Kahoot! challenge autumn 2023

Data & Insights Intern

Brage Helgesen Jåsund

GitHub-link: https://github.com/BrageJs/Kahoot-challenge.git

Table of Contents

Introduction

The objective of this analysis was to provide insights to Kahoot!'s market research team on which universities, countries, or regions to target for marketing and sales efforts. Utilizing the "World University Rankings" dataset available on Kaggle, we analyzed rankings of universities around the world from three different publishers: Times, Shanghai, and CWUR. This dataset also provided supplementary demographic data, which was leveraged to draw correlations and insights. The aim was to identify top universities and regions, understand the differences in rankings, and correlate university rankings with factors such as direct expenditure on education.

Data Import and Cleaning

Datasets Overview:

Six datasets were utilized for this analysis:

1. CWUR Rankings Data (cwurData.csv): Provides university rankings based on the Center for World University Rankings.
2. Times Rankings Data (timesData.csv): Contains university rankings from the Times Higher Education World University Rankings.
3. Shanghai Rankings Data (shanghaiData.csv): Offers university rankings from the Academic Ranking of World Universities (often referred to as the Shanghai Rankings).
4. Education Expenditure Data (education_expenditure_supplementary_data.csv): Gives information on direct expenditure on education for various countries.
5. Educational Attainment Data (educational_attainment_supplementary_data.csv): Contains data on educational attainment levels for different countries.
6. School and Country Data (school_and_country_table.csv): Maps universities to their respective countries.

Data Cleaning:

The initial step in the analysis involved cleaning and preprocessing the data to ensure accuracy and reliability of the results. The following steps were undertaken:

Handling missing values: Rows with missing or incomplete data were identified and dealt with appropriately to ensure the integrity of the analysis.
Preprocessing the 'world_rank' column: Some rankings were presented as ranges (e.g., 101-150). These were processed to obtain a single representative rank for each university.

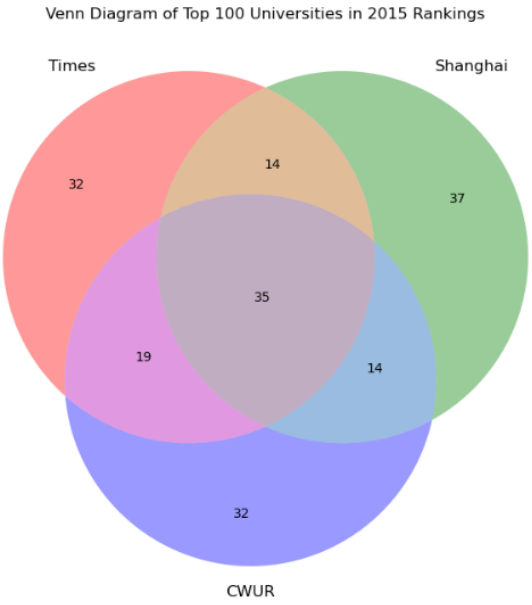
Top 100 Universities Analysis

Methodology:

The analysis began by extracting the top 100 universities from each of the three ranking systems for the year 2015. This provided a snapshot of the leading institutions as per each ranking system.

Findings:

- **Common Universities:** An overlap analysis revealed that several universities consistently ranked in the top 100 across the three ranking systems. This indicates the global recognition and reputation of these institutions.
- **Venn Diagram Analysis:** A Venn diagram was created to visualize the overlaps between the top 100 universities from each ranking. This provided a clear picture of universities that are universally recognized versus those that are favored by a particular ranking system.



