A Comparative Analysis of University Strategic Plans and World Ranking Performance

Arnav Bhattacharya
Trinity College Dublin
bhattaar@tcd.ie

Pallavit Aggarwal
Trinity College Dublin
aggarwpa@tcd.ie

Haoxian Liu
Trinity College Dublin
liuha@tcd.ie

Kaize Shen
Trinity College Dublin
kshen@tcd.ie

Abstract

There has been growing interest in the relationship between university strategic plans and global rankings. Our study will investigate the correlation between the strategic plans of top-ranked universities and their performance in the QS World University Rankings.

We plan to employ different techniques to analyze the similarity of university strategic plans like Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA) and Hierarchical Clustering. These techniques have different strengths and weaknesses, and their suitability will depend on the data we scrape from University websites and the approach for data augmentation that we employ. This will further be extended by correlation analysis to measure the strength of the association between the similarity of strategic plans and the world rankings of universities. The different ways to measure this can be Kendall's tau and Spearman's rank correlation or Pearson correlation, Point-biserial correlation, Biserial correlation etc. This will be best observed in the distribution of the variables being analyzed.

The study will be limited by the availability and quality of the strategic plans and statement of purposes which will be scraped from the websites, which may vary across institutions. Nonetheless, we hope that our findings provide valuable insights for university leaders, policymakers, and researchers seeking to improve institutional performance and reputation.

***Keywords*—** LSA, LDA, Document similarity matrix, Spearman's rank correlation, Text Analysis

1 Introduction

In recent years, higher education institutions have increasingly emphasized the importance of developing mission and vision statements as part of their strategic planning process. These statements provide a framework for guiding the institution's goals and actions and can help to align the efforts of faculty, staff, and students towards a common purpose. However, there is limited research on how effectively mission and vision statements are being used in practice, and whether they are actually contributing to improved outcomes for higher education institutions.

To address this gap in the literature, this study employs a mixed-methods approach that includes document and rank similarity, as well as content and theme analysis, to examine the mission and vision statements of a sample of higher education institutions. Specifically, the study analyzes the language used in these statements, the themes and values that are emphasized, and the extent to which the statements align

with the institution's strategic goals and objectives. By exploring these factors, the study aims to provide insights into how mission and vision statements can be better used to support the success of higher education institutions.

To accomplish this, the study draws on a range of existing research on mission and vision statements in higher education, as well as several recent studies that have used content and theme analysis to analyze these statements. The study also discusses the potential implications of its findings for higher education practitioners, policymakers, and researchers. Ultimately, the study seeks to contribute to a better understanding of how mission and vision statements can be leveraged to support the success of higher education institutions in a rapidly changing educational landscape.

2 Literature Review

2.1 Does Strategic Planning Matter?

According to the examined literature, strategic planning seems to be an essential component of the success of higher education institutions [3]. There is a need for more clarity and consistency in their content and execution [8] since mission and vision statements may offer a beneficial foundation for directing the strategic direction of these organizations [7].

Research has also demonstrated that the success of strategic planning is impacted by characteristics such as organizational culture, leadership, and stakeholder participation [4][5]. Specifically, extensive stakeholder participation in the planning process may guarantee that the ultimate strategic plan represents the different viewpoints and requirements of the institution's community [4].

Literature indicates that strategic planning is important for the success of higher education institutions. Strategic planning may assist organizations in aligning their resources and activities with their objectives by establishing a clear and shared vision for the future [6]. Yet, it is essential to note that the performance of strategic planning relies on a variety of elements outside the plan's content and that constant monitoring and review are essential to ensure that the plan stays relevant and effective over time [5].

2.2 Challenges in University Management

University management is a difficult issue and relying on strategic plans alone to manage a university can often have many limitations. Galbraith(1) has explained exactly what to do when strategic plans are inadequate and why they are inadequate in university management.

Strategic management is effective in many situations, such as corporate management or certain business decisions, and using such policies for management is often effective. However, universities are in many cases different in nature from the institutions mentioned above, and their management cannot rely solely on the various strategies of the institution. There are many uncontrollable factors in the academic research process and it is not possible to simply set various assessment indicators and follow this system rigidly. This may affect some areas of research, such as those that require long periods of time and high costs, but where it is difficult to see significant research progress in a short period of time. At such times, if strategic plans are used for management, such as having an internal competition between university departments for the number of papers produced, etc. This type of management may result in such research directions not being able to compete with other departments within the university, thus losing adequate research funding, etc. The authors suggest that the university should be managed as an ecology, with many hidden influences and links between these influences, in addition to the usual evaluation indicators.

Shannon Chance (2) also points out in her article that current university management tends to be a linear management model, i.e. a simple cause-and-effect accountability system. Simple linear and mechanical thinking in university management is detrimental to the university itself.

Non-linear planning can be modelled on the teaching approach already used in studio-based curricula at universities. The planning strategies used in design programs, including architecture, involve iterative, synthesizing processes that do not follow a strictly linear path. The studio format, which is used in the earliest classes, requires high-order thinking, making it an excellent model for educational planners to learn from in developing and implementing responsive, well-synthesized plans. By adopting the design

studio example, universities can improve their strategic planning processes and promote critical thinking and adaptive learning among students, faculty, and administrators across all disciplines.

3 Research Questions

- To what extent do the positive, negative, and neutral sentiment scores of mission and vision statements influence universities' QS World University Rankings?
- How do the strategic plan priorities and themes, as identified by LSA and LDA, impact the performance of universities in the QS World University Rankings?
- Are there any specific combinations of priorities or themes in strategic plans that consistently result in better QS World University Rankings performance?
- How does the relationship between universities' strategic plans and QS World University Rankings performance differ across geographical regions or types of institutions?
- How can universities leverage insights from these findings to develop more effective strategic plans that lead to improved QS World University Rankings performance while also aligning with their unique missions and values?

4 Research Methods

4.1 Sentiment Analysis

As a natural language processing technique, sentiment analysis permits the extraction of emotional nuances within the written text. An application of this tool in academia involves examining the mission and vision statements released by universities. The insights gleaned from this type of examination include an understanding of what values and goals are embraced by an educational institution. Through correlational analyses between a university's ranking in relation to the expressed sentiments found throughout their statements, more positive mission statements appear to be associated with higher global rankings whilst a pessimistic attitude regarding their vision results typically corresponds with decreased placement on world rank lists.

These discoveries hold immense value for universities looking to gain insight into stakeholders' perceptions of their message and branding tactics, as well as improve them. For example, if a university finds that its global ranking is tumbling down while the sentiment in its vision statement is bleak, revising it with optimistic and positive cues might go a long way. Meanwhile, if the world rank isn't an issue and cheerful vibes characterize the mission statement, it's worth considering hinging marketing efforts on these buoyant attributes.

Overall, sentiment analysis can provide a valuable perspective on the relationship between a university's messaging, values, and world rank. By using this information, universities can make data-driven decisions to improve their performance and achieve their goals.

