Objective : 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

After the recent funding programmes, 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.

Analysis Approach :

Data Collection and Cleaning : Collecting the required data and cleaning the data like Null value removal, Unnecessary column removal

Visualizing The Data : Visualizing the data and analyzing how independent variables are corelated with each other.

Treating Outliers : Finding the outliers from each column , removing outliers which do not effect the problem statement

Scaling Data : standardizing all columns

Hopkin Test: This Test is use to know is this data is has tendency for clustering or not .

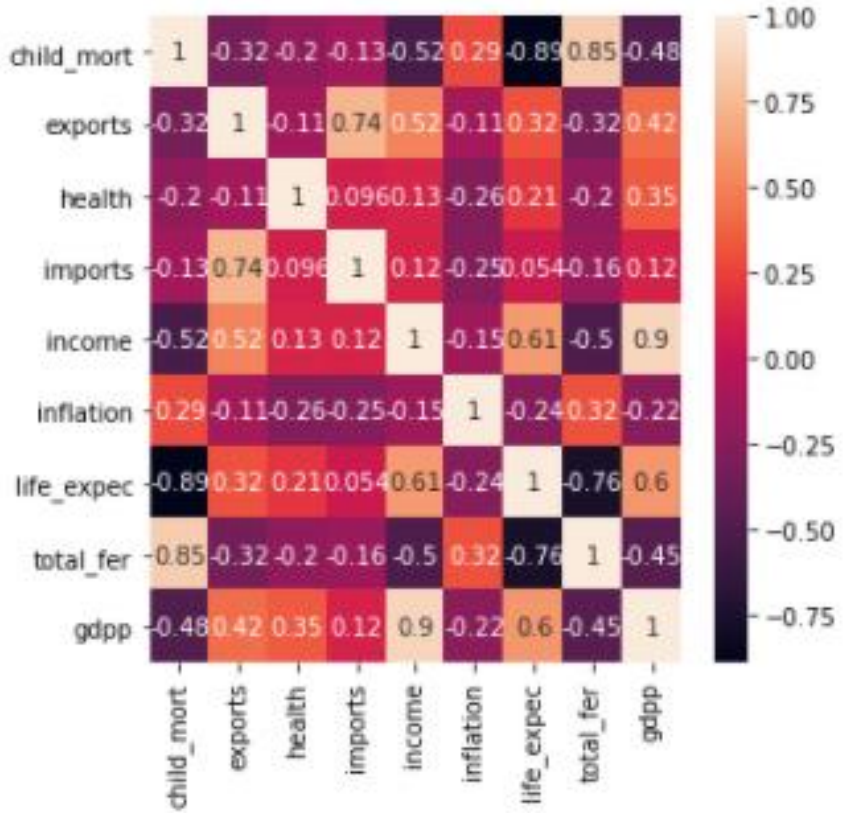
K-mean clustering : Identifying the k value through silhouette score and elbow curve method,

Then forming the clusters on scaled data and adding cluster ID on original data for better interpretation of data .

Hierarchical Clustering :Identifying the Optimal value of K by dendogram , Then forming the clusters on scaled data and adding cluster ID on original data for better interpretation of data .

