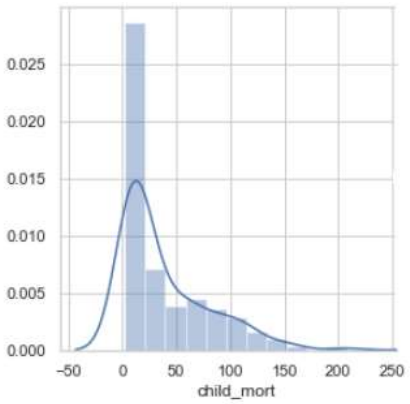
# Humanitarian Financial Aid

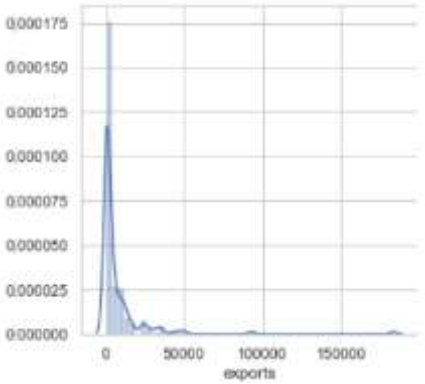
BY SABYASACHI PARIDA

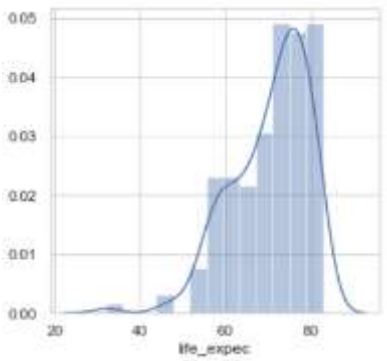
# Objective

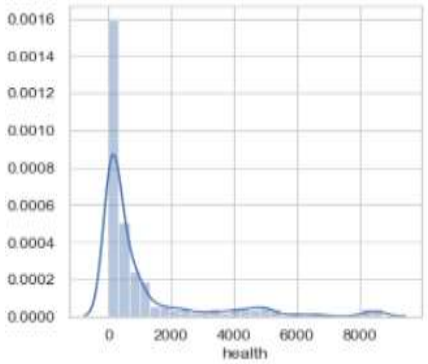
I. categorize the countries using some socio-economic and health factors that determine the overall development of the country
II. Determine the top 5 countries which are in direst need of aid.