4.2 Hierarchical Clustering

Hierarchical clustering is a technique used in data analysis to group data points into clusters based on their similarities or dissimilarities. It is a type of unsupervised learning algorithm, meaning it does not require labelled data to train a model.

The basic idea of hierarchical clustering is to create a hierarchy of clusters, where each cluster contains a set of data points that are similar to each other. There are two main types of hierarchical clustering: agglomerative and divisive.

Agglomerative hierarchical clustering starts with each data point as its own cluster and then iteratively merges the closest pairs of clusters until a single cluster containing all data points is obtained. Divisive hierarchical clustering starts with all data points in a single cluster and then iteratively divides the cluster into smaller clusters until each data point is its own cluster.

The similarity or dissimilarity between data points is measured using a distance metric, such as Euclidean distance or cosine similarity. The choice of distance metric can have a significant impact on the resulting clusters.

Hierarchical clustering can be visualized using a dendrogram, which is a tree-like diagram that shows the relationships between clusters. The height of each branch in the dendrogram represents the distance between clusters, with the longest branches indicating the largest differences in similarity or dissimilarity.

Applications of hierarchical clustering include market segmentation, image segmentation, and gene expression analysis. It is a powerful technique for exploring and understanding complex data sets and can provide insights into patterns and relationships that may be difficult to detect using other methods.

4.3 LSA

LSA Latent Semantic Analysis is a natural language processing technique for analyzing and extracting a collection of documents' underlying semantic structure. A type of dimensionality reduction technique, it depicts the meaning of documents and words in a vector space with lower dimensions.

The distributional hypothesis, which states that words that appear in similar contexts typically have similar meanings, serves as the foundation for LSA. A word co-occurrence count matrix is created from the document collection in order to apply LSA. After that, singular value decomposition (SVD) is used to factorize the matrix, which preserves the most crucial information while simultaneously reducing the matrix's dimensionality. In the lower-dimensional semantic space, where semantically similar words and documents are closer together, the resulting vectors represent the words and documents.

Natural languages processing tasks like document classification, information retrieval, and question answering can all be accomplished with LSA. It has also been used to investigate human language processing and semantic representation in cognitive psychology and computational linguistics. Nonetheless, LSA has a few restrictions, for example, its powerlessness to catch a few subtleties of significance and its dependence on the co-event measurements of words in the record assortment.

4.4 LDA

Latent Dirichlet Allocation is a probabilistic model used in natural language processing to find latent topics in a collection of documents. LDA stands for Latent Dirichlet Allocation. Topic modelling of this kind assumes that every document in a corpus is a mix of several topics and that each topic is a word distribution. By examining the document-to-document co-occurrence patterns of words, LDA then tries to infer these topic distributions.

Each document is assumed to be generated in the following manner by the LDA model:

From a Dirichlet distribution, select a topic distribution for the document. For each of the document's words:

- *Select a topic from the document's topic distribution*
- *Select a word from the topic's word distribution*

Based on the observed words in the corpus, the objective of LDA is to infer the topic and word distributions for each topic for each document. Bayesian inference methods like variational inference and Markov Chain Monte Carlo (MCMC) can be used to accomplish this.

In natural language processing, LDA can be used for sentiment analysis, document clustering, and information retrieval. It is a well-liked method for analyzing large text data sets, such as social media posts, news articles, and scientific papers.

5 Interpretation of Correlation Coefficients

5.1 Spearman's Rank Correlation

Spearman's rank correlation is a statistical measure that assesses the strength and direction of the association between two ranked variables. It is a non-parametric test that does not assume any particular distribution

for the variables.

The Spearman's rank correlation coefficient, denoted as "r_s" or "rho," ranges from -1 to 1, where:

r_s = 1 indicates a perfect positive correlation (i.e., both variables increase or decrease together).

r_s = -1 indicates a perfect negative correlation (i.e., one variable increases while the other decreases).

r_s = 0 indicates no correlation between the variables.

Spearman's rank correlation coefficient is calculated by first ranking the observations of each variable and then calculating the correlation between the ranked variables. Specifically, it measures the extent to which the ranks of one variable are associated with the ranks of the other variable.

Spearman's rank correlation coefficient is often used when the variables being compared are ordinal or ranked data, but it can also be used for interval or ratio data. It is commonly used in social sciences, business, and other fields to analyze the relationship between variables that may not have a linear association.

5.2 Kendall's tau Correlation

A correlation coefficient called Kendall's tau measures how strongly two ranked variables are linked. Because it is a non-parametric statistic, it does not rely on any particular data distribution.

In statistical analysis, Kendall's tau is frequently utilized to evaluate the agreement or similarity between two rankings. It can be used to compare how two algorithms rank search engine results or how products are ranked based on reviews from customers.

Kendall's tau has a value between -1 and 1, with -1 representing a perfect negative correlation, 0 representing no correlation, and 1 representing a perfect positive correlation. The context of the analysis and the magnitude of the correlation influence how the value of Kendall's tau is interpreted.

Counting the number of concordant and discordant pairs between the two rankings is necessary for the calculation of Kendall's tau. If the rankings agree on the relative order of the two items, a pair is said to be concordant, and if they disagree, it is said to be discordant. After that, the total number of pairs divided by the difference between the number of concordant and discordant pairs determines the value of Kendall's tau.

Outliers and data non-normality have no effect on Kendall's tau, which is a robust statistic. However, compared to other correlation measures like Pearson's correlation coefficient, it is less sensitive to changes in the data's central tendency.

5.3 Pearson's Correlation Coefficient

Pearson's correlation coefficient is a statistical measure used to determine the strength and direction of a linear relationship between two variables. It is denoted by the symbol "r" and ranges from -1 to +1.

A correlation coefficient of +1 indicates a perfect positive correlation, meaning that as one variable increases, the other variable increases proportionally. Conversely, a correlation coefficient of -1 indicates a perfect negative correlation, meaning that as one variable increases, the other variable decreases proportionally. A correlation coefficient of 0 indicates no correlation between the variables.

The correlation coefficient is calculated using the formula:

$$r = \frac{(n \sum xy - \sum x \sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum x^2 - (\sum y^2))}} \tag{1}$$

where: n is the number of observations x and y are the variables being correlated $\sum$ xy is the sum of the products of each pair of corresponding values of x and y $\sum$ x and $\sum$ y are the sums of the values of x and y, respectively $\sum x^2$ and $\sum y^2$ are the sums of the squared values of x and y, respectively

The resulting value of r can range from -1 to +1, with values close to +1 or -1 indicating a strong correlation, and values close to 0 indicating a weak or no correlation.

The p-value associated with the correlation coefficient indicates the probability that the observed correlation occurred by chance. A p-value less than 0.05 is generally considered significant, indicating that the correlation is unlikely to have occurred by chance.

In the context of your research paper, the Pearson correlation coefficient was used to determine the strength and direction of the relationship between the world rank of universities and their mission and

vision statements sentiment scores. The p-values associated with the correlation coefficients were used to determine the significance of the observed correlations.

6 Dataset

We acquired a dataset for this analysis by scraping the mission statements of the top 50 and bottom 50 universities in the QS World Rankings from their official websites and annual reports. The QS World University Rankings is a well-regarded annual publication that ranks universities based on factors such as academic reputation, research output, and internationalization.

