, as the final output of this research focuses solely on the Buckingham campus. Additionally, this process involved addressing inconsistencies within the datasets, such as correcting spelling errors in categorical fields.

- **Handling Missing Values**: Following the data cleaning process, it was observed that data for the "unknown" category was missing for several admission years. To address this, columns were created and filled with zero values, as some admission years contained data for the "unknown" category while others did not. Creating these missing columns and filling them with zero (indicating no intakes for the "unknown" category during those years) was deemed more appropriate than deleting the columns entirely.

- **Unpivoting the Datasets**: To restructure the cleaned datasets for different types of analysis, unpivoting techniques were employed. Unpivoting was used to flatten the admission year data for easier processing, converting it to admission dates for the different January and September datasets. This transformation facilitated generating insights at multiple aggregation levels, such as by admission date, department, or academic level.

- **Merging the Datasets**: The merging process involved combining the eight different datasets for each January and September admission year to create two final datasets: one for application and acceptance data and another for registration data. The distinction between these datasets was due to the format of the original data collected. The registration numbers were already in an unpivoted format and were not differentiated by category, whereas the applications and acceptances data were categorised and presented in a pivoted format. To merge these into a single dataset for predictive modelling, the category data had to be removed from the applications and acceptances dataset. Although this combined dataset was used for building the predictive model, it was not used for creating the interactive dashboard framework for visualisation.

- **Data Cleaning (After Merging the Datasets)**: After merging the final application and acceptance dataset with the registration dataset for the purpose of building the predictive model, inconsistencies were identified in the registration numbers, as some registration numbers were higher than their corresponding application numbers, which is logically impossible and could negatively affect predictive modelling. To address this, a function was developed to replace any registration numbers that exceeded their application numbers with the respective application numbers. This was deemed the most logical solution, as the only way to determine the correct registration numbers would be to recollect the data from the university and cross-check it, which was not feasible at the time due to data sensitivity concerns.

- **Feature Engineering**: Feature engineering was conducted to create new attributes that would potentially enhance the analysis and visualisation within the interactive dashboard framework. For example, attributes such as "month" and "year" were derived from the admissions date column to simplify the analysis and facilitate visualisations that display specific admission periods. Additionally, a new feature called "main level" was created from the admissions level column to streamline the number of academic levels in the datasets, resulting in more comprehensible visuals.
- **Data Validation**: After preprocessing, the final datasets underwent a thorough validation process to ensure that all transformations were accurately implemented and that the datasets were suitable for visualisation and modelling. This process included verifying that missing values were correctly handled, confirming the consistency of data after merging, and ensuring that no information was lost during data cleaning and transformation.

3.4 Predictive Modelling

Predictive modelling, although not the main objective, is a key component of this research, aimed at estimating the number of student applications that should be accepted to achieve a target number of student registrations at the University of Buckingham. The predictive model developed in this study was designed to accomplish this goal, thereby helping in making data-driven decisions regarding student enrolment, targeting recruitment efforts, and optimising resource allocation. It is important to note that the model was built using the available datasets, which were extracted during a period when the number of acceptances had been streamlined. This means that the acceptance numbers in the datasets are not the original values, as they had been reduced due to some students not meeting acceptance conditions or withdrawing their offers. Consequently, this impacts the model's output, but with access to the original acceptance data, the model could provide more accurate predictive analysis. The modelling process involved the following steps:

- **Model Selection and Approach**: The predictive modelling approach used in this research was based on statistical methods combined with machine learning evaluation features. The primary model employed was a statistical estimation model, where the predicted number of acceptances was calculated by dividing the target number of registrations by the probability of student registration. This approach was implemented in Tableau using parameters and calculated fields, allowing for dynamic adjustments and what-if scenario analyses. This model enables university administrators and stakeholders to understand how changes in the number of acceptances could impact the likelihood of reaching the desired registration target. Predictive machine learning models could not be used in this research due to the format of the provided university records datasets and the unavailability of predictive features in the datasets.
- **Implementation in Tableau**: The predictive model was implemented using Tableau's parameter functionality and calculated fields. A parameter named Target Registration was created, allowing users to input the desired number of students they wanted to register at the university. The model then used this value, divided by the probability of student registration (calculated based on the

provided historical datasets), to predict the number of students who needed to be accepted to achieve that target. This approach made the predictive model highly interactive and user-friendly for non-technical stakeholders, as they could easily visualise different outcomes by adjusting the parameter.