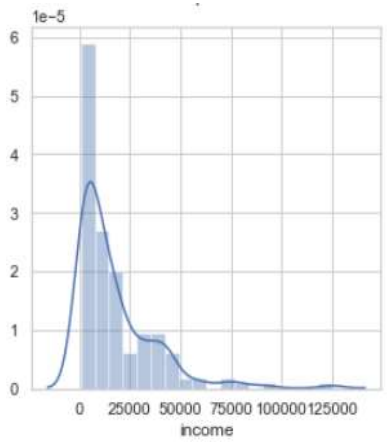
# Data Distribution

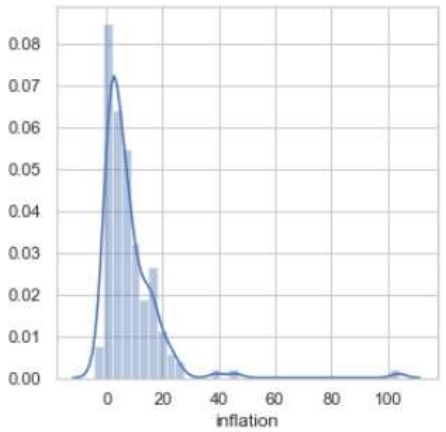


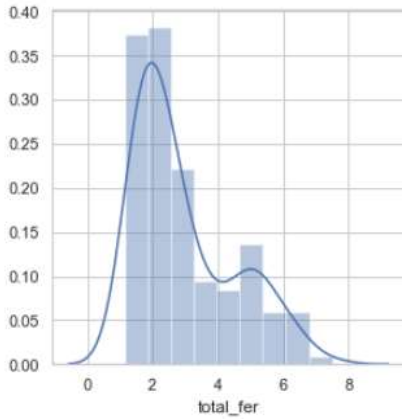



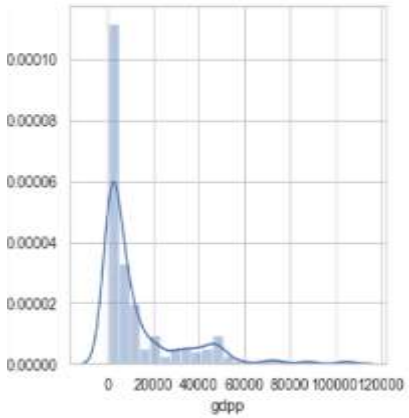



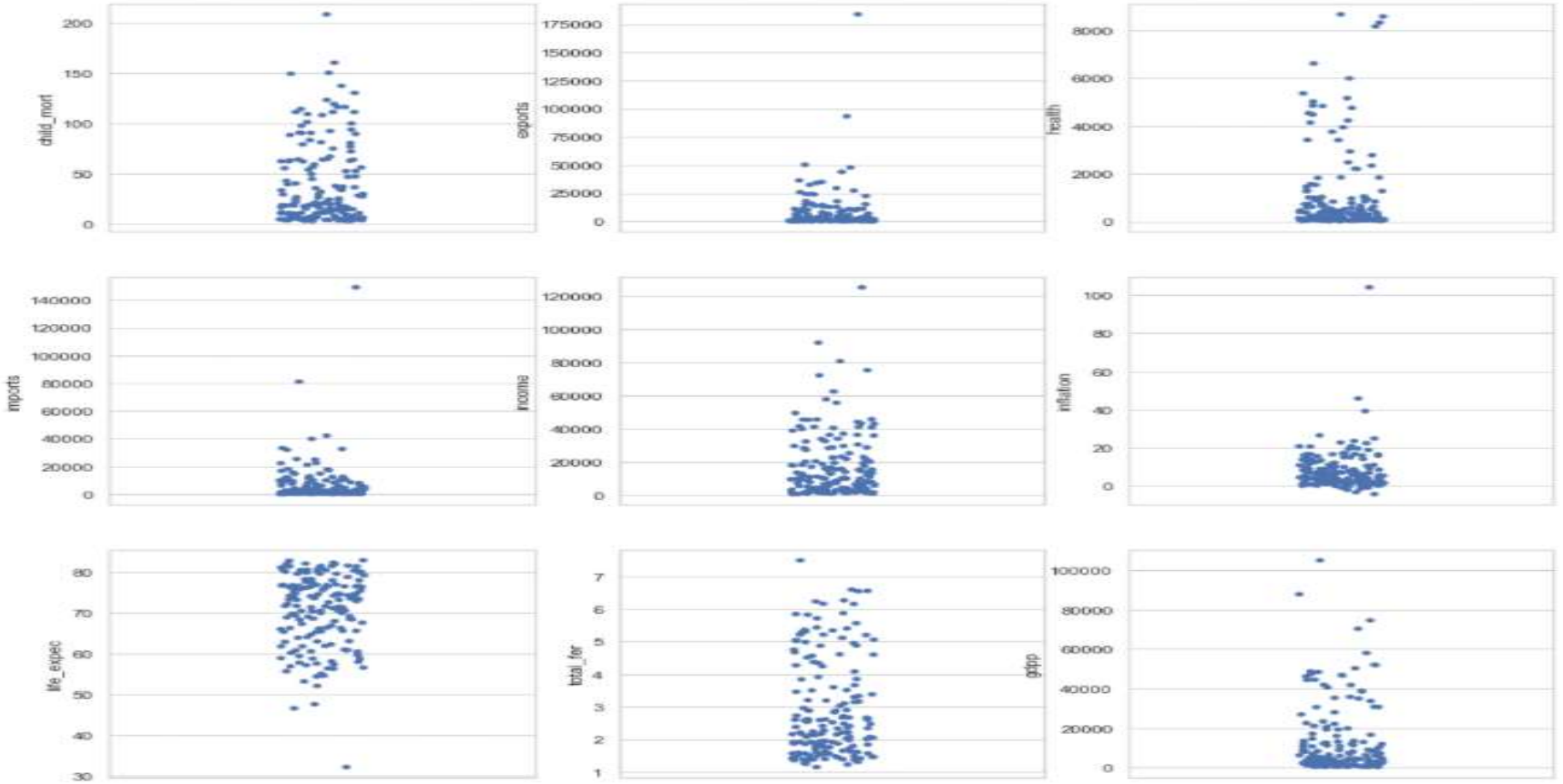




**Observation**

I. Child mortality and export are right skewed distribution.
II. Life expectancy is left skewed .
III. Child mortality , Income, total fertility and GDPP doesn't seem to be following normal distribution which can be used for cluster profiling later.

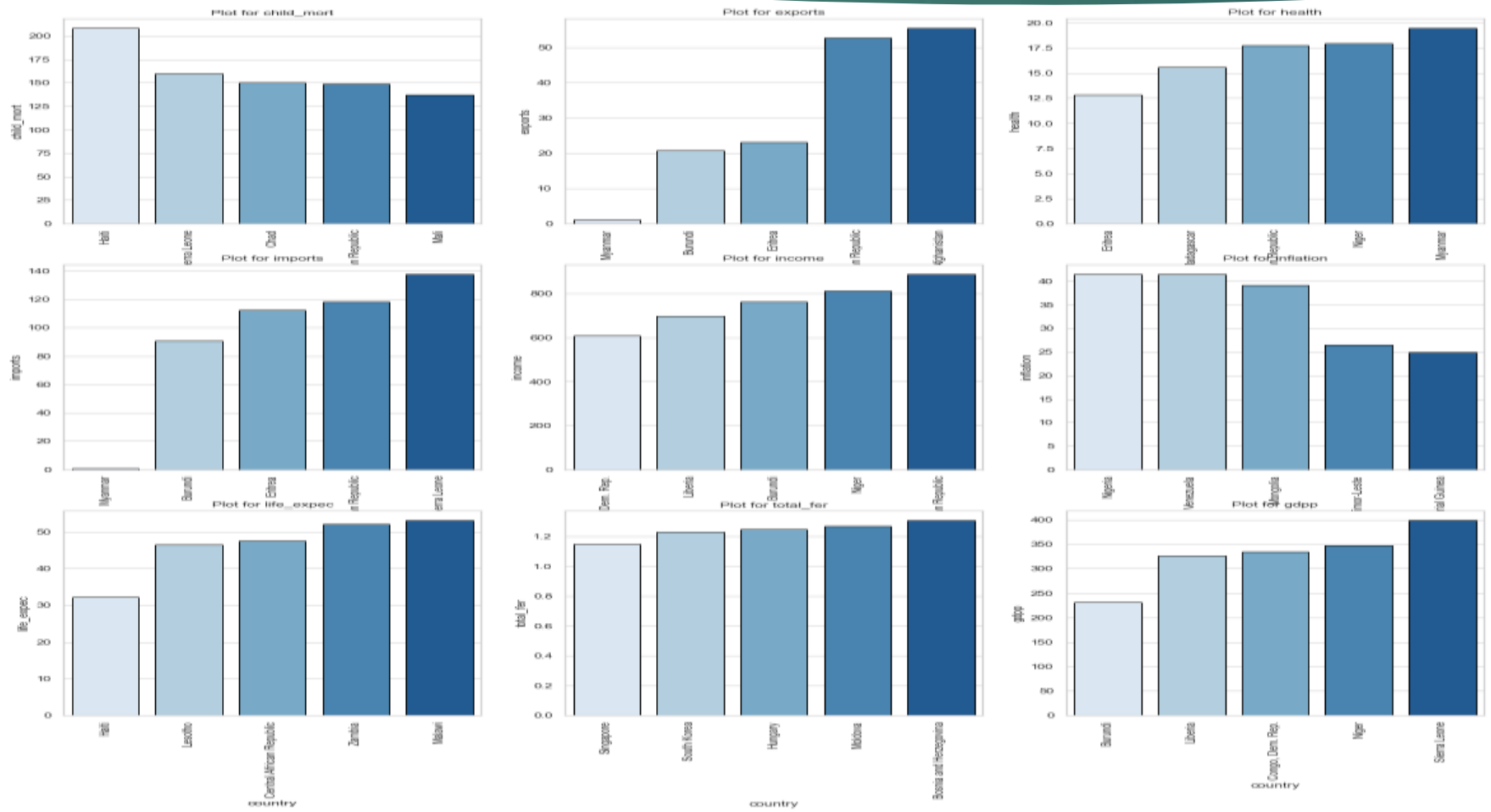
# Univariate Analysis




**Observation**

I. Most of the values of Life expectancy is above 50
II. GDPP of most of the countries is less than 60000.
III. Inflation of most of the countries lie below 20.
IV. Imports and Exports of most of the countries are below 50000.