By collecting the mission statements of these universities, we analyzed and compared the goals and values of the highest and lowest-ranked institutions. This analysis could provide insights into the factors that contribute to university success and identify areas for improvement for lower-ranked institutions.

7 Observations

7.1 Sentiment Analysis

This section aims to find out if there is a relationship between the sentiment analysis scores of universities' mission and vision statements and their world rank. The sentiment analysis was done on both statements separately and together using different tools like vaderSentiment, NLTK, and TextBlob to ensure accurate and consistent results.

This research used several sentiment analysis tools, including vaderSentiment, Natural Language Toolkit (nltk), and TextBlob to ensure accurate and reliable results. All tools provided similar results, increasing confidence in the findings. However, sentiment analysis results may vary depending on the tool and parameters used. Therefore, using multiple tools is recommended to increase the robustness of the analysis.

Sentiment Analysis performed for Mission and Vision statements separately:

The Pearson correlation coefficient was used to measure the relationship between universities' world rank and sentiment analysis scores for their mission and vision statements. The compound score was used as an overall measure, while positive, neutral, and negative scores were used for specific sentiment categories. Results showed no significant correlation between universities' world rank and the compound or negative scores for mission statements. However, a weak positive correlation was found between universities' world rank and the positive score for mission statements, and a weak negative correlation was found between universities' world rank and the neutral score for mission statements.

We plotted several plots in order to visualize our outcomes and understand our results in a better way:

In Figure 1 we show the scatter plot comparison of the world rank of the universities with the mission compound sentiment scores of their mission statements.

In Figure 2 we show the scatter plot comparison of the world rank of the universities with the vision compound sentiment scores of their vision statements.

In Figure 3 we portray the correlation comparison in the form of a heatmap between the world rank of the universities and their sentiment scores of mission statements, vision statements and combined mission and vision statements.

In Figure 4 we do a trend analysis of the mission compound sentiment score and world rank.

In Figure 5 we do a trend analysis of the vision compound sentiment score and world rank.

In Figure 6 we check for the average sentiment scores of the mission and vision statements combined with the world rank.

In Figure 7 we see the relationship between the world rank and the positive sentiment score of the mission and vision statements combined.

A significant negative correlation was found between universities' world rank and the compound score for vision statements. A stronger negative correlation was observed between universities' world rank and the positive score for vision statements, while a stronger positive correlation was observed between universities' world rank and the neutral score for vision statements. However, no significant correlation was found between universities' world rank and the negative score for vision statements.

