# FORMATTING INSTRUCTIONS FOR ICLR 2021 CONFERENCE SUBMISSIONS

**Anonymous authors**
Paper under double-blind review

## 1 BACKGROUND AND MOTIVATION

Since around 2004, world university rankings (WUR) have gained popularity among higher education institutions worldwide.

Although QS releases rankings with scores for various criteria , these scores are relative values. The top university in a given criterion would receive a full score of 100 points, while other universities are assigned relative scores based on their gap from the top university. These relative values are calculated using a hidden methodology and specific metrics. What we do know are the weights for the criteria scores, which are used to tally the overall scores.

According to the rules officially announced by QS, specific intrepretation of these criteria is shown in figure 1.

| Indicator | Description | Weight | Source |
|---|---|---|---|
| **Academic Reputation** | Peer review survey designed to evaluate the perceptions of global academics on which institutions are demonstrating academic excellence | 30% | QS Academic Reputation Survey |
| **Citations per Faculty** | Average number of citations obtained per faculty member (self-citations excluded and citations are normalized; citations in 5 subject areas weighted equally) | 20% | Scopus abstract and citation database |
| **Employer Reputation** | Employer survey with design focused to evaluate the perceptions of global employers on which institutions are providing the most job-ready graduates | 15% | QS Employer Reputation Survey |
| **Employment Outcomes** | An indicator of the ability of institutions to ensure a high level of employability for graduates, while also nurturing future leaders who go on to make an impact in their respective fields | 5% | Two metrics from QS Graduate Employability Rankings |
| **Faculty Student Ratio** | Ratio of academic staff to students and intended to evaluate commitment to learning and teaching | 10% | Data provided by institutions |
| **International Faculty Ratio** | Proportion of international faculty members, providing a proxy for how internationally attractive a university is to academic staff | 5% | Data provided by institutions |
| **International Research Network** | This indicator assesses the richness and diversity of an institution's international research partnerships | 5% | QS calculation - adaptation of Margalef Index |
| **International Students Ratio** | Proportion of international students, providing a proxy for how internationally attractive a university is to students | 5% | Data provided by institutions |
| **Sustainability** | Evaluation of the social and environmental impact of universities as centres of education and research | 5% | Extraction from analysis of QS Sustainability Ranking |

Figure 1: official description of criteria

## 2 How Do The Scores of World University Rankings Distribute?

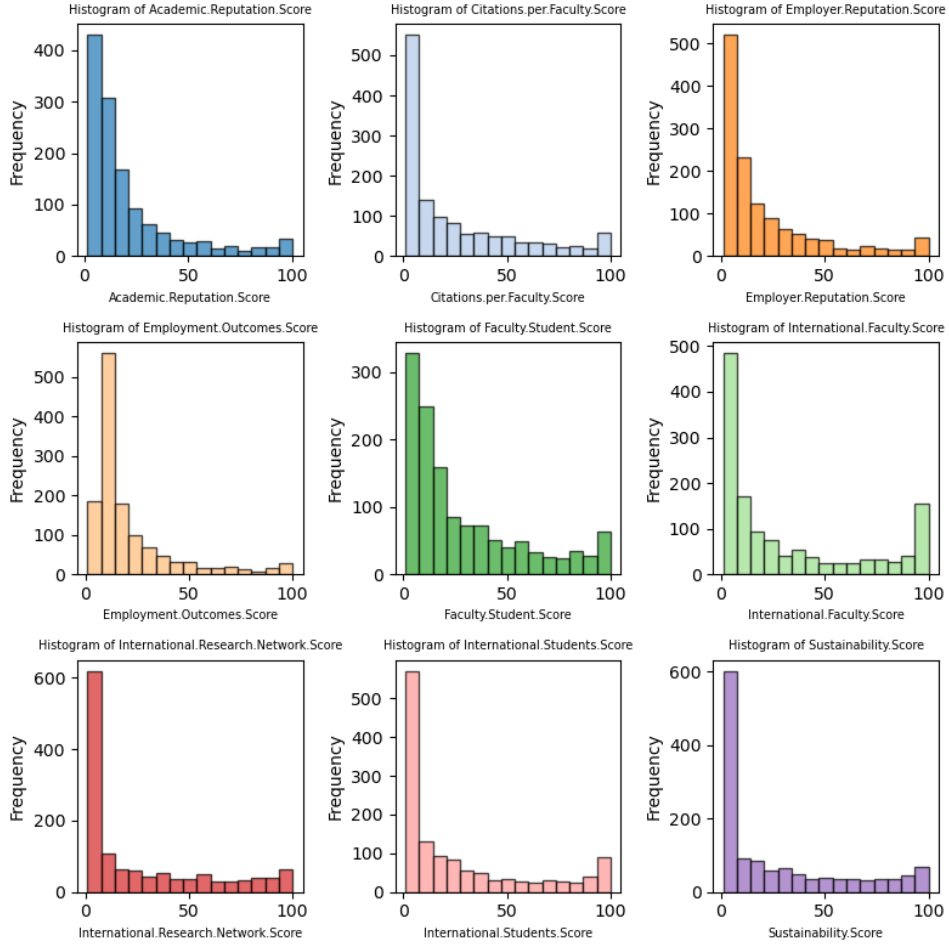



Figure 2: statistical distributions about WUR scores.

