



# Help International Funding Recommendations

By: Karandeep Malik

# Introduction

## Help International Funding Recommendations

### Background

- ❖ **Help International has been able to raise around \$ 10 million in funds and a decision needs to be made about how to spend this money strategically and effectively on countries that are in dire need of financial aid**

### Objective

- ❖ **Analyse the socio – economic and health factors of the given countries, determine the overall development of the countries and categorise the given countries based on that.**
- ❖ **Based on analysis, suggest a set of countries out of the given list of countries that are in dire need of financial aid.**

### Benefits

- ❖ **Recommendations will help to spend the raised money on operational projects in these countries**

# Problem Solving Methodology

## Business and Data Understanding

- 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. It runs a lot of operational projects from time to time along with advocacy drives to raise awareness as well as for funding purposes.
- After the recent funding programmes, they have been able to raise around \$ 10 million. Now the NGO needs to decide how to use this money strategically and effectively

## Data Pre-Processing

- Data was sourced from Upgrad and was cleaned
  - ❖Outliers were checked and treated based on problem statement
  - ❖Checked for missing values, garbage values, duplicate rows

## PCA, K Means Clustering and Hierarchical Clustering

- PCA was carried out on the data set to reduce the dimensionality and get components capturing maximum variance. Based on scree plot and recommendations 4 Principal Components were selected that captured close to 85% variance
- K Means Clustering was carried out on the Principal Components. The value of K chosen after analysing results of Elbow Curve method and Business Justification was 3.
- Hierarchical Clustering was carried out on the Principal Components. The value of no of clusters was chosen as 5 after seeing the formation of clusters

## Conclusion and Recommendations

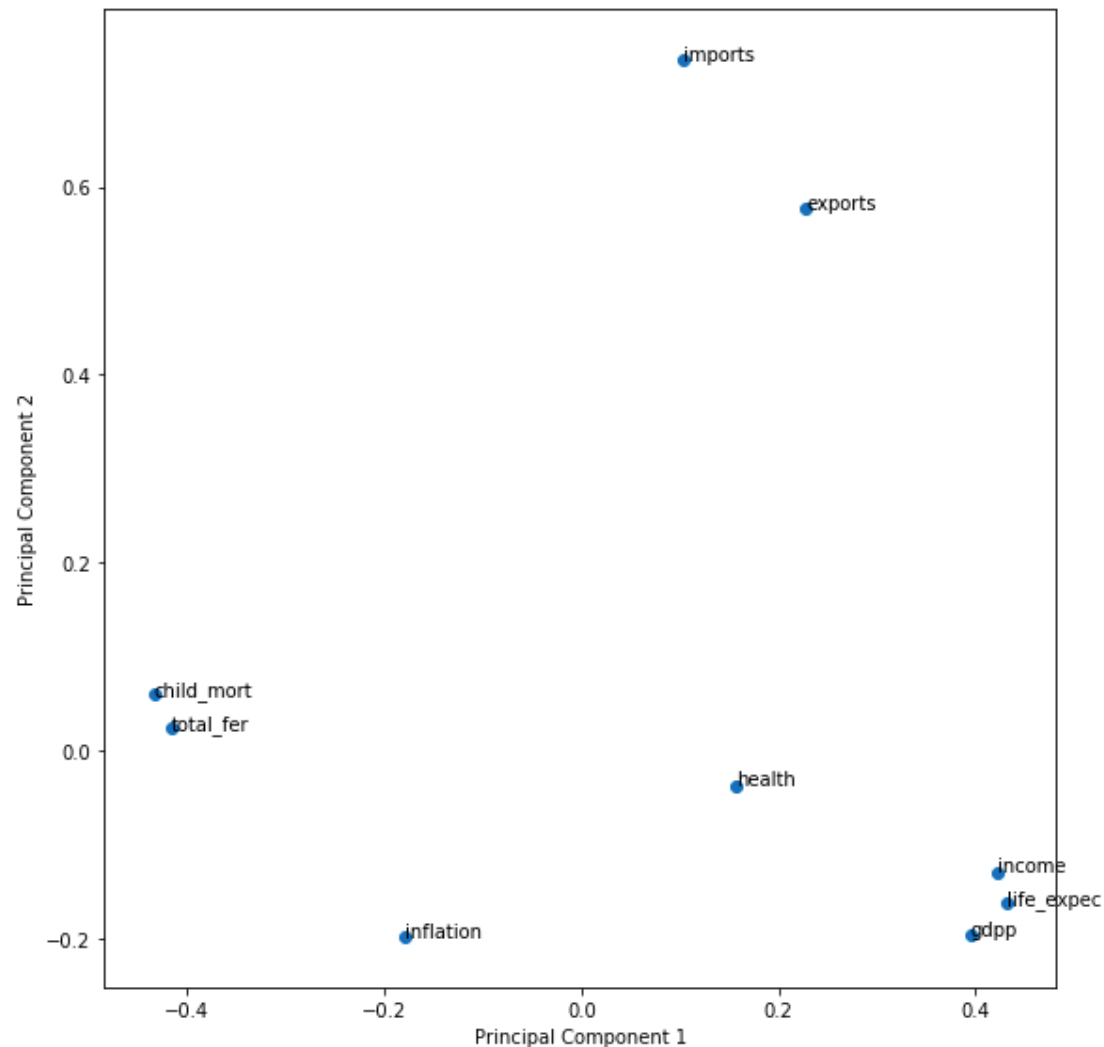
- In K Means clustering the cluster 2 was found to represent the countries with lower life expectancy, higher child mortality rate, lower gdpp and lower net income per person. This cluster contained countries like **Afghanistan, Benin, Burundi, Burkina Faso, Central African Republic, Chad, Congo Democratic Republic etc**
- In Hierarchical clustering the cluster 0 was found to represent the under developed countries containing countries like etc. **Afghanistan, Benin, Burundi, Central African Republic, Chad, Congo Democratic Republic .**
- The recommended set of countries that are in dire need of aid are: - **Afghanistan, Central African Republic, Burundi, Congo Democratic Republic, Haiti, Sierra Leone, Niger et**

