HELP International

A STUDY TO FIND THE COUNTRIES IN DIREST NEED OF AID

-BHAVINI D. GADHIWALA

Intent of the study

Help International, the NGO has raised $10 million and wishes to donate this money strategically and effectively

The study intends to identify from the list of countries in the direst need for aid

The socio-economic and health factors or the countries will be observed based on given dataset

The overall development of the country will also be taken into account

Data Overview

Total countries: **167**

Parameters available for study are as follows:

- Gross Domestic Product (GDPP)
- Child Mortality
- Exports per capita
- Imports per capita
- Total health spending per capita
- Net Income per person
- Inflation (Annual growth rate)
- Total Fertility
- Life Expectancy

Observations from EDA

- Heavy skewness found in the overall data
- Outliers found in most parameters towards higher side except life expectancy
- We treated the outliers with the intent of study in mind
- We capped the lower child mortality, inflation and total fertility values because the higher values were essential for the study
- Among the other parameters, we capped the high values

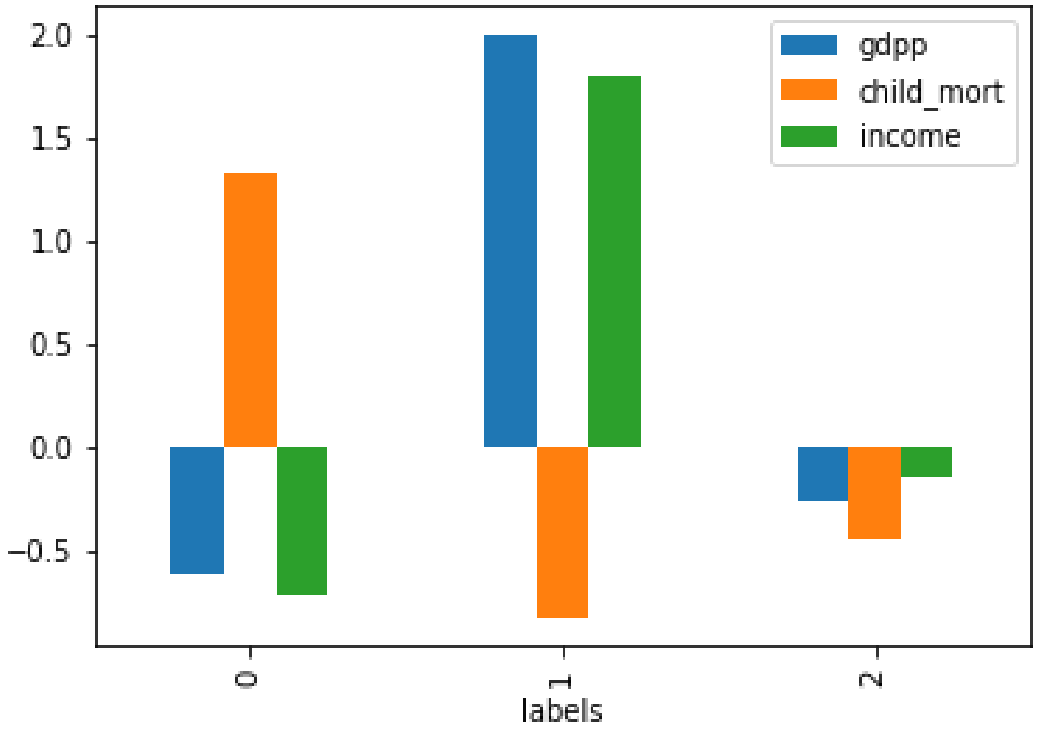
Clustering Models

- ✓ Hopkins Statistic performed on dataset to check the clustering tendency
- ✓ Data found highly favorable for clustering with over 90% Hopkins score
- ✓ Data standardization done for modelling based on Standard Scaler
- ✓ Optimum number of clusters identified based on dual observation of Silhouette Score and SSD – Elbow curve analysis
- ✓ Clustering done using K means clustering model
- ✓ Clustering verified using Hierarchical clustering model
- ✓ Data outcome of both models verified

Observations from Clustering Results

- ✓The cluster 0 has high child mortality and low GDPP and low income.
- ✓The cluster 1 has high GDPP and high income and low child mortality
- ✓The cluster 2 has low GDPP and low income and low child mortality

