CLUSTERING & PCA

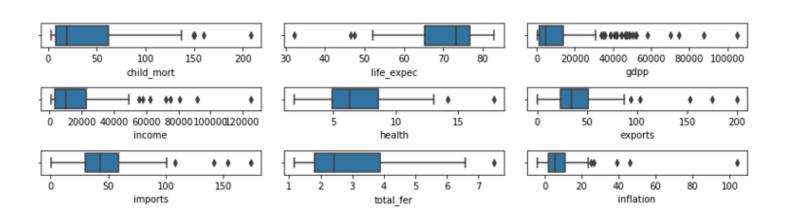
ASSIGNMENT

Objective

Task is to cluster the countries by the factors mentioned below to filter out 5 countries out of 167 who are in direst need of help

- Child mortality rate
- Life expectancy
- The GDP per capita
- Spend on health
- o Income
- Exports
- Imports
- Total fertility rate
- Inflation

All the factors at glance



Outlier Treatment

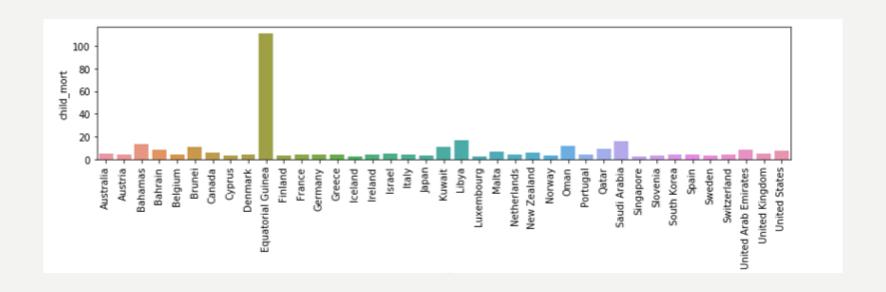
We will keep the outliers for below factors because we are looking for countries lower on these parameters as they are in direst need or help

```
#Child mortality
#life_expectancy
#total_fertility
```

We will remove outlies higher on below factors due to well off and are not the focus of study

```
#gdpp – Gross Domestic Product Per Capita #income
```

Dropped outlier /well off countries Child mortality stats

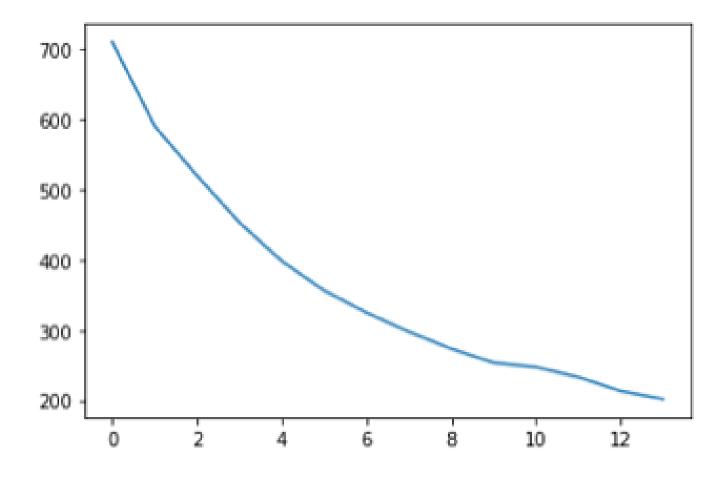


Child Mortality rate of outlier/high on gdpp countries are low except Equatorial Guinea however is oil rich country also added back to data for analysis

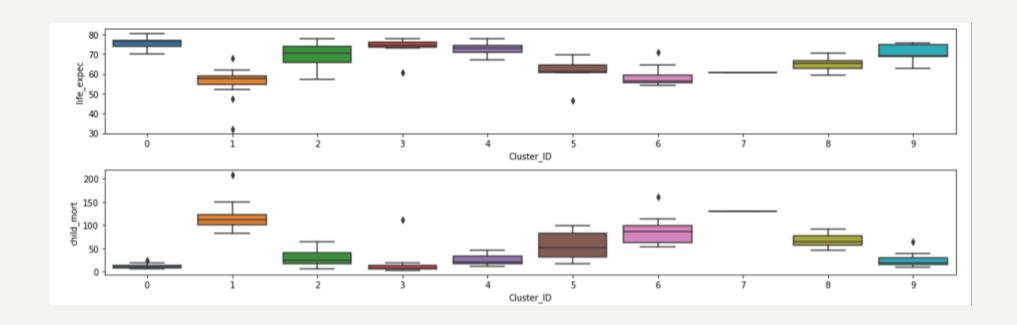
Using PCA to compress the Data in 5 components responsible of 95% variance