## Quick Summary

### Key Statistics

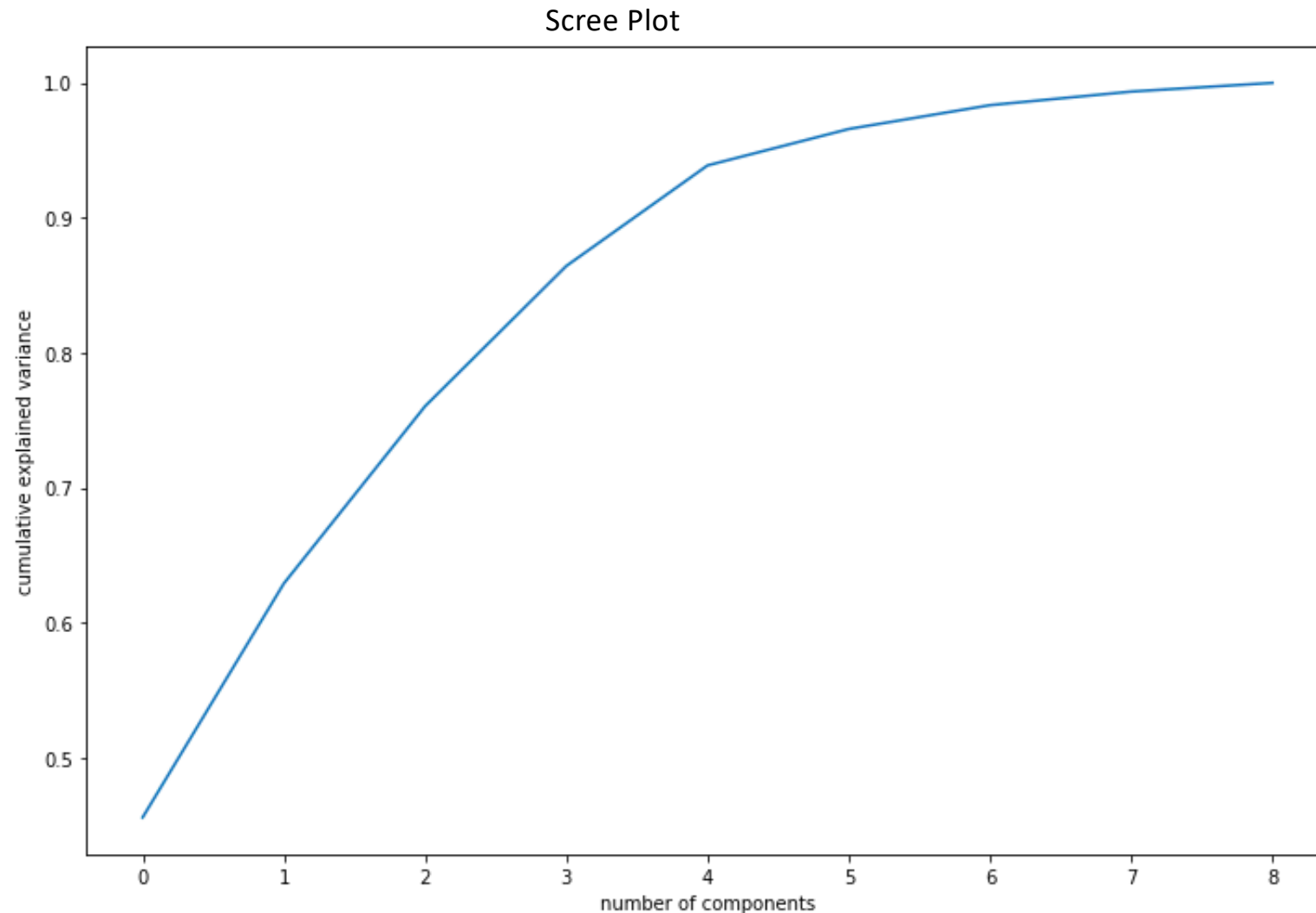
- Total No Of Countries Data in Dataset :- **167**
- The Socio Economic and Health Factors Analysed :- **child mortality, exports as percentage of GDP, health spending as percentage of GDP, imports as percentage of GDP, net income per person, annual growth rate of total GDP, life expectancy, total fertility rate, GDP per capita**
- Average Net Income : - **17144**
- Average Life Expectancy :- **70 yrs**
- Average GDP per Capita :- **12964**
- Average Child Mortality Rate :- **38.27**

# Principal Component Analysis (PCA)



The feature loading in PCA1 was mostly dominated by life expectancy (life\_expec), net income (income), gdp per capita (gdpp), exports and health expenditure. The feature loading in PC2 was dominated by imports, exports and child mortality (child\_mort)

