HELP INTERNATIONAL

WE CARE

AIM:

- Finding countries which are in the direst need of help based on their socio-economic and health factors.
- Distribution of recent funding accordingly

DATA UNDERSTANDING:

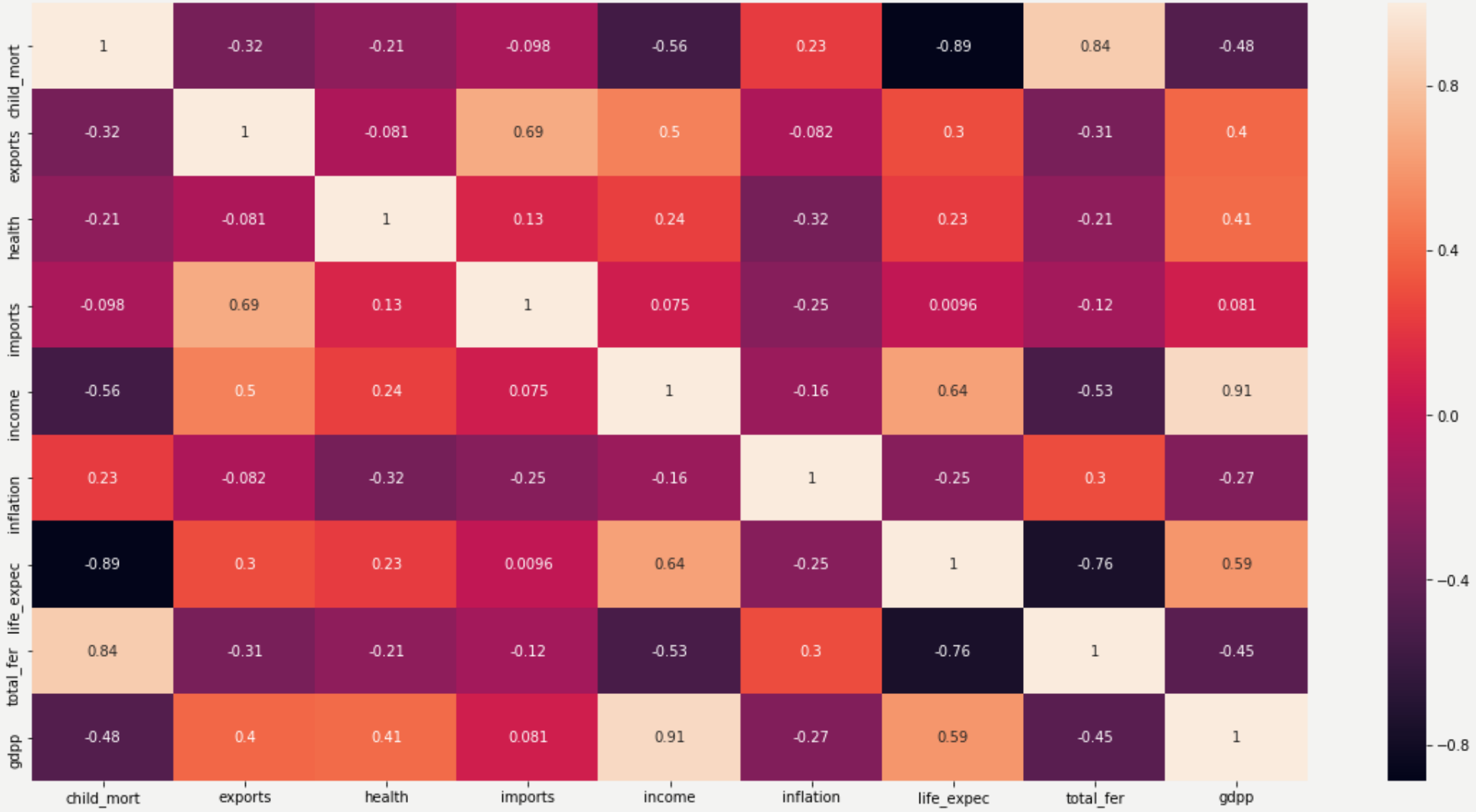
Data Dctionary

Column Name	Description
country	Name of the country
child_mort	Death of children under 5 years of age per 1000 live births
exports	Exports of goods and services. Given as %age of the Total GDP
health	Total health spending as %age of Total GDP
imports	Imports of goods and services. Given as %age of the Total GDP
Income	Net income per person
Inflation	The measurement of the annual growth rate of the Total GDP
life_expec	The average number of years a new born child would live if the current mortality patterns are to remain the same
total_fer	The number of children that would be born to each woman if the current age-fertility rates remain the same.
gdpp	The GDP per capita. Calculated as the Total GDP divided by the total population.