- **Model Training and Testing (Evaluation)**: The dataset for the predictive model was split into training and testing data. Due to the limited number of rows, the data was split to the 50:50 ratio so that both the training and testing data would capture all the important information necessary for prediction and visualisation. The predictive model was trained using the training data to make the acceptance and registration estimations. The evaluation of the model's performance was done using the evaluation metric root mean square error (RMSE). This metric was used because the model functions similarly to a machine learning regression predictive model, where continuous quantities are predicted rather than classification labels. RMSE provided a measure of the model's accuracy by quantifying the average deviation between predicted and actual values.

- **Model Interpretation and Insights**: The results of the predictive model will be presented to university administrators and stakeholders through an interactive dashboard in Tableau. The dashboard will allow them to input different target registration values and immediately see the corresponding predictions for the number of acceptances required. This level of interactivity will not only provide transparency into how the predictions are made but also enable decision-makers to simulate different scenarios and plan accordingly.

- **Handling Uncertainty**: Predictive modelling in higher education is fundamentally uncertain due to the dynamic nature of student behaviours and external factors such as policy changes and economic conditions. To address uncertainties in this research, sensitivity analysis was conducted using the Tableau model. Sensitivity analysis allows for the exploration of how changes in key parameters, such as the probability of registration, impact the predicted number of acceptances. This helps the university assess the robustness of its recruitment strategies under different scenarios.

- **Model Deployment**: The final predictive model was deployed as part of a broader Tableau dashboard, which included other visual analytics related to student applications, acceptances, and registrations. This dashboard was designed to be accessible to university administrators and stakeholders, providing them with the tools needed to explore data-driven predictions, understand trends, and make informed decisions that align with their strategic enrolment goals.

```
IF SUM([Number of Acceptances (train new final records.csv)]) > 0
THEN [Target Registrations] / [Probability of Registrations]
ELSE NULL
END
```

Figure 1: Calculation Field on Tableau Used to Develop the Predictive Model

3.5 Data Visualisation

Data visualisation played a pivotal role in this research, transforming complex data insights into an intuitive format that could be easily understood by administrator and

stakeholders at the University of Buckingham. The goal of the data visualisation was not only to present the results of data analysis and predictive modelling but also to provide an interactive tool that the administrators and stakeholders could use to explore trends and insights on their own. This section elaborates on the data visualisation approach and tools used in this research.

- **Tool Selection**: Tableau was chosen as the primary data visualisation tool due to its capability to create interactive, visually appealing dashboards. Tableau's versatility allowed for the integration of the complex datasets and the developed statistical predictive model, providing dynamic visual representations that would be easily understood by different university administrators, stakeholders, as well as marketing teams and admissions officers.
- **Dashboard Design**: An interactive dashboard with a couple of pages were created to display the trends and insights derived from the student enrolment data. The dashboards were designed to be user-friendly, allowing non-technical administrators and stakeholders to explore the data and draw conclusions without needing specialised data analysis skills. The key components of the dashboards included:
    - ➢ **Enrolment Trends Over Time**: A line chart was used to represent trends in applications, acceptances, and registrations over various academic years. This visualisation enables university administrators and stakeholders to identify seasonal peaks and trends, facilitating better planning for recruitment campaigns.
    - ➢ **Departmental Analysis**: Stacked bar charts and a tree maps were used to display the breakdown of applications, acceptances, and registrations by academic departments and their breakdown in each academic year. This allows the university administrators and stakeholders to compare performance across the departments, identify areas with lower acceptance rates, and strategise to improve them.
- **Visual Representation of Key Metrics**: To provide a comprehensive view of the enrolment metrics, key performance indicators (KPIs) were highlighted on the dashboard. These included the application-to-acceptance ratio (acceptance rate) and the acceptance-to-registration ratio (registration rate). The KPIs were displayed using gauge charts and summary cards to ensure that administrators and stakeholders could quickly grasp the overall performance of the university's enrolment processes.
- **Actionable Insights for the University**: The dashboards were designed with the primary goal of delivering actionable insights. Filters were added to allow university administrators and stakeholders to drill down into specific segments, such as comparing undergraduate versus postgraduate levels or domestic UK-based (home) versus international (overseas) students. This functionality enables the admissions and marketing teams to create more focused and effective strategies based on detailed insights.
- **Predictive Model Integration**: The results of the predictive model were integrated into the dashboard to provide dynamic, scenario-based insights. By using adjustable parameters, administrators and stakeholders could modify the target number of registrations and immediately visualise the required number

