HELP International.

Fighting Poverty.

Problem Statement:

- The client is an international humanitarian NGO and they have raise around $ 10 million to help fight poverty and providing the people of backward countries with basic amenities and relief during the time of disasters and natural calamities.

- The CEO of the NGO needs to decide how to use this money strategically and effectively.

- This analysis categorizes the countries using some socio-economic and health factors that determine the overall development of the country and suggest the countries which are in dire need of aid to the CEO to focus on the most.

Road-map of our Analysis.

This analysis is broadly divided into 5 parts:

1. **Data Inspection and Data-Cleaning:** Getting the feel of Country-data.csv, identified data types, missing values, identified outliers, number of outliers and the respective countries etc.

2. **Exploratory Data Analysis**: Univariate and bivariate analysis to study the data distribution, imbalance and the mutual relationship amongst variables to derive inferences.

3. **Data Preparation**: Hopkins test to check validity of data for clustering and rescaling of variables for usage in model building.

4. **Model**: KMeans clustering for different values of K, Finding optimal number of clusters(k) using elbow method and Silhouette method and hierarchical clustering using single and complete linkage with 3 and 4 clusters to visualize and derive insights from the clusters.

5. **Conclusion:** Based on EDA and Model section, suggested the list of countries in dire need of financial aid.

Data Inspection and Data-Cleaning :

- Shape of the dataset: 167 rows and 10 columns.
- Data-types of columns.
- Identifying null values.
- Identifying and quantifying outliers.

Data-types of columns, Identifying null values :

```
#data-types of the variables
ngo.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 167 entries, 0 to 166
Data columns (total 10 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   country     167 non-null    object 
 1   child_mort  167 non-null    float64
 2   exports     167 non-null    float64
 3   health      167 non-null    float64
 4   imports     167 non-null    float64
 5   income      167 non-null    int64  
 6   inflation   167 non-null    float64
 7   life_expec  167 non-null    float64
 8   total_fer   167 non-null    float64
 9   gdpp        167 non-null    int64  
dtypes: float64(7), int64(2), object(1)
memory usage: 13.2+ KB
```

All other variables in the dataframe are either of int or float type other than 'country'.

```
#Checking missing values
ngo.isnull().sum()
```

```
country       0
child_mort    0
exports       0
health        0
imports       0
income        0
inflation     0
life_expec    0
total_fer     0
gdpp          0
dtype: int64
```

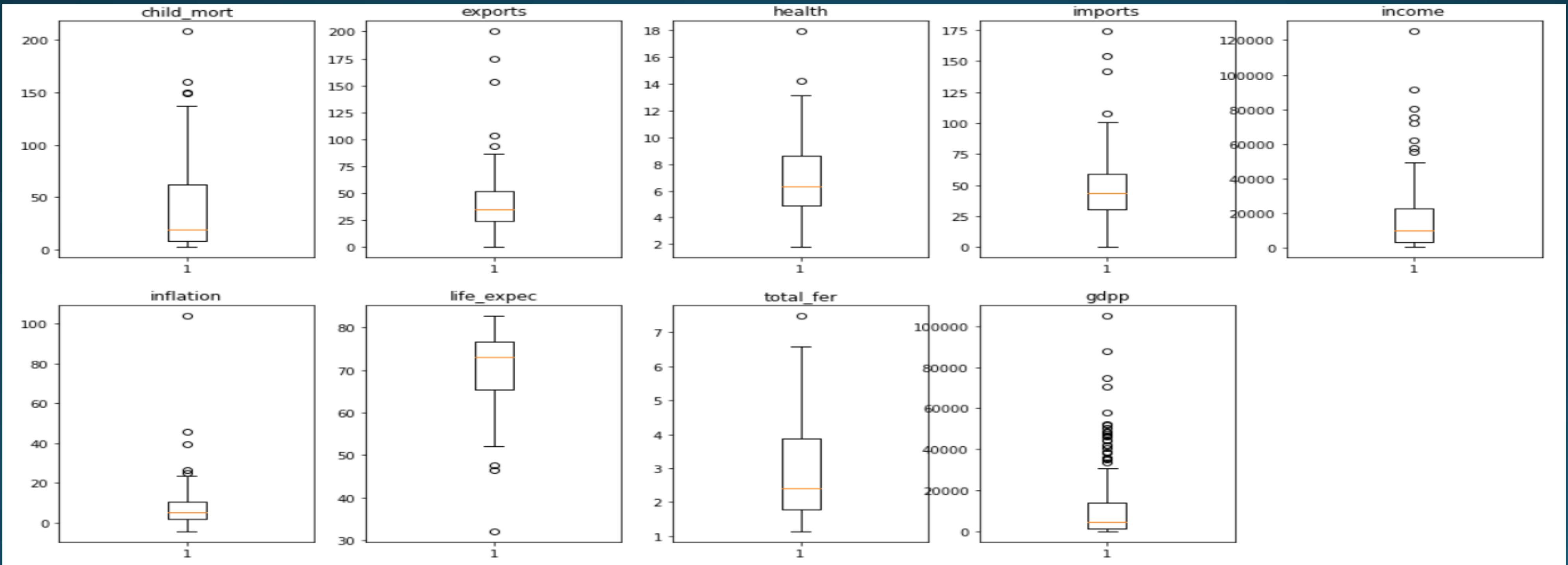
There are no null values in the dataframe.

- Analysed the data-types of the column values

- Finding out the missing values in the data: There are none.

Identifying if any Outliers in the variables.

```
#Visualizing if any outliers using boxplots.
plt.figure(figsize=(18,10))
j=1
for i in ngo_columns.drop('country'):
    plt.subplot(2,5,j)
    plt.boxplot(ngo[i])
    plt.title(i)
    j+=1
plt.show()
```



There are visible outliers in gdpp, income, exports and inflation along with very few outliers in other variables.

Number of outliers in each variable:

Defining a function to calculate 25th percentile, 75th percentile, IQR and number of outliers in each variable:

```
#Quantifying outliers: q1, q3, IQRs and number of outliers for all variables
def num_outliers(dataframe):
    q1=dataframe.quantile(0.25)
    q3=dataframe.quantile(0.75)
    iqr=q3-q1
    outlier=((dataframe < (q1 - 1.5 * iqr)) | (dataframe > (q3 + 1.5 * iqr))).sum()
    print('q1')
    print(q1)
    print('\nq3')
    print(q3)
    print('\nIQR ')
    print(iqr)
    print('\nNumber of outliers')
    return outlier
outlier=num_outliers(ngo[ngo_columns.drop('country')])
outlier
```

```
Number of outliers

child_mort     4
exports        5
health         2
imports        4
income         8
inflation      5
life_expec     3
total_fer      1
```

Based on the IQRs, q1 and q3, found the countries having outliers for each variable.

Results of Univariate and Multivariate Analyses together for better visualization:

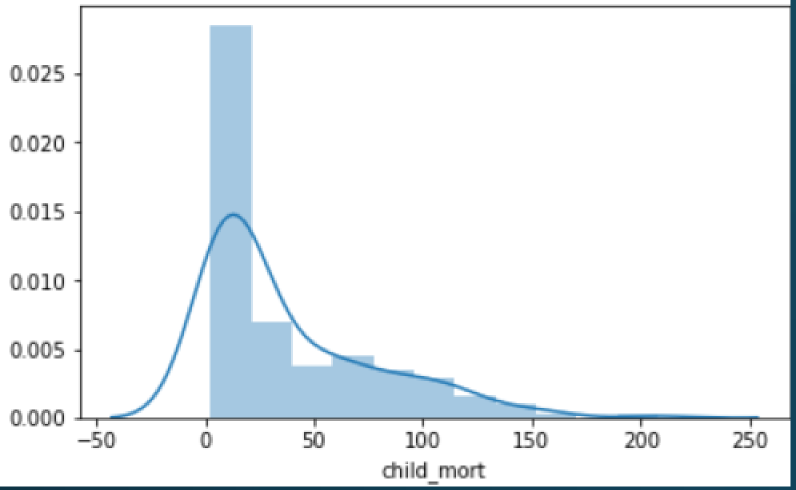
Univariate Analysis

Countries with maximum and minimum values for each variable for better univariate analysis

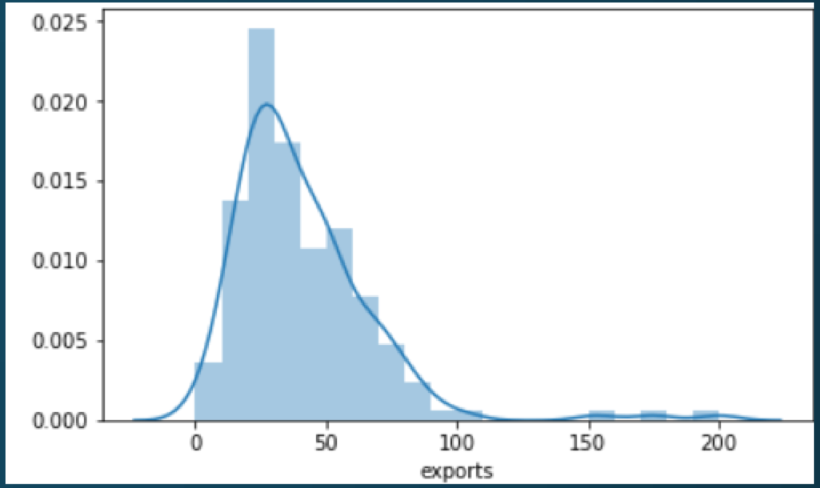
```
#Function to determine Country with maximum and minimum values for different columns
def max_min_col(col):
    dict1={}
    max_country=ngo[ngo[col]==max(ngo[col])].country
    dict1[max_country.iloc[0]]=max(ngo[col])
    min_country=ngo[ngo[col]==min(ngo[col])].country
    dict1[min_country.iloc[0]]=min(ngo[col])
    return dict1
```

```
for i in ngo[ngo_columns.drop('country')]:
    my_dict=max_min_col(i)
    print(i, my_dict)
```

