Graphical user interface, text

Description automatically generated

## Final Customer Report

### 

### Predicting the Rank of the School

**Submitted to**:

**Professor**: **Dr. Brent Furneaux**

BSMM-8750: Predictive Modeling and Decision Making  
Odette School of Business | University of Windsor

**Submitted by:**

**Team Outliers (Group 3)**

| **Rashi Gupta** | 110100410 |
| --- | --- |
| **Methu Damodaran** | 110100926 |
| **Phuong Loan Pham** | 110094008 |
| **Rashmi Kumari** | 110093914 |

**Submission Date:**

6th December 2023

Contents

[**Final Customer Report 1**](#_38suo8u8czs)

[Predicting the Rank of the School 1](#_tzbjsgz5dees)

[**Project Overview 3**](#_gjdgxs)

[Predicting Master of Management School Rank Enhancement 3](#_p9vslk1bvr5t)

[**Objectives 3**](#_30j0zll)

[**Given Data Overview 3**](#_1fob9te)

[**Data Preprocessing 4**](#_3znysh7)

[Data Cleaning 4](#_2et92p0)

[Eliminate inaccuracies and inconsistencies in the dataset. 4](#_hhw28usafr3x)

[Handling Missing Values 4](#_tyjcwt)

[Feature Engineering 4](#_sgaltipfg6hv)

[Explanatory Data Analysis 5](#_1t3h5sf)

[Key Trends and Pattern 5](#_4d34og8)

[Factor Analysis: Uncontrollable vs. Controllable Factors 5](#_2s8eyo1)

[Longitudinal Analysis 6](#_17dp8vu)

[Why Regression? 7](#_3rdcrjn)

[● Continuous Prediction 7](#_urymfs26n8j6)

[● Interpretability 7](#_2pq5fe88q3h)

[● Capturing Linear Relationships 7](#_lbr110i8wwsi)

[● Multifactorial Analysis 7](#_pvq2258lt0mu)

[Data Collection 8](#_dd3hu55jmcjc)

[Feature Derivation 8](#_lnxbz9)

[Assumptions 8](#_35nkun2)

[Formula 9](#_1ksv4uv)

[Feature Importance 9](#_44sinio)

[Modelling 10](#_2jxsxqh)

[SVM Model 10](#_z337ya)

[Random Forest Model 10](#_3j2qqm3)

[Ultimately Gradient Boosting 11](#_1y810tw)

[Refining the Model performance 11](#_4i7ojhp)

[**Recommendations 12**](#_2xcytpi)

[**Appendix 13**](#_te9c1z8wk27o)

## Project Overview

### Predicting Master of Management School Rank Enhancement

The primary objective of this project is to conduct a thorough and detailed analysis of the factors influencing university performance in the Financial Times 100 (FT100) rankings. This project is driven by the critical need to unravel the dynamics that shape these rankings, emphasizing the identification of actionable strategies that universities can employ to improve their standing.

## 

## Objectives

* Improve Model Accuracy
  + Enhance the accuracy of the current ranking prediction model.
  + Incorporate advanced analytical methods to improve the precision of predictions.
* Explore New Data Sources
  + Identify and explore opportunities to collect new data relevant to university performance in the FT100 rankings.
  + Integrate additional data to enrich the predictive model.
* Identify Key Predictive Features
  + Conduct a thorough analysis to identify and prioritize key features contributing significantly to university performance in the FT100 rankings.
* Model Transparency and Interpretability
  + Prioritize transparency and interpretability in the model's decision-making process.
  + Ensure stakeholders have a clear understanding of how the model arrives at its predictions.

## 

## Given Data Overview

Our project commenced with a comprehensive dataset encompassing five years of Financial Times (FT) rankings (2018-2022) for the top 100 schools annually. This ranking data serves as a key foundation for understanding the dynamic performance trends in higher education.

Accompanying this dataset, we were provided with crucial insights into the FT ranking methodology, detailing the criteria and metrics shaping school evaluations. This transparency guides our analytical approach, aligning our strategies with the parameters emphasized by FT.

In addition to rankings, the dataset includes detailed information about university courses. This course-specific data augments our analysis, allowing exploration of correlations between program characteristics and overall university performance.

## 

## 

## 

## Data Preprocessing

In the data preparation phase of the project, we meticulously curated a comprehensive dataset spanning three years of Financial Times 100 (FT100) rankings and then identified the common columns. This unified dataset of three years was subsequently stored in a separate Excel file for streamlined analysis. After which data preprocessing starts. The process involves systematic data cleaning, handling missing values judiciously, and strategic feature engineering to enrich the dataset.

