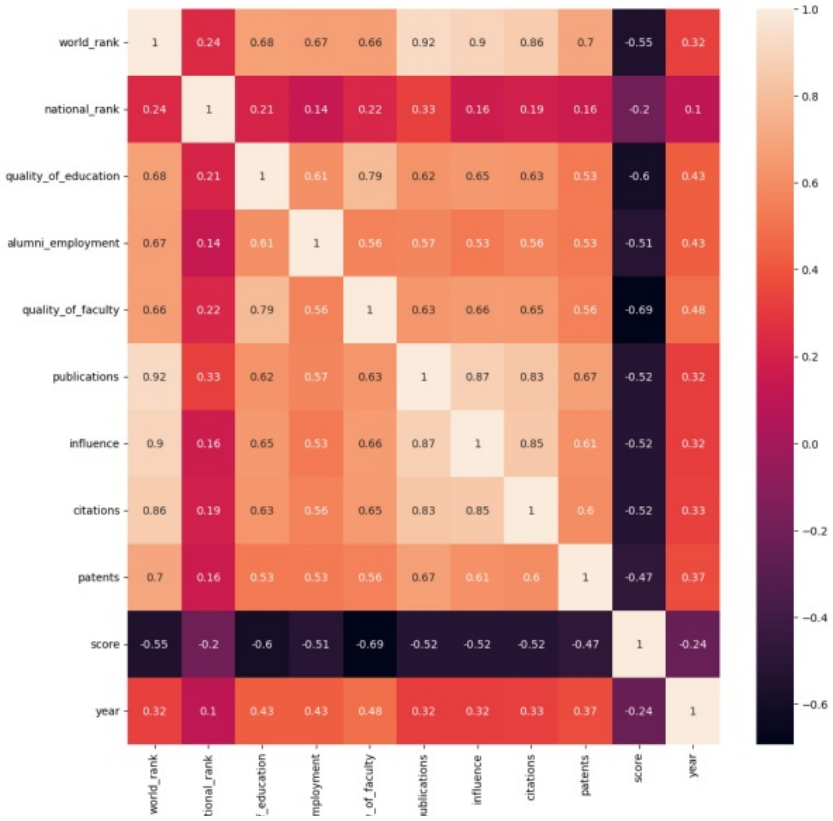
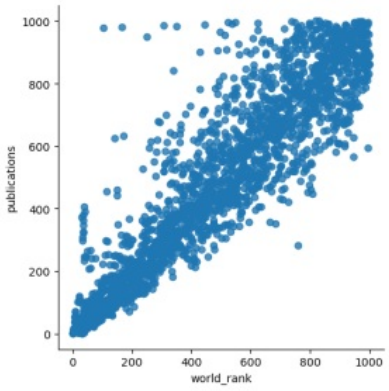
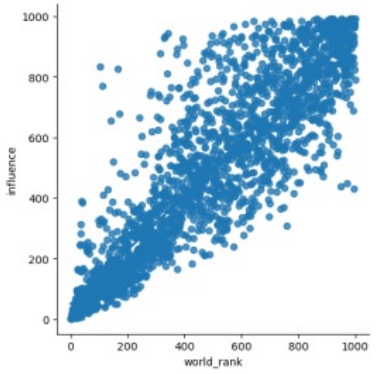
University Ranking

Based on the Center for World University Rankings







For data exploration, we first wanted to look for linear relationships.

We started looking at the correlation between the factors that determine the ranking of the universities via a heat map. Based on the results, we decided that there is a strong relationship between publications and influence with world rank.

The plots show a strong positive relationship between publications and world rank and influence and world rank. As world rank increases, so do publications. The same can be said for the relationship between influence and world rank.

The upward trend led us to the following hypothesis:

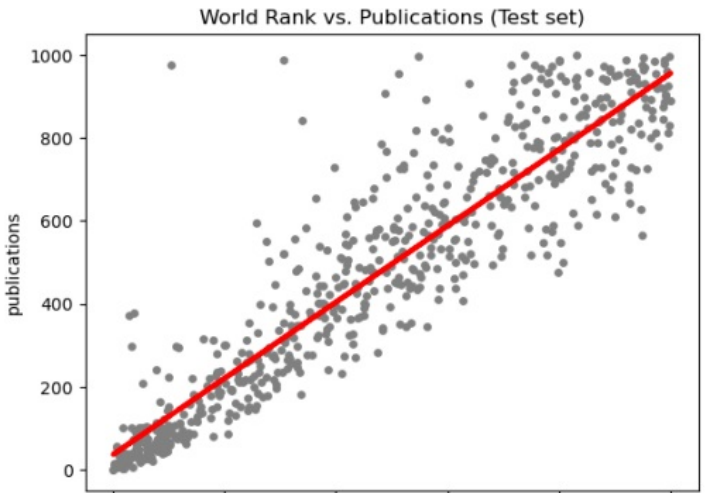
The higher your publication and influence, the higher your world rank

