Report : "Data Analysis of The Times Higher Education and World Development Indicators: Insights and Findings"

DOPP Group 11: Ilija Filipovski, Marko Pajkovski, Dario Jugo, Mirjana Sadikovikj

1. Introduction

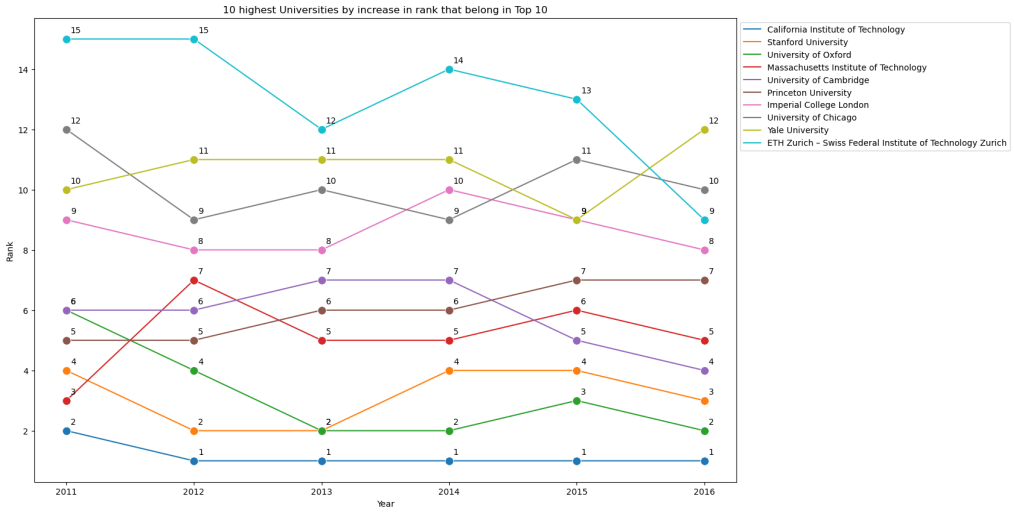
This report aims to present the key insights and findings from the data analysis of two datasets: The Times Higher Education (THE) and World Development Indicators (WDI). The goal of the analysis is to answer a specific question and build a predictive model. The THE dataset provides performance data on universities, while the WDI dataset includes a wide range of development indicators.

2. Data Preprocessing

Data preprocessing was performed on both datasets in order to clean, reshape, and add new columns. For the WDI dataset, the data was reshaped from wide to long format and then pivoted back to wide format. The 'Year' and 'Country Name' columns were cleaned and a new column called 'continent' was added by mapping the country name to the continent name. For the THE dataset, wrong country names were replaced and matching country codes were added using the 'pycountry' library. Missing values were handled by replacing them with median or mean values of the specific column grouped by "country code" and use of fillna() function. Additionally, the optimal weights for dependent variables and a new column "total_score" was created to store the calculated score.

3. Key Insights and Findings

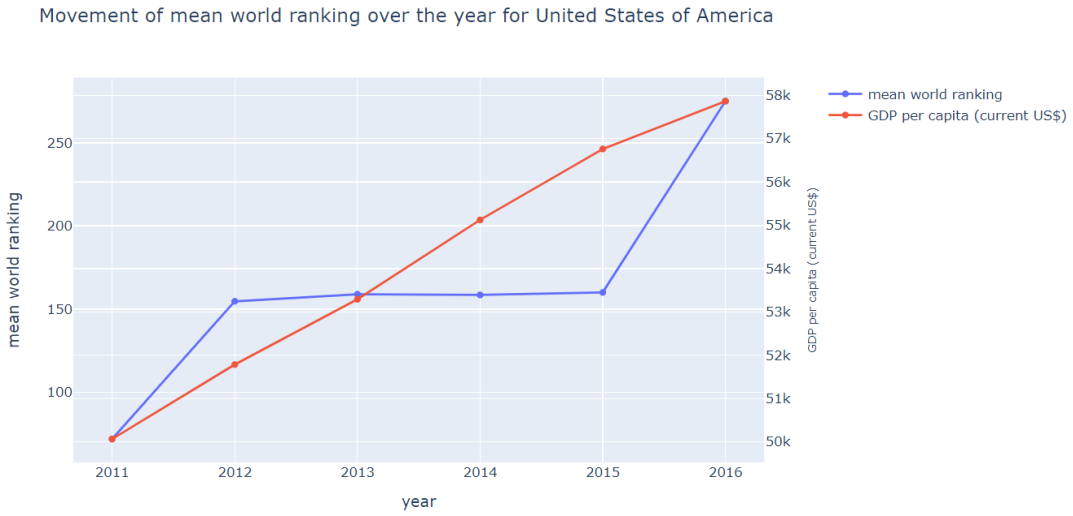
It can be seen that ETH Zurich and University of Oxford have had the highest increase in ranking for the universities belonging to the top 10, with ETH Zurich increasing its rank by 6 places while Oxford 4 places.