	PC1	PC2	PC3	PC4	PC5
0	-2.890823	-0.426885	-0.552403	0.513809	-0.372145
1	1.159019	0.243935	-1.012387	-0.566024	0.447850
2	0.790881	1.457252	0.416624	-0.476767	0.420697
3	-2.045701	0.106737	2.859533	0.025867	-0.144246
4	2.790349	-0.098338	0.194710	0.075022	-0.643379

Choosing clusters
=10 based on
below elbow curve
and seeing need
small set of 5
countries out of
167



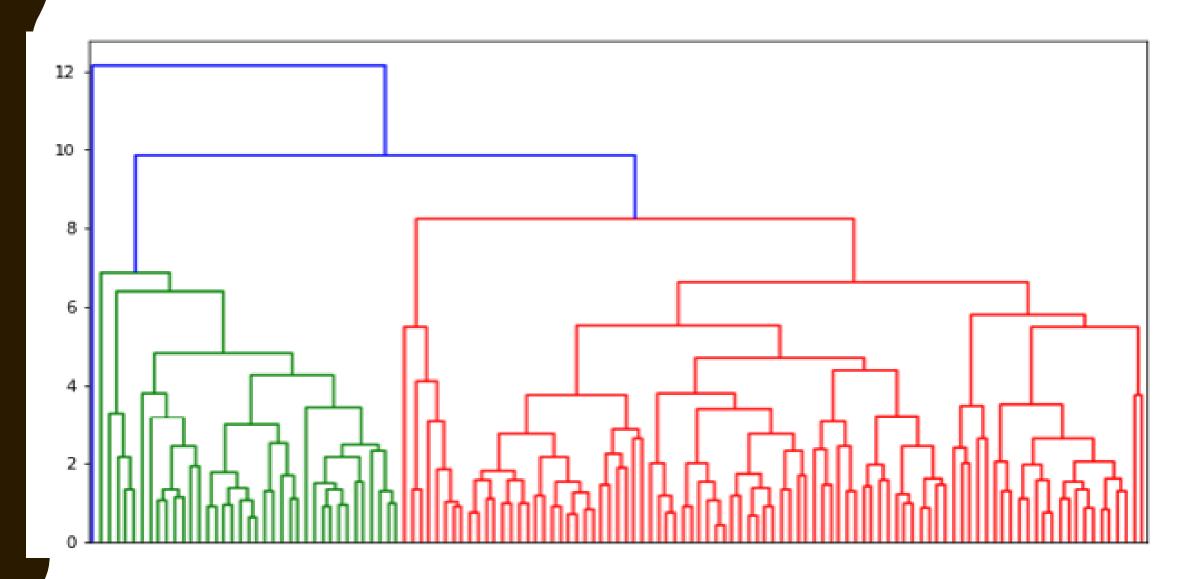
Cluseter-I with low life expectancy, high child mortality rate and low GDP per capita



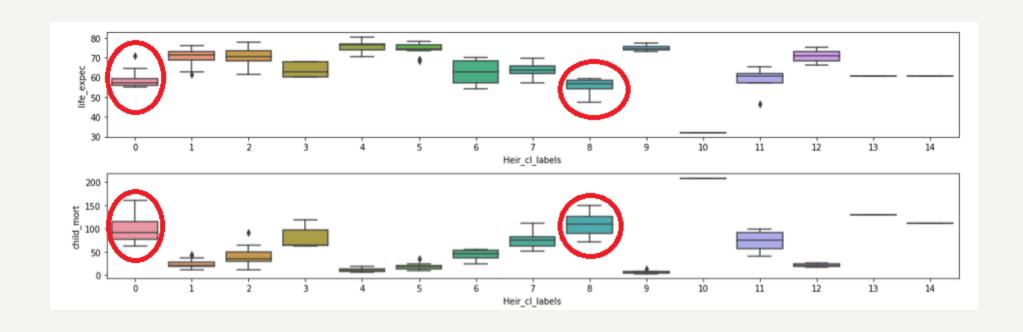
Complete list of Cluster I countries for review based on K Means(cluster=10)