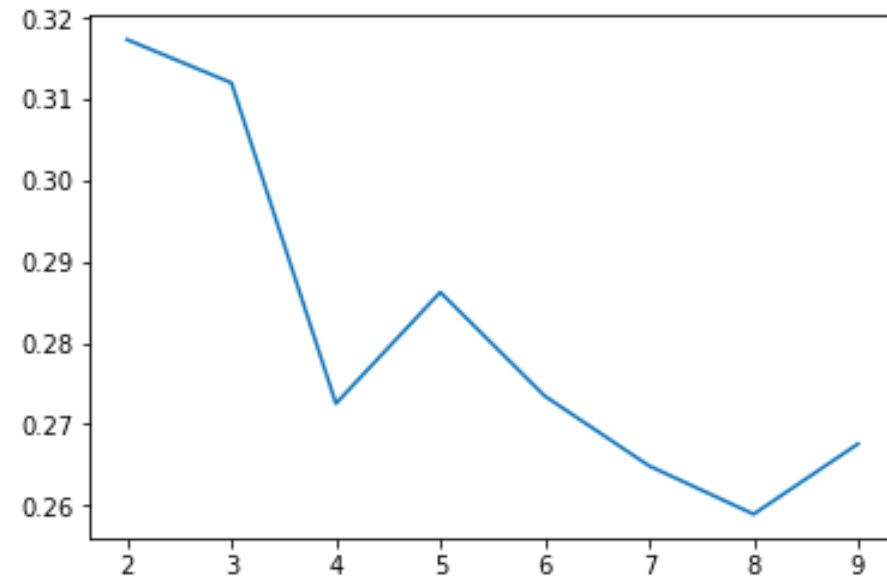
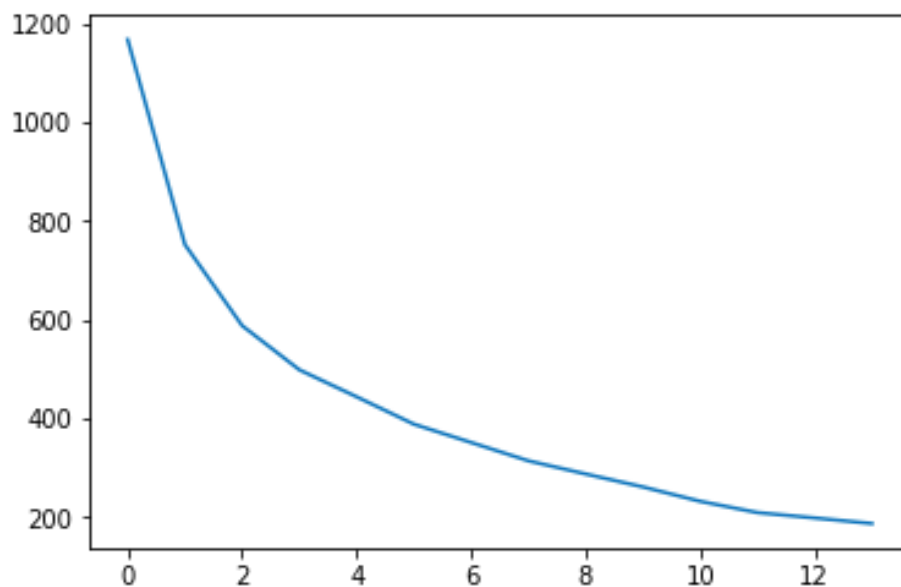
# Principal Component Analysis (PCA)



From Scree plot it can be seen that 4 components seems a good choice as it captures around 85 % cumulative variance and reduces dimensionality to half.

# K Means Clustering

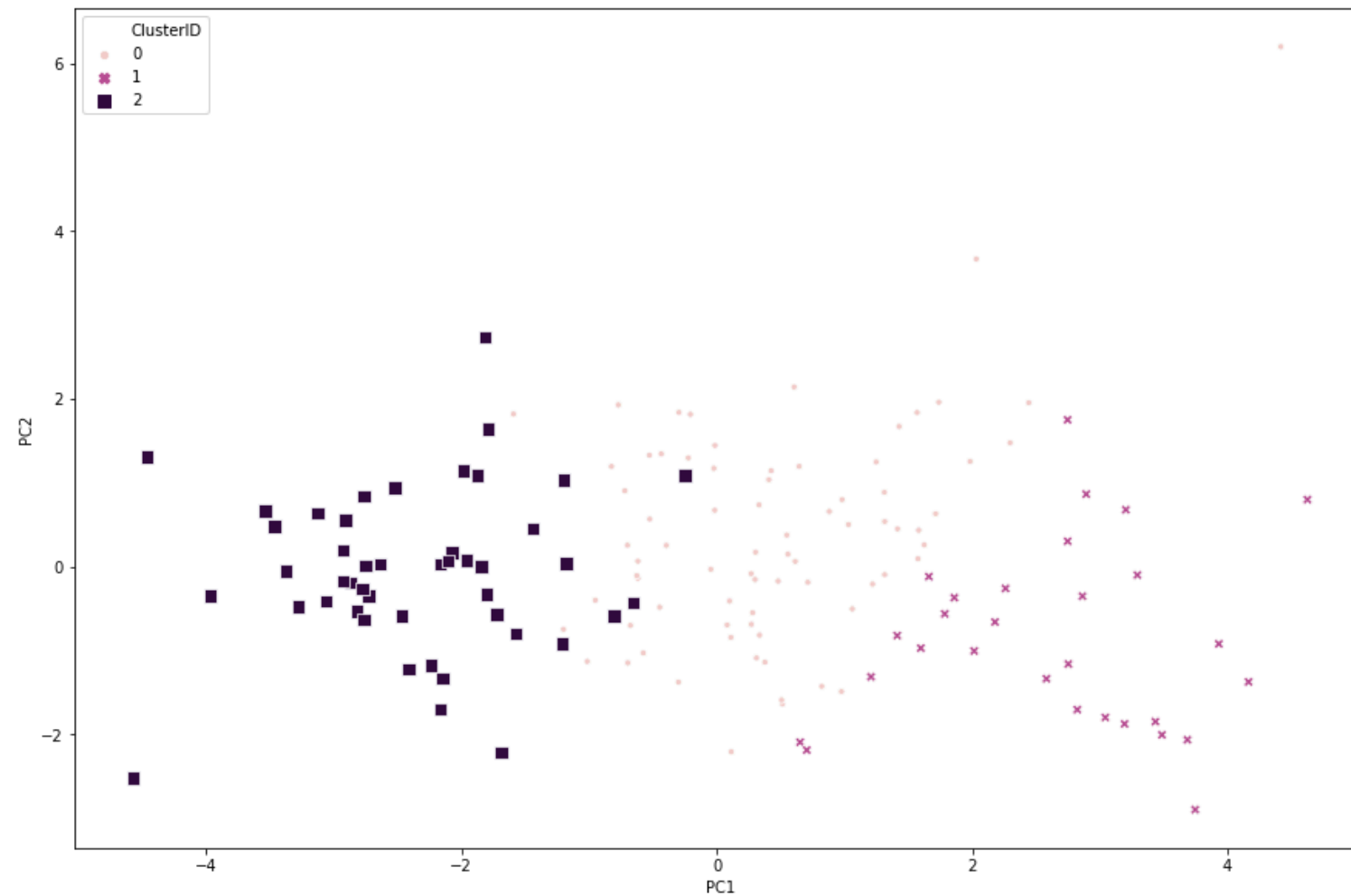
## Elbow curve Method and Silhouette Coefficient Analysis



