# Global Rankings of Universities Analysis using R Studio

Kothuru Devi Prathyusha, Arvind Devkate, Himanshu Panwar, Aaryan Maheshwari

## Introduction

- In this analysis, we explore the global rankings of universities, examining how institutions from around the world are positioned based on various metrics. University rankings are often used by students, faculty, and policymakers to evaluate the prestige, quality of education, research capabilities, and overall global standing of universities. These rankings are compiled by different organizations, like Quacquarelli Symonds (QS) and Times Higher Education (THE) etc. and typically take into account factors such as academic reputation, research output, faculty quality, and international diversity and more. We also aim to identify how these top universities are spread across the globe and lastly, where do IITs stand in these rankings and the potential areas of their improvement.

- The dataset used for this analysis includes ranking data from **QS rankings 2025** and **Times Higher Education (THE)** rankings for the top **1000** universities of the world. Through data cleaning, transformation, and visualization techniques in R, we aim to uncover meaningful insights that help understand the dynamics of university rankings. We have also used GDP and population data of various countries, as listed by **worldometer**.

Here is a sample of the finally obtained dataset of the first five Universities and Overall Ranking.

```
##                                    Title Overall
## 1 Massachusetts Institute of Technology (MIT)   100.0
## 2                      Imperial College London    98.5
## 3                        University of Oxford    96.9
## 4                          Harvard University    96.8
## 5                      University of Cambridge    96.7
```

## Data Scraping

For data scraping we used the following sites:

- **Times Higher Education**: We used this site for scraping University Names and Female Student Ratio **(in %)**.

- **Top Universities**: We used this site for scraping the following: Overall Score, Title, Country, Region, Rank, Academic Reputation, Citation per Faculty, Faculty Student Ratio, Employer Reputation, International Student Ratio, International Research Network, International Faculty Ratio and Sustainability of the QS rankings 2025.

- **Worldometer**: We used this site for scraping the following: Countries, GDP, and Population.

## Data Cleaning and Transformation

- Firstly, to compile all the data obtained into a single dataset, we had to **rename** some columns and country names (for example, from China(Mainland) to China) so that all datasets were compatible with each other.

- Then we **merged** the data about female student ratio from THE rankings data into the QS rankings data. In this task, there were some missing values, as the female student ratio data was not exhaustive. So we **replaced** the null values by the **median** of the other obtained values.

- Lastly, we found out the **GDP per capita** data by dividing the gross GDP and population of each country. Then we **scaled** the GDP per Capita data on a scale of 0-100 so as to compare this with the scores of other columns, all of which were scored out of 100. For this we firstly applied a **log** transformation and then **linearly** scaled these values.

## Potential Biases in the Data

Here are some potential biases that we think could have been in our data:

- **Regional Bias**: QS and other ranking systems can favor universities in certain regions, often Western countries. This can result from differences in funding, research output visibility, and the resources available to institutions in wealthier regions.

- **Alumni Outcomes and Employability**: Rankings that consider employment outcomes and salary levels post-graduation may favor universities in countries with stronger job markets, thereby skewing the ranking toward institutions in these locations.

- **Temporal Bias**: QS Rankings and others may not account fully for recent changes, improvements, or declines in universities, as reputation and citation data can take years to reflect the current status of a university.

- **Exclusion of Smaller or Less Prestigious Institutions**: By focusing only on the top 1000, the rankings may have inherently excluded universities that might excel in specific areas or serve unique communities but don't rank highly overall. This can our analysis by highlighting the features of already well-established institutions while overlooking smaller or newer universities that might offer high-quality education and innovation in certain fields.

## Questions to ask from the data

Here are some interesting questions that could be asked from the data:

- What is the geographical spread of the top 1000 universities of the world?

- How do different parameters such as academic reputation, employer reputation ,international student ratio etc. affect ranking and what are their general trends with rank?

- How does GDP and population of a country influence the ranking of its top universities?

- What is the composition of female students across top universities?

## Important Visualizations

- **Number of Universities v/s Country graph, colored by region**

  A country-wise number of universities graph provides a visual comparison of the number of higher education institutions across different nations. It is also coloured by the regions of the countries.

  As we can see, countries like **United States**, **United Kingdom** and **China** have a higher number of universities, with the U.S. leading globally with 100+ institutions out of the top 1000. Also, notably, countries from all regions are present among the top countries except from the **African** region. This shows that the quality of university in Africa tends to be worse than the quality of their counterparts from other regions of the world.