of acceptances needed to achieve that target. This interactive feature empowers administrators and stakeholders to simulate different scenarios and make informed decisions based on data-driven predictions.

- **User Interactivity**: User interactivity was a key feature of the visualisations created in Tableau. Parameters and filters allowed administrators and stakeholders to explore different data segments, compare trends over time, and test different scenarios. For example, the administrators and stakeholders can use the created parameter to set target registration numbers and see how different values impacted the number of required acceptances. This hands-on approach enabled the university to take ownership of the data and make decisions based on evidence rather than intuition.

- **Real-Time Data Updates**: The dashboards in Tableau were designed to be updated in real-time as new data became available. This ensures that administrators and stakeholders always had access to the most current information, allowing for timely adjustments to recruitment strategies and more effective responses to changes in the enrolment landscape.

- **Ease of Use and Accessibility**: The Tableau dashboards were customised to be accessible to users of all technical backgrounds. Efforts were made to keep the design clean, with intuitive navigation and explanatory tooltips to help users understand the visual elements. This approach ensures that university administrators and stakeholders can effectively navigate the dashboards and utilise the insights for decision-making without requiring additional training sessions.
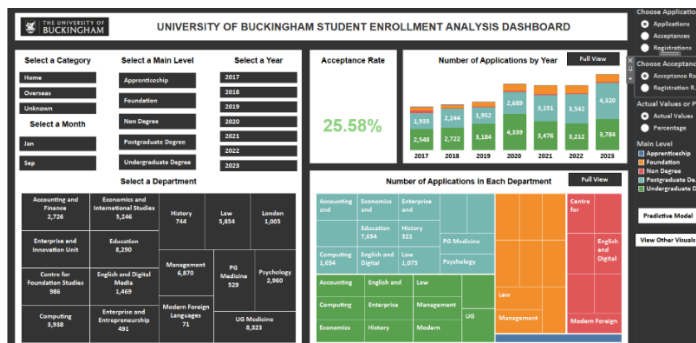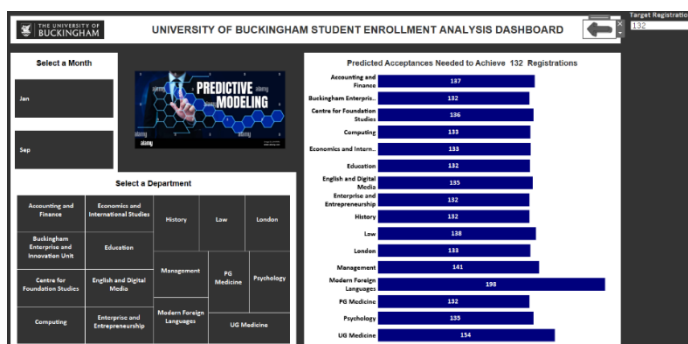

Figure 2: Main Dashboard Page Visualisation

Figure 2: Predictive Model Visualisation

CHAPTER FOUR: EXPERIMENT, RESULT, AND DISCUSSION

4.1 Exploratory Data Analysis (EDA)

The exploratory data analysis (EDA) conducted for this research was a crucial step in understanding the dataset's structure and identifying underlying patterns and trends that could impact the student enrolment process. The preprocessed dataset used for this analysis included various features such as main academic level, admission date, category, and number of applications, acceptances, and registrations. The key objectives of the EDA were to uncover meaningful insights that could inform the development of the visualisation framework for the University of Buckingham's administrators and stakeholders. This framework aimed to enhance their understanding of student enrolment processes and support the creation of the predictive model designed to optimise recruitment strategies.

