```
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import matplotlib.cm as cm
import seaborn as sns
import plotly.graph_objects as go
from plotly.subplots import make_subplots
import plotly.express as px
import plotly.offline

data1 = pd.read_csv(r"C:\Users\SAMEERM. CHANNE\Downloads\datascience
fbt notes\projects\AI_index_db.csv")

data1
```

```
                      Country  Talent  Infrastructure  Operating
Environment  \
0   United States of America  100.00           94.02
64.56
1                     China   16.51          100.00
91.57
2             United Kingdom   39.65           71.43
74.65
3                    Canada   31.28           77.05
93.94
4                    Israel   35.76           67.58
82.44
..                      ...     ...             ...
...
57                 Sri Lanka    6.27           34.64
35.79
58                     Egypt    1.11           38.84
0.00
59                     Kenya    0.75           14.11
29.84
60                   Nigeria    2.74            0.00
50.10
61                  Pakistan    8.00            2.43
12.48

    Research  Development  Government Strategy  Commercial  Total
score  \
0     100.00       100.00                77.39      100.00
100.00
1      71.42        79.97                94.87       44.02
62.92
2      36.50        25.03                82.82       18.91
40.93
3      30.67        25.78               100.00       14.88
```

```
40.19
4      32.63         27.96                 43.91        27.33
39.89
..       ...           ...                   ...          ...           ..
.
57      0.12          0.95                 35.57         0.09
6.62
58      2.08          1.54                 68.72         0.31
4.83
59      0.07         12.15                  7.75         0.31
2.30
60      0.45          2.06                  7.75         0.33
1.38
61      2.17          1.09                 13.92         0.27
0.00

          Region                Cluster  Income group     Political
regime
0       Americas          Power players          High    Liberal
democracy
1   Asia-Pacific          Power players  Upper middle     Closed
autocracy
2         Europe  Traditional champions          High    Liberal
democracy
3       Americas  Traditional champions          High    Liberal
democracy
4    Middle East           Rising stars          High    Liberal
democracy
..           ...                    ...           ...
...
57  Asia-Pacific                Nascent  Lower middle  Electoral
democracy
58   Middle East                Nascent  Lower middle  Electoral
autocracy
59        Africa                Nascent  Lower middle  Electoral
autocracy
60        Africa                Nascent  Lower middle  Electoral
autocracy
61  Asia-Pacific                Nascent  Lower middle  Electoral
autocracy

[62 rows x 13 columns]

data1.head()

                     Country  Talent  Infrastructure  Operating
Environment  \
0  United States of America  100.00           94.02
64.56
1                     China   16.51          100.00
```

```
91.57
2            United Kingdom   39.65           71.43
74.65
3                    Canada   31.28           77.05
93.94
4                    Israel   35.76           67.58
82.44

   Research  Development  Government Strategy  Commercial  Total score
\
0    100.00       100.00                77.39      100.00       100.00

1     71.42        79.97                94.87       44.02        62.92

2     36.50        25.03                82.82       18.91        40.93

3     30.67        25.78               100.00       14.88        40.19

4     32.63        27.96                43.91       27.33        39.89


         Region                Cluster  Income group   Political
regime
0      Americas          Power players          High  Liberal
democracy
1  Asia-Pacific          Power players  Upper middle   Closed
autocracy
2        Europe  Traditional champions          High  Liberal
democracy
3      Americas  Traditional champions          High  Liberal
democracy
4   Middle East           Rising stars          High  Liberal
democracy

data1.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 62 entries, 0 to 61
Data columns (total 13 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Country               62 non-null     object
 1   Talent                62 non-null     float64
 2   Infrastructure        62 non-null     float64
 3   Operating Environment  62 non-null     float64
 4   Research              62 non-null     float64
 5   Development           62 non-null     float64
 6   Government Strategy   62 non-null     float64
 7   Commercial            62 non-null     float64
 8   Total score           62 non-null     float64
```

```
 9   Region                  62 non-null     object
 10  Cluster                 62 non-null     object
 11  Income group            62 non-null     object
 12  Political regime        62 non-null     object
dtypes: float64(8), object(5)
memory usage: 6.4+ KB
```

```
data1.describe()
```

```
           Talent  Infrastructure  Operating Environment                 
Research  \
count   62.000000       62.000000              62.000000   62.000000

mean    16.803065       63.503710              66.925484   16.610000

std     15.214963       20.217525              20.000424   17.413996

min      0.000000        0.000000               0.000000    0.000000

25%      7.365000       55.857500              58.107500    3.032500

50%     13.445000       65.230000              69.505000   12.930000

75%     24.567500       75.947500              80.500000   25.412500

max    100.000000      100.000000             100.000000  100.000000


       Development  Government Strategy  Commercial  Total score
count    62.000000            62.000000   62.000000    62.000000