Number of Universities by Country (All Regions)
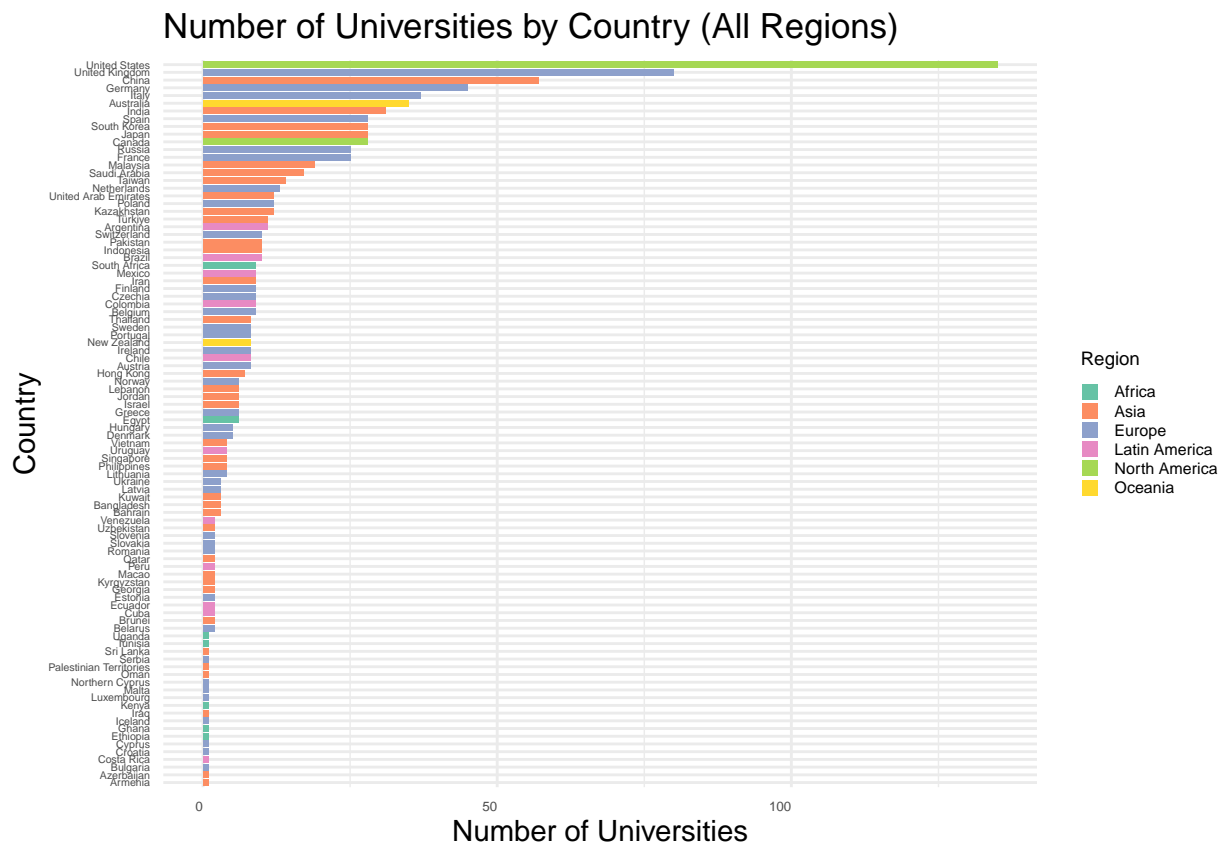
- **Academic Reputation vs Rank Plot**

  As we can see that the trend is **decreasing** with increase in rank. This means the two variables Rank and Academic Reputation have a **highly negative correlation**.

