ECEN 689 – Applied Information Science

Population Data Visualisation of Countries

Harish Chigurupati, Mahalakshmi Sridharan, Ranjith Tamilselvan

Objective:

To visualize the population correlation prediction between different countries from the year 2000 to 2016 based on the population from year 1960 to 1999.

Primary Analysis:

The primary analysis was conducted by visualizing the population_parameter.csv file using Gephi. The countries correspond to the nodes and the edges correspond to the correlation between the countries based on the coefficients computed for each of the countries.

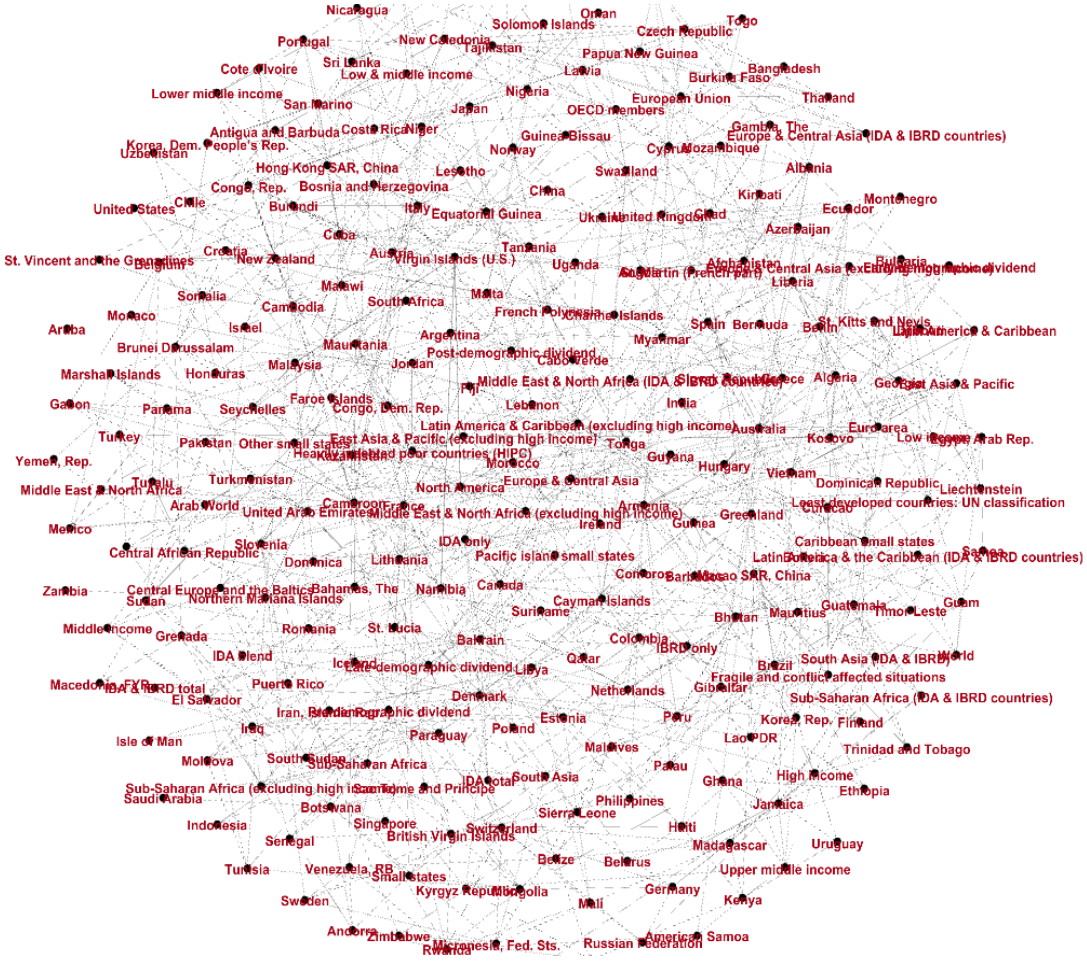

Figure 1: Graphical Fruchterman-Reingold 3D layout of 258 country nodes

Case 1: Population Correlation between United Kingdom and other countries

On analysis of the correlation of population of United Kingdom with the countries that influence its population prediction, we infer that countries like Canada which have stable population and negligible population increase through the years have a strong influence. Hence, the predictions are validated

