

HELP INTERNATIONAL

CASE STUDY ANALYSIS

JASDEEP SINGH MONGA

What is HELP International?

- **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.

Problem Statement

- Help International recently has received a funding of 10 million dollars.
- Now, CEO of the organization needs to decide how to use this money strategically and effectively.
- The significant issues that come while making this decision are mostly related to choosing the countries that are in the direst need of aid.
- So, as a Data Analyst I need to provide him with the countries that are in dire need of the aid based on the socio economic factors.

Analysis Approach Taken

1. Reading and Visualising the Data set

- Importing the relevant libraries along with the Data set

2. Cleaning the Data set.

- Data set was redefined by changing certain Columns into their appropriate values.

3. Preparing the Data for Modelling

- By Using the IQR process(5% and 95%) most of the outliers were removed.

4. Modelling

- K-Means and HCA Process.

5. Final Analysis and Recommendations

- Relevant approach was selected for better Results.

Clustering Methods Used

Clustering: It is a method by which I can divide the given data points into a group based on their similarity. That means the points which are more similar to each other will fall in one group.

There are 2 types of Clustering:

- **K-Means Clustering**
- **Hierarchical Clustering Algorithm**

K-Means Algorithm

1. In **K-Means Algorithm**, I initialise the No. of clusters in the beginning to be 4.
2. Then the code will select the random centroid for these clusters. **K=4**
3. Now we will assign each point to its centroid.
4. Once, the centroid is assigned now we compute the centroid of this new cluster again.
5. We will repeat step 3 and 4 unless we stop observing any new change in the given structure.

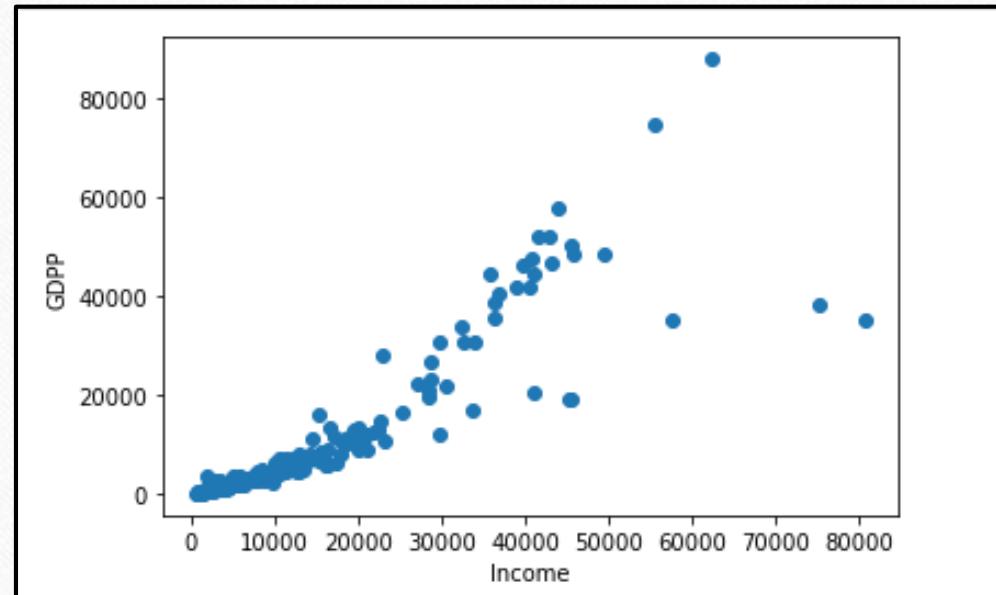
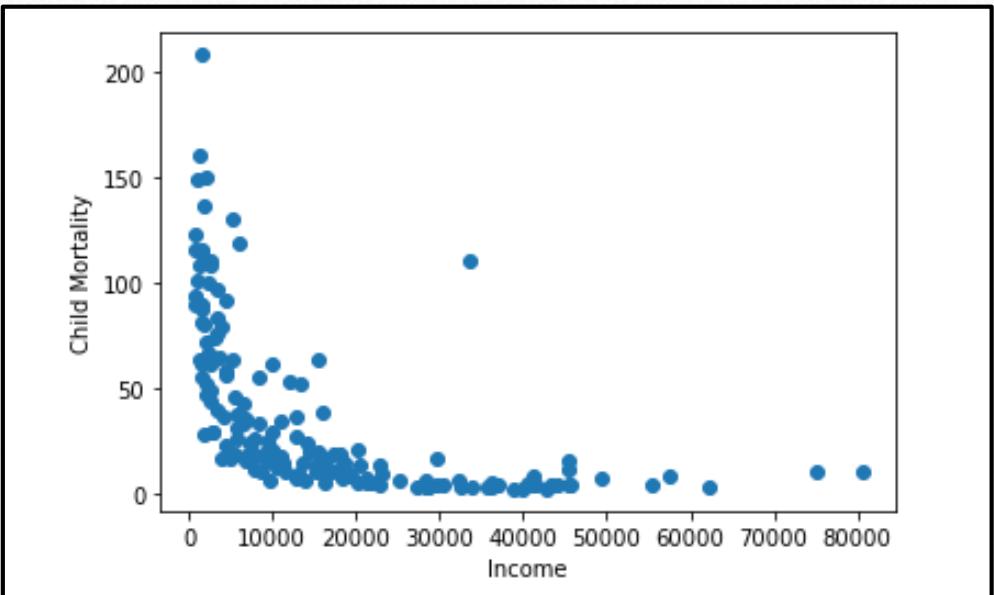
Hierarchical Clustering Algorithm