Our business requirements suggest that the cluster 0 is the most appropriate cluster that is in the dire need of the aids based on its properties



Please note the data in the graph is scaled for modeling purpose

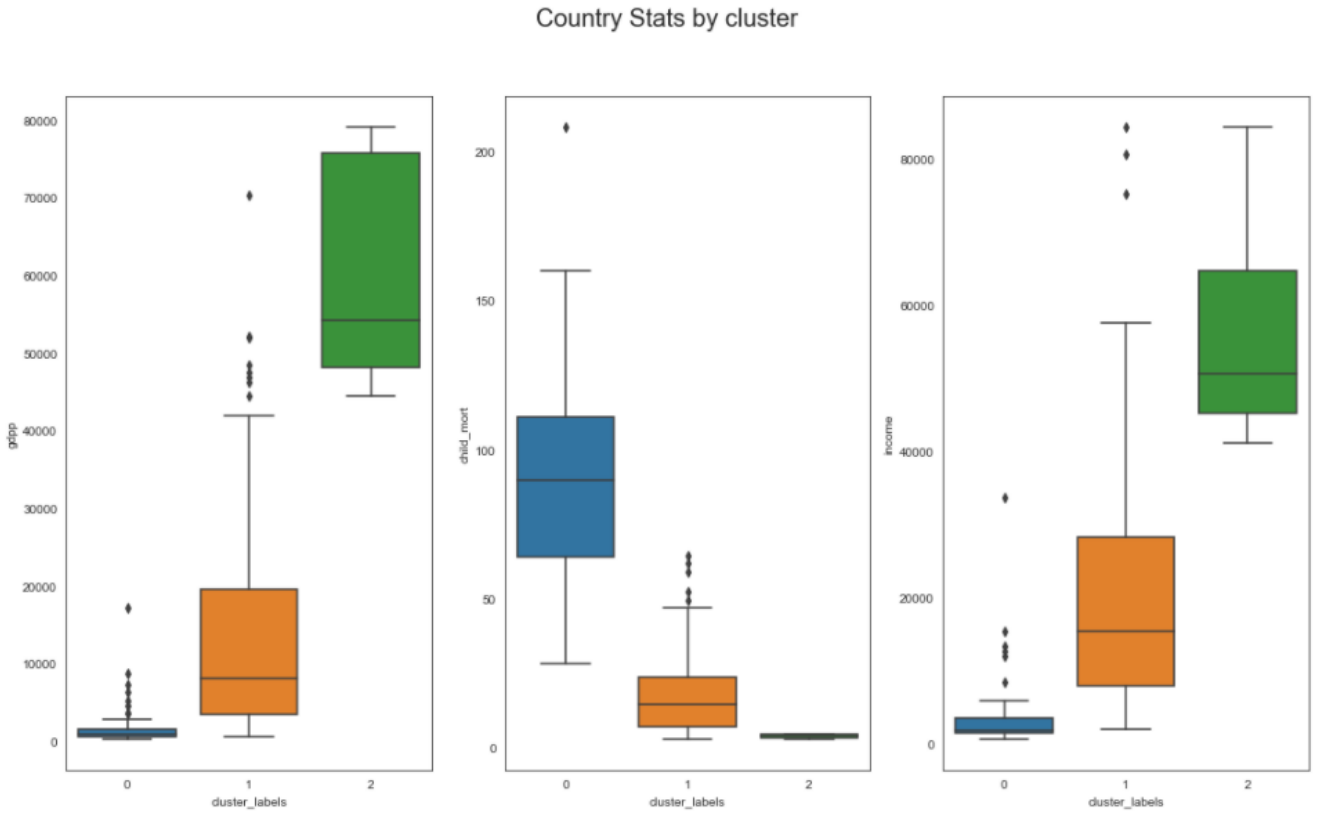
Summary

Both the models resulted in the same country distribution in the cluster

The cluster of our interest was the one with high child mortality, low GDPP and low income

This cluster had 48 countries

Post clustering we sorted the countries based on GDPP, child mortality and income in ascending, descending and ascending order respectively to filter out top 10 countries



TOP 10 countries

1. Burundi
2. Liberia
3. Congo Dem. Rep.
4. Niger
5. Sierra Leone
6. Madagascar
7. Mozambique
8. Central African Republic
9. Malawi
10. Eritrea

Thank You!