From the silhouette analysis method we see that 2 or 3 could be the right value of clusters. We could cluster countries into Under Developed, Developing and #Developed Countries based on problem statistics, so cluster value of 3 makes sense.

# K Means Clustering

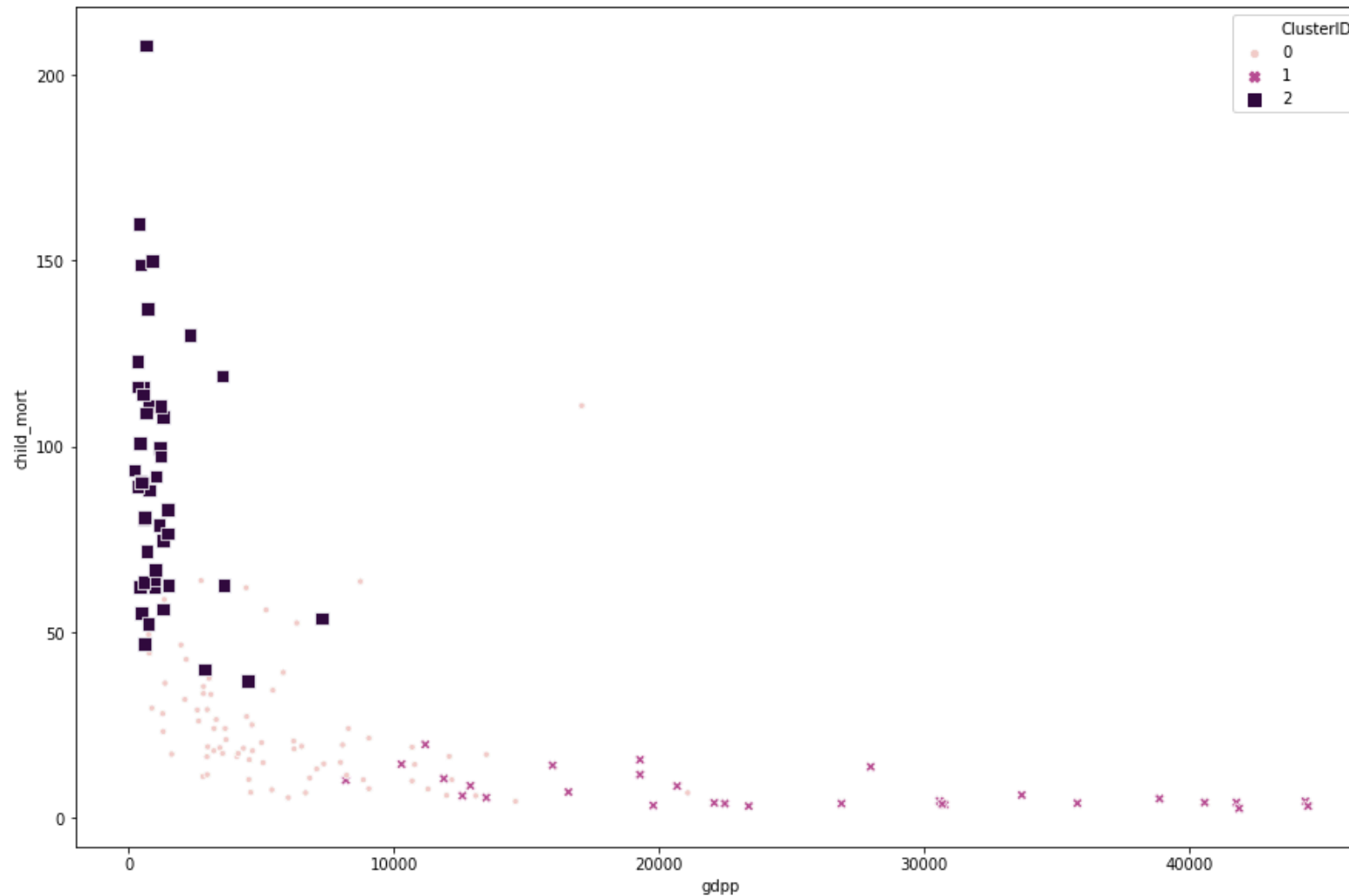
Clusters Formation





# K Means Clustering

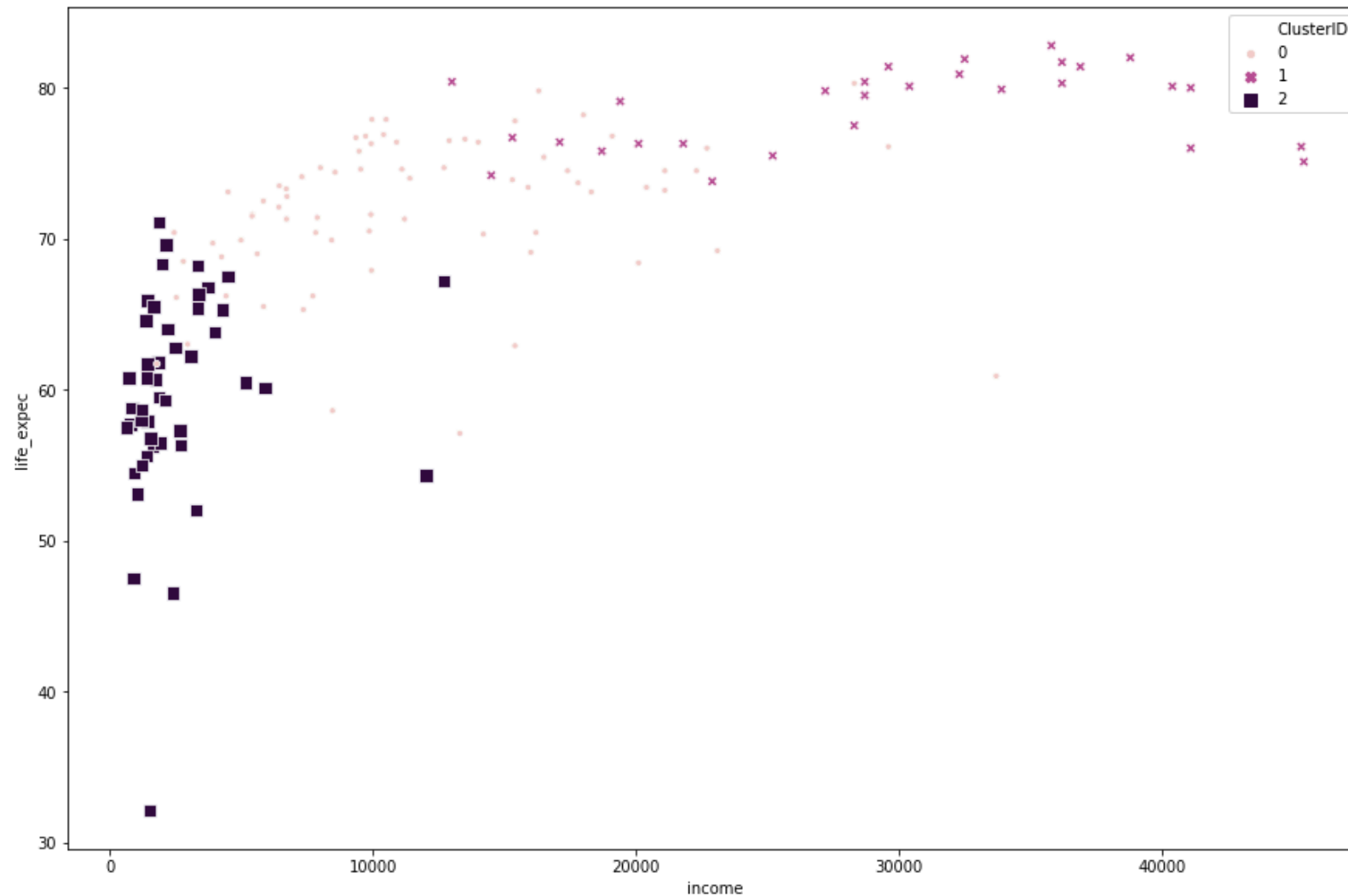
Clusters association with child mortality and per capita



The countries in cluster 2 seem to have a high child mortality rate and low gdp per capita

