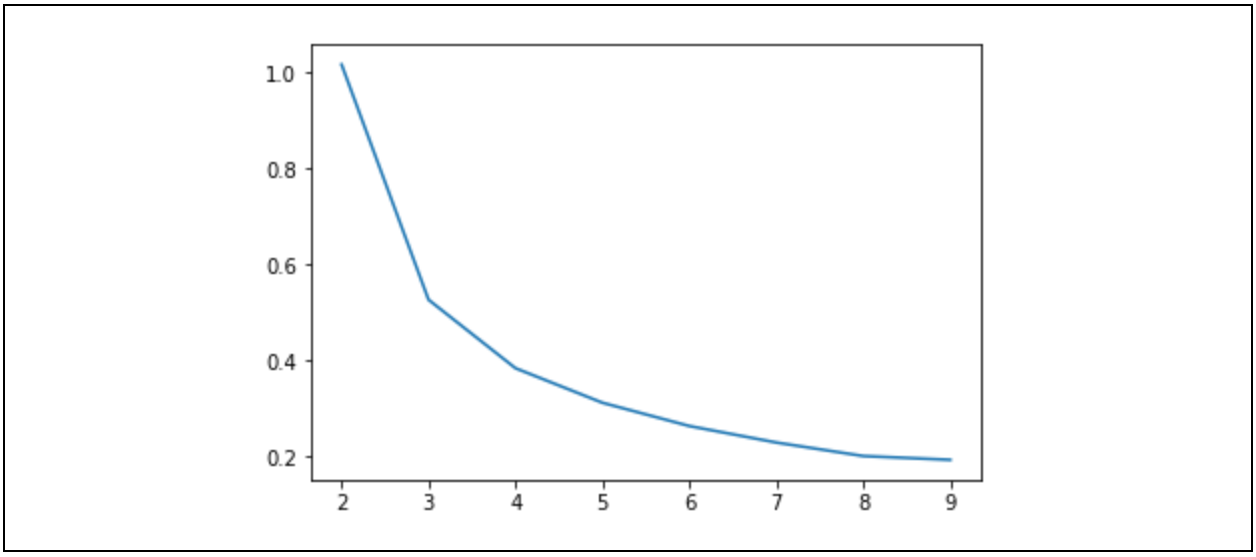
1 Lloyd's algorithm

1- Please see the notebook at the end of this document

2- Please see the notebook at the end of this document

3- The results for k = 2, 3, ..., 9 is considered drawing the elbow graph. Please see the notebook at the end of this document

4-



5- Based on the above cure, k = 3 or 4 may be selected. Both of these values are close to 3 which is the actual number of classes.

6-

- 100% of samples that belong to category #1 fall into same cluster (cluster 1).
- 96% of samples that belong to category #2 fall into the same cluster (cluster 2).
- 72% of samples that belong to category #3 fall into the same cluster (cluster 3).

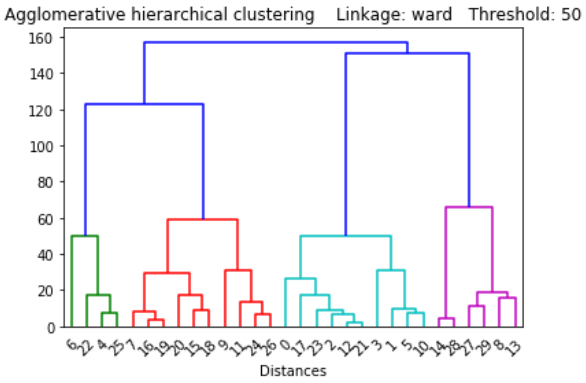
The geometric mean of these ratios equals 0.8841675596736928, which indicates that generally, 88% of the samples fall into the right category.

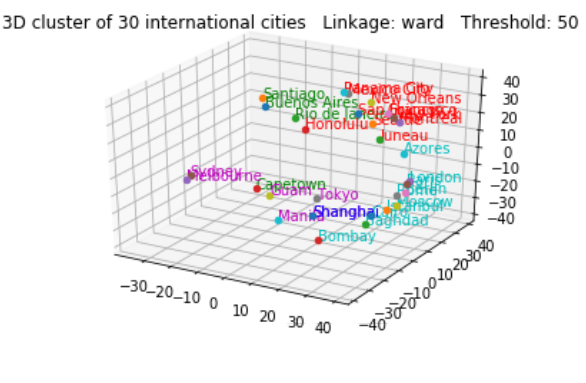
2 Hierarchical clustering

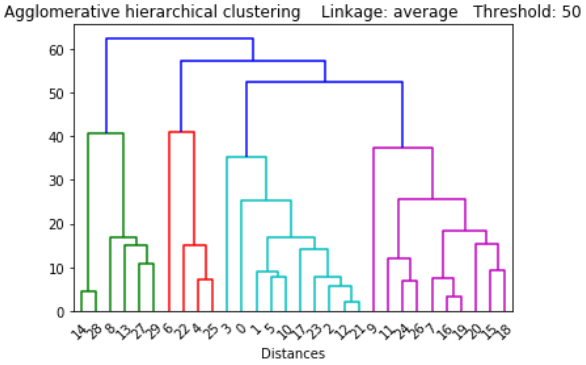
1- -
2- Please see the notebook at the end of this document
3- Please see the notebook at the end of this document
4- Please see the notebook at the end of this document

