SDGs Datasets

Camille, Chetan, Nina, Selim

2 + 1 DATASETS

Detailed SDGs

Initially - 1 091 179 rows × 38 columns

After preprocessing - 184 rows × 173 columns

Summarized SDGs

Initially - 193 rows × 486 columns

After preprocessing - 193 rows × 20 columns

HDI Scores

184 Country, Scores (0 - 1)

Detailed SDGs dataset

Indicators x SGD x year x country

Summarized SDGs dataset

Country	GlobalSDG	SDG1	SDG2	SDG3	SDG4	SDG5	SDG6	SDG7	SDG8	SDG9	SDG10	SDG11	SDG12	SDG13	SDG14	SDG15	SDG16	SDG17
Afghanistan	49.6	NaN	42.93	38.71	23.67	21.35	39.40	56.81	43.20	4.12	NaN	34.20	98.22	96.96	NaN	52.08	46.67	41.87
Angola	51.3	45.16	47.78	33.74	37.84	53.57	45.96	48.75	59.88	5.65	57.18	45.96	93.66	90.89	47.84	64.97	41.72	51.91
Albania	70.3	96.85	46.06	82.21	93.40	53.30	77.25	88.64	62.24	27.97	59.78	78.57	83.18	92.08	40.99	79.96	67.34	64.83
Andorra	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
United Arab Emirates	69.7	NaN	59.81	86.15	85.70	56.33	55.75	90.69	70.62	60.84	85.89	75.99	40.92	33.41	62.50	45.38	81.52	100.00
Argentina	72.4	97.03	60.83	80.19	94.27	77.41	80.23	91.42	72.21	40.54	39.66	83.51	79.12	92.15	39.94	52.98	62.96	85.52
Armenia	68.8	91.21	56.53	78.52	89.75	56.92	66.93	95.35	63.16	33.89	50.06	67.25	90.12	95.01	NaN	61.22	75.44	57.05
Antigua and Barbuda	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Australia	73.9	99.03	52.41	96.45	92.81	78.90	96.98	91.00	81.09	84.20	76.96	80.59	40.85	33.90	56.29	47.76	85.69	61.14
Austria	81.1	99.19	71.58	94.93	96.56	79.05	94.90	93.83	82.03	80.24	87.40	85.82	45.52	84.29	NaN	71.44	92.03	68.02
Azerbaijan	70.5	100.00	58.22	75.29	90.76	53.91	62.80	91.01	68.10	37.30	68.31	83.08	89.11	90.64	22.45	66.88	70.16	69.76
Burundi	51.5	0.00	44.09	47.65	61.02	63.03	54.29	0.00	44.02	3.54	67.07	58.53	96.57	99.35	NaN	73.94	48.86	59.82
Belgium	78.9	99.54	70.24	94.14	94.74	83.90	79.25	91.94	81.41	75.88	93.42	82.33	46.65	82.94	30.65	84.96	86.87	62.29
Benin	50.9	18.31	55.47	47.15	46.95	40.36	46.64	13.36	70.73	8.79	36.21	52.59	94.87	97.35	49.73	84.71	49.01	52.28
Burkina Faso	52.4	24.26	51.89	47.99	27.03	34.62	44.15	9.07	69.21	12.43	78.08	57.85	93.67	90.79	NaN	82.07	57.03	56.21
Bangladesh	60.9	58.21	51.15	59.60	79.75	48.23	65.48	55.97	77.10	15.44	76.38	51.90	96.52	97.12	51.88	60.85	50.48	38.83
Bulgaria	74.5	96.98	58.20	80.18	71.03	69.20	78.01	90.80	80.44	41.22	61.80	81.34	66.58	87.70	65.34	93.31	68.79	75.90
Bahrain	68.7	NaN	67.42	90.37	93.46	50.02	54.54	93.66	82.26	43.23	NaN	52.01	72.08	65.69	57.79	54.79	67.09	60.27
Bahamas, The	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Bosnia and Herzegovina	69.4	99.72	65.01	80.26	99.38	39.86	72.58	80.28	62.19	24.94	82.20	71.81	89.04	72.28	8.67	61.91	72.57	96.91