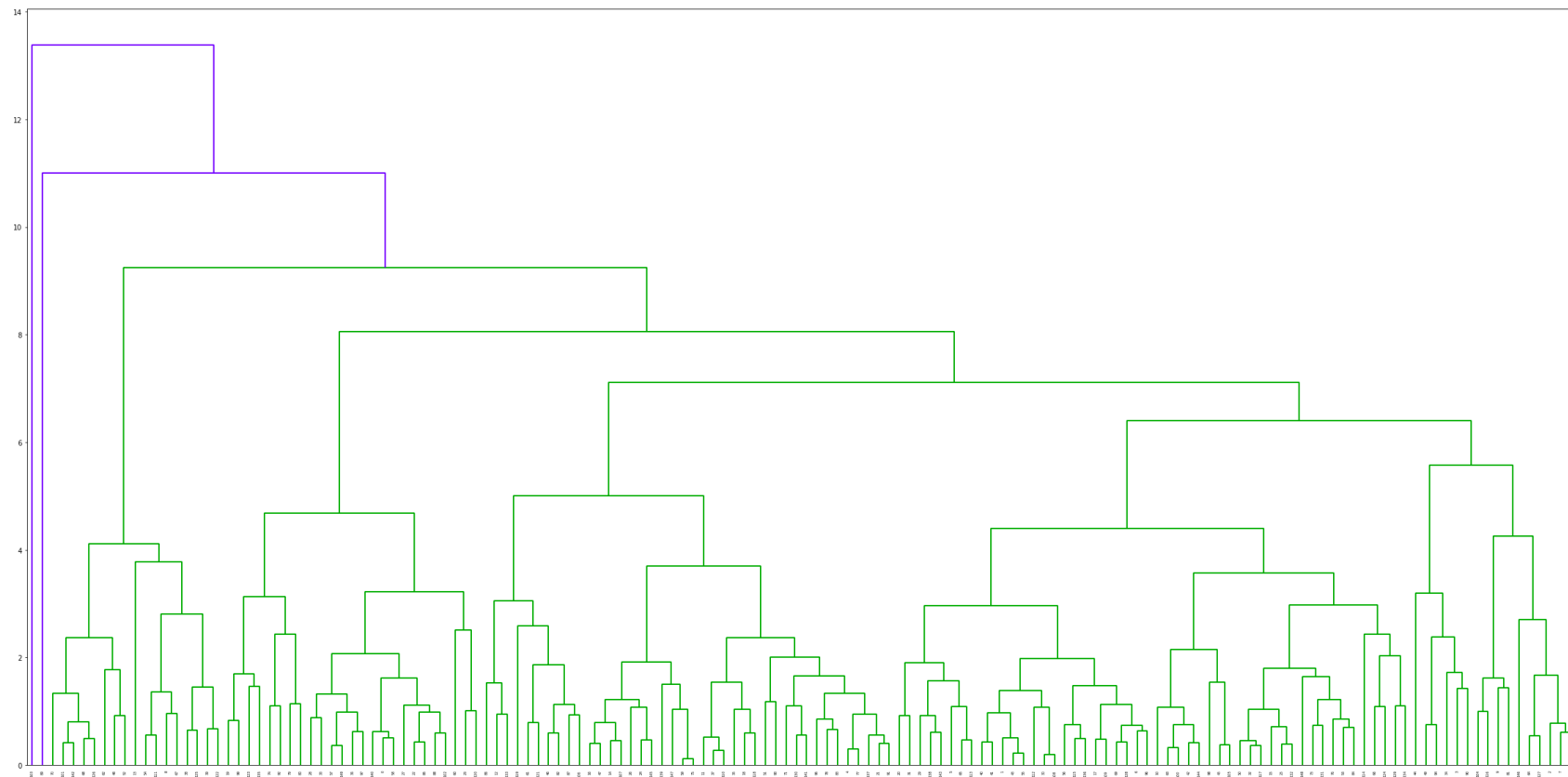
# K Means Clustering

Clusters association with life expectancy and low income per capita



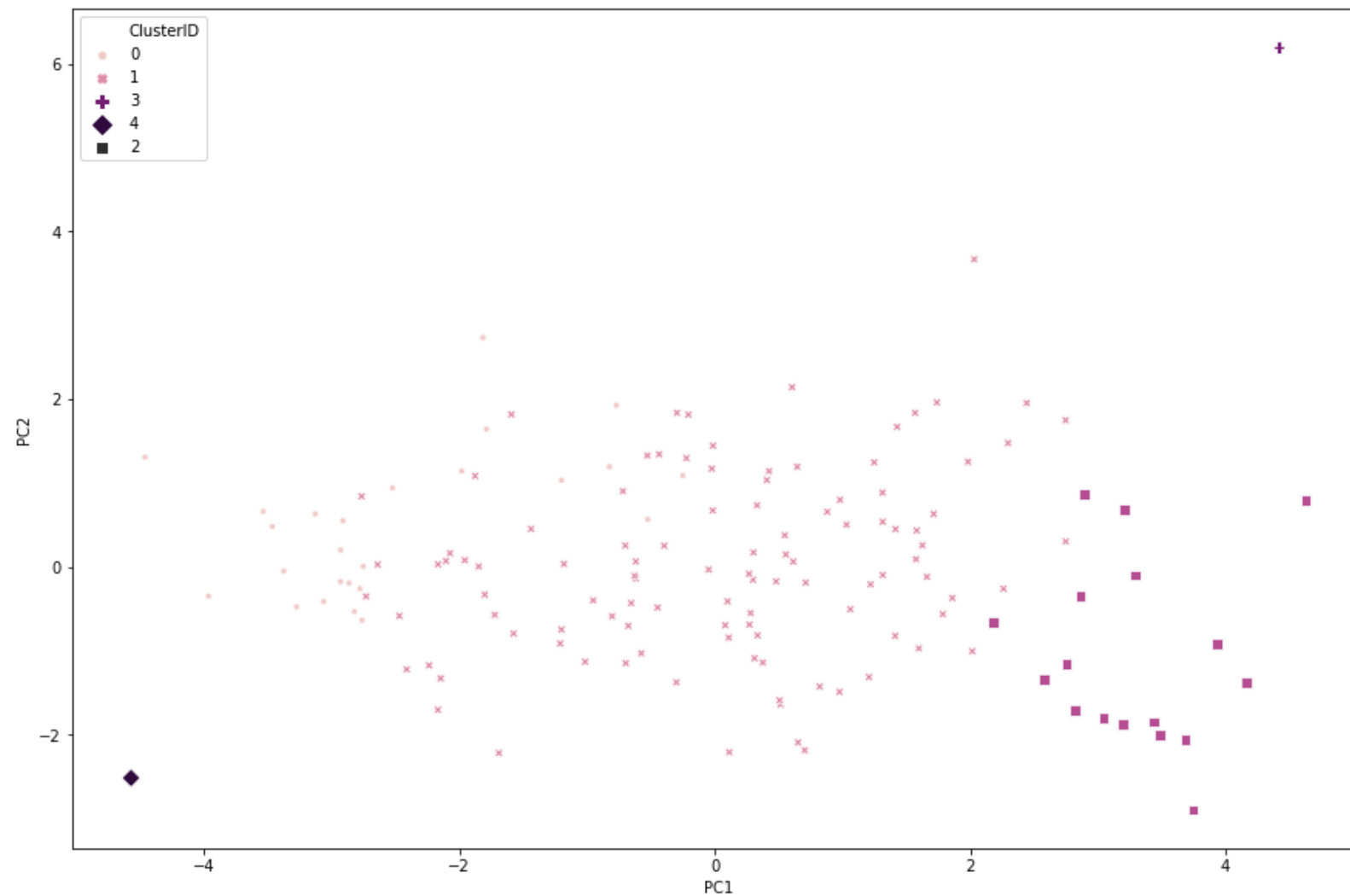
The companies in cluster 2 seem to have low life expectancy and low net income per capita

