

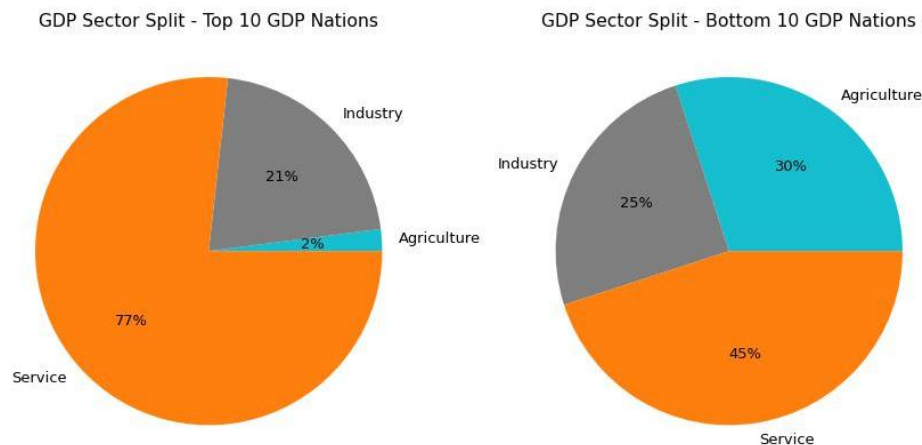
Applied Data Science 1

Dataset description

This study is performed on Kaggle's [countries of the world](#) dataset which was originally taken from CIA's The World Factbook. This dataset includes records of 227 countries across the world. Each country in the dataset has been described with the help of 20 attributes.

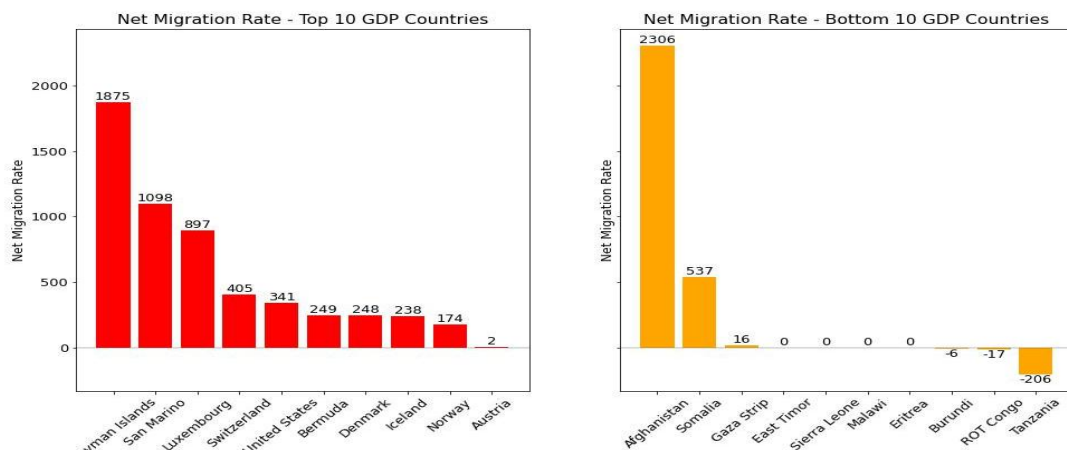
Economic Split Gross Domestic Product (GDP of top 10 nations with highest GDP and top 10 nations with the lowest GDP)

The first observation is that among the nations with highest GDP on an average only 2% economy depends on agriculture. This implies that in these nations majority of the agricultural products are imported. The remaining 98% of the economy depends upon industries and services. On an average 77% of the economy is driven by the service sector and the remaining 21% by the industrial activities. The reason behind marginal contribution of agricultural sector is the size of the nation as well. Except the United States in this category the remaining countries are either island nations or are smaller in size.



In case of nations with lowest GDP, on an average 30, 25 and 45 per cent of the economy in these nations depends upon agriculture, industries and services. The reason behind agricultural sector huge contribution to economy that population in these nations is does not have adequate facilities and capital power to boost their economy at a fast pace. 7 out of 10 nations are part of Sub – Saharan Africa. This explains the 25% contribution to the economy of industrial sector. Lastly, majority of the populations consists of middle class and is employed in various service based companies.