1. In this Algorithm, I will assign all the points to an individual cluster.
2. Then I'll combine all those points which are having the smallest distance between them.
3. So, I'll repeat this step until I am left with a single cluster.
4. Once, I run commands on this in python I will get a **Dendrogram** which help me get my K value from the graph itself.
5. Now, the points where I cut my **Dendrogram** will be my K-value.

Clustering Results

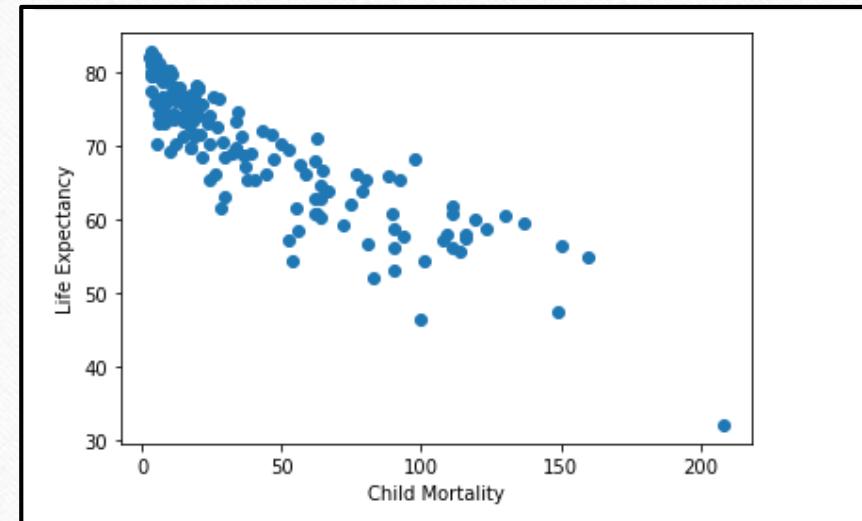
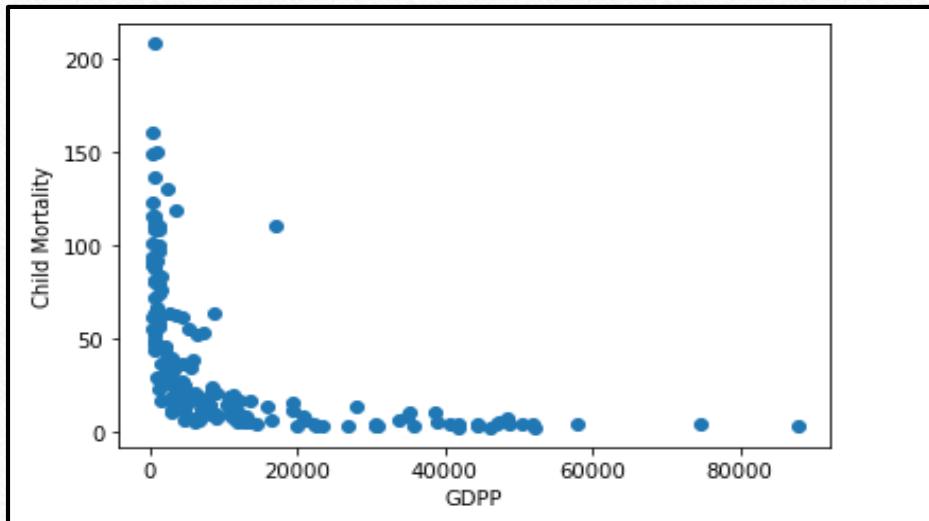
- In **K-Means Algorithm**, we found out that there are **48 countries** which are in need of aid from the NGO.
 - There is a major impact on GDP, Imports, Exports , Income of the person and Life expectancy of the children in these countries.
 - In this Algorithm we started with **K=4** but eventually we found out that, as per the **elbow curve K value which is best suited is 3**.
- ❖ Using **Hierarchical Clustering algorithm**, we found out that there are **126 countries** which need assistance from the NGO.
 - ❖ K value here is found out using the **bottom up approach**, in which we club all the points which show similarity and have less distance between them in a single cluster.
 - ❖ So, here also we took the number of clusters to be 3 i.e. **K=3**.