# Hierarchical Clustering– Dendrogram



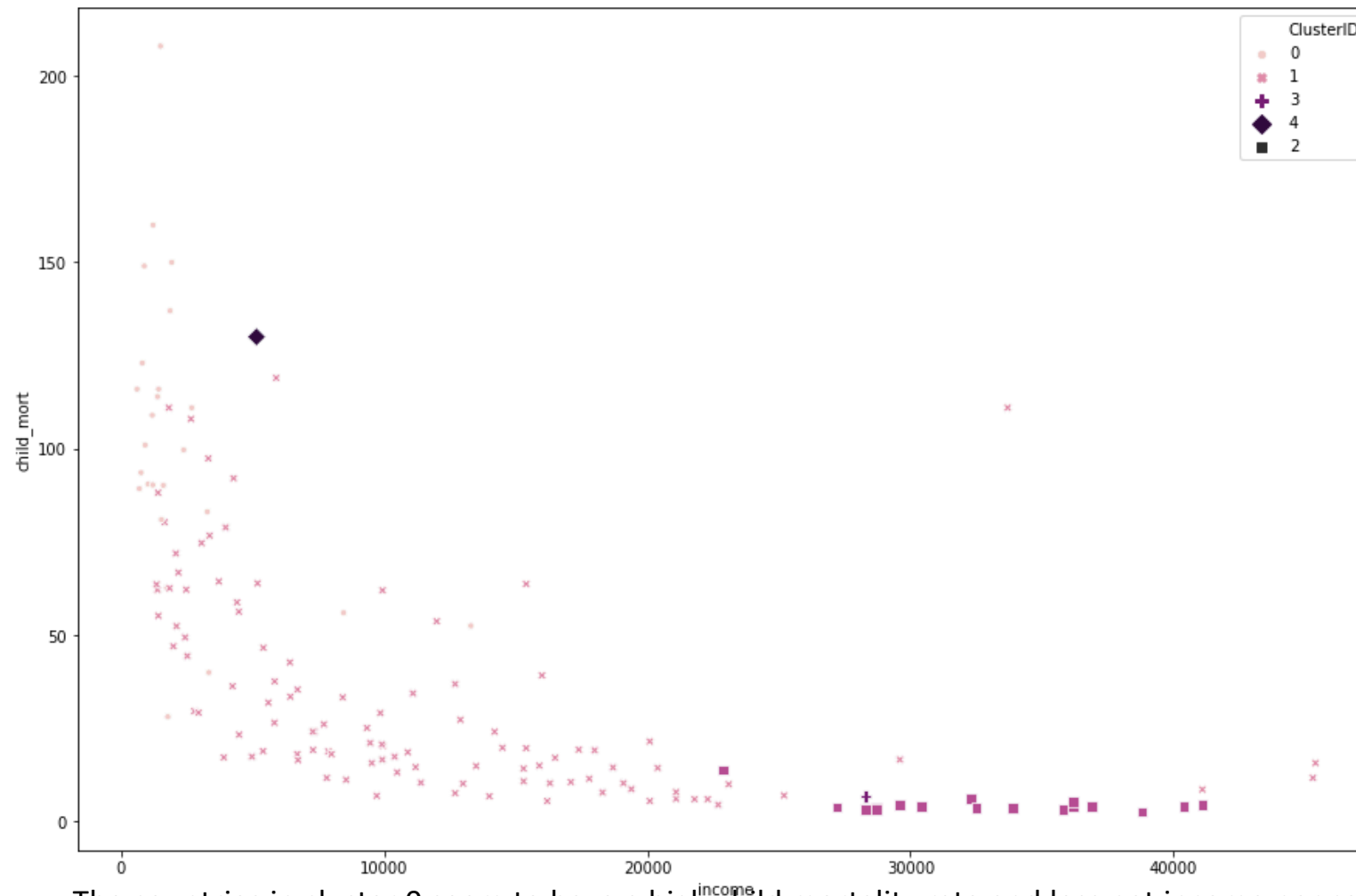
The business justification says we need to break into 3 well formed popular clusters representing under developed, developing and developed countries. but on observing the linkage we find that we get three well defined populated clusters in the dendrogram only when we take at least 5 clusters because two of these 5 clusters captures outlier countries and have 1 and 4 countries. so it makes sense to take 5 as no of clusters for Hierarchical Clustering