## Scatter Plot of Rank vs Academic Reputation
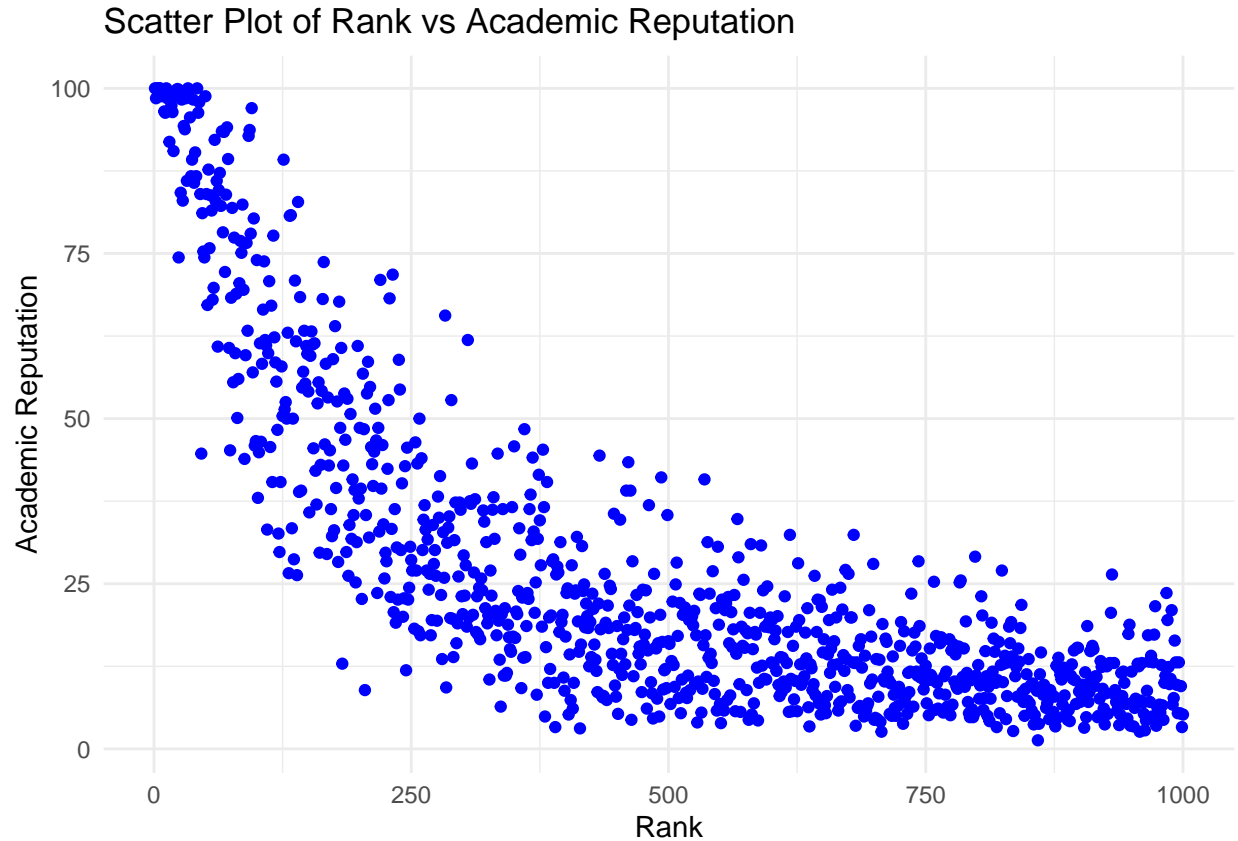


- **Citations per Faculty vs Rank Plot, with top 5 IITs coloured in red**

  We can see that overall there is a **decreasing** trend, but there are many **outliers**. This suggests that in lower ranked universities, even if their overall ranking is low, citations per faculty can be high. This implies that there are some labs in these universities with considerable research output but **other labs lag behind**, leading to high citations but low overall score. We believe this is also the case with **IITs** in general, as can be observed (IITB, IITD, IITK, IITM, IITKGP are coloured in red.)