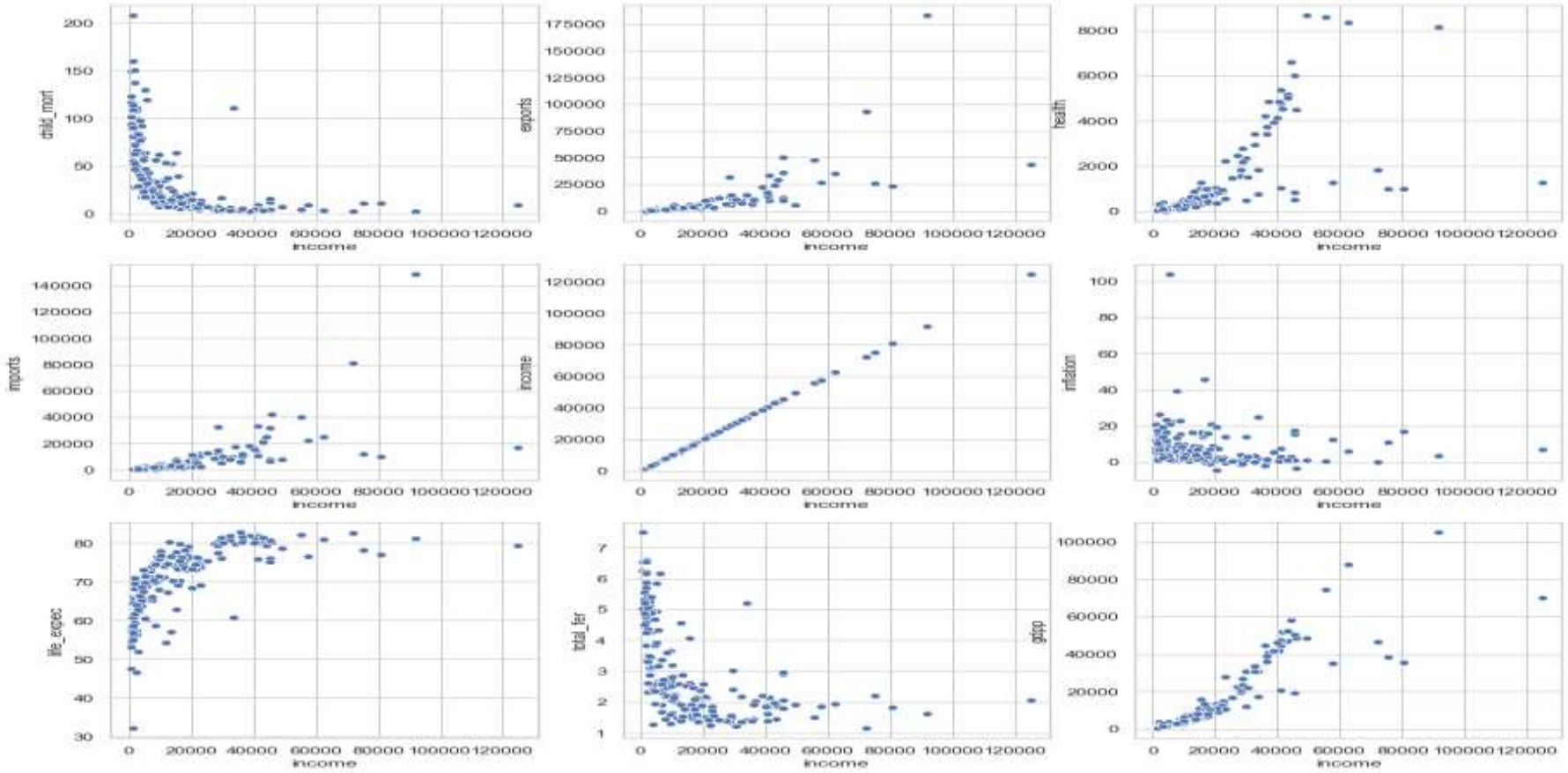
# Univariate Analysis(top 5 from bottom in each field)




**Observation**

I. Haiti has the highest child mortality rate and lowest Life expectancy.
II. Myanmar does the least import as well as export.
III. Eritria spends the least in health per person.
IV. Congo has the least income .
V. Nigeria has the highest Inflation.
VI. Burundi has the least GDPP .

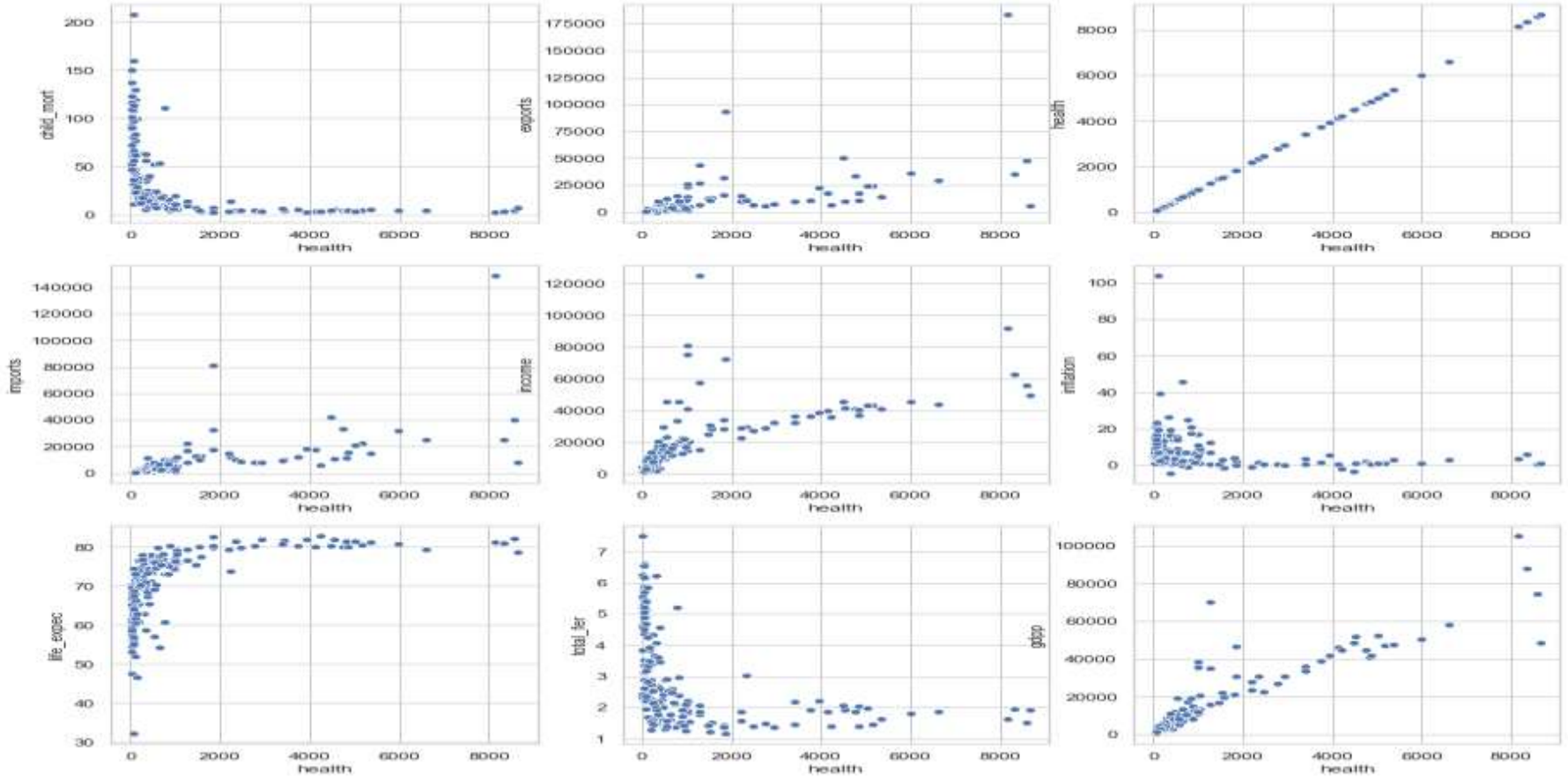
# Bivariate Analysis (with respect to income)