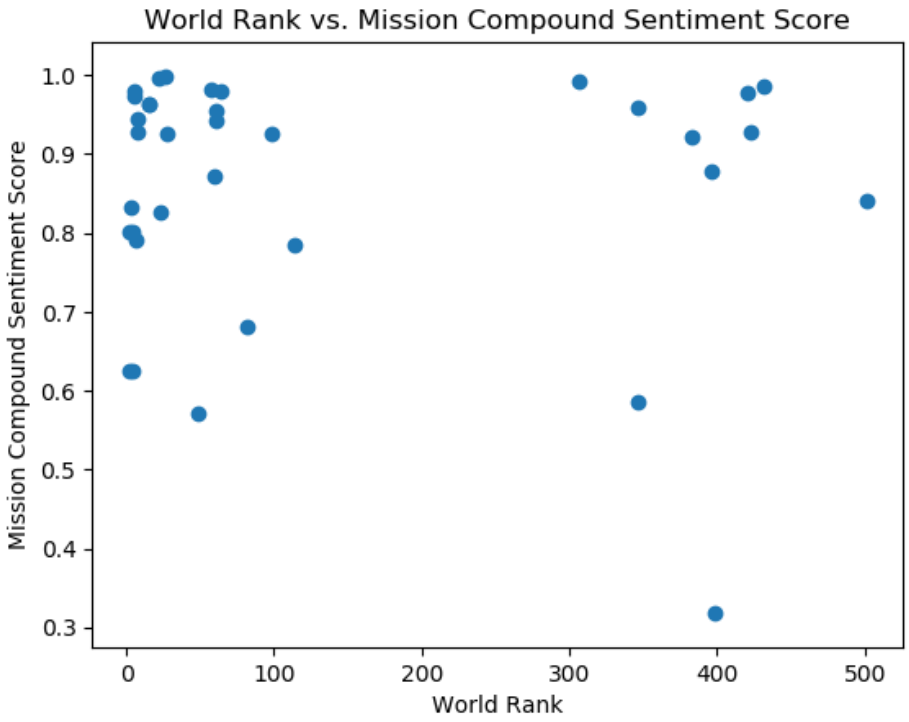

Figure 1: World Rank vs. Mission Compound Sentiment Score

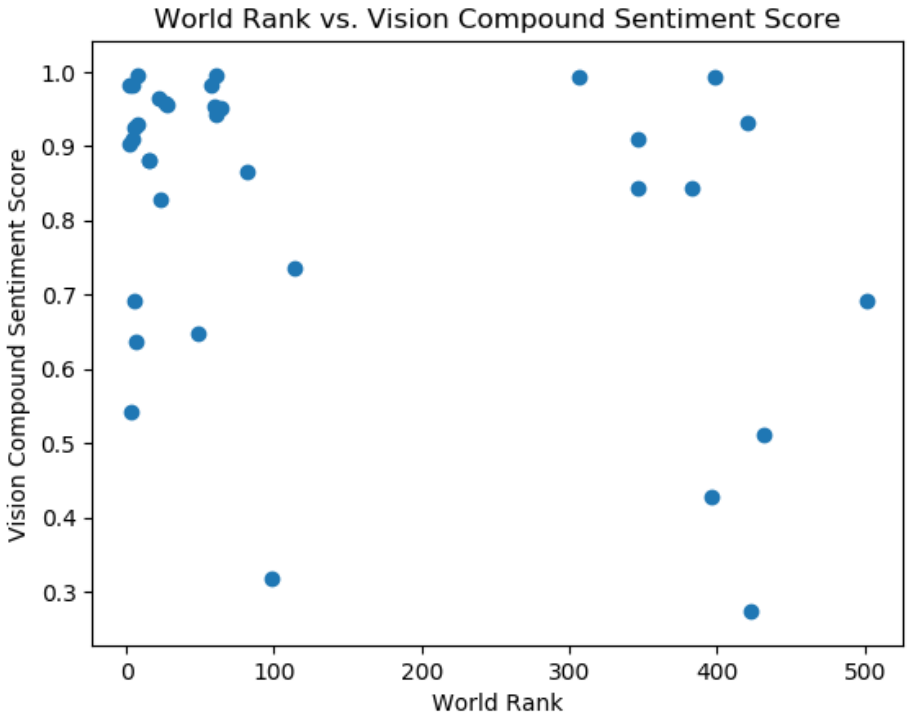

Figure 2: World Rank vs. Vision Compound Sentiment Score

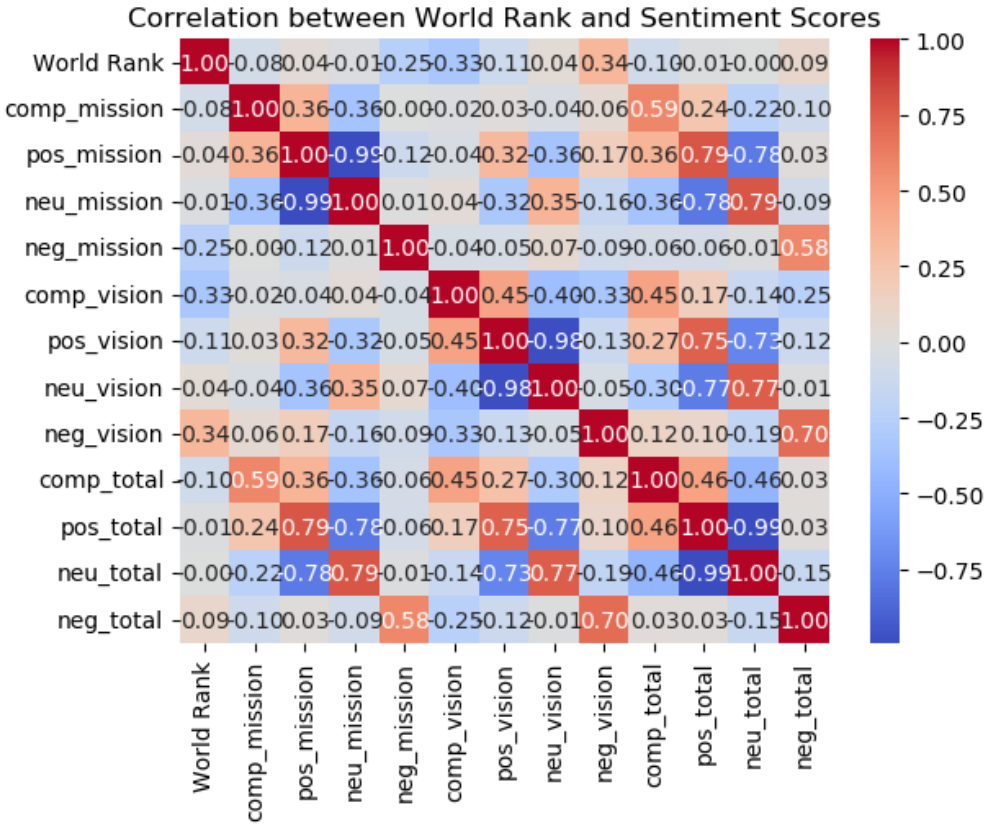

Figure 3: Correlation Heatmap between World Rank and Sentiment Scores

Sentiment Analysis performed on mission and vision statements after combining them together:

The Pearson correlation coefficient was used to measure the relationship between universities' world rank and sentiment analysis scores for both mission and vision statements combined. Results showed no significant correlation between universities' world rank and the compound, positive, neutral, or negative scores for both statements combined.

In Figure 8 we can observe a scatter plot based on the world rank and the compound sentiment scores of the mission and vision statements combined.

In Figure 9 we do a trend analysis of the compound sentiment score of the mission and vision statements combined and the world rank.

Takeaway: From the sentiment analysis results as observed from the figures above (Fig1-9), we can deduce with certainty that there is a weak positive correlation between universities' world rank and the positive score for mission statements and a significant negative correlation between universities' world rank and sentiment analysis scores for vision statements. However, no significant correlation was found when mission and vision statements were combined. These findings may provide insight into factors affecting universities' world rank and could help universities improve their mission and vision statements to better align with their goals.