## Scatter Plot of Rank vs Citations per Faculty
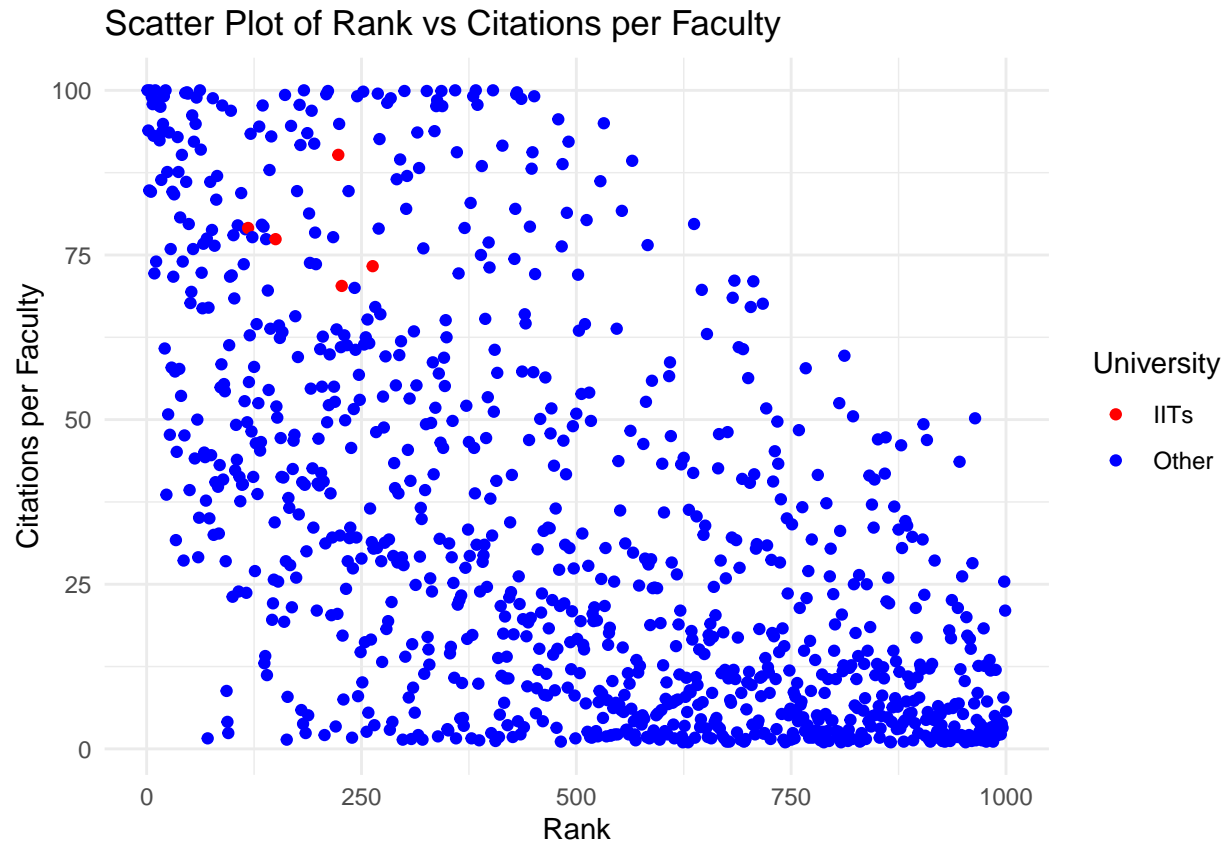


- **Employer Reputation vs Rank plot, top 5 IITs colored in red**

  Here also there is an overall **decreasing** trend but many **outliers**. This suggests that some brilliant students, though are very talented, cannot obtain admissions to top universities but prove to be excellent employees, as is shown by high employer reputation. We believe **IITs** have a similar status.