country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	Cluster_ID
Angola	119.0	62.3	2.85	42.9	5900	22.400	60.1	6.16	3530	1
Benin	111.0	23.8	4.10	37.2	1820	0.885	61.8	5.36	758	1
Burkina Faso	116.0	19.2	6.74	29.6	1430	6.810	57.9	5.87	575	1
Cameroon	108.0	22.2	5.13	27.0	2660	1.910	57.3	5.11	1310	1
Central African Republic	149.0	11.8	3.98	26.5	888	2.010	47.5	5.21	446	1
Chad	150.0	36.8	4.53	43.5	1930	6.390	56.5	6.59	897	1
Congo, Dem. Rep.	116.0	41.1	7.91	49.6	609	20.800	57.5	6.54	334	1
Cote d'Ivoire	111.0	50.6	5.30	43.3	2690	5.390	56.3	5.27	1220	1
Guinea	109.0	30.3	4.93	43.2	1190	16.100	58.0	5.34	648	1
Haiti	208.0	15.3	6.91	64.7	1500	5.450	32.1	3.33	662	1
Malawi	90.5	22.8	6.59	34.9	1030	12.100	53.1	5.31	459	1
Mali	137.0	22.8	4.98	35.1	1870	4.370	59.5	6.55	708	1
Mauritania	97.4	50.7	4.41	61.2	3320	18.900	68.2	4.98	1200	1
Mozambique	101.0	31.5	5.21	46.2	918	7.640	54.5	5.56	419	1
Niger	123.0	22.2	5.16	49.1	814	2.550	58.8	7.49	348	1
Togo	90.3	40.2	7.65	57.3	1210	1.180	58.7	4.87	488	1
Zambia	83.1	37.0	5.89	30.9	3280	14.000	52.0	5.40	1460	1
	Angola Benin Burkina Faso Cameroon Central African Republic Chad Congo, Dem. Rep. Cote d'Ivoire Guinea Haiti Malawi Mali Mauritania Mozambique Niger	Angola 119.0 Benin 111.0 Burkina Faso 116.0 Cameroon 108.0 Central African Republic 149.0 Chad 150.0 Congo, Dem. Rep. 116.0 Cote d'Ivoire 111.0 Guinea 109.0 Haiti 208.0 Malawi 90.5 Mali 137.0 Mauritania 97.4 Mozambique 101.0 Niger 123.0 Togo 90.3	Angola 119.0 62.3 Benin 111.0 23.8 Burkina Faso 116.0 19.2 Cameroon 108.0 22.2 Central African Republic 149.0 11.8 Chad 150.0 36.8 Congo, Dem. Rep. 116.0 41.1 Cote d'Ivoire 111.0 50.6 Guinea 109.0 30.3 Haiti 208.0 15.3 Malawi 90.5 22.8 Mali 137.0 22.8 Mauritania 97.4 50.7 Mozambique 101.0 31.5 Niger 123.0 22.2 Togo 90.3 40.2	Angola 119.0 62.3 2.85 Benin 111.0 23.8 4.10 Burkina Faso 116.0 19.2 6.74 Cameroon 108.0 22.2 5.13 Central African Republic 149.0 11.8 3.98 Chad 150.0 36.8 4.53 Congo, Dem. Rep. 116.0 41.1 7.91 Cote d'Ivoire 111.0 50.6 5.30 Guinea 109.0 30.3 4.93 Haiti 208.0 15.3 6.91 Malawi 90.5 22.8 6.59 Mali 137.0 22.8 4.98 Mauritania 97.4 50.7 4.41 Mozambique 101.0 31.5 5.21 Niger 123.0 22.2 5.16 Togo 90.3 40.2 7.65	Angola 119.0 62.3 2.85 42.9 Benin 111.0 23.8 4.10 37.2 Burkina Faso 116.0 19.2 6.74 29.6 Cameroon 108.0 22.2 5.13 27.0 Central African Republic 149.0 11.8 3.98 26.5 Chad 150.0 36.8 4.53 43.5 Congo, Dem. Rep. 116.0 41.1 7.91 49.6 Cote d'Ivoire 111.0 50.6 5.30 43.3 Guinea 109.0 30.3 4.93 43.2 Haiti 208.0 15.3 6.91 64.7 Malawi 90.5 22.8 6.59 34.9 Mali 137.0 22.8 4.98 35.1 Mauritania 97.4 50.7 4.41 61.2 Mozambique 101.0 31.5 5.21 46.2 Niger 123.0 22.2 5.16 49.1 Togo 90.3 40.2 7.65 57.3	Angola 119.0 62.3 2.85 42.9 5900 Benin 111.0 23.8 4.10 37.2 1820 Burkina Faso 116.0 19.2 6.74 29.6 1430 Cameroon 108.0 22.2 5.13 27.0 2660 Central African Republic 149.0 11.8 3.98 26.5 888 Chad 150.0 36.8 4.53 43.5 1930 Congo, Dem. Rep. 116.0 41.1 7.91 49.6 609 Cote d'Ivoire 111.0 50.6 5.30 43.3 2690 Guinea 109.0 30.3 4.93 43.2 1190 Haiti 208.0 15.3 6.91 64.7 1500 Malawi 90.5 22.8 6.59 34.9 1030 Mali 137.0 22.8 4.98 35.1 1870 Mauritania 97.4 50.7 4.41 61.2 3320 Mozambique 101.0 31.5 