mean     14.824677            57.865645    6.171935    23.914677
std      19.419279            26.252448   14.029632    15.123586
min       0.000000             0.000000    0.000000     0.000000
25%       1.202500            41.030000    0.697500    14.805000
50%       9.005000            63.930000    2.585000    23.220000
75%      19.980000            77.952500    5.307500    30.487500
max     100.000000           100.000000  100.000000   100.000000
```

We knows that

There are no missing values in the dataset , with 62 countries (records) presented there to describe their AI potential as of 2023 There are 8 numeric and 4 categorical variables in the dataset All numeric variables have their values in the range from 0 to 100 Total score variable can be considered to be the target variable whereas other variables can describe the various effects on it

```
data1.isna().sum()
```

```
Country                   0
Talent                    0
Infrastructure            0
```

```
Operating Environment    0
Research                 0
Development              0
Government Strategy      0
Commercial               0
Total score              0
Region                   0
Cluster                  0
Income group             0
Political regime         0
dtype: int64
```

```
data1.isnull()
```

```
    Country  Talent  Infrastructure  Operating Environment
Research  \
0     False   False           False                  False     False

1     False   False           False                  False     False

2     False   False           False                  False     False

3     False   False           False                  False     False

4     False   False           False                  False     False

..      ...     ...             ...                    ...       ...

57    False   False           False                  False     False

58    False   False           False                  False     False

59    False   False           False                  False     False

60    False   False           False                  False     False

61    False   False           False                  False     False


    Development  Government Strategy  Commercial  Total score  Region
\
0         False                False       False        False   False

1         False                False       False        False   False

2         False                False       False        False   False

3         False                False       False        False   False

4         False                False       False        False   False
```

```
..          ...                 ...         ...          ...     ...

57        False               False       False        False   False

58        False               False       False        False   False

59        False               False       False        False   False

60        False               False       False        False   False

61        False               False       False        False   False
```

```
    Cluster  Income group  Political regime
0     False         False             False
1     False         False             False
2     False         False             False
3     False         False             False
4     False         False             False
..      ...           ...               ...
57    False         False             False
58    False         False             False
59    False         False             False
60    False         False             False
61    False         False             False

[62 rows x 13 columns]
```

```
data1.isnull().sum()
```

```
Country                   0
Talent                    0
Infrastructure            0
Operating Environment     0
Research                  0
Development               0
Government Strategy       0
Commercial                0
Total score               0
Region                    0
Cluster                   0
Income group              0
Political regime          0
dtype: int64
```

```
data1.isnull().sum().sum()
```

```
0
```

```
data1.head()
```

```
                    Country  Talent  Infrastructure  Operating
Environment  \
0  United States of America  100.00           94.02
64.56
1                     China   16.51          100.00
91.57
2            United Kingdom   39.65           71.43
74.65
3                    Canada   31.28           77.05
93.94
4                    Israel   35.76           67.58
82.44

   Research  Development  Government Strategy  Commercial  Total score
\
0    100.00       100.00                77.39      100.00       100.00

1     71.42        79.97                94.87       44.02        62.92

2     36.50        25.03                82.82       18.91        40.93

3     30.67        25.78               100.00       14.88        40.19

4     32.63        27.96                43.91       27.33        39.89


         Region                Cluster  Income group   Political
regime
0      Americas          Power players          High  Liberal
democracy
1  Asia-Pacific          Power players  Upper middle   Closed
autocracy
2        Europe  Traditional champions          High  Liberal
democracy
3      Americas  Traditional champions          High  Liberal
democracy
4   Middle East           Rising stars          High  Liberal
democracy
```

```
data1.drop_duplicates(keep= 'first',inplace =True)
```

```
data1
```

```
                    Country  Talent  Infrastructure  Operating
Environment  \
0   United States of America  100.00           94.02
64.56
1                     China   16.51          100.00
91.57
2             United Kingdom   39.65           71.43
74.65
```

```
3                      Canada   31.28           77.05
93.94
4                      Israel   35.76           67.58
82.44
..                        ...     ...             ...