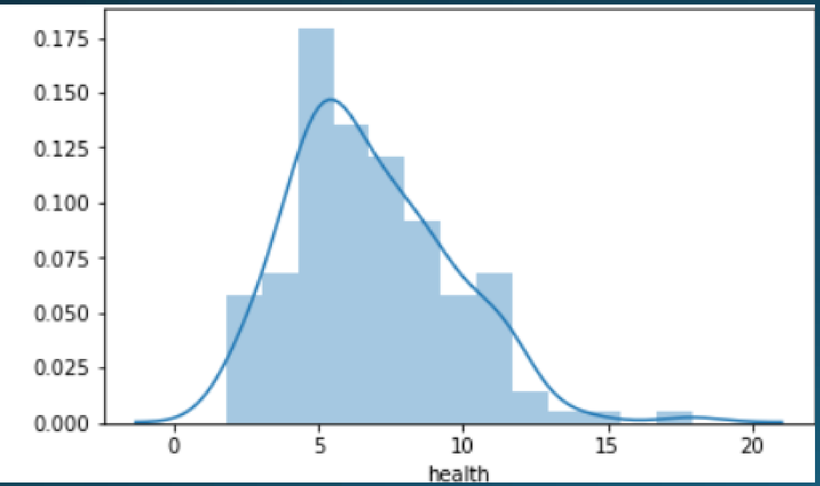
```
child_mort {'Haiti': 208.0, 'Iceland': 2.6}
exports {'Singapore': 200.0, 'Myanmar': 0.109}
health {'United States': 17.9, 'Qatar': 1.81}
imports {'Singapore': 174.0, 'Myanmar': 0.0659}
income {'Qatar': 125000, 'Congo, Dem. Rep.': 609}
inflation {'Nigeria': 104.0, 'Seychelles': -4.21}
life_expec {'Japan': 82.8, 'Haiti': 32.1}
total_fer {'Niger': 7.49, 'Singapore': 1.15}
gdpp {'Luxembourg': 105000, 'Burundi': 231}
```



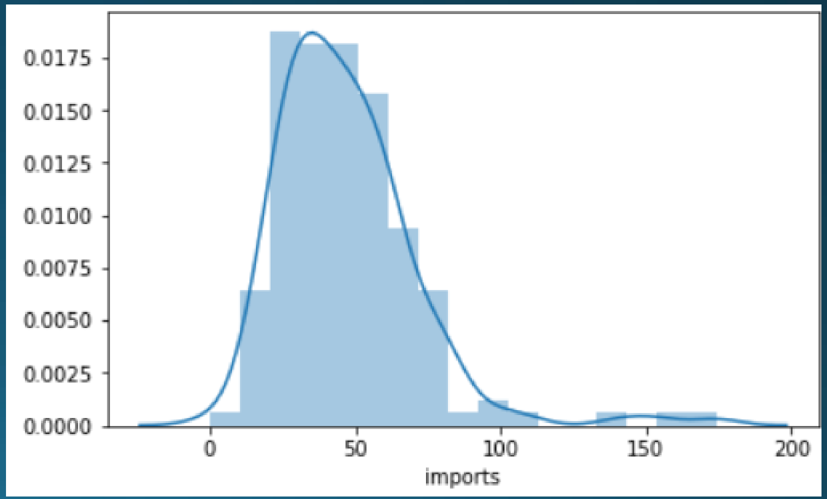
On an average 19 out of 1000 children die before the age of 5. The maximum mortality rate in a country is 208-Haiti and the minimum is Iceland-2.6. Haiti is in pretty bad condition as far as child mortality is considered.



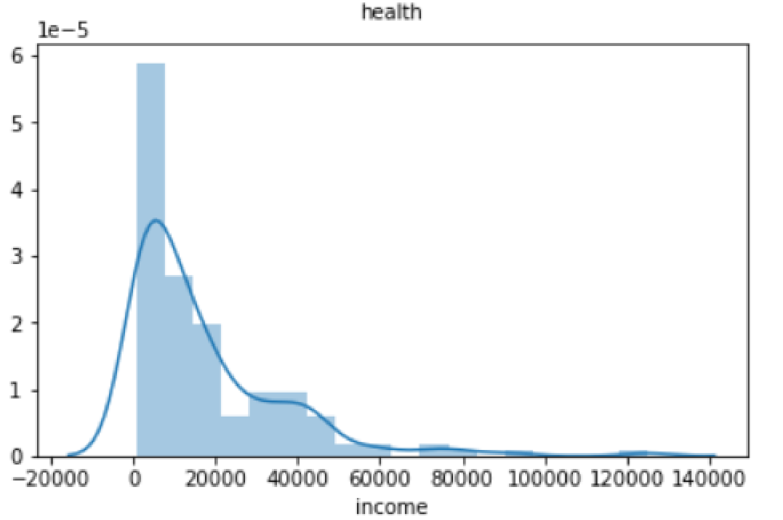
On an average exports is 35% of the GDP per capita comprises of exports accross the world. The maximum % is Singapore- 200 which i clearly an outlier(Some other sector must be nullifying this 200% export % in gdp per capita) and the minimum is Myanmar-0.109.



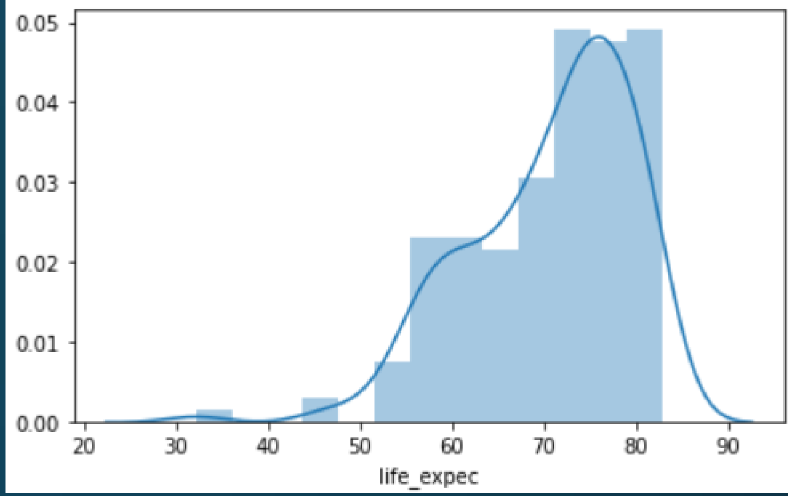
On an average Health spending is only 6.32% in gdp per capita accross the world. The maximum % is United States- 17.9% and the minimum is Qatar- 1.81%. But as per prior knowledge Qatar is an effluent country.



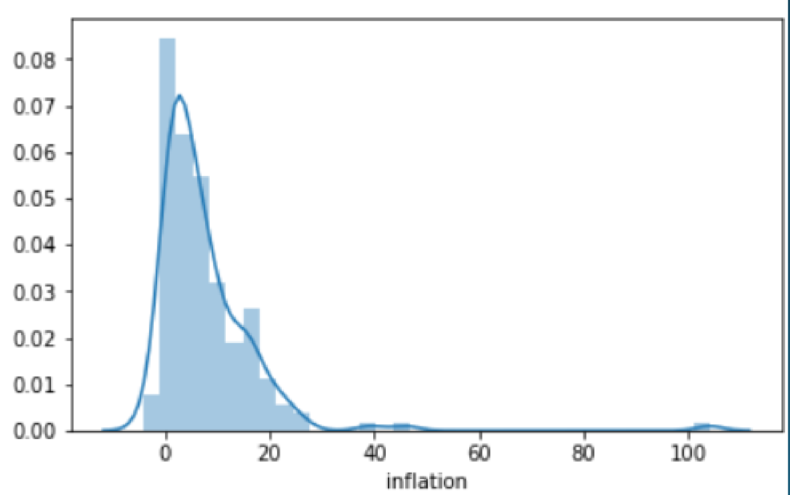
On an average imports contribute only 43% of the gdp per capita accross the world. The maximum % is Singapore- 174 which is clearly an outlier(Some other sector must be nullifying this 174% import % in gdp per capita) and the minimum is Myanmar-0.0659.



