

CLUSTERING AND PCA ASSIGNMENT

BY – Rahul Mahajan

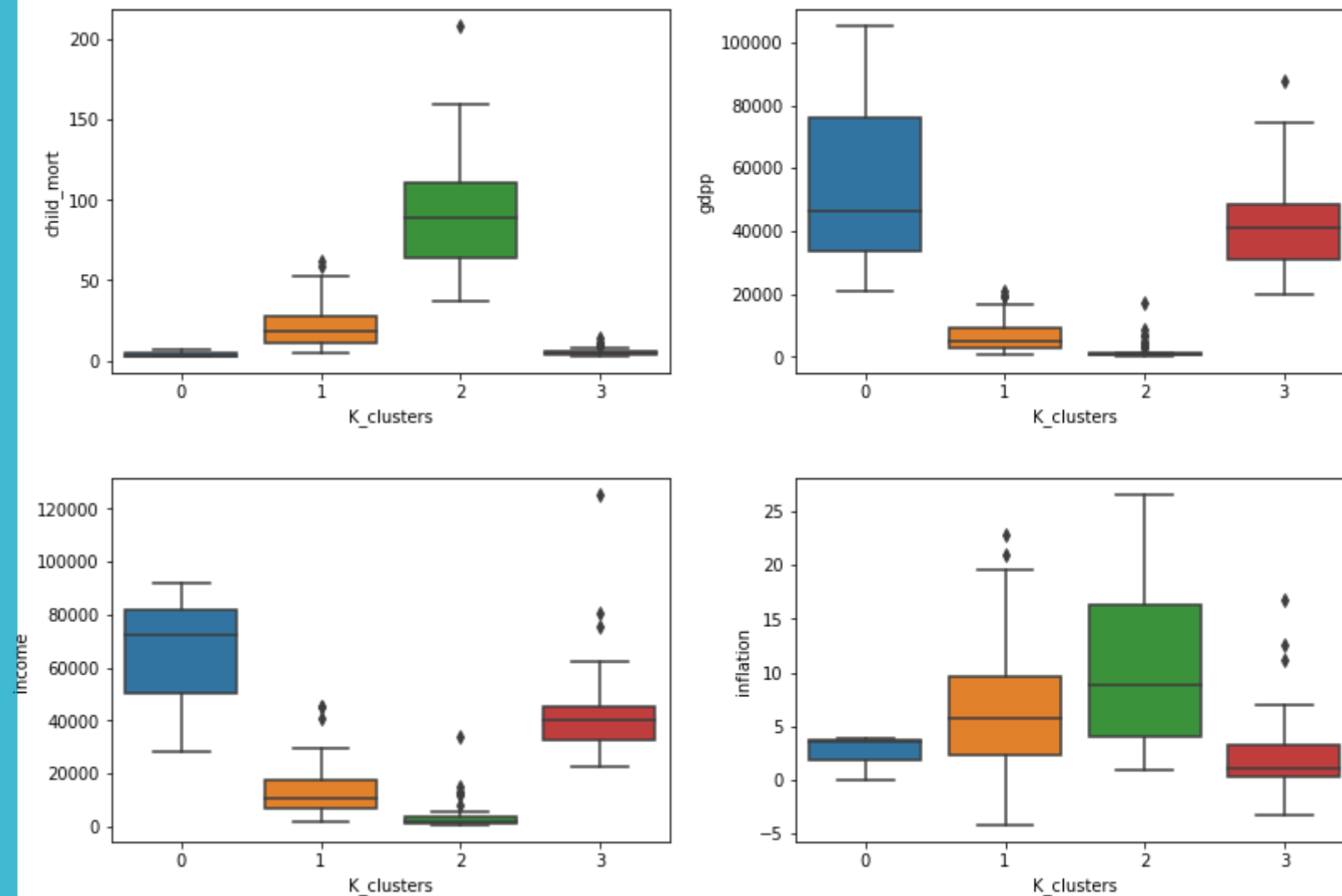
PROBLEM STATEMENT

- HELP International an NGO focused on fighting poverty around the world has managed to raise \$10 Million funds.
- It wants to identify countries in dire need of funds so as to prioritize fund spending

ANALYSIS APPROACH

- The socio economic parameters like child mortality , GDP per person, income, inflation etc. can be used to classify countries.
- Countries doing very poor on these paraments can be identified and targeted for fund spending.
- List of 167 countries along with their socio economic parameters was collected and used for analysis.
- 3 countries with high inflation were left out of analysis as its inflation figure were way too high and had a significant impact in analysis.
- Remaining 166 countries were grouped using cluster analysis under different categories.
- Though we are able to cluster countries into groups of 3 and 4 using different clustering techniques, the top 10 countries in dire need of funds remains same in both the analysis.

