

**upGrad**

---

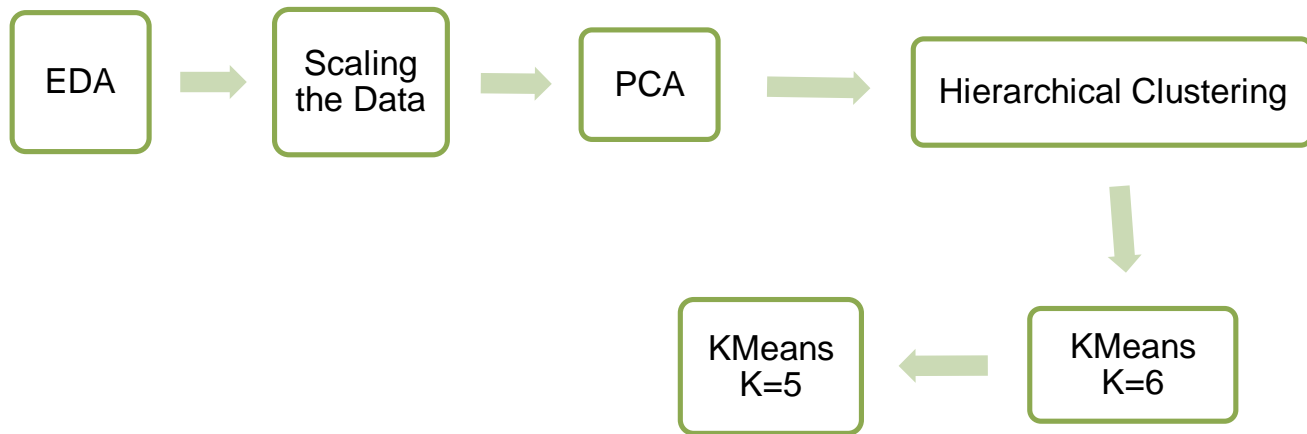
Online Power Learning

# Assignment

Clustering & PCA

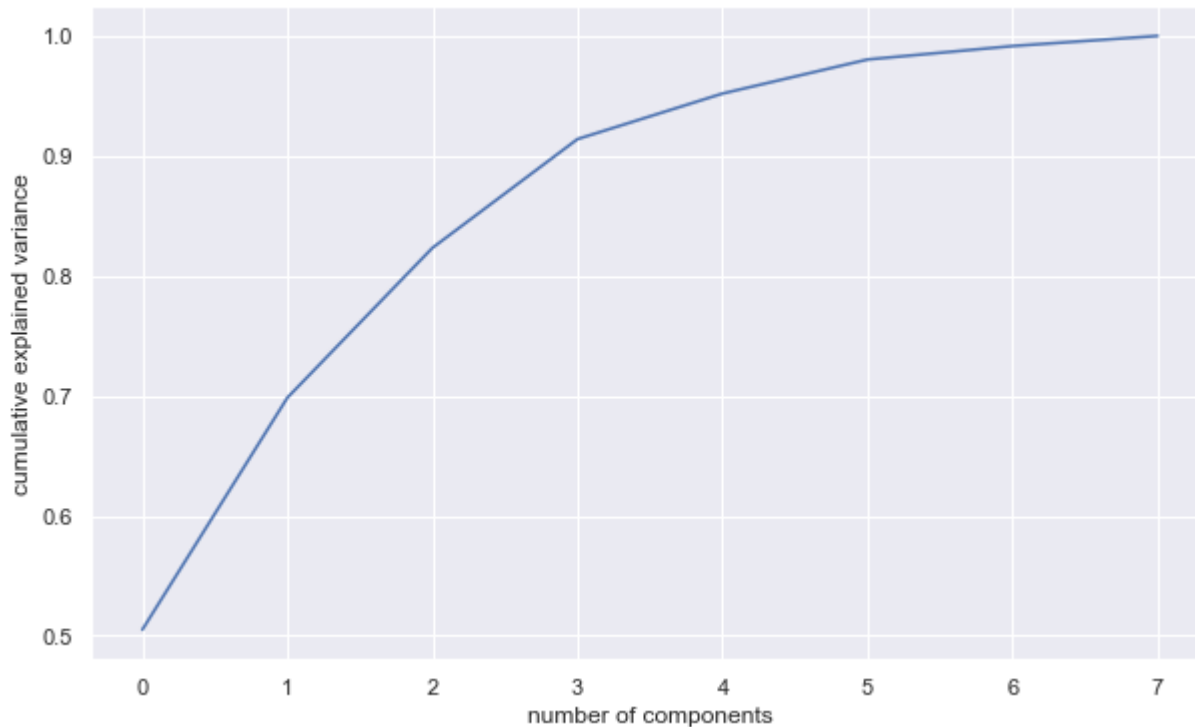
HELP International needs to decide how to utilise their funds. The main decision revolves which countries needs the help most.