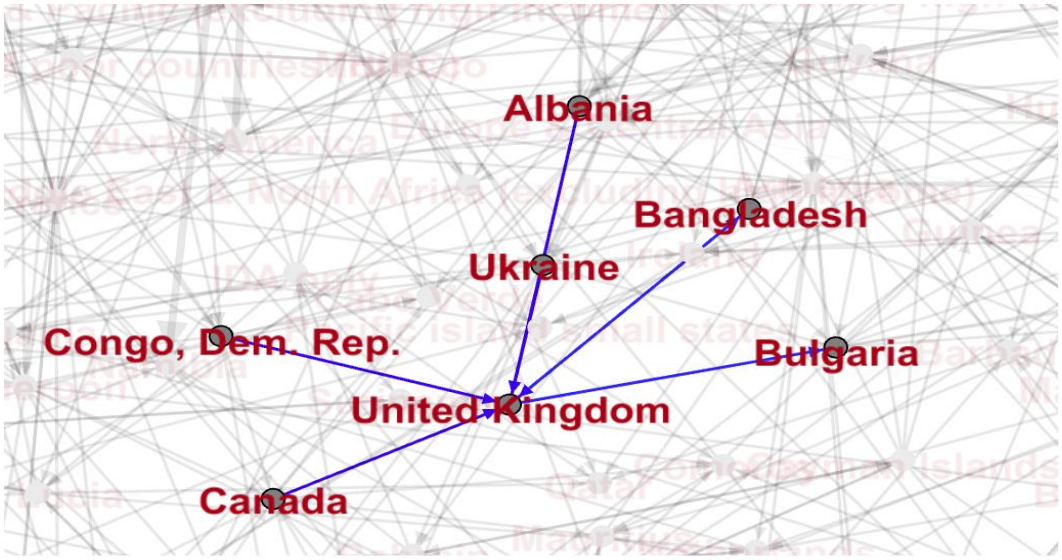

Figure 2: Population correlation of United Kingdom

Case 2: Population Correlation between Senegal and other countries

Prediction of Senegal reveals that it is based on demographics of neighbouring countries. Thus, the demographics of a country can be related to the demographics of neighbouring countries.

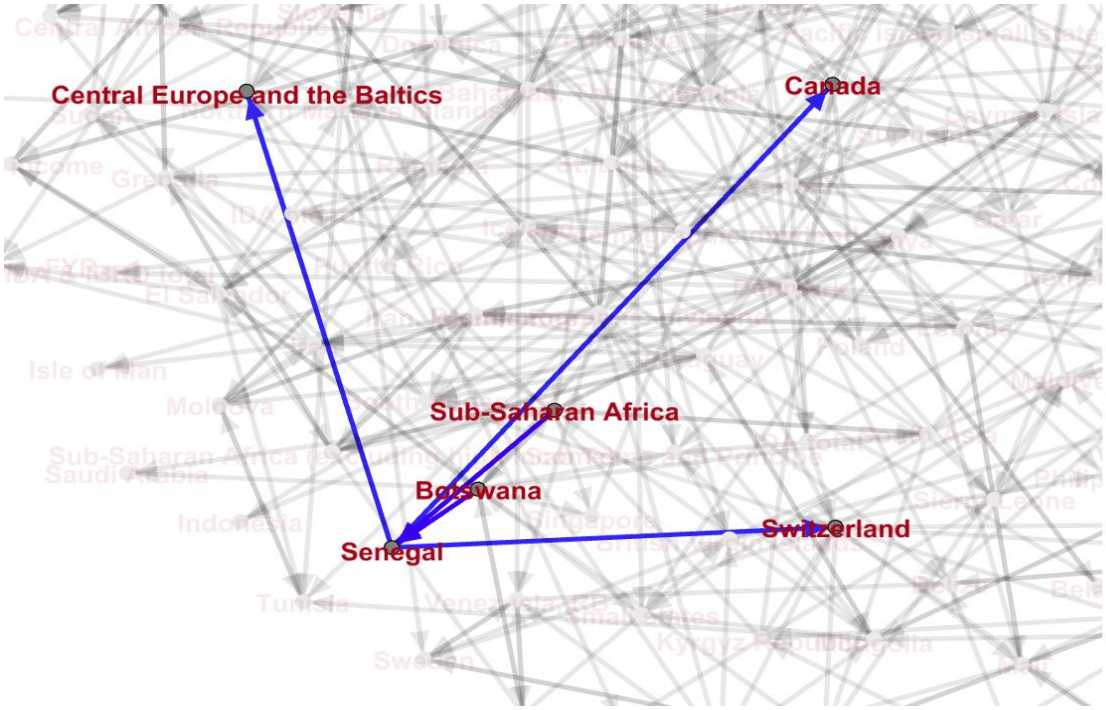

Figure 3: Population correlation of Senegal

Case 3: Population Correlation between Canada and other countries

Similarly, population correlation analysis has been conducted on Canada. We find that demographics of countries can also be related to the economic stability and life-expectancy of the country – hence developed countries like Canada show strong correlation with other developed countries like United Kingdom, Poland.