**Observation**

I. With increase in income child mortality decreases.
II. With increase in income Export, import, health spending per person , Life expectancy , GDPP increases .

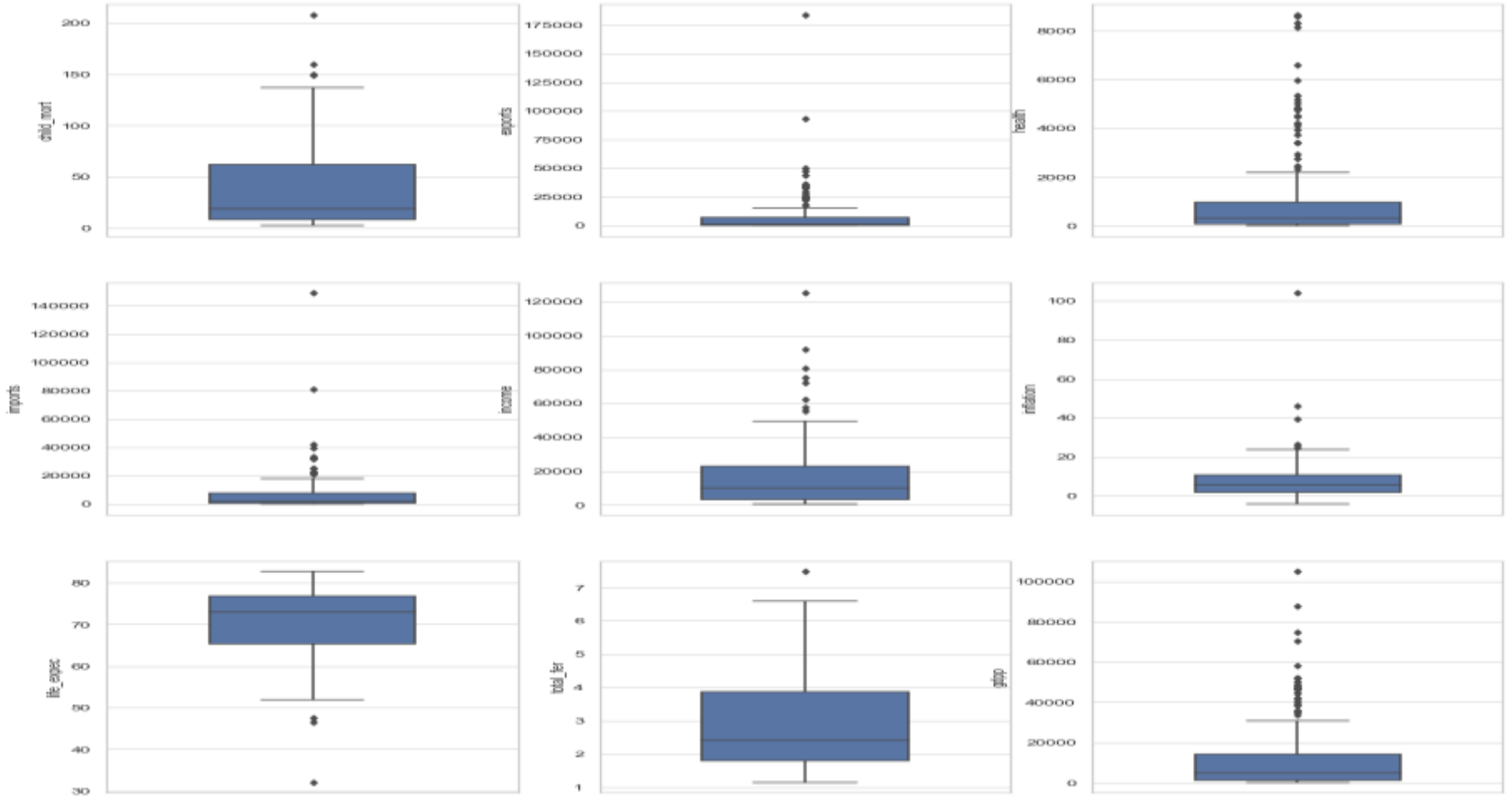
# Bivariate Analysis (with respect to Spending in Health)




**Observation**

I. Child mortality decreases with increase in spending in health.
II. Life expectancy increases with increase in spending in health.
III. GDPP also increases with increasing in spending in health .

# Outlier Analysis

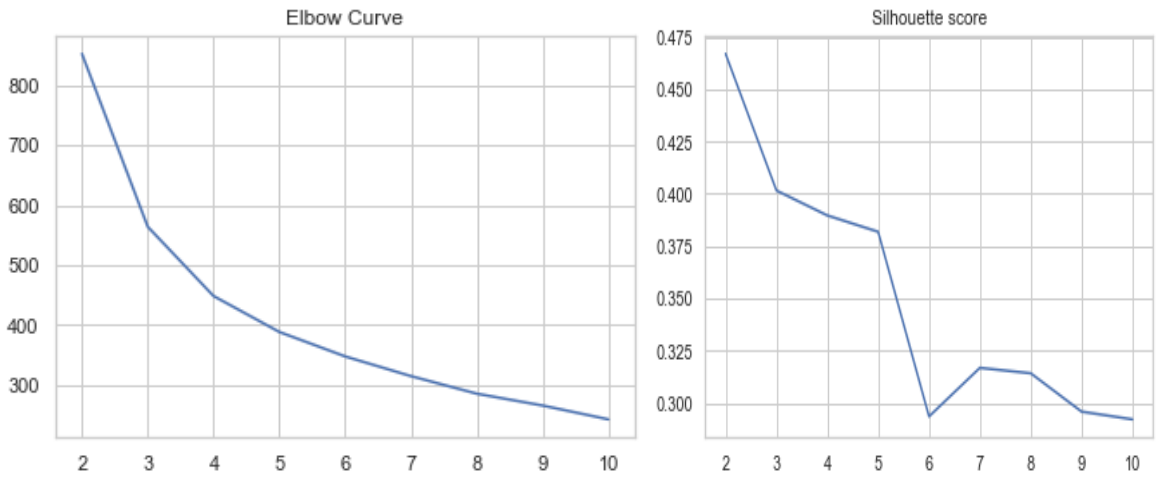



**Observation**

I. There are outliers for all the variables .
II. We will not be capping Life expectancy, GDPP and child mortality rates as those might be the data for countries which would be in direst need of Aid.
III. As very less data is present we will not delete any outliers.