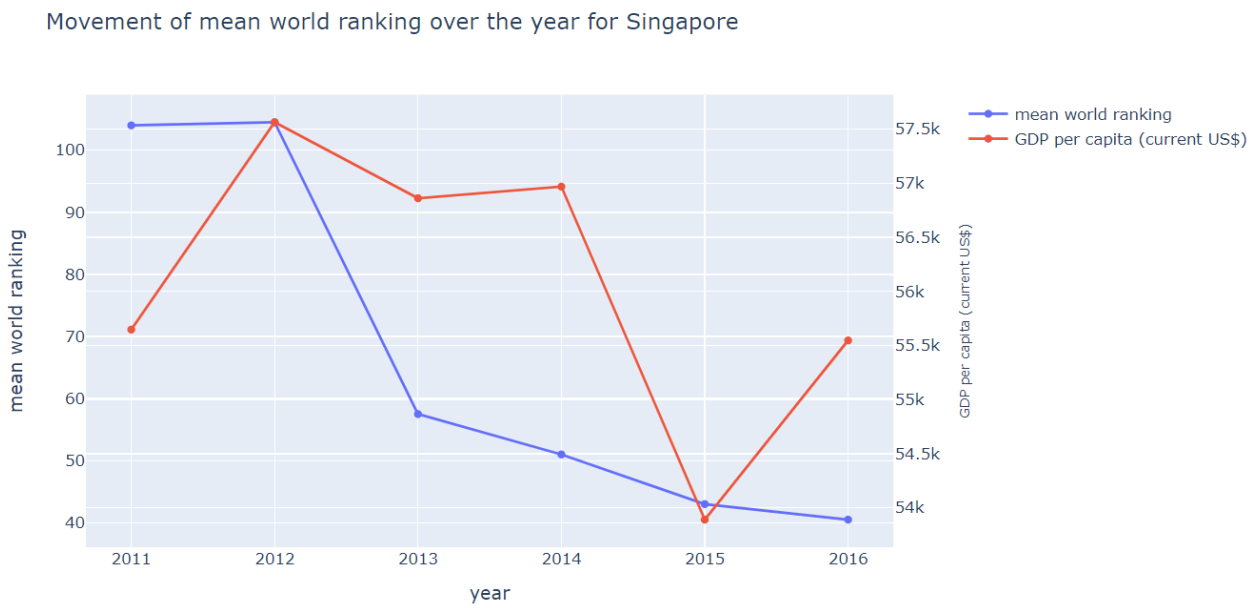
In the following plots we will introduce a new variable which is called “mean world rank” for each country. Here we are grouping instances by the country (for each year separately) and calculating the mean ranking for the years. We want to look at how different indicators are impacting this mean world rank. In order to perform this we took three different countries to present this idea (USA, Singapore and Switzerland) and we took following indicators (only second one will be shown in the report):

1. GDP per capita (current US$)
2. Government expenditure on education, total (% of GDP)
3. Current education expenditure, total(% of total expenditure in public institutions)

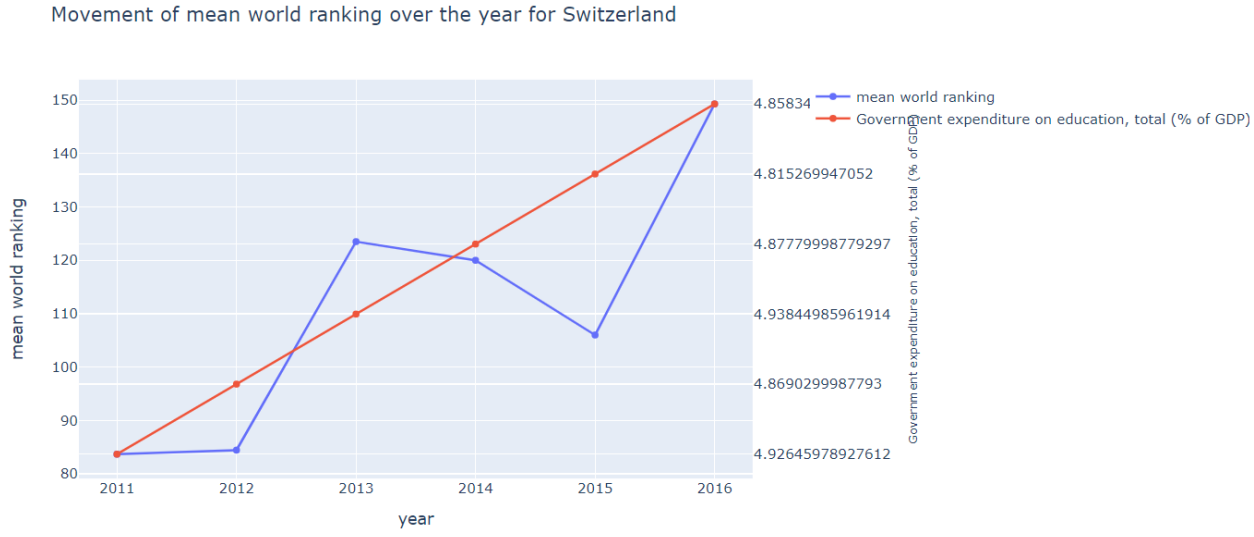
First we will look at universities in the USA. From the plot below we can see that even if the USA has bigger Government expenditure on education, total (% of GDP) over the years their rankings are also getting bigger (which is not a positive effect because that means that universities are not improving).