We have a dataset with some health and economic parameters for 167 countries.



# SCREEPLOT

This plot clearly shows that 95% of variation captures by 5 (4<sup>th</sup> marker in plot) components.



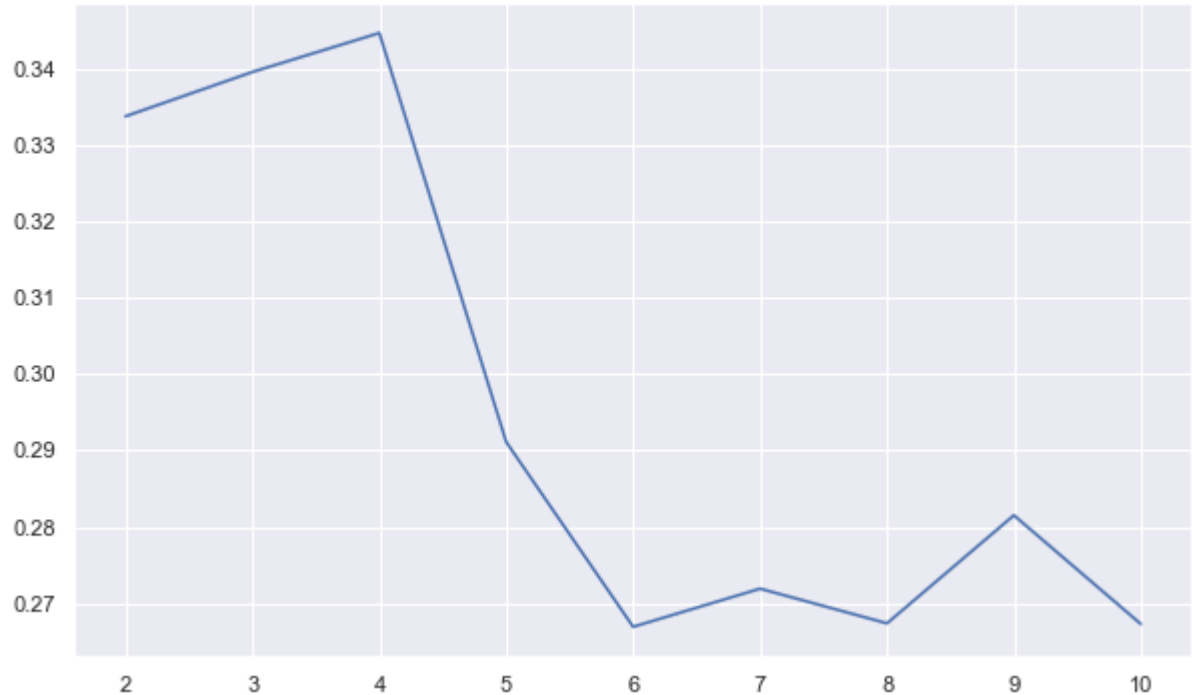
# Correlation Plot

This plot tells us that correlation among the principal components is very insignificant which is very good for modelling.



