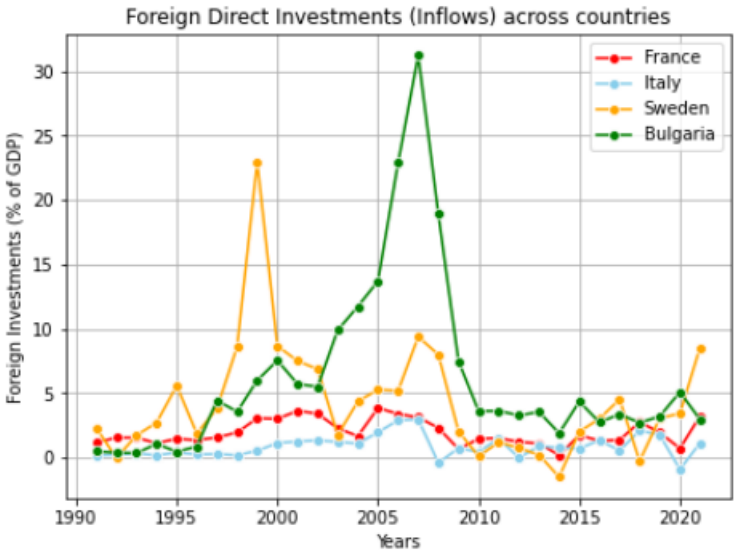
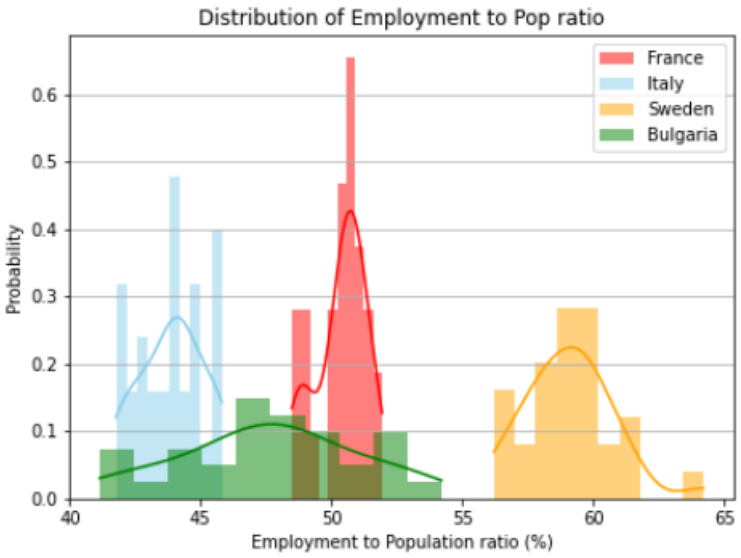
GDP Analysis: Comparison of various parameters affecting GDP and its projections.

Introduction:

GDP (Gross Domestic Product) is a standard practice across the world to measure the value generated by a country in terms of income through the production of goods and services during a period. Measuring GDP involves considering a lot of parameters out of which we have considered the following GDP per capita, employment to population ratio, foreign direct investments inflows and outflows, expense, and exports of goods and services. Furthermore, the data set comprises GDP along with all the above factors and a few countries (Bulgaria, France, Sweden, and Italy) from 1991 to 2021.