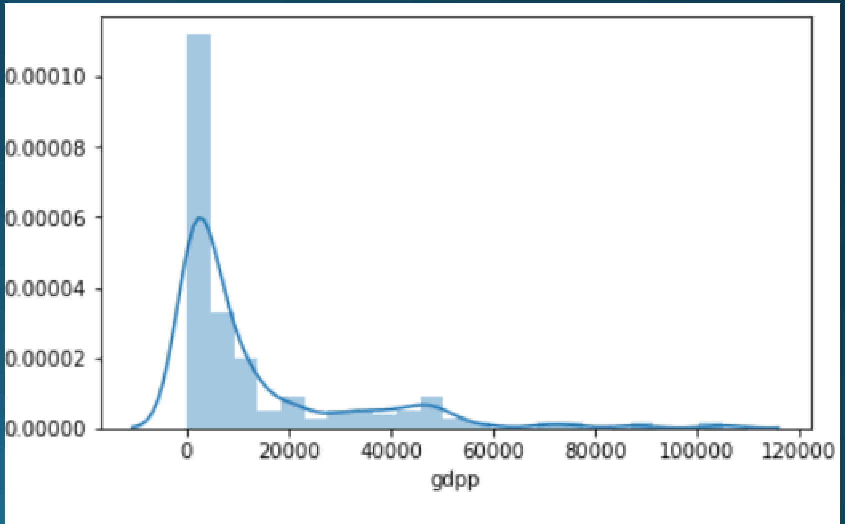
On an average the net income of a person accross the world is 9960 units. The maximum net income of a person is in Qatar- 125000 which says much about people's financial well being in Qatar and minimum is in Congo- 609 whuch is much less than the 25th percentile. People in Congo Demoocratic Republic are financially miserable.



The life expectancy average accross the world is 73.10. The minimum life expectancy is in Haiti-32.1 which is very less compared to the average life expectancy(Haiti is also in the worst situation as fara as Child Mortality is considered) and the maximum life expectancy is in Japan- 82.8



On an average the inflation accross the world is 5.39% accross the world. The maximum inflation is in Nigeria: 104% which is troublesome for them and minimum inflation is in Seychelles: -4.21%.



The average gdpp accross the world is 4660. The maximum gdpp is observed in Luxembourg: 105000 which is clearly an outlier and explains the development there and Burundi has the least 231.

Bivariate Analysis

- Analysing the relationship between the variables amongst themselves using pair-plots, correlation matrix and Variance Inflation factor.

Variance Inflation Factor:

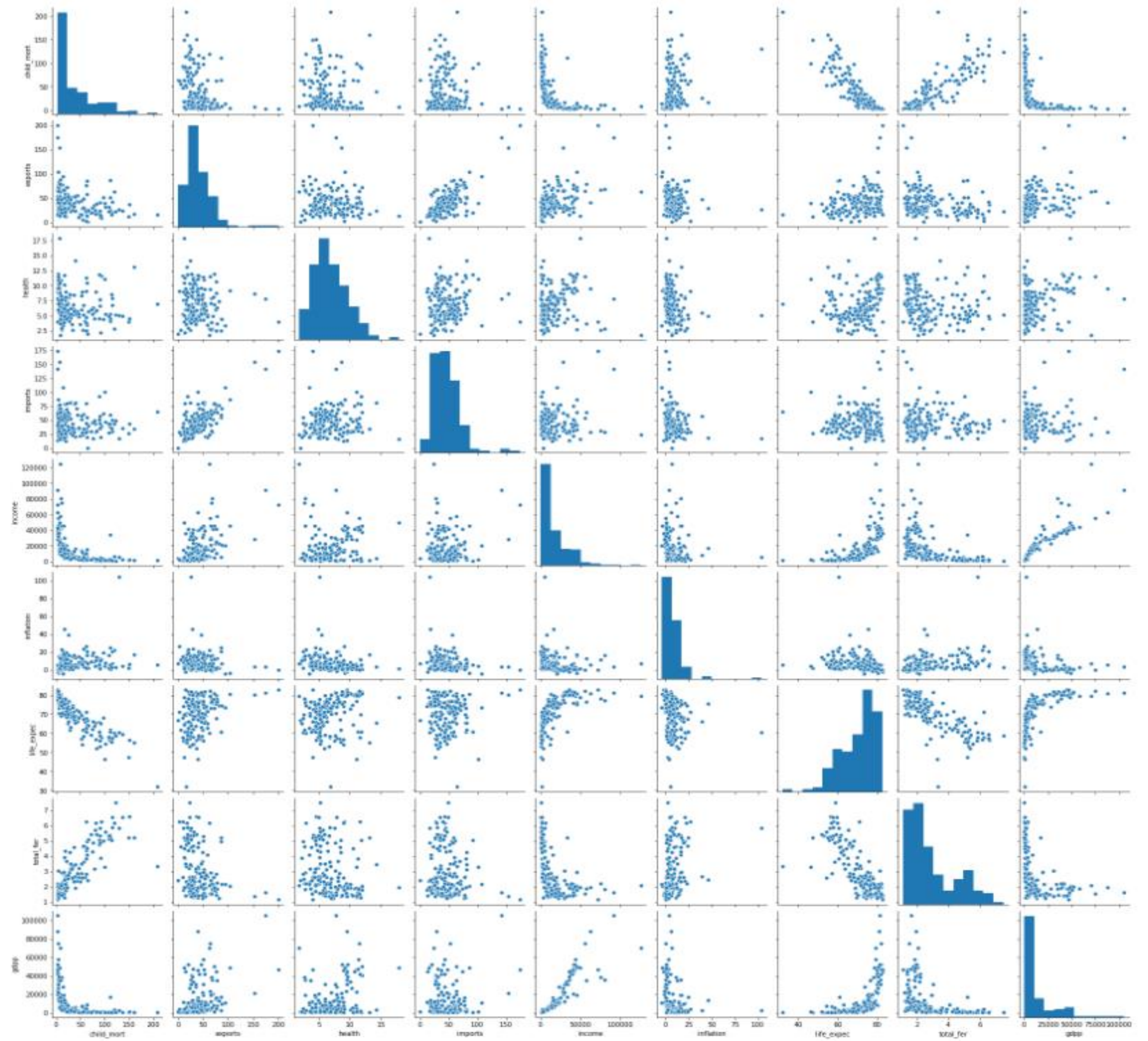
```
#Calculating the variance inflation factor and sort according to the most multi collinear variables to see the relationship of
#a variable along with all other variables.
from statsmodels.stats.outliers_influence import variance_inflation_factor
vif=pd.DataFrame()
vif['Features']=ngo_columns.drop('country')
vif['VIF']=[variance_inflation_factor(ngo[ngo_columns.drop('country')].values, i) for i in range(ngo[ngo_columns.drop('country')
vif['VIF']=round(vif['VIF'],2)
vif=vif.sort_values(by='VIF', ascending=False)
vif
```

	Features	VIF
6	life_expec	21.92
7	total_fer	17.67
3	imports	17.28
1	exports	16.08
4	income	13.41
2	health	12.42
8	gdpp	10.30
0	child_mort	8.08
5	inflation	1.99

Life expectancy and total fertility are the most collinear variables with all other variables which in a way is indicative of country's situation.

Pair-plot

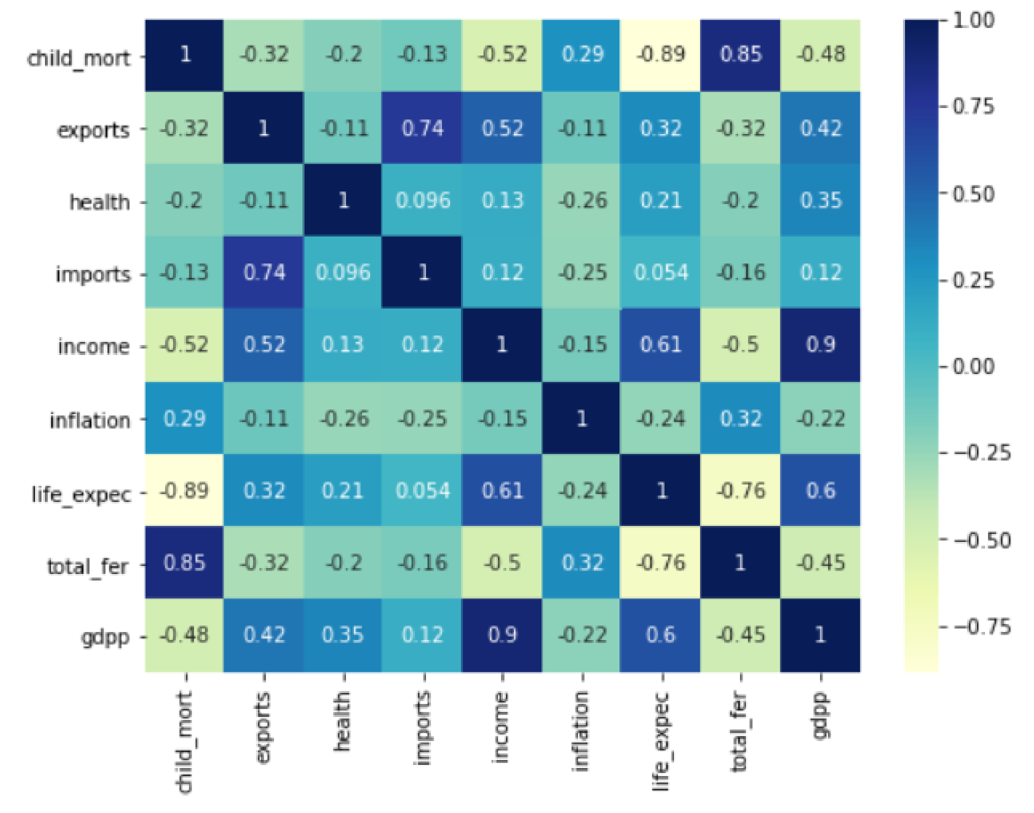
```
# Visualizing the pairplot for the entire dataframe to understand the relationship amongst them:
sns.pairplot(ngo[ngo_columns.drop('country')])
plt.show()
```



Correlation matrix

```
#Calculating the correlation matrix.
correlation_matrix=ngo[ngo_columns.drop('country')].corr()

#Plotting the correlation matrix in a heatmap.
plt.figure(figsize=(8,6))
sns.heatmap(correlation_matrix, annot = True,cmap="YlGnBu")
plt.show()
```



- Exports and imports are highly related to each other
- A person's income and country's gdpp is highly related.
- Life expectancy is inversely related to child mortality and total fertility

Data Preparation:

- **Hopkins Statistics**: To check whether our data is fit for clustering.