DATA PREPARATION:

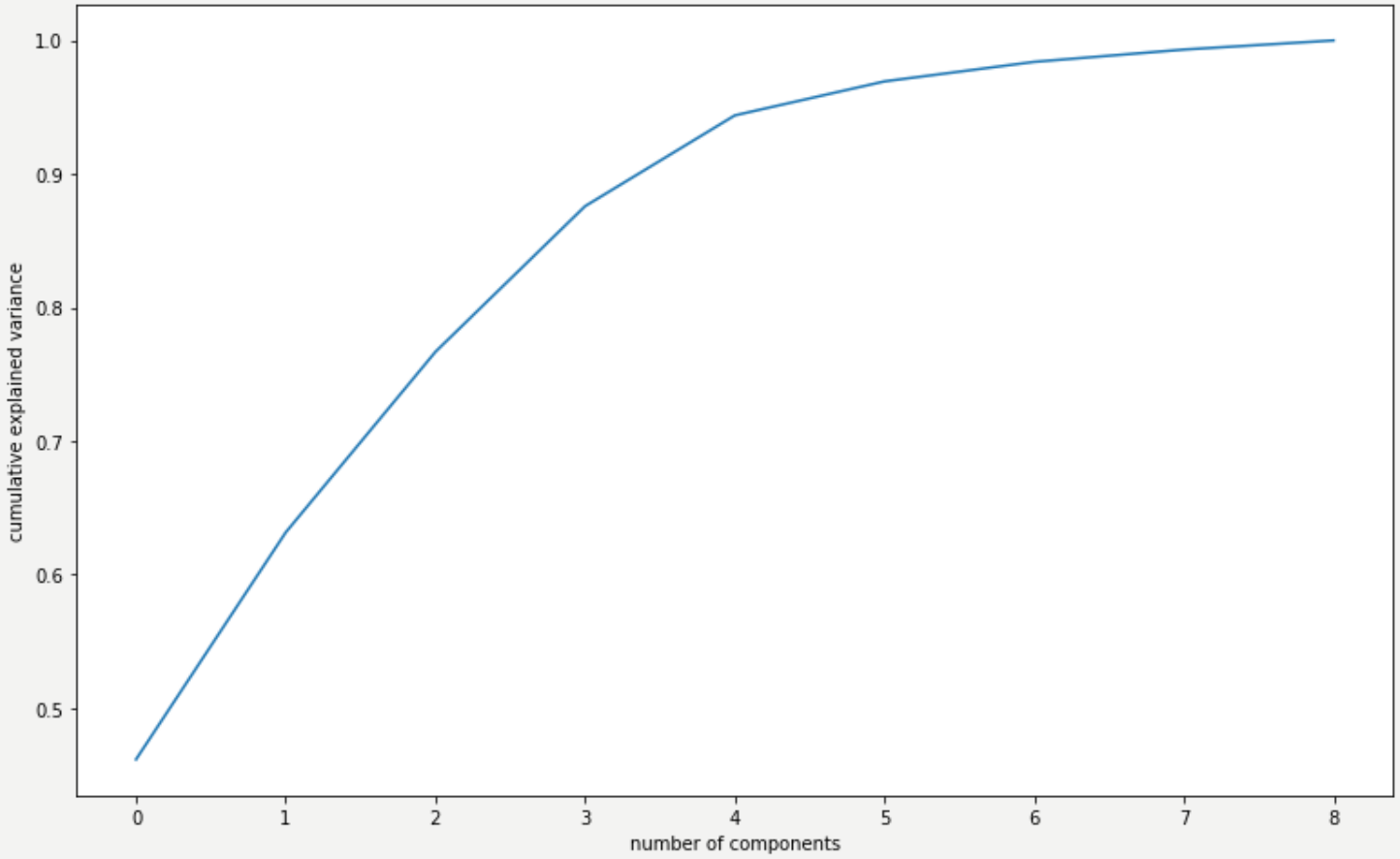
- Data from 167 countries were collected
- Data set has 0 missing values
- Data set has some outliers which were treated
- Continuous variables were standardized