Now we will present a positive effect for universities from Singapore which had higher Government expenditure on education, total (% of GDP) which is also influencing mean world rank in a negative way (which means that rank is getting smaller which actually means that universities are getting better).

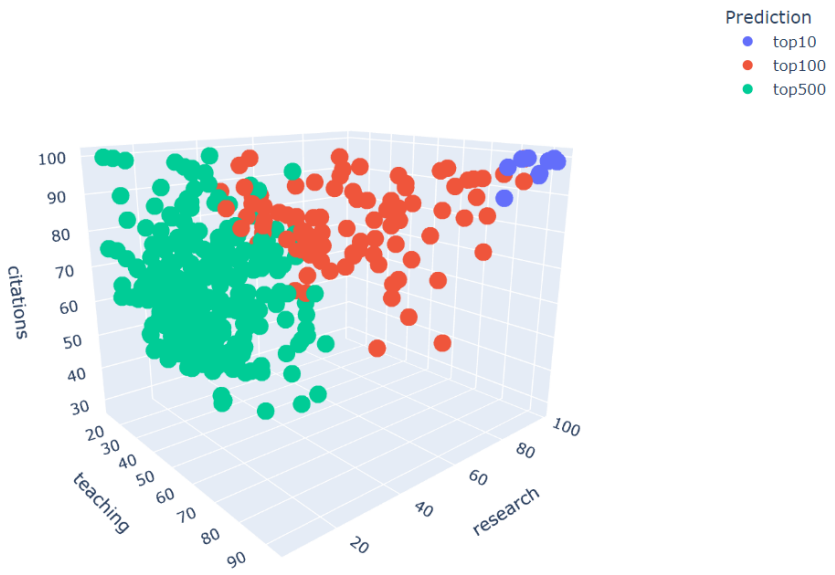


Here we are showing how the ranking of Switzerland has changed over the years. In this example we cannot see a clear trend in terms of ranking and Government expenditure on education, total (% of GDP). We can see that in 2015 rank got better with growth of Government expenditure on education, total (% of GDP) but next year the rank got worse.



The primary factors that determine a university's ranking are its performance in teaching, research, and citations. These factors are consistent across countries and are considered to be the most significant predictors of a university's overall performance. Analysis of the top 10

universities in the rankings reveals that they consistently exhibit high scores in all three areas, indicating exceptional performance in each field. For universities that are ranked within the top 100, it is essential to either excel in one or two of these areas, or to maintain above-average scores across all three.



4. Conclusion

Using this analysis it can be seen that European universities are increasing their ranking through the years, with 11 out of 20 of the top growers in the top 11-500 universities in the world being located in Europe. Moreover, from the analysis it can be seen that in order to be in a top 100 ranking universities should increase their teaching capabilities and strive to provide high-quality education to students, by investing in faculty development, student support services and innovative teaching methods. Moreover, they should invest in research infrastructure, and encourage faculty and students to engage in high-impact research. This will help to increase the number of publications and citations, which are key indicators of a university's research performance.

Moreover, it was expected that increasing investment in education would lead to an improvement in mean world rank. However, this trend is not observed in the case of the United States. The mean world rank for the United States has actually decreased over the years, despite a significant increase in investment in education. The reason for this phenomenon is not entirely clear, but it may be due to the increasing number of universities being established in the United States, many of which may be starting at lower ranks. Additionally, the United States has the highest number of universities in the dataset, which could also be affecting the mean world rank. On the other hand, a clear trend is observed in the case of Singapore, where an increase in investment in education is positively correlated with an improvement in mean world rank. Further investigation is required to understand the underlying factors that contribute to this phenomenon.