Correlation matrix

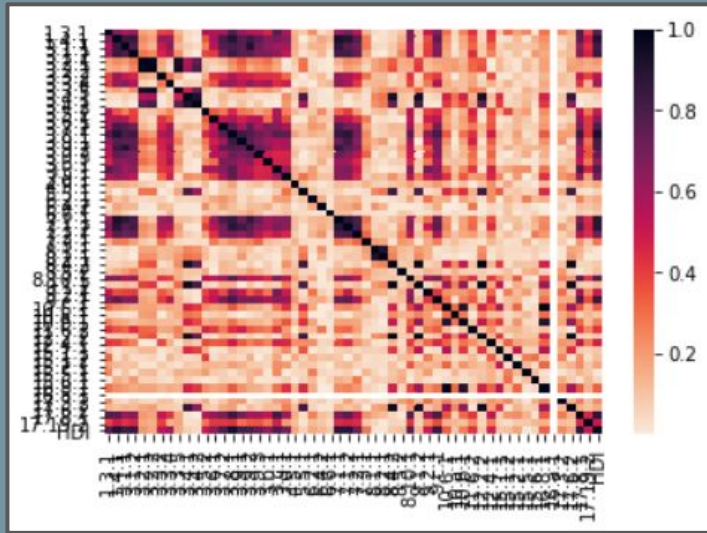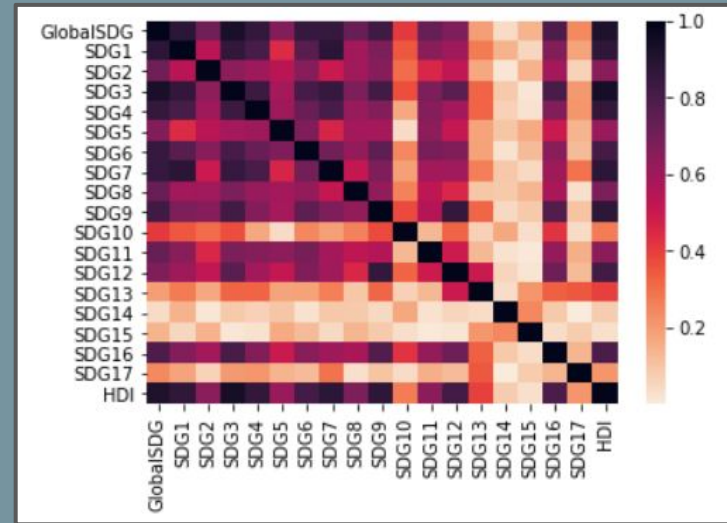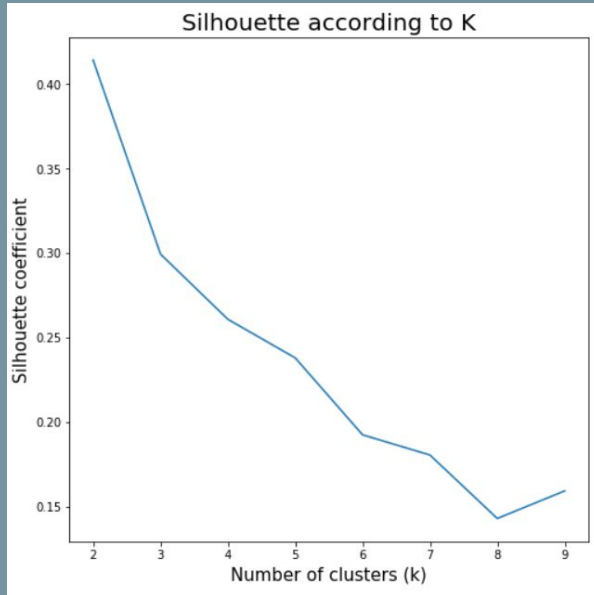**Detailed SDGs dataset**
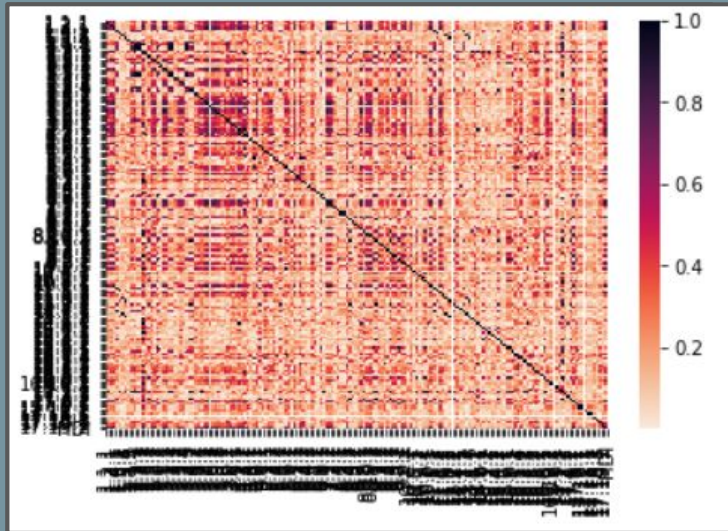
Summarized SDGs dataset
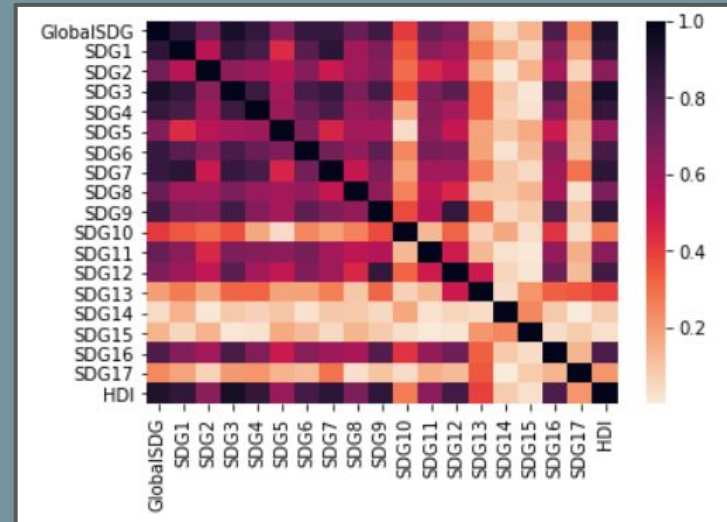
Correlation matrix

