

HELP International NGO CASE STUDY

SUBMISSION

Submitted By : Shivali Dalmia

PROBLEM STATEMENT

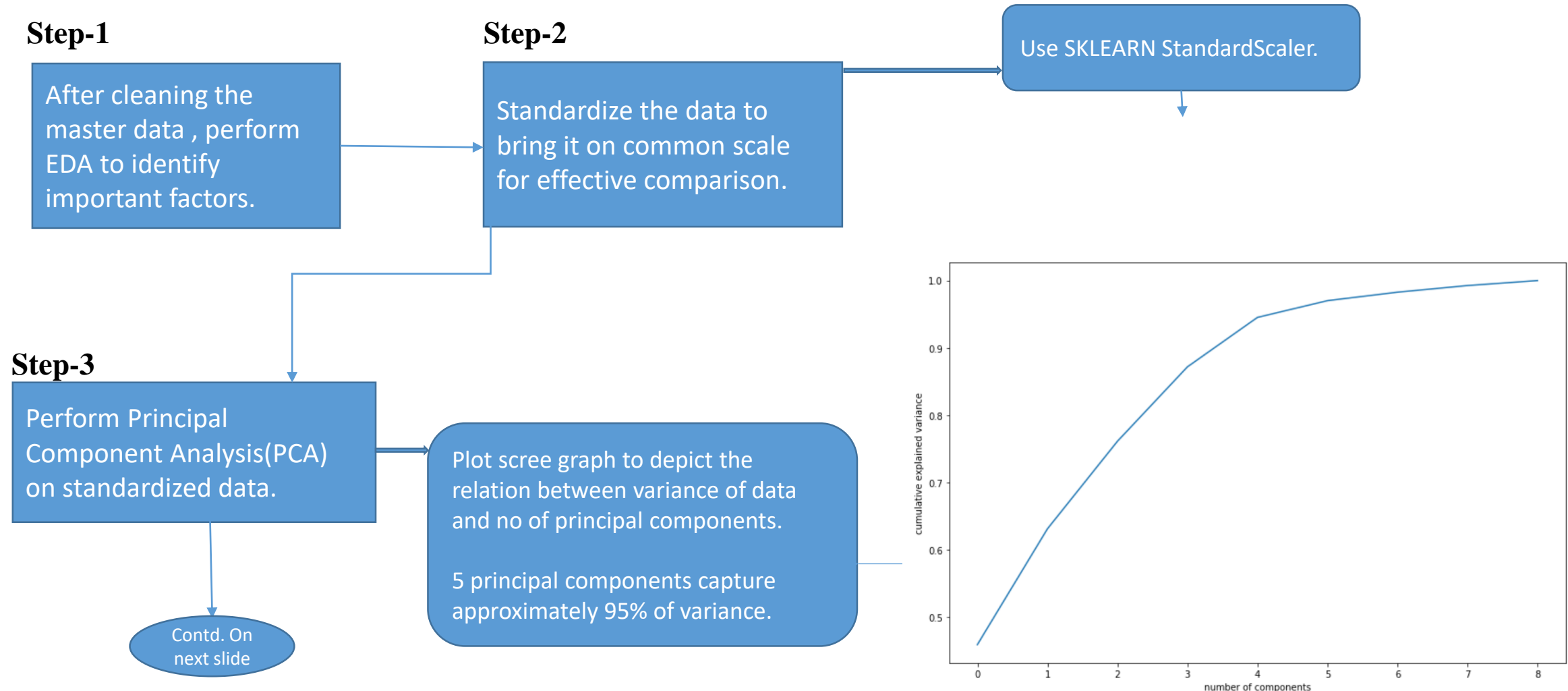
Problem:

HELP international NGO needs to know the countries which are in direst aid of need. They want to channelize their money strategically and effectively to help the areas in need.

Goal of Analysis:

- Categorize countries by analysing certain socio economic and health factors that determine the overall development of the country.
- Identify the countries which are in most need of help to progress on their overall development.
- Recommend some countries to invest in.

Analysis Approach



Step - 4

Perform outlier analysis on PCA's chosen (PC1 to PC5).