CORRELATION BETWEEN DIFFERENT VARIABLES:



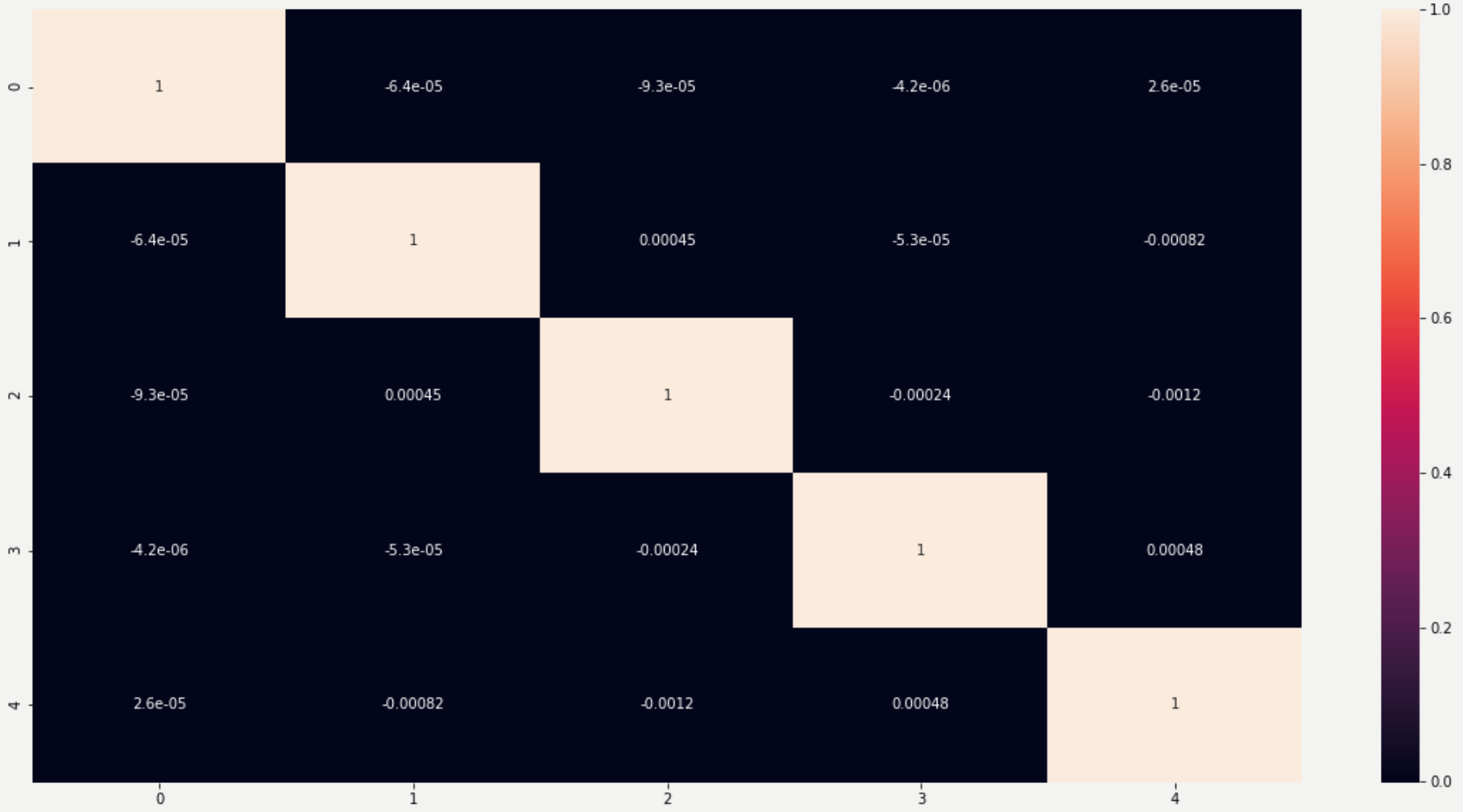
As correlation is found among variables, PCA is performed

PCA ANALYSIS:



from the scree plot, it is found that 5 number of components can describe around 95% variance in the data set