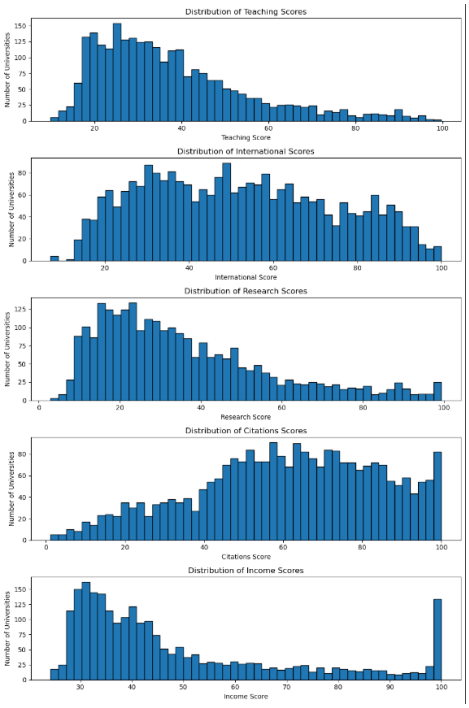
Score Distributions

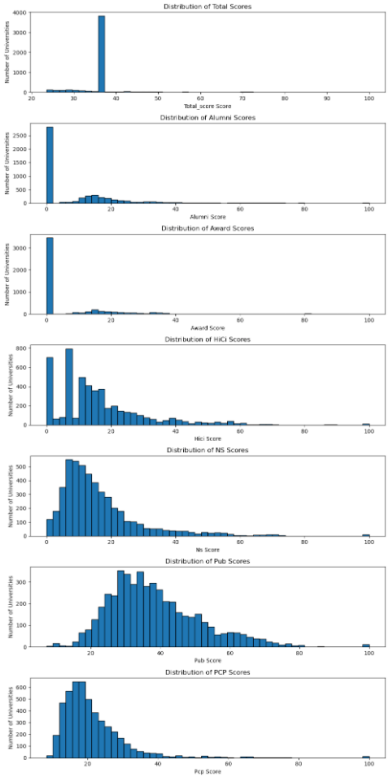
Overview:

To understand the distribution of scores that contribute to the overall rankings, we visualized the score distributions for various metrics across the three ranking systems.

Findings:

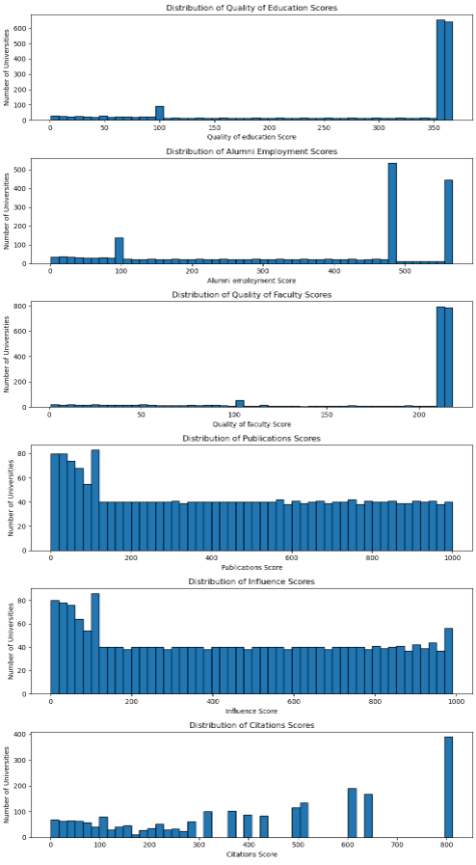
- **Times Dataset:** The distributions for teaching, international, research, citations, and income scores were plotted. This provided insights into the spread and concentration of scores for each metric.





- **Shanghai Dataset:** Distributions for total score, alumni, award, HiCi, NS, Pub, and PCP scores were visualized. The plots highlighted the range and distribution of scores that universities achieved in each category.

- CWUR Dataset: The distributions for quality of education, alumni employment, quality of faculty, publications, influence, and citations scores were showcased. The visualizations helped in understanding the spread of scores and identifying any patterns or outliers.



Correlation Analysis

Objective:

The goal was to identify which metrics most influence the overall ranking or score for each ranking system. A correlation analysis was performed to achieve this.

Findings:

- Methodology: Pearson correlation coefficients were calculated between the overall ranking/score and other metrics for each ranking system.
- Top Correlations: For each ranking system, the top 3 most correlated metrics were identified. These metrics are pivotal in determining the overall ranking of a university and provide insights into areas of strength or improvement for institutions.

```
Top 3 Correlations for Times Dataset:
citations   -0.803925
research    -0.759342
teaching    -0.716650
Name: average_rank, dtype: float64

Top 3 Correlations for Shanghai Dataset:
award    0.560591
ns       0.529876
pcp      0.488366
Name: total_score, dtype: float64

Top 3 Correlations for CWUR Dataset:
score                -0.549098
broad_impact          0.554597
quality_of_faculty    0.663864
Name: world_rank, dtype: float64

Top 3 Correlations Overall:
publications    0.923037
influence       0.895871
citations       0.856573
dtype: float64
```

