



Clustering Assignment

by
Dheeraj Mallampati



About the Problem

We need to provide the people of backward countries with basic amenities and relief during the time of disasters and natural calamities.

We have raised funds using the recent funding programmes and now we need to decide how to use this money strategically and effectively.

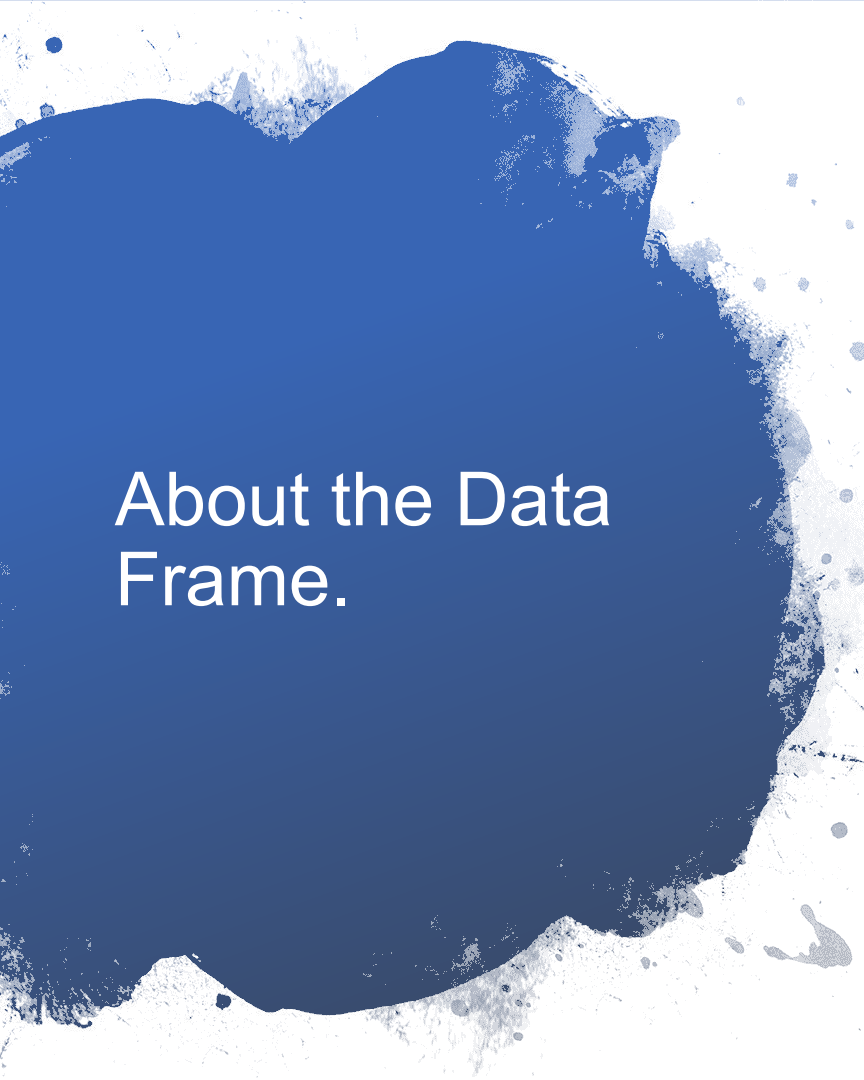
We need to find the countries which are in dire need of aid and the countries that don't need any aid from the NGO.



Solution Strategy.

I chose KMeans clustering to create different unique cluster of countries based on their socio-economic and health factors.

After clustering we got a cluster of countries which are performing very bad, considering the socio-economic and the health factors. These countries are in the dire need of aid from the NGO.



About the Data Frame.

We have a data frame that consists the economic and health factors of 167 different countries.

