Clustering for HELP International

—

By Madhurima Deekonda

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.

They have been able to raise around $ 10 million. Now the CEO of the NGO needs to decide how to use this money strategically and effectively. The significant issues that come while making this decision are mostly related to choosing the countries that are in the direst need of aid.

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

The Approach

We used Clustering to identify the countries which are in need of aid.

Clustering allows us to categorise the countries using some socio-economic and health factors such as Income, Child Mortality Rate and GDPP.

Steps Involved

- Data Understanding

- Data Preparation

- Outlier Analysis

- Modelling

- Clustering - K Means and Hierarchical.

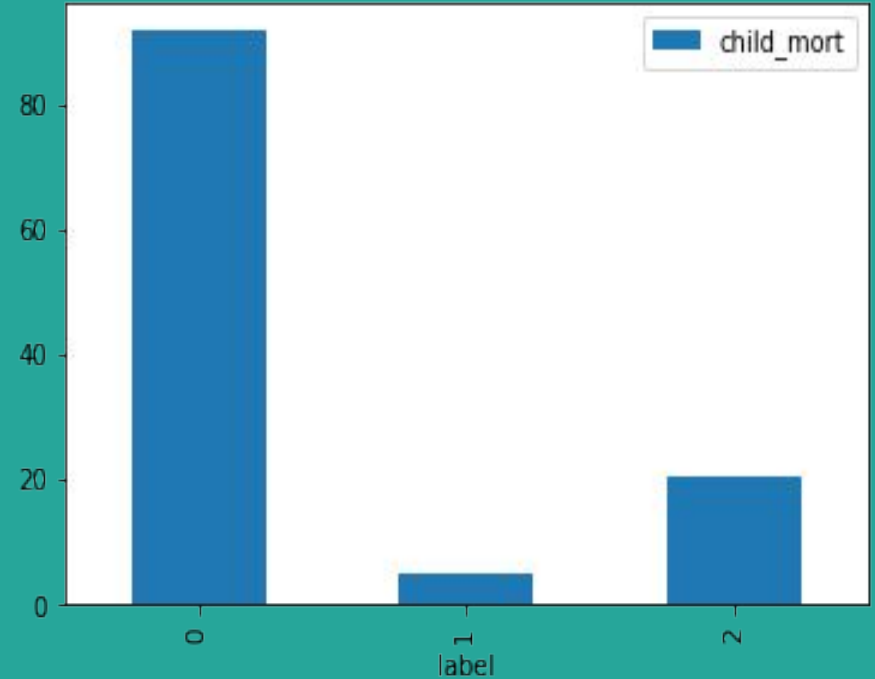- Cluster Visualisation
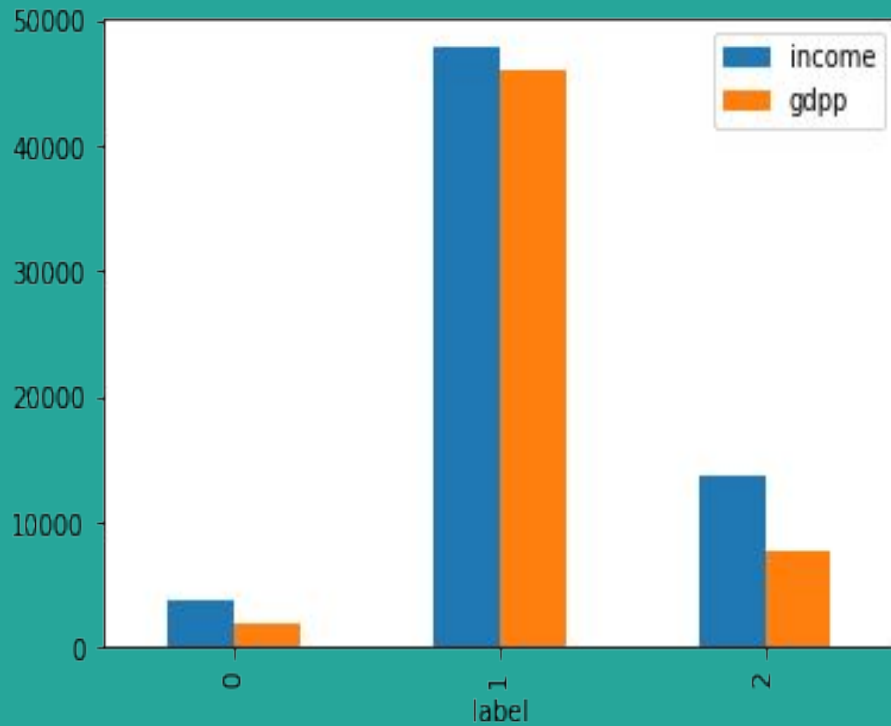
- Country Identification

K-Means Clustering

K Means Clustering is a method that aims to partition 'n' observations of a data into 'k' clusters.