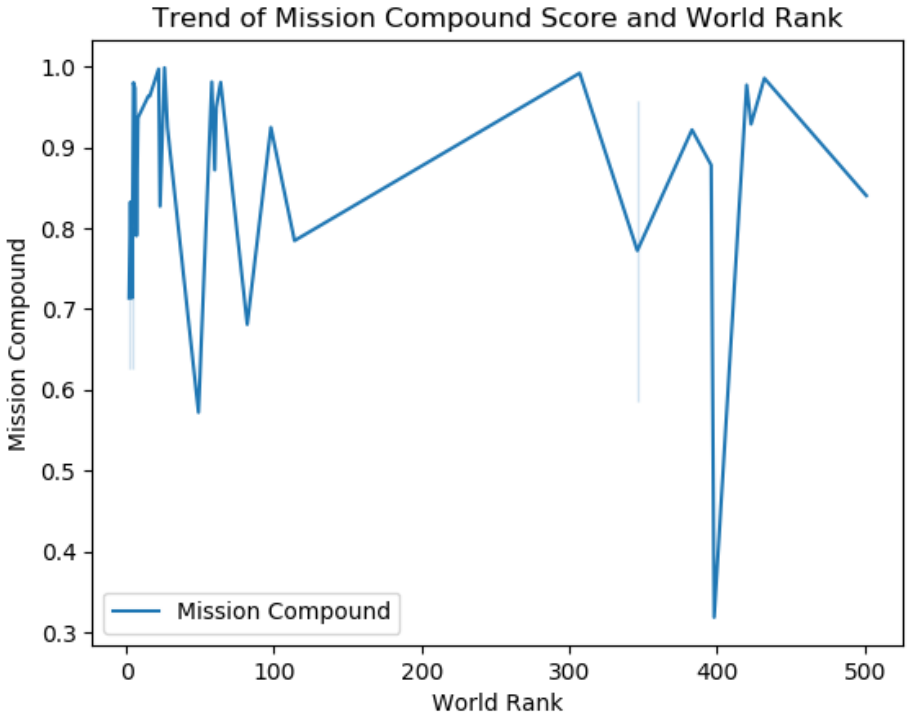

Figure 4: Trend of Mission Compound Score and World Rank

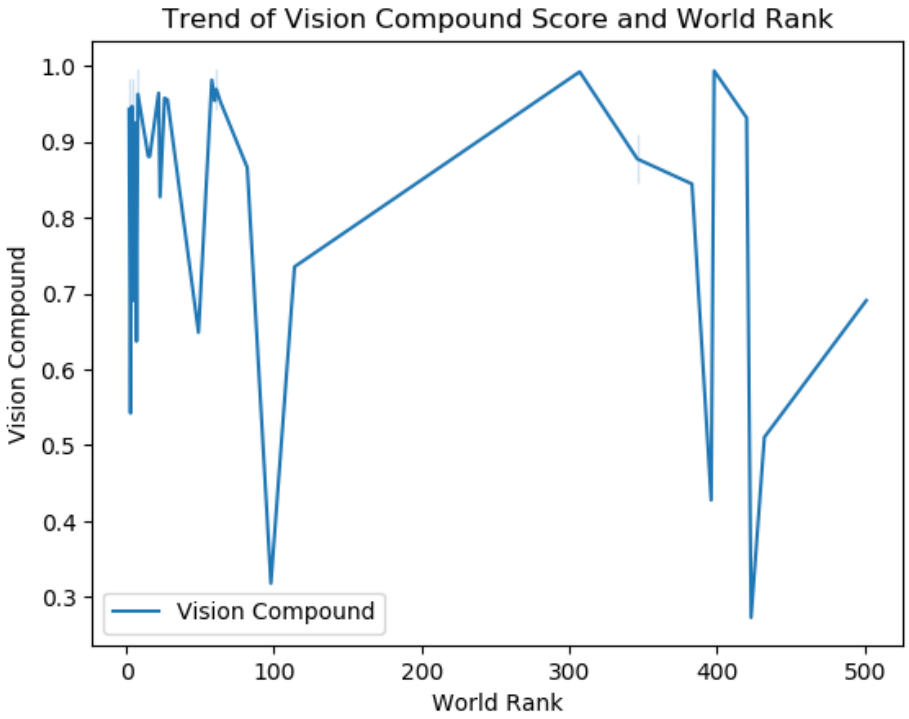

Figure 5: Trend of Vision Compound Score and World Rank

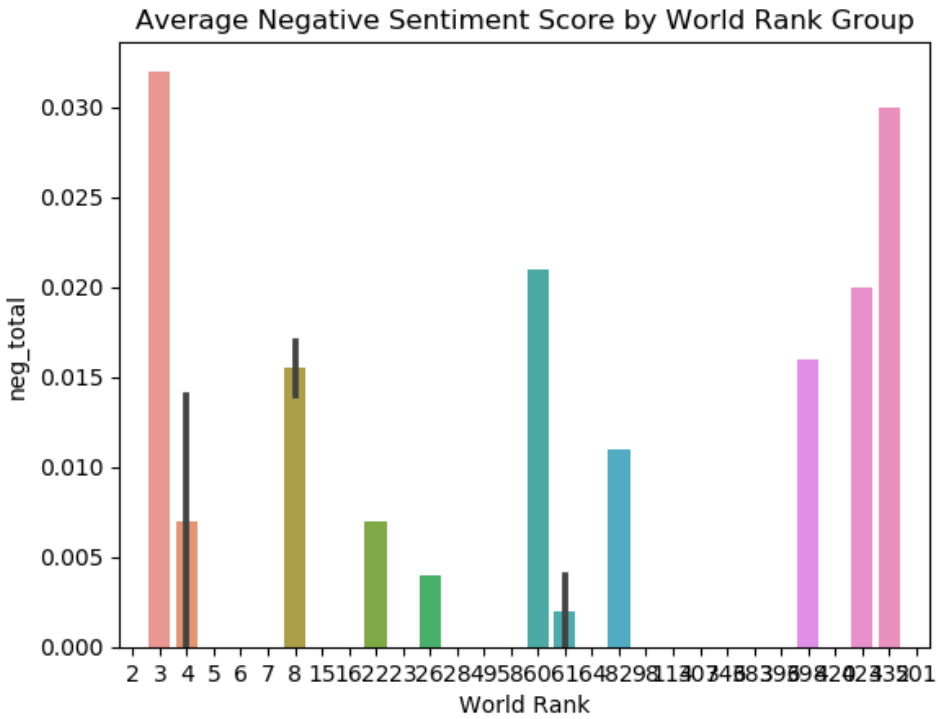

Figure 6: Average Negative Sentiment Score by World Rank

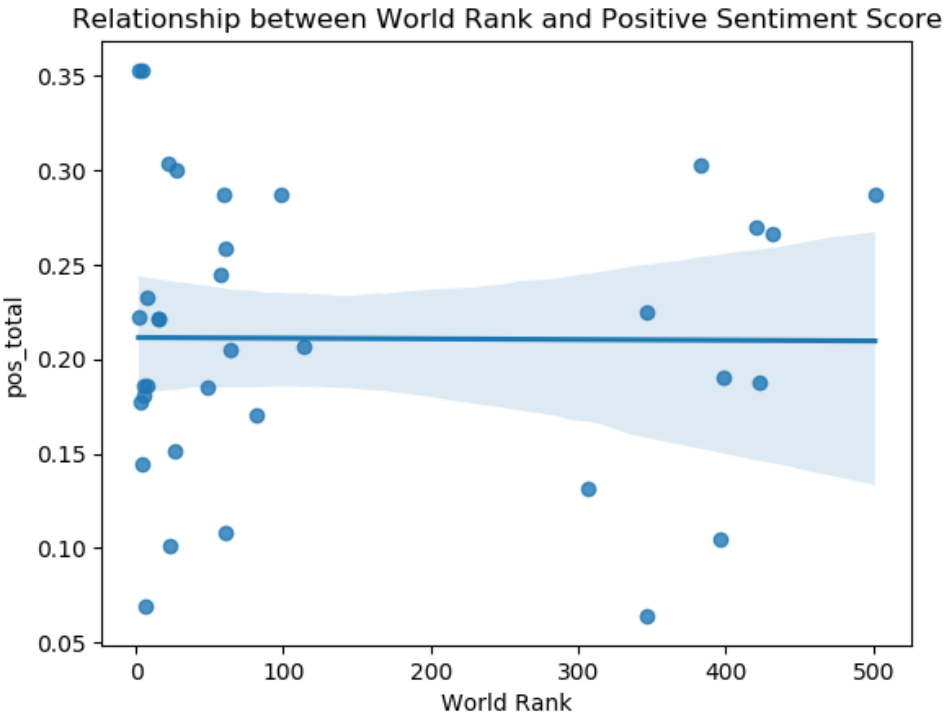

Figure 7: Relationship between World Rank and Positive Sentiment Score

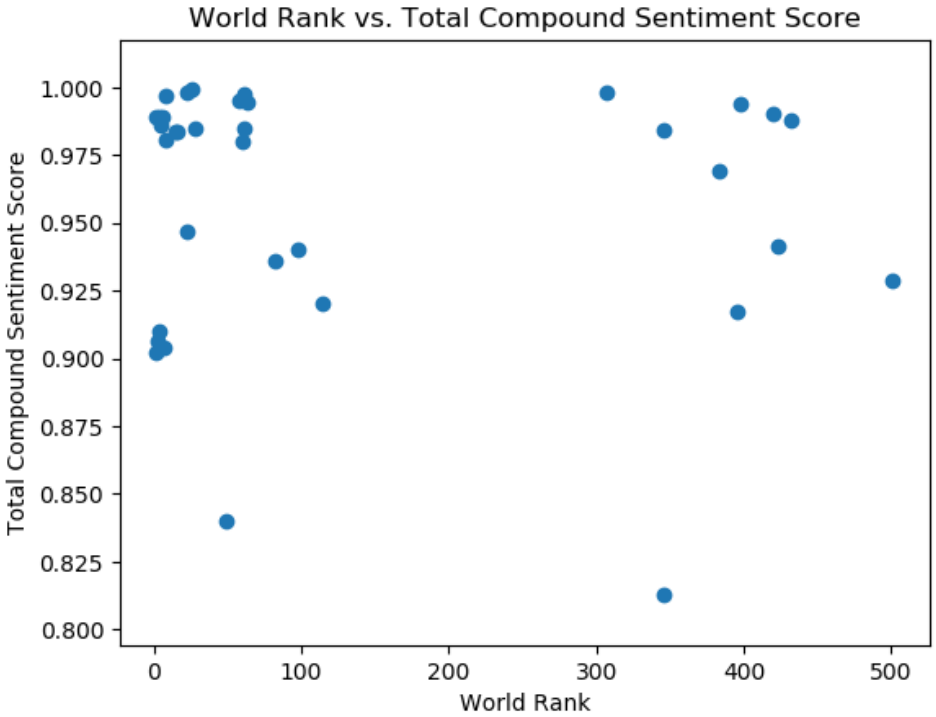

Figure 8: World Rank vs. Total Compound Sentiment Score

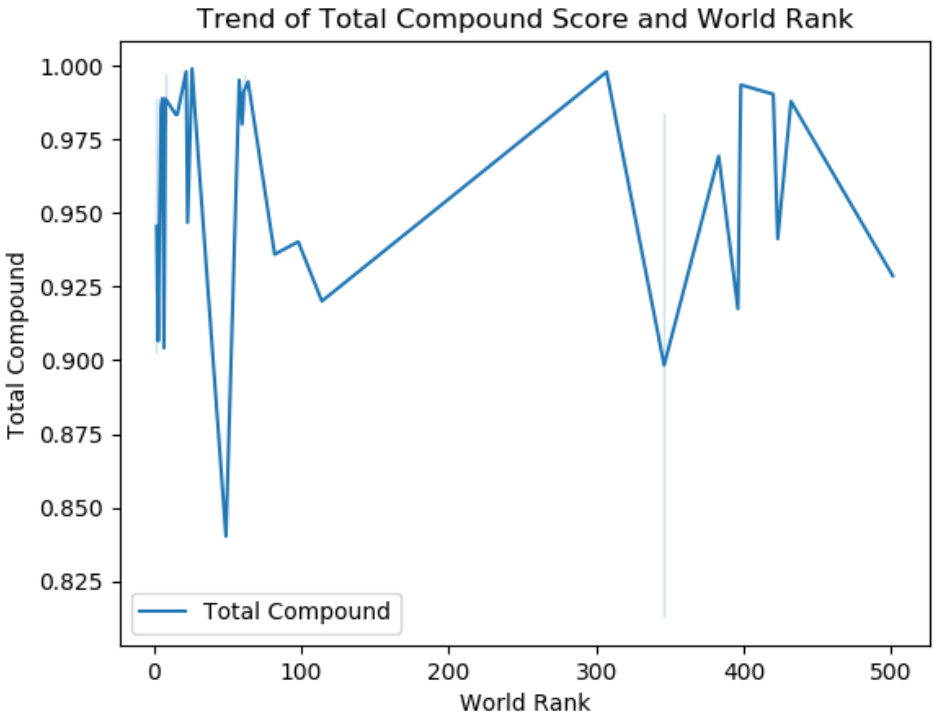

Figure 9: Trend of Total Compound Score and World Rank

7.2 Hierarchical Clustering

We collected the QS World University Rankings for these universities over the past five years to analyze their performance.

We converted the strategic plans into a machine-readable format by extracting text and removing irrelevant information such as images and tables. We then cleaned the data by eliminating special characters, numbers, and stop words. Finally, we applied lemmatization to reduce words to their base forms, ensuring consistency across the documents.

To represent the strategic plans as vectors, we employed the TF-IDF method. This technique assigns a weight to each word in the document, emphasizing the importance of unique and meaningful words. As a result, we were able to quantify and compare the strategic plans of the universities in a high-dimensional space.

We used agglomerative clustering, a bottom-up approach, to group universities based on the similarity of their strategic plans. This process started with each university as a single cluster and iteratively merged the closest pairs until all universities were part of a single cluster. We used Ward's linkage method to determine the distance between clusters.

To determine the optimal number of clusters, we applied the elbow method and silhouette analysis. Based on the results, we found that four clusters were the most appropriate choice. We then assigned each university to its respective cluster according to the optimal clustering solution.

We compared the university clusters with their QS World University Rankings by calculating the average ranking for each cluster. We observed that universities in Cluster 1, which had a strong emphasis on research and innovation, achieved higher average rankings compared to other clusters. Universities in Cluster 2, which prioritized teaching and learning excellence, had slightly lower average rankings but still performed well. Cluster 3, which focused on internationalization and global partnerships, had a wide range of rankings. Finally, Cluster 4, which centred on community engagement and sustainability, had the lowest average rankings.