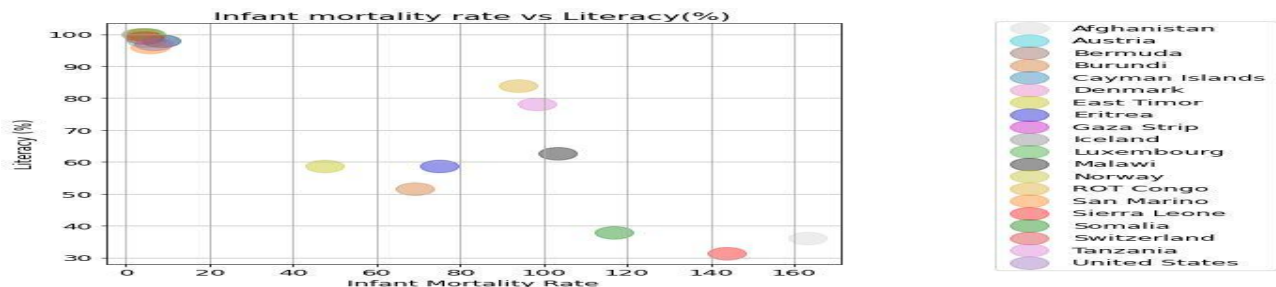
Net Migration Rate



In the above graphs it is clear that the highest GDP nations have NMR of more than 100 per 1000 individuals except Austria where the NMR is only two.

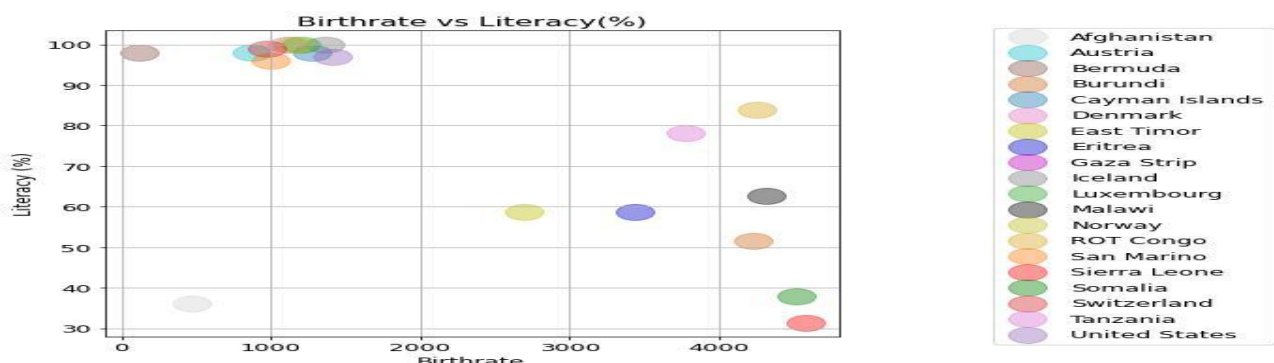
On the contrary in the nations with weakest economy except three nations Afghanistan, Somalia and Gaza Strip, the NMR per 1000 individuals is either 0 or negative which indicates that number of emigrants is either equal to or less than number of immigrants.

Infant Mortality Rate VS Literacy



On comparing Infant Mortality Rate per thousand children and literacy per cent, Afghanistan, Sierra Leone and Somalia have the least literate population as well as the highest infant mortality rate in the world. The graph clearly indicated that there is a negative correlation between the IMR and Literacy(%).

On contrary on comparing Literacy(%) with the Birthrate, the observation is that there is no direct correlation emerging from the graph. However, there are three separate groups emerging from analysis. The first one with more than 90% literacy in each country and their respected birth rates are less than 1500.



Whereas, the second group is scattered along literacy, indicating each of the countries have different literacy levels. Despite this observation the birthrate is more than 4000. The last group consists of countries where literacy rate lies between 55% and 80% and their birth rates are between 2500 and 4000. Lastly, an outlier in the analysis is Afghanistan where both literacy rate and the birth rates are least among the twenty countries under study. This highlights the lack of infrastructure in Afghanistan and also the food scarcity.

References

- <https://www.ers.usda.gov/amber-waves/2013/august/developing-countries-dominate-world-demand-for-agricultural-products/> pppp
- <https://www.washingtonpost.com/news/worldviews/wp/2016/06/24/afghan-babies-have-been-dying-in-huge-numbers-for-decades-now-something-is-changing/>
- https://www.openintro.org/data/index.php?data=cia_factbook
- https://en.wikipedia.org/wiki/Net_migration_rate
- Saurabh, Suman, Sonali Sarkar, and Dhruv K. Pandey. "Female literacy rate is a better predictor of birth rate and infant mortality rate in India." *Journal of family medicine and primary care* 2.4 (2013): 349.