4.1.1 Key Findings from EDA

- **Patterns in Application Data by Admission Date**: The EDA identified significant trends in student application numbers during the January and September admission periods. The data revealed a notable peak in applications for the January admission entry in 2020 and 2023. Furthermore, in 2023, the total number of admissions for the January entry surpassed those of the September entry. These findings highlight the potential need for more targeted marketing campaigns and optimised resource allocation for January admission periods to maximise student applications and enrolment outcomes.
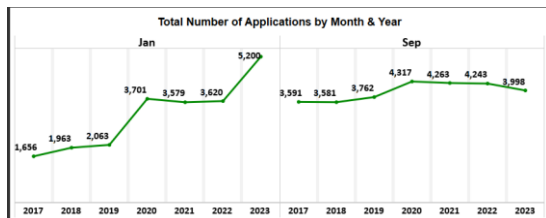

Figure 4: Image of peak application numbers in January 2020 and 2023

- **Geographical Differences**: The analysis highlighted geographical variations in application behaviour, revealing a significant annual decline in applications from domestic UK-based (home) students, contrasted with a steady annual increase in applications from international (overseas) students. This trend suggests that implementing tailored marketing strategies focused on international students could be highly effective in driving further growth in application numbers.
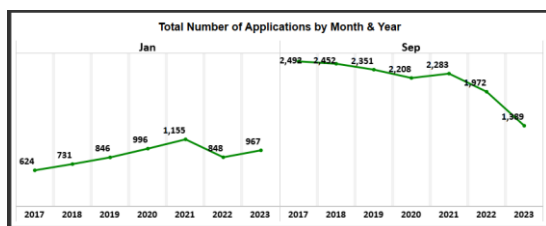

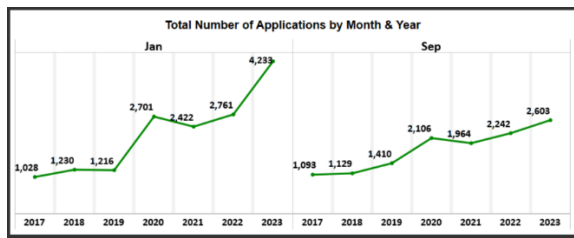Figure 5: Image of decline in application numbers for domestic students

Figure 6: Image of increase in application numbers for international students

- **Departmental Acceptance Rate**: The analysis revealed variations in acceptance rates across different departments. Departments such as London and Education demonstrated higher acceptance rates, exceeding 50%, compared to other departments. This trend may indicate that applicants to these departments are of a higher calibre. As such, focused marketing efforts and tailored communication strategies should target these students to ensure they accept their offers and enrol, ultimately enhancing the overall quality of students at the university.

- **Registration Rate Analysis**: An analysis of registration rates was conducted to assess the proportion of students who, after being accepted, ultimately enrol at the university. The results indicated consistently high registration rates across all variables in this study. However, as outlined in the methodology chapter, the visualisations and predictive model were developed using datasets extracted during a period when acceptance numbers had been streamlined. This adjustment reflects reductions due to some students failing to meet acceptance conditions or withdrawing their offers. Consequently, the registration rates in this research are artificially inflated, with values approaching 100%, and may not fully represent the original data. However, this is due to the unavailability of the original data, which was inaccessible due to data sensitivity concerns. With the current structure of this research, the framework is expected to be more effective and accurate if utilised by someone with access to the complete and unaltered datasets.

4.1.2 Visual Techniques Used in EDA

- **Line Graphs**: Line graphs were utilised to analyse trends in applications, acceptances, and registrations over time, providing a temporal perspective on enrolment patterns across the admission years. This visualisation was particularly effective in identifying peak periods, enabling targeted interventions and strategic planning. Additionally, it offered insights into long-term trends, helping university administrators and stakeholders understand the dynamics of student enrolment and informing efforts to optimise recruitment and resource allocation.

- **Stacked Bar Charts**: Bar charts were employed to compare the number of applications, acceptances, and registrations across various admission years for each department. This visualisation effectively highlights trends and variations over time, enabling university administrators and stakeholders to identify patterns in departmental performance and make data-driven decisions to enhance enrolment strategies and resource allocation.

