



5/31/2024

CSEL 302

Final Project: DATA ANALYSIS AND VISUALIZATIONS



Topic:
World University Rankings 2023 Analysis

Submitted by:

BSCS -2A

ALCANTARA, ALYSSA MAE D.

PANSOY, ANDREA CLARISSE A.

Table of CONTENTS

01

Project Overview

Key user attributes : 'University Rank', 'Name of University', 'Location', 'No of student', 'No of student per staff', 'International Student', 'Female:Male Ratio', 'OverAll Score', 'Teaching Score', 'Research Score', 'Citations Score', 'Industry Income Score', 'International Outlook Score'

02

Libraries and Data Handling

Libraries used : Pandas, Matplotlib, Seaborn.

Data Loading and preprocessing : Loading from CSV, data cleaning, and categorical data.

03

Analysis Technique Data

Descriptive statistics : Mean, median, count, standard deviation.

Visualization methods : Bar charts, pie charts, and heatmaps

04

Key Findings

User Demographics : Distributions of Students by each locations. Top 10 University based on Citation and Overall Scores. Number of students by each University.

05

Advance Analysis

Academic/ Faculty Reputation, Citation Scores, and the number of International Students



PROJECT OVERVIEW

The World University Ranking 2023 Analysis project aims to comprehensively evaluate the performance of global higher education institutions based on various key metrics. This analysis involves a detailed examination of the latest rankings, considering factors such as 'No of student per staff', 'International Student', 'Female:Male Ratio', 'OverAll Score', 'Teaching Score', 'Research Score', 'Citations Score', 'Industry Income Score', 'International Outlook Score'

By leveraging data from renowned ranking sources, the project seeks to identify trends, strengths, and areas for improvement among top universities worldwide. Additionally, the analysis will highlight regional differences, shifts in rankings compared to previous years, and the impact of new methodologies or changes in criteria. The ultimate goal is to provide insights and recommendations for universities striving to enhance their global standing and to inform prospective students, educators, and policymakers about the current landscape of higher education excellence.

1. **University**- is an institution of higher education and research that awards academic degrees in various disciplines. It provides undergraduate and postgraduate education, conducts research, and often includes various faculties, colleges, and schools. Offers degree programs and courses for students. Research: Conducts academic and applied research.
2. **Ranks** - University ranks in the World University Rankings serve multiple critical purposes for various stakeholders in the higher education ecosystem. For prospective students and their families, these rankings provide a comparative analysis of institutions, helping them make informed decisions about where to pursue their studies based on factors like academic quality, research output, and international reputation. For universities, rankings act as a benchmark for measuring their performance against peers globally, driving them to improve their educational offerings, research capabilities, and international collaborations. Policymakers and educational authorities use these rankings to assess the

effectiveness of their higher education policies and to allocate resources more efficiently.

3. **Location** - The location of a university can influence its accessibility, cultural diversity, and opportunities for international collaboration.
4. **International Students**- The proportion of international students reflects the institution's global appeal and diversity.
5. **Teaching Score**- The teaching score evaluates the quality of education, faculty qualifications, and student-to-faculty ratios, indicating the learning environment.
6. **Research Score** - The research score measures the volume, income, and reputation of a university's research output, showcasing its contribution to innovation and knowledge.
7. **Academic Evaluation**- Academic evaluation combines several metrics, such as peer reviews and academic reputation, providing a holistic view of institutional excellence.
8. **Citations Score**- Citations score assesses the impact and reach of a university's research, highlighting its influence in the academic community.
9. **Industry Income Score** - industry income score gauges the institution's success in attracting funding from industry partnerships, reflecting its practical applications and contributions to the economy. Together, these metrics offer a comprehensive understanding of a university's strengths and areas for improvement, guiding stakeholders in making informed decisions.
10. **International Outlook Score** - The International Outlook Score in university world rankings serves as a critical measure of a university's global engagement and attractiveness to international students and staff. This score encompasses several key factors, including the proportion of international students and faculty, as well as the university's ability to collaborate on research with institutions in other countries.