Each cluster belongs to the nearest cluster centroid.

The countries in the given dataset have been divided into 3 clusters, based on which we have attained the list of countries which required the most attention from HELP International.

The Bar Plots above show that the countries in cluster labelled 0 have the lowest Income, lowest GDPP and the highest Child Mortality rate.
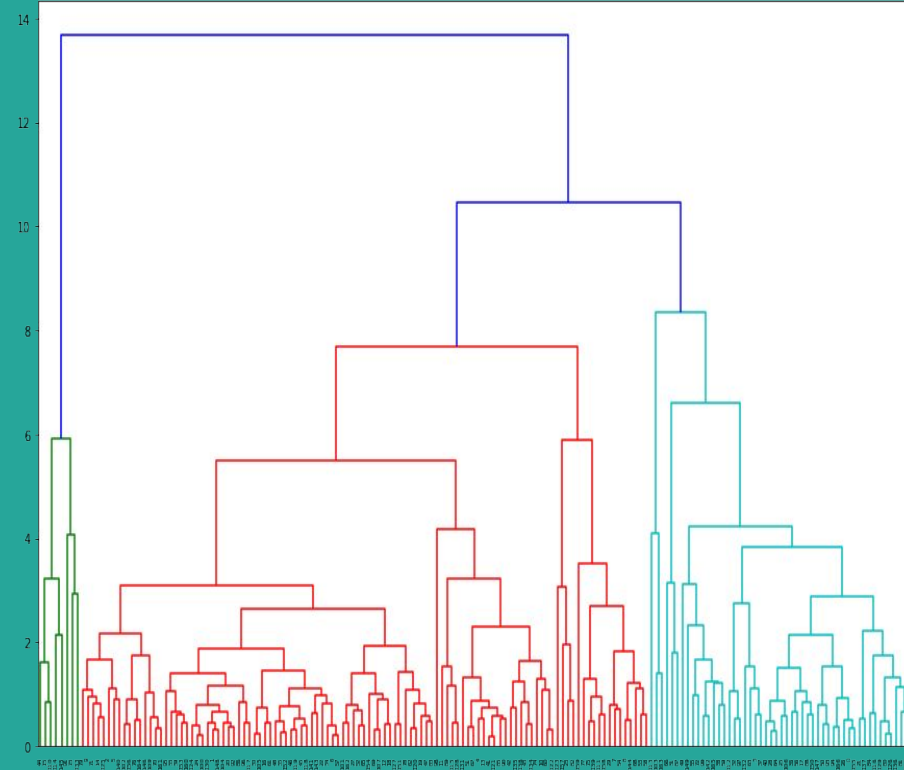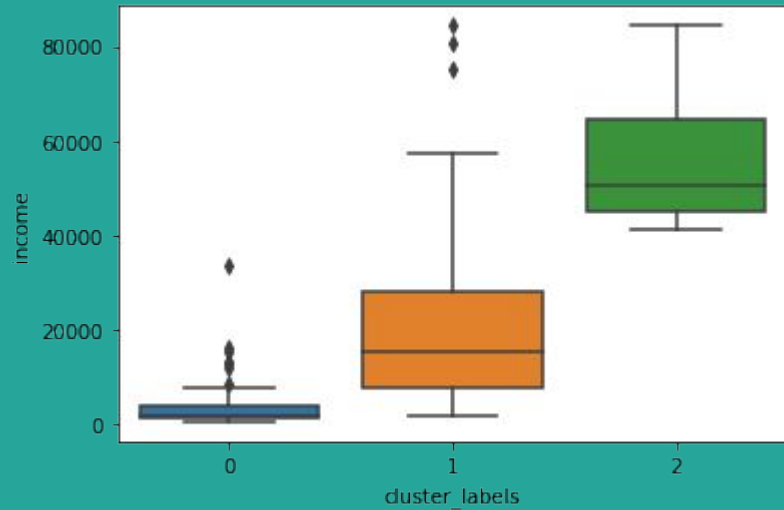
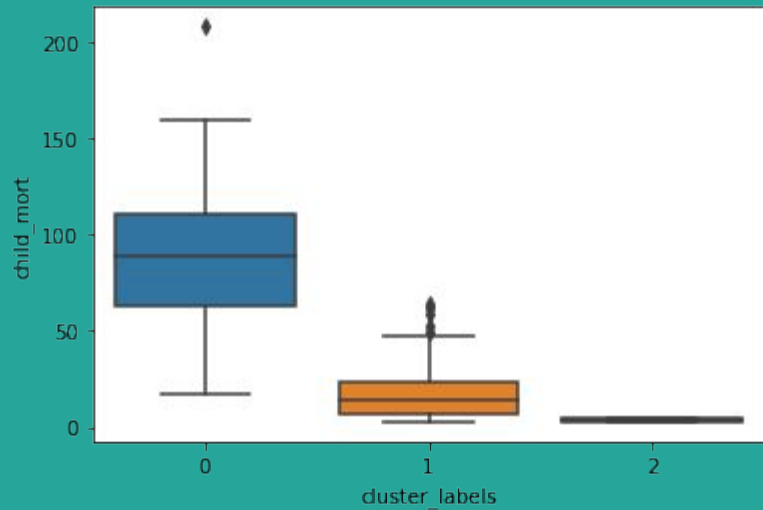# Hierarchical Clustering