...
57                  Sri Lanka    6.27           34.64
35.79
58                      Egypt    1.11           38.84
0.00
59                      Kenya    0.75           14.11
29.84
60                    Nigeria    2.74            0.00
50.10
61                   Pakistan    8.00            2.43
12.48

    Research  Development  Government Strategy  Commercial  Total
score  \
0     100.00       100.00                77.39      100.00
100.00
1      71.42        79.97                94.87       44.02
62.92
2      36.50        25.03                82.82       18.91
40.93
3      30.67        25.78               100.00       14.88
40.19
4      32.63        27.96                43.91       27.33
39.89
..       ...          ...                   ...         ...           ..
.
57      0.12         0.95                35.57        0.09
6.62
58      2.08         1.54                68.72        0.31
4.83
59      0.07        12.15                 7.75        0.31
2.30
60      0.45         2.06                 7.75        0.33
1.38
61      2.17         1.09                13.92        0.27
0.00

          Region                Cluster  Income group     Political
regime
0       Americas          Power players          High    Liberal
democracy
1   Asia-Pacific          Power players  Upper middle     Closed
autocracy
2         Europe  Traditional champions          High    Liberal
```

```
democracy
3       Americas  Traditional champions          High    Liberal
democracy
4    Middle East           Rising stars          High    Liberal
democracy
..           ...                    ...           ...
...
57  Asia-Pacific                Nascent  Lower middle  Electoral
democracy
58   Middle East                Nascent  Lower middle  Electoral
autocracy
59        Africa                Nascent  Lower middle  Electoral
autocracy
60        Africa                Nascent  Lower middle  Electoral
autocracy
61  Asia-Pacific                Nascent  Lower middle  Electoral
autocracy

[62 rows x 13 columns]
```

```
data1['Cluster'] = data1['Cluster'].str.lower()
```

```
data1
```

```
                     Country  Talent  Infrastructure  Operating
Environment  \
0   United States of America  100.00           94.02
64.56
1                      China   16.51          100.00
91.57
2             United Kingdom   39.65           71.43
74.65
3                     Canada   31.28           77.05
93.94
4                     Israel   35.76           67.58
82.44
..                       ...     ...             ...
...
57                Sri Lanka    6.27           34.64
35.79
58                    Egypt    1.11           38.84
0.00
59                    Kenya    0.75           14.11
29.84
60                  Nigeria    2.74            0.00
50.10
61                 Pakistan    8.00            2.43
12.48

    Research  Development  Government Strategy  Commercial  Total
```

```
score  \
0     100.00       100.00                77.39      100.00
100.00
1      71.42        79.97                94.87       44.02
62.92
2      36.50        25.03                82.82       18.91
40.93
3      30.67        25.78               100.00       14.88
40.19
4      32.63        27.96                43.91       27.33
39.89
..       ...          ...                  ...         ...          ..
.