LIBRARIES AND DATA HANDLING

1. **Pandas** - Pandas is a powerful data manipulation and analysis library that provides data structures like Data Frames and Series. It simplifies the process of data cleaning, transformation, and analysis by offering a wide range of functions to handle missing data, filter and sort datasets, merge and join data from different sources, and perform group operations. Its intuitive and flexible data structures allow for easy handling of structured data, making it a cornerstone for data analysis tasks.
2. **Matplotlib** - Matplotlib is a versatile plotting library for creating static, animated, and interactive visualizations in Python. It allows for the creation of a wide variety of plots, such as line charts, bar charts, scatter plots, histograms, and more. Matplotlib provides extensive customization options, enabling users to tweak every aspect of their plots to produce publication-quality graphics. It serves as the foundation for other visualization libraries, providing the basic building blocks for more complex visualizations.
3. **Seaborn** - Built on top of Matplotlib, Seaborn is a statistical data visualization library that makes it easier to create aesthetically pleasing and informative visualizations. Seaborn comes with built-in themes and color palettes to enhance the visual appeal of plots. It specializes in creating complex visualizations with less code, such as heatmaps, violin plots, box plots, and pair plots. Seaborn also integrates well with Pandas DataFrames, allowing for quick and straightforward plotting of categorical and aggregate data.
4. **Numpy** - Numpy is the fundamental package for scientific computing with Python, providing support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays. It is essential for numerical computations and serves as the underlying engine for data structures in Pandas and other scientific libraries. Numpy enables efficient manipulation of numerical data and is crucial for performance-intensive tasks in data analysis.

Data Cleaning and Reprocessing - Data Cleaning and Reprocessing is an essential step in the data analysis process, particularly in projects like university world ranking analysis. This process involves identifying and correcting errors, inconsistencies, and missing values within the dataset to ensure its accuracy and

reliability. Data may come from various sources, and it's common for it to contain discrepancies or be in a format unsuitable for analysis.

Uploading CSV files - Moreover, CSV files enable seamless collaboration among team members, as they can easily share and exchange datasets for analysis purposes. By leveraging CSV file uploads, analysts can streamline the data analysis process, facilitating the extraction of valuable insights and supporting informed decision-making in university ranking assessments.

Data Analysis and Techniques

Data analysis encompasses examining, cleansing, transforming, and modeling data to uncover valuable insights and support decision-making. Techniques like descriptive statistics summarize dataset features, while inferential statistics make predictions based on samples using hypothesis testing and regression analysis. Exploratory data analysis (EDA) involves visually exploring data with histograms, scatter plots, and box plots to reveal hidden patterns and relationships.

1. **Descriptive statistics** - provide a summary of the main characteristics of a dataset, allowing analysts to understand its central tendency, variability, and distribution.
2. **Mean and median** - are measures of central tendency used to represent the average or typical value of a dataset, with the mean being sensitive to extreme values and the median being more robust to outliers.
3. **Count** - simply enumerates the number of observations in a dataset, providing a basic understanding of its size or sample population.
4. **Standard deviation** - quantifies the dispersion or spread of data points around the mean, providing insights into the variability within the dataset and aiding in the interpretation of the distribution's shape and range.

These statistical measures are fundamental in data analysis and visualization, offering valuable insights into the characteristics and trends present in the data.

Visual Insights

Data visualization plays a crucial role in elucidating the complexities of university world rankings, providing stakeholders with intuitive insights into the relative performance of higher education institutions. Through graphical representations such as bar charts, line graphs, and heatmaps, key metrics such as Universities Location, and Overall Scores can be effectively communicated.

1. **Bar Charts:** Bar charts are commonly used to compare the performance of universities across different metrics or criteria. For example, a bar chart could display the rankings of top universities based on factors like academic reputation, research output, or international outlook. Each bar represents a university, and the height of the bar corresponds to its score or rank in a particular category. Bar charts allow for easy comparison between universities and highlight their relative strengths and weaknesses in various areas.
2. **Pie Charts:** Pie charts are useful for illustrating the composition or distribution of a particular attribute within a dataset. In the context of university rankings, a pie chart could display the percentage of international students or faculty at different universities, providing insight into their global diversity. Similarly, it could represent the distribution of universities by country or region. Pie charts offer a visually intuitive way to understand proportions and percentages, making them valuable for showcasing demographic or geographic data related to university rankings.
3. **Heatmaps:** Heatmaps are effective for visualizing patterns or correlations in large datasets, particularly when dealing with multidimensional data. In the context of university rankings, a heatmap could depict the relationships between various ranking criteria or the performance of universities across different regions. Each cell in the heatmap is color-coded based on the value of the data it represents, allowing viewers to quickly identify trends or outliers. Heatmaps can reveal complex relationships and provide insights into which factors contribute most

significantly to overall rankings, aiding decision-making and strategic planning for universities and policymakers.