CORRELATION AFTER PCA



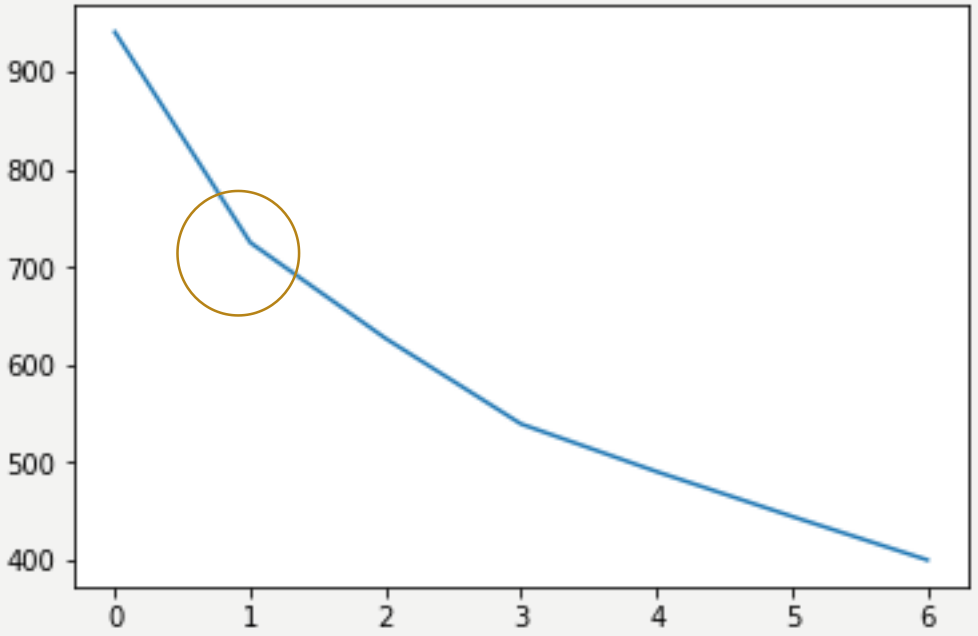
Now we have variables with ‘0’ correlation among them

CLUSTERING:

- Clustering is performed in order to group and identify countries which are in direst need of help from other countries
- This helps in understanding and proper utilization of the funding received recently

CLUSTERING (BY K-MEANS):

SSD/ Elbow-curve

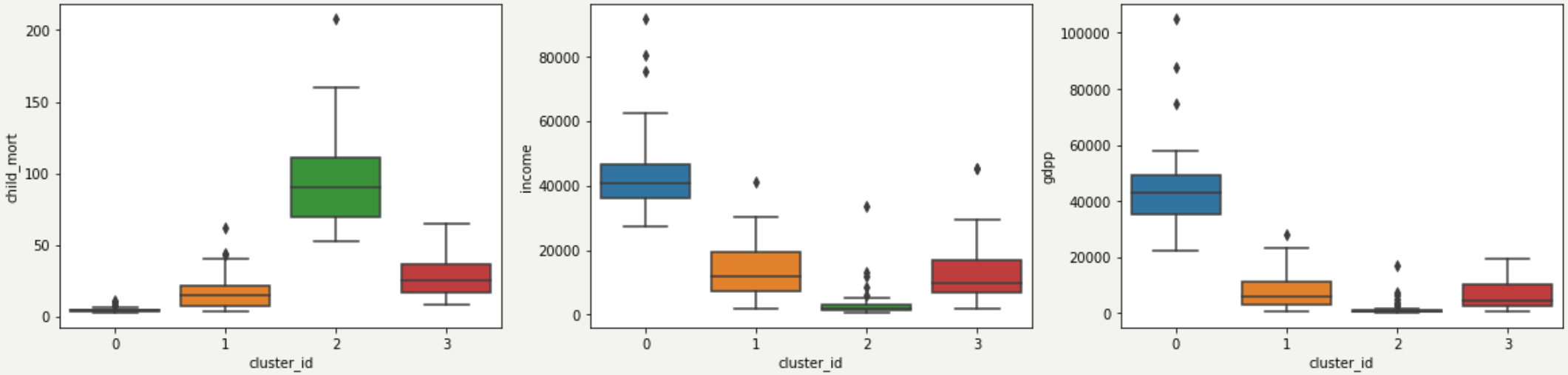


silhouette analysis

For n_clusters=2, the silhouette score is 0.2947265724525269
For n_clusters=3, the silhouette score is 0.2953926549102592
For n_clusters=4, the silhouette score is 0.2588650094831585
For n_clusters=5, the silhouette score is 0.26788564884683724
For n_clusters=6, the silhouette score is 0.27054453015052826
For n_clusters=7, the silhouette score is 0.26546933771700437
For n_clusters=8, the silhouette score is 0.27351801371482 38