5.21 46.2 918 Niger 123.0 22.2 5.16 49.1 814 Togo 90.3 40.2 7.65 57.3 1210	Angola 119.0 62.3 2.85 42.9 5900 22.400 Benin 111.0 23.8 4.10 37.2 1820 0.885 Burkina Faso 116.0 19.2 6.74 29.6 1430 6.810 Cameroon 108.0 22.2 5.13 27.0 2660 1.910 Central African Republic 149.0 11.8 3.98 26.5 888 2.010 Chad 150.0 36.8 4.53 43.5 1930 6.390 Congo, Dem. Rep. 116.0 41.1 7.91 49.6 609 20.800 Cote d'Ivoire 111.0 50.6 5.30 43.3 2690 5.390 Guinea 109.0 30.3 4.93 43.2 1190 16.100 Haiti 208.0 15.3 6.91 64.7 1500 5.450 Malawi 90.5 22.8 6.59 34.9 1030 12.100 Mali 137.0 22.8 4.98 35.1 1870 4.370 Mauritania 97.4 50.7 4.41 61.2 3320 18.900 Mozambique 101.0 31.5 5.21 46.2 918 7.640 Niger 123.0 22.2 5.16 49.1 814 2.550 Togo 90.3 40.2 7.65 57.3 1210 1.180	Angola 119.0 62.3 2.85 42.9 5900 22.400 60.1 Benin 111.0 23.8 4.10 37.2 1820 0.885 61.8 Burkina Faso 116.0 19.2 6.74 29.6 1430 6.810 57.9 Cameroon 108.0 22.2 5.13 27.0 2660 1.910 57.3 Central African Republic 149.0 11.8 3.98 26.5 888 2.010 47.5 Chad 150.0 36.8 4.53 43.5 1930 6.390 56.5 Congo, Dem. Rep. 116.0 41.1 7.91 49.6 609 20.800 57.5 Cote d'Ivoire 111.0 50.6 5.30 43.3 2690 5.390 56.3 Guinea 109.0 30.3 4.93 43.2 1190 16.100 58.0 Haiti 208.0 15.3 6.91 64.7 1500 5.450 32.1 Malawi 90.5 22.8 6.59 34.9 1030 12.100 53.1 Mali 137.0 22.8 4.98 35.1 1870 4.370 59.5 Mauritania 97.4 50.7 4.41 61.2 3320 18.900 68.2 Mozambique 101.0 31.5 5.21 46.2 918 7.640 54.5 Niger 123.0 22.2 5.16 49.1 814 2.550 58.8	Angola 119.0 62.3 2.85 42.9 5900 22.400 60.1 6.16 Benin 111.0 23.8 4.10 37.2 1820 0.885 61.8 5.36 Burkina Faso 116.0 19.2 6.74 29.6 1430 6.810 57.9 5.87 Cameroon 108.0 22.2 5.13 27.0 2660 1.910 57.3 5.11 Central African Republic 149.0 11.8 3.98 26.5 888 2.010 47.5 5.21 Chad 150.0 36.8 4.53 43.5 1930 6.390 56.5 6.59 Congo, Dem. Rep. 116.0 41.1 7.91 49.6 609 20.800 57.5 6.54 Cote d'Ivoire 111.0 50.6 5.30 43.3 2690 5.390 56.3 5.27 Guinea 109.0 30.3 4.93 43.2 1190 16.100 58.0 5.34 Haiti 208.0 15.3 6.91 64.7 1500 5.450 32.1 3.33 Malawi 90.5 22.8 6.59 34.9 1030 12.100 53.1 5.31 Mali 137.0 22.8 4.98 35.1 1870 4.370 59.5 6.55 Mauritania 97.4 50.7 4.41 61.2 3320 18.900 68.2 4.98 Mozambique 101.0 31.5 5.21 46.2 918 7.640 54.5 5.56 Niger 123.0 22.2 5.16 49.1 814 2.550 58.8 7.49 Togo 90.3 40.2 7.65 57.3 1210 1.180 58.7 4.87	Angola 119.0 62.3 2.85 42.9 5900 22.400 60.1 6.16 3530 Benin 111.0 23.8 4.10 37.2 1820 0.885 61.8 5.36 758 Burkina Faso 116.0 19.2 6.74 29.6 1430 6.810 57.9 5.87 575 Cameroon 108.0 22.2 5.13 27.0 2660 1.910 57.3 5.11 1310 Central African Republic 149.0 11.8 3.98 26.5 888 2.010 47.5 5.21 446 Chad 150.0 36.8 4.53 43.5 1930 6.390 56.5 6.59 897 Congo, Dem. Rep. 116.0 41.1 7.91 49.6 609 20.800 57.5 6.54 334 Cote d'Ivoire 111.0 50.6 5.30 43.3 2690 5.390 56.3 5.27 1220 Guinea 109.0 30.3 4.93 43.2 1190 16.100 58.0 5.34 648 Haiti 208.0 15.3 6.91 64.7 1500 5.450 32.1 3.33 662 Malawi 90.5 22.8 6.59 34.9 1030 12.100 53.1 5.31 459 Malawi 90.5 22.8 4.98 35.1 1870 4.370 59.5 6.55 708 Mauritania 97.4 50.7 4.41 61.2 3320 18.900 68.2 4.98 1200 Mozambique 101.0 31.5 5.21 46.2 918 7.640 54.5 5.56 419 Niger 123.0 22.2 5.16 49.1 814 2.550 58.8 7.49 348 Togo 90.3 40.2 7.65 57.3 1210 1.180 58.7 4.87 488