We have used the below variables to understand the performance of countries in different clusters

- Child mortality rate
- Income per capita
- GDP per capita

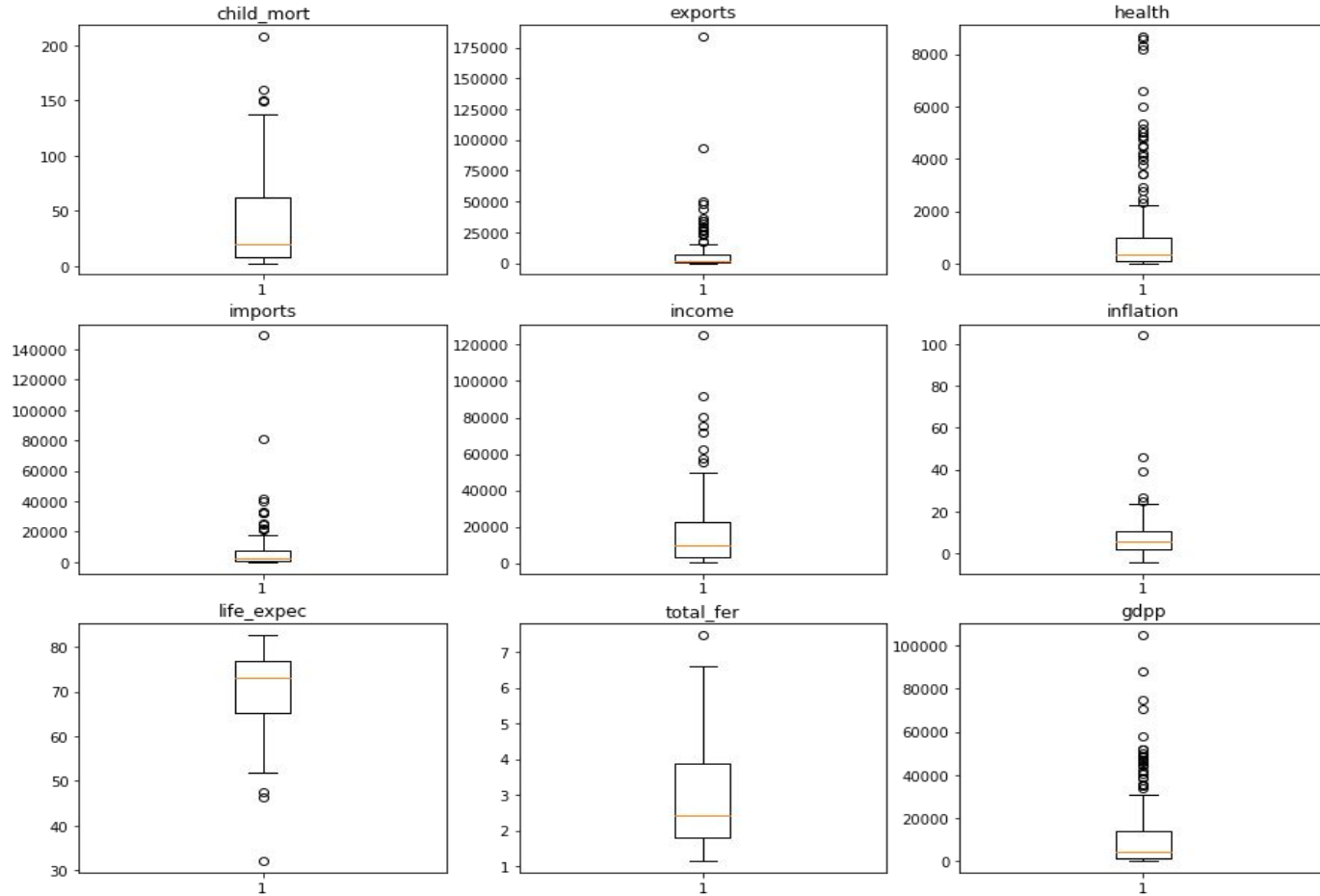
EDA

EDA performed on the Data frame was very minimal because it is already well structured.