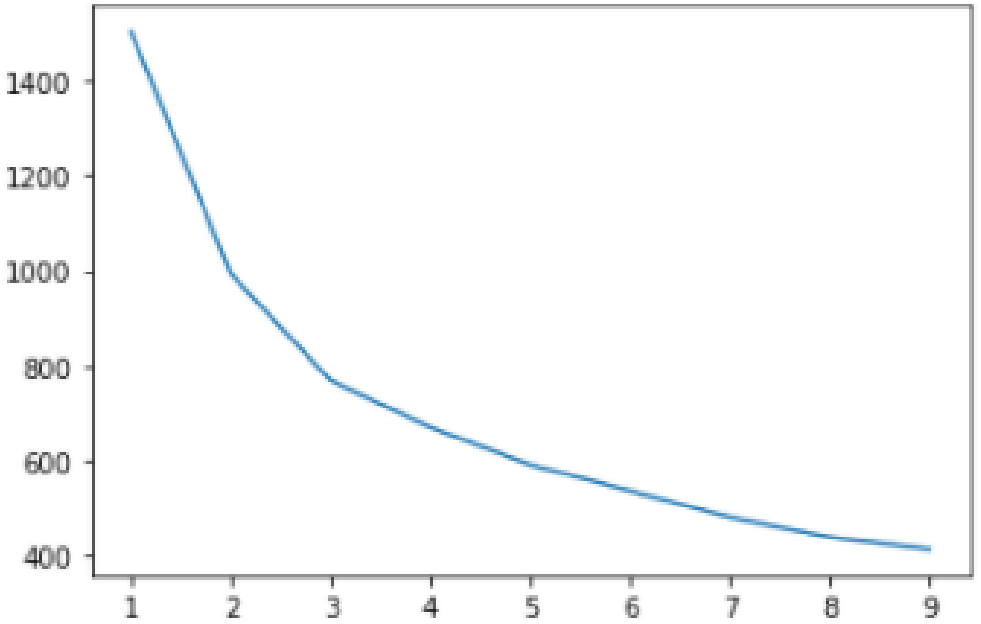
Conclusion / Decision : identifying top 10 countries which are dire of aid