# Hierarchical Clustering– Clusters Formation



# Hierarchical Clustering

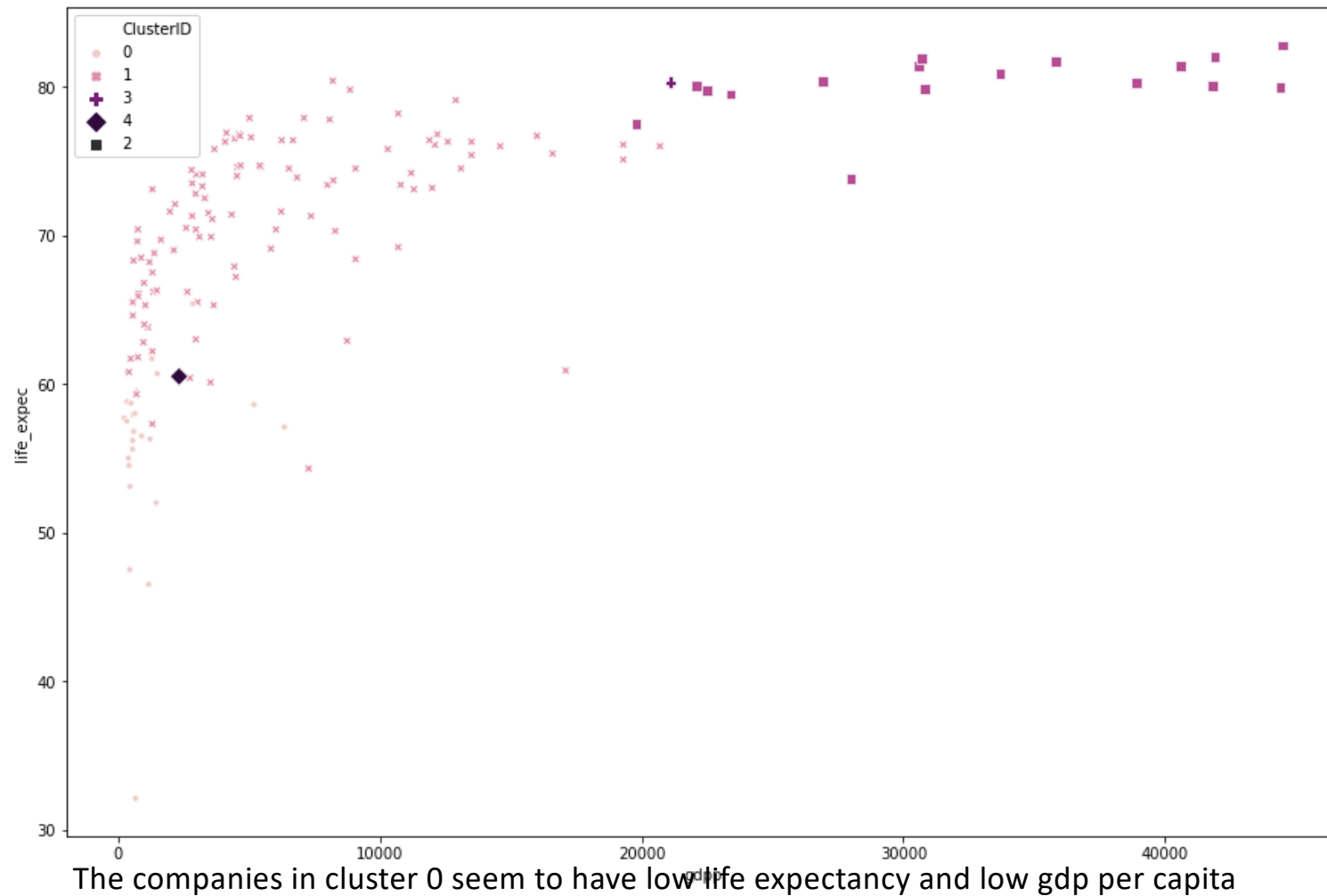
Clusters association with child mortality and per capita income



The countries in cluster 0 seem to have a high child mortality rate and less net income per capita

# Hierarchical Clustering

Clusters association with life expectancy and low income per capita



## Conclusion :

The cluster 2 in K-Means Clustering and cluster 0 in Hierarchical Clustering seem to represent the cluster of economically weakest and hence countries in direst need of funds

The companies in cluster 2 in K Means are: **Afghanistan, Benin, Burundi, Burkina Faso, Central African Republic, Chad, Congo Republic, Congo Democratic Republic etc.**

The companies in cluster 0 of hierarchical clustering are:- **Afghanistan, Benin, Burkina Faso, Botswana, Burundi, Central African Republic, Chad, Congo Democratic Republic, Kenya etc.**

## Recommendations :

Serial No	Country	GDP per capita	Mortality rate	Life expectancy	Net income per person
1	<b>Afghanistan</b>	553	90.2	56.2	1610
2	<b>Central African Republic</b>	446	149	47.5	888
3	<b>Congo Democratic Republic</b>	334	116	60.4	609
4	<b>Burundi</b>	231	93.6	57.7	764
5	<b>Haiti</b>	662	208	32.1	1500
6	<b>Sierra Leone</b>	399	160	55	1220
7	<b>Niger.</b>	348	123	58.8	814
	<b>Average</b>	<b>12964</b>	<b>38.27</b>	<b>70</b>	<b>17144</b>