

HELP –INTERNATIONAL – CLUSTERING

SUBMITTED BY

DHARANI ARUMUGAM

PROBLEM STATEMENT :

- HELP International is an international humanitarian NGO that is committed to fighting poverty and providing the people of backward countries with basic amenities and relief during the time of disasters and natural calamities. It runs a lot of operational projects from time to time along with advocacy drives to raise awareness as well as for funding purposes.

Objective :

- To categorise the countries using some socio-economic and health factors that determine the overall development of the country. Then you suggest the countries which the CEO needs to focus on the most.



APPROACH - I:

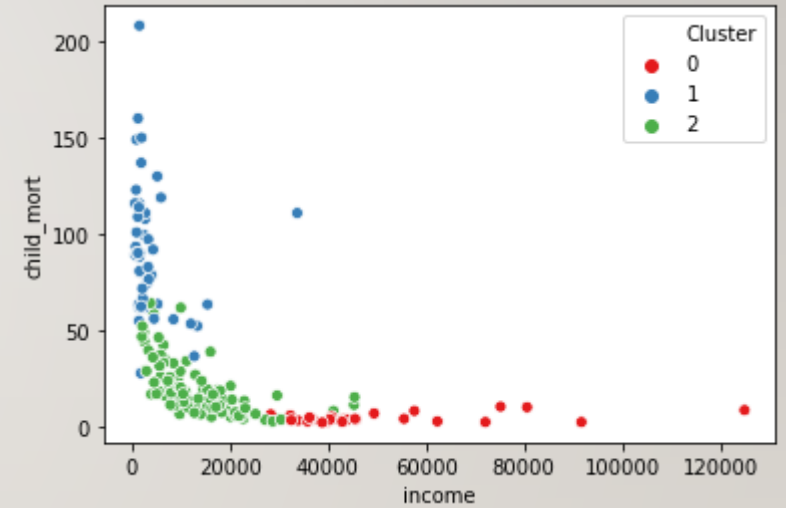
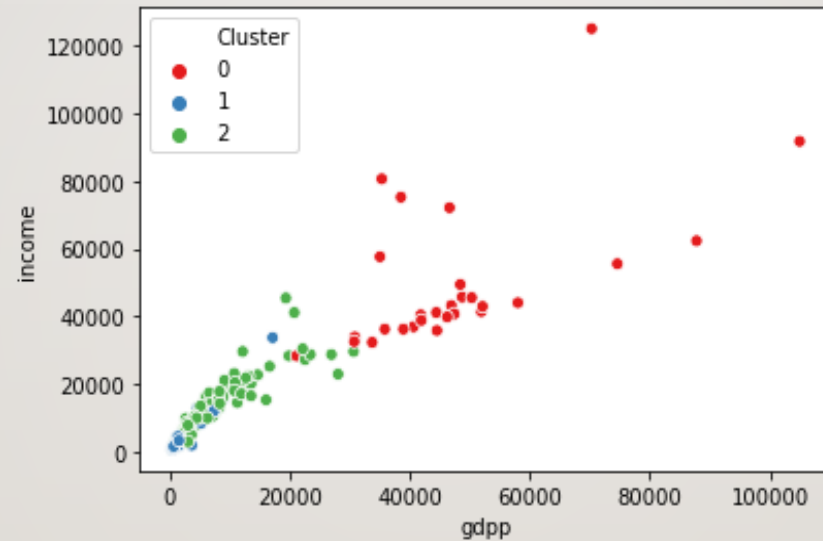
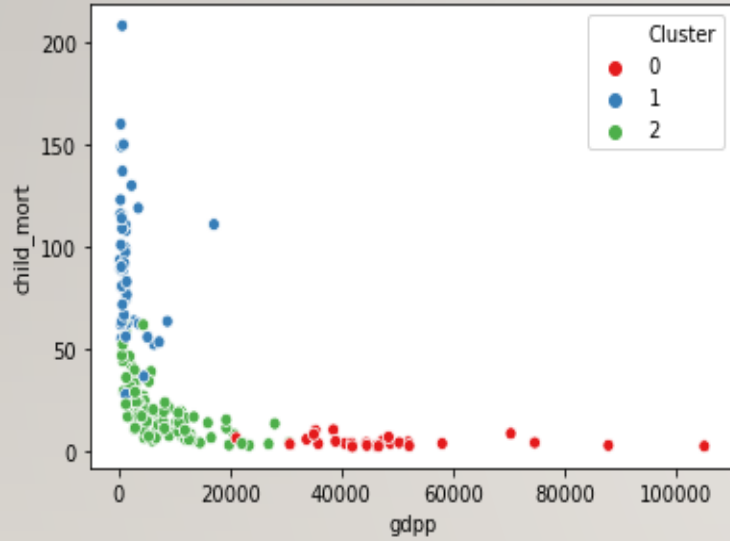
K-Means Clustering :