Analyzing How variables are corelated each other

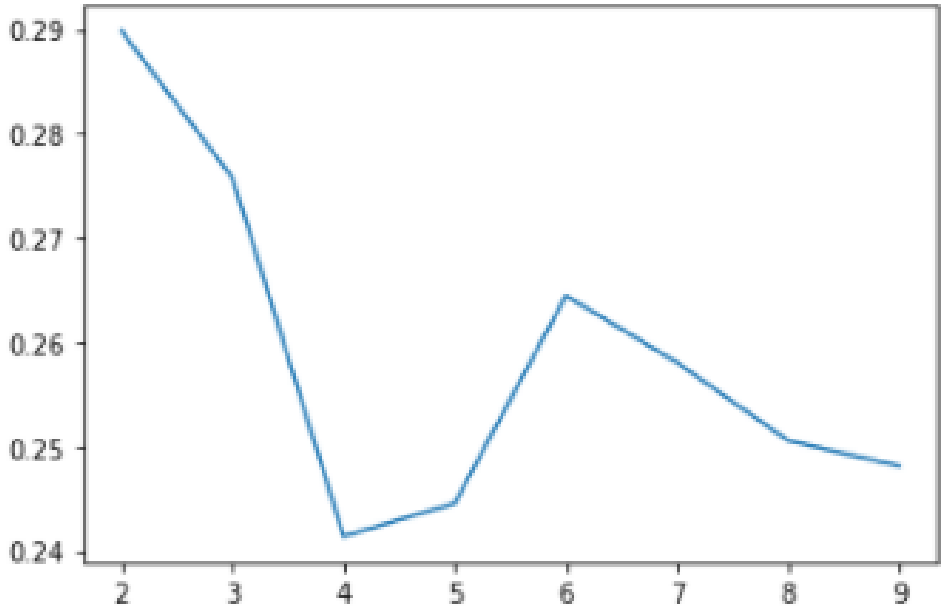


* Income and GDDP are highly corelated with each other

K- Mean Clustering



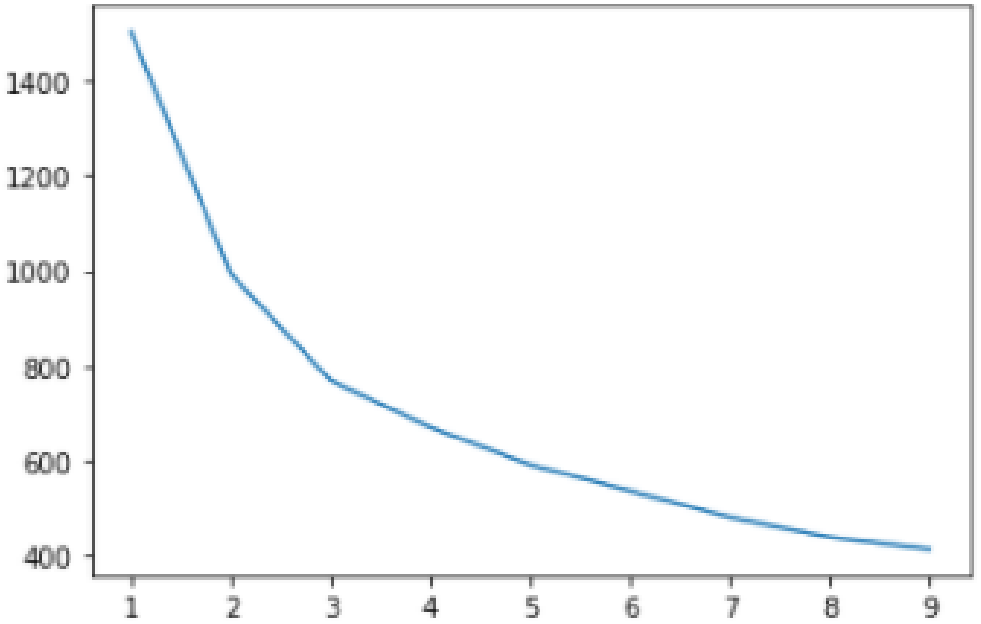
Elbow Curve



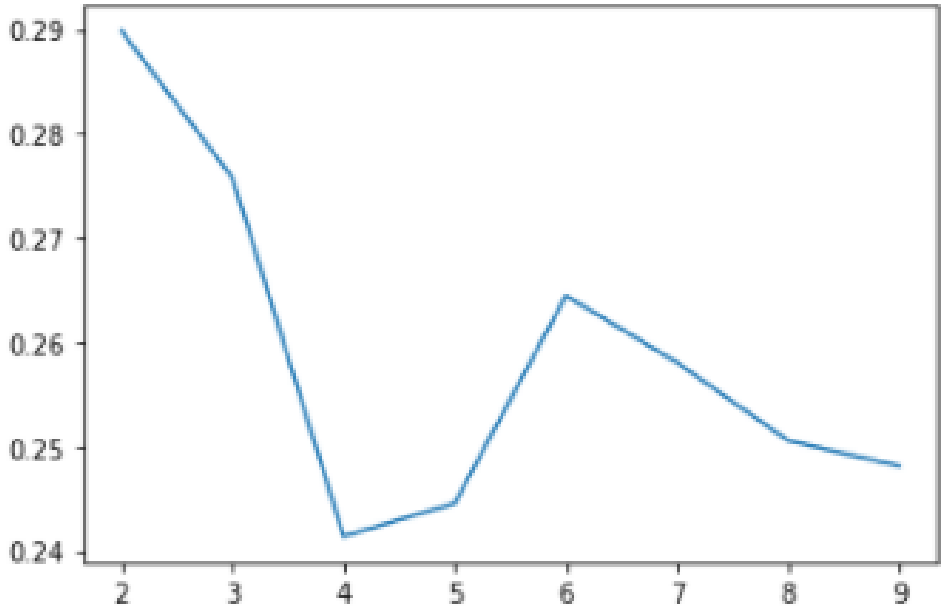
Silhouette Score

From above graph we can conclude that K=4 will be a good number