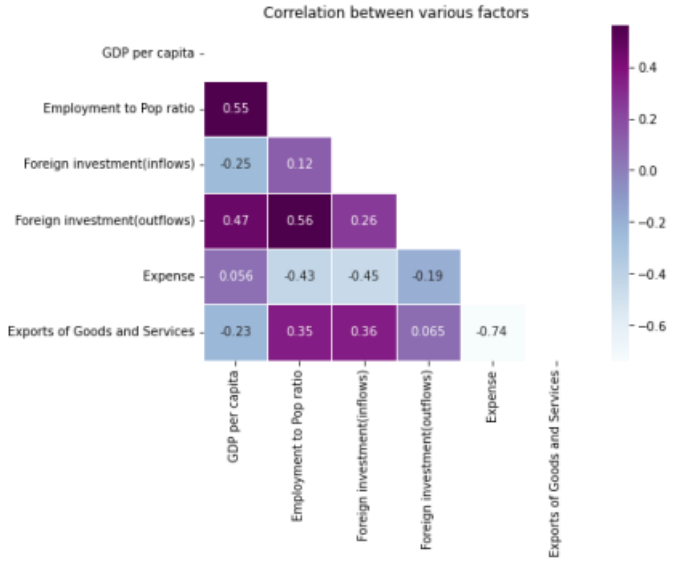
Starting the analysis with a line graph that depicts the comparison between foreign investments inflows (% ofGDP) across the considered countries. The percentage of foreign investment was highest from 2005 to 2010 at 33% for Bulgaria, which would mean that the country's inflow was higher than the outflow investments. Whereas Italy showed the least value in terms of the percentage of foreign investment from which it is safe to assume that during this period Italy's outflows were higher than the inflow investments. Sweden also showed similar trends in foreign investments as Bulgaria from 1995 to 2000. On the other hand, France had an almost constant foreign investment from the beginning of 1990 to the end of 2020. Now, for further investigation, the employment-to-population ratio histogram will be analysed.



The histogram depicts the distribution of employment to population ratio of the countries. This is an important step in the analysis of the GDP of a country as the employment ratio gives the quality of labour in a country which in turn depicts the number of Foreign Direct Investments (outflows). Hence, giving a comprehensive idea of the GDP of a country. In the following histogram, it Is evident that the employment ratio in Sweden is high when compared with the rest of the countries, and in any given year it had at least 56 of 100 people employed. France has an employee ratio spread between 48 to 52 people out of 100. Italy has the lowest of all employment-to-population ratios, 42 to 46 are employed per 100 people all over these years. Bulgaria's data shows a vast range, few years it had 41 employees per 100 people but again in a few years it had 54 employees per 100 population and for a lot of years it was between 45 to 50 people employed per every 100 people available in the country. A look into this data is relevant is reflected further down this report.

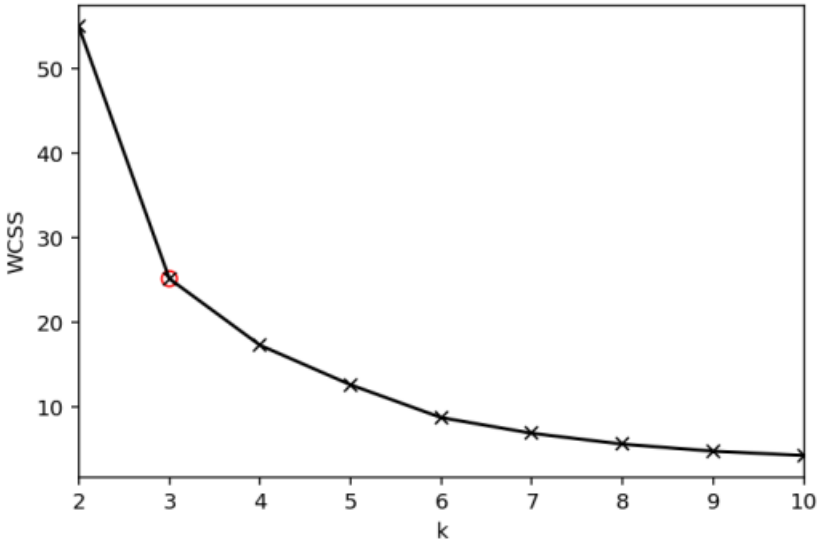
Now, let's check some statistics of the data, the mean of the GDP per capita for all the countries is 27894 $, and the minimum value is 1148 $ whereas the maximum is 61647 $. The Skewness provides insights into the distribution of the data. For GDP and Expense, the value is near to 0 (-0.01 and -0.08 respectively) which indicates that the values are relatively normally distributed across the sample. The Foreign Direct Investments (Inflows) show a high positive skewness (3.42) which indicates the distribution is right skewed.