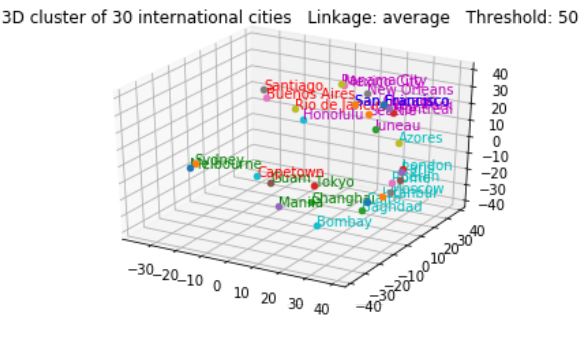
5- Please see the notebook at the end of this document

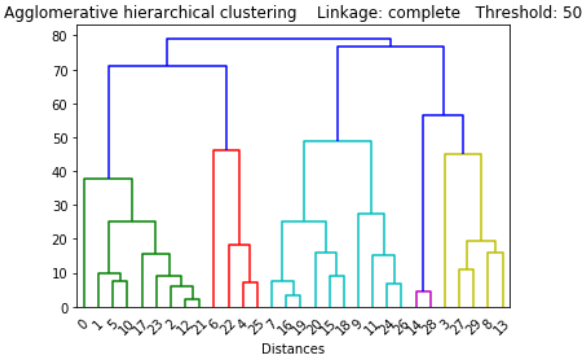
6-

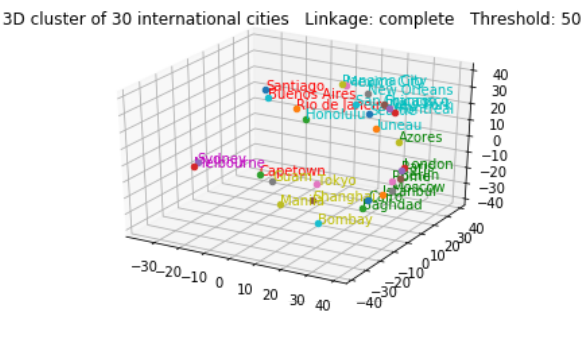


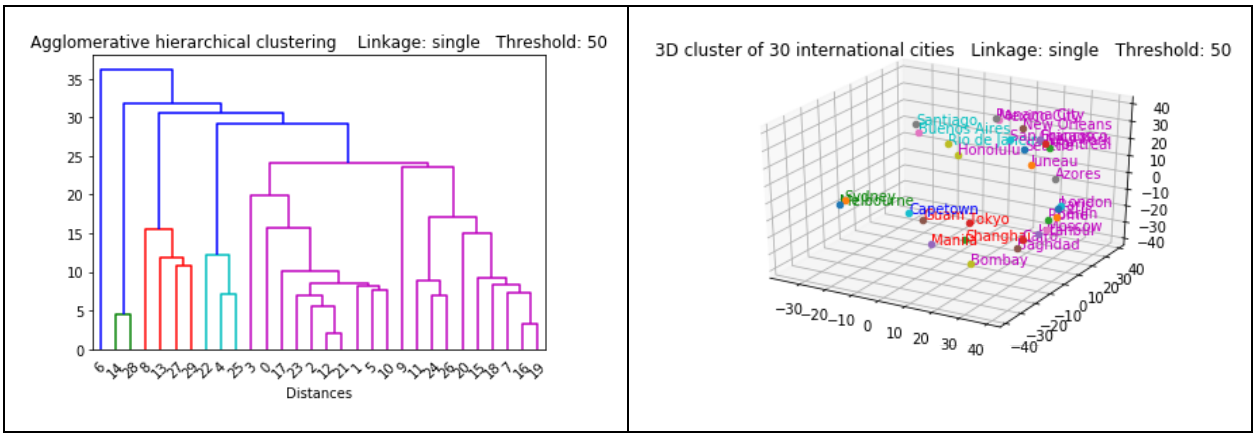




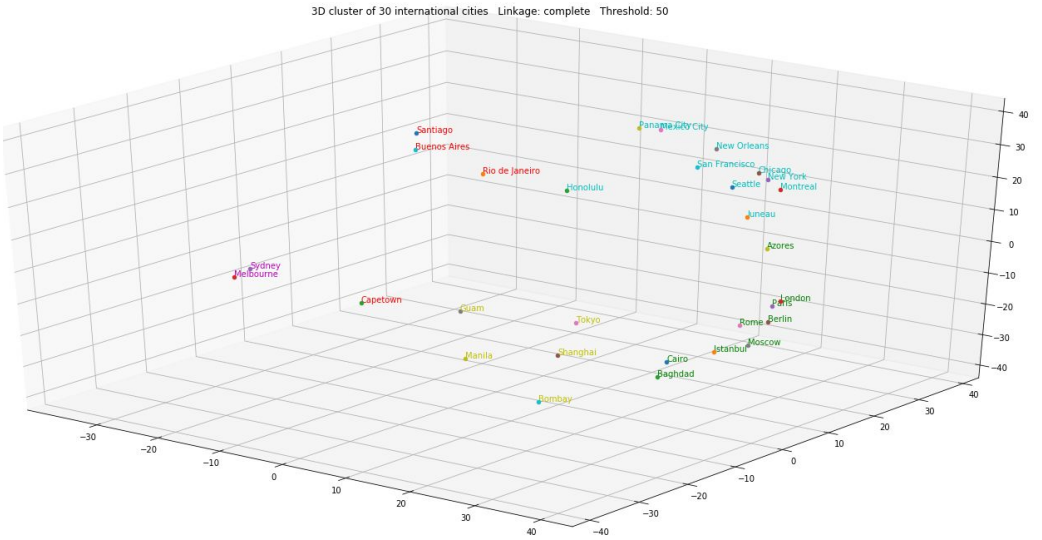








Most realistic: complete; the clusters seem resealable according to their positions in the space.



Least realistic: single; the cluster with text color violet can be divided into two parts.