K- Mean Clustering



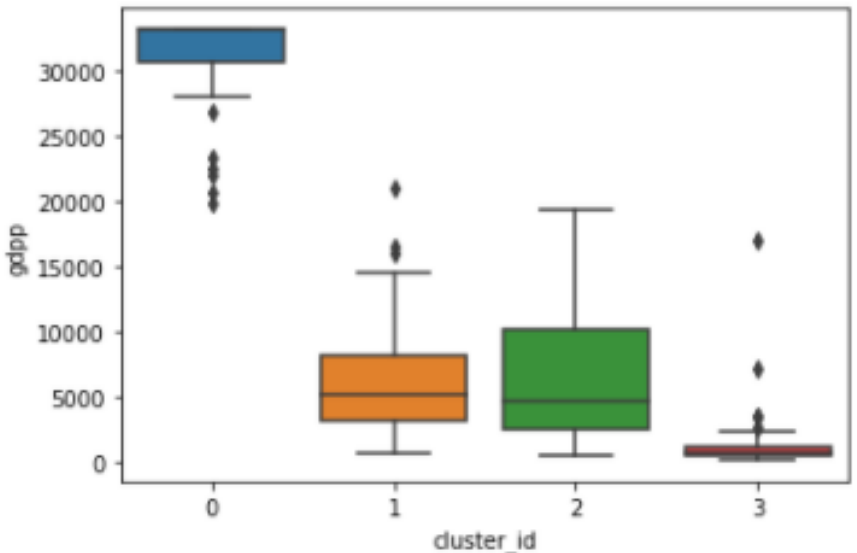
Elbow Curve

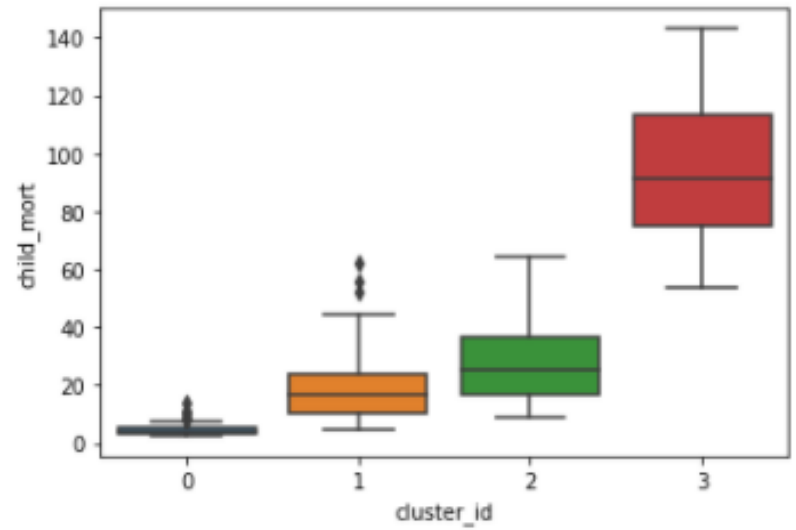


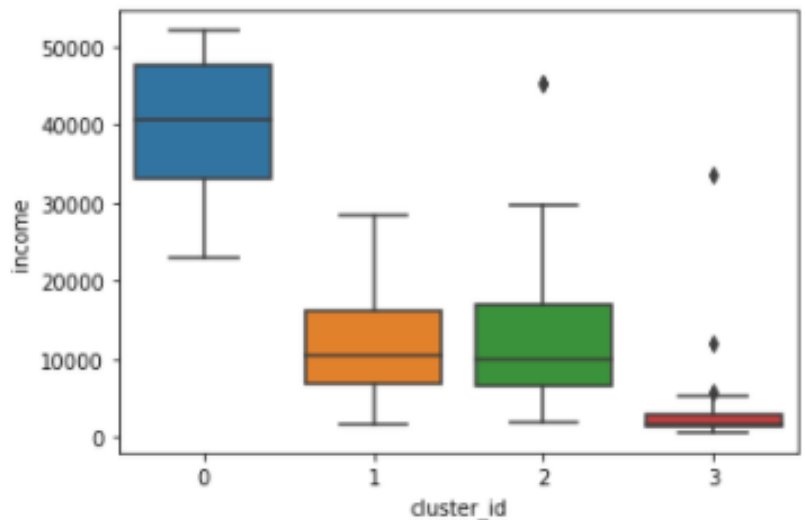
Silhouette Score

From above graph we can conclude that K=4 will be a good number

For Cluster 3 , GDPP and Income are weak , and more Child mortality , So Focus cluster is Cluster3 .