Detailed SDGs dataset

Summarized SDGs dataset
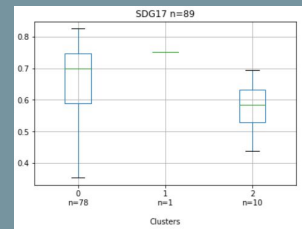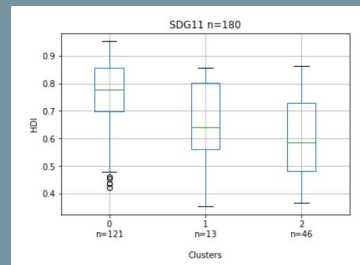
Finding best k for clustering on summarized dataset
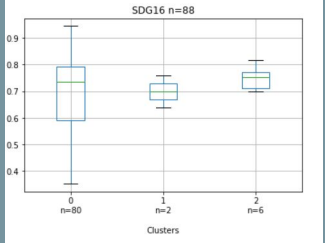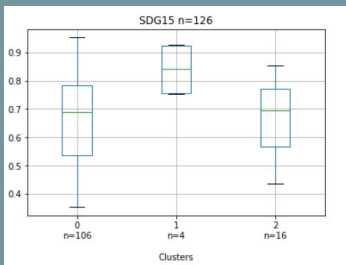