### Data Cleaning

#### Eliminate inaccuracies and inconsistencies in the dataset.

* **Variable Value Standardization**
  + **Objective**: Enhance uniformity by addressing inconsistencies in variable values.
  + **Implementation**: Data transformation tools, such as Python's replace() function, were used to consistently normalize variable values. This phase provides consistent representation, allowing for accurate comparisons and analysis. Because our dataset was manageable in size, we decided to execute Variable Value Standardization in Excel as well. This decision was driven by the ease of use and efficiency offered by Excel for this specific task.
* **Handling Non-numeric Characters**
  + **Objective**: Address non-numeric characters in the "Weighted salary" column for uniformity and numerical consistency.
  + **Implementation**: Executed code to remove non-numeric characters and convert the column to a float data type. This crucial step ensures numerical uniformity in the "Weighted salary" data, enabling meaningful quantitative analysis of salary-related metrics.

df["Weighted salary"] = df["Weighted salary"].str.replace('[^\d.]', '', regex=True).astype(float)

#### Handling Missing Values

* **Imputation of Missing Values**

Following the identification of missing values in columns such as 'Employed at three months' and 'GMAT/GRE Required,' our data imputation technique included leveraging the median values for the specified columns. This method ensures that central tendencies are preserved and that future analysis are strong.

* **Handling Columns with High Missing Values**

With the goal of reducing noise and bias in our dataset, we strategically implemented the removal of columns with a high number of missing values. This process involved setting a defined threshold (threshold = 250) beyond which columns were dropped, focusing on retaining those with sufficient data for generating meaningful insights.

#### 

#### 

#### Feature Engineering

Enhance dataset with new, impactful features for predictive modeling.

* **Deriving New Metrics**

Create new metrics that capture nuanced aspects of university performance. We applied mathematical operations to derive two features.

* + **Diversity Score:** Calculated as the average of international student percentage, international faculty percentage, and international board percentage. This metric offers a comprehensive measure of international diversity within the university.
  + **Female Empowerment Score:** Computed using a weighted combination of female student percentage, female faculty percentage, and women on the board (0.4, 0.4 and 0.2 respectively). The weighting reflects their varying impacts on female empowerment.
* **Programming Implementation**
  + Leveraged programming languages like Python, utilizing relevant libraries (e.g., pandas, NumPy) to implement feature engineering. This ensured systematic and reproducible creation of new features across the dataset.

### 

### Explanatory Data Analysis

#### Key Trends and Pattern

The FT100 rankings place a substantial emphasis on financial metrics, with up to 31% of the ranking weight dedicated to finance-related factors, in contrast to THE and QS, which prioritize academic and research metrics, accounting for up to 40% of their total weight.

A key focus of this analysis is to examine whether the FT100's methodology, particularly in terms of weight and percentage allocations, exhibits any consistent trends over the three-year period from 2020 to 2022. The data consistently indicates that programs with a higher weighted salary in the US tend to achieve higher rankings in the FT100. Additionally, metrics such as the percentage of alumni achieving their goals post-graduation (termed 'Aim Achieved') and the percentage of international students also significantly influence rankings. This consistency in trends over the years provides valuable insights for universities like School X, guiding them to prioritize and focus on these key areas to enhance their standing in future FT100 rankings.

#### Factor Analysis: Uncontrollable vs. Controllable Factors

The EDA also delves into a detailed factor analysis, categorizing influencing factors into 'uncontrollable' and 'controllable' elements. Uncontrollable factors include geographical location, which inherently affects international student and faculty ratios, and historical reputation, which evolves over long periods. On the other hand, controllable factors encompass aspects like quality of faculty, research output, and student support services, which universities can directly enhance to improve their rankings.

* **Uncontrollable Factors Analysis**

A key uncontrollable factor is the geographical location of an institution, which was found to be strongly correlated with the percentage of international students, internship opportunities, and value-for-money rankings. The data reveals a dominance of European schools in the rankings, particularly French programs, which account for 21% of the observations. This is followed by the UK and Germany, with 11% and 7% respectively.

Additionally, there's a notable rise in Asian schools, exemplified by the Indian Institute of Management Bangalore and Taiwan’s National Sun Yat-sen University, which have shown significant improvements in their rankings. This trend is attributed to the recent entry of Asian universities into the rankings and the economic and educational advancements in Asia. In contrast, European schools have been consistently dominant for over a decade.