Country Analysis

Data Merging:

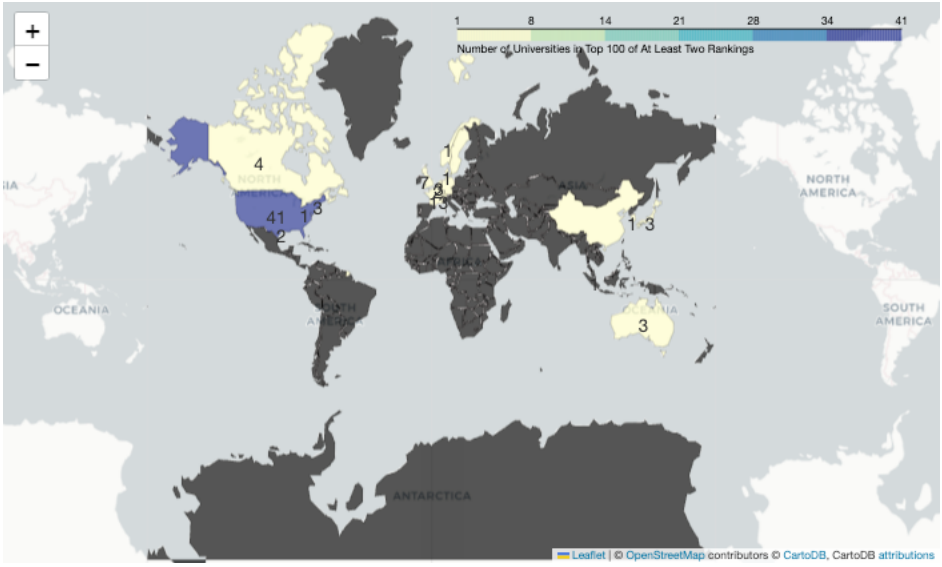
To gain a geographical perspective, the ranking datasets were merged with supplementary data to associate universities with their respective countries.

Findings:

- **Correlation with Expenditure:** A correlation analysis was performed to understand the relationship between university rankings and direct expenditure on education for each country.

```
(-0.03568328806187922, -0.006975814890098521, -0.033979651415346375)
```

- **Top Universities Visualization:** A world map was created to visualize the number of top universities in each country, providing a global view of educational excellence.



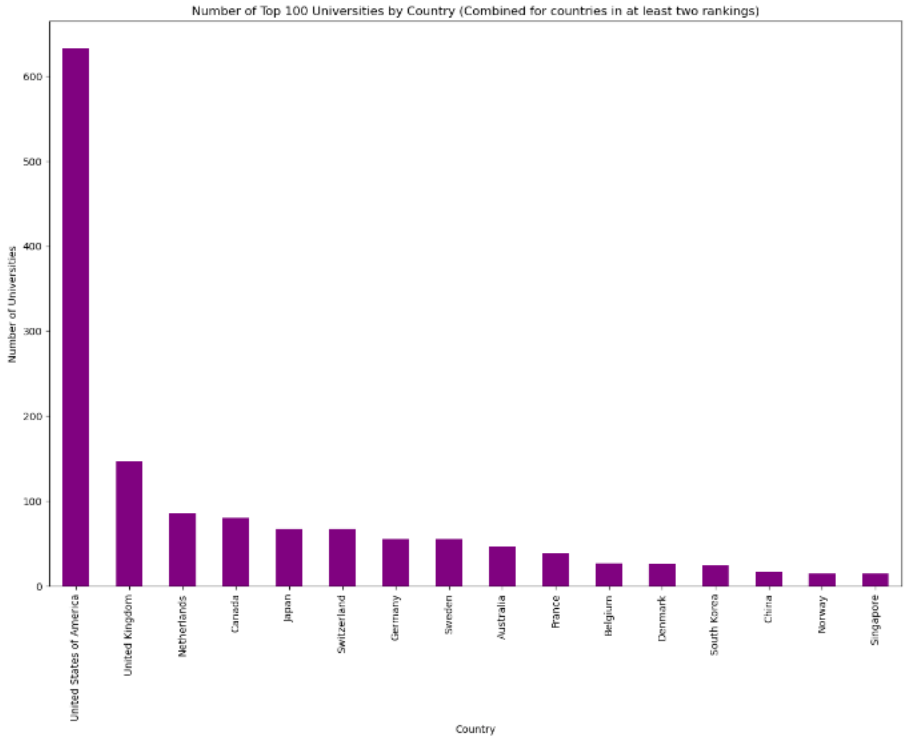
- **Top 10 Countries:** The top 10 countries with the highest average university rankings and the highest direct expenditure on education were identified. This analysis provided insights into countries that prioritize education and have leading institutions

```
country
United States of America     71.847222
France                       80.250000
Hong Kong                    80.500000
Switzerland                  83.666667
Republic of Ireland          85.000000
Australia                    90.571429
Japan                        91.400000
South Korea                 101.500000
Finland                     102.000000
Singapore                   104.000000
Name: world_rank, dtype: float64
```