Visualisations



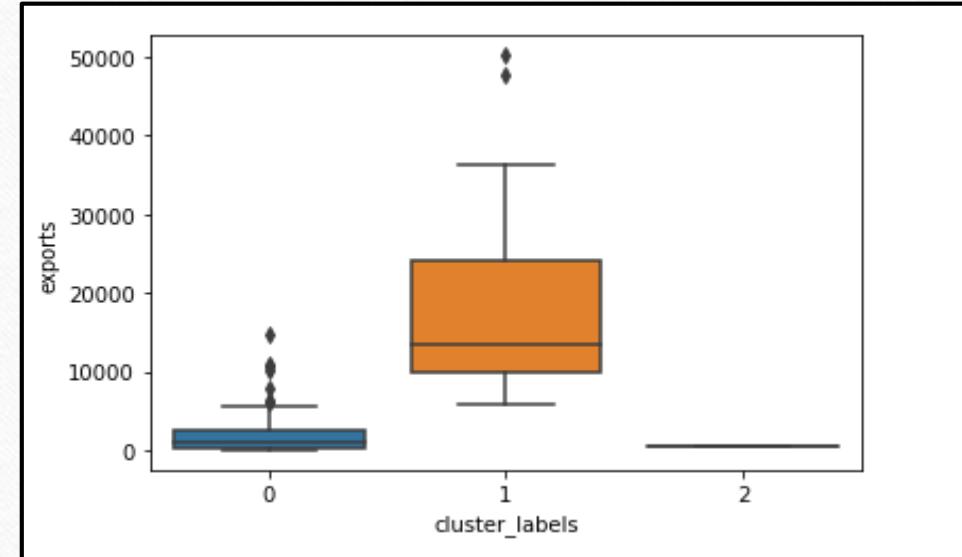
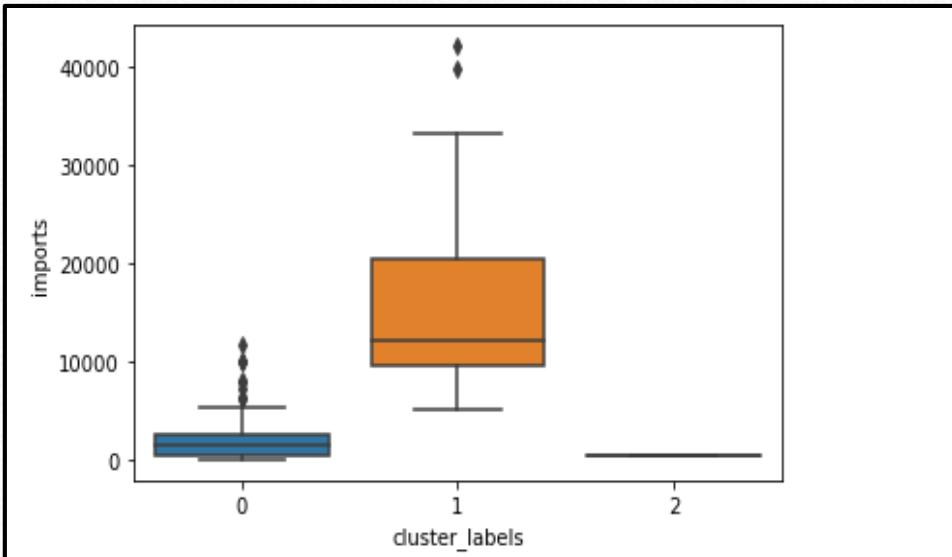
- Now the above two plots describe the relation of Income with Child Mortality and Income with GDPP Mortality.
- **Focus of the NGO should be on Countries where the Income is less than \$10k-\$20k.**
- **Those are the countries where there is High Child Mortality Rate and low GDP growth**