## 3 Cluster Analysis

**It appears that the existing technique using the overall score to rank the universities seems to be simple as the overall score is just a linear combination of the variables with their respective weightings.** Since the rankings data are of multivariate nature, therefore it is possible to use multivariate statistical techniques like cluster analysis.

### 3.1 Cophenetic Correlation Coefficient

Clustering is performed on the basis of similarities or dissimilarities. Distance as dissimilarity measure provide an insight of how close observations are to each other in a set. **Euclidean distance and Gower's**

**distance are used in our analysis.** Gower's distance can be used to measure how different two records are. It takes values between 0 and 1 with 0 representing maximum similarity and 1 no similarity. We use the nine score indicator variables to compute the Euclidean distance and Gower's distance. The Gower's distance are computed with weights:

$$(0.3, 0.2, 0.15, 0.05, 0.1, 0.05, 0.05, 0.05, 0.05)$$

The conventional linkage methods are single linkage (minimum distance or nearest neighbour), complete linkage (maximum distance or farthest neighbour), average linkage (average distance), centroid method and ward method.

Consequently, we are faced with an arbitrary choice both of method and of distance metric and it is usual to try all or as many as possible choices of these and then compare the results. In order to ensure a good clustering, we can employ the cophenetic correlation coefficient technique for comparing dendrograms. **The cophenetic correlation coefficient serves as a measure of degree of fit of a clustering to a dataset and as well as a criterion for evaluating the efficiency of various clustering techniques.**

$$c = \frac{\sum_{w<y}[d(\mathbf{w},\mathbf{y}) - d][t(\mathbf{w},\mathbf{y}) - t]}{\sqrt{\sum_{w<y}[d(\mathbf{w},\mathbf{y}) - d)^2]\sum_{w<y}[t(\mathbf{w},\mathbf{y}) - t]^2}}$$

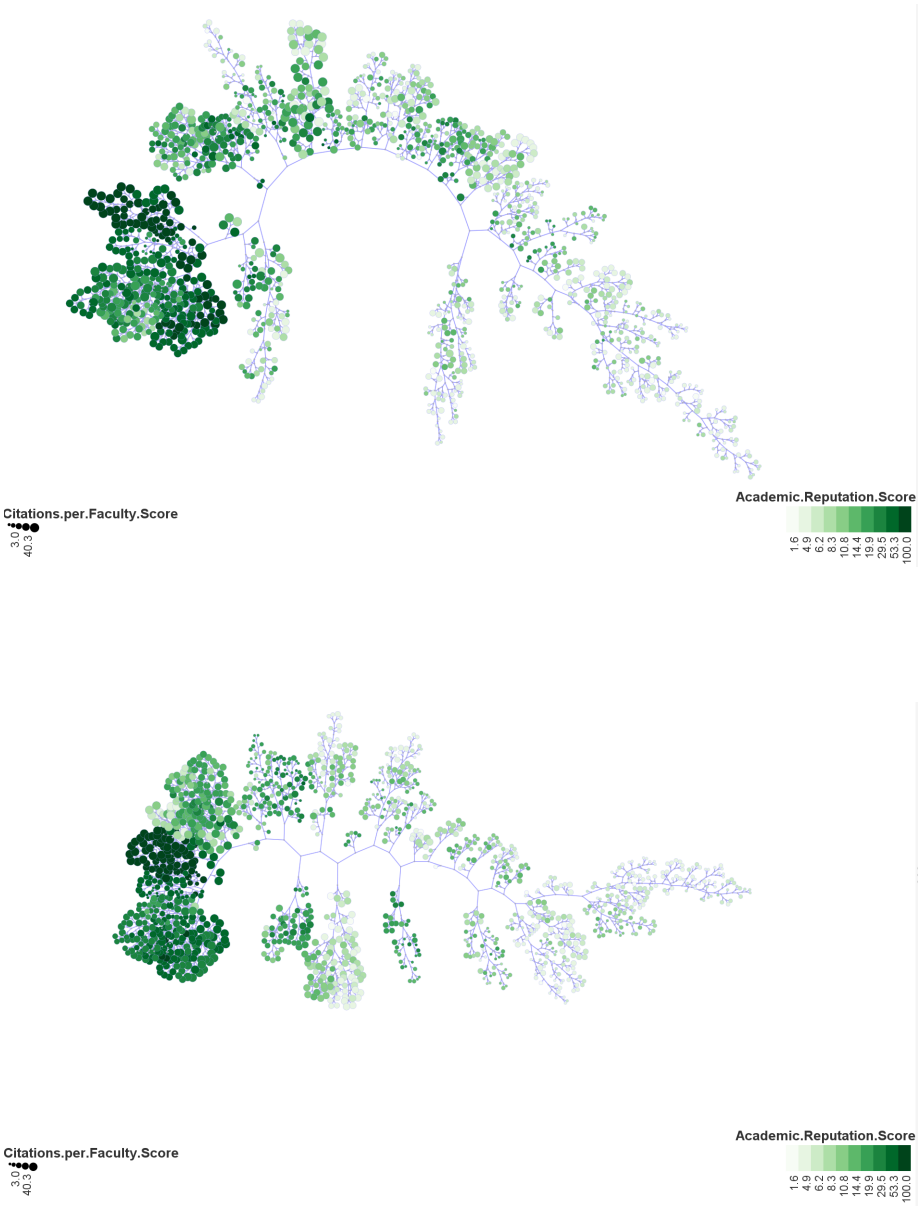
The values of c for the possible combinations of linkage methods and distance metrics are presented below.