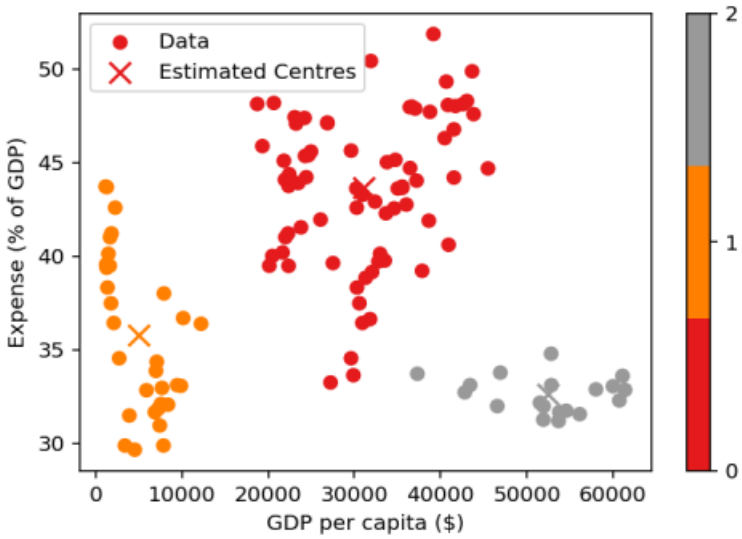
The Kurtosis value provides the nature of the tail and peak of their distributions. The GDP and Employment to population ratio have negative kurtosis values which are -0.7 and -0.9 respectively infers that they have fewer extreme values.

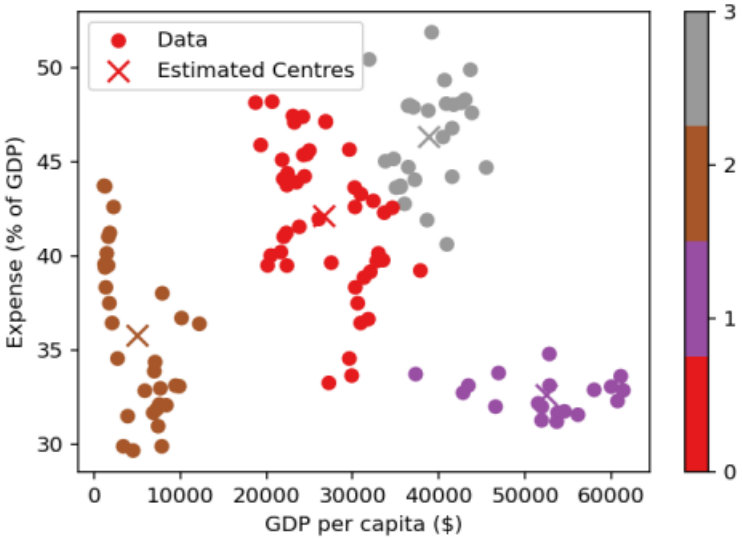


The heat map provides an elaborate view of the correlation between various factors to the GDP of a country. From the heat map, we can depict that the correlation value of employment to population ratio and GDP per capita is 0.55 whereas the correlation value for exports and GDP is -0.23 which indicates that the growth of GDP is more dependent on the employment ratio rather than exports. Moreover, it can be said the higher the employment ratio, the higher the GDP. The heat map also depicts that the FDI Inflows(Foreign Direct Investment) vs GDP of the countries is negative(-0.25), which means that the GDP may decrease with an increase in FDI Inflows. Furthermore, the correlation value for expense to GDP is 0.056 which might mean the countries are importing more than exporting hence, increment in the expenses. Also, we can observe that the Foreign Direct Investments (outflows) have a correlation value of 0.56 with the employment-to-population ratio, this shows that they are positively related which is FDI (outflows) will increase the employment.