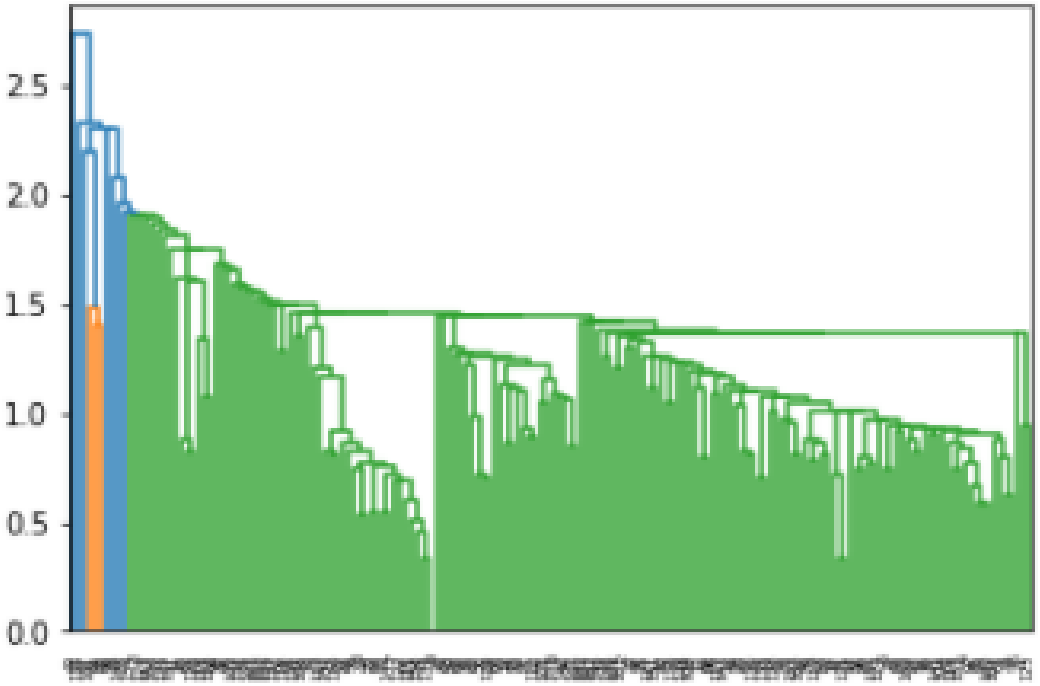


Top 10 countries obtained by K mean clustering by sorting Child mort in descending and income and gdpp in ascending way