# Hopkins Score , Elbow curve , Silhouette score

| Average Hopkins Score |
| --- |
| 0.915160168 |







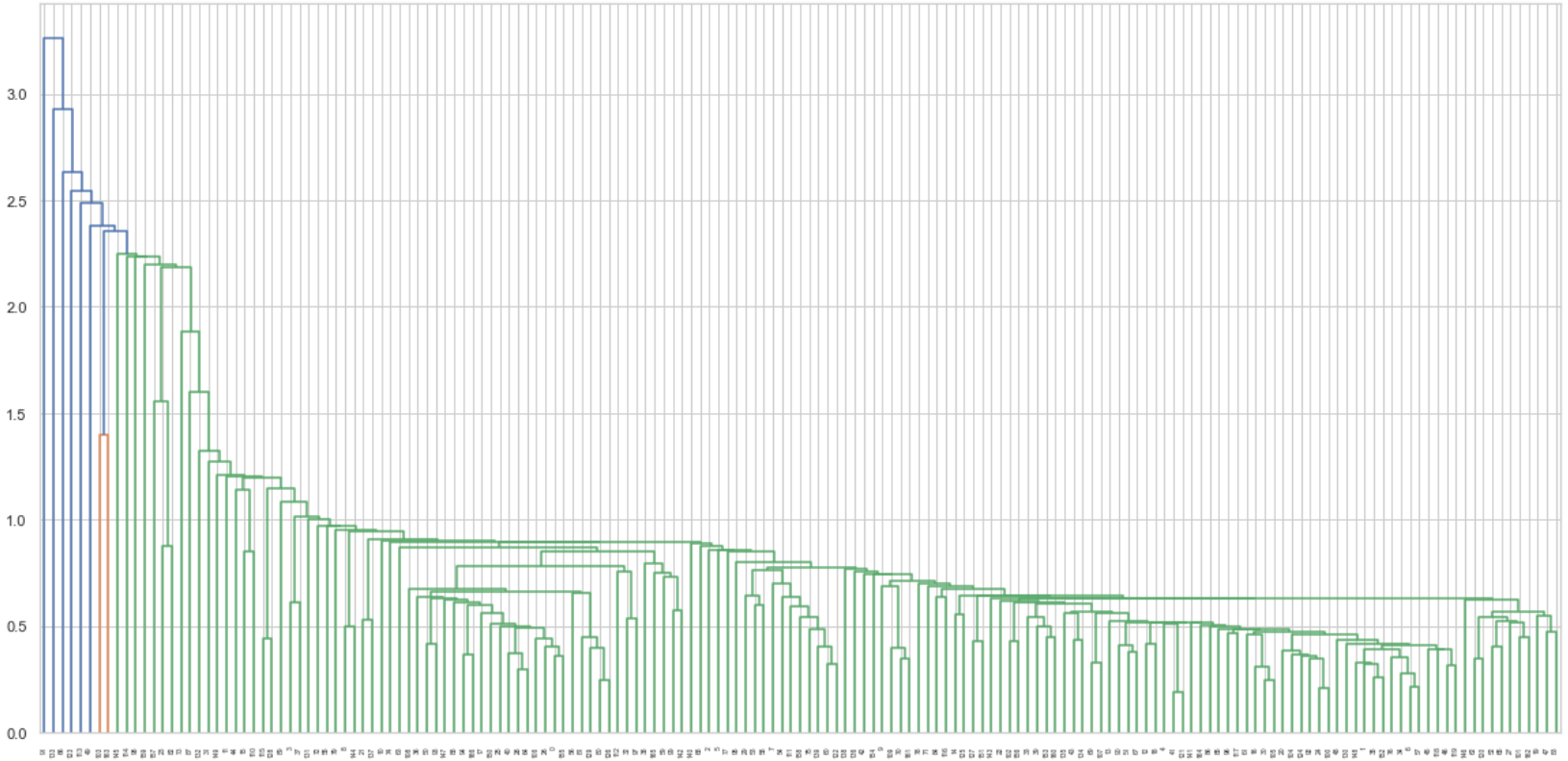
**Observation**

I. A Hopkins score above 0.7 indicates a clustering tendency at 90% confidence level. Our average score is 0.91.

II. From both Elbow curve and Silhouette score statistically it is suggested to take 3 as number of clusters, but we will take 4 and 5 to test as both of them has a good score as 3 would be very less number of clusters as per my understanding.