Economic factors also play a significant role in this context. For instance, Germany's strong economy, the UK's status as a major financial hub, and France's global corporate presence contribute to their universities' ability to attract business students and offer extensive internship opportunities. These factors, while influential, are beyond the direct control of the universities and significantly impact their rankings in the FT100.

* **Controllable Factors Analysi**s

Regarding the controllable factors influencing university rankings in the Financial Times 100 (FT100), Career Service Rank stands out as a critical element. While it contributes just 4% to the FT100 score, its impact is disproportionately significant. The effectiveness of Career Services in aiding student recruitment, as rated by alumni, is highly correlated with Overall Satisfaction and Weighted Salary (US), which together account for up to 20% of the total score.

Furthermore, top-tier institutions often boast a Career Service Rank with 66% effectiveness in recruitment support, attracting premier companies for campus recruitment. Furthermore, these schools typically offer robust entrepreneurial programs and specialized services designed to enhance student career prospects. For instance, London Business School organizes exclusive recruitment events, while EMlyon Business School offers unique courses focused on business startups.

On another hand, Value for Money, another key metric, is calculated based on alumni salaries, tuition, and other costs. The economic backdrop of the country where the university is located, such as Germany's robust economy or the UK's financial prowess, also plays a role in this calculation. Another noteworthy observation is that schools with a high Career Service Rank often facilitate faster employment of students at higher salaries, often through internships. In our analysis, approximately 48% of the total observations over three years included internships, with only 11.8% not offering this opportunity. Additionally, the presence of faculty with doctorates is observed to positively influence rankings, indicative of a higher quality of education. This factor is a controllable element that universities can leverage to improve their standings in the FT100 rankings.

In summary, focusing on enhancing Career Services, providing valuable internships, and recruiting highly qualified faculty are strategic areas where universities can exert control to improve their rankings and overall educational quality. These insights offer actionable paths for educational institutions seeking to elevate their positions in the competitive landscape of global university rankings.

#### Longitudinal Analysis

The purpose of longitudinal analysis of university rankings and metric changes over a three-year period is to provide a clear and accessible evaluation of how certain universities have progressed in the Financial Times 100 (FT100) rankings. From the data, it is evident that certain universities have experienced significant changes in their rankings. For example, some institutions have shown remarkable improvement, climbing the rankings, while others have seen a decline.

Take Luiss University as example, over the course of three years, from 2020 to 2022, Luiss University has demonstrated a commendable trajectory of improvement in its academic and operational standings as per the Financial Times 100 (FT100) rankings. In 2020, Luiss University held the rank of 83, which by 2022, had advanced to 55th rank. A longitudinal analysis of the university's performance metrics reveals significant shifts:

* **International Course Experience Rank**: The university has improved its position by nearly 15 ranks, a testament to its expanding global reach and international program offerings.
* **Career Services Rank:** There has been a notable rise by approximately 10 ranks, reflecting the university's strengthened commitment to student career progression and recruitment support.
* **Aims Achieved and Employment at Three Months:** These metrics have remained relatively stable, suggesting consistent outcomes in terms of meeting alumni expectations and post-graduation employment.
* **Value for Money Rank:** A notable increase of 14 ranks in this metric signifies the university's improved economic proposition in terms of return on investment for its students.
* **International Faculty and Faculty with Doctorates:** Both metrics have seen an increase in the percentage score by approximately 6 ranks, reflecting the university's investment in high-caliber, international, and highly qualified academic staff.

These metrics collectively paint a picture of Luiss University's strategic focus areas and their outcomes. Improvements in internationalization, career services, and academic excellence have contributed to the university's elevated ranking position. For stakeholders, these findings are indicative of the dynamic environment in which educational institutions operate. The data suggests that universities can, and do, make significant movements within a relatively short time frame. For an institution like Luiss University, the improvement in diversity and international metrics may reflect strategic initiatives aimed at enhancing global presence and inclusivity — factors that are increasingly important in higher education.

### 

### Why Regression?

In the landscape of predicting university rankings, the selection of regression as our primary modeling technique is purposeful. In the project having scattered data points representing various factors influencing university rankings. Regression finds the best-fitting line through these points, allowing us to understand the changes in one factor might predict changes in another. It helps us make predictions based on patterns in the data.