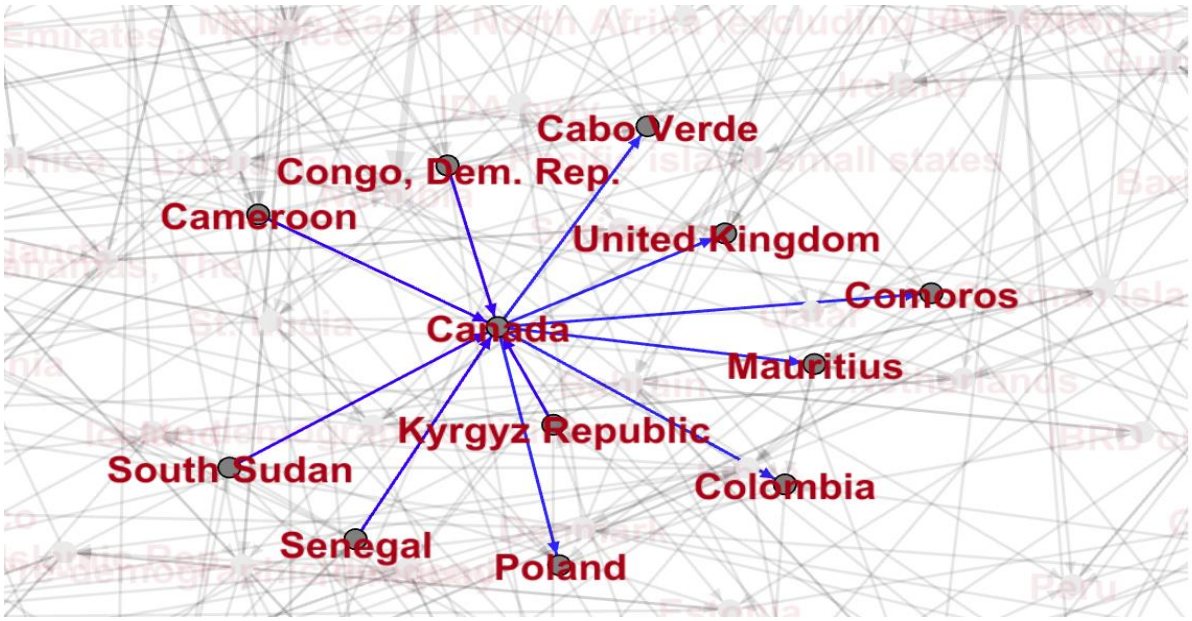

Figure 4: Population correlation of Canada

Case 4: Population Correlation between Tajikistan and other countries

On analysis of Tajikistan, we notice the similarity of population data with that of other developing and under-developing countries like Cambodia, Papua New Guinea, Burkina Faso.

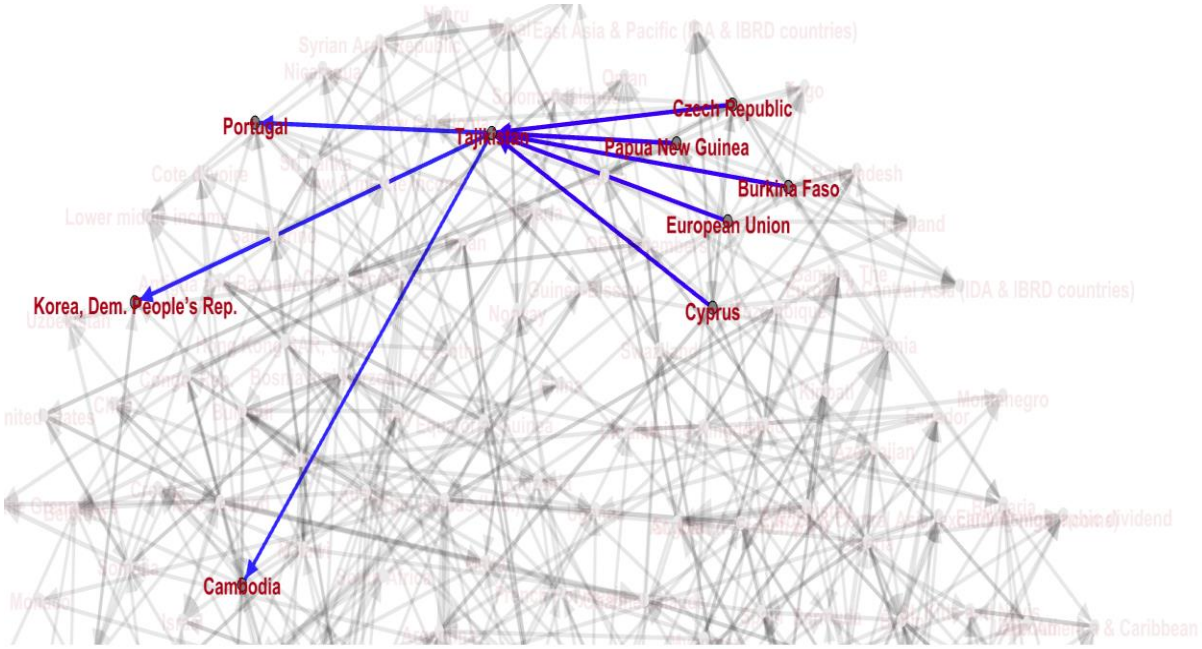

Figure 5: Population correlation of Tajikistan

Conclusion:

Fruchterman-Reingold model of visualization was utilized to graphically interpret the population correlation between countries. This led us to infer that the countries that have population correlation between them have similar demographics, geographical location or economic conditions.

References:

[1] Benjamin Bengfort and Konstantinos Xirogiannopoulos (2015, April 6). Visual Discovery of Communication Patterns in Email Networks. Retrieved from http://www.cs.umd.edu/~bengfort/papers/visual-discovery-email-networks.pdf

[2] Sebastian Raschka, Python Machine Learning. Retrieved from http://liuchengxu.org/books/src/Machine%20Learning/Python-Machine-Learning.pdf