57      0.12         0.95                35.57        0.09
6.62
58      2.08         1.54                68.72        0.31
4.83
59      0.07        12.15                 7.75        0.31
2.30
60      0.45         2.06                 7.75        0.33
1.38
61      2.17         1.09                13.92        0.27
0.00

          Region                Cluster  Income group     Political
regime
0       Americas          power players          High    Liberal
democracy
1   Asia-Pacific          power players  Upper middle     Closed
autocracy
2         Europe  traditional champions          High    Liberal
democracy
3       Americas  traditional champions          High    Liberal
democracy
4    Middle East           rising stars          High    Liberal
democracy
..           ...                   ...           ...
...
57  Asia-Pacific               nascent  Lower middle  Electoral
democracy
58   Middle East               nascent  Lower middle  Electoral
autocracy
59        Africa               nascent  Lower middle  Electoral
autocracy
60        Africa               nascent  Lower middle  Electoral
autocracy
61  Asia-Pacific               nascent  Lower middle  Electoral
autocracy
```

```
[62 rows x 13 columns]
```

```
data1['Operating Environment'].fillna(0, inplace=True)
```

```
data1
```

```
                     Country  Talent  Infrastructure  Operating
Environment  \
0   United States of America  100.00           94.02
64.56
1                      China   16.51          100.00
91.57
2             United Kingdom   39.65           71.43
74.65
3                     Canada   31.28           77.05
93.94
4                     Israel   35.76           67.58
82.44
..                       ...     ...             ...
...
57                 Sri Lanka    6.27           34.64
35.79
58                     Egypt    1.11           38.84
0.00
59                     Kenya    0.75           14.11
29.84
60                   Nigeria    2.74            0.00
50.10
61                  Pakistan    8.00            2.43
12.48

    Research  Development  Government Strategy  Commercial  Total
score  \
0     100.00       100.00                77.39      100.00
100.00
1      71.42        79.97                94.87       44.02
62.92
2      36.50        25.03                82.82       18.91
40.93
3      30.67        25.78               100.00       14.88
40.19
4      32.63        27.96                43.91       27.33
39.89
..       ...          ...                   ...         ...          ..
.
57      0.12         0.95                35.57        0.09
6.62
58      2.08         1.54                68.72        0.31
4.83
```

```
59      0.07        12.15                  7.75        0.31
2.30
60      0.45         2.06                  7.75        0.33
1.38
61      2.17         1.09                 13.92        0.27
0.00

          Region                Cluster  Income group     Political
regime
0       Americas          power players          High    Liberal
democracy
1   Asia-Pacific          power players  Upper middle     Closed
autocracy
2         Europe  traditional champions          High    Liberal
democracy
3       Americas  traditional champions          High    Liberal
democracy
4    Middle East           rising stars          High    Liberal
democracy
..           ...                    ...           ...
...
57  Asia-Pacific                nascent  Lower middle  Electoral
democracy
58   Middle East                nascent  Lower middle  Electoral
autocracy
59        Africa                nascent  Lower middle  Electoral
autocracy
60        Africa                nascent  Lower middle  Electoral
autocracy
61  Asia-Pacific                nascent  Lower middle  Electoral
autocracy

[62 rows x 13 columns]
```

```
fig = make_subplots(rows=2, cols=4, subplot_titles=('<b>Distribution
of Talent</b>',
                                                    '<b>Distribution
of Infrastructure</b>',
                                                    '<b>Distribution
of Operating Environment</b>',
                                                    '<b>Distribution
of Research</b>',
                                                    '<b>Distribution
of Development</b>',
                                                    '<b>Distribution
of Government Strategy</b>',
                                                    '<b>Distribution
of Commercial</b>',
                                                    '<b>Distribution
of Total score</b>'
```

```
                                                    ))

fig.add_trace(go.Histogram(x=data1['Talent'], nbinsx=30), row=1,
col=1)
fig.add_trace(go.Histogram(x=data1['Infrastructure']), row=1, col=2)
fig.add_trace(go.Histogram(x=data1['Operating Environment'],
nbinsx=30), row=1, col=3)
fig.add_trace(go.Histogram(x=data1['Research'], nbinsx=30), row=1,
col=4)

fig.add_trace(go.Histogram(x=data1['Development'], nbinsx=30), row=2,
col=1)
fig.add_trace(go.Histogram(x=data1['Government Strategy']), row=2,
col=2)
fig.add_trace(go.Histogram(x=data1['Commercial'], nbinsx=30), row=2,
col=3)
fig.add_trace(go.Histogram(x=data1['Total score'], nbinsx=30), row=2,
col=4)


fig.update_layout(
    showlegend=False,
    width=800,
    height=500,
    autosize=False,
    margin=dict(t=15, b=0, l=5, r=5),
    template="plotly_white",
)

fig.update_coloraxes(colorbar_tickfont_size=10)
fig.update_annotations(font_size=10)
fig.update_traces(opacity=0.75)

fig.show()
```


```
{"config":{"plotlyServerURL":"https://plot.ly"},"data":
[{"nbinsx":30,"opacity":0.75,"type":"histogram","x":
[100,16.51,39.65,31.28,35.76,39.38,14.54,33.83,27.63,28.32,25.43,29.93
,24.99,27.07,21.66,15.18,45.27,25.63,28.21,17.56,17.61,16.97,18.74,12.