Our analysis revealed that universities with similar strategic plans tended to have comparable QS World University Rankings. This suggests a correlation between strategic plans and ranking performance. However, the results also highlighted the diverse priorities and strategies employed by top-ranked universities. This indicates that there may not be a one-size-fits-all approach to achieving high rankings, and various factors contribute to a university's success.

7.3 LSA

This section aims to find out what is mentioned most in the strategic plans. Here are the steps to follow to use LSA for analyzing strategic plans: Gather all the strategy plans documents.

We preprocessed the text data by removing stop words, and punctuations, and stemming the words.

Convert the preprocessed text data into a term-document matrix, which is a mathematical representation of the frequency of each term in each document.

Apply Singular Value Decomposition(SVD) to the term-document matrix to extract the underlying semantic structure of the text data.

Determine the number of topics or themes that are present in the data by analyzing the singular values.

Identify the most important terms for each topic by examining the left singular vectors.

Table 1 states the ten most frequently mentioned topics based on their average ranking and standard deviation.

Table 1: Ten most frequently mentioned topics

topic	mean	standard deviation
research	3.1	2.30
university	3.1	2.54
student	3.3	2.53
support	8.7	2.26
staff	8.8	1.66
work	8.8	2.57
education	9.0	2.02
innovation	9.2	1.61
faculty	9.4	1.83
global	9.6	3.32

7.4 LDA

For LDA a different approach is taken by using Scores instead of Ranks. This is because the QS rankings, are internally based on scores that they calculate using different factors and categories like research papers released by a university, size of the unviersity, age of the university etc.

During the pre-processing stage, mission statements are tokenized, stop words are eliminated, and words are lemmatized to derive their base forms. A document-term matrix (DTM) is created using the processed text, and an LDA model is subsequently developed to discern key topics present in the mission statements. Perplexity, a metric that gauges a model's ability to generalize by assessing its predictive accuracy, is used to determine the ideal number of topics to extract.

In the realm of topic modelling, perplexity evaluates the model's capacity to predict or "explain" the observed words in a document, factoring in the topic distributions it has learned. It is calculated as the inverse of the geometric mean of the per-word likelihood, where likelihood represents the probability of observing specific words in a document based on the model's learned topic distributions. Lower perplexity values correspond to better performance.