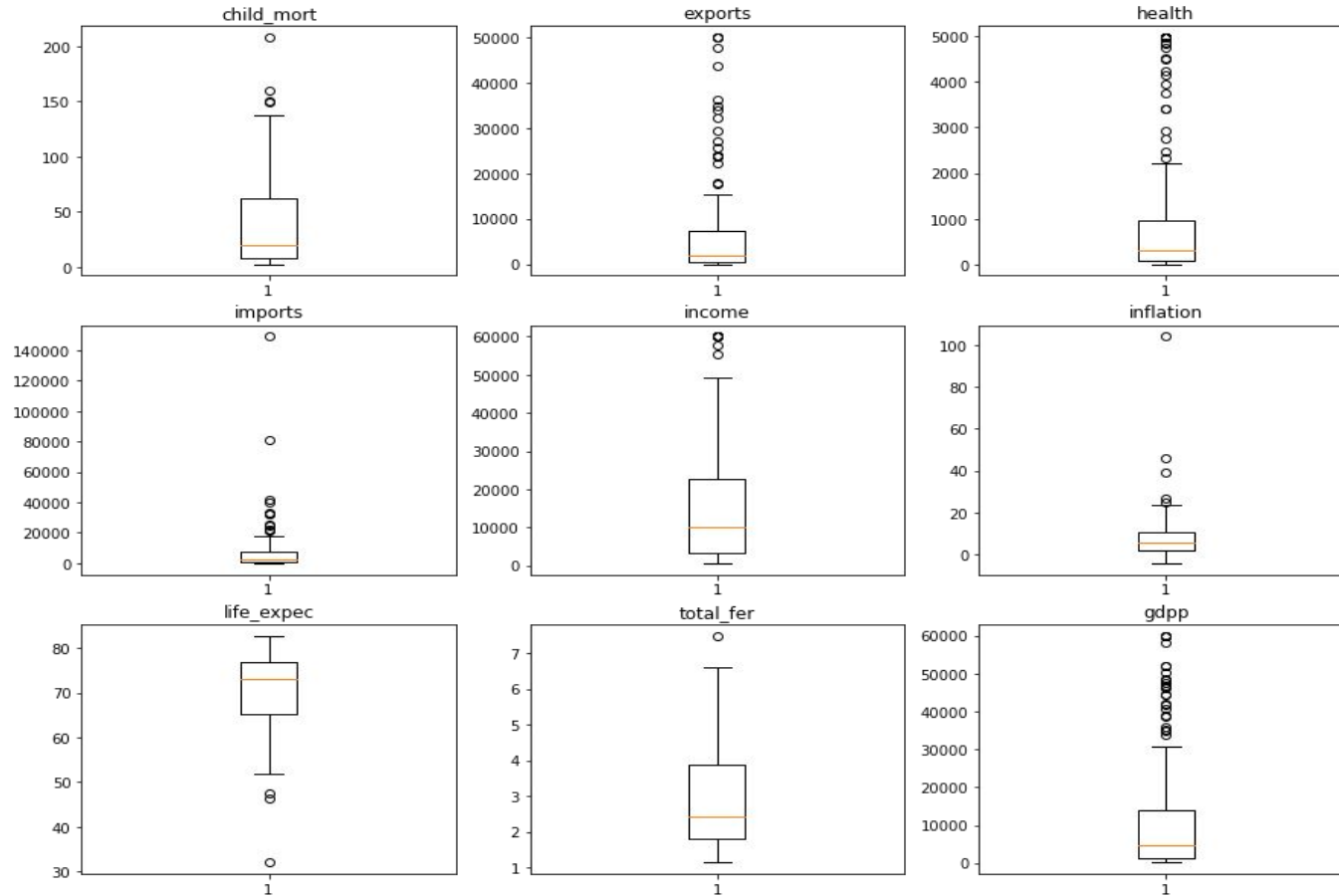
There are no missing values in any columns.

We had to cap the outliers in some of the variables so that the clustering does not get affected. Check the next slides to see the difference in high values after outliers treatment.

Numerical variables before outliers treatment.

