

# Clustering Assignment

-Tanvi Joshi

# Issue at hand

- ▶ The HELP International Humanitarian NGO has been able to raise \$10 million to help countries who are in dire need of aid.
- ▶ Data needs to be analysed based on socio-economic and health factors
- ▶ We need to find out which countries are in need and direct the funds for their aid.

# Factors on which the countries are analysed

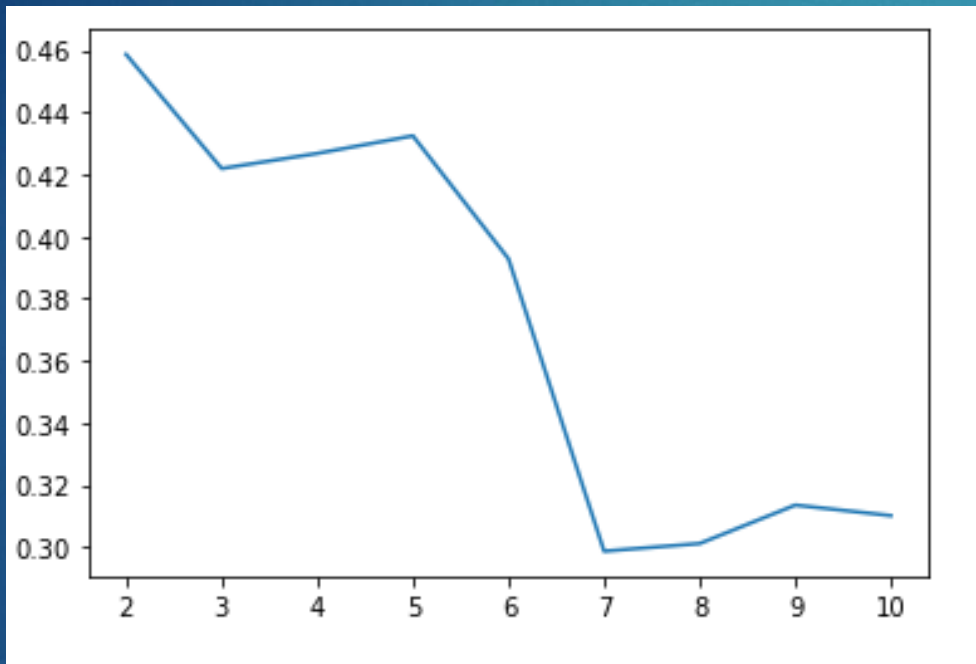
- ▶ 1) Inflation
- ▶ 2) Child mortality rate
- ▶ 3) Exports
- ▶ 4) Imports
- ▶ 5) GDP per capita
- ▶ 6) Income
- ▶ 7) Total Fertility
- ▶ 8) Life expectancy