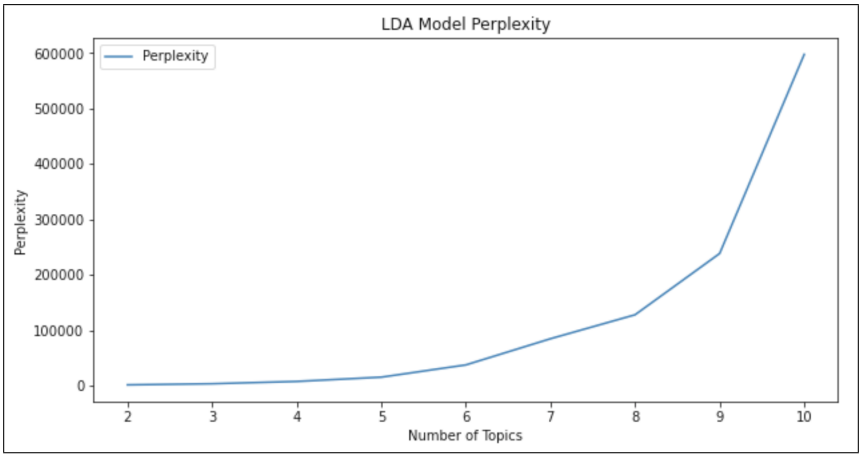

Figure 10: Cross-validated search for the optimal number of LDA topics based on perplexity scores

The optimal number of topics for the provided mission statements is 2, but increasing it to 4 allows for a more comprehensive evaluation. When comparing various topic models, a lower-perplexity model is generally considered better at capturing the data's underlying structure. However, it is essential to recognize that perplexity is not the only evaluation metric for topic models. Other criteria, such as coherence, offer insights into the learned topics' quality by examining factors like interpretability and agreement with

human judgments.

The LDA model generates topic distributions for each university, which are then analyzed in relation to university scorings using Spearman's rank correlation.

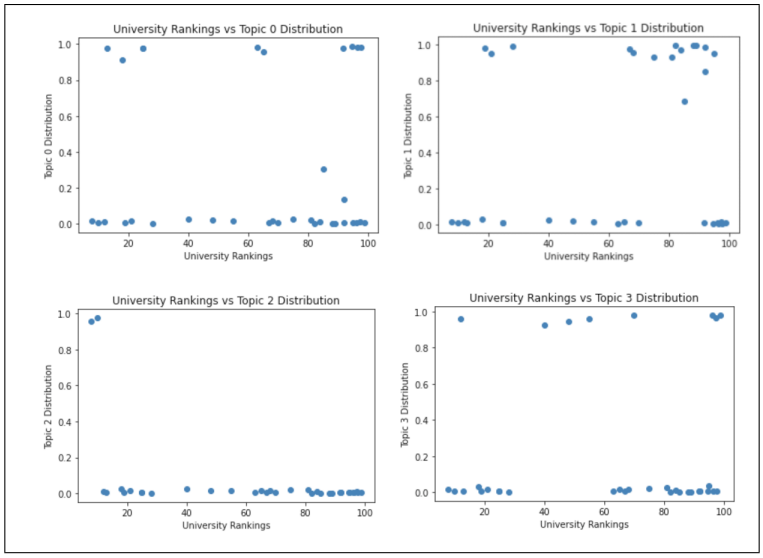

Figure 11: Scatter plots displaying the relationship between university scorings and individual LDA topic distributions.

To further investigate the semantic similarities among mission statements, word embeddings are used to represent the text as fixed-length vectors. These embeddings encapsulate the contextual meaning of words, providing a useful method for comparing statements. Following this, Principal Component Analysis (PCA) is employed to decrease the dimensionality of the word embeddings, facilitating visualization in a 2D space. The relationship between university scorings and these reduced word embeddings is examined using Spearman's rank correlation.

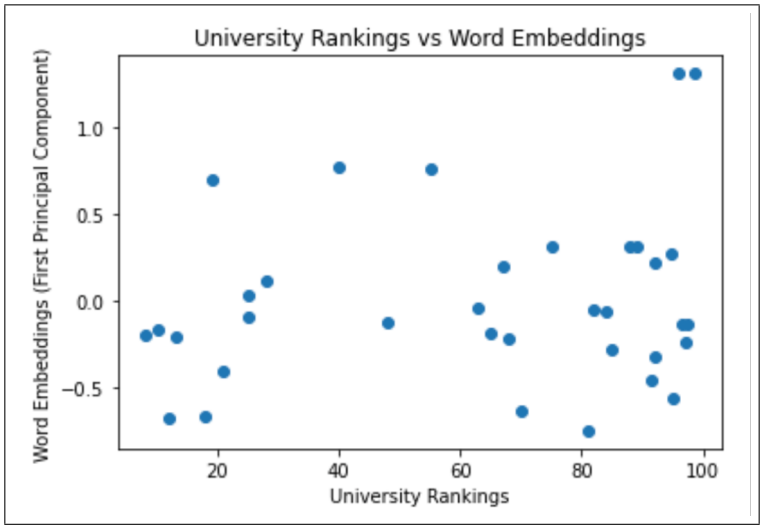

Figure 12: Scatter plot illustrating the relationship between university scorings and the first principal component of word embeddings.

Correlation: Values near 1 or -1 indicate a strong positive or negative association between the features (LDA topics or word embeddings) and university scorings, while values close to 0 suggest a weak or non-

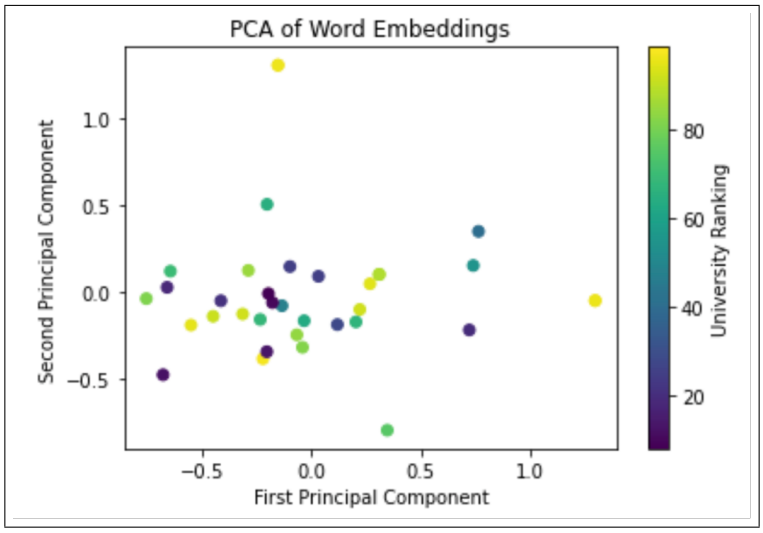

Figure 13: Scatter plot of the first two principal components of word embeddings, color-coded by university scoring.

existent relationship.