# Silhouette Score

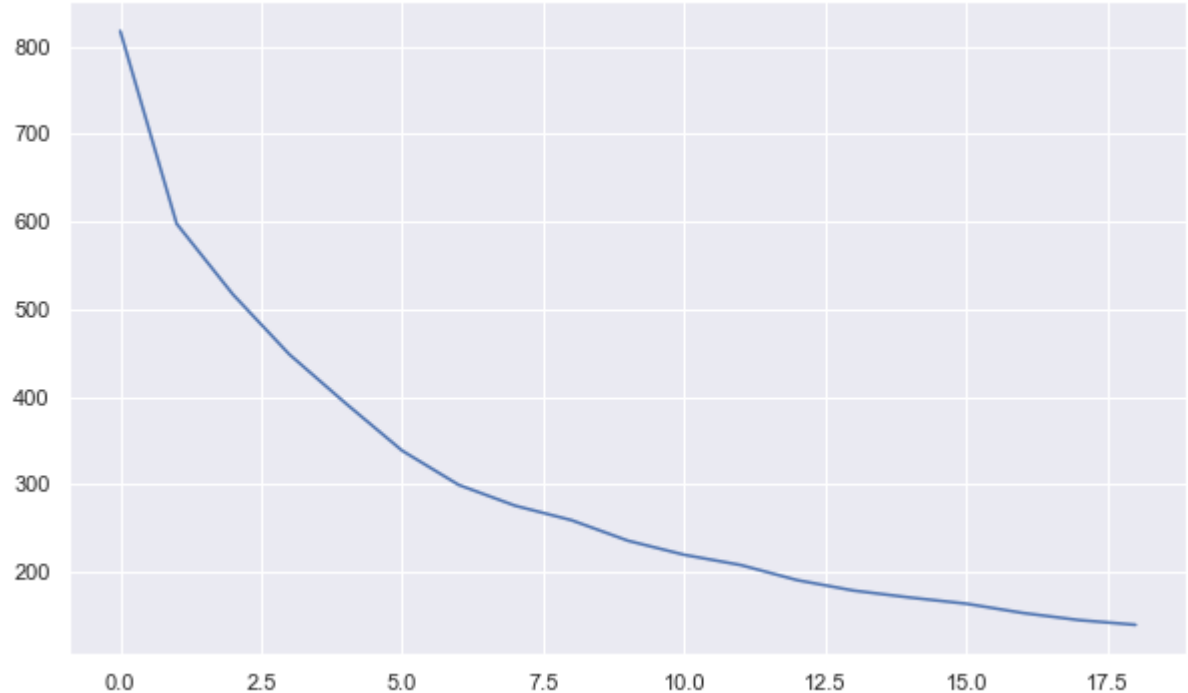
This suggests us that an optimum number of clusters would be 6.