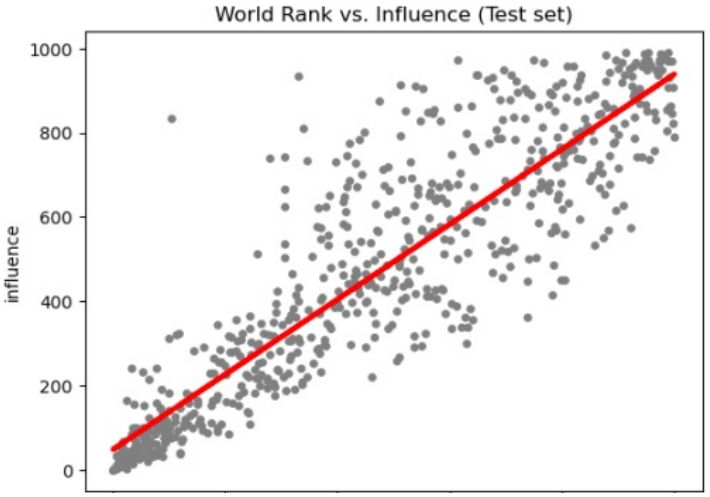
The higher your publication and influence, the higher your world rank score.

To test this hypothesis, we conducted a linear regression.

The results showed that **85%** of the trend line for the world rank vs. publications and **80%** for the world rank vs. influence. The relationship between the two sets of variables is almost entirely linear and both the publications and influence are a good determinant of the world rank score of each university.

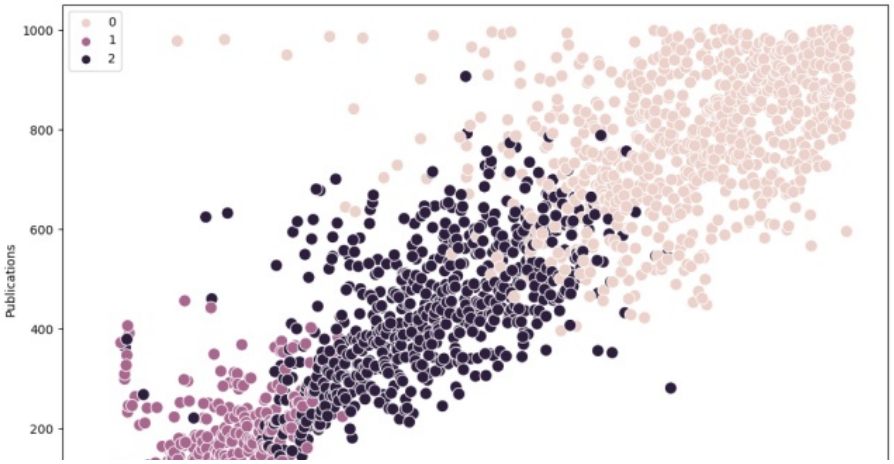
However, some points fall beyond the regression line, and there is a high density of data points when the world rank versus publications or influence is low. As a linear regression isn't enough to fully explain the data, we must try another approach.