Visualisations



- The above 2 plots describe the relation between **GDPP v/s Child Mortality** and **Life Expectancy v/s Child Mortality**.
- Now the main focus of the NGO should be on those countries where due to low **GDPP** (less than \$15k) we have **high Child Mortality Rate**.
- Also, NGO needs to focus on all those countries where we have life expectancy between 50-60 years and **Child Mortality Rate range is 50 to 200**.

Visualisation



- The above 2 boxplots show us the Import and Export in the Hierarchical Clustering Algorithm.
- Results were more or less similar in K-Means as well.
- Now the point here is NGO here needs to focus on countries which are in cluster label 0 and cluster label 2.
- Improving these 2 factors will improve the GDPP, which will in return give more monetary benefits to the people and then we will have low Child Mortality Rate and correspondingly we will have high life expectancy of every individual.

Final List of All Countries.

- My preferred approach here was to go ahead with the **K-Mean Clustering** because it gave me the right set of cluster with any over analysis done on the data set.

country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	cluster_id	c
Afghanistan	90.2	55.3000	41.9174	248.297	1610	9.440	56.2	5.82	553	2	
Angola	119.0	2199.1900	100.6050	1514.370	5900	22.400	60.1	6.16	3530	2	
Benin	111.0	180.4040	31.0780	281.976	1820	0.885	61.8	5.36	758	2	
Botswana	52.5	2768.6000	527.0500	3257.550	13300	8.920	57.1	2.88	6350	2	
Burkina Faso	116.0	110.4000	38.7550	170.200	1430	6.810	57.9	5.87	575	2	
Burundi	93.6	20.6052	26.7960	90.552	764	12.300	57.7	6.26	231	2	
Cameroon	108.0	290.8200	67.2030	353.700	2660	1.910	57.3	5.11	1310	2	
Central African Republic	149.0	52.6280	17.7508	118.190	888	2.010	47.5	5.21	446	2	
Chad	150.0	330.0960	40.6341	390.195	1930	6.390	56.5	6.59	897	2	
Comoros	88.2	126.8850	34.6819	397.573	1410	3.870	65.9	4.75	769	2	
Congo, Dem. Rep.	116.0	137.2740	26.4194	165.664	609	20.800	57.5	6.54	334	2	
Congo, Rep.	63.9	2331.7400	67.4040	1498.780	5190	20.700	60.4	4.95	2740	2	
Cote d'Ivoire	111.0	617.3200	64.6600	528.260	2690	5.390	56.3	5.27	1220	2	
Equatorial Guinea	111.0	14671.8000	766.0800	10071.900	33700	24.900	60.9	5.21	17100	2	
Eritrea	55.2	23.0878	12.8212	112.306	1420	11.600	61.7	4.61	482	2	
Gabon	63.7	5048.7500	306.2500	1653.750	15400	16.600	62.9	4.08	8750	2	
Gambia	80.3	133.7560	31.9778	239.974	1660	4.300	65.5	5.71	562	2	
Ghana	74.7	386.4500	68.3820	601.290	3060	16.600	62.2	4.27	1310	2	
Guinea	109.0	196.3440	31.9464	279.936	1190	16.100	58.0	5.34	648	2	
Guinea-Bissau	114.0	81.5030	46.4950	192.544	1390	2.970	55.6	5.05	547	2	
Haiti	208.0	101.2860	45.7442	428.314	1500	5.450	32.1	3.33	662	2	
Iraq	36.9	1773.0000	378.4500	1534.500	12700	16.600	67.2	4.56	4500	2	
Kenya	62.2	200.1690	45.9325	324.912	2480	2.090	62.8	4.37	967	2	
Kiribati	62.7	198.1700	168.3700	1190.510	1730	1.520	60.7	3.84	1490	2	
Lao	78.9	403.5600	50.9580	562.020	3980	9.200	63.8	3.15	1140	2	
Lesotho	99.7	460.9800	129.8700	1181.700	2380	4.150	46.5	3.30	1170	2	
Liberia	89.3	62.4570	38.5860	302.802	700	5.470	60.8	5.02	327	2	
Madagascar	62.2	103.2500	15.5701	177.590	1390	8.790	60.8	4.60	413	2	
Malawi	90.5	104.6520	30.2481	160.191	1030	12.100	53.1	5.31	459	2	
Mali	137.0	161.4240	35.2584	248.508	1870	4.370	59.5	6.55	708	2	