III. Will choose the best between 4 or 5 and will proceed with cluster profiling.

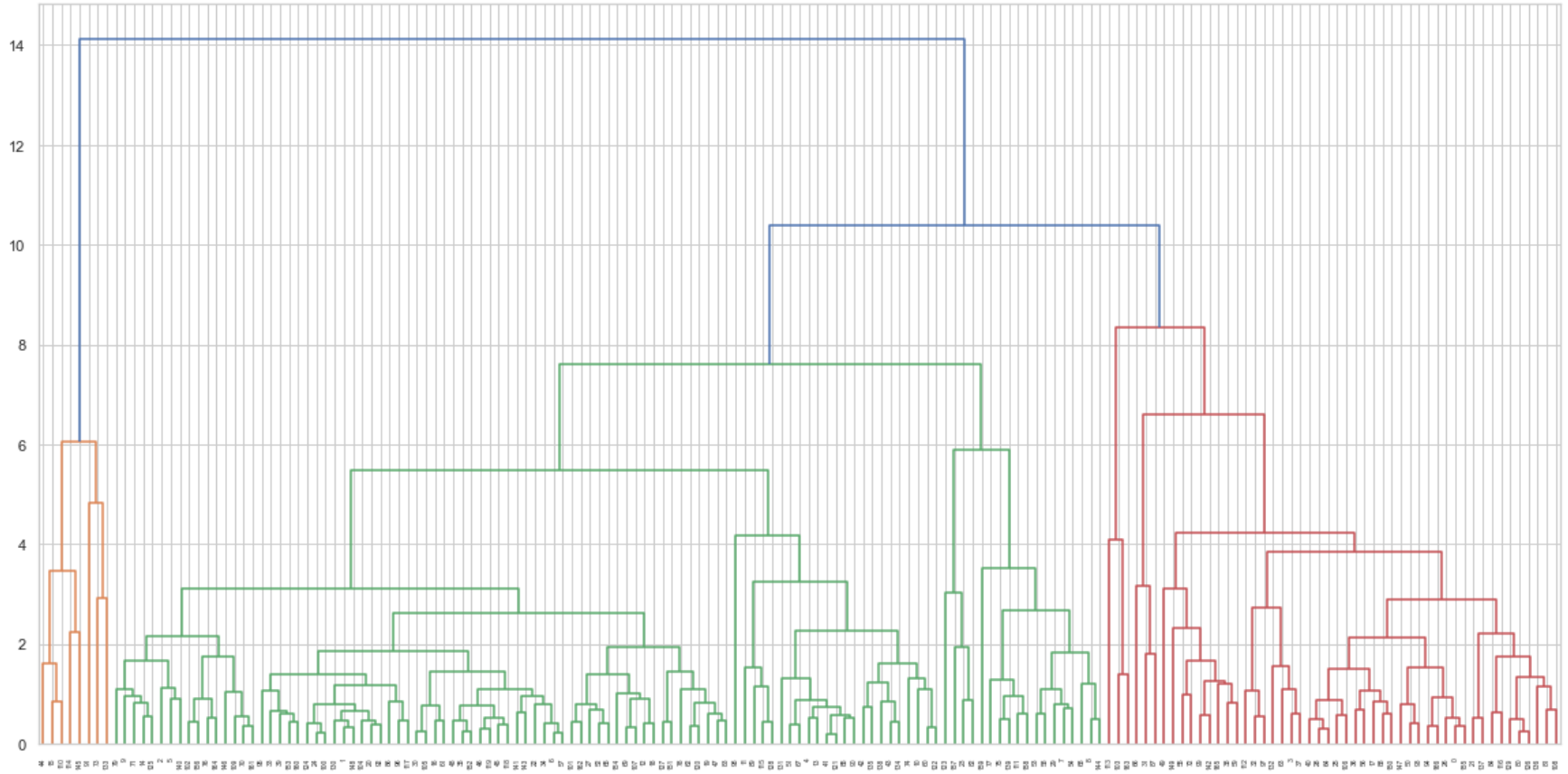
# Hierarchical Clustering (Single linkage)




**Observation**

I. We Cannot find much distinctive clusters hence it is recommended to try Hierarchical Clustering Complete Linkage.

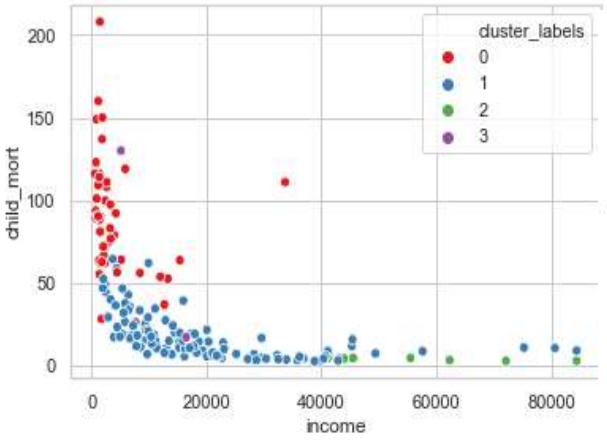
# Hierarchical Clustering (Complete linkage)

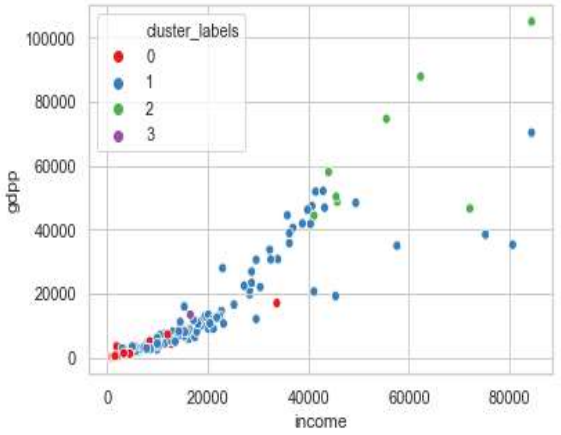


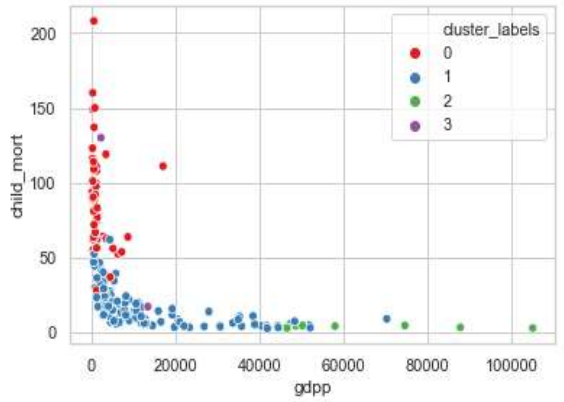

**Observation**

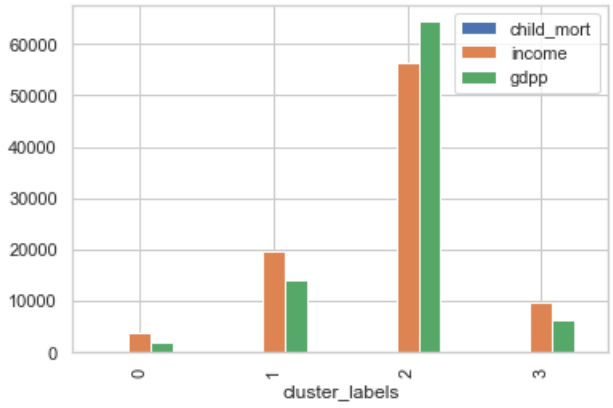
I. We can Cut the dendrogram at 8 i.e. for 4 clusters ,as we can visualize 4 distinctive clusters.