4 Clusters Analysis – Graphical representation



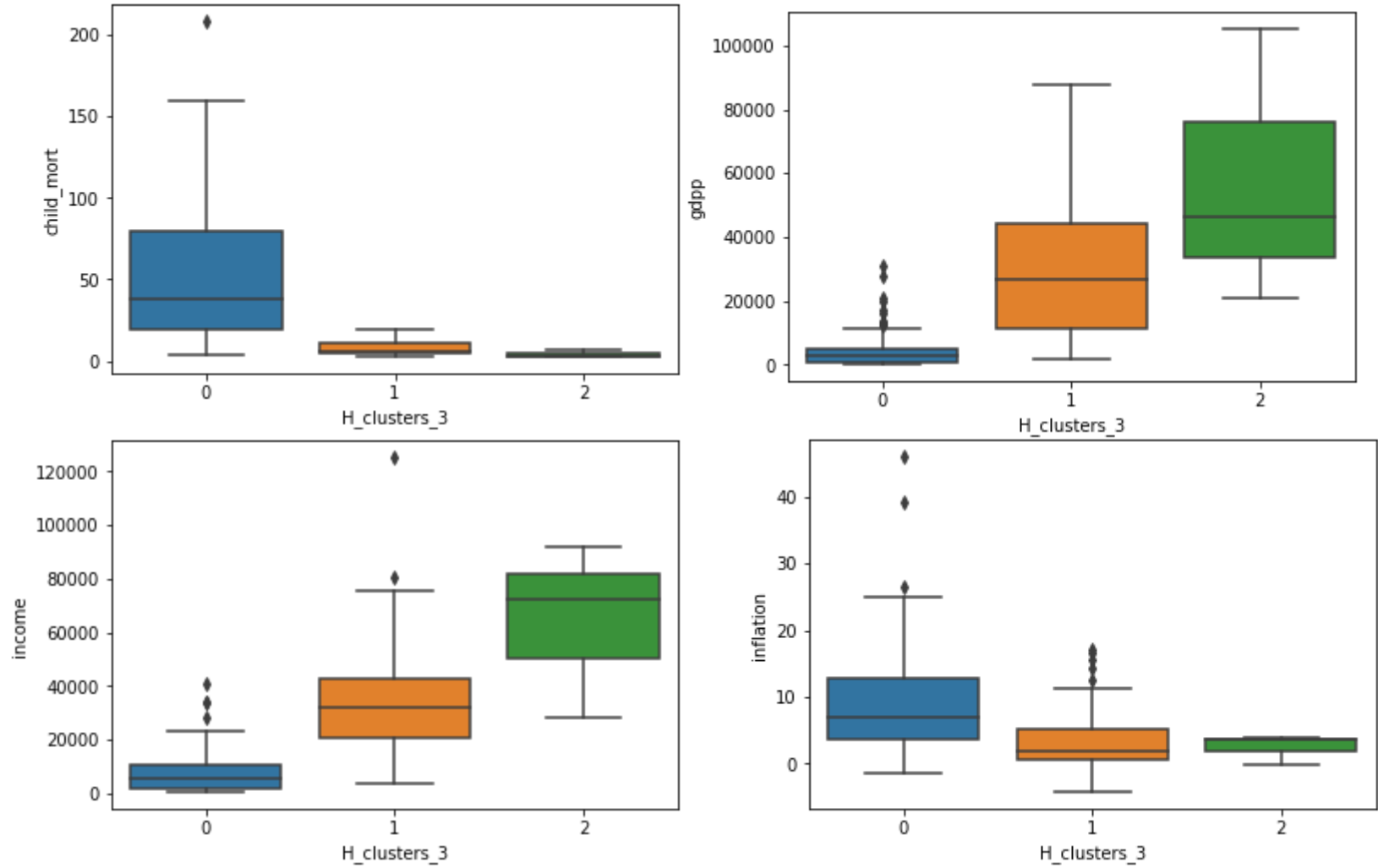
4 CLUSTER APPROACH - Interpretation

- 1. Cluster 2 - Very High child mort , Very Low income , Very Low GDPP. Very High inflation.
- 2. Cluster 1 - High Child mort, Low income , low GDPP. High inflation.
- 3. Cluster 3 - Low child mort, Average income and Average GDPP. Average Inflation
- 4. Cluster 0 - Low child mort, High income and High GDPP

Clusters	Child_mort	Health	Income	Inflation	Life_expect	GDPP
0 (3 countries)	4.13	6.79	64033.33	2.47	81.43	57566.67
1(82 countries)	21.17	6.27	12794.39	6.76	73.00	6549.07
2(47 countries)	91.57	6.32	3911.98	9.96	59.32	1893.83
3(32 countries)	5.18	9.09	44021.88	2.51	80.08	42118.75

- From our business objective we should focus on cluster 2 then cluster 1. Cluster 0 and 3 are countries doing well

3 CLUSTER APPROACH - Graphical representation



3 CLUSTER APPROACH - Interpretation

- Cluster 0 - Very High child mort , Very Low income , Very Low GDPP. High inflation.
- Cluster 1 - Low Child mort, Average income , Average GDPP. Average inflation.
- Cluster 2 - Low child mort, High income and High GDPP. Low Inflation

CLUSTERS	Child_mortality	Health	Income	Inflation	Life_expec	GDPP
0 (112 Countries)	52.31	6.21	8000.92	8.99	66.88	4365.04
1 (51 Countries)	7.65	8.19	34702.16	3.55	78.18	29433.33
2 (3 Countries)	4.13	6.79	64033.33	2.47	81.43	57566.67

- From our business objective we should focus on Cluster 0 . Cluster 2 countries are doing well

FINAL RECOMMEND ATION - LIST OF TOP 10 CUNTRIES IN DIRE NEED OF AID

COUNTRY	Child_Mortality	Health	Income	Inflation	Life_Expectancy	GDPP
Haiti	208	6.91	1500	5.45	32.1	662
Sierra Leone	160	13.1	1220	17.2	55	399
Chad	150	4.53	1930	6.39	56.5	897
Central African Republic	149	3.98	888	2.01	47.5	446
Mali	137	4.98	1870	4.37	59.5	708
Niger	123	5.16	814	2.55	58.8	348
Angola	119	2.85	5900	22.4	60.1	3530
Congo, Dem. Rep.	116	7.91	609	20.8	57.5	334
Burkina Faso	116	6.74	1430	6.81	57.9	575
Guinea- Bissau	114	8.5	1390	2.97	55.6	547

NOTE : 3 Countries with hyper inflation were left out of analysis. Out of thiese 3 Nigeria and Mongolia should also they have low income and gdpp