- **Tree maps**: Tree maps were utilised to compare departmental performance in terms of applications, acceptances, and registrations. This visualisation was also employed to assess the performance of domestic UK-based (home) students versus international (overseas) students. By providing a clear illustration of disparities, the tree maps helped identify areas where targeted marketing efforts and resource allocation could be optimised to address gaps and enhance overall performance.

4.2 Predictive Model Results

The predictive model was developed to estimate the number of student applications that should be accepted to reach the target number of registrations. The approach involved dividing the target registrations by the probability of student registration. The model leveraged historical data to compute these probabilities, providing a data-driven approach to achieving enrolment targets, which was elaborated on in the methodology chapter.

4.2.1 Model Performance Evaluation

Root Mean Squared Error (RMSE): The RMSE of the model was calculated to evaluate the difference between the predicted and actual number of acceptances from the testing data. The RMSE value obtained was 8.9372, indicating that the predicted values closely aligned with the actual acceptance numbers. A lower RMSE value signifies a smaller average prediction error, reflecting the model's ability to accurately estimate the number of acceptances needed to achieve the desired registration targets.

The implementation of the predictive model within Tableau allows for dynamic parameter adjustment, enabling the university administrators and stakeholders to test different target scenarios and observe how changes would affect the required number of acceptances. By providing a user-friendly interface, the predictive model enables the university to make data-driven decisions, ensuring effective and efficient student enrolment management.

However, as outlined in the methodology chapter, the predictive model was developed using datasets extracted during a period when acceptance numbers had been streamlined. This adjustment reflects reductions due to some students failing to meet acceptance conditions or withdrawing their offers. Consequently, the RMSE value might be very good because of this reason. However, this is due to the unavailability of the original data, which was inaccessible due to data sensitivity concerns. With the current structure of this research, the predictive model is expected to be more effective and accurate if utilised by someone with access to the complete and unaltered datasets.

4.3 Visual Insights

The interactive dashboard developed in Tableau played a crucial role in communicating insights derived from the data and the predictive model. These visual insights will empower the university administrators and stakeholders by providing them with an intuitive data-driven framework to explore and interact with the enrolment data.

The visualisations used provided both a macro-level overview and granular insights into the dynamics of student applications, acceptances, and registrations.

4.3.1 Key Visualisations

- **Line Graphs**: Line graphs were utilised to analyse trends in applications, acceptances, and registrations over time, providing a temporal perspective on enrolment patterns across the admission years. This visualisation was particularly effective in identifying peak periods, enabling targeted interventions and strategic planning. Additionally, line graphs offered insights into long-term trends, helping university administrators and stakeholders understand the dynamics of student enrolment and informing efforts to optimise recruitment and resource allocation.
- **Stacked Bar Charts**: Stacked bar charts were employed to compare the number of applications, acceptances, and registrations across various admission years for each department. This visualisation effectively highlighted trends and variations over time, enabling university administrators and stakeholders to identify patterns in departmental performance. The stacked bar charts provided a clear comparison between departments, helping decision-makers make data-driven decisions to enhance enrolment strategies and allocate resources more efficiently. By visualising the trends over time, stakeholders can determine which departments require more focused efforts to improve their performance.
- **Tree Maps**: Tree maps were utilised to compare departmental performance in terms of applications, acceptances, and registrations. This visualisation was also employed to assess the performance of domestic UK-based (home) students versus international (overseas) students. By providing a clear illustration of disparities, the tree maps helped identify areas where targeted marketing efforts and resource allocation could be optimised to address gaps and enhance overall performance. This visual approach allows university administrators and stakeholders to quickly identify departments that need additional attention or support, facilitating better decision-making.

4.4 Discussion

The findings of this research reinforce the significant potential of data-driven enrolment strategies in enhancing student application, acceptance, and registration processes at the University of Buckingham. By leveraging predictive modelling and interactive visualisations, the research was able to provide a comprehensive understanding of the insights into the university's current enrolment processes. These insights are crucial for developing more targeted, data-informed recruitment strategies and for allocating resources more effectively.

4.4.1 Benefits of the Data-Driven Strategies

- **Enhanced Decision-Making**: The integration of predictive modelling with Tableau dashboards will greatly enhance the decision-making process for the University of Buckingham. The potential for the model's high accuracy (RMSE), coupled with the interactive nature of the dashboards, will enable administrators

and stakeholders to make informed decisions in real time. The ability to simulate different enrolment scenarios and visualise the impact of potential strategies will empower the university to optimise planning and allocate resources more efficiently.