| Linkage Method | Euclidean | Gower |
|---|---|---|
| Single | 0.7344228 | 0.7396017 |
| Complete | 0.828204 | 0.8314316 |
| Average | 0.8620978 | 0.8805992 |
| Centroid | 0.8191313 | 0.8615349 |
| Ward | 0.592691 | 0.623357 |

From the table, it suggests that the **average linkage** has the highest cophenetic correlation coefficient among the two distance metrics. Thus we choose the average linkage method for the agglomerative hierarchical clustering using two different distance metrics.

### 3.2 TREEANDLEAF

When n is large (in this case, n=1310), it is difficult to visualize the clusters. Thus we proceeded to generate the TreeAndLeaf graphs supported by R package "TreeAndLeaf" for the two hierarchical clustering respectively.

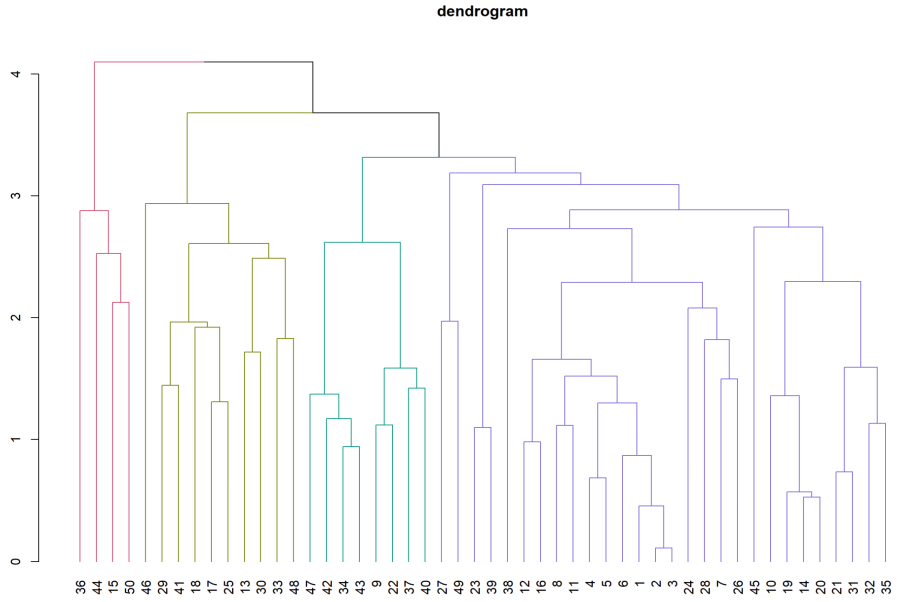



The color represents the university's academic reputation score and the size of the point represents the citations per faculty score. These two indicators have the highest weights in the QS ranking.

**The two graphs perform similarly: there are several branches in the left having the greenest color and the largest point size. The point in the back is getting lighter and smaller, especially in the tail.** Secondly, there is a very obvious green branch in the left of the two graphs. Thus we consider to do the hierarchical clustering only for the first 50 universities.