#### Continuous Prediction

Our goal involves predicting university rankings, a continuous variable. Regression, designed for predicting numerical outcomes, aligns seamlessly with the nature of our target variable, offering a suitable framework to model the intricate relationships within the data.

#### Interpretability

Regression models provide a clear and interpretable understanding of the relationships between input features and the predicted output. This interpretability is crucial for our analysis, enabling us to discern the impact of different factors on university rankings and facilitating actionable insights.

#### Capturing Linear Relationships

Regression is adept at capturing linear relationships between variables. Given the complexity of university performance dynamics, the ability to model linear dependencies allows us to uncover trends and patterns that might influence rankings.

#### Multifactorial Analysis

Regression's ability to handle both numerical and categorical variables allows us to simultaneously analyze the impact of diverse factors, providing a holistic understanding of their contributions to rankings.

In essence, the strategic choice of regression is tailored to the intricate nature of our project, where diverse factors converge to influence university rankings. The adaptability, interpretability, and proven effectiveness of regression make it the ideal tool for unraveling the multifaceted dimensions of university performance.

### Data Collection

Data collection is the systematic gathering of information crucial for our analysis, encompassing Financial Times rankings, course details, and methodology insights. This diverse dataset serves as the bedrock for our exploration into university performance, facilitating a nuanced understanding of key factors influencing rankings and informing strategic decision-making in higher education.

In our data collection process, we employed various methods, including the manual review of school websites. This meticulous approach ensured a comprehensive gathering of information, allowing us to capture nuanced details beyond conventional datasets and enriching our analysis with firsthand insights.

### Feature Derivation

In our pursuit of a comprehensive understanding of university performance, we extended our dataset by deriving three new features, each offering unique insights into aspects not fully captured by traditional metrics. These features were crafted with careful consideration of assumptions and weighted formulae to ensure nuanced representation.

#### Assumptions

* **GMAT/GRE Requirement**

GMAT/GRE requirements as a feature in our analysis is underpinned by a set of thoughtful assumptions aimed at capturing the academic rigor and standards of university programs. The assumptions are as follows:

* + **Binary Representation**: We assume that the inclusion of standardized tests, such as GMAT or GRE, is a meaningful indicator of the academic rigor and competitiveness of a program. As such, if GMAT/GRE is required, we assign a value of 1; if not required, a value of 2 is assigned.
  + **Data Availability**: Recognizing potential data gaps, we have considered the scenario where information about GMAT/GRE requirements might be unavailable. In such instances, we denote this as 0, ensuring transparency in our representation of the dataset.
* **Diversity Score**

Each component contributing to diversity—international student percentage, international faculty percentage, and international board percentage—is assumed to carry equal importance in the overall Diversity Score. The choice of equal weightage for components contributing to the Diversity Score is rooted in the pursuit of a fair and balanced representation of a university's internationalization efforts. This fosters inclusivity, to avoid biases, ensures comprehensive representation across different dimensions, and enhances simplicity and transparency in the calculation of the Diversity Score.

* **Female Empowerment Score**

Assigning different weightages (40% to female students, 40% to female faculty, and 20% to women on the board) in the Female Empowerment Score is strategic. It captures the unique impacts of each group—reflecting the varying influences of female students, faculty, and board members on fostering an empowered academic environment. This approach aligns with industry standards, mirroring the weightings used in the FT methodology (female faculty % and female students % with a weight of 5 each, and women on board % with a weight of 1), ensuring consistency.

#### Formula

* **GMAT/GRE Requirement**

GMAT/GRE required = 1

GMAT/GRE not required = 2

GMAT/GRE required no information = 0

* **Diversity Score**

Diversity Score = (International Student (%) + International Faculty (%) + International Board (%))/3

* **Female Empowerment Score**

Female Empowerment Score = ( 0.4 × Female Student (%) + 0.4 × Female Faculty (%) + 0.2 × Women on Board (%))

### 

### Feature Importance

In machine learning, not all features contribute equally to the accuracy of a model; some may be highly influential, some may have no influence at all, and some can even have a misleading influence. We had several columns from the original dataset and 3 feature engineered columns as well so we wanted to assign a score to input features based on how useful they are at predicting a target variable. Understanding feature importance can help in several ways such as feature selection, model interpretability and model performance.