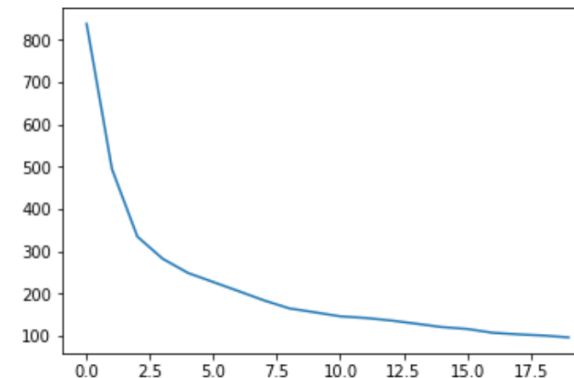
Use IQR method to remove the outliers completely which can disrupt the cluster formations.

Step - 5

Apply K-means algorithm by deciding initial no of clusters (k=3/5).

- Check how well is data clustered using Hopkins statistics (0.73 close to 1)
- Plot elbow curve to decide on optimum no of clusters.
- K=5 is optimum as after that graph becomes flat.

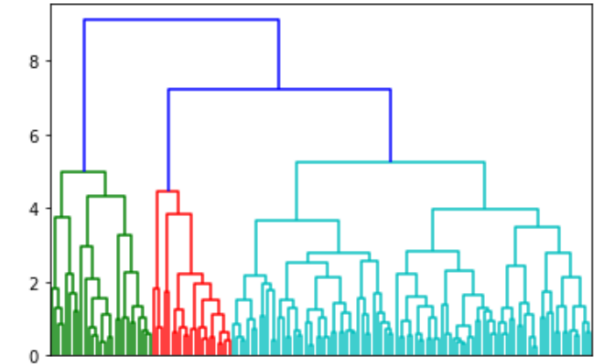
K=5 is chosen as optimum no of clusters to be formed.



Step - 6

Apply hierarchical clustering to decide on no of clusters to be chosen.

K=5 comes to be the optimum value for no of clusters by cutting dendrogram by a horizontal line.



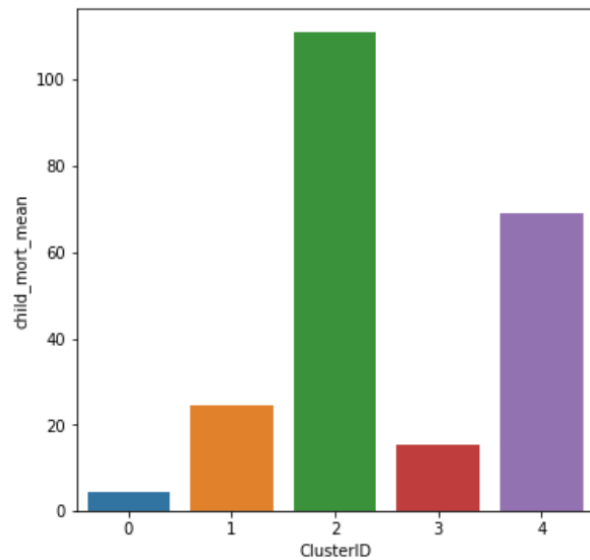
Step - 7

Combine the master country data set with principal components derived during PCA

Perform mean analysis for important principal components and original columns for all 5 clusters.

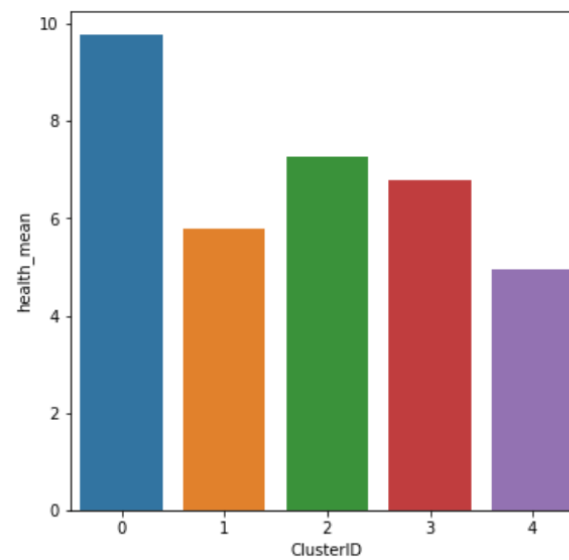
Feature mean analysis

- PC1, PC2 , PC3 , PC4 , PC5 , child mortality rate, total health expenditure , total fertility rate are chosen to be most important parameters to decide on clusters which needs help.
- From below plots we can see that PC1 can be seen as a linear combination of income, life expectancy and gdp parameters as they all are highest in cluster 0 and lowest in cluster 2.
- PC2 can be seen related to inflation original parameter as both are highest in cluster 4 and lowest in cluster 0.



