CLUSTERING ASSIGNMENT

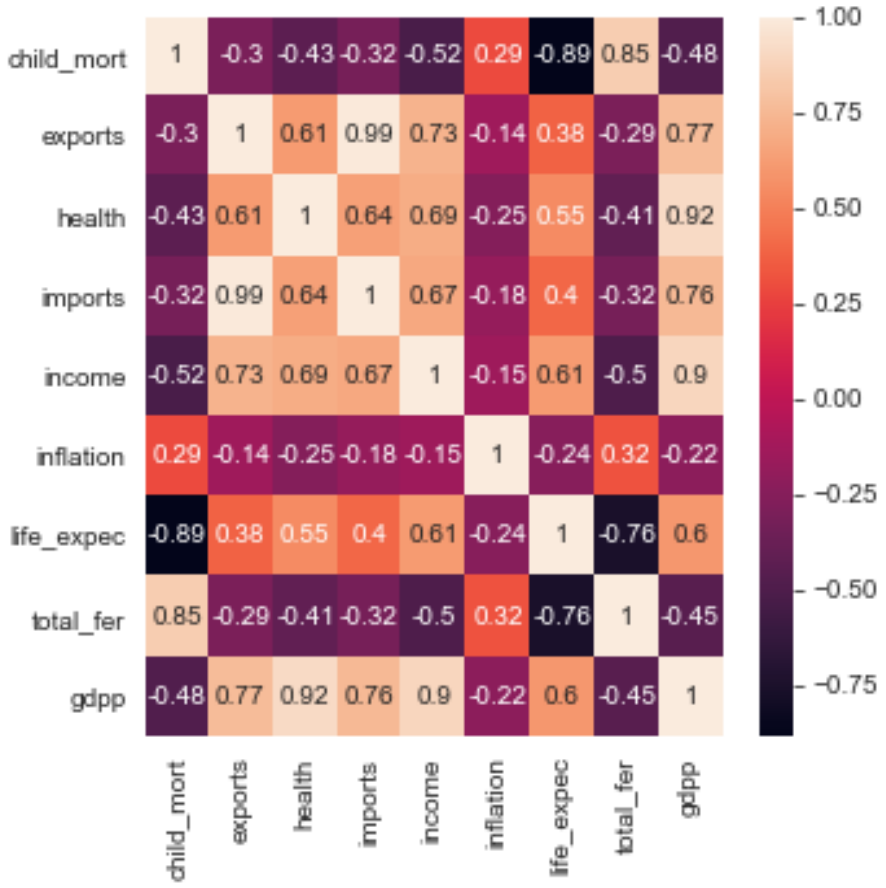
- By Chayan Banerjee

VISUALISATION OF THE VARIABLES

From the heatmap on the right we can conclude below points –

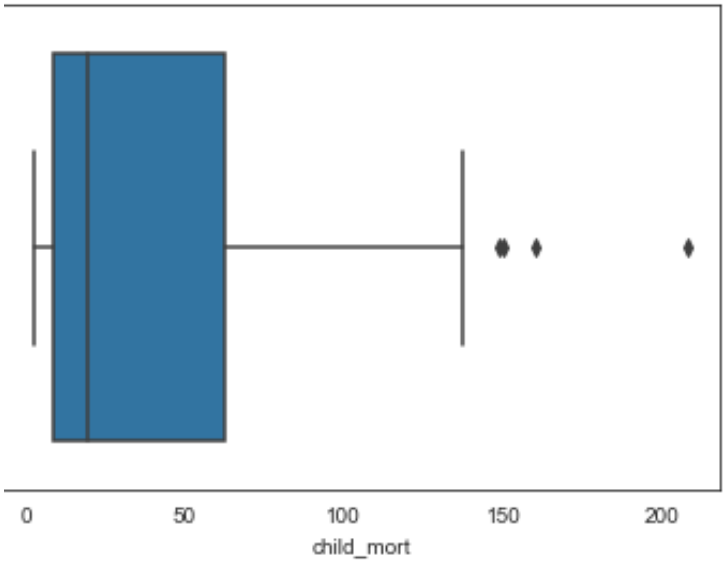
Some of the variables are having high positive correlation like income and gdpp, health, imports etc. Some of them have high negative correlation like child mortality and life expectancy, gdpp, health etc.