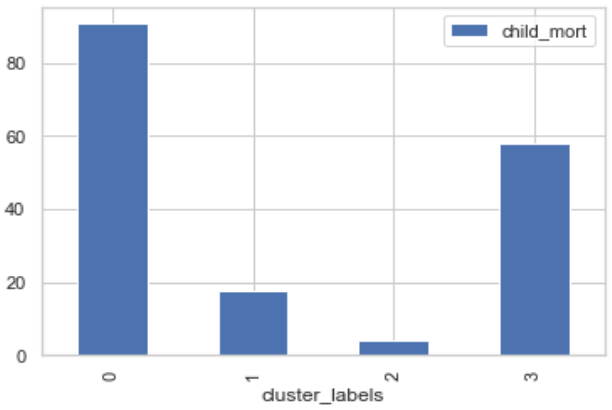
# Hierarchical Clustering (Complete linkage)




**Observation**

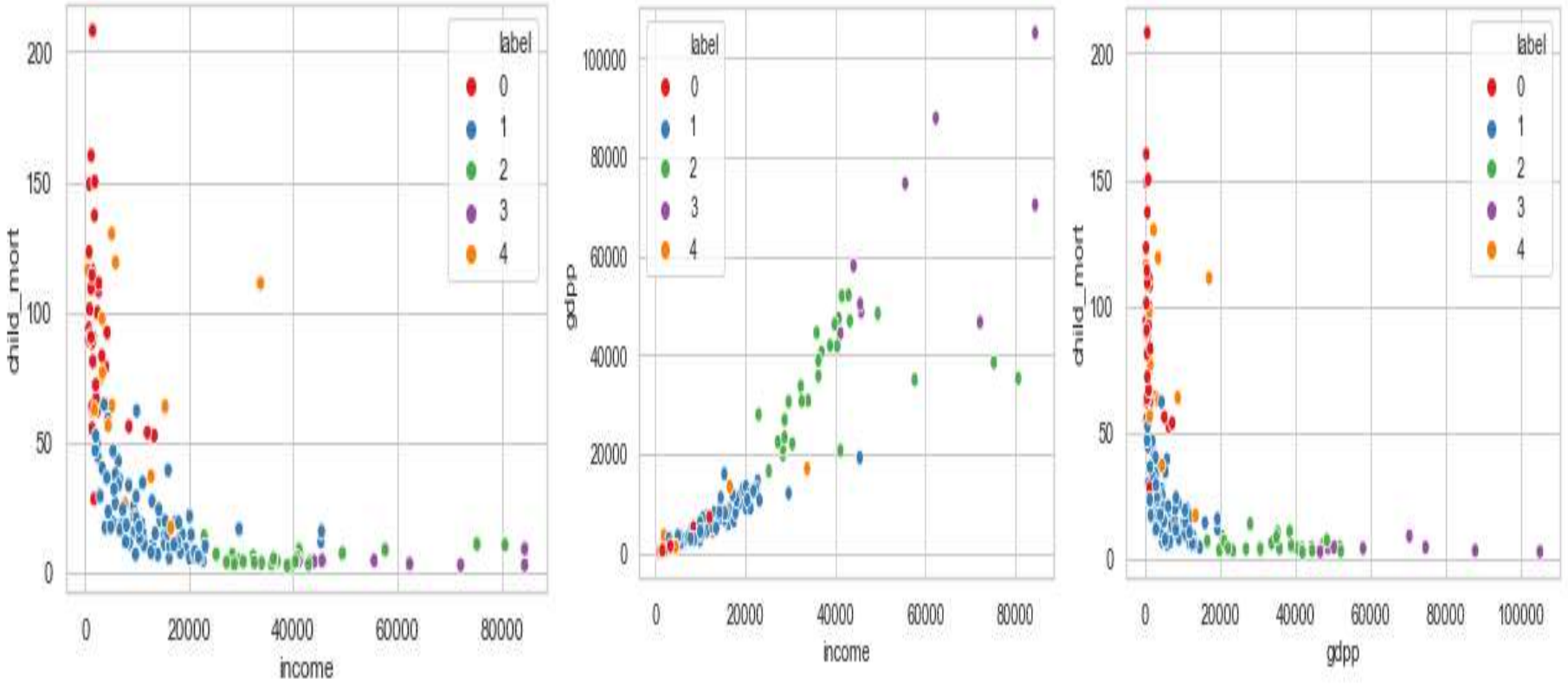
I. From the plots we can infer that the Cluster **0** has low income, low GDPP and High Child mortality .

II. Hence Cluster 0 Should be considered for the Aid .

III. Below are the Country names with respective details.

| country | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp |
|---|---|---|---|---|---|---|---|---|---|
| Burundi | 93.6 | 20.6052 | 26.7960 | 90.552 | 764.0 | 12.30 | 57.7 | 6.26 | 231 |
| Liberia | 89.3 | 62.4570 | 38.5860 | 302.802 | 700.0 | 5.47 | 60.8 | 5.02 | 327 |
| Congo, Dem. Rep. | 116.0 | 137.2740 | 26.4194 | 165.664 | 609.0 | 20.80 | 57.5 | 6.54 | 334 |
| Niger | 123.0 | 77.2560 | 17.9568 | 170.868 | 814.0 | 2.55 | 58.8 | 7.49 | 348 |
| Sierra Leone | 160.0 | 67.0320 | 52.2690 | 137.655 | 1220.0 | 17.20 | 55.0 | 5.20 | 399 |

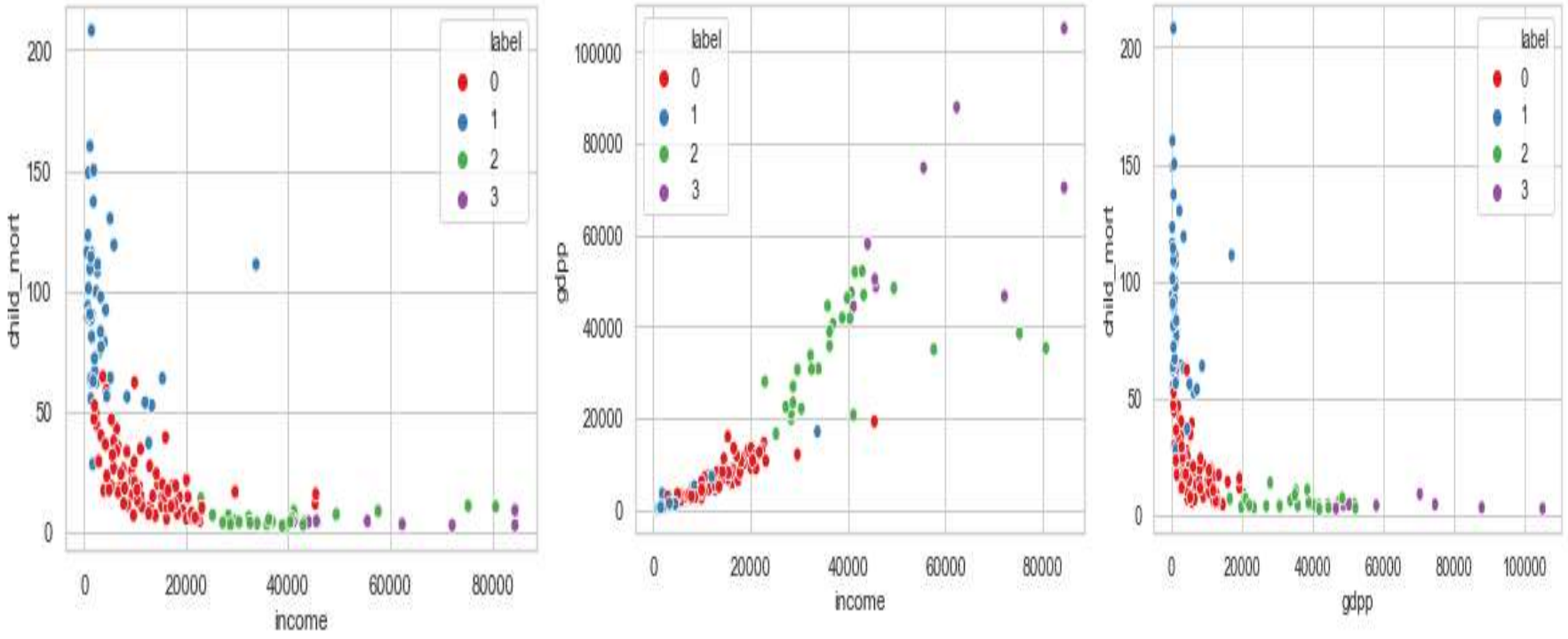
# K-Means Clustering (K=5)