range_n_clusters = [2, 3, 4, 5, 6, 7, 8]

from elbow curve and silhouette analysis, we are taking k =4,

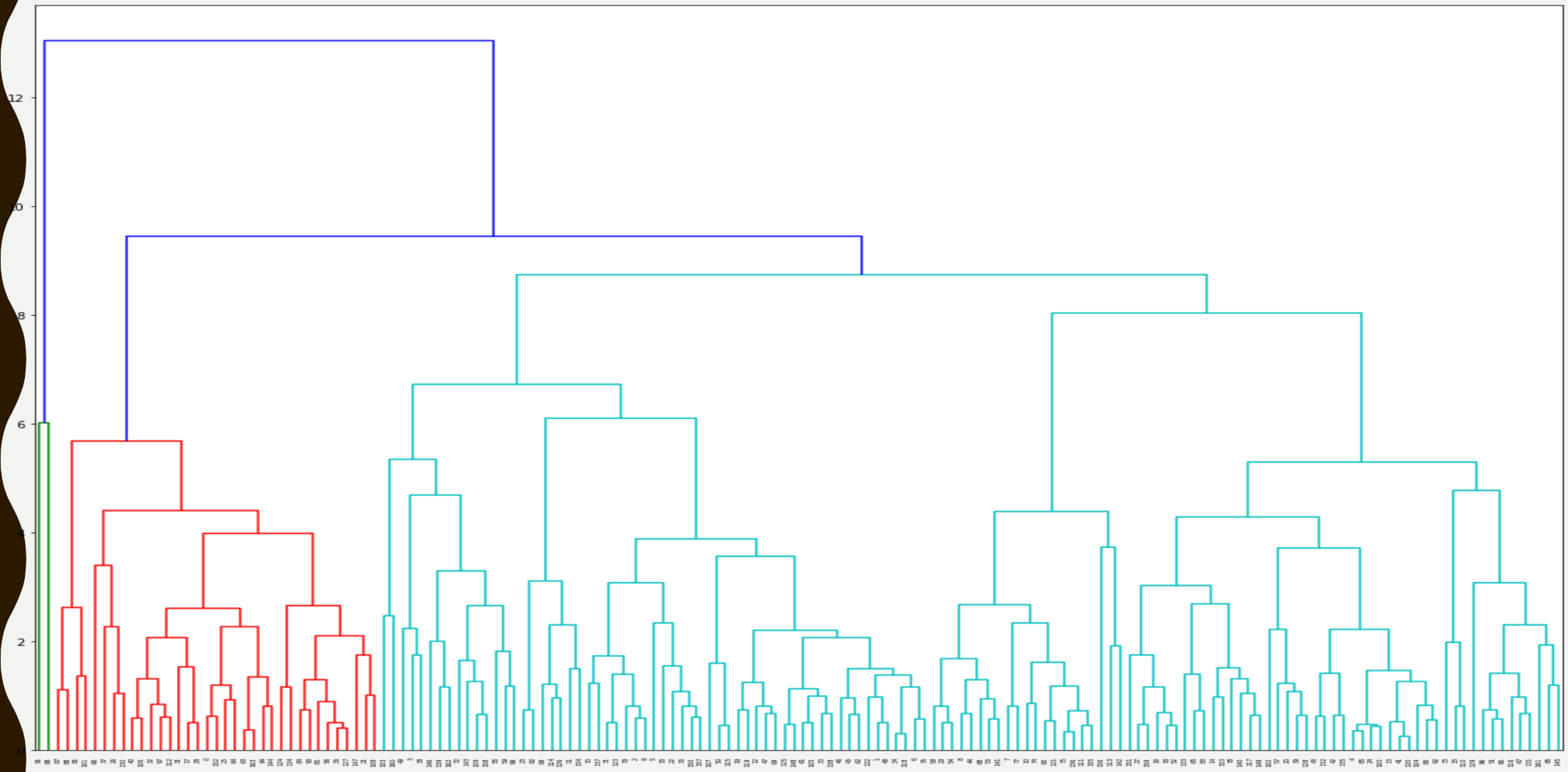


Analysis :