Hierarchical Clustering is another algorithm that groups similar objects into clusters.

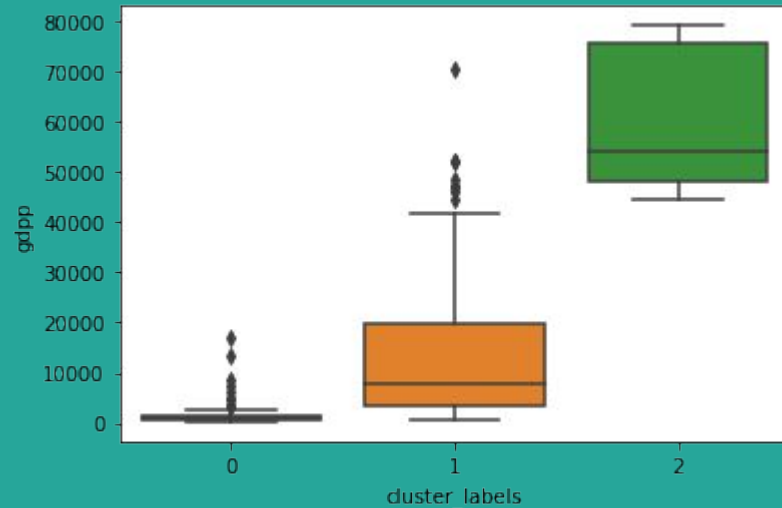The endpoint is a set of clusters, where each cluster is distinct from each other, and the objects within each cluster are broadly similar to each other.

Based on the Hierarchical Clustering of the given data, as seen in the image, the countries have been partitioned into 3 Clusters.

The box plots here show that the countries in cluster labelled as '0' have the lowest Income, lowest GDPP and the highest Child Mortality rate.

Final List of Countries

country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	cluster_labels
Congo, Dem. Rep.	116.0	137.2740	26.4194	165.664	609.0	20.800	57.5	6.54	334.0	0
Liberia	89.3	62.4570	38.5860	302.802	700.0	5.470	60.8	5.02	327.0	0
Burundi	93.6	20.6052	26.7960	90.552	764.0	12.300	57.7	6.26	231.0	0
Niger	123.0	77.2560	17.9568	170.868	814.0	2.550	58.8	7.49	348.0	0
Central African Republic	149.0	52.6280	17.7508	118.190	888.0	2.010	47.5	5.21	446.0	0
Mozambique	101.0	131.9850	21.8299	193.578	918.0	7.640	54.5	5.56	419.0	0
Malawi	90.5	104.6520	30.2481	160.191	1030.0	12.100	53.1	5.31	459.0	0
Guinea	109.0	196.3440	31.9464	279.936	1190.0	16.100	58.0	5.34	648.0	0
Togo	90.3	196.1760	37.3320	279.624	1210.0	1.180	58.7	4.87	488.0	0
Sierra Leone	160.0	67.0320	52.2690	137.655	1220.0	17.200	55.0	5.20	399.0	0
Rwanda	63.6	67.5600	59.1150	168.900	1350.0	2.610	64.6	4.51	563.0	0
Guinea-Bissau	114.0	81.5030	46.4950	192.544	1390.0	2.970	55.6	5.05	547.0	0
Madagascar	62.2	103.2500	15.5701	177.590	1390.0	8.790	60.8	4.60	413.0	0
Comoros	88.2	126.8850	34.6819	397.573	1410.0	3.870	65.9	4.75	769.0	0
Eritrea	55.2	23.0878	12.8212	112.306	1420.0	11.600	61.7	4.61	482.0	0

Conclusion

After Performing the K-Means and the Hierarchical clustering, we found that there were around 50 countries that were underdeveloped and that they are still in the process of being developed.

But, according to some socio-economic and health factors, there were 5 countries that are on top of the list which are in dire need of aid, as their GDPP and Income were seen to be very low and also the child mortality rate in these countries are very high, as per the given data.

We suggest that HELP International should focus on providing their services to these countries.

The Top 5 countries are:

- Congo, Dem. Rep.
- Liberia
- Burundi
- Niger
- Central African Republic