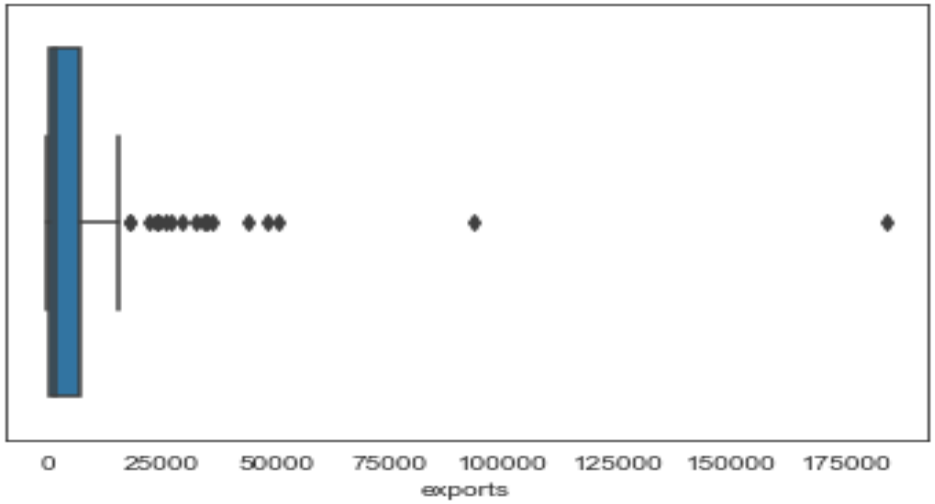
This shows that the dataset is having multicollinearity.

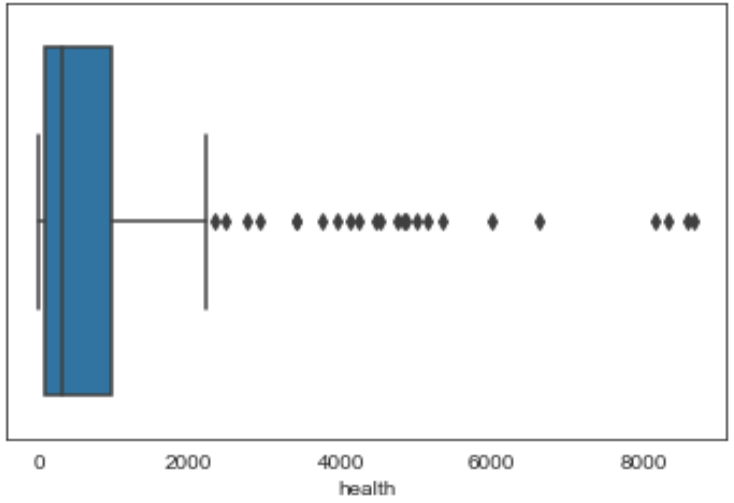


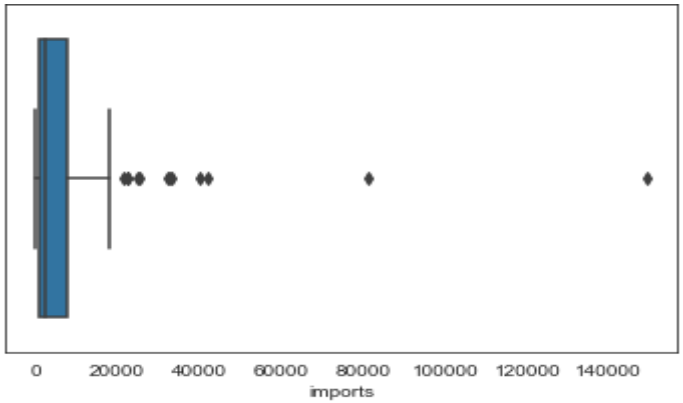
VISUALIZATION OF OUTLIERS

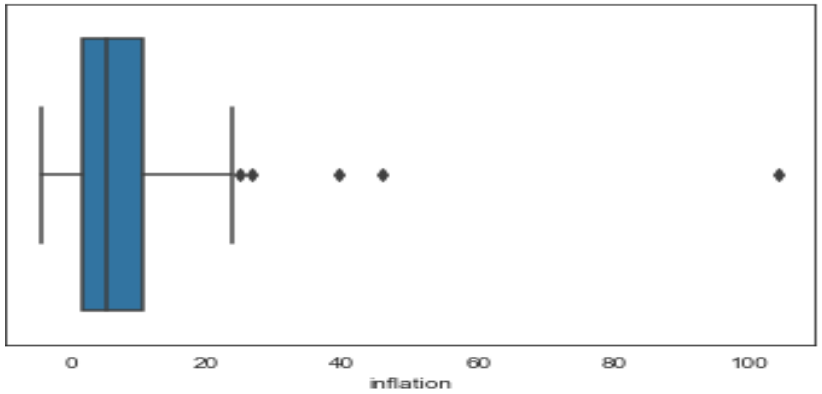
As we can see that all the boxplots created for the variables are having decent amount of outliers

