

1.4 Programming Challenge

Question 1.4.11: Data Collection

The dataset for the problem was acquired from the following web page:

<https://simplemaps.com/data/world-cities>

The data contains latitudes and longitudes for the world's cities. I then used the Python package Geopy to calculate the geodesic distances between the cities. See Code appendix 1.4.11 for Python code.

Question 1.4.12: Classical MDS

Input: Distance matrix D

Output: 2-dimensional embedding X

Data: City distances

1. Compute a distance matrix D according to the steps outlined in part 1.4.11
2. Compute a similarity matrix from $D : S \leftarrow \text{similaritymatrix}(D)$
3. Compute Eigen-decomposition of $S : [D, Q] \leftarrow \text{Eig}(S)$
4. Order D in descending order of eigenvalues magnitude and Q correspondingly
5. Ensure elements of D and Q are real
6. Compute embedding: $X \leftarrow I_{2 \times 102} D Q^T$

The results of the MDS are shown below. Cities from corresponding continents are correctly clustered together. The placement of Africa (represented by the dark blue dots) is incorrect. South America is shown at roughly the correct longitude but at the same latitude as North America. Asia, Europe and North America are very close to the correct placement on the map. See Code appendix 1.4.12 for Python code showing MDS function and plot.

Question 1.4.13: Metric MDS

Using the Metric MDS package from SKlearn yielded similar results to the classical MDS function defined above. Increasing the number of iterations in the parameter `max_iter` increased the accuracy of the map.

The parameter 'dissimilarity' can be set to either 'euclidean' or 'precomputed'. Since the distance matrix was already calculated using geodesic distances in part 1.4.11 we can use 'precomputed' as the parameter.

In this result we see Asia, Europe, North America and Africa in roughly the correct position. However South America is shown to the west of North America instead of to the south.

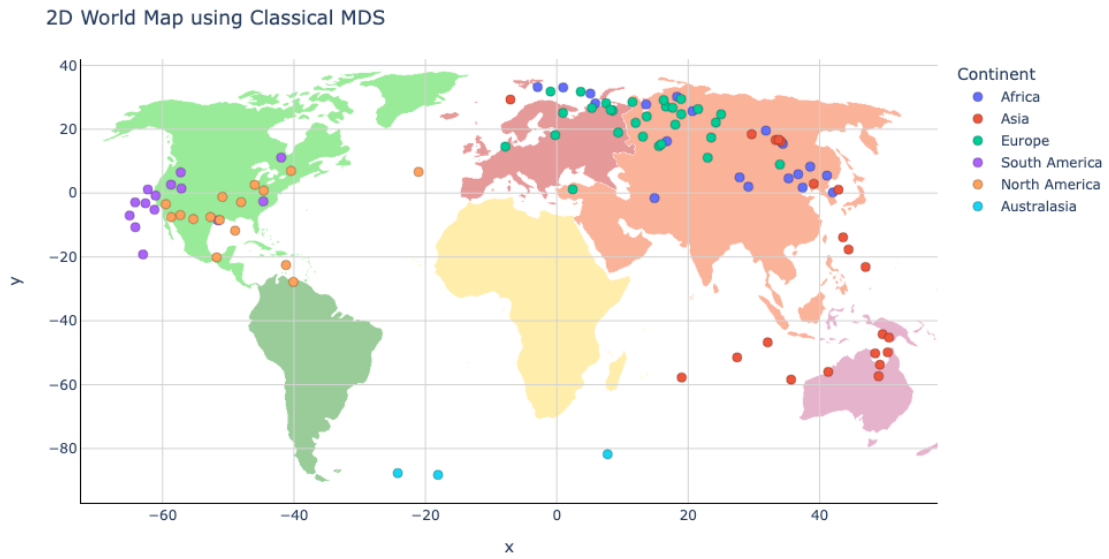


Figure 2: Classical MDS results plotted on a world map

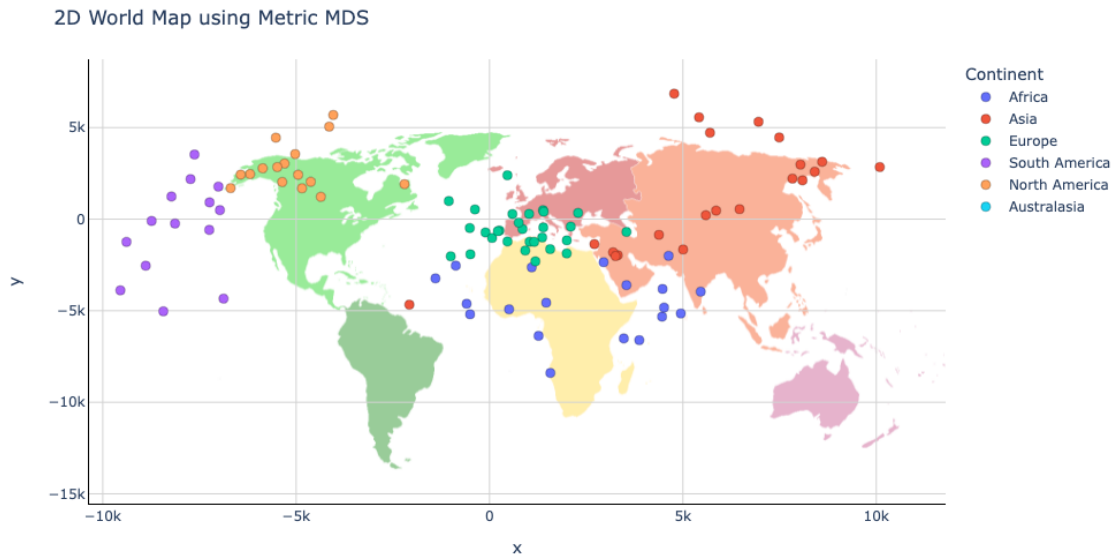


Figure 3: Metric MDS results plotted on a world map