  Null Hypothesis: Dataset is uniformly distributed. Hence no meaningful clusters. **Average H<=0.85**

  Alternate Hypotheis: dataset is not uniformly distributed. Hence It contains meaningful clusters. **Average H > 0.85**

  where H is Hopkins Statisctic value.

  After performing the Hopkins test, average value of H > 0.85.

  Hence Null Hypothesis is rejected. There are meaningful clusters in our dataset.

- **Rescaling**: To standardize the variables using StandardScaler irrespective of the difference in their values due to different units and to optimize the modelling process.

Modelling:

- Kmeans Clustering

  (i) Finding optimal value of clusters(k): Elbow Method.

  (ii) Verifying optimal value of clusters(k): Silhouette Score.

  (iii)Fitting the model using best values of k and analyse the clusters to derive insights.

- Hierarchical clustering

  (i) Hierarchical clustering using single linkage.

  (ii) Hierarchical clustering using complete linkage with best values of k decided visualizing the dendrogram.

  (iii) Analyse the clusters to derive insights.

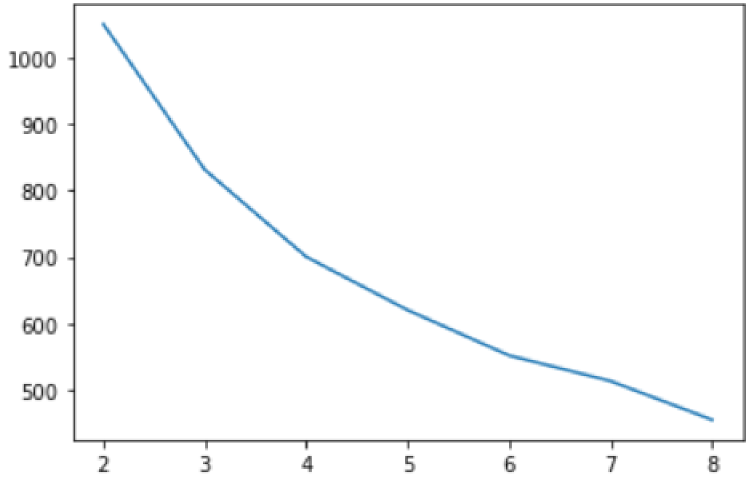
KMeans Clustering

Elbow Method:

Finding the optimal value of k using the elbow method. The average sum squared of Euclidian distances amongst the datapoints is plotted against the number of clusters.

```
#Finding the ssd for clusters iterating over number of clusters from 2 to 8.
num_clusters=[2,3,4,5,6,7,8]
ssd=[]
for i in num_clusters:
    kmeans=KMeans(n_clusters=i, init='k-means++', n_init=20)
    kmeans.fit(ngo_model)
    ssd.append(kmeans.inertia_)

#Plotting the graph of ssd- The elbow curve
plt.plot([2,3,4,5,6,7,8], ssd)
plt.show()
```



Based on elbow method, we can conclude that the optimal number of K must be 3 or 4 based on the optimal cluster patterns.

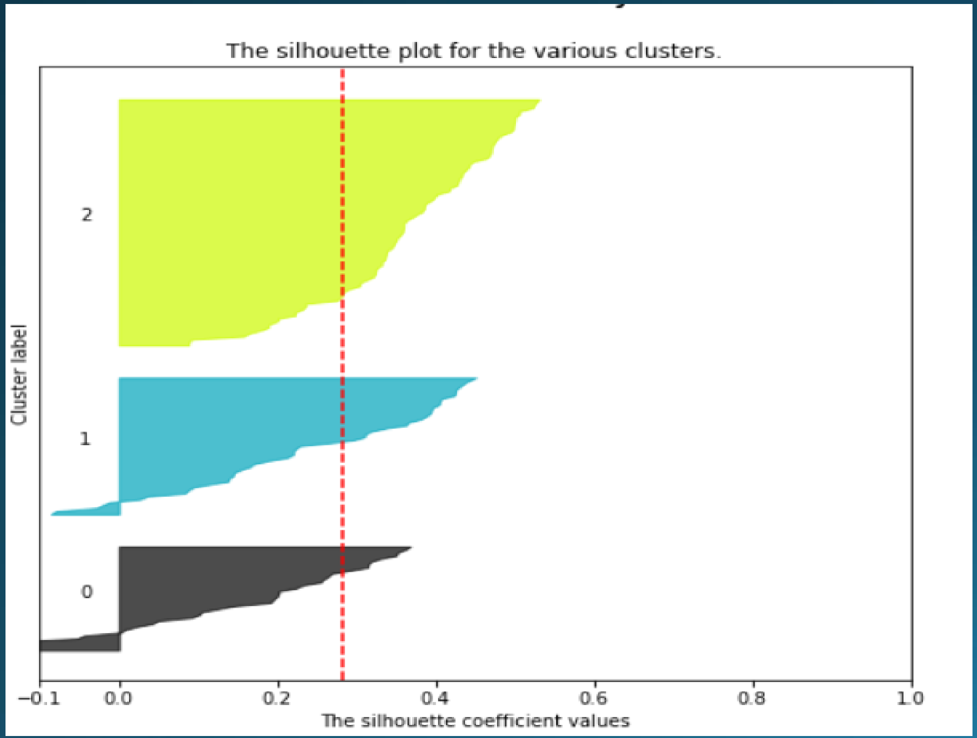
Silhouette Score and Silhouette Coefficient:

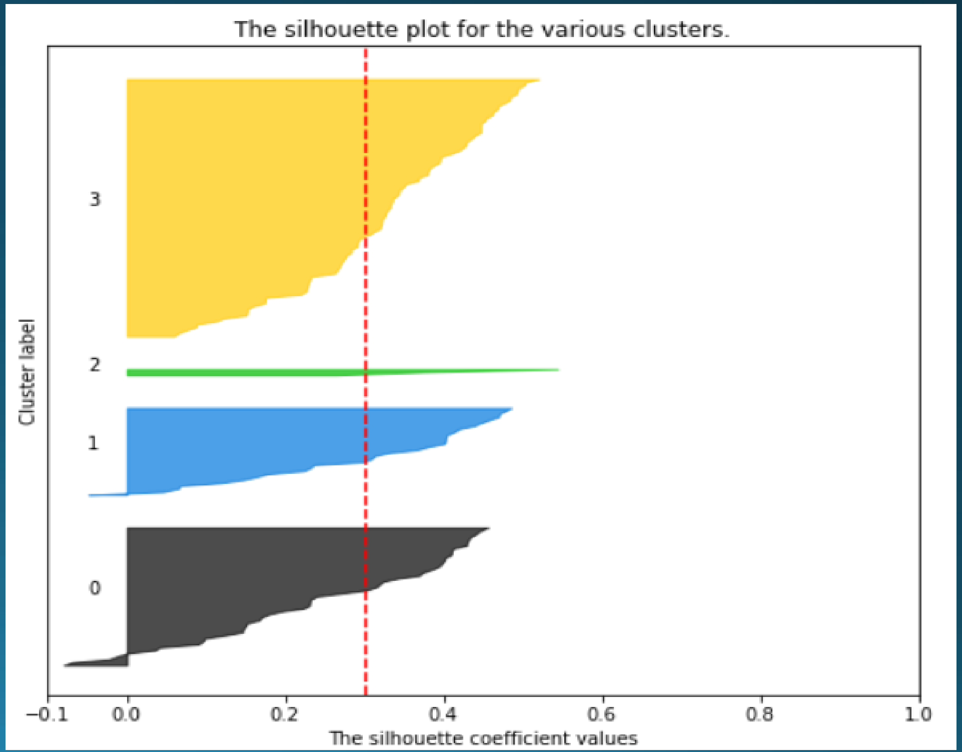
Silhouette scores for number of clusters from 2-8 are:

```
For n_clusters = 2 The average silhouette_score is : 0.2873566892140671
For n_clusters = 3 The average silhouette_score is : 0.28329575683463126
For n_clusters = 4 The average silhouette_score is : 0.301375962376881
For n_clusters = 5 The average silhouette_score is : 0.22327899566511256
For n_clusters = 6 The average silhouette_score is : 0.22412825679578482
For n_clusters = 7 The average silhouette_score is : 0.25049570233679747
For n_clusters = 8 The average silhouette_score is : 0.23627177323933837
```

<u>We can infer that the silhouette score rises from K=3 to 4 but the graph of silhouette coefficient goes negative for both k=3 and k=4. Hence will consider both k=3 and k=4 and analyse cluster profiles for both</u>.