Clustering can be performed on the data to identify the groups of objects. It's an unsupervised classification. Here, we perform clustering between GDP and Expense as the correlation value is 0.05 which means the data points are scattered across the plot. This Graph depicts the Elbow plot of the dataset which is to determine the optimal number of clusters(k) for k-means clustering. Also, we have calculated the silhouette score which compares how close data points are within a cluster and the separation between clusters. In the graph, the red circle shows the cluster number given by the silhouette



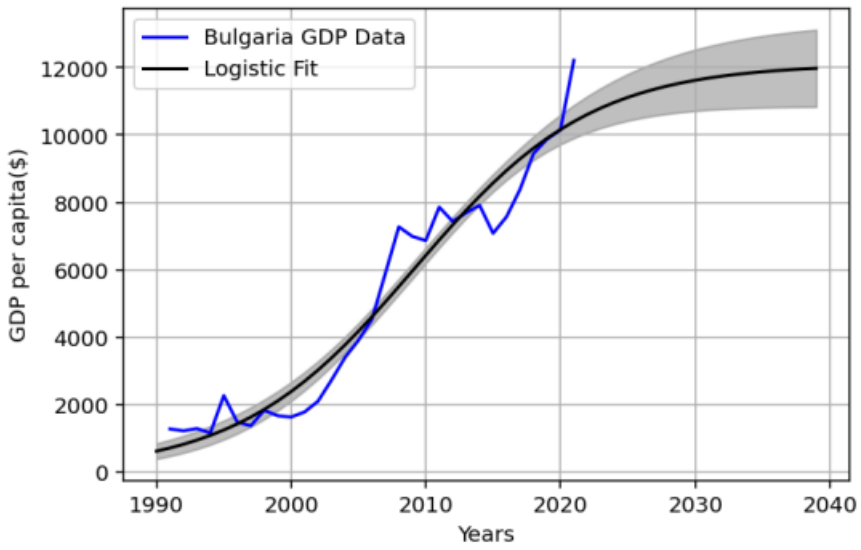




method which is 3. However, we can identify that the elbow method shows the number of clusters as 4 because increasing the number of clusters from 4 to 5 does not yield as much improvement in the compactness of the clusters compared to previous increments which is the WCSS score (Within-Cluster Sum of Square).

By looking at the clusters that have been created by the K means clustering algorithm. We can say that either 3 or 4 clusters seem to be appropriate. However, we can say that four clusters seem to be more appropriate as there are four different countries in the data set, and we can say that each cluster may belong to a country.

The graph shows the actual GDP growth, the logistic fit for the GDP, and the GDP predicted projections by 2040 for the country Bulgaria. There are upswings and downswings in the real GDP growth. There are instances where the logistic fit is exceeded or falls short of the real data, signifying booms, or busts in the economy, respectively. According to the logistic fit, GDP growth will eventually reach carrying capacity and then slow down. This could be caused by several variables, including resource scarcity, market saturation, or economic policies. The graph suggests that GDP will eventually stabilize or expand more slowly rather than continuing to rise at an unabated rate. Finally, from the projections of the fitting curve, the GDP per capita will reach $11,000 by 2040 with a variation of $1000.



Conclusion:

In conclusion, a comparative analysis of GDP in Bulgaria, France, Sweden, and Italy using various economic indicators such as Employment to Population Ratio, Foreign Direct Investments, and Expenses reveals that there is a significant relationship between these variables and the nation's economic performance. The employment-to-population ratio is a key factor in GDP growth than exports of goods and services. Furthermore, clustering analysis aids in identifying the similarities and differences between different nations, therefore offering a personalized view of economic plans and policies. The projections which were given by Logistic fit imply that GDP growth may not continue at its current pace due to various factors, pointing towards more restrained economic development in Bulgaria.

Referred Dataset :

https://databank.worldbank.org/source/world-development-indicators

GitHub Link:

https://github.com/pratapponnam/ADS-Clustering-and-Fitting

Author: Pratap Bhargav Ponnam

ID:23027654