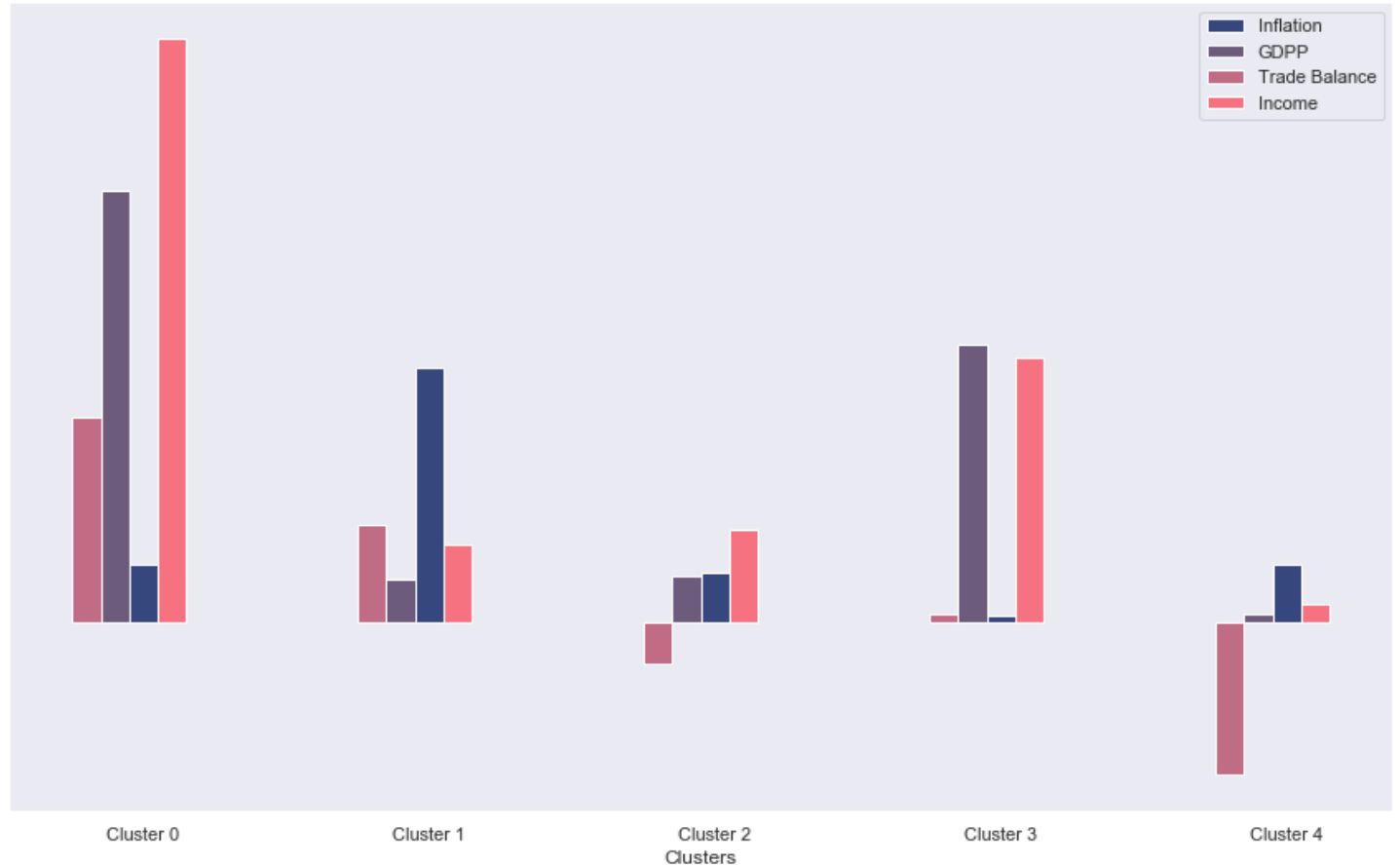
# Elbow Curve

This suggests us that an optimum number of clusters would be between 5 and 7.5.