## Scatter Plot of Rank vs Employer Reputation
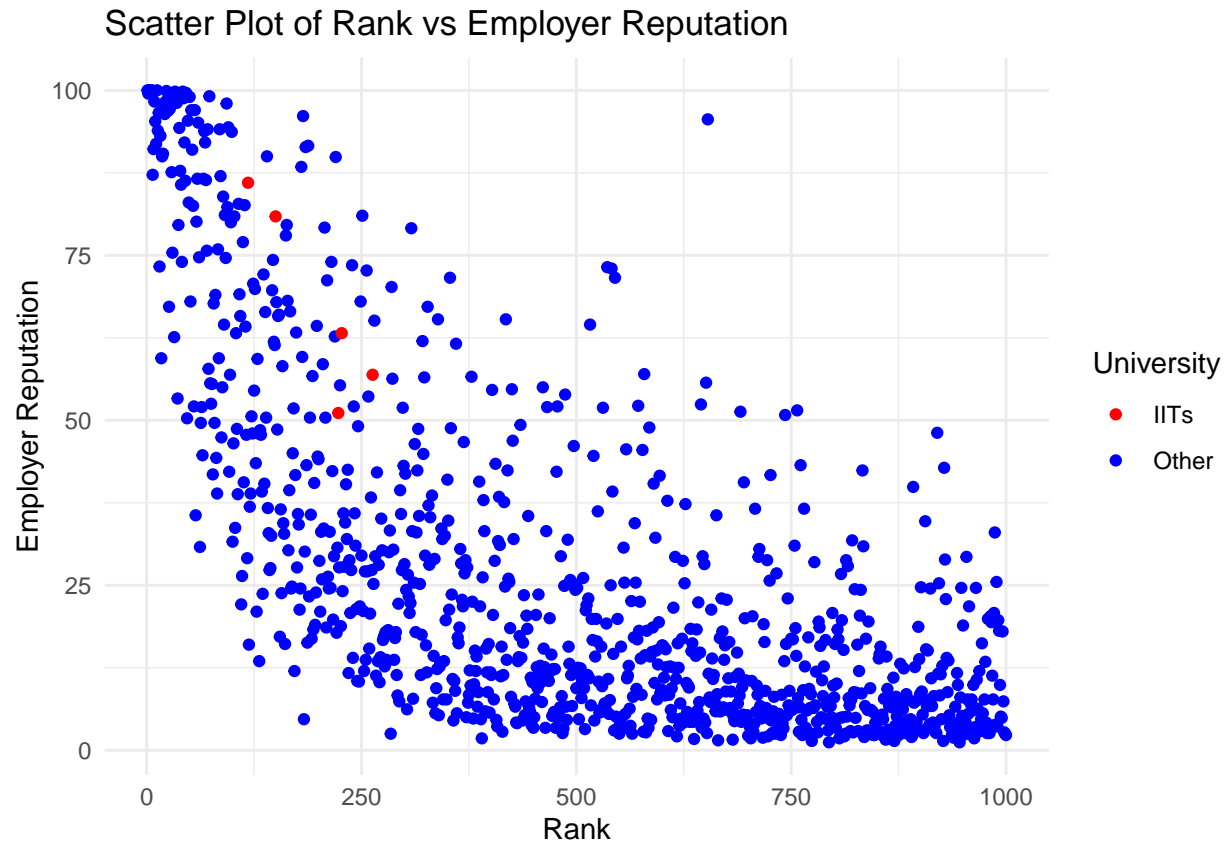


- **International Student Ratio vs Rank, with IITs coloured in red**

  We can see that all 5 IITs have a **very poor score** with respect to top ranked universities. This suggests an **area of improvement** to boost the rankings of our premier institutes. A similar observation is obtained when we plot **international student ratio** with rank.