34,27.61,4.49,15.17,14.21,13.02,23.3,11.09,12.46,15.87,2.65,13.43,11.1
1,18.45,14.3,13.46,7.62,8.55,10.34,10.44,1.72,10.56,8.4,0,8.49,4.97,7.
28,4.99,6.34,5.51,8.94,4.61,3.36,6.69,6.27,1.11,0.75,2.74,8],"xaxis":"
x","yaxis":"y"},{"opacity":0.75,"type":"histogram","x":
[94.02,100,71.43,77.05,67.58,84.3,85.23,81.99,77.22,77.15,63.43,89.5,7
1.6,74.08,94.88,84.58,33.91,78.43,75.19,96.11,73.32,64.49,63.65,77.86,
76.2,70.8,65.1,70.96,72.06,69.78,64.76,62.59,67.12,79.16,64.2,64.26,72
.45,63.19,62.61,55.44,65.36,69.17,62.04,41.85,61.97,56.15,67.97,40.74,
54.8,58.77,60.39,55.76,47.52,38.59,45.73,44.88,37.84,34.64,38.84,14.11
,0,2.43],"xaxis":"x2","yaxis":"y2"},
{"nbinsx":30,"opacity":0.75,"type":"histogram","x":
```

```
[64.56,91.57,74.65,93.94,82.44,43.15,68.86,88.05,70.22,80.02,61.23,70.
15,78.76,85.39,66.96,57.53,77.3,44.14,66.77,59.5,75.36,76.3,88.67,56.6
7,36.65,100,64.08,99.56,94.55,90.35,83.25,52.85,70.96,72.12,80.66,76.9
7,41.19,80.67,72.82,83.58,88.71,58.01,73.24,97.03,56.73,75.95,62.58,64
.34,62.37,70.75,60.9,37.61,51.18,62.32,58.43,60.17,58.4,35.79,0,29.84,
50.1,12.48],"xaxis":"x3","yaxis":"y3"},
{"nbinsx":30,"opacity":0.75,"type":"histogram","x":
[100,71.42,36.5,30.67,32.63,37.67,26.66,25.54,35.84,25.48,32.63,16.79,
25.21,26.01,19.39,22.51,18.92,38.24,27.61,31.51,18.6,23.56,11.75,25.71
,21.18,13.63,22.15,10.6,19.1,12.23,20.3,14.21,5.96,5.13,8.96,11.26,18.