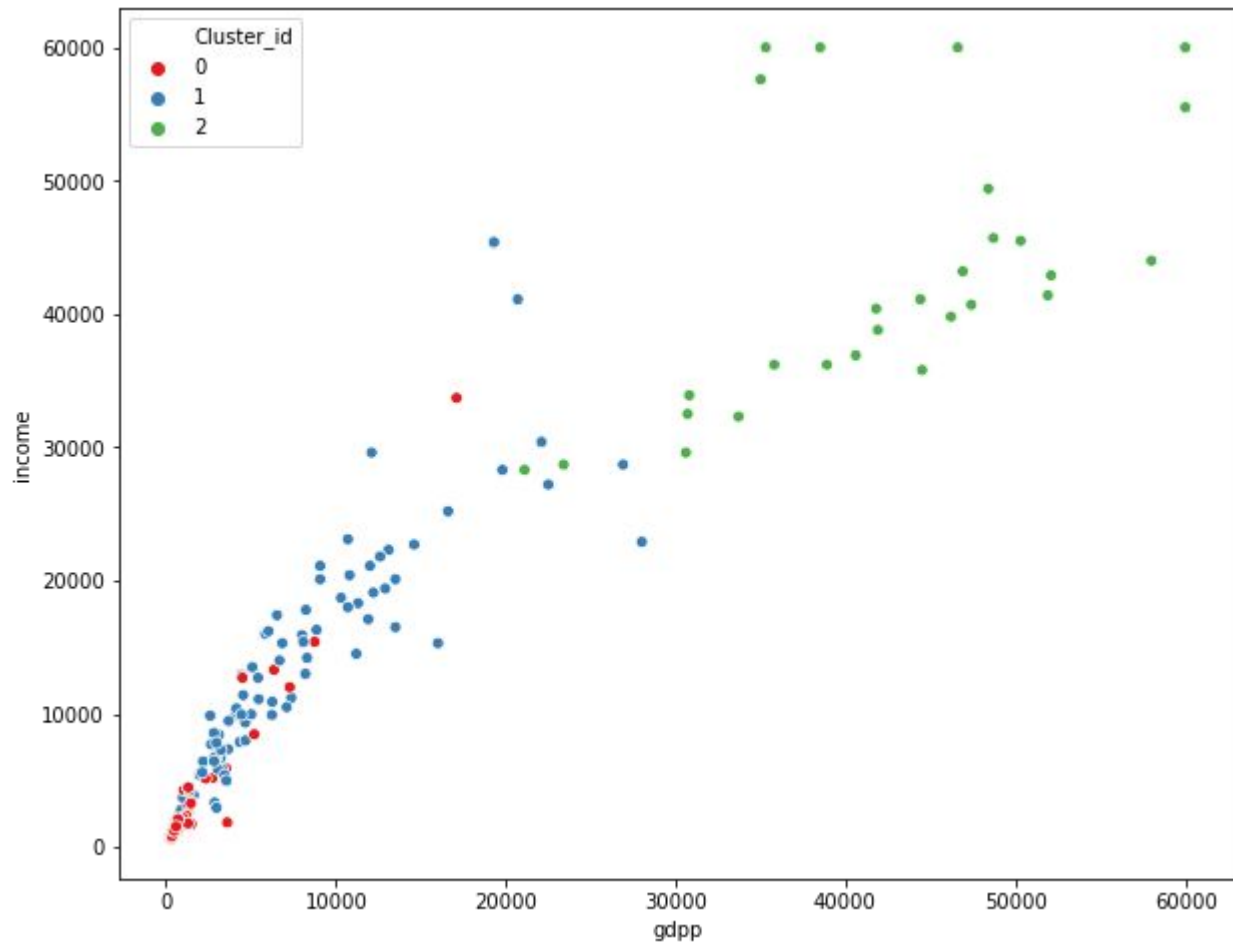


Numerical variables after outliers treatment.