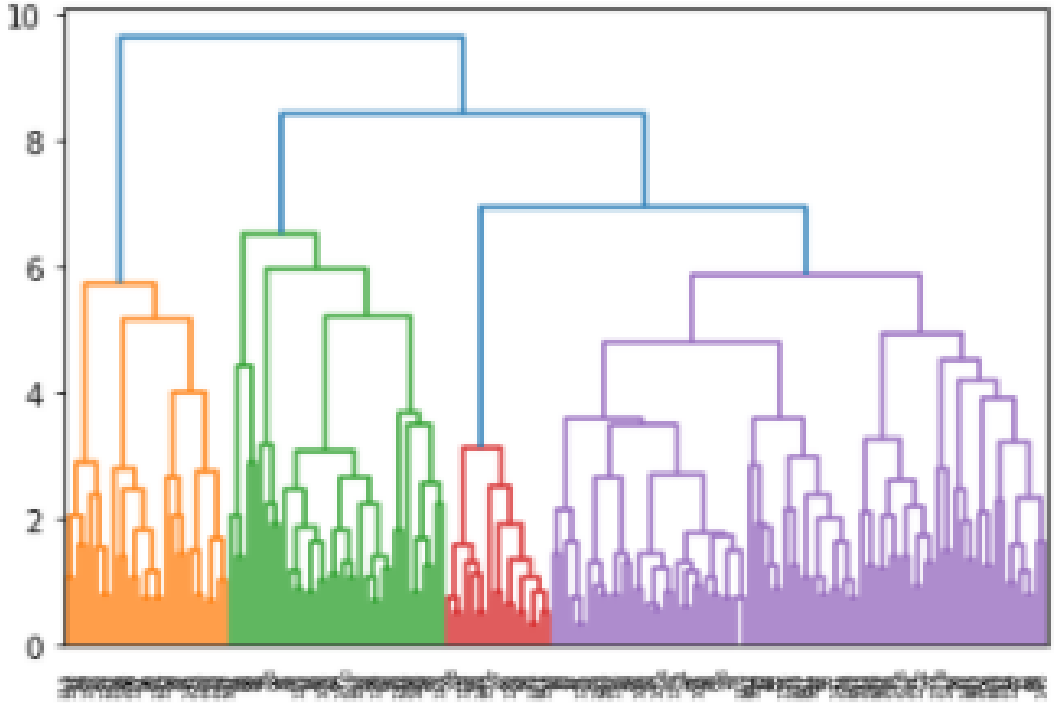
	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	cluster_id
0	Sierra Leone	142.875	16.8	13.10	34.5	1220.0	17.20	55.00	5.2000	399.0	3
1	Central African Republic	142.875	11.8	3.98	26.5	888.0	2.01	48.05	5.2100	446.0	3
2	Haiti	142.875	15.3	6.91	64.7	1500.0	5.45	48.05	3.3300	662.0	3
3	Chad	142.875	36.8	4.53	43.5	1930.0	6.39	56.50	6.5900	897.0	3
4	Mali	137.000	22.8	4.98	35.1	1870.0	4.37	59.50	6.5500	708.0	3
5	Nigeria	130.000	25.3	5.07	17.4	5150.0	24.16	60.50	5.8400	2330.0	3
6	Niger	123.000	22.2	5.16	49.1	814.0	2.55	58.80	7.0075	348.0	3
7	Angola	119.000	62.3	2.85	42.9	5900.0	22.40	60.10	6.1600	3530.0	3
8	Congo, Dem. Rep.	116.000	41.1	7.91	49.6	609.0	20.80	57.50	6.5400	334.0	3
9	Burkina Faso	116.000	19.2	6.74	29.6	1430.0	6.81	57.90	5.8700	575.0	3