Key Findings

Number of Ranked Universities and Locations – Based on the Findings, the top 5 countries with the most number of ranked universities are : United States with 173 ranked universities, Japan with 150, United Kingdom with the number of 149, India with 91 and China has 82. Contrary with that, Puerto Rico, Mozambique, Mauritius, Namibia and Kosovo only has 1 university each that are included in the ranking.

Distribution of Students by Location – United States and other Western Countries have the most student per university and locations. Some of the countries aside from the countries that mentioned with the most ranked universities, Turkey, Brazil, Iran, Spain, Nigeria, Pakistan and Germany also have a significant number of University students.

Top Universities based on Overall Score – Based on the results, the Top 1 university is the University of Oxford in United Kingdom, Top 2 is Harvard University in United States, and Top 3 is the University of Cambridge in United Kingdom. With this findings, the scores of each university in Citations, Research, Industrial Income and also the No. of students and International students have a correlation why they are in the Top Universities across the world.

Influence on Business Decisions or Strategies

Talent Acquisition: High-ranking universities often attract top-tier talent, including graduates with sought-after skills and expertise. Businesses may use university rankings as a guide for recruiting strategies, targeting graduates from prestigious institutions known for producing high-caliber professionals

Partnerships and Collaborations: Businesses may seek collaborations with top-ranked universities to leverage their research capabilities, expertise, and resources.

Educational and Professional Development Programs: Businesses may collaborate with top-ranked universities to offer educational and professional development programs tailored to their employees' needs. These programs can range from executive education courses and

leadership development programs to customized training initiatives focused on specific industry trends or technologies.

Strategic Location and Expansion: Companies looking to establish new operations, research facilities, or innovation hubs may consider the presence of top-ranked universities as a key factor in selecting strategic locations.

Overall, world university rankings can serve as a valuable reference point for businesses in making informed decisions and formulating strategies related to talent acquisition, partnerships, brand positioning, innovation, and geographic expansion. By leveraging the reputation and capabilities of top-ranked universities, businesses can gain a competitive edge and drive sustainable growth in an increasingly globalized and knowledge-driven economy.

Advanced Analysis

Geographical Insights: Through the application of geographical analysis, it is possible to find places with untapped potential and excellence clusters by analyzing spatial trends in university rankings. The distribution of highly regarded universities across geographical areas can be seen using heatmaps, which also emphasize areas of rising academic significance.

Temporal Trends: Analyzing temporal trends in university rankings over time can unveil evolving patterns and trajectories of academic excellence. Time-series analysis can track the rise and fall of individual universities, identifying drivers of change and assessing the impact of strategic initiatives or policy reforms.

Predictive Modeling: Leveraging advanced statistical techniques and machine learning algorithms, predictive modeling can forecast future trends in university rankings based on historical data and key performance indicators. Predictive analytics can help universities and policymakers anticipate shifts in rankings, identify opportunities for improvement, and formulate proactive strategies to enhance competitiveness and reputation.

Network Analysis: Network analysis techniques can uncover hidden relationships and interdependencies among universities, faculty, research institutions, and industry partners, elucidating collaborative networks and knowledge exchange pathways.

Conclusion

In conclusion, the World University Ranking 2023 Analysis offers a deep dive into the global landscape of higher education, revealing trends, strengths, and areas needing attention among the world's top-ranked institutions. By delving into key metrics such as, 'No of student per staff', 'International Student', 'Female:Male Ratio', 'OverAll Score', 'Teaching Score', 'Research Score', 'Citations Score', 'Industry Income Score', 'International Outlook Score' this analysis provides a nuanced understanding of what makes universities excel.

It's clear that global engagement and diversity play pivotal roles in fostering academic brilliance and driving innovation. Universities that embrace internationalism, boasting diverse student bodies and faculty cohorts, hold a distinct advantage in attracting top talent and nurturing a dynamic intellectual atmosphere conducive to cross-cultural collaboration.

In addition, the analysis highlights geographic variations and new trends in higher education, providing stakeholders, politicians, and academic institutions with useful information to improve their reputation internationally. Through the identification of strengths and areas that require enhancement, this analysis functions as a valuable tool for academic institutions aiming to prosper in the dynamic global academic landscape.

In the end, the World University Ranking 2023 Analysis serves as an invaluable resource for anyone navigating the complexity of higher education, enabling them to strategically invest and make well-informed decisions that promote equitable growth, excellence, and innovation on a global scale. In the future, colleges will need to keep analyzing and working together in order to achieve even more and make significant contributions to society and the growth of knowledge.