There are several feature selection methods we used to understand which features we should incorporate for model training. The feature selection models used were Lasso Regression, Random forest Model, SHAP, Recursive feature elimination model. We went ahead with several combinations because features that are important across multiple models tend to be more robust. Additionally, considering the model's overall performance and not just feature importance, as a balance between a model's accuracy and interpretability is often crucial.

As per the outputs of different models we incorporated these features for further modelling process.

metrics = [

"Weighted salary", "Salary percentage increase", "Value for money rank",

"Career progress rank", "Aims achieved (%)", "Effective Emp rate",

"International work mobility rank", "International course experience rank",

"Faculty with doctorates (%)", "Internship", "Overall Satisfaction",

"Diversity", "GMAT/GRE Required", "Female Empowerment Score"

]

This was an iterative process of selecting some features which were fit for modelling and dropping which yielded no results.

### 

### Modelling

#### 

#### SVM Model

A Support Vector Machine (SVM) is a powerful and versatile supervised machine learning algorithm used for classification and regression.We wanted to try Support Vector machine model for predicting rank for our project because for small to medium-sized datasets, SVM can be an excellent choice for ranking due to its effectiveness in finding a margin of separation and handling high-dimensional data. However when we passed the metrics to our SVM model the results we got were different from expected.

**Analysing the results from SVM:**

MSE (Mean Squared Error- 493893.56): A lower MSE value is better as it indicates that the model's predictions are closer to the actual data. Our MSE is quite high.

MAE (Mean Absolute Error- 563.300): This is the average of the absolute differences between the predictions and the actual observations. It gives an idea of how wrong the predictions were. The MAE is also quite high in our model.

R² (R-squared- -573.407): This is the coefficient of determination, which measures how well the regression predictions approximate the real data points. An R² of 1 indicates that the regression predictions perfectly fit the data. Our model has a negative R², which indicates that the model fits the data worse than a horizontal line (mean of the observed data). This is an indicator of a very poor model fit.

The main reason we believe we achieved these poor results were because of the high dimensionality. Scalability is often a drawback of SVM model, Scalability in this context refers to the ability of the method to handle increasing amounts of features ie: we had 15 features or computation efficiently. Also, SVMs make certain assumptions. For instance, SVMs assume that the data is at least roughly linearly separable when a linear kernel is used. If the actual underlying relationship is highly non-linear and complex, a linear SVM will perform poorly.

#### 

#### Random Forest Model

A Random Forest is an ensemble machine learning algorithm that combines the predictions from multiple decision trees to produce a more accurate and stable prediction. It is a form of bagging, where models are trained on different parts of the data and then averaged, which helps to reduce variance and avoid overfitting.By training a Random Forest as a regression model, however we can predict continuous scores for items. These scores can then be used to rank the items.This essentially became the idea for choosing this model.

We passed the same set of metrics into this model and compared to SVM it yielded pretty good results.

**Analysing the results from Random Forest Model:**

MSE (Mean Squared Error): Our Random Forest model has an MSE of approximately 153.76, which suggests that on average, the squared error of the predictions is 153.76.

MAE (Mean Absolute Error): With an MAE of approximately 9.99, our model's predictions are, on average, about 9.99 units away from the actual values.

R² (R-squared): An R² of 0.821 means that approximately 82.1% of the variability in the outcome data can be explained by the model's inputs. This is generally considered a good fit, indicating that the model explains a large proportion of the variability in the data.

As this model performed well we wanted to try a better version of the Random Forest model to see if it outperforms our current model. Gradient Boosting Models (GBMs) are another type of ensemble learning technique, which can be considered an alternative to Random Forests, and in many cases, they can outperform Random Forests. They both create a strong predictive model by combining multiple weaker models—decision trees in most cases.

#### 

#### Ultimately Gradient Boosting

* **Predictive Power**: When the primary goal is predictive power, and you have the computational resources to perform the necessary hyperparameter tuning, GBMs often emerge as the stronger model.
* **Ranking Tasks**: GBMs have been used effectively for ranking tasks. For instance, variants like Lambda MART are specifically designed for learning to rank and have been successful in information retrieval tasks.
* **Large Datasets with Noisy Features**: GBMs can handle large datasets effectively, and they can focus on features that really matter when predicting the target variable, even if the dataset has many noisy features.