	country	2011
229	Denmark	7.9
236	Iceland	7.7
241	Korea, Republic of	7.6
245	New Zealand	7.5
7	Denmark	7.5
24	Norway	7.3
238	Israel	7.3
256	United States	6.9
14	Iceland	6.9
225	Belgium	6.6

- **Comparison of Rankings:** A bar chart showcased the number of top 100 universities by country, combining data from all three rankings. This visualization highlighted

countries with a consistent presence in the top 100 across different ranking systems.



Conclusion

The comprehensive analysis of the "World University Rankings" dataset provided valuable insights into the global landscape of higher education. Several key findings emerged:

1. **Consistent Top Performers:** Certain universities consistently ranked in the top 100 across different ranking systems, highlighting their global recognition and excellence in education.
2. **Score Distributions:** The visualizations of score distributions for various metrics across the ranking systems provided insights into the spread and concentration of scores, revealing areas of strength and potential improvement for institutions.
3. **Influential Metrics:** The correlation analysis identified pivotal metrics that significantly influence the overall ranking of a university. These metrics are crucial for institutions aiming to improve their global standing.
4. **Geographical Insights:** The country analysis revealed nations that prioritize education both in terms of quality (as indicated by top university rankings) and quantity (as indicated by direct expenditure on education).

Recommendations

Based on the findings, the following recommendations are proposed for Kahoot!'s marketing and sales efforts:

Universities:

1. Consistently Top-Ranked Universities:

- Universities that consistently appear in the top 100 across all three ranking systems should be primary targets. These institutions have a global reputation for excellence and associating with them can significantly enhance Kahoot!'s brand identity.
    - Examples: Harvard University, Stanford University, University of Cambridge, and Massachusetts Institute of Technology (MIT).

Countries:

1. High Direct Expenditure on Education:

- Countries that have a high direct expenditure on education indicate a strong commitment to the educational sector. These countries are likely to be more receptive to e-learning platforms like Kahoot!.
    - Examples: Norway, Switzerland, and the United States.

2. High Average University Rankings:

- Countries with a high average university ranking indicate a strong educational infrastructure and a commitment to academic excellence.
    - Examples: United States, United Kingdom, and Germany.

3. Presence in Top 100 Across Multiple Rankings:

- Countries that have multiple universities in the top 100 across different ranking systems show a consistent commitment to quality education.
    - Examples: United States, China, and the United Kingdom.

Regions:

1. **North America:**
    a. Given the consistent high rankings of universities in the United States and Canada, North America remains a prime region for marketing and sales efforts.
2. Western Europe:
    a. Countries like the United Kingdom, Germany, Switzerland, and the Netherlands have universities that consistently rank high. Western Europe, with its strong educational infrastructure and high expenditure on education, is a key region to target.
3. Asia:
    a. With the rise of universities in China, Singapore, and Japan in global rankings, Asia presents a significant opportunity. The region's emphasis on education

and technological advancement makes it a promising market for e-learning platforms.

Reasoning:

- **Reputation & Branding:** Associating with top-ranked universities can significantly enhance Kahoot!'s brand identity as one of the best e-learning platforms.
- **Market Potential:** Countries and regions with a strong emphasis on education and technological advancement present a larger market potential for e-learning platforms.
- **Strategic Positioning:** Targeting regions with multiple top-ranked universities can provide a strategic advantage, allowing Kahoot! to establish a strong foothold in the global educational sector.