## Scatter Plot of Rank vs International Student Ratio
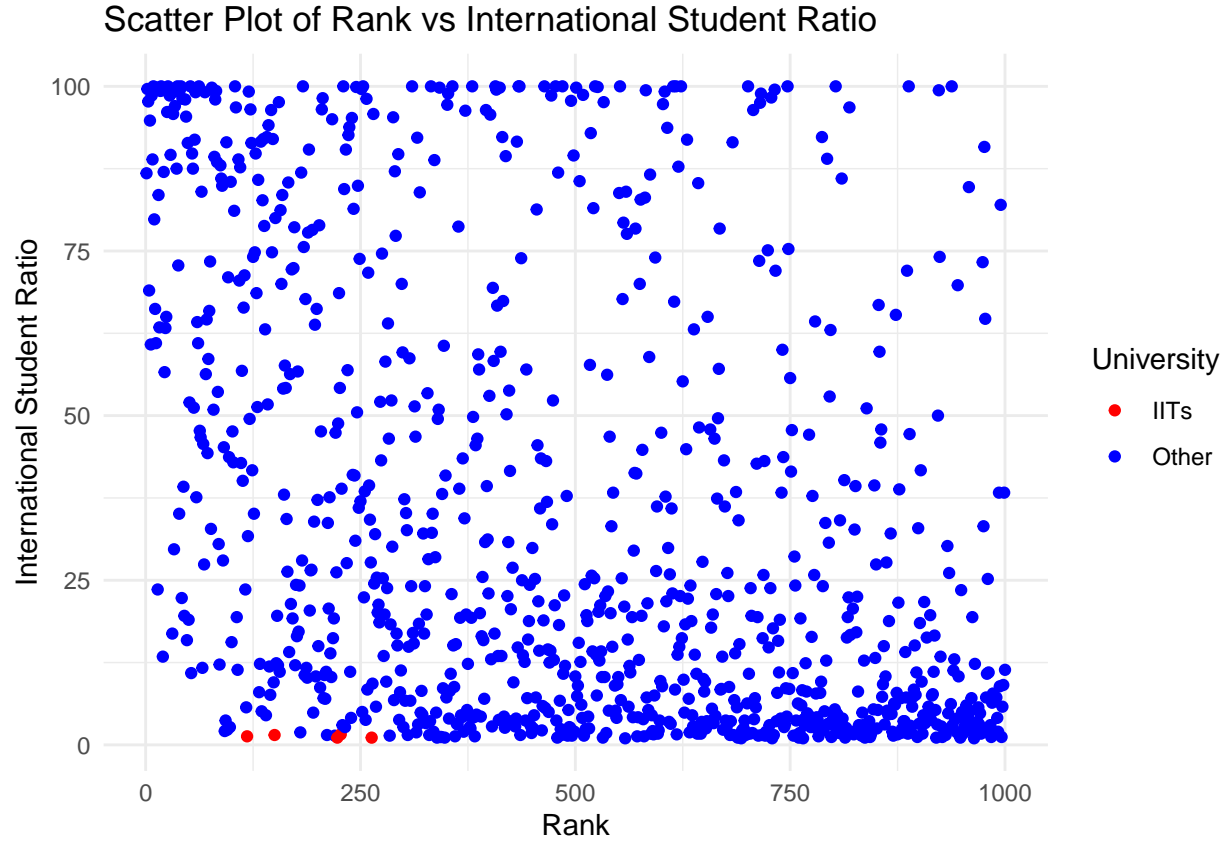


- **Average University Ranking Vs GDP Per Capita Plot**

  We have plotted the Country and the GDP Per Capita (by firstly creating a dataframe) in the order of the Ranking of the Universities. We can see the Average ranking of the University of each Country is decreasing, the GDP is also decreasing.

| Country | GDP | Average Rank |
| --- | --- | --- |
| United States | 89.24293 | 152.5000 |
| United Kingdom | 78.90087 | 169.7143 |
| United Kingdom | 78.90087 | 174.2308 |
| United States | 89.24293 | 184.2500 |
| United Kingdom | 78.90087 | 197.0000 |
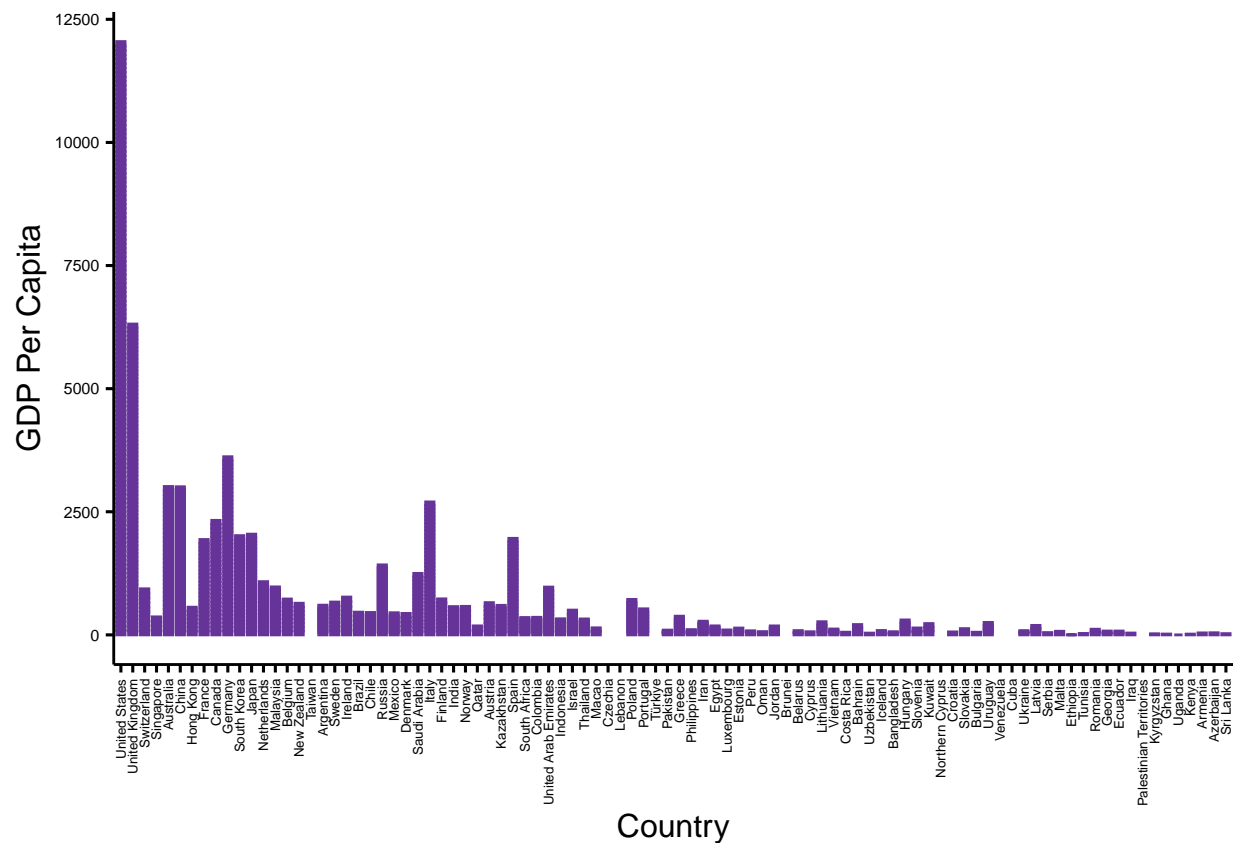
- **Plot for Country and the Average Rank of University for each Country.**