Modelling

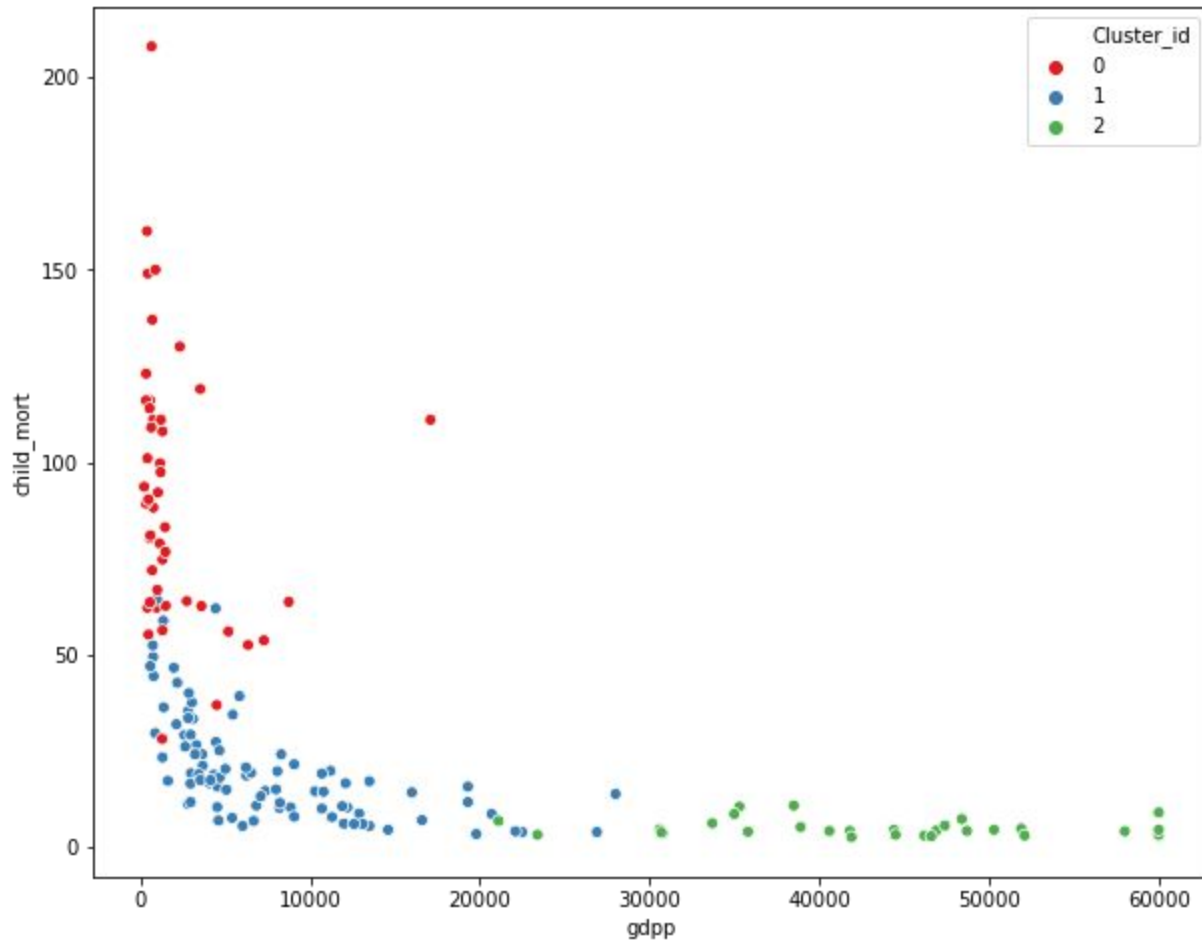
- I tried both KMeans and Hierarchical Clustering to create unique clusters of countries.
- Analyzing the results of both the methods I chose to select KMeans clustering method to create a model that will form different and unique clusters of countries.
- We got 3 clusters using the model
 - Cluster 0 - Under developed countries.
 - Cluster 1 - Developing Countries.
 - Cluster 2 - Well developed Countries