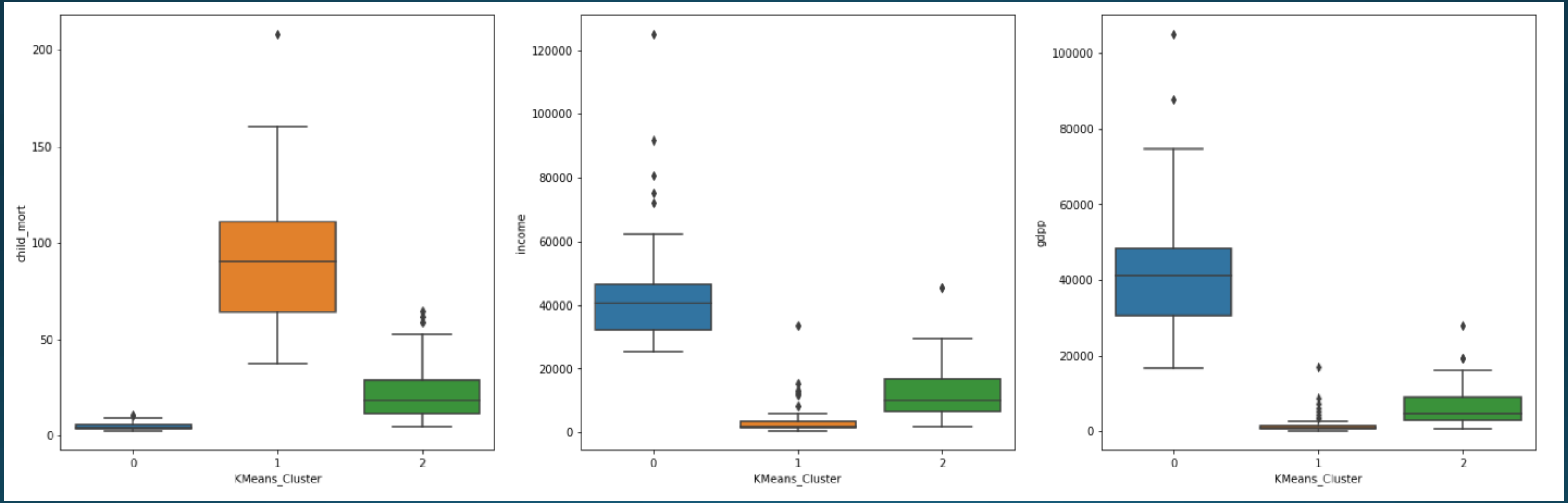
Since the Silhouette score is maximum for k=3 and 4, performing Silhouette Analysis for them.





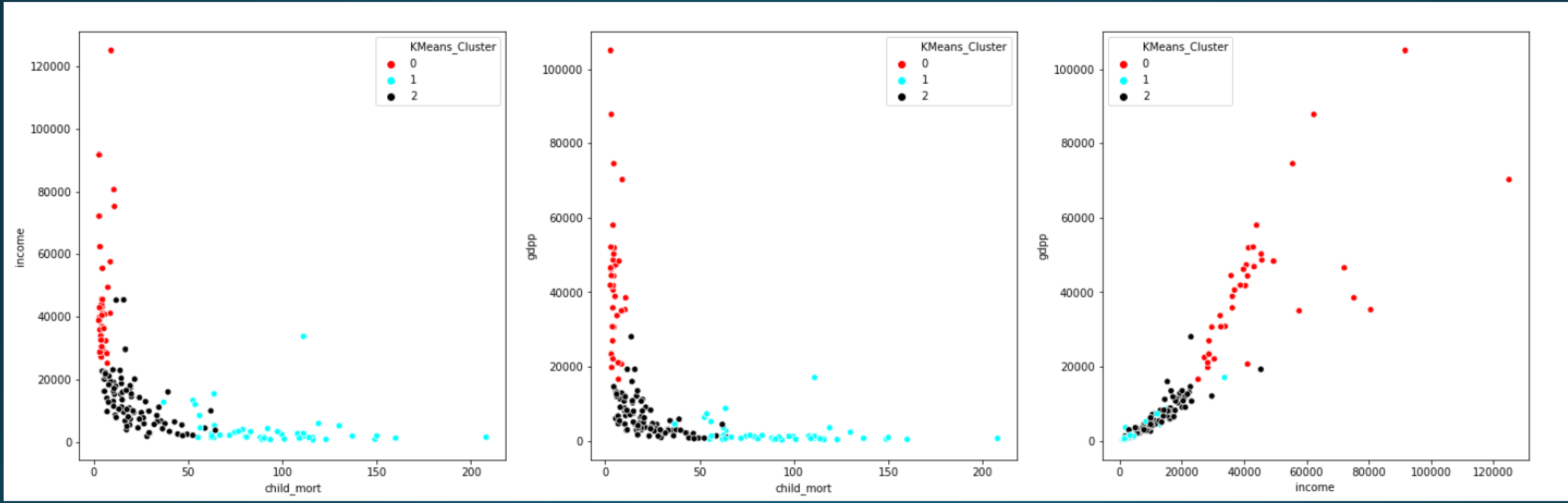
KMeans clustering with k=3

- Analysing the clusters:
- Comparing the variable values for 3 clusters.



- Cluster 0: Very low on Child mortality rate, Best on income, Best on gdpp.
- Cluster 1: High on child mortality rate, lowest on income, lowest on gdpp.
- Cluster 2: Moderate on Child mortality rate, Moderate on income, Moderate on gdpp.

Visualizing the clusters across variables Income-Child mortality, Gdpp-Child mortality and Gdpp-Income.



There is efficient clustering when visualizing them in terms of income-child mortality and gdpp-child mortality. The clusters somewhat overlap when gdpp-income is considered which can be slightly evident from the box-plots as well.

Insights:

- ❖ None of the countries in cluster 0 requires funding. They are the best performing nations.
- ❖ Countries in cluster 1 are in trouble in all terms and are in dire need of funds.
- ❖ Few countries in Cluster 2 might also be in a bad state and might need help. Further analysis will be required.

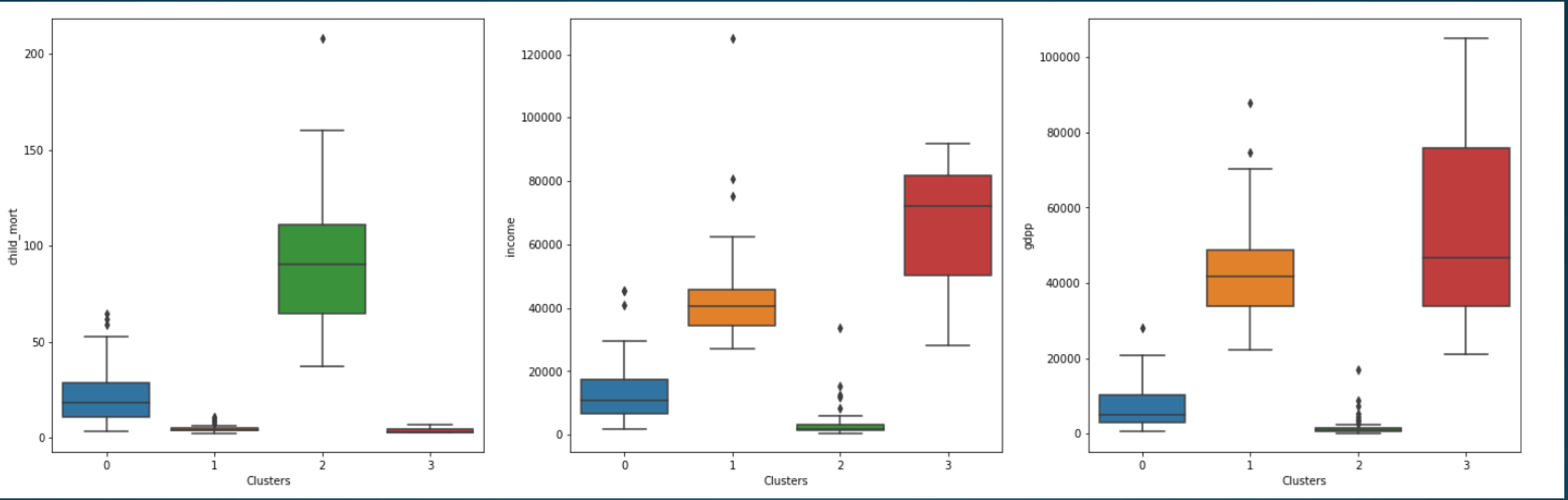
- ❖ **<u>The best performing countries</u>** based on child mortality rate, income and gdpp are Iceland, Luxembourg, Singapore, Sweden, Finland etc.