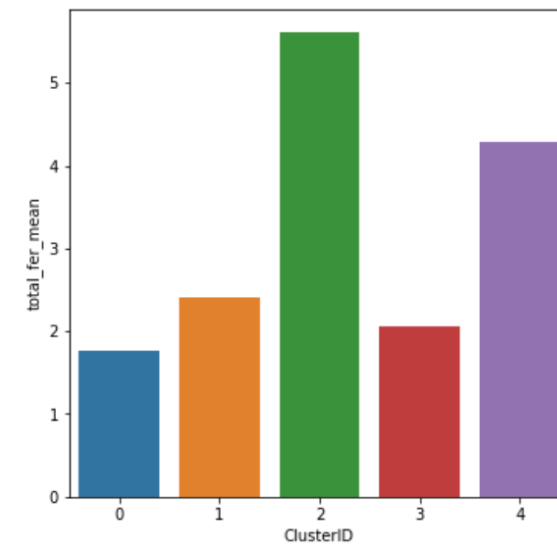
Cluster 0 – Lowest (Good)

Cluster 2 – Highest (Bad)



Cluster 0 – Highest

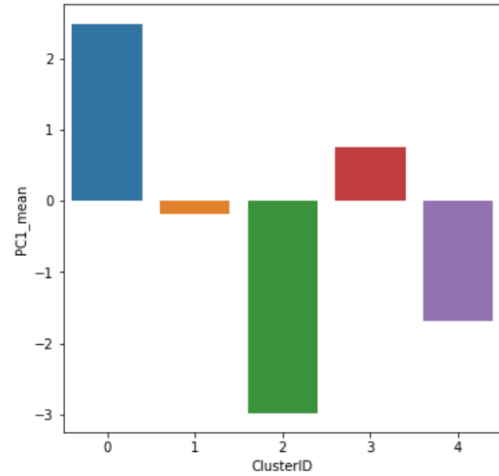
Cluster 4 – Lowest



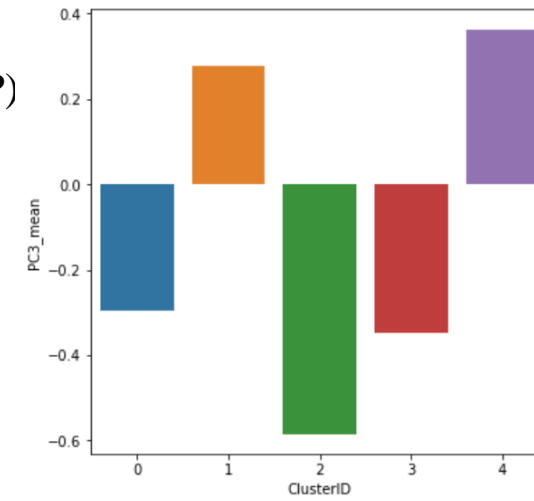
Cluster 0 – Lowest (Good)

Cluster 2 – Highest (Bad)