Silhouette according to K

We decided k=3 to have low hdi, middle hdi, and high hdi clusters

Clustering of each SDGs on Detailed data set
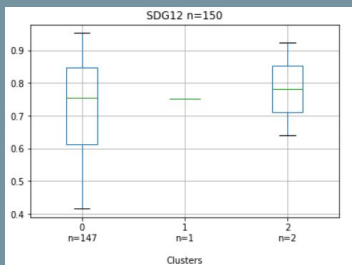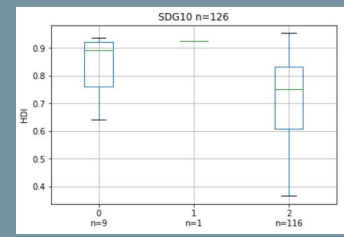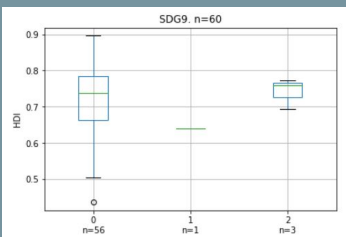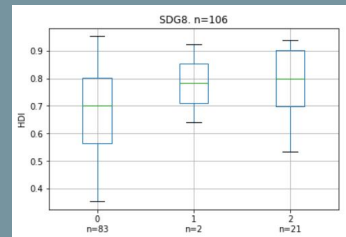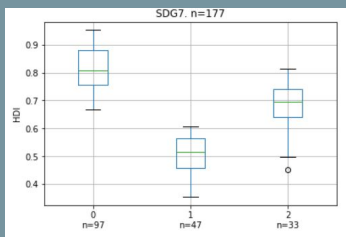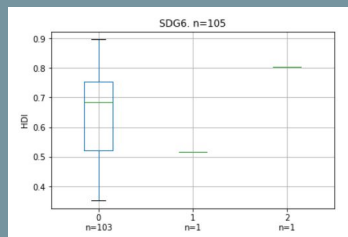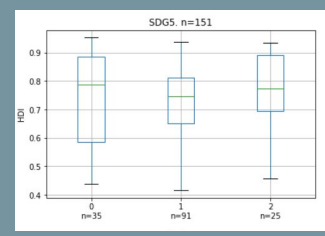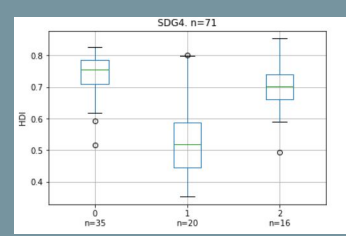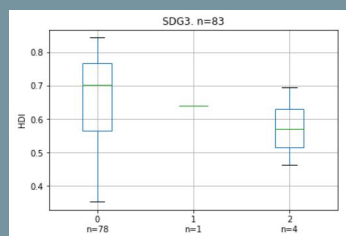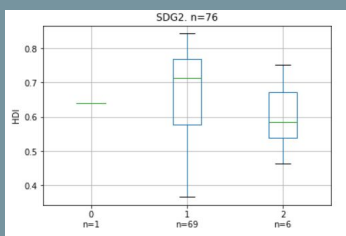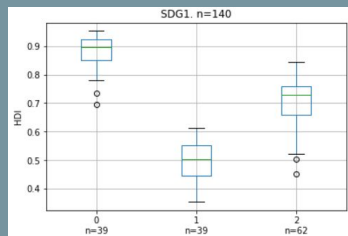
Clustering of each SDGs on Detailed data set