- ❖ **<u>The countries in dire need of funding</u>** according to these clusters are:
  1. Congo, Dem. Rep. : Because of its lowest income, gdpp and relatively high child mortality rate.
  2. Central African Republic: Because of its low income, gdpp and relatively higher child mortality rate.
  3. Burundi: Low income, gdpp and high child mortality.
  4. Liberia: Low income, gdpp and high child mortality.
  5. Niger: Low income, gdpp and relatively high child mortality.
  6. Haiti: Highest child mortality, low income and gdpp.
  7. Sierra Leone: Very high child mortality, low income and gdpp.

**Also: Liberia, The Central African Republic, Burundi, The Democratic Republic of the Congo, Niger form the top 5 poorest countries in the world with the lowest Gross National Income(GNI Index)*

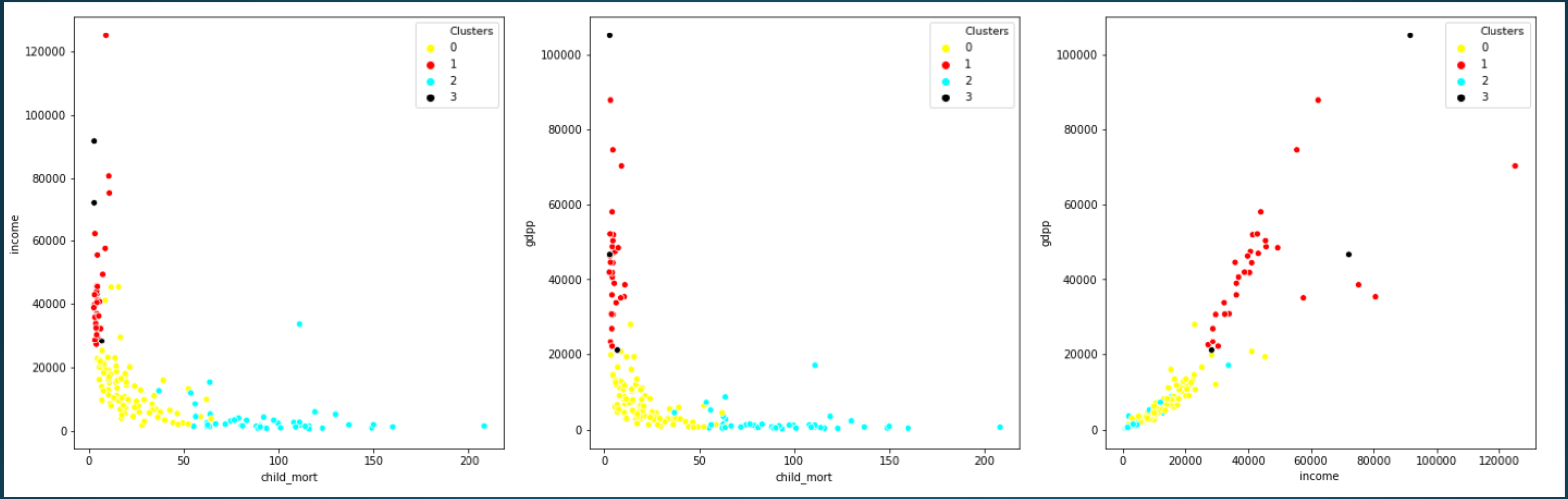
KMeans clustering with k=4

- Analysing the clusters:
- Comparing the variable values for 4 clusters.



- Cluster 2: Maximum on Child mortality rate, lowest on income, lowest on gdpp.
- Cluster 1: Low child mortality rate, moderate income, moderate gdpp.
- Cluster 0: Moderate on Child mortality rate, low on income, low on gdpp.
- Cluster 3: Lowest on child mortality rate, maximum on income, maximum on gdpp.

Visualizing the clusters across variables Income-Child mortality, Gdpp-Child mortality and Gdpp-Income.



There is efficient clustering when visualizing them in terms of income-child mortality and gdpp-child mortality. The datapoints in cluster 3 is spreadout when considering gdpp and income.

Insights:

- None of the countries in cluster 3 and 1 requires funding. They are the best and moderately performing nations respectively.
- Many countries in cluster 2 are in trouble in every terms.
- Very few countries in Cluster 0 might also be in a bad state and might need help. Further analysis will be required.

- **The best performing countries** based on child mortality rate, income and gdpp are Luxembourg, Malta and Singapore.

- **The countries in dire need of funding** according to these clusters are:
  1. Congo, Dem. Rep. : Because of its lowest income, gdpp and relatively high child mortality rate.
  2. Central African Republic: Because of its low income, gdpp and relatively higher child mortality rate.
  3. Burundi: Low income, gdpp and high child mortality.
  4. Liberia: Low income, gdpp and high child mortality.
  5. Niger: Low income, gdpp and relatively high child mortality.
  6. Haiti: Highest child mortality, low income and gdpp.
  7. Sierra Leone: Very high child mortality, low income and gdpp.

**Also: Liberia, The Central African Republic, Burundi, The Democratic Republic of the Congo, Niger form the top 5 poorest countries in the world with the lowest Gross National Income(GNI Index)*

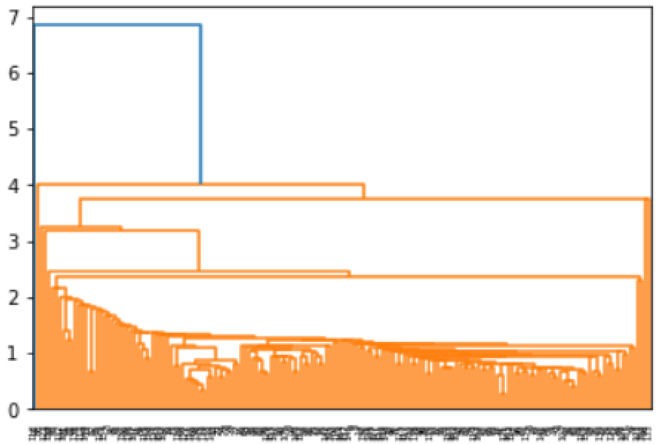
Hierarchical Modelling

Hierarchical clustering using single linkage.

- Dendrogram developed: Just a look at the dendrogram indicates that the clusters will serve no purpose in this case. Further cutting the dendrogram with 3 clusters, the labels indicate that all the data points are assigned to cluster 0 except 2 datapoints assigned to cluster 1 and 2 each.

```
# Developing the linkages with Single linkage and making the dendogram

mergings=linkage(ngo_model, method='single', metric='euclidean')
dendrogram(mergings)
plt.show()
```




```
# Cutting the dendogram to split the datapoints into 3 clusters.
hier_cluster=cut_tree(mergings, n_clusters=3).reshape(-1,)
hier_cluster
```

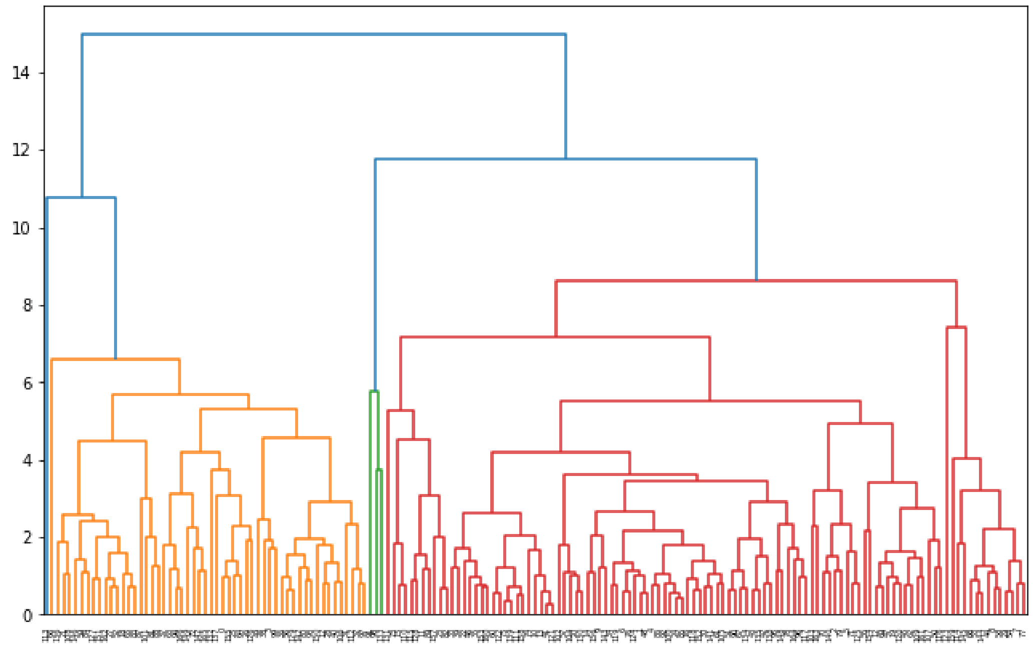
```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0])
```

- The only 2 countries separated from the data-points in the entire space are Luxembourg and Nigeria.
- Hence discarding the result developed by Hierarchial clustering using Single Linkage.