29,3.22,4.83,15.12,2.97,4.31,5.57,8.11,1.49,1.25,11.94,9.53,0,0.73,2.5
3,2.03,0.98,3.9,0.83,1.46,0.28,0.12,2.08,7.0e-
2,0.45,2.17],"xaxis":"x4","yaxis":"y4"},
{"nbinsx":30,"opacity":0.75,"type":"histogram","x":
[100,79.97,25.03,25.78,27.96,22.55,77.25,30.17,24.79,21.44,41.15,30.85
,18.32,8.92,19.95,34.47,30.86,23.11,17.81,8.63,10.87,17.81,9.31,19.99,
13.56,14.38,19.81,9.09,1.06,5.96,14.66,19.48,11.72,15.53,3.92,2.7,0.19
,6.18,5.07,2.21,0.34,5.4,0.88,4.46,0.67,3.19,0,1.02,0.89,0.26,0,0.3,3.
52,0,7.52,5.0e-
2,0.33,0.95,1.54,12.15,2.06,1.09],"xaxis":"x5","yaxis":"y5"},
{"opacity":0.75,"type":"histogram","x":
[77.39,94.87,82.82,100,43.91,79.82,87.5,62.35,84.65,91.2,82.11,69.44,8
5.99,74.23,66.69,71.96,58.83,12.18,40.35,33.29,91.28,72.14,72.08,55.97
,59.05,91.63,63.58,78.14,80.38,47.62,61.43,90.4,70.49,81.38,70.69,70.2
9,22.15,64.28,67.72,22.15,43.07,55.01,47.6,54.21,60.5,54.94,33.49,67.4
5,85.29,30.92,17.72,68.86,59.99,12.18,0,15.9,14.4,35.57,68.72,7.75,7.7
5,13.92],"xaxis":"x6","yaxis":"y6"},
{"nbinsx":30,"opacity":0.75,"type":"histogram","x":
[100,44.02,18.91,14.88,27.33,15.07,5.41,4.97,8.29,7.65,6.72,3.94,4.64,
3.46,4.68,7.31,7.39,7.76,4.51,5.3,3.08,3.08,12.51,2.53,3.95,4.73,5.31,
2.25,0.61,2.49,2.64,1.38,4.3,3.22,2.05,1.75,5.74,1.77,1.36,0.92,0.67,1
.08,0.63,0.78,2.95,0.34,0,0.95,0.5,0.43,0.24,0.31,0.91,0.15,2.03,0.1,1
.37,9.0e-2,0.31,0.31,0.33,0.27],"xaxis":"x7","yaxis":"y7"},
{"nbinsx":30,"opacity":0.75,"type":"histogram","x":
[100,62.92,40.93,40.19,39.89,38.67,38.6,36.35,36.04,34.42,33.86,33.04,
31.36,30.87,30.73,30.53,30.36,30.25,29.85,29.11,26.95,26.89,26.6,25.79
,25.77,25.6,25.52,25.2,25.19,24.88,24.45,21.99,21.85,21.17,20.89,20.31
,19.81,19.59,18.89,17.33,17.24,17,16.66,15.48,15.33,15.24,14.66,14.39,
13.85,13.27,11.79,11.62,11.47,9.87,9.71,8.87,8.49,6.62,4.83,2.3,1.38,0
],"xaxis":"x8","yaxis":"y8"}],"layout":{"annotations":[{"font":
{"size":10},"showarrow":false,"text":"<b>Distribution of
Talent</b>","x":0.10625,"xanchor":"center","xref":"paper","y":1,"yanch
or":"bottom","yref":"paper"},{"font":
{"size":10},"showarrow":false,"text":"<b>Distribution of
Infrastructure</b>","x":0.36875,"xanchor":"center","xref":"paper","y":
1,"yanchor":"bottom","yref":"paper"},{"font":
{"size":10},"showarrow":false,"text":"<b>Distribution of Operating
Environment</b>","x":0.6312500000000001,"xanchor":"center","xref":"pap
er","y":1,"yanchor":"bottom","yref":"paper"},{"font":
```

{"size":10},"showarrow":false,"text":"Distribution of Research","x":0.89375,"xanchor":"center","xref":"paper","y":1,"yanchor":"bottom","yref":"paper"},{"font":{"size":10},"showarrow":false,"text":"Distribution of Development","x":0.10625,"xanchor":"center","xref":"paper","y":0.375,"yanchor":"bottom","yref":"paper"},{"font":{"size":10},"showarrow":false,"text":"Distribution of Government Strategy","x":0.36875,"xanchor":"center","xref":"paper","y":0.375,"yanchor":"bottom","yref":"paper"},{"font":{"size":10},"showarrow":false,"text":"Distribution of