- Perform EDA and Outlier Analysis
- Hopkins Statistics check
- K-Means Algorithm to find optimal no of clusters
- Run K-Means with the chosen cluster
- Visualizing the cluster
- Cluster Profiling using child_mort, income, gdpp
- Finalize the top 5 countries that are in need

CLUSTER FORMED

- On Performing the Analysis – Chose optimal no of clusters as 3:
 - Cluster 0 => Developed Countries => Countries who has high income, high gdpp , low child_mort
 - Cluster 1 => Under -Developed Countries => Countries who has low income, low gdpp, high child_mort
 - Cluster 2 => Developing countries => Countries who has moderate income, moderate gdpp, medium child_mort

VISUALIZING THE CLUSTER FORMED



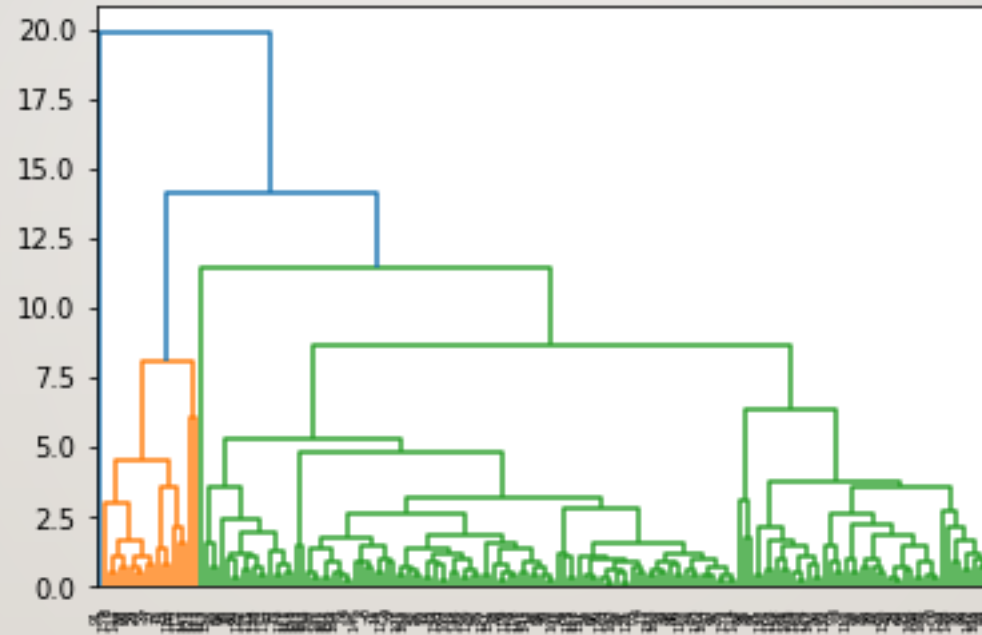
COUNTRIES WHICH I CHOOSE AS OF K-MEANS CLUSTERING ARE :

- Haiti
- Sierra Leone
- Chad
- Central African Republic
- Mali

APPROACH -2

- Hierarchical Clustering:
 - Perform scaling on the data
 - Create single linkage and complete linkage for data
 - Check for dendrogram to decide optimal cluster
 - Visualizing the cluster
 - Cluster Profiling using child_mort, income,gdpp
 - Finalize the top 5 countries that are in need