Cluster 0 -
Low income and gdp per capita

Cluster 1 -
Medium income and medium GDP per capita

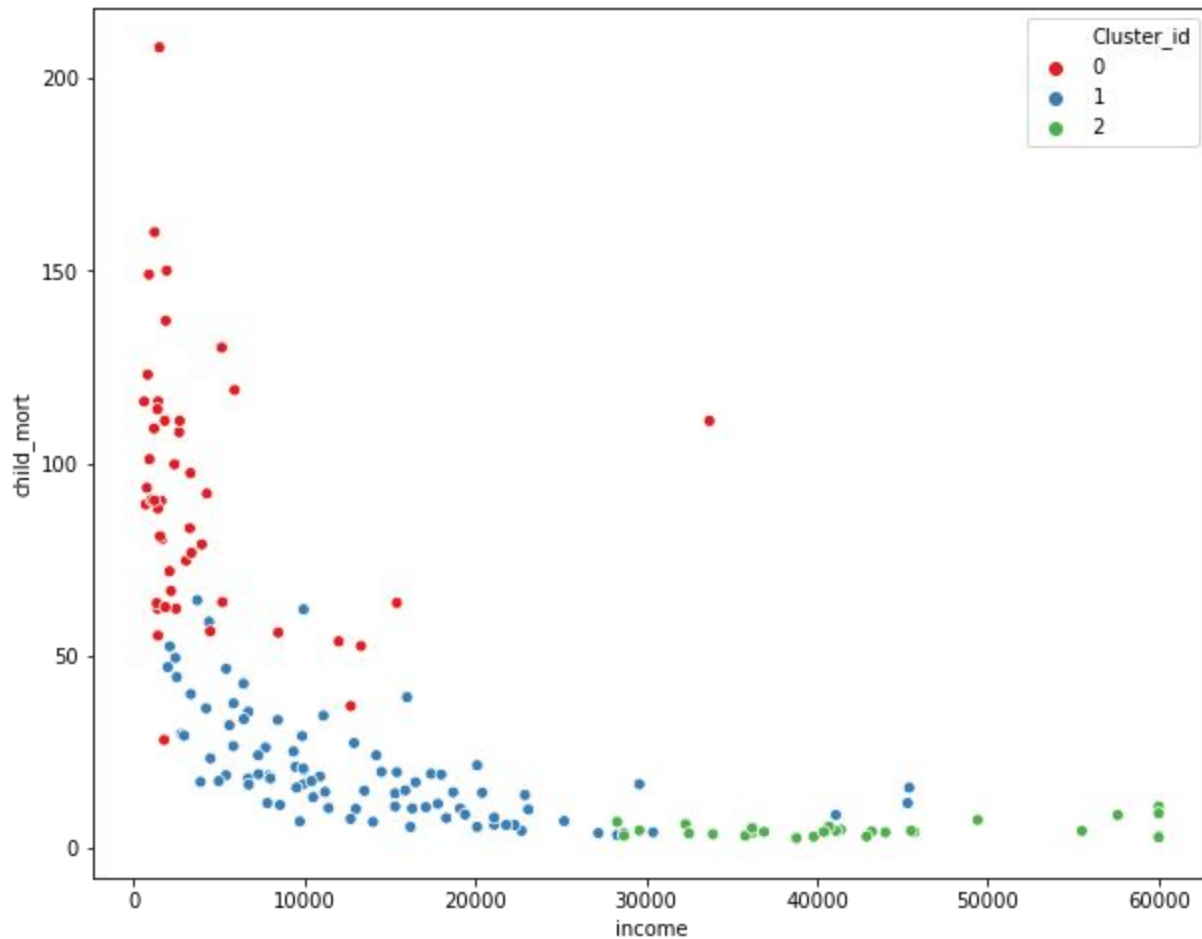
Cluster 2 -
Very high Income and very high GDP per capita