  We can see there is a direct correlation with the GDP and the Average Rank of the Universities for each Country.
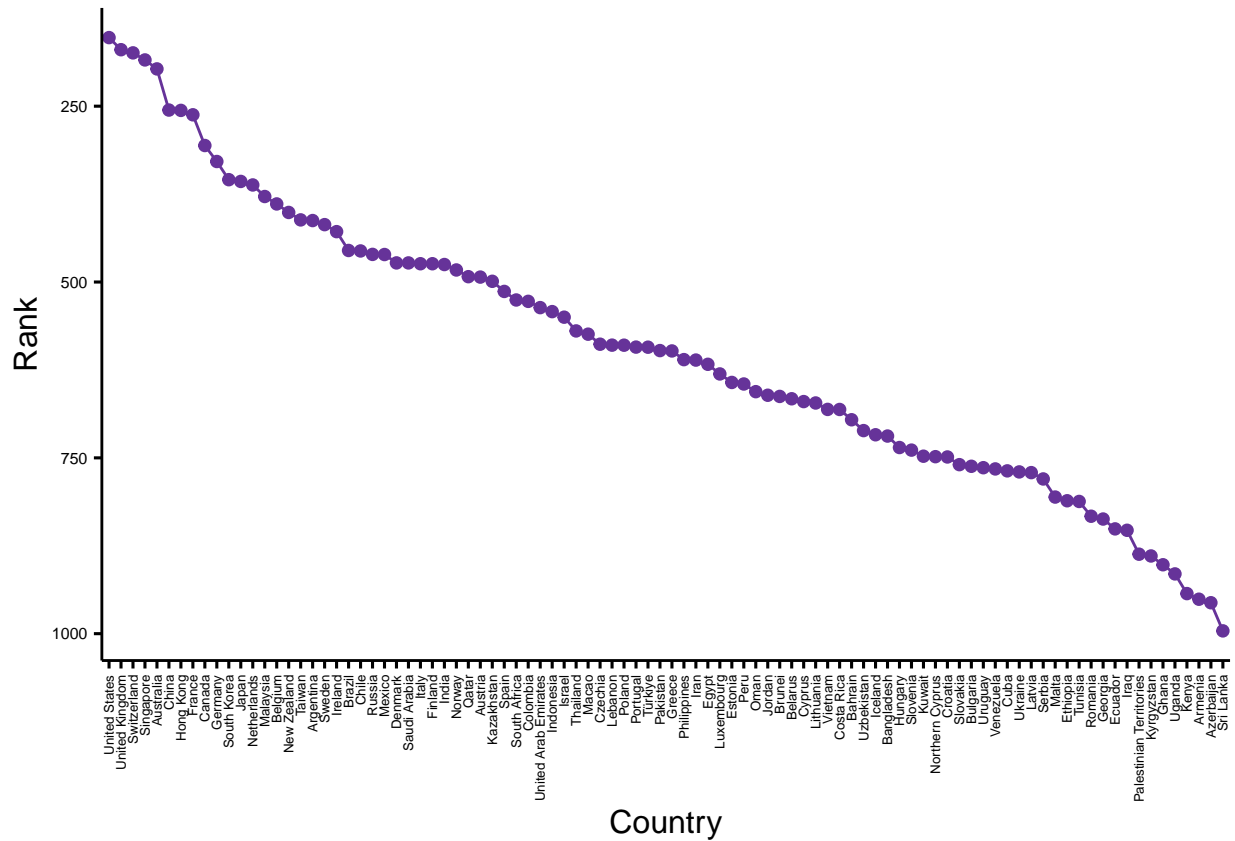
- **Density Plot for Female Student Ratio for the first 50 Universities**