Commercial","x":0.6312500000000001,"xanchor":"center","xref":"paper","y":0.375,"yanchor":"bottom","yref":"paper"},{"font":{"size":10},"showarrow":false,"text":"Distribution of Total score","x":0.89375,"xanchor":"center","xref":"paper","y":0.375,"yanchor":"bottom","yref":"paper"}],"autosize":false,"coloraxis":{"colorbar":{"tickfont":{"size":10}}},"height":500,"margin":{"b":0,"l":5,"r":5,"t":15},"showlegend":false,"template":{"data":{"bar":[{"error_x":{"color":"#2a3f5f"},"error_y":{"color":"#2a3f5f"},"marker":{"line":{"color":"white","width":0.5},"pattern":{"fillmode":"overlay","size":10,"solidity":0.2}},"type":"bar"}],"barpolar":[{"marker":{"line":{"color":"white","width":0.5},"pattern":{"fillmode":"overlay","size":10,"solidity":0.2}},"type":"barpolar"}],"carpet":[{"aaxis":{"endlinecolor":"#2a3f5f","gridcolor":"#C8D4E3","linecolor":"#C8D4E3","minorgridcolor":"#C8D4E3","startlinecolor":"#2a3f5f"},"baxis":{"endlinecolor":"#2a3f5f","gridcolor":"#C8D4E3","linecolor":"#C8D4E3","minorgridcolor":"#C8D4E3","startlinecolor":"#2a3f5f"},"type":"carpet"}],"choropleth":[{"colorbar":{"outlinewidth":0,"ticks":""},"type":"choropleth"}],"contour":[{"colorbar":{"outlinewidth":0,"ticks":""},"colorscale":[[0,"#0d0887"],[0.1111111111111111,"#46039f"],[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],[0.8888888888888888,"#fdca26"],[1,"#f0f921"]],"type":"contour"}],"contourcarpet":[{"colorbar":{"outlinewidth":0,"ticks":""},"type":"contourcarpet"}],"heatmap":[{"colorbar":{"outlinewidth":0,"ticks":""},"colorscale":[[0,"#0d0887"],[0.1111111111111111,"#46039f"],[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],[0.8888888888888888,"#fdca26"],[1,"#f0f921"]],"type":"heatmap"}],"heatmapgl":[{"colorbar":{"outlinewidth":0,"ticks":""},"colorscale":[[0,"#0d0887"],[0.1111111111111111,"#46039f"],[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],[0.6666666666666666,"#ed7953"],

```
[0.7777777777777778,"#fb9f3a"],[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"heatmapgl"}],"histogram":[{"marker":{"pattern":
{"fillmode":"overlay","size":10,"solidity":0.2}},"type":"histogram"}],
"histogram2d":[{"colorbar":{"outlinewidth":0,"ticks":""},"colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"histogram2d"}],"histogram2dcontour":
[{"colorbar":{"outlinewidth":0,"ticks":""},"colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"histogram2dcontour"}],"mesh3d":[{"colorbar":
{"outlinewidth":0,"ticks":""},"type":"mesh3d"}],"parcoords":[{"line":
{"colorbar":{"outlinewidth":0,"ticks":""}},"type":"parcoords"}],"pie":
[{"automargin":true,"type":"pie"}],"scatter":[{"fillpattern":
{"fillmode":"overlay","size":10,"solidity":0.2},"type":"scatter"}],"sc
atter3d":[{"line":{"colorbar":{"outlinewidth":0,"ticks":""}},"marker":
{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scatter3d"}],"scattercarpet":
[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scattercarpet"}],"scattergeo":
[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scattergeo"}],"scattergl":
[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scattergl"}],"scattermapbox":