Cluster 0 -
High Child mortality rate and very low
gdp per capita

Cluster 1 -
Medium Child mortality rate and
medium GDP per capita

Cluster 2 -
Low Child mortality rate and very high
GDP per capita

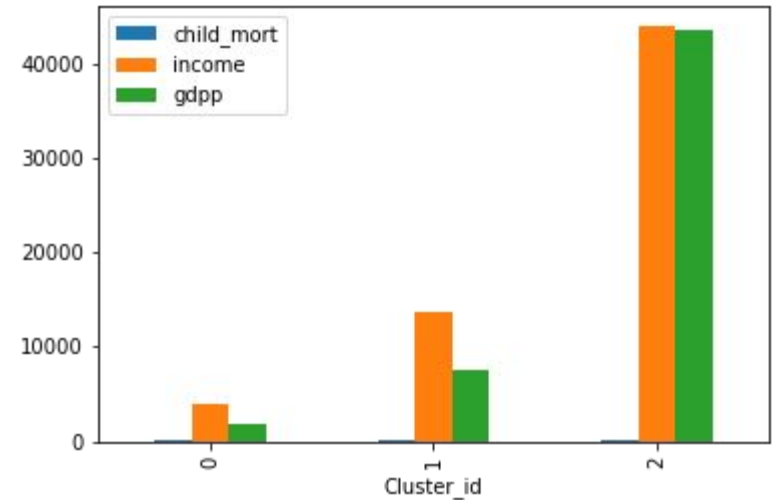
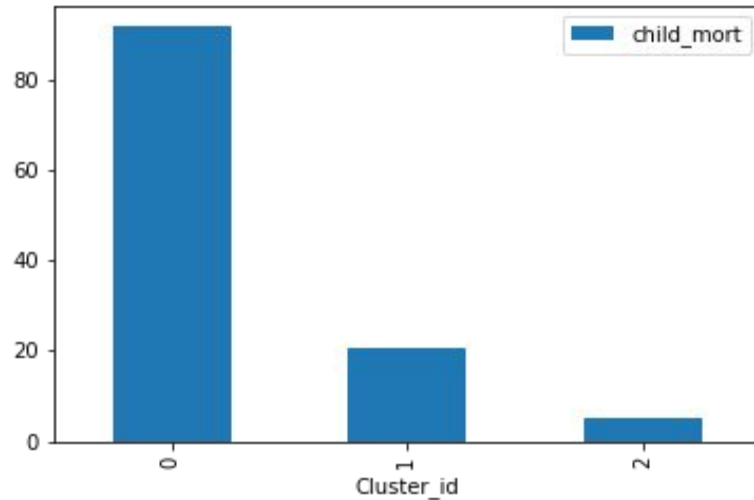


Cluster 0 -
High Child mortality rate and very low
income

Cluster 1 -
Medium Child mortality rate and
medium income

Cluster 2 -
Low Child mortality rate and very high
income

**Below 2 barcharts clearly explains the points mentioned in the previous 3 slides.
Child mortality rate is very high in the countries of cluster 0 and the income and GDP per capita are very low in the cluster.
So, countries belonging to cluster 0 are in dire need of aid from the NGO.**



Countries that Belong to Cluster 0
Underdeveloped countries.
These countries are in dire need of aid from NGO.

Afghanistan, Angola, Benin, Botswana, Burkina Faso, Burundi,
Cameroon, Central African Republic, Chad, Comoros, Congo,
Dem. Rep., Congo, Rep., Cote d'Ivoire, Equatorial Guinea,
Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Haiti,
Iraq, Kenya, Kiribati, Lao, Lesotho, Liberia, Madagascar,
Malawi, Mali, Mauritania, Mozambique, Namibia, Niger,
Nigeria, Pakistan, Rwanda, Senegal, Sierra Leone, Solomon
Islands, South Africa, Sudan, Tanzania, Timor-Leste, Togo,
Uganda, Yemen, Zambia

Countries that Belong to Cluster 1

Developing countries.

These countries which are not in dire need but still needs some assist.

Albania, Algeria, Antigua and Barbuda, Argentina, Armenia, Azerbaijan, Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belize, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Cambodia, Cape Verde, Chile, China, Colombia, Costa Rica, Croatia, Czech Republic, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Fiji, Georgia, Greece, Grenada, Guatemala, Guyana, Hungary, India, Indonesia, Iran, Jamaica, Jordan, Kazakhstan, Kyrgyz Republic, Latvia, Lebanon, Libya, Lithuania, Macedonia, FYR, Malaysia, Maldives, Mauritius, Micronesia, Fed. Sts., Moldova, Mongolia, Montenegro, Morocco, Myanmar, Nepal, Oman, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russia, Samoa, Saudi Arabia, Serbia, Seychelles, Slovak Republic, South Korea, Sri Lanka, St. Vincent and the Grenadines, Suriname, Tajikistan, Thailand, Tonga, Tunisia, Turkey, Turkmenistan, Ukraine, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam.

Countries that Belong to Cluster 2
Well-developed countries.
These countries need no assist from the NGO.

Australia, Austria, Belgium, Brunei, Canada, Cyprus,
Denmark, Finland, France, Germany, Iceland, Ireland,
Israel, Italy, Japan, Kuwait, Luxembourg, Malta,
Netherlands, New Zealand, Norway, Qatar, Singapore,
Slovenia, Spain, Sweden, Switzerland, United Arab
Emirates, United Kingdom, United States

THANK you!!!!!!