# Is the data good for clustering?

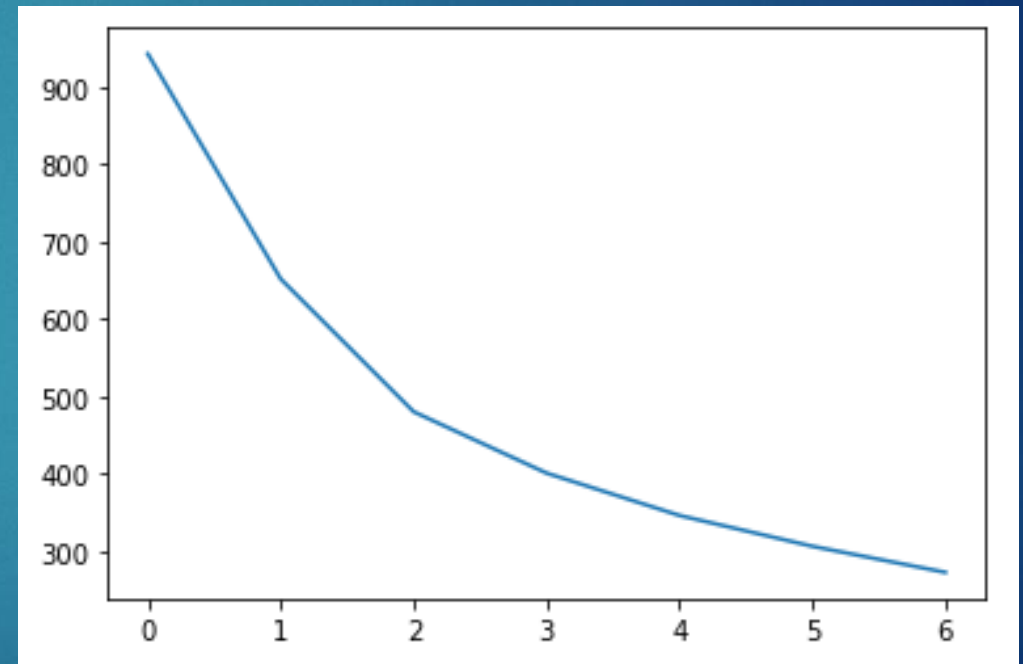
- ▶ Yes ! As the Hopkins score is greater than 80