- **Targeted Recruitment Campaigns**: The analysis of domestic UK-based (home) students and international (overseas) students provided valuable insights by clearly illustrating disparities in application and enrolment patterns. These findings identified specific areas where targeted marketing efforts and optimised resource allocation could address gaps, enhance recruitment strategies, and improve overall performance.
- **Interactive Scenario Analysis**: The parameter-driven predictive model implemented in Tableau provides the University of Buckingham with a powerful tool for interactive scenario analysis. By allowing the university administrators and stakeholders to input different target registration values and observe the predicted acceptances required to meet those targets, the model provides a dynamic, scenario-based approach to decision-making. This feature proves to be both intuitive and impactful, enabling the university to explore various scenarios and select the most effective strategies for achieving their enrolment goals.
- **Improving Enrolment Rate Efficiency**: The registration rate analysis conducted during the EDA revealed significant insights in the rate of acceptance to registration conversion. As explained in the methodology chapter and EDA section, with the structure of this research, the framework is expected to be more effective and accurate if utilised by someone with access to the complete and unaltered datasets. The university can then implement targeted interventions to improve their enrolment rates.

4.4.2 Challenges and Considerations

While the results are promising, several challenges must be addressed to fully leverage data-driven approaches in higher education:

- **Data Quality and Consistency**: One of the primary challenges in implementing data-driven strategies is ensuring data quality and consistency. In this research, considerable effort was invested in preprocessing the data, including cleaning and handling missing values. However, maintaining the data accuracy and consistency will be a challenge, especially as new data will be continuously collected. To address this issue, the University of Buckingham must invest in robust data governance practices to ensure that data remains reliable and that any inconsistencies are promptly identified and rectified.
- **Data Privacy and Ethical Considerations**: The use of student data for predictive modelling raises important ethical considerations regarding privacy and data protection. The University of Buckingham must ensure that student data is collected, stored, and used in compliance with data protection regulations. Transparency in data collection practices is essential, and students should be informed about how their data will be used. Additionally, appropriate cyber security safeguards must be implemented to protect sensitive information and prevent misuse.

- **Organisational Culture and Adoption**: For data-driven strategies to be effective, they must be embraced at all levels of the organisation. This requires fostering a cultural shift towards evidence-based decision-making, where data and analytics are integral to planning and strategy. To support this transition, training and capacity-building initiatives should be implemented to ensure that the University of Buckingham has personnel with the necessary skills and knowledge to effectively leverage data. This will enable the institution to derive actionable data-driven insights and improve its enrolment strategies.

The integration of predictive modelling and interactive data visualisation has the potential to significantly enhance student enrolment management at the University of Buckingham. However, to fully realise the benefits, challenges related to data quality, privacy, ethical use, and organisational adoption must be addressed. By investing in data governance, fostering a culture of evidence-based decision-making, and ensuring transparency in data practices, the University of Buckingham can unlock the full potential of data-driven strategies to transform recruitment and enrolment processes. This transformation will not only enhance the efficiency of enrolment management but also contribute to a more equitable and inclusive educational environment, ultimately leading to improved outcomes for both students and the institution.

CHAPTER FIVE: CONCLUSION AND FUTURE WORK

5.1 Conclusion

The primary aim of this research was to develop a data-driven framework for enhancing student applications, acceptances, and registrations at the University of Buckingham. By leveraging predictive modelling and interactive data visualisation, the study sought to provide a comprehensive understanding of their student enrolment processes and to equip the university administrators and stakeholders with a tool for making informed, data-driven decisions.

The predictive model built in this study demonstrated the effectiveness of using historical data to estimate future enrolment outcomes. With a potential high level of accuracy and if provided with the original data, the model is expected to provide reliable estimates of the number of acceptances needed to achieve specific registration targets. This capability allows for more strategic planning and targeted interventions, optimising recruitment and allocation of resources. The use of Tableau dashboards added a valuable layer of interactivity, enabling administrators and stakeholders to visualise different scenarios and better understand the enrolment landscape. Through these dashboards, the university can interact with the data, simulate different enrolment scenarios, and make evidence-based decisions to optimise outcomes.