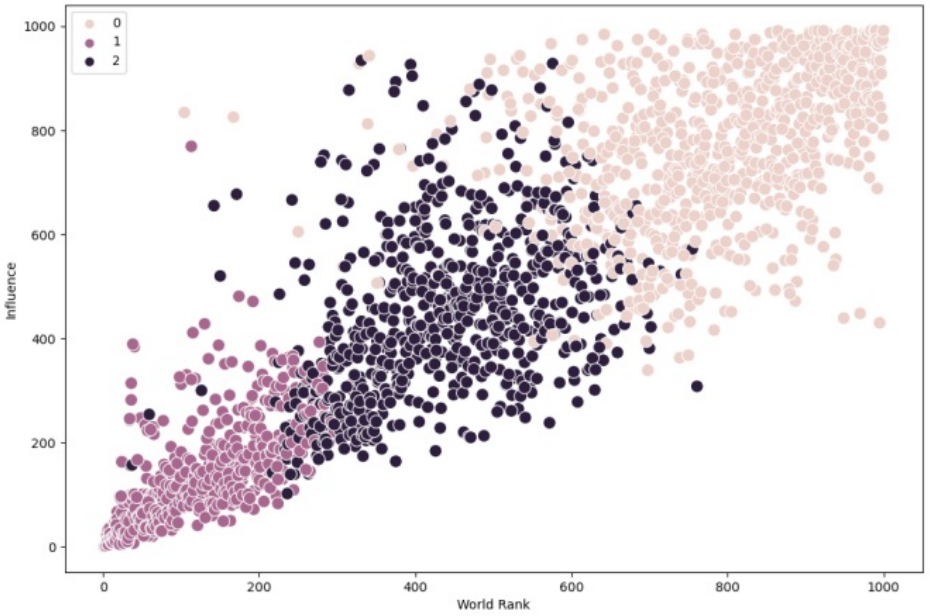




Because a linear regression wasn't enough o prove our hypothesis, we needed a non-linear approach. So we conducted a cluster analysis.

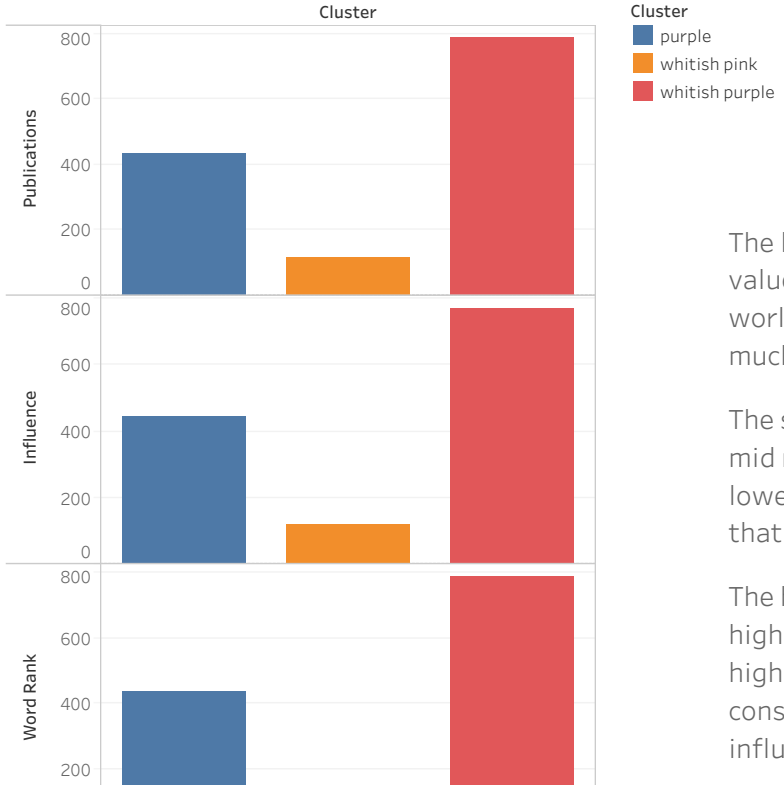
A cluster analysis groups data points into "clusters." We can then compare the groups to uncover new patterns and trends.





Our cluster analysis yielded three district groups of data points, which are represented in different colors on this scatterplot - whitish pink, light purple, and purple.