# How many clusters to be formed?

The silhouette score suggests 3

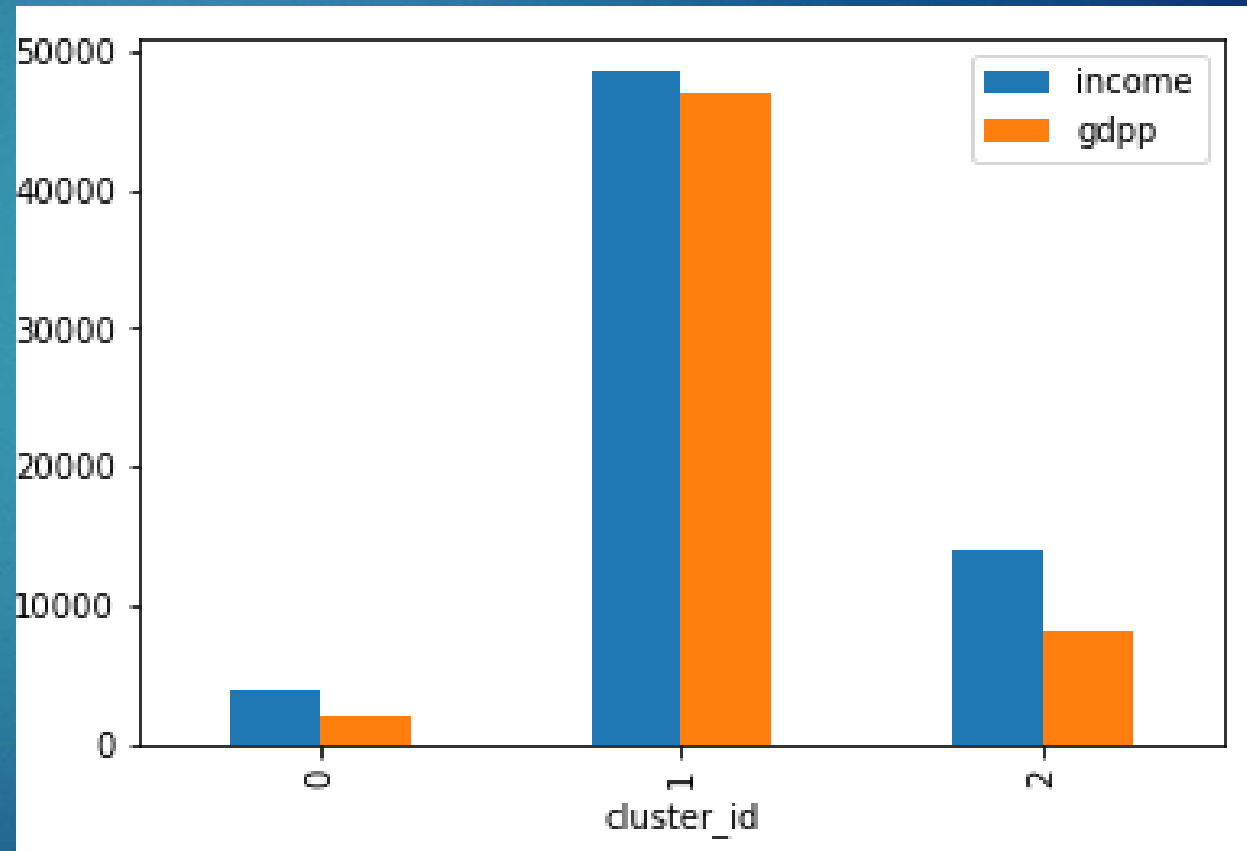


The elbow curve suggests 3



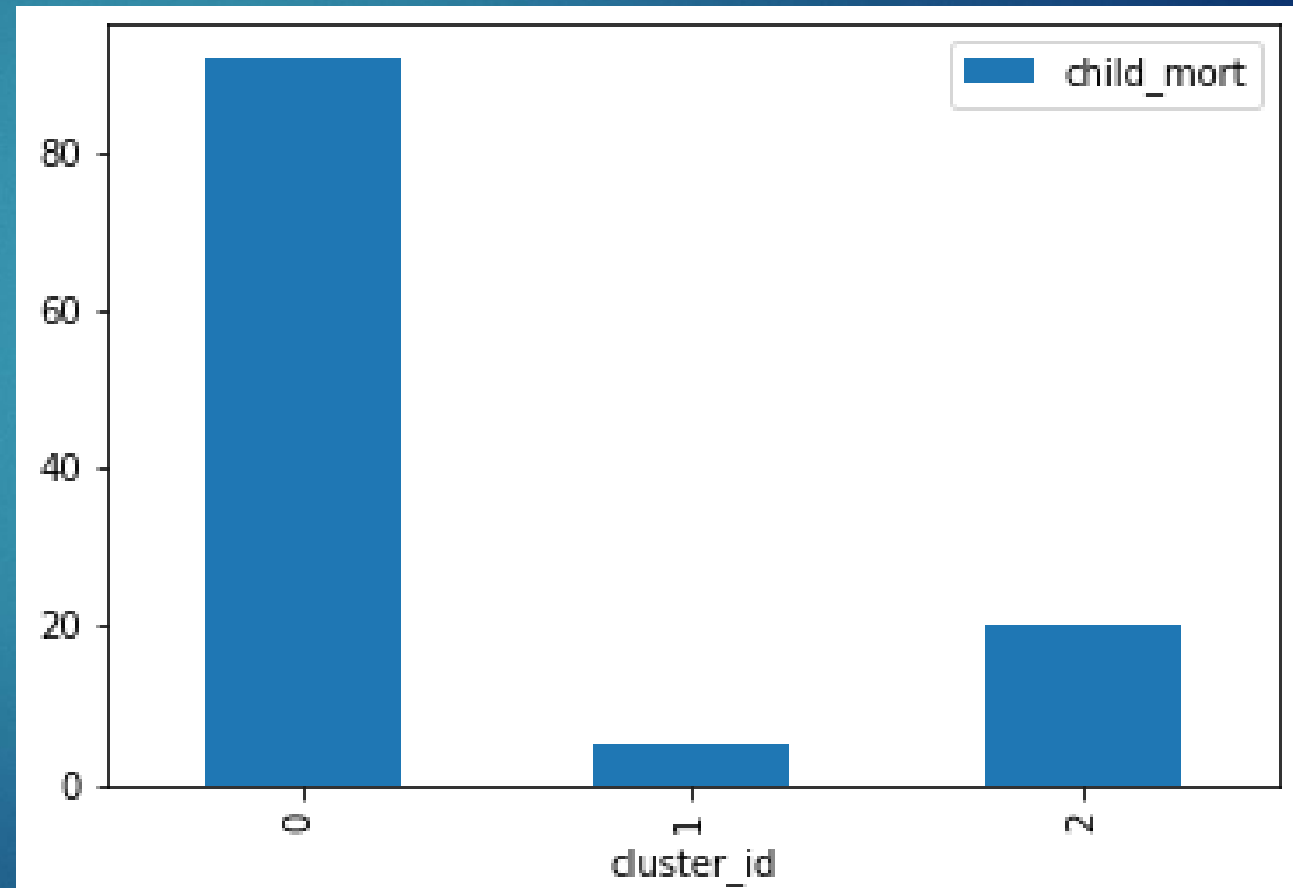
# K – means clustering performed

- This cluster profiling