- On the basis of 'child_mort' clusters are classified as 2 > 3 > 1 > 0
- On the basis of 'income' clusters are classified as 0 > 1 > 3 > 2
- On the basis of 'gdpp' clusters are classified as 0 > 1 > 3 > 2

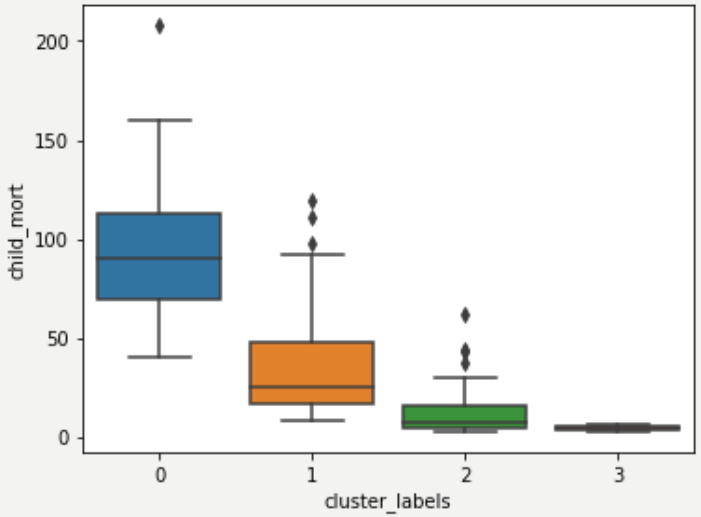
- Hence, cluster 2 has low 'gdpp', low 'income, and high 'child_mort'
- Which signifies that condition of countries belongs to cluser-2 is bad and help should be provided here.

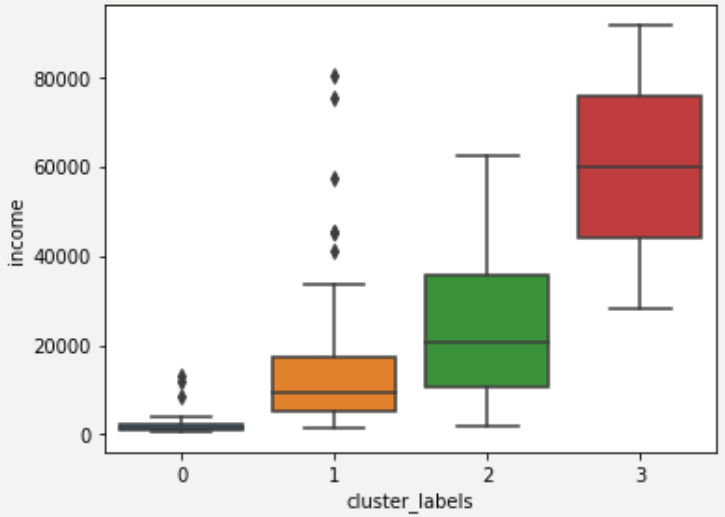
HIERARCHICAL CLUSTERING:

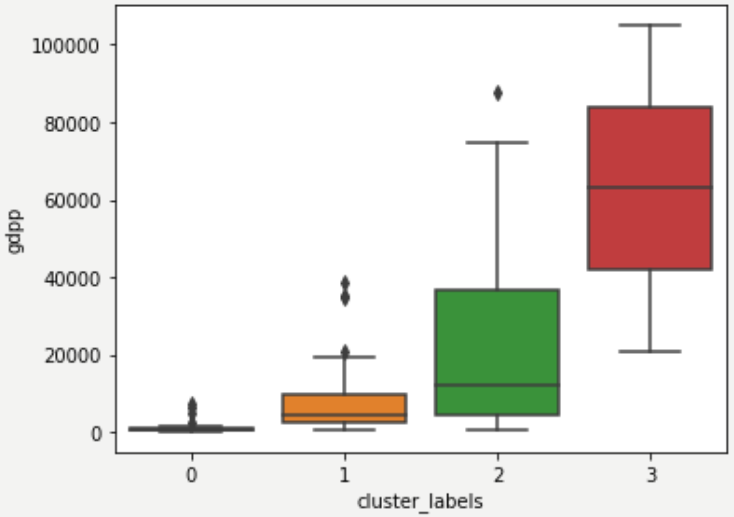


Complete linkage Dendogram

- Number of cluster is taken as 4 in order to compare with results of k-means clustering