The Exploratory Data Analysis (EDA) provided key insights into the underlying patterns within the data. The analysis highlighted key insights that can help improve student enrolment. Using these insights, the research enables the University of Buckingham to develop more targeted recruitment campaigns and support initiatives.

Moreover, the study highlighted the potential of data-driven strategies in analysing domestic UK-based (home) and international (overseas) students, offering valuable insights by clearly illustrating disparities in application and enrolment patterns. By identifying and understanding these disparities, the university can implement targeted interventions to provide additional support, fostering a more equitable and inclusive enrolment process.

The implementation of data-driven strategies has the potential to significantly enhance enrolment management processes at the University of Buckingham. However, to fully realise these benefits, several challenges must be addressed. Ensuring data quality is crucial, as predictive models are only as reliable as the data they are based on. Maintaining privacy and ethical standards is also critical, as the use of student data must comply with data protection regulations and respect student privacy. Additionally, fostering a culture of evidence-based decision-making is essential for the successful adoption of data-driven strategies.

5.2 Future Work

While the current study has made significant contributions to understanding and optimising student enrolment, there are several areas where future research and development could further enhance the framework:

- **Incorporating Machine Learning Techniques**: Future work could explore the use of machine learning models to predict enrolment outcomes. Machine learning techniques such as decision trees or random forests could potentially provide even more accurate predictions and uncover complex relationships within the data. Ensemble methods that combine multiple machine learning models could be tested to see if they improve prediction accuracy and robustness.
- **Expanding the Dataset**: The current study was limited to data from specific admission years and attributes. Expanding the dataset to include additional variables, such as geographical data, individual student data, and date and time of application and enrolment, could provide a more holistic view of student behaviour and improve the accuracy of predictive models. Including data from external sources, such as economic indicators or high school performance metrics, could also add valuable context for understanding trends in applications and registrations.
- **Real-Time Data Integration**: Implementing real-time data integration could further enhance the responsiveness of the predictive model. By incorporating up-to-date information, the university could make more timely decisions and adjust recruitment strategies as needed.
- **Addressing Data Privacy and Ethical Concerns**: Future research should also focus on developing robust frameworks for ensuring data privacy and ethical use of student information. This includes creating transparent policies for data collection and use, obtaining informed consent from students, and implementing best practices for data anonymisation.
- **Longitudinal Analysis**: Conducting a longitudinal analysis to track the impact of data-driven interventions over time would provide valuable insights into the effectiveness of these strategies. By assessing how changes in recruitment campaigns, resource allocation, and student support services impact enrolment outcomes over multiple academic years, the university can refine its approach and continuously improve its enrolment management practices. Longitudinal studies could also help identify lasting effects of specific interventions and highlight areas where persistent challenges remain.
- **Incorporating Student Feedback**: Another area of future work could involve incorporating student feedback into the data-driven framework. Understanding students' perspectives on the enrolment process, their motivations for applying, and their reasons for accepting or rejecting offers can provide additional qualitative insights that enhance predictive models. Surveys, focus groups, and sentiment analysis of student communications could be used to gather this information.

5.3 Final Remarks

The integration of predictive modelling, data visualisation, and data-driven decision-making has the potential to transform enrolment management at the University of Buckingham. By utilising insights from historical data and leveraging advanced analytics, the university can not only optimise recruitment strategies but also better serve its student population. Data-driven approaches enable the identification of key

drivers of student success, allowing for more personalised support and targeted interventions, which can help improve student outcomes and overall satisfaction.

Addressing the challenges identified in this study is crucial for fully realising the potential benefits of data-driven strategies. Ensuring data quality, maintaining ethical standards, and fostering a culture of data literacy across the institution are key factors in the successful adoption of these methods. Investment in data infrastructure will be essential for sustaining the momentum gained through this research. Implementing robust data governance frameworks will help maintain the accuracy and reliability of the data. Ethical considerations must also remain a priority, with transparent data usage policies and clear communication to students about how their data is being utilised.