**Observation**

I. The clustering with 5 clusters doesn't seem to be the best cluster as we can see a lot of data points from cluster 0 and 4 seem to be very near by and seems like they should belong to same cluster.

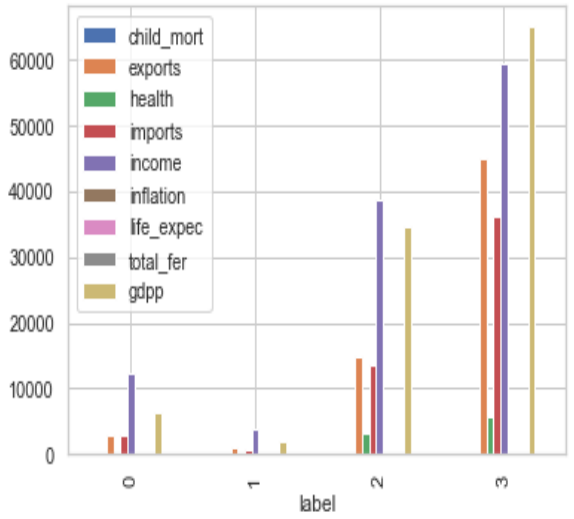
# K-Means Clustering (K=4)

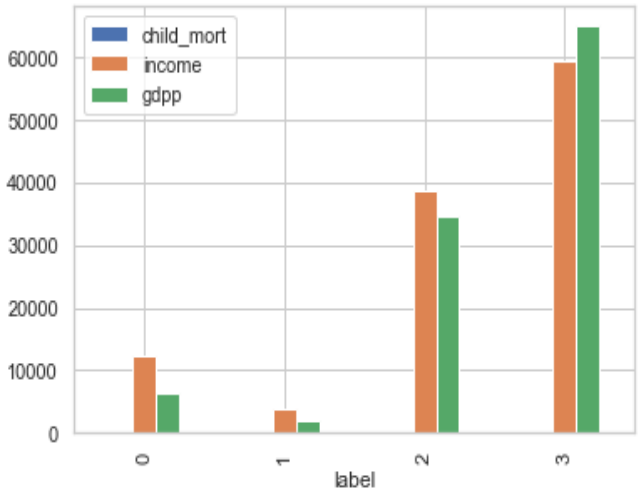


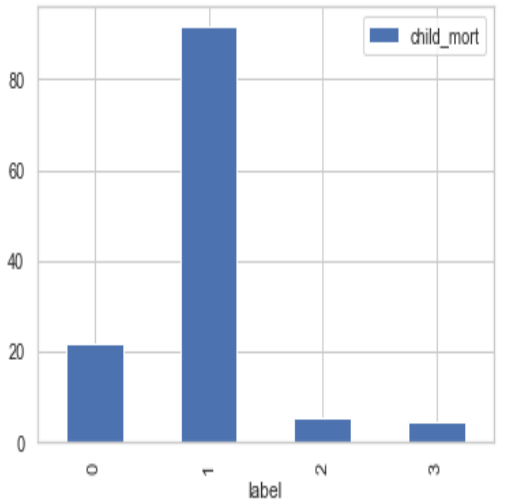

**Observation**

I. The Cluster 1 has low GDPP and high child mortality rate
II. Cluster 1 has low GDPP and low income
III. Cluster 1 has low income and High child mortality rate.

# Cluster Profiling




**Observation**

I. The Cluster 1 has low GDPP and high child mortality rate
II. Cluster 1 has low GDPP and low income
III. Cluster 1 has low income and High child mortality rate.
IV. Hence Cluster 1 is in direst need of Aid

**Top 5 which need Aid**

On the basis of low GDPP, Income and high child mortality

| | country | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp | label |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **26** | Burundi | 93.6 | 20.6052 | 26.7960 | 90.552 | 764.0 | 12.30 | 57.7 | 6.26 | 231 | 1 |
| **88** | Liberia | 89.3 | 62.4570 | 38.5860 | 302.802 | 700.0 | 5.47 | 60.8 | 5.02 | 327 | 1 |
| **37** | Congo, Dem. Rep. | 116.0 | 137.2740 | 26.4194 | 165.664 | 609.0 | 20.80 | 57.5 | 6.54 | 334 | 1 |
| **112** | Niger | 123.0 | 77.2560 | 17.9568 | 170.868 | 814.0 | 2.55 | 58.8 | 7.49 | 348 | 1 |
| **132** | Sierra Leone | 160.0 | 67.0320 | 52.2690 | 137.655 | 1220.0 | 17.20 | 55.0 | 5.20 | 399 | 1 |

# Conclusion

- On the basis of K-means clustering technique the clusters were formed.
- We have taken the following 3 features as the important socio economic factors for our strategic decisions.
  - GDPP
  - Income
  - Child Mortality
- The following are the top 5 countries which are in direst need of Aid(arranged on the basis of low GDPP, Low income and high child mortality).

| country | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp |
|---|---|---|---|---|---|---|---|---|---|
| Burundi | 93.6 | 20.6052 | 26.7960 | 90.552 | 764.0 | 12.30 | 57.7 | 6.26 | 231 |
| Liberia | 89.3 | 62.4570 | 38.5860 | 302.802 | 700.0 | 5.47 | 60.8 | 5.02 | 327 |
| Congo, Dem. Rep. | 116.0 | 137.2740 | 26.4194 | 165.664 | 609.0 | 20.80 | 57.5 | 6.54 | 334 |
| Niger | 123.0 | 77.2560 | 17.9568 | 170.868 | 814.0 | 2.55 | 58.8 | 7.49 | 348 |
| Sierra Leone | 160.0 | 67.0320 | 52.2690 | 137.655 | 1220.0 | 17.20 | 55.0 | 5.20 | 399 |