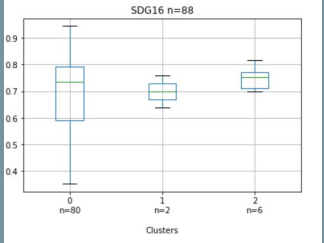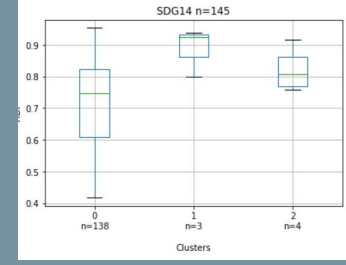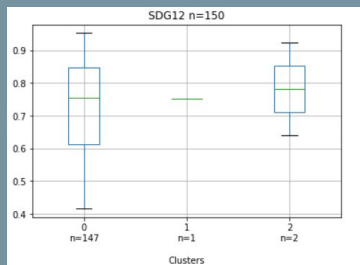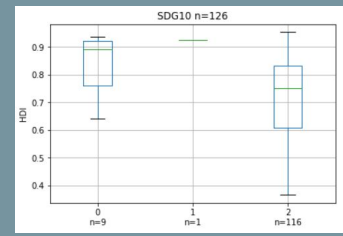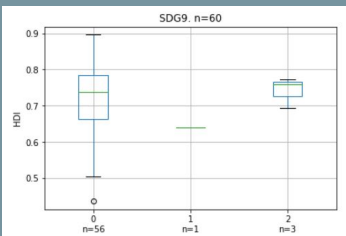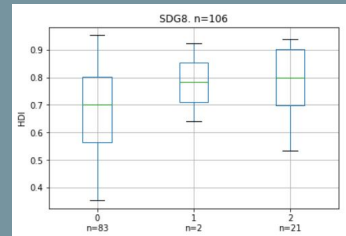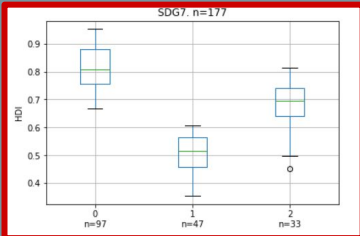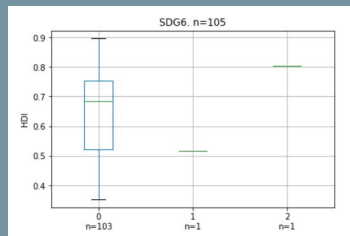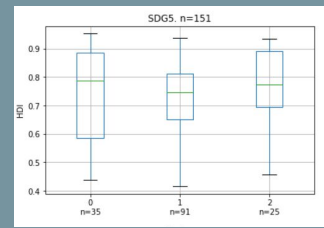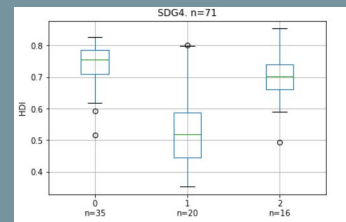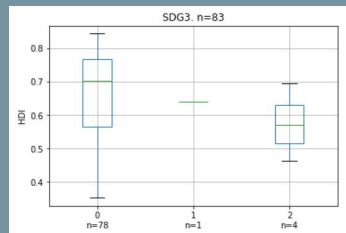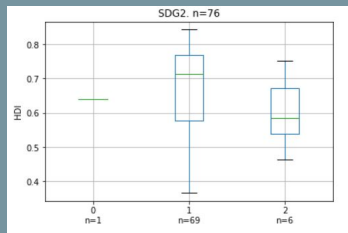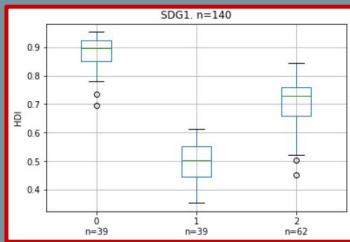# Clustering of all 17 SDGs on Summarized data set
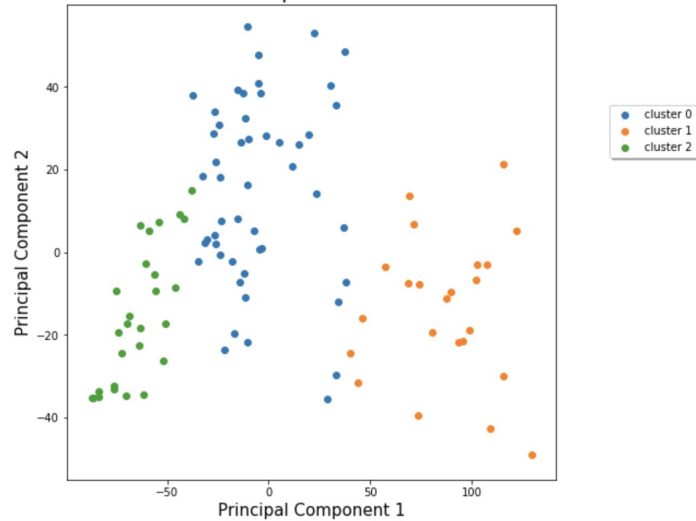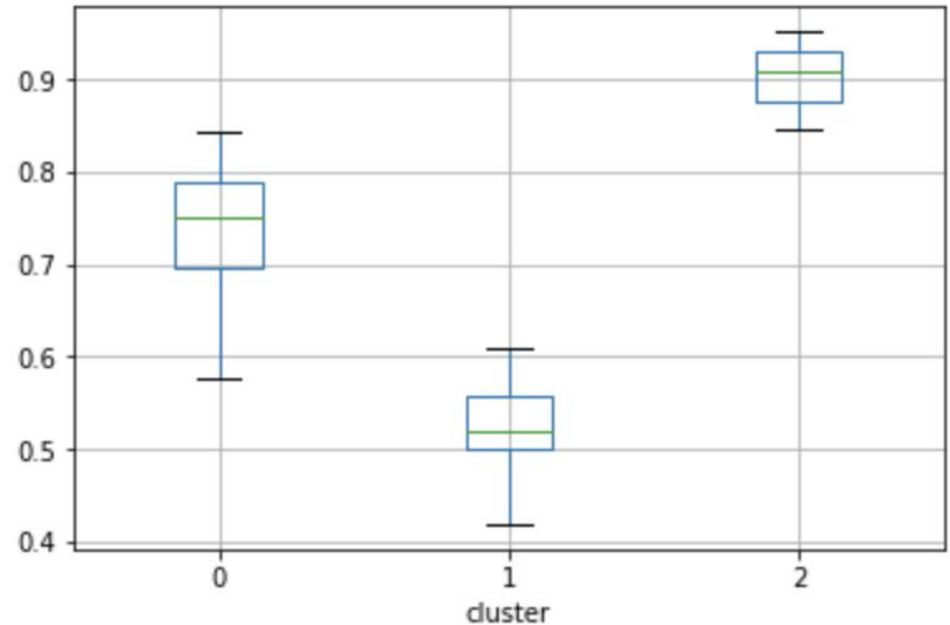
Clustering of all 17 SDGs on Summarized data set

```
Number of rows: 100
                SDG1          SDG2          SDG3          SDG4          SDG5          SDG6
cluster
0           89.207800     53.223400     74.096600     84.775400     61.308200     70.922000
1           36.467826     45.544783     45.371304     46.724783     45.908696     49.218261
2           99.167037     61.970000     93.420000     96.086667     76.942222     86.619630

                SDG7          SDG8          SDG9          SDG10         SDG11         SDG12
cluster
0           82.173200     72.961600     32.955800     49.506200     77.174200     81.897200
1           33.516957     64.732174     12.355652     57.636957     55.054348     93.991739
2           93.232222     81.912222     69.879630     78.091852     84.624074     50.187037

                SDG13         SDG14         SDG15         SDG16         SDG17           HDI
cluster
0           89.163600     49.385200     60.667600     65.737600     67.256600      0.735840
1           93.393043     53.915217     62.086957     54.073913     57.068696      0.521478
2           82.145556     50.011111     75.105556     83.465556     62.238148      0.903222
```