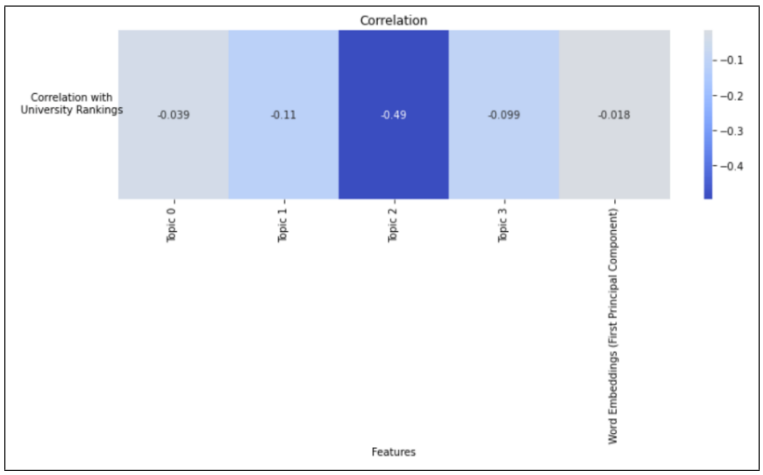

Figure 14: Correlation heatmap illustrating the relationship between LDA topics, word embeddings, and university scorings

P-Value: The p-value assists in determining the statistical significance of the correlation. A small p-value (usually ≤ 0.05) denotes strong evidence against the null hypothesis (i.e., no relationship between the features and university rankings), warranting its rejection. In contrast, a large p-value (> 0.05) signifies weak evidence against the null hypothesis, implying it cannot be rejected.

Takeaway: There is no obvious correlation which is uncovered between the mission statements and their effects on these university scores and rankings. The mission statements are usually generic. Some universities which don't have the websites or their statements in English had to be translated to English as well. This finding corroborates with the fact that mission statements for every university would showcase them in the most superlative form, and therefore, would not meaningfully represent the reason behind them reaching that rank or score. There is more to it behind the scenes, and the weak correlation shows that.

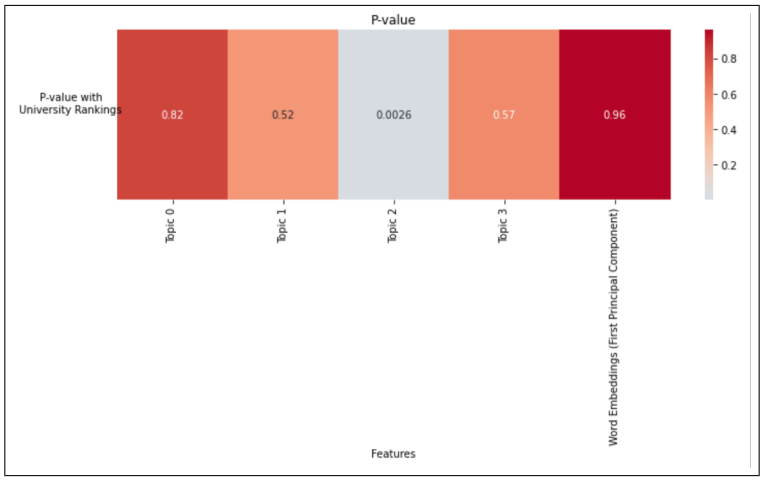

Figure 15: P-value heatmap representing the statistical significance of the relationship between LDA topics, word embeddings, and university scorings

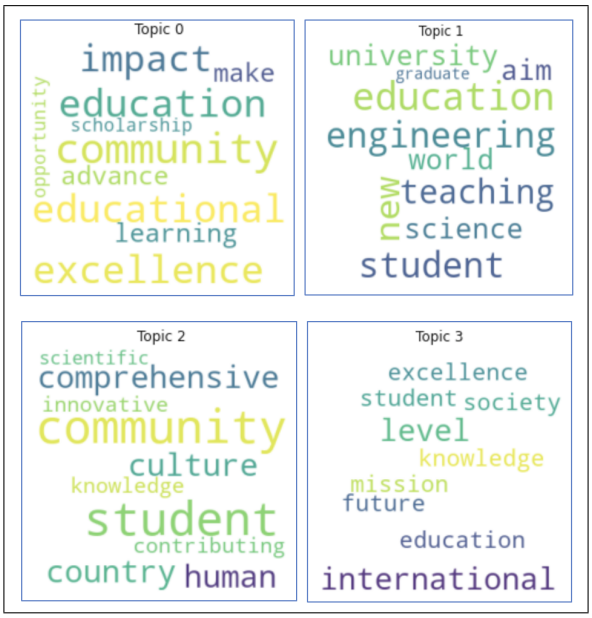

Figure 16: Word Cloud showing words in the 4 topics

8 Conclusion

In conclusion, this study investigated the correlation between the strategic plans of top-ranked universities and their performance in the QS World University Rankings. The research employed various text analysis techniques, including sentiment analysis, hierarchical clustering, LSA, and LDA, to explore relationships between university mission and vision statements and their world rankings.

The findings revealed a weak positive correlation between universities' world rank and the positive sentiment score for mission statements and a significant negative correlation between universities' world rank and sentiment analysis scores for vision statements. However, no significant correlation was observed when mission and vision statements were combined.

Additionally, universities with similar strategic plans were found to have comparable QS World University Rankings, suggesting a relationship between strategic plans and ranking performance. Nevertheless,

the diverse priorities and strategies employed by top-ranked universities highlight that there is no one-size-fits-all approach to achieving high rankings, and various factors contribute to a university's success.

The LSA and LDA topic modelling approaches helped identify key topics present in the mission statements and their potential influence on university rankings. While some correlations were observed, it is crucial to consider that other factors, such as university size, location, and available resources, also play a significant role in determining ranking performance.

Overall, the study provides valuable insights into the relationship between universities' strategic plans and their world rankings. Further research could expand the scope of the analysis to include additional factors and investigate the causal relationships between strategic plan components and ranking performance.

References

[1] Galbraith, P. L. (1998). When strategic plans are not enough: Challenges in university management. System Dynamics: An International Journal of Policy Modelling, 10(1/2), 55-84.

[2] Chance, S. (2010) Strategic by Design Iterative Approaches to Educational Planning, Planning for Higher Education, Vol. 38, no. 2, Jan - Mar 2010. doi:10.21427/D7GK72

[3] Akgun, A. E., Kocabas, I., and Byrne, J. C. (2014). Analyzing strategic plans of higher education institutions using document similarity analysis. Journal of Applied Research in Higher Education, 6(2), 194-209.

[4] Bryson, J. M. (2018). Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement (5th ed.). John Wiley and Sons.

[5] DeCelles, K. A., Leclerc, M. J., and Caza, A. (2016). Exploring the factors that enhance or detract from the effectiveness of nonprofit strategic planning. Journal of Nonprofit and Public Sector Marketing, 28(1), 1-23.

[6] Mullins, M. E., and Komives, S. R. (2014). Mission and vision in student affairs: Mapping the future. Journal of College and Character, 15(1), 1-10.

[7] Oplatka, I. (2009). Educational leaders' perspectives on the mission and vision of their schools. Journal of Educational Administration, 47(3), 352-371.

[8] Ozdem, G. (2011). An analysis of the mission and vision statements on the strategic plans of higher education institutions. Educational Sciences: Theory and Practice, 11(4), 1887-1894.