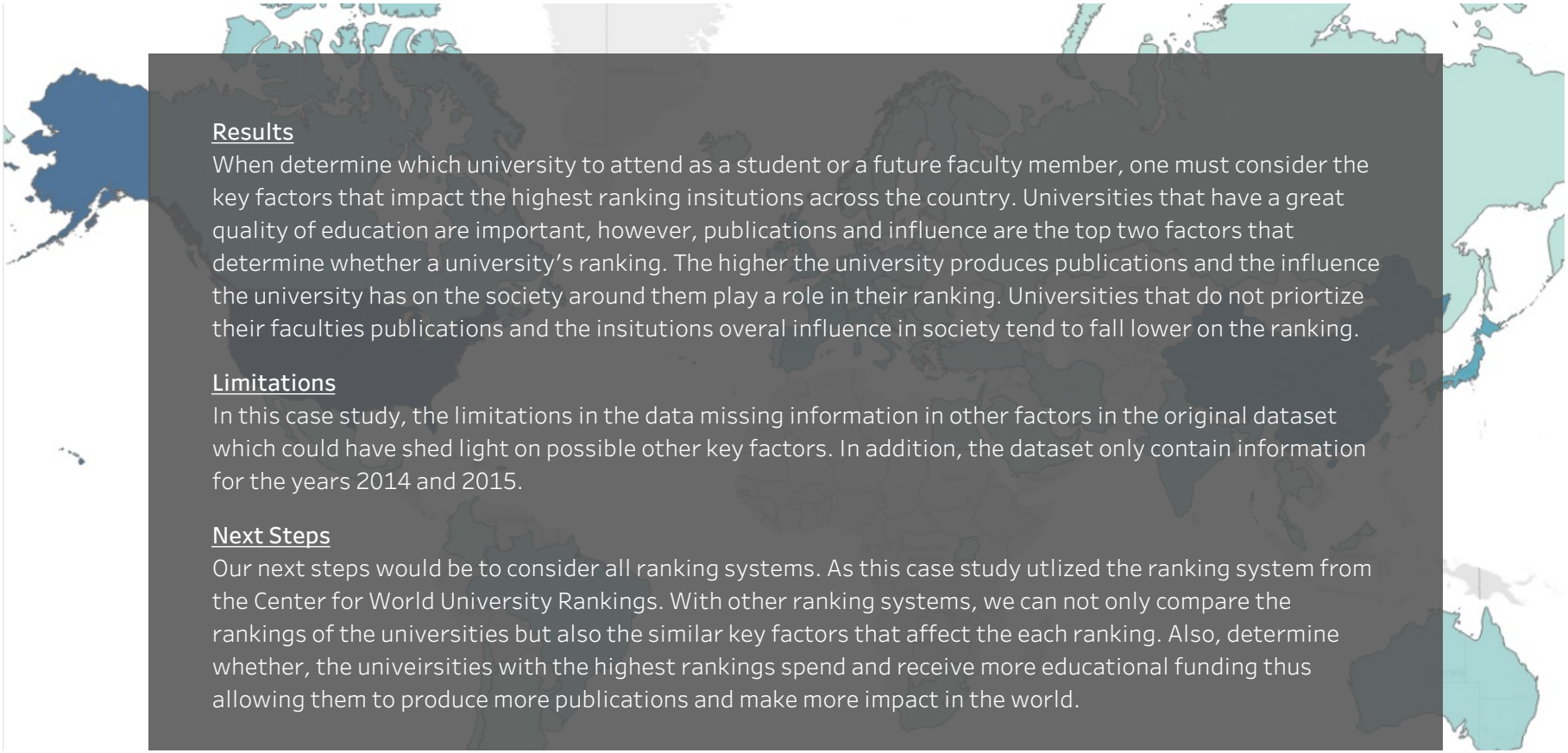
Difference between Clusters



The lowest world ranked cluster (**whitish pink**) is the cluster with least value of publications and influence. This could result in a low average world ranking. We can assume these are universities that do you have much publucations or influence in society.

The seond lowest world ranked cluster (**purple**) is the cluster with the mid range of publications and influence. The average is higher than the lowest world ranked cluster. We can assume that these are universities that have some publications and influence impact in the society.

The highest world ranked cluster (**whitish purple**) is the cluster with the highest amount of publications and influence. The average is much higher than the first two clusters. We can assume these universities are consistently produce publications and have a steady but impactly influence in society.



Results

When determine which university to attend as a student or a future faculty member, one must consider the key factors that impact the highest ranking insitutions across the country. Universities that have a great quality of education are important, however, publications and influence are the top two factors that determine whether a university's ranking. The higher the university produces publications and the influence the university has on the society around them play a role in their ranking. Universities that do not priortize their faculties publications and the insitutions overal influence in society tend to fall lower on the ranking.

Limitations

In this case study, the limitations in the data missing information in other factors in the original dataset which could have shed light on possible other key factors. In addition, the dataset only contain information for the years 2014 and 2015.

Next Steps

Our next steps would be to consider all ranking systems. As this case study utlized the ranking system from the Center for World University Rankings. With other ranking systems, we can not only compare the rankings of the universities but also the similar key factors that affect the each ranking. Also, determine whether, the univeirsities with the highest rankings spend and receive more educational funding thus allowing them to produce more publications and make more impact in the world.