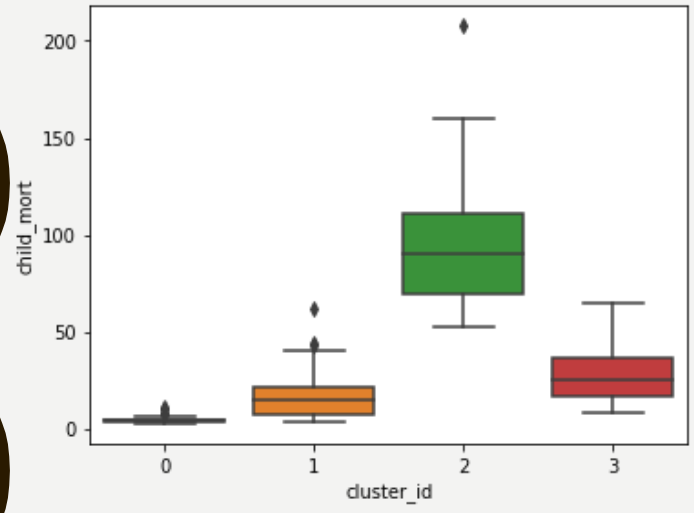
Analysis :

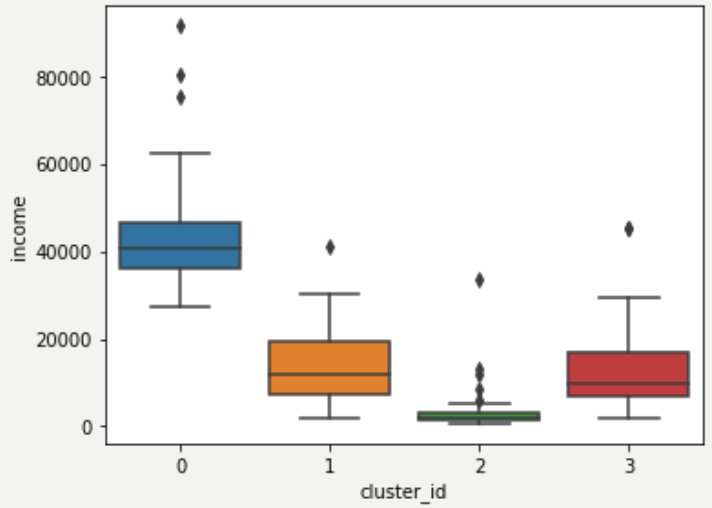
- On the basis of 'child_mort' clusters are classified as 0 > 1 > 2 > 3
- On the basis of 'income' clusters are classified as 3 > 2 > 1 > 0
- On the basis of 'gdpp' clusters are classified as 3 > 2 > 1 > 0

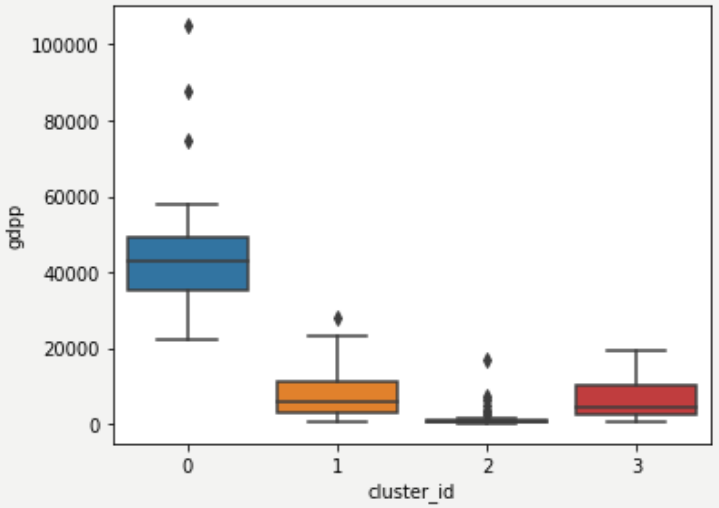
- Hence, cluster 2 has low 'gdpp', low 'income, and high 'child_mort'