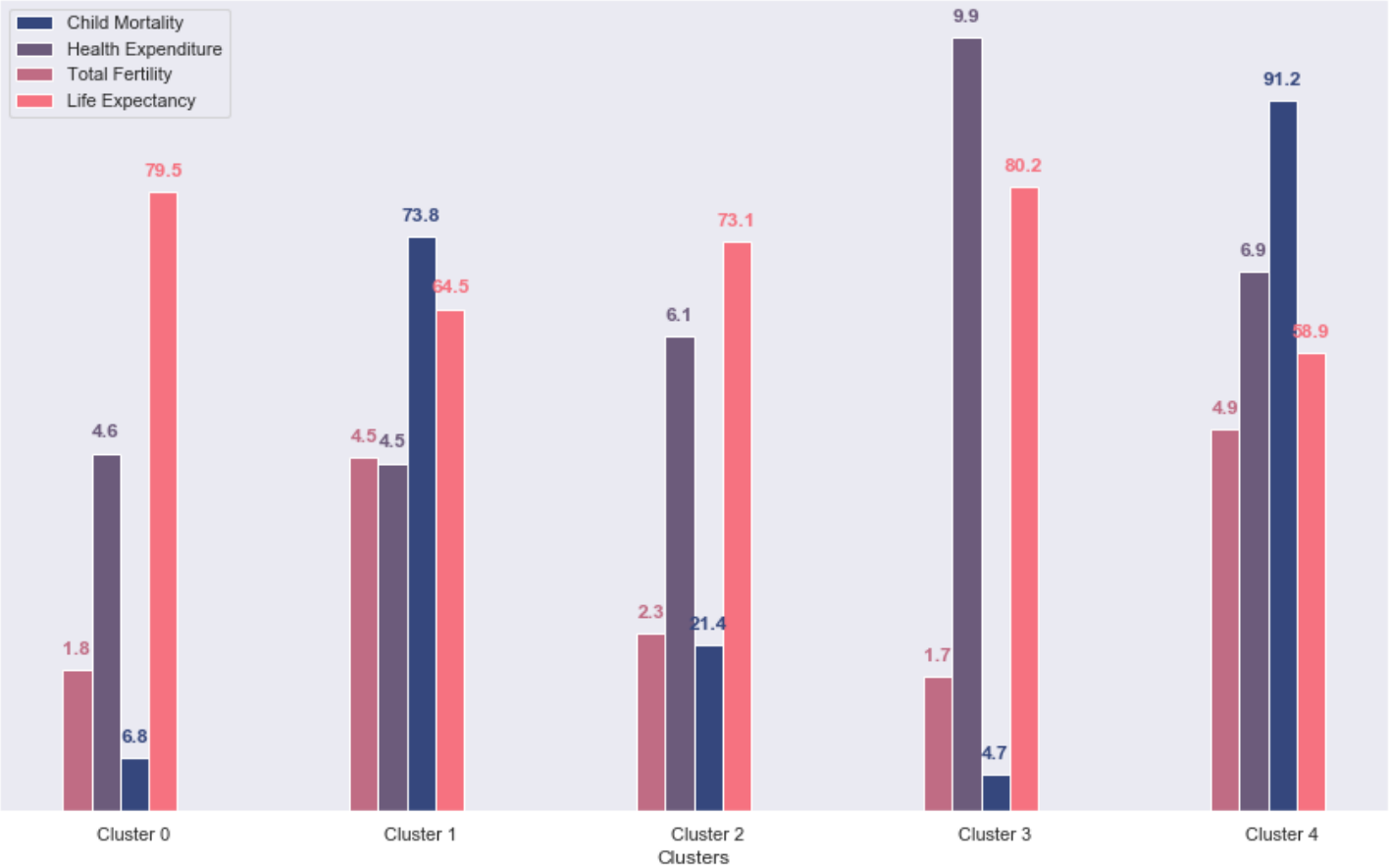
Along with the previous analysis we would go with  $K=6$ .



The final clusters showing a comparison of **economic** factors in a grouped bar chart.



The final clusters showing a comparison of **health** factors in a grouped bar chart.





## Explanation for PCA and Clustering Results

**Cluster 0** - High Per capita income, low inflation, highest gdpp and highest trade balance. - Economically Very Well  
Second highest life expectancy, Health expenditure is fine, low child mortality, total fertility is low - Good Health wise

**Cluster 1** - Cluster 1 has high inflation and other factors are moderate - Economically not so well.  
Second highest child mortality and second lowest life expectancy, other factors are also moderate - Health wise needs attention.

**Cluster 2** - Second worst scenario economically but cluster 4 needs more help than this one. Health wise also it is second worse scenario.

**Cluster 3** - This cluster is second best economically and best health wise. Surely this group of countries does not need the help.

**Cluster 4** - This the cluster that needs CEO's attention. These group of countries have worst trade balance, gdpp and per capita income is lowest, has highest child mortality, lowest life expectancy. Other Factors are moderate like inflation, health expenditure.