shows that cluster 0 has lowest income and lowest GDP



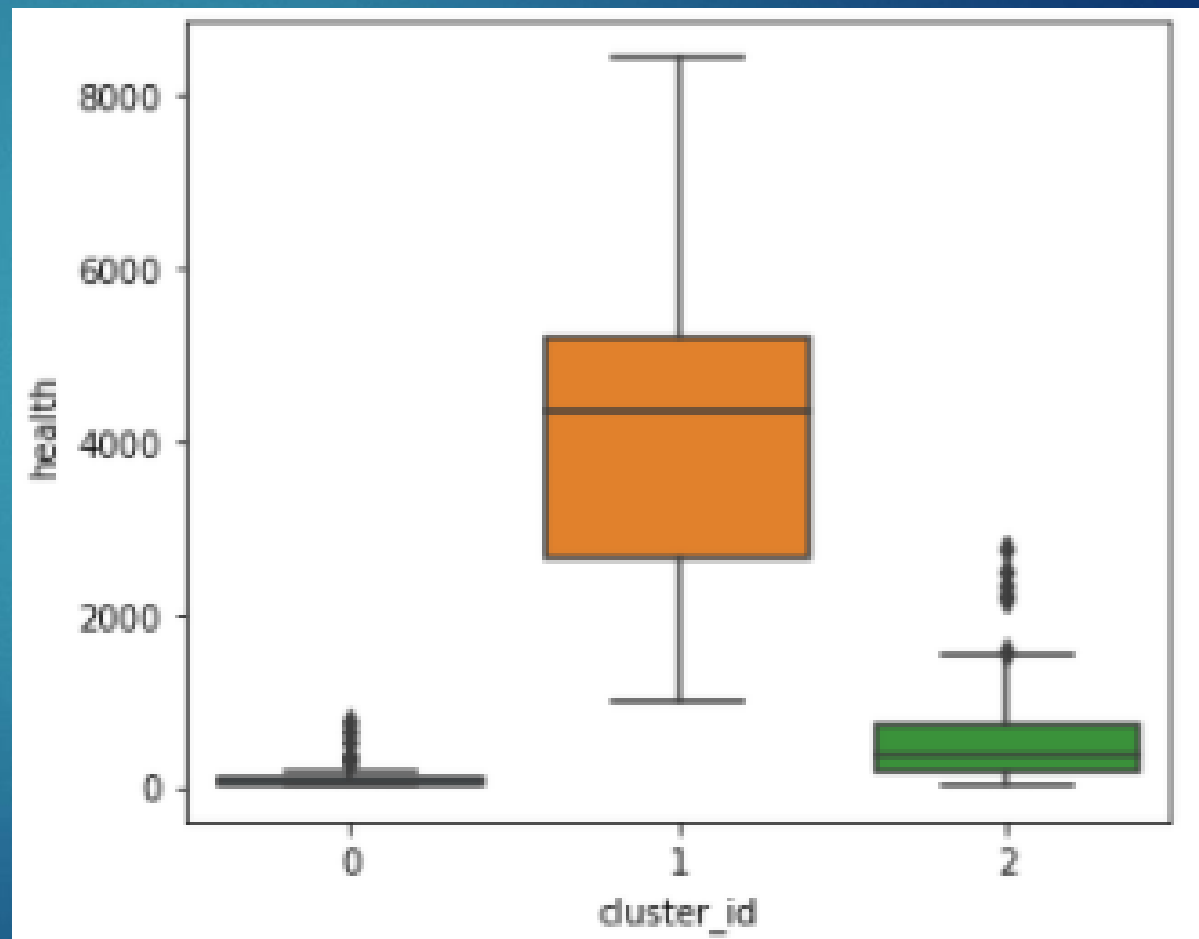


- This Cluster profile shows that Cluster 0 has highest Child mortality rate



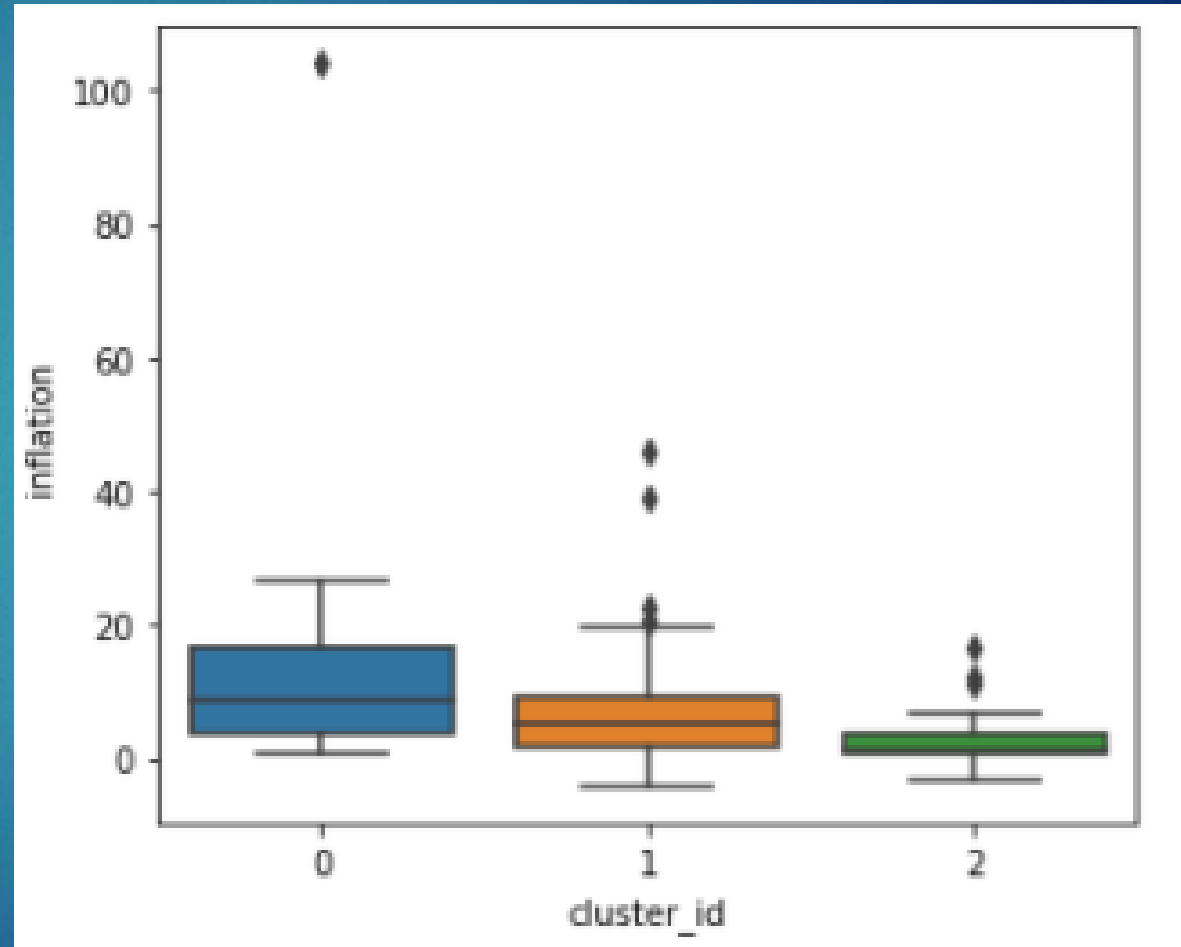