K-MEANS AND HIERARCHICAL CLUSTERING

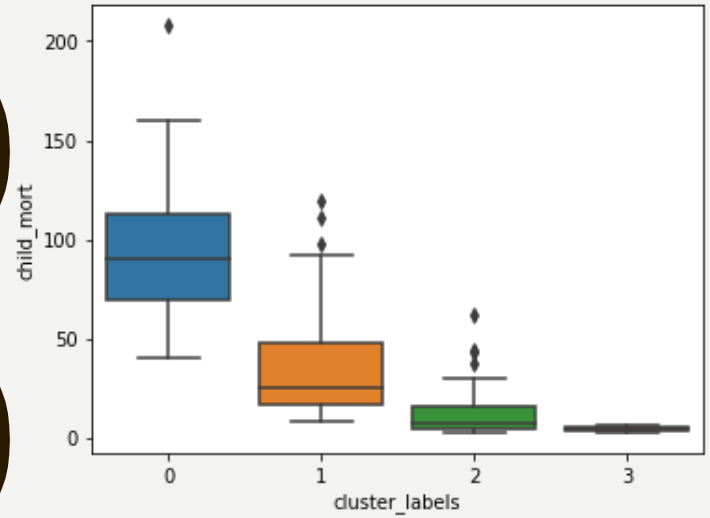
k-means clustering

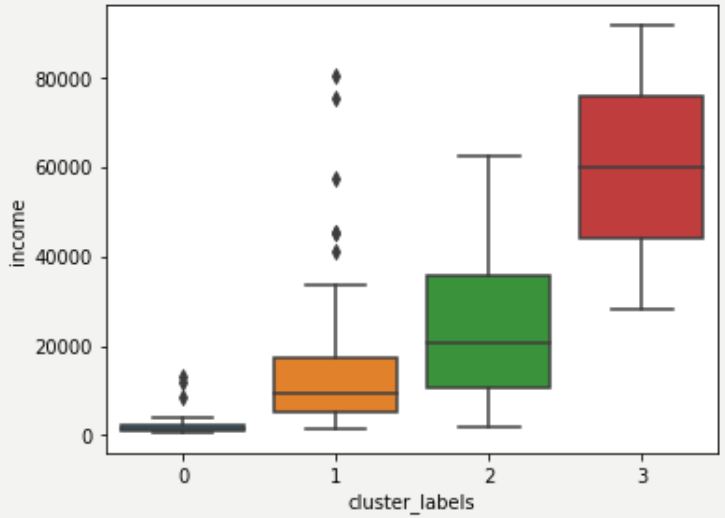


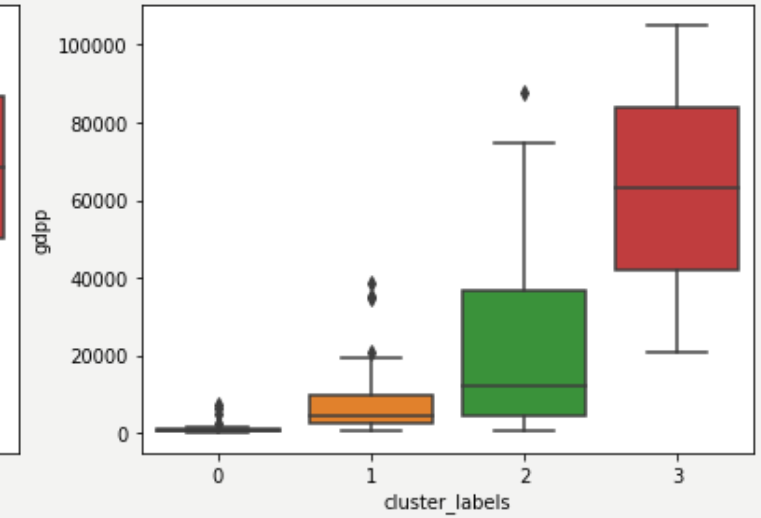




hierarchical clustering







BY COMPARING BOXPLOTS OF CLUSTERS CREATED BY BOTH K-MEANS AND HIERARCHICAL CLUSTERING ,WE FOUND THAT K-MEANS CLUSTERING PRODIVES BETTER RESULTS. HENCE WE ARE CONSIDERING CLUSTER-2 COUNTRIES

Top – 10 Countries which are direst need of help

1. Haiti
2. Sierra Leone
3. Chad
4. Central African Republic
5. Mali
6. Niger
7. Angola
8. Congo, Dem. Rep.
9. Burkina Faso
10. Guinea-Bissau

This list is created on the basis of 'Child_mort – high'.This is because if children are dying then this indicates:

- **Poor healthcare system**
- **Poor education**
- **Poor environment , etc**

So money can be used in this countries to improve the quality of their lives.

THANK YOU