The future of enrolment management at the University of Buckingham lies in the ability to continuously adapt and innovate. By building on the insights gained from this research and incorporating emerging technologies and methodologies, the university can stay ahead of changing trends in higher education and provide a more inclusive, supportive environment for all students. Ultimately, the successful integration of data-driven strategies will contribute to a more effective enrolment process, enhanced student experiences, and improved institutional performance, positioning the University of Buckingham as a leader in the use of analytics for educational excellence.

References

1. Apraxine, D., and Stylianou, E. (2017) 'Business intelligence in a higher educational institution: The case of University of Nicosia', *2017 IEEE Global Engineering Education Conference (EDUCON)*, pp. 1735-1746.

2. Baker, R. S., and Yacef, K. (2009) 'The State of Educational Data Mining in 2009: A Review and Future Visions', *Journal of Educational Data Mining*, 1(1), pp. 3-17.

3. BBC (2024) *Drop in foreign student visas worrying for UK universities.* Available at: https://www.bbc.co.uk/news/articles/cdd4p62nyg8o (Accessed: 25 November 2024)

4. BBC (2024) *Overseas student applications to UK universities rise again.* Available at: https://www.bbc.co.uk/news/education-68294078 (Accessed: 25 November 2024)

5. BBC (2024) *Universities' international admissions to be reviewed.* Available at: https://www.bbc.co.uk/news/education-68186670 (Accessed: 25 November 2024)

6. BBC (2021) *Record numbers of UK students head for university.* Available at: https://www.bbc.co.uk/news/education-58478227 (Accessed: 25 November 2024)

7. BBC (2023) *Fall in students accepted into university in UK.* Available at: https://www.bbc.co.uk/news/education-66867075 (Accessed: 25 November 2024)

8. BBC (2024) *Top-grade unis soak up students, Ucas data shows.* Available at: https://www.bbc.co.uk/news/articles/cp816v37jm5o (Accessed: 25 November 2024)

9. Chen, N., Yin, C., Isaias, P., and Psotka, J. (2020) 'Educational big data: extracting meaning from data for smart education', *Interactive Learning Environments*, 28(2), pp. 142-147.

10. Cope, B., and Kalantzis, M. (2016) 'Big Data Comes to School: Implications for Learning, Assessment, and Research'. *AERA Open*, 2(2), pp. 1-19.

11. Daniel, B. (2015) 'Big Data and Analytics in Higher Education: Opportunities and Challenges'. *British Journal of Educational Technology*, 46(5), pp. 904-920.

12. Dart, S., and Cunningham, S. (2023) 'Using institutional data to drive quality improvement and innovation', *Technology-Enhanced Learning and the Virtual University*.

13. Haque, W. (2018) 'An Interactive Framework to Allocate and Manage Teaching Workload using Hybrid OLAP Cubes', *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pp. 694-700.

14. Krein, U., and Schiefner-Rohs, M. (2021) 'Data in Schools: (Changing) Practices and Blind Spots at a Glance', *Sec. Leadership in Education,* 6.

15. McFarland, D. A., Khanna, S., Domingue, B. W., and Pardos, Z. A. (2021) 'Education Data Science: Past, Present, Future'. *AERA Open*, 7(1), pp. 1-12.

16. Park, Y. (2020) 'Uncovering trend-based research insights on teaching and learning in big data', *J Big Data,* 7(93).

17. Schildkamp, K. (2019) 'Data-based decision-making for school improvement: Research insights and gaps', *Educational Research*, 61(3), pp. 257–273.

18. Sequeira, R., Reis, A., Alves, P., and Branco, F. (2023) 'Roadmap for Implementing Business Intelligence Systems in Higher Education Institutions: Systematic Literature Review', *MDPI*.

19. Stojanov, A., and Daniel, B. K. (2023) 'A decade of research into the application of big data and analytics in higher education: A systematic review of the literature' *Education and Information Technologies,* 29, pp. 5807-5831.

20. Soncin, M., and Cannistra, M. (2021) 'Data analytics in education: are schools on the long and winding road?', *Qualitative Research in Accounting & Management,* 19(3), pp. 1176-6093.

21. Widodo, F. F., Rofiq, A., and Aprilianto, A. (2022) 'Student Admission Strategy In Increasing The Quality of Students', *Chalim Journal of Teaching and Learning,* 2(2), pp. 124-132.