# Comparing the other metrics

- ▶ Cluster 0 shows lowest on health score

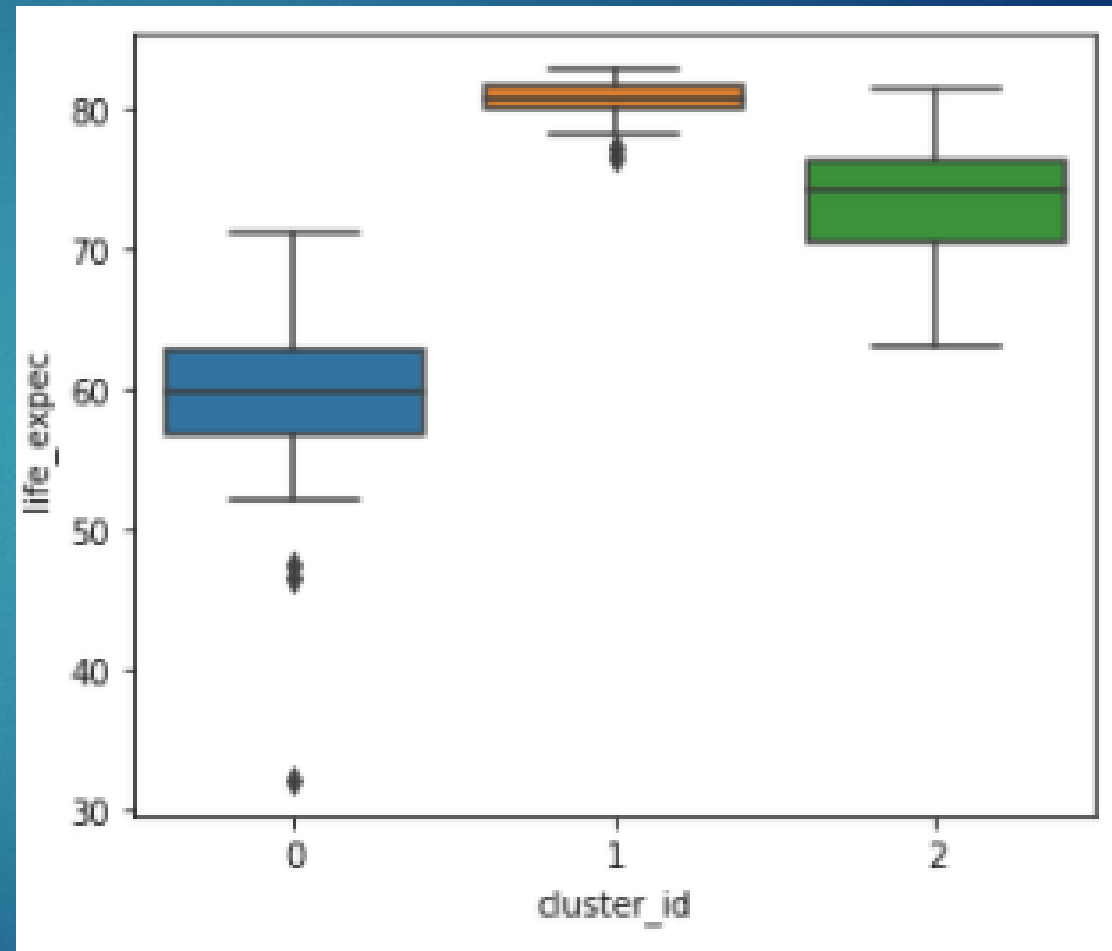




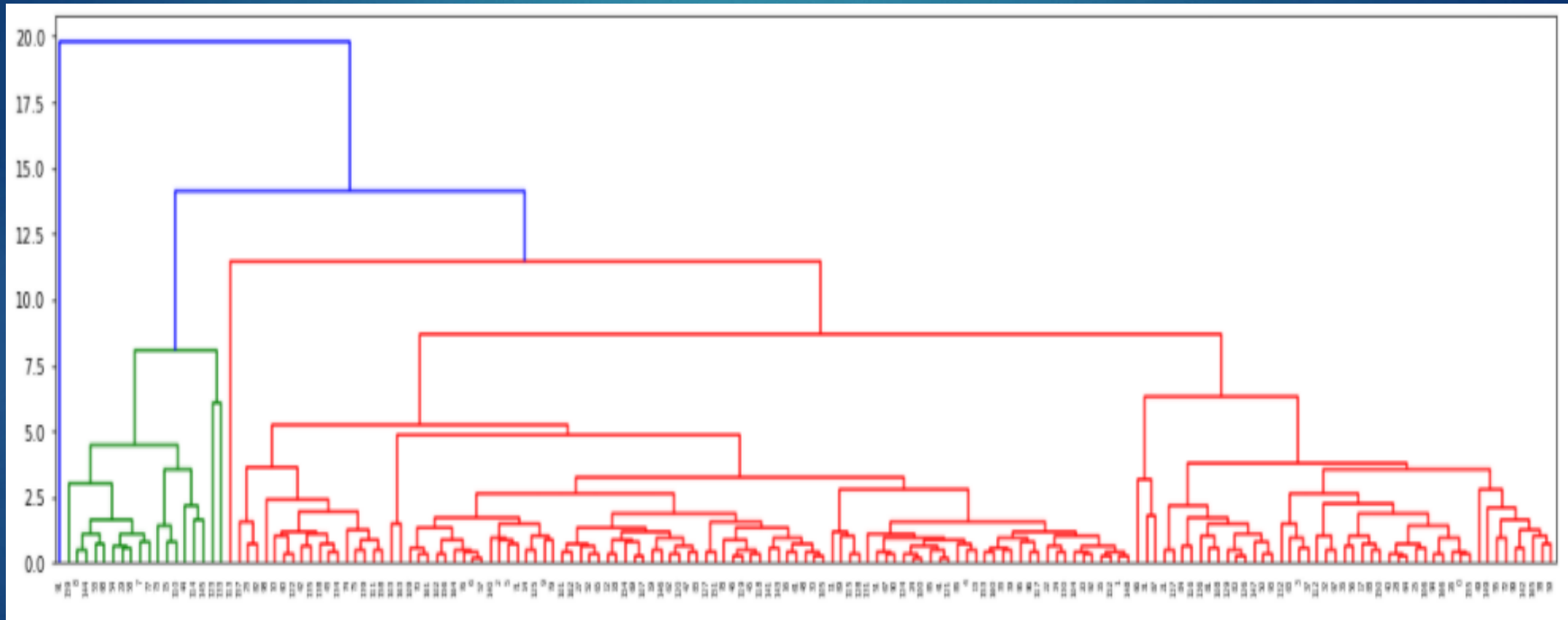
- ▶ Cluster 0 shows high inflation



- ▶ Cluster 0 shows least life expectancy



100



# Final Factors

► We have arrived at a final list based on the following factors

- 1) Highest child mortality rate
- 2) Lowest Income
- 3) Lowest GDP

# Final List of countries

- ▶ 1) Haiti
- ▶ 2) Sierra Leone
- ▶ 3) Chad
- ▶ 4) Central African Republic
- ▶ 5) Mali

# Recommendations

- ▶ The data can be further analysed and checked for other factors
- ▶ We can arrive at a different list based on the factors that are more important to be addressed by the NGO



Thank you