Clustering of highest correlated SDGs to HDI on Summarized data set (1,2,3,4,5,6,7,8,9,11,12,16)

Clustering of highest correlated SDGs to HDI on Summarized data set (1,2,3,4,5,6,7,8,9,11,12,16)

```
Number of rows: 135
            SDG1          SDG2          SDG3          SDG4          SDG5          SDG6  \
cluster
0        86.848406     52.970145     72.454783     84.245797     60.531159     69.653913
1        25.938750     46.042813     41.924062     41.485000     48.832813     47.819687
2        99.194412     62.711176     93.340588     95.822647     76.498529     87.777059

            SDG7          SDG8          SDG9         SDG11         SDG12         SDG16  \
cluster
0        80.484928     72.350145     30.046232     75.602464     82.755072     64.918696
1        24.538750     62.629063     11.283750     54.670938     94.726562     52.637187
2        92.581176     81.356471     70.237059     85.530588     49.480588     83.799412
```

Clustering of highest correlated SDGs to HDI on Summarized data set (1,2,3,4,5,6,7,8,9,11,12,16)

```
Number of rows: 135
            SDG1         SDG2         SDG3         SDG4         SDG5         SDG6   \
cluster
0       86.848406    52.970145    72.454783    84.245797    60.531159    69.653913
1       25.938750    46.042813    41.924062    41.485000    48.832813    47.819687
2       99.194412    62.711176    93.340588    95.822647    76.498529    87.777059

            SDG7         SDG8         SDG9        SDG11        SDG12        SDG16   \
cluster
0       80.484928    72.350145    30.046232    75.602464    82.755072    64.918696
1       24.538750    62.629063    11.283750    54.670938    94.726562    52.637187
2       92.581176    81.356471    70.237059    85.530588    49.480588    83.799412
```
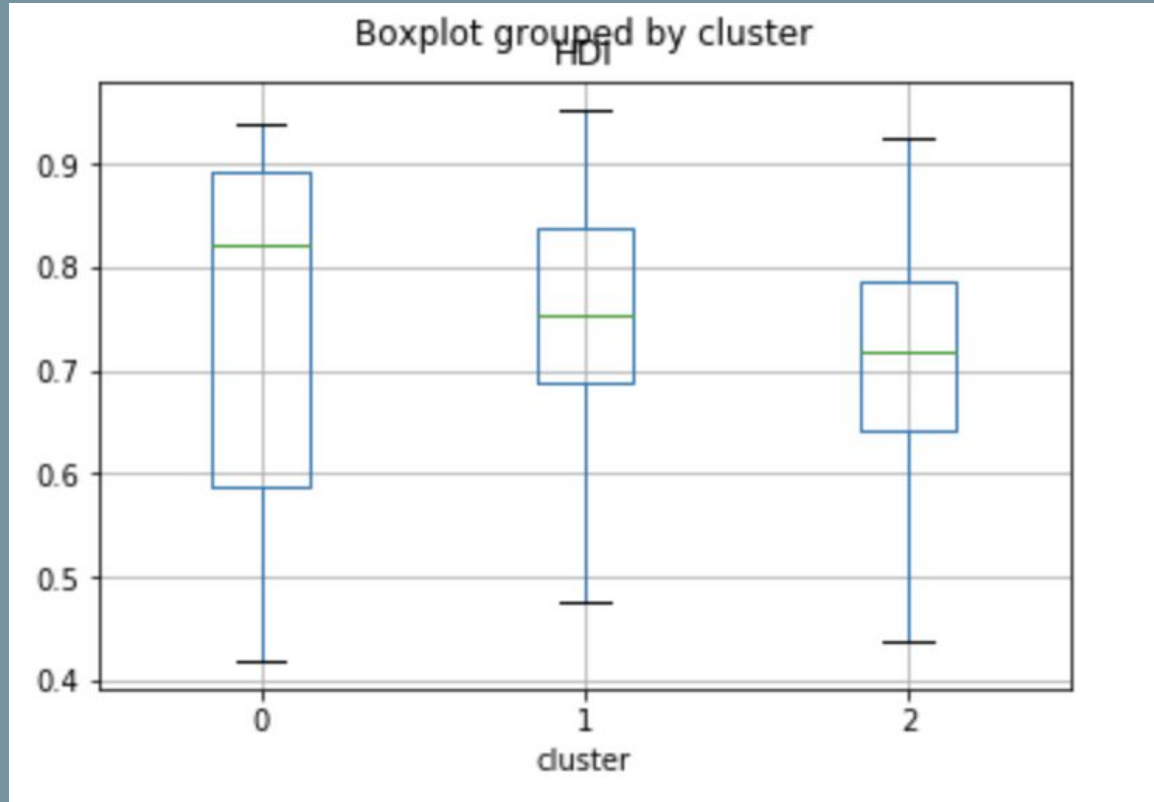
Clustering of lowest correlated SDGs to HDI on Summarized data set (10,13,14,15,17)

Clustering of lowest correlated SDGs to HDI on Summarized data set (10,13,14,15,17)

```
Number of rows: 101
          SDG10        SDG13        SDG14        SDG15        SDG17          HDI
cluster
0      71.718409    89.206818    53.256136    73.823409    59.232500     0.752386
1      75.123478    80.332609    37.479130    49.767391    75.368261     0.746174
2      32.702353    90.729706    56.377059    62.992059    62.239412     0.699088
```

Conclusion

- There is a correlation between SDGS and development index of a country.

- This is especially true for sdgs number 1,2,3,4,5,6,7,8,9,11,16