Hierarchical clustering using complete linkage. Number of clusters=3

Dendrogram developed: The dendrogram formed here looks very well defined.

```python
# Developing the linkages with Complete linkage and making the dendogram
mergings=linkage(ngo_model, method='complete', metric='euclidean')
plt.figure(figsize=(11,7))
dendrogram(mergings)
plt.show()
```




```python
# Cutting the dendogram to split the datapoints into 3 clusters.
hier_cluster=cut_tree(mergings, n_clusters=3).reshape(-1,)
hier_cluster
```

```
array([0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0,
       1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1,
       1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,
       0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0,
       0, 1, 1, 2, 1, 0, 0, 1, 1, 0, 2, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1,
       1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1,
       0, 2, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1,
       1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0])
```

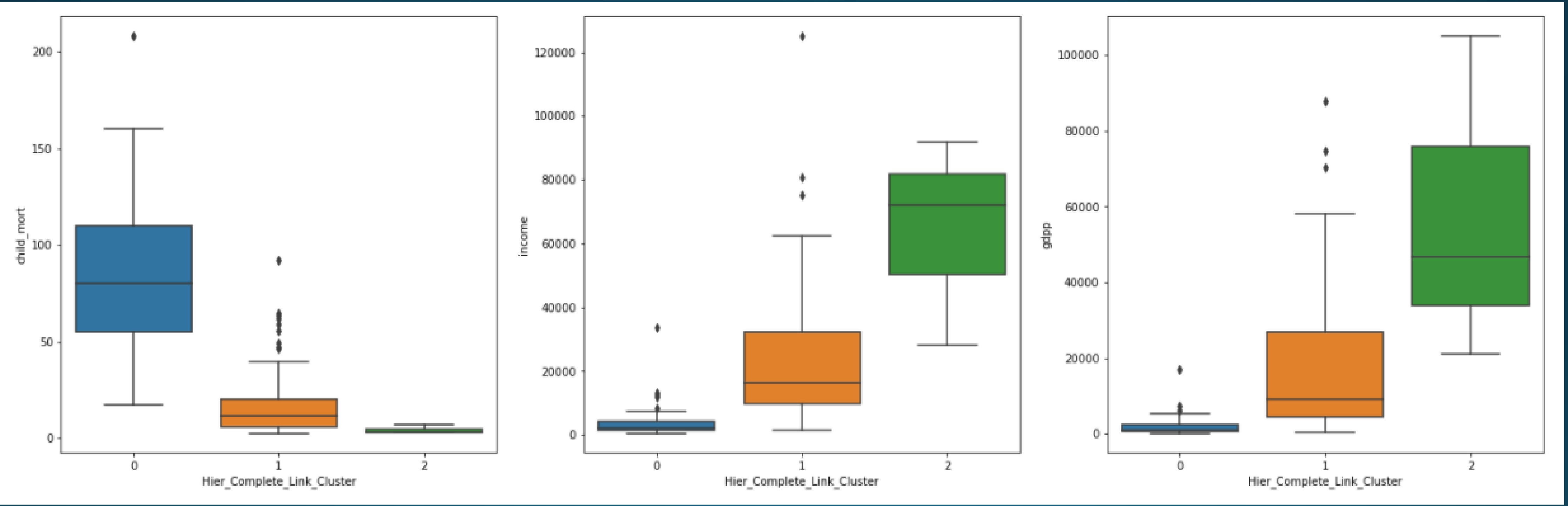
- The distribution of the data-points looks good when the dendrogram is cut into 3 clusters.
- Number of data=points in each cluster:

```python
#Finding the count of datapoints falling under each cluster.
ngo.Hier_Complete_Link_Cluster.value_counts()
```

```
1    109
0     55
2      3
```

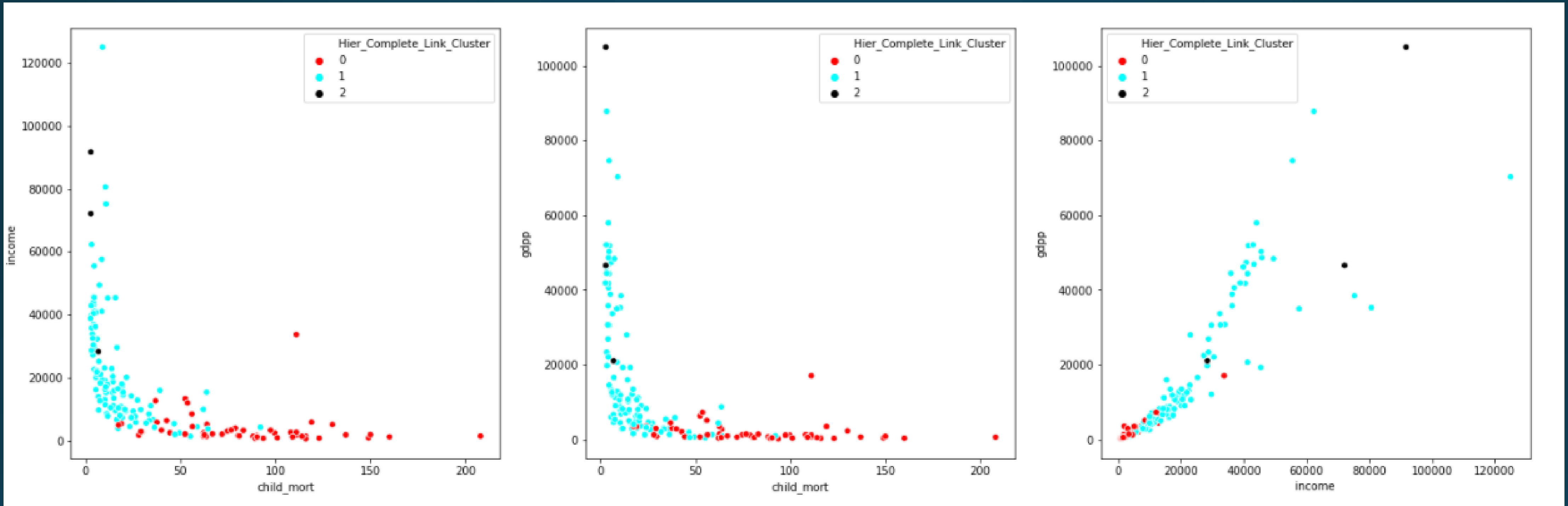
Hierarchical clustering with complete linkage and k=3

- Analysing the clusters:
- Comparing the variable values for 3 clusters.



- Cluster 0: Highest on child mortality rate, lowest on income, lowest on gdpp.
- Cluster 1: Moderate on Child mortality rate, Moderate on income, Moderate on gdpp.
- Cluster 2: Very low on Child mortality rate, Best on income, Best on gdpp.

Visualizing the clusters across variables Income-Child mortality, Gdpp-Child mortality and Gdpp-Income.



There is efficient clustering when visualizing them in terms of income-child mortality, gdpp-child mortality and gdpp-income. But the The datapoints in cluster 2 is spreadout.

Insights:

- ❖ None of the countries in cluster 2 requires funding. They are the best performing nations.
- ❖ Many countries in cluster 0 are in trouble in every terms.
- ❖ Some countries in Cluster 1 might also be in a bad state and might need help. Further analysis will be required.

- ❖ **The best performing countries** based on child mortality rate, income and gdpp are Luxembourg, Malta and Singapore etc.

- ❖ **The countries in dire need of funding** according to these clusters are:
  1. Congo, Dem. Rep. : Because of its lowest income, gdpp and relatively high child mortality rate.
  2. Central African Republic: Because of its low income, gdpp and relatively higher child mortality rate.
  3. Burundi: Low income, gdpp and high child mortality.
  4. Liberia: Low income, gdpp and high child mortality.
  5. Niger: Low income, gdpp and relatively high child mortality.
  6. Haiti: Highest child mortality, low income and gdpp.
  7. Sierra Leone: Very high child mortality, low income and gdpp.

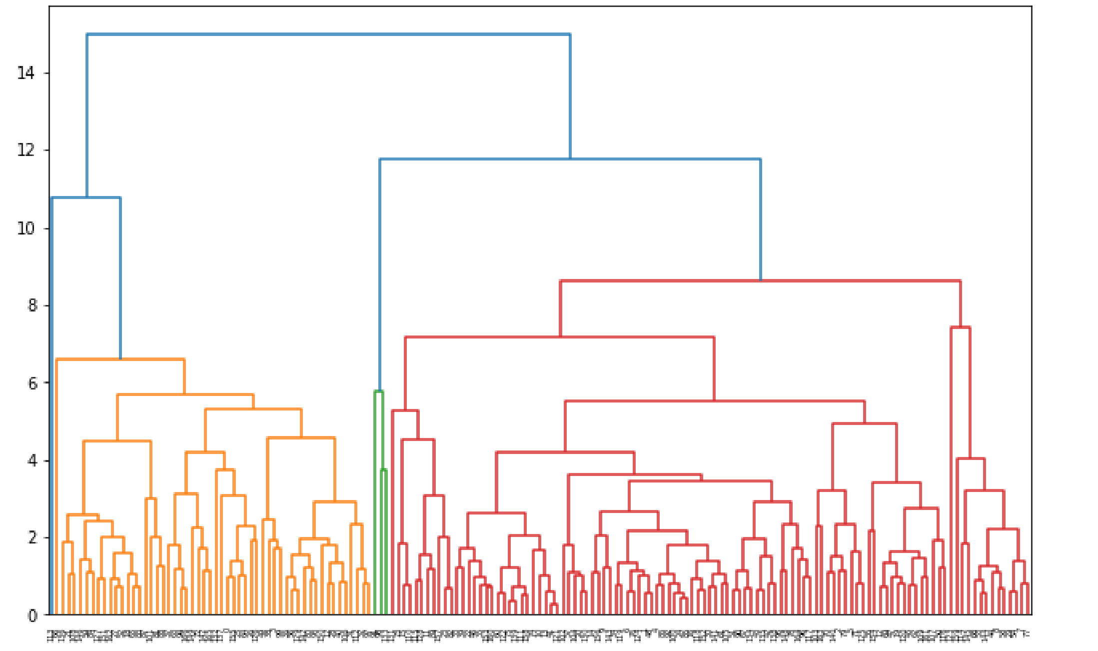
**Also: Liberia, The Central African Republic, Burundi, The Democratic Republic of the Congo, Niger form the top 5 poorest countries in the world with the lowest Gross National Income(GNI Index)*