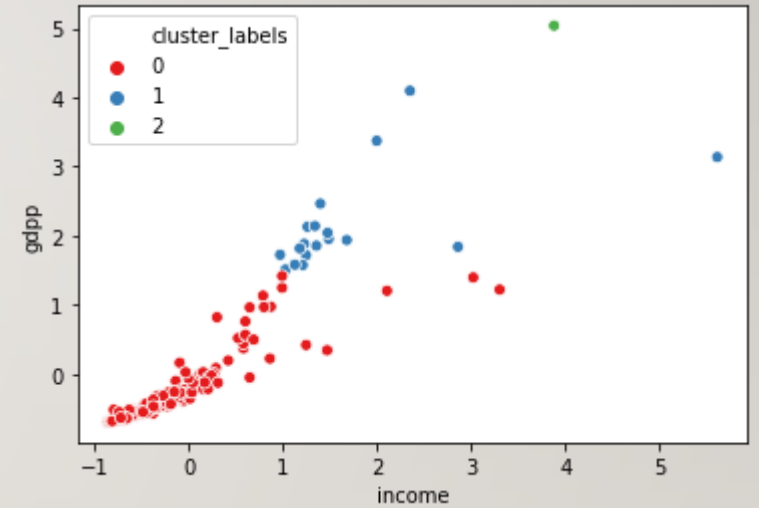
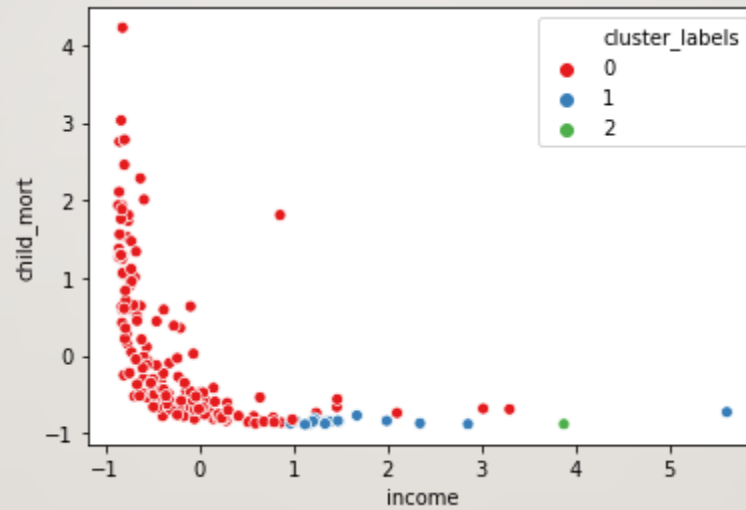
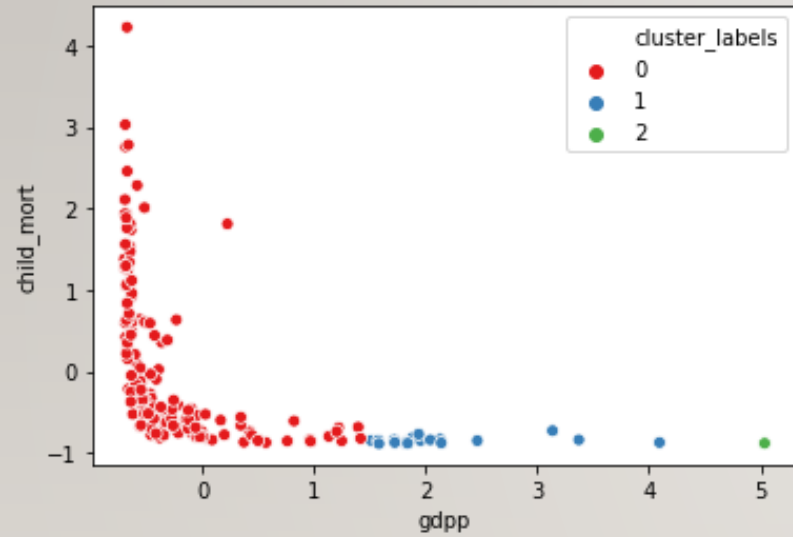
DENDROGRAM



CLUSTERS FORMED :

- Cluster 2 => Developed Countries => Countries who has high income, high gdpp , low child_mort
- Cluster 1 => Developing countries => Countries who has moderate income, moderate gdpp, medium child_mort
- Cluster 0 => Under -Developed Countries => Countries who has low income, low gdpp, high child_mort

VISUALIZING THE CLUSTERS:



COUNTRIES AS OF HIERARCHICAL CLUSTERING:

- - Haiti
- - Sierra Leone
- - Chad
- - Central African Republic
- - Mali