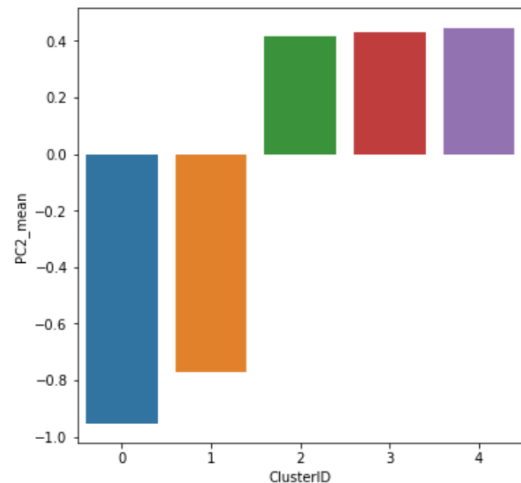
Principal Component Mean Analysis



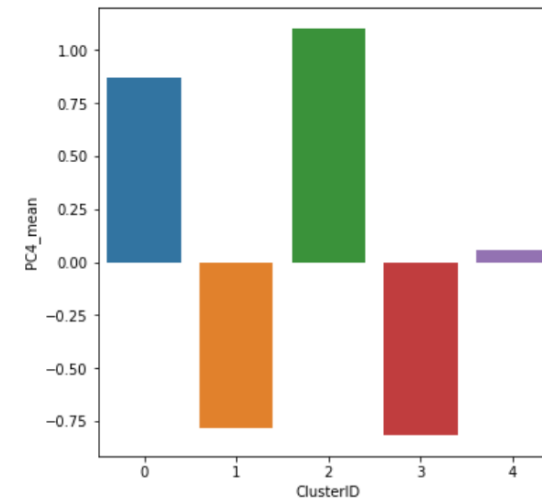
Cluster 0 – Highest (High income, life_expec ,GDP)
Cluster 2 – Lowest (Low income, life_expec,GDP)



Cluster 4 – Highest
Cluster 2 - Lowest



Cluster 0 – Lowest (Low inflation)
Cluster 2/3/4 – Highest (High inflation)



Cluster 2 - Highest
Cluster 3 - Lowest

Inferences

Cluster ID	Analysis
0	Has lowest child mortality rate as due to highest income in this sector they have the ability to avail effective health benefits. As a result inflation is lowest in this cluster and life expectancy is highest .GDP is highest.
1	Has moderately low child mortality rate but inflation is high , but gdp is quiet low.
2	Has highest child mortality rate due to lowest income in this sector. As a result inflation is high in this cluster and life expectancy is lowest. GDP is lowest.
3	Child mortality rate is second lowest , health expenditures are high , inflations are low , life expectancy rate is quiet high , gdp is second highest in this group.
4	Second highest child_mortality_rate , low health expenditures , low income, highest inflation, low life_expec in comparison to other clusters, low gdp.

Conclusion : As per above table countries belonging to cluster 2 are the most direst one in need of investment in every sector , to intensify their growth of development.

Countries to be invested in.

country	ClusterID	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp
Afghanistan	2	90.2	10	7.58	44.9	1610	9.44	56.2	5.82	553
Burkina Faso	2	116	19.2	6.74	29.6	1430	6.81	57.9	5.87	575
Burundi	2	93.6	8.92	11.6	39.2	764	12.3	57.7	6.26	231
Chad	2	150	36.8	4.53	43.5	1930	6.39	56.5	6.59	897
Congo, Dem. Rep.	2	116	41.1	7.91	49.6	609	20.8	57.5	6.54	334
Cote d'Ivoire	2	111	50.6	5.3	43.3	2690	5.39	56.3	5.27	1220
Guinea	2	109	30.3	4.93	43.2	1190	16.1	58	5.34	648
Guinea-Bissau	2	114	14.9	8.5	35.2	1390	2.97	55.6	5.05	547
Haiti	2	208	15.3	6.91	64.7	1500	5.45	32.1	3.33	662
Malawi	2	90.5	22.8	6.59	34.9	1030	12.1	53.1	5.31	459
Mali	2	137	22.8	4.98	35.1	1870	4.37	59.5	6.55	708
Mozambique	2	101	31.5	5.21	46.2	918	7.64	54.5	5.56	419
Niger	2	123	22.2	5.16	49.1	814	2.55	58.8	7.49	348
Rwanda	2	63.6	12	10.5	30	1350	2.61	64.6	4.51	563
Sierra Leone	2	160	16.8	13.1	34.5	1220	17.2	55	5.2	399
Tanzania	2	71.9	18.7	6.01	29.1	2090	9.25	59.3	5.43	702
Togo	2	90.3	40.2	7.65	57.3	1210	1.18	58.7	4.87	488
Uganda	2	81	17.1	9.01	28.6	1540	10.6	56.8	6.15	595
Zambia	2	83.1	37	5.89	30.9	3280	14	52	5.4	1460