  Here we can see the data is mostly scattered, but there is a huge **peak** at around 55%. This means there is a high percentage of females in the top 50 universities. This fact is correlated to the fact that most top 50 universities are from **North America** and **Europe**, which are more favourable places for women in STEM, as compared to other regions.
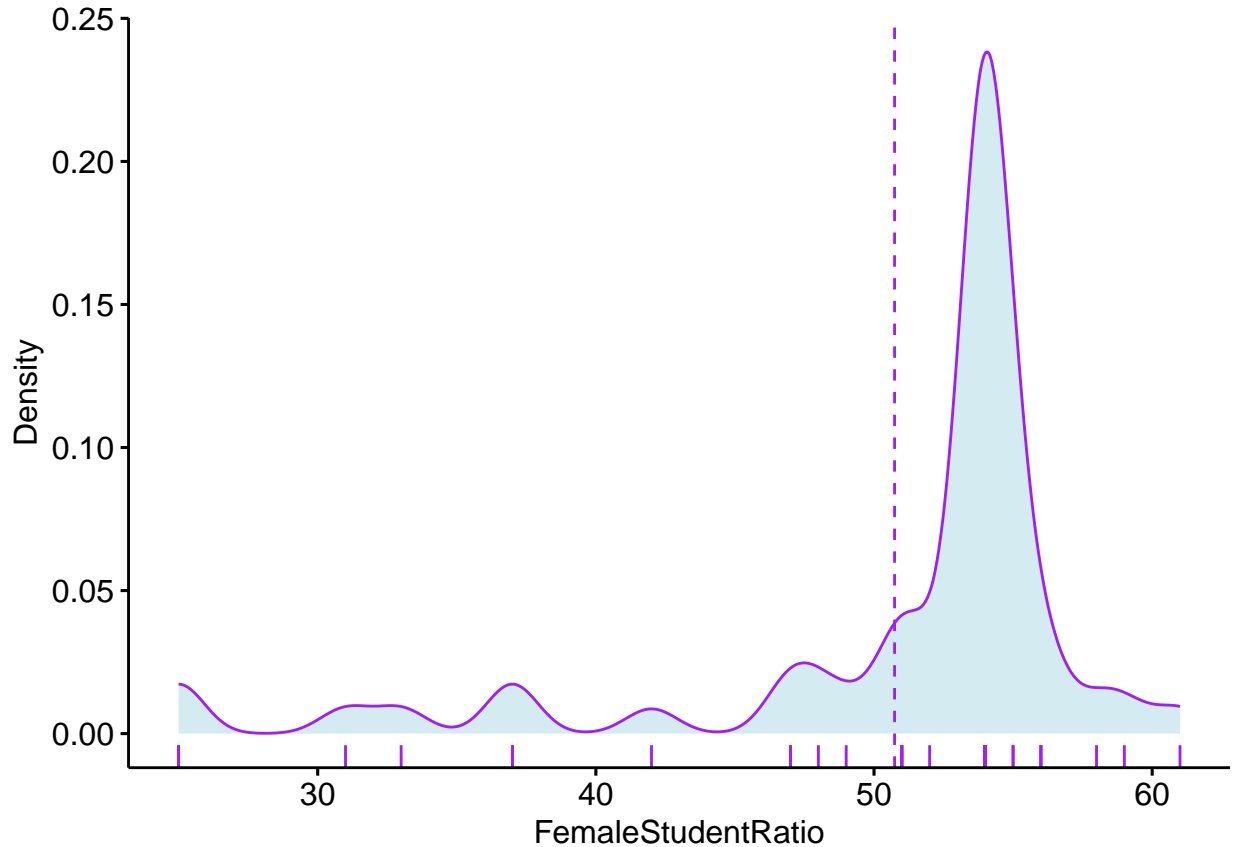
## Final Conclusions

- Most top universities of the world are concentrated in a few countries (viz. USA and UK). African universities lack severely in quality and are in no competition to their american or european counterparts. This also represents the fact that larger and well-funded institutions typically perform better in rankings, as they have more resources for research, faculty hiring, and infrastructure. Smaller universities with fewer resources can struggle to meet the same benchmarks, regardless of quality.

- Rankings give more weight to academic and employer reputation. There are many outliers though, which have high academic and employer reputation but still have poor ranking due to their low scores in other categories. It is also interesting to note that this reputation-bias can overshadow newer universities or those with regional influence, potentially distorting perceptions of quality and impact.

- For IITs to improve their rankings, they should increase their focus on parameters such as international student ratio and international faculty ratio, as both of these prove to be detrimental to the overall ranking of our prestigious institutions. These can be achieved by increasing exchange programmes and enabling cross country collaboration, etc.

## References

- Times Higher Education.

- Top Universities.

- Worldometer.