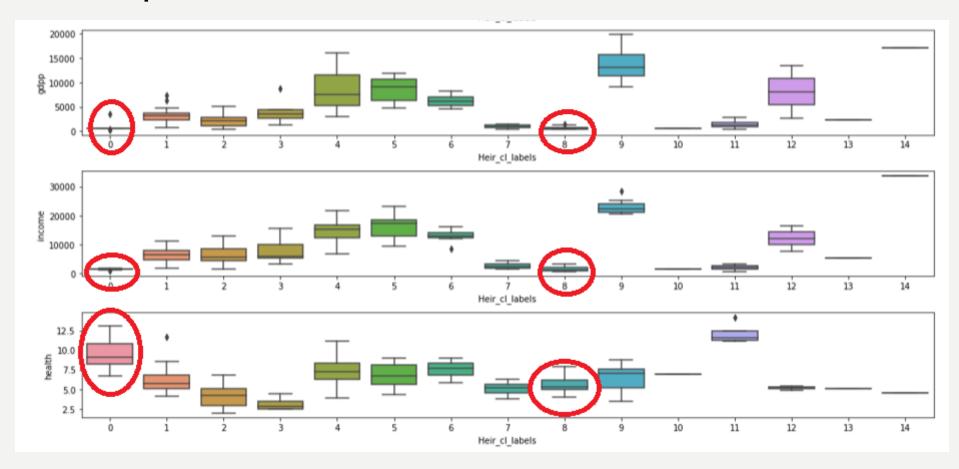
Cut tree for 15 clusters based on below tree structure to get set of countries with lower number of count to select 5 out of them: Hierarchical Clustering