Hierarchical Clustering

Single Linkage

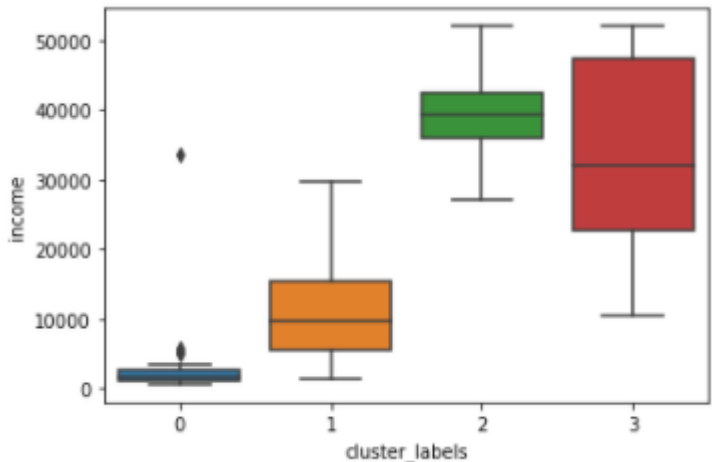


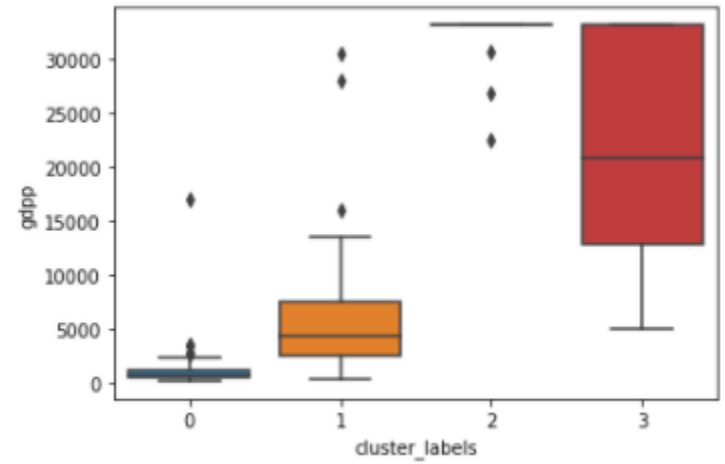
Complete Linkage

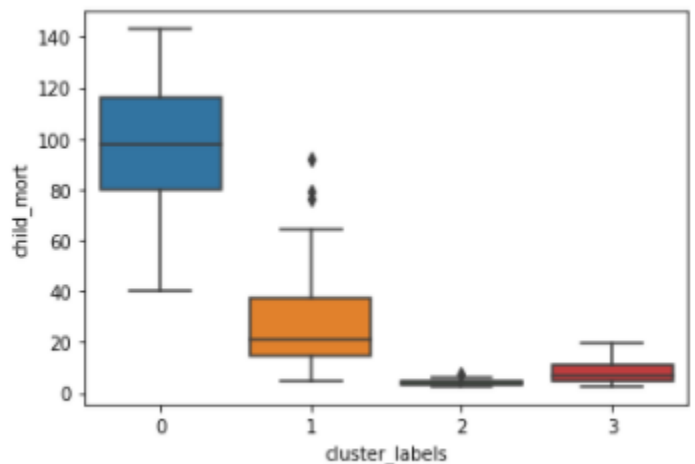


From Two of linkages method , Complete Linkage method is good enough to find out K value , K = 4

For Cluster 0 , GDPP and Income are weak , and more Child mortality , So Focus cluster is Cluster 0 .







Top 10 countries obtained by Hierarchical clustering by sorting Child mort in descending and income and gdpp in ascending way

	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	cluster_labels
0	Sierra Leone	142.875	16.8	13.10	34.5	1220.0	17.20	55.00	5.2000	399.0	0
1	Central African Republic	142.875	11.8	3.98	26.5	888.0	2.01	48.05	5.2100	446.0	0
2	Haiti	142.875	15.3	6.91	64.7	1500.0	5.45	48.05	3.3300	662.0	0
3	Chad	142.875	36.8	4.53	43.5	1930.0	6.39	56.50	6.5900	897.0	0
4	Mali	137.000	22.8	4.98	35.1	1870.0	4.37	59.50	6.5500	708.0	0
5	Nigeria	130.000	25.3	5.07	17.4	5150.0	24.16	60.50	5.8400	2330.0	0
6	Niger	123.000	22.2	5.16	49.1	814.0	2.55	58.80	7.0075	348.0	0
7	Angola	119.000	62.3	2.85	42.9	5900.0	22.40	60.10	6.1600	3530.0	0
8	Congo, Dem. Rep.	116.000	41.1	7.91	49.6	609.0	20.80	57.50	6.5400	334.0	0
9	Burkina Faso	116.000	19.2	6.74	29.6	1430.0	6.81	57.90	5.8700	575.0	0

Top 10 countries obtained by Hierarchical clustering and By Kmean Clustering to be focused to dire of aid.

- Sierra Leone
- Central African Republic
- Haiti
- Chad
- Mali
- Nigeria
- Niger
- Angola
- Congo, Dem. Rep
- Burkina Faso