### 3.3 DENDROGRAM

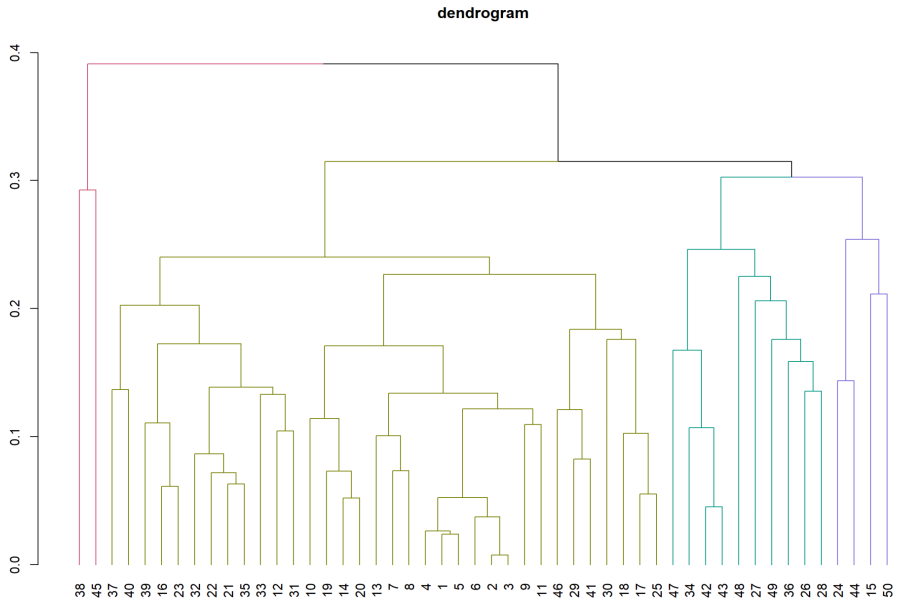
The results of agglomerative clustering for the top 50 universities can be well visualized in the form of a tree diagram known as a dendrogram. **A dendrogram shows the nested structure of the partitions and how the various clusters are linked at each stage.** Here shows the dendrograms for the two hierarchical clustering for the top 50 universities.




From the dendrogram, we can see that the **distance between the rank 2 and rank 3 universities are the closest and so they joined together foremost in the euclidean distance metric.** Then it comes to universities rank 1,6,4,5.

This is consistent with the 2024 and 2023 QS university rankings showed below, **the top six were all produced in these schools**.

| 2024 Rank | 2023 Rank | Institution Name |
|---|---|---|
| 1 | 1 | Massachusetts Institute of Technology (MIT) |
| 2 | 2 | University of Cambridge |
| 3 | 4 | University of Oxford |
| 4 | 5 | Harvard University |
| 5 | 3 | Stanford University |
| 6 | 6 | Imperial College London |




Again, the top six quickly combined together in this dendrogram. **It is clear that the first 6 universities are in a compact group and more similar to each other.** Whereas, the rank 7 university seems has a longer distance to the top six universities. In our intuition, we may consider the top five or top ten as the first tier, **however from the natural clusters it is evident that the rank 7 university is in fact does not belong to the first tier.**

### 3.4 RANKING

For the two distance measures, we can **rank universities based on the distance** between their scores and the full scores for each score indicator variable.

| 2024 Rank | 2023 Rank | Euclidean | Gower |
|---|---|---|---|
| 1 | 1 | 3 | 1 |
| 2 | 2 | 1 | 2 |
| 3 | 4 | 2 | 3 |
| 4 | 5 | 5 | 4 |
| 5 | 3 | 6 | 5 |
| 6 | 6 | 4 | 6 |
| 7 | 9 | 8 | 8 |
| 8 | 11 | 7 | 7 |
| 9 | 8 | 11 | 9 |
| 10 | 27 | 16 | 11 |
| 11 | 10 | 14 | 10 |
| 12 | 13 | 9 | 13 |

The results are consistent with the 2024 and 2023 QS rankings. What's more, **the top six are still fluctuating among these schools**, which is consistent with the results of the previous analysis. Compute the

correlation coefficient between the overall score and distance in the following table. **The strong linear relationship between the overall score and distance verifies the agreement between distance rankings and the existing QS rankings.**

| | Euclidean | Gower1 |
|---|---|---|
| Correlation coefficient | -0.9800178 | -0.9987225 |

In fact, using the linear combination of the scores of each indicator for ranking is equivalent to using the **weighted absolute distance** between each university and the full score to rank. **Hence, it is natural to use other distance metric to rank universities.**

### 3.5 FURTHER DEVELOPMENT

**Current methodology for university rankings has bias towards the natural and life sciences.** Measures such as citation counts do favour universities which are strong in the fields of medicine and natural sciences, where there is a strong publishing and citation culture. Arts and humanities-focused universities produce many fewer research papers than science-focused ones. **Sub-rankings can be an option.**

Since gower's distance can apply to both continuous and discrete variables, **next step we can try to include discrete variables, such as university's status, age, research intensity, focus, size, into the distance measures.** This will be a huge advantage of distance rankings.

## REFERENCES