**The countries are similar to those in KMeans clustering.*

Hierarchical clustering using complete linkage. Number of clusters=4

Using the same dendrogram developed. The distribution of the data-points looks good when the dendrogram is cut into 4 clusters but only 1 data-point is assigned to the cluster 3 which will have to be analyzed if making any sense for our analysis.

```python
# Developing the linkages with Complete linkage and making the dendogram
mergings=linkage(ngo_model, method='complete', metric='euclidean')
plt.figure(figsize=(11,7))
dendrogram(mergings)
plt.show()
```




```python
# Cutting the dendogram to split the datapoints into 4 clusters.
hier_cluster=cut_tree(mergings, n_clusters=4).reshape(-1,)
hier_cluster
```

```
array([0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0,
       1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1,
       1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,
       0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0,
       0, 1, 1, 2, 1, 0, 0, 1, 1, 0, 2, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1,
       1, 1, 0, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1,
       0, 2, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1,
       1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0])
```

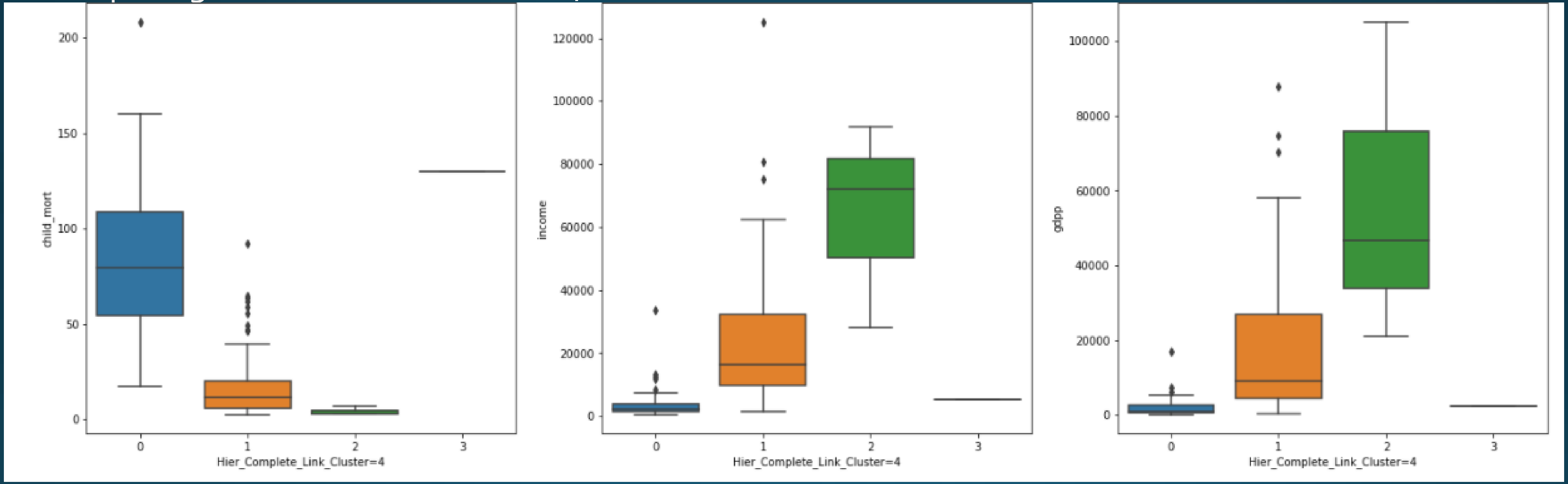
- Number of data=points in each cluster:

```python
#Finding the count of datapoints falling under each cluster.
ngo['Hier_Complete_Link_Cluster=4'].value_counts()
```

```
1    109
0     54
2      3
3      1
Name: Hier_Complete_Link_Cluster=4, dtype: int64
```

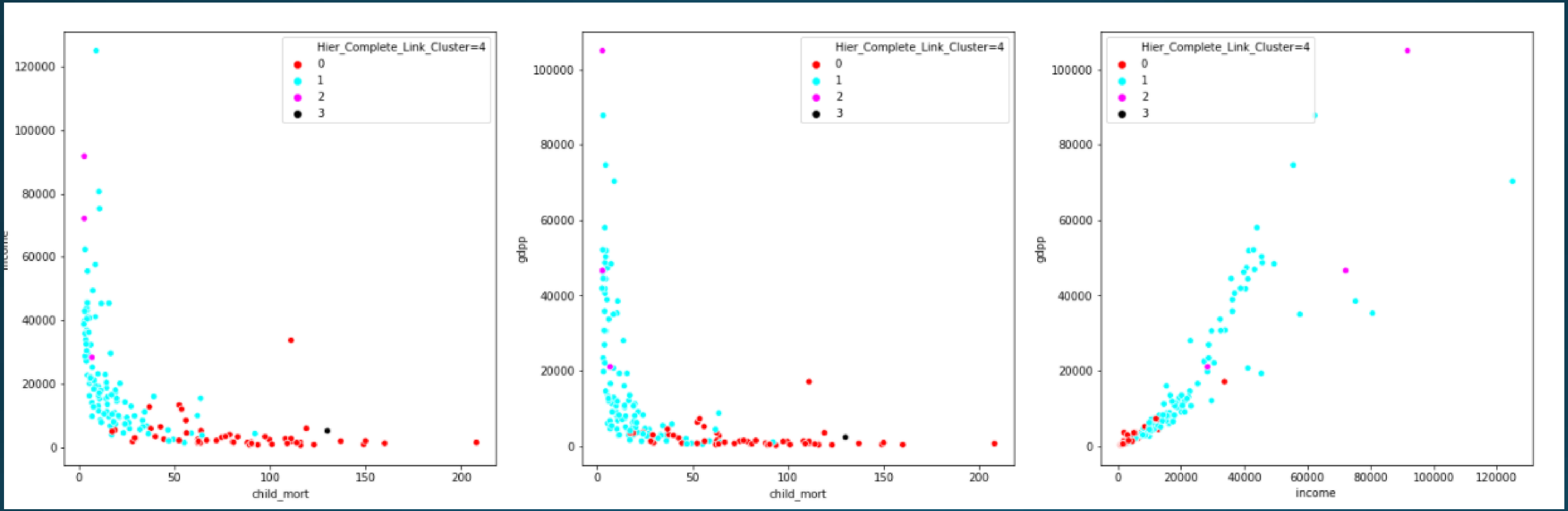
Hierarchical clustering with complete linkage and k=4

- Analysing the clusters:
- Comparing the variable values for 4 clusters.



- Cluster 0: Highest on child mortality rate, lowest on income, lowest on gdpp.
- Cluster 1: Moderate on Child mortality rate, Moderate on income, Moderate on gdpp.
- Cluster 2: Very low on Child mortality rate, Best on income, Best on gdpp.
- Cluster 3: Very high on child mortality rate, very low on income, very low on gdpp. **This cluster is very similar to cluster 0 and is making no sense to us as far as child mortality, income and gdpp are concerned.*

Visualizing the clusters across variables Income-Child mortality, Gdpp-Child mortality and Gdpp-Income.



- None of the countries in cluster 2 requires funding. They are the best performing nations.
- Many countries in cluster 0 and the only country in cluster 3 are in trouble in every terms.
- Some countries in Cluster 1 might also be in a bad state and might need help.

Hierarchical clustering with 4 clusters is serving no additional purpose to us. Hence not analyzing further.

Conclusion and Recommendations:

In our case, KMeans clustering with k=4 and Hierarchical clustering with complete linkage and number of clusters=3 gave us the best clusters- enclosing the 3 most outlier countries: Luxembourg, Malta and Singapore together and making the countries in dire need of financial aid easily identifiable.

Based on the inferences derived earlier while analysing the clusters formed by KMeans k=3 and 4 and Hierarchical clustering with complete linkage(number of cluster=3), we have got common inferences:

The countries in dire need of financial aid are:

1. Congo, Dem. Rep. : Because of its lowest income, gdpp and relatively high child mortality rate.
2. Central African Republic: Because of its low income, gdpp and relatively higher child mortality rate.
3. Burundi: Low income, gdpp and high child mortality.
4. Liberia: Low income, gdpp and high child mortality.
5. Niger: Low income, gdpp and relatively high child mortality.
6. Haiti: Highest child mortality, low income and gdpp.
7. Sierra Leone: Very high child mortality, low income and gdpp.

***Also: Liberia, The Central African Republic, Burundi, The Democratic Republic of the Congo, Niger form the top 5 poorest countries in the world with the lowest Gross National Income(GNI Index)**