**Analysing the results from GBM Model:**

MSE: 137.31731978132183

MAE: 8.93878269273592

RMSE: 11.71824729988755

R^2: 0.8402973981971944

Comparing these metrics to the earlier ones for the Random Forest model, the Gradient Boosting model has a slightly lower MSE and RMSE and a slightly higher R-squared value, all of which suggest a better fit to the data. The MAE is also lower, indicating that the average error in predictions is smaller. This comparison suggests that for this dataset and problem, the Gradient Boosting model is outperforming the Random Forest model.

The choice between Gradient Boosting and Random Forest models can depend on various factors, but these metrics suggest that in this case, Gradient Boosting has a slight edge in predictive performance.

#### Refining the Model performance

The values mentioned above can be enhanced by hyperparameter tuning. Hyperparameter tuning is a critical step in maximizing the performance of a Gradient Boosting Machine (GBM). After an initial run of the GBM, hyperparameter tuning helps to refine the model to make more accurate predictions.

We used a randomised search CV function in python. RandomizedSearchCV is available as part of the sklearn.model\_selection module Randomized Search Cross-Validation (RandomizedSearchCV) is a hyperparameter tuning technique in machine learning.

The updated metrics post-tuning are considerably better than both the previous SVM and Random Forest models we described earlier:

MSE (Mean Squared Error): It has decreased to around 102.83 from the previous models, indicating a reduction in the average squared difference between the predicted and actual values.

R² (R-squared: An increase to approximately 0.880 suggests that the model now explains 88% of the variance in the target variable, which is a significant improvement and indicates a strong predictive power.

MAE (Mean Absolute Error): This has decreased to around 7.66, reflecting a lower average error in the model's predictions.

RMSE (Root Mean Squared Error): The RMSE has reduced to about 10.14, indicating a smaller dispersion of the prediction errors.

The improvements in these metrics suggest that the hyperparameter tuning has been effective in enhancing the model's performance, making the predictions more accurate and the model more reliable.

## 

## Recommendations

* Improving a university’s performance in the FT100 Rankings requires a diverse and comprehensive strategy. Firstly, a significant emphasis should be placed on Career Services. By enhancing these services, especially in terms of recruitment support and forging strong connections with leading companies for campus placements, universities can directly influence alumni satisfaction and weighted salaries, both critical factors in the FT100 rankings. Following is prioritizing internships and practical experience. By offering substantial internship opportunities that pave the way for faster and more lucrative employment, universities can directly impact key metrics of the FT100 rankings.
* Another crucial aspect involves the recruitment and retention of faculty. The presence of faculty members with doctorates and robust research backgrounds positively impacts rankings, necessitating investment in attracting and retaining such talent to bolster educational and research excellence. Moreover, understanding and strategically adapting to the FT100 ranking methodology is also vital, universities need to focus on areas with higher weightage in the rankings, such as financial metrics, international student ratios, and alumni achievements, to make informed and strategic improvements.
* On the other hand, focusing on optimizing value for money is another strategic direction. Universities should strive to balance tuition costs with the quality of education and career prospects offered, enhancing the return on investment for students. Additionally, investing in strong alumni relations is crucial. This not only aids in networking and fundraising but also influences rankings through alumni surveys. Strengthening alumni ties can significantly contribute to a university's ongoing success and reputation.
* As academic reputation is heavily influenced by research quality, universities should also encourage faculty publications in prestigious journals and foster collaborations with industry and academia. Developing strategic partnerships with leading global institutions can provide mutual benefits, such as research collaboration and resource sharing, thereby elevating the university's international standing.

## Appendix

A graph with blue and green bars

Description automatically generated

**Chart 1. Correlation between Metric Change and Change in Rankings**

A screenshot of a computer

Description automatically generated

**Chart 2. Uncontrollable Factors Example: Location**

**A screenshot of a computer

Description automatically generated**

**Chart 3. Uncontrollable Factors Example: Location and its Correlation with Value For Money**

**A screenshot of a graph

Description automatically generated**

**Chart 4. Controllable Factors Example: Internship and Its observation**

**A graph of different colored columns

Description automatically generated with medium confidence**

**Chart 5. Applied Longitidual Analysis on Dataset for Ranking Comparison Between 2020 and 2022**

**A graph with green squares and black text

Description automatically generated**

**Chart 6. Example of Changing in Ranking of Factors Affected Ranking Improved of Luiss University**

**A group of blue graphs

Description automatically generated**

**Chart 7. Distribution of Features in FT100 Dataset in 3 years**

**A graph of multiple colored bars

Description automatically generated with medium confidence**

**Chart 8. Average Rank by Cohort Across 3 Years from 2020 to 2022**