Spotting Cluster 0 and 8 based on low life_expec and high child_mort



Cluster 0 and 8 low on GDPP, Income and health Spend



Cluster8 with total 12 countries with high child_mort, lower GDPP and health spend based on hierarchical clustering

	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	Cluster_ID	Heir_cl_labels
32	Chad	150.0	36.8	4.53	43.5	1930	6.39	56.5	6.59	897	1	8
31	Central African Republic	149.0	11.8	3.98	26.5	888	2.01	47.5	5.21	446	1	8
97	Mali	137.0	22.8	4.98	35.1	1870	4.37	59.5	6.55	708	1	8
112	Niger	123.0	22.2	5.16	49.1	814	2.55	58.8	7.49	348	1	8
37	Congo, Dem. Rep.	116.0	41.1	7.91	49.6	609	20.80	57.5	6.54	334	1	8
40	Cote d'Ivoire	111.0	50.6	5.30	43.3	2690	5.39	56.3	5.27	1220	1	8
63	Guinea	109.0	30.3	4.93	43.2	1190	16.10	58.0	5.34	648	1	8
106	Mozambique	101.0	31.5	5.21	46.2	918	7.64	54.5	5.56	419	1	8
94	Malawi	90.5	22.8	6.59	34.9	1030	12.10	53.1	5.31	459	1	8
150	Togo	90.3	40.2	7.65	57.3	1210	1.18	58.7	4.87	488	1	8
166	Zambia	83.1	37.0	5.89	30.9	3280	14.00	52.0	5.40	1460	1	8
147	Tanzania	71.9	18.7	6.01	29.1	2090	9.25	59.3	5.43	702	8	8

Cluster 0 with total 8 countries with high child_mort, lower gdpp and little better health spend than cluster 8

	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	Cluster_ID	Heir_cl_labels
132	Sierra Leone	160.0	16.80	13.10	34.5	1220	17.20	55.0	5.20	399	6	0
25	Burkina Faso	116.0	19.20	6.74	29.6	1430	6.81	57.9	5.87	575	1	0
64	Guinea-Bissau	114.0	14.90	8.50	35.2	1390	2.97	55.6	5.05	547	6	0
26	Burundi	93.6	8.92	11.60	39.2	764	12.30	57.7	6.26	231	6	0
0	Afghanistan	90.2	10.00	7.58	44.9	1610	9.44	56.2	5.82	553	6	0
155	Uganda	81.0	17.10	9.01	28.6	1540	10.60	56.8	6.15	595	6	0
126	Rwanda	63.6	12.00	10.50	30.0	1350	2.61	64.6	4.51	563	6	0
149	Timor-Leste	62.6	2.20	9.12	27.8	1850	26.50	71.1	6.23	3600	6	0

Selecting final 5 countries 3 from cluster 8 and 2 from cluster 0 of hierarchical clustering based on some subjectivity like severity and stability in these countries . These are part of K Means cluster I as well.

country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp	Cluster_ID	Heir_cl_labels
Sierra Leone	160.0	16.8	13.10	34.5	1220	17.20	55.0	5.20	399	6	0
Chad	150.0	36.8	4.53	43.5	1930	6.39	56.5	6.59	897	1	8
Central African Republic	149.0	11.8	3.98	26.5	888	2.01	47.5	5.21	446	1	8
Burkina Faso	116.0	19.2	6.74	29.6	1430	6.81	57.9	5.87	575	1	0
Malawi	90.5	22.8	6.59	34.9	1030	12.10	53.1	5.31	459	1	8

THANK YOU