Mauritania	97.4	608.4000	52.9200	734.400	3320	18.900	68.2	4.98	1200	2
Mozambique	101.0	131.9850	21.8299	193.578	918	7.640	54.5	5.56	419	2
Namibia	56.0	2480.8200	351.8820	3150.330	8460	3.560	58.6	3.60	5190	2
Niger	123.0	77.2560	17.9568	170.868	814	2.550	58.8	7.49	348	2
Nigeria	130.0	589.4900	118.1310	405.420	5150	104.000	60.5	5.84	2330	2
Pakistan	92.1	140.4000	22.8800	201.760	4280	10.900	65.3	3.85	1040	2
Rwanda	63.6	67.5600	59.1150	168.900	1350	2.610	64.6	4.51	563	2
Senegal	66.8	249.0000	56.6000	403.000	2180	1.850	64.0	5.06	1000	2
Sierra Leone	160.0	67.0320	52.2690	137.655	1220	17.200	55.0	5.20	399	2
Solomon Islands	28.1	635.9700	110.2950	1047.480	1780	6.810	61.7	4.24	1290	2
South Africa	53.7	2082.0800	650.8320	1994.720	12000	6.350	54.3	2.59	7280	2
Sudan	76.7	291.5600	93.5360	254.560	3370	19.600	66.3	4.88	1480	2
Tanzania	71.9	131.2740	42.1902	204.282	2090	9.250	59.3	5.43	702	2
Timor-Leste	62.6	79.2000	328.3200	1000.800	1850	26.500	71.1	6.23	3600	2
Togo	90.3	196.1760	37.3320	279.624	1210	1.180	58.7	4.87	488	2
Uganda	81.0	101.7450	53.6095	170.170	1540	10.600	56.8	6.15	595	2
Yemen	56.3	393.0000	67.8580	450.640	4480	23.600	67.5	4.67	1310	2
Zambia	83.1	540.2000	85.9940	451.140	3280	14.000	52.0	5.40	1460	2

Countries that Require Immediate Aid

Here are the Names of the Top 10 countries that require our Immediate Attention

country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	cluster_id	cluster_labels
Afghanistan	90.2	55.3000	41.9174	248.297	1610	9.440	56.2	5.82	553	0	0
Benin	111.0	180.4040	31.0780	281.976	1820	0.885	61.8	5.36	758	0	0
Burkina Faso	116.0	110.4000	38.7550	170.200	1430	6.810	57.9	5.87	575	0	0
Burundi	93.6	20.6052	26.7960	90.552	764	12.300	57.7	6.26	231	0	0
Cameroon	108.0	290.8200	67.2030	353.700	2660	1.910	57.3	5.11	1310	0	0
Central African Republic	149.0	52.6280	17.7508	118.190	888	2.010	47.5	5.21	446	0	0
Chad	150.0	330.0960	40.6341	390.195	1930	6.390	56.5	6.59	897	0	0
Comoros	88.2	126.8850	34.6819	397.573	1410	3.870	65.9	4.75	769	0	0
Congo, Dem. Rep.	116.0	137.2740	26.4194	165.664	609	20.800	57.5	6.54	334	0	0
Cote d'Ivoire	111.0	617.3200	64.6600	528.260	2690	5.390	56.3	5.27	1220	0	0

Final List of Countries

- In the previous slide we have the list of **48 countries** which are in need of Aid.
- **But 1st our attention should be on the 10 countries which are having imports less than 1000.**
- The means by which we can reduce the no. of countries from that cluster is by improving the below mentioned factors:
- **So, as a Data Analyst my suggestion to the CEO would be to improve → the trading business(import and export) in these countries which will increase the GDPP → that will increase the economy and an individual will have a better income in hand-→ this will improve the child mortality rate as people will focus more on the health of their children that are below 5 years of age and ultimately → this will lead to increase in life expectancy of the person.**

Thank You