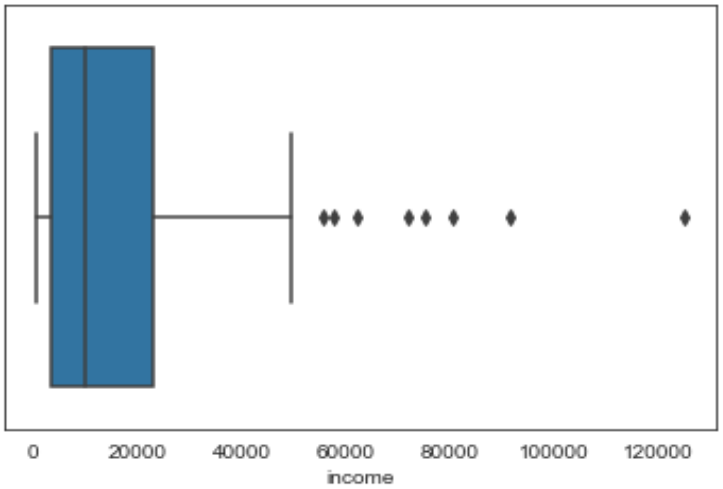


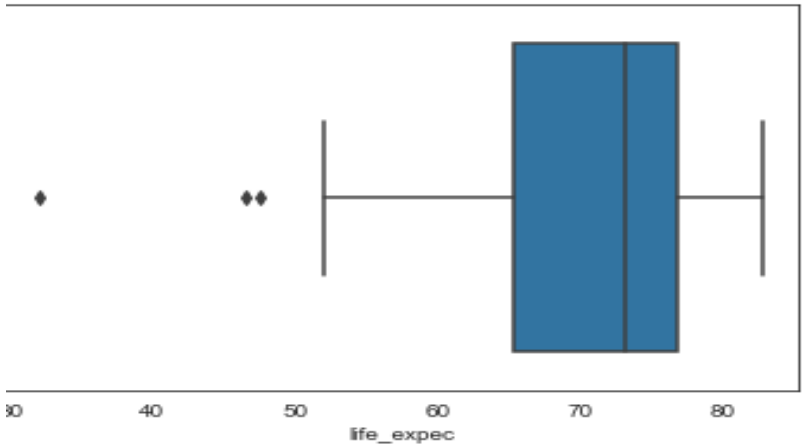


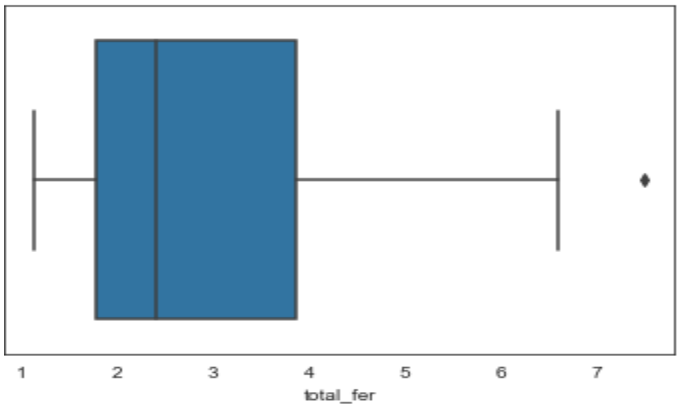


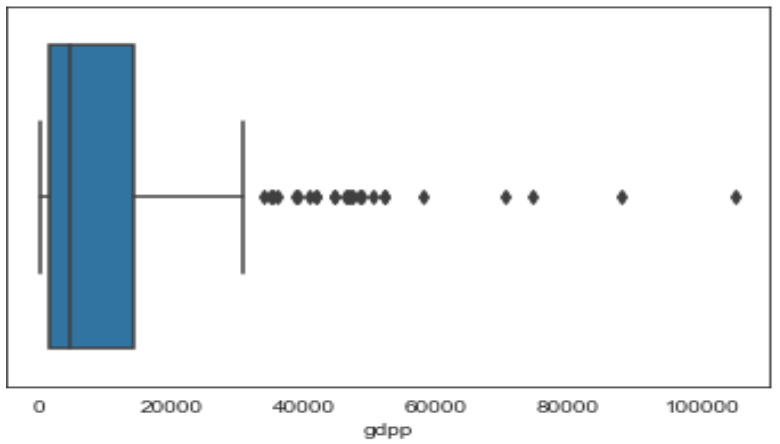

0
20000
40000
60000
80000
100000
120000
140000
imports


0
20
40
60
80
100
inflation


0
20000
40000
60000
80000
100000
120000
income


30
40
50
60
70
80
life_expec

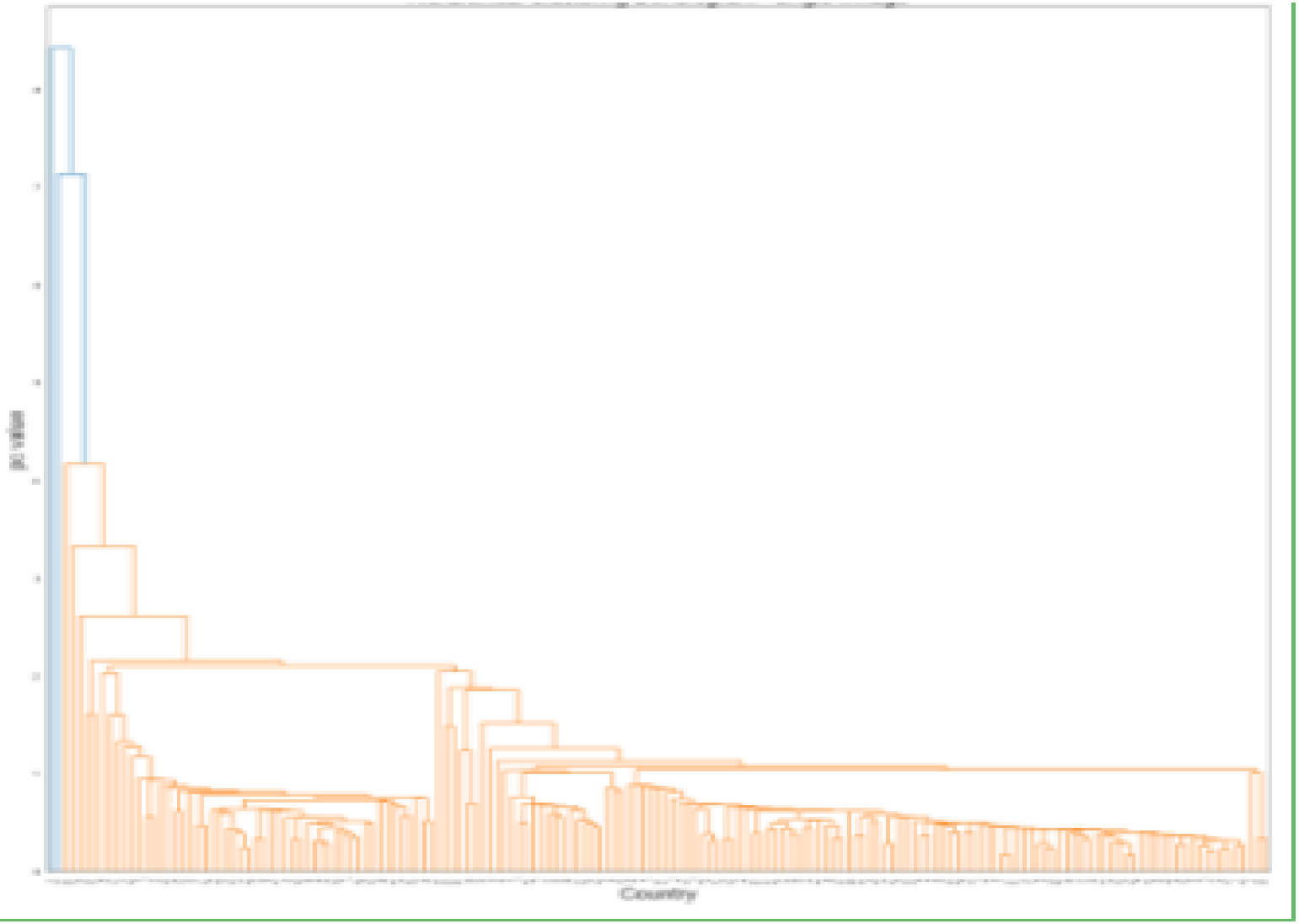

1
2
3
4
5
6
7
total_fer


0
20000
40000
60000
80000
100000
gdpp

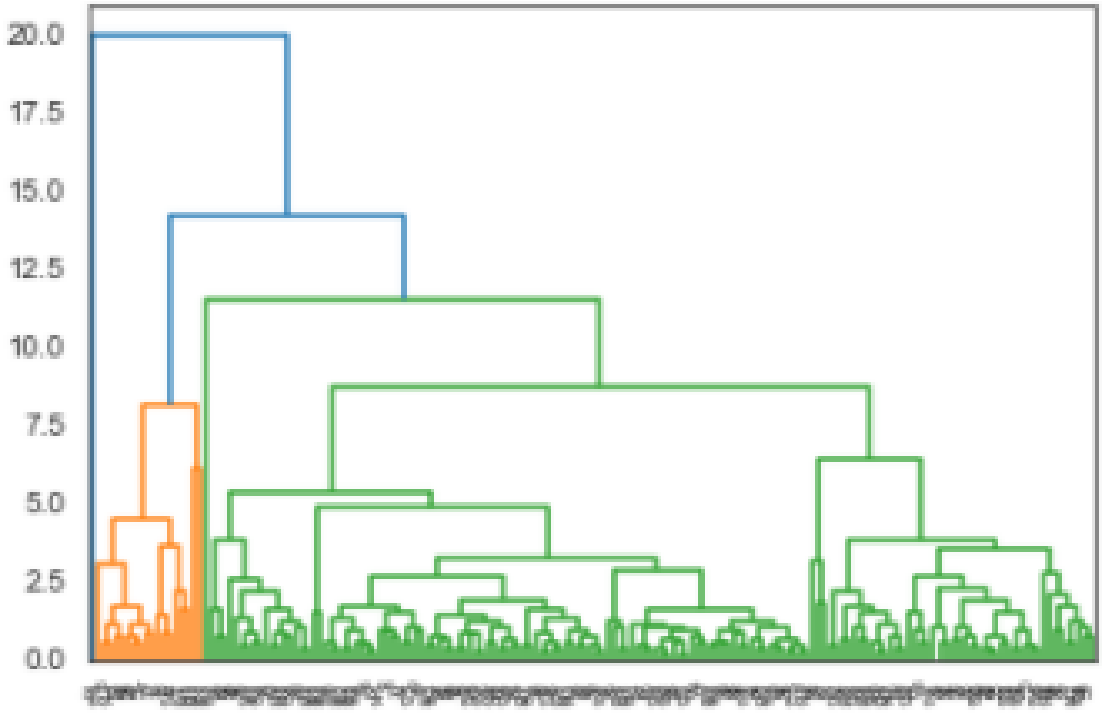
K-Means analysis

- We are trying to find the optimum value of the k-value based on the business requirements.
- So, to achieve this we used silhouette analysis to find the score of range of cluster values.
- We found below silhouette scores:
- [2, 0.4586330603547626 4]
- [3, 0.4218615812599681]
- [4, 0.4267335739770451 4]
- [5, 0.4307751339677046 7]
- [6, 0.3360998381358960 6]
- [7, 0.3094054909508284]
- [8, 0.3220272610230326]
- [9, 0.3093287457421382 6]
- [10, 0.2979658104477083]]

Hierarchical clustering
Single (Linkage)



Hierarchical clustering complete(Linkage)



CONCLUSIONS

- Among the two-conclusion drawn as per the business requirements i.e. the countries which are having low socio-economic and health factors.

- The final list of 47 countries name needs to focus on the most are mentioned below :
- Afghanistan, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, C, Rwanda, Senegal, Sierra Leone, South Africa, Sudan, Tanzania, Timor-Leste, Togo, Uganda, Yemen and Zambiahad, Comoros, Congo, Dem. Rep., Congo, Rep., Cote d'Ivoire, Equatorial Guinea, Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, Iraq, Kenya, Kiribati, Lao, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Pakistan

Thank You