[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scattermapbox"}],"scatterpolar"
:[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scatterpolar"}],"scatterpolargl
":[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scatterpolargl"}],"scatterterna
ry":[{"marker":{"colorbar":
{"outlinewidth":0,"ticks":""}},"type":"scatterternary"}],"surface":
[{"colorbar":{"outlinewidth":0,"ticks":""},"colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"surface"}],"table":[{"cells":{"fill":
{"color":"#EBF0F8"},"line":{"color":"white"}},"header":{"fill":
{"color":"#C8D4E3"},"line":
{"color":"white"}},"type":"table"}]},"layout":{"annotationdefaults":
{"arrowcolor":"#2a3f5f","arrowhead":0,"arrowwidth":1},"autotypenumbers
```

```
":"strict","coloraxis":{"colorbar":
{"outlinewidth":0,"ticks":""}},"colorscale":{"diverging":
[[0,"#8e0152"],[0.1,"#c51b7d"],[0.2,"#de77ae"],[0.3,"#f1b6da"],
[0.4,"#fde0ef"],[0.5,"#f7f7f7"],[0.6,"#e6f5d0"],[0.7,"#b8e186"],
[0.8,"#7fbc41"],[0.9,"#4d9221"],[1,"#276419"]],"sequential":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],[1,"#f0f921"]],"sequentialminus":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],[1,"#f0f921"]]},"colorway":
["#636efa","#EF553B","#00cc96","#ab63fa","#FFA15A","#19d3f3","#FF6692"
,"#B6E880","#FF97FF","#FECB52"],"font":{"color":"#2a3f5f"},"geo":
{"bgcolor":"white","lakecolor":"white","landcolor":"white","showlakes"
:true,"showland":true,"subunitcolor":"#C8D4E3"},"hoverlabel":
{"align":"left"},"hovermode":"closest","mapbox":
{"style":"light"},"paper_bgcolor":"white","plot_bgcolor":"white","pola
r":{"angularaxis":
{"gridcolor":"#EBF0F8","linecolor":"#EBF0F8","ticks":""},"bgcolor":"wh
ite","radialaxis":
{"gridcolor":"#EBF0F8","linecolor":"#EBF0F8","ticks":""}},"scene":
{"xaxis":
{"backgroundcolor":"white","gridcolor":"#DFE8F3","gridwidth":2,"lineco
lor":"#EBF0F8","showbackground":true,"ticks":"","zerolinecolor":"#EBF0
F8"},"yaxis":
{"backgroundcolor":"white","gridcolor":"#DFE8F3","gridwidth":2,"lineco
lor":"#EBF0F8","showbackground":true,"ticks":"","zerolinecolor":"#EBF0
F8"},"zaxis":
{"backgroundcolor":"white","gridcolor":"#DFE8F3","gridwidth":2,"lineco
lor":"#EBF0F8","showbackground":true,"ticks":"","zerolinecolor":"#EBF0
F8"}},"shapedefaults":{"line":{"color":"#2a3f5f"}},"ternary":{"aaxis":
{"gridcolor":"#DFE8F3","linecolor":"#A2B1C6","ticks":""},"baxis":
{"gridcolor":"#DFE8F3","linecolor":"#A2B1C6","ticks":""},"bgcolor":"wh
ite","caxis":
{"gridcolor":"#DFE8F3","linecolor":"#A2B1C6","ticks":""}},"title":
{"x":5.0e-2},"xaxis":
{"automargin":true,"gridcolor":"#EBF0F8","linecolor":"#EBF0F8","ticks"
:"","title":
{"standoff":15},"zerolinecolor":"#EBF0F8","zerolinewidth":2},"yaxis":
{"automargin":true,"gridcolor":"#EBF0F8","linecolor":"#EBF0F8","ticks"
:"","title":
{"standoff":15},"zerolinecolor":"#EBF0F8","zerolinewidth":2}}},"width"
:800,"xaxis":{"anchor":"y","domain":[0,0.2125]},"xaxis2":
{"anchor":"y2","domain":[0.2625,0.475]},"xaxis3":
{"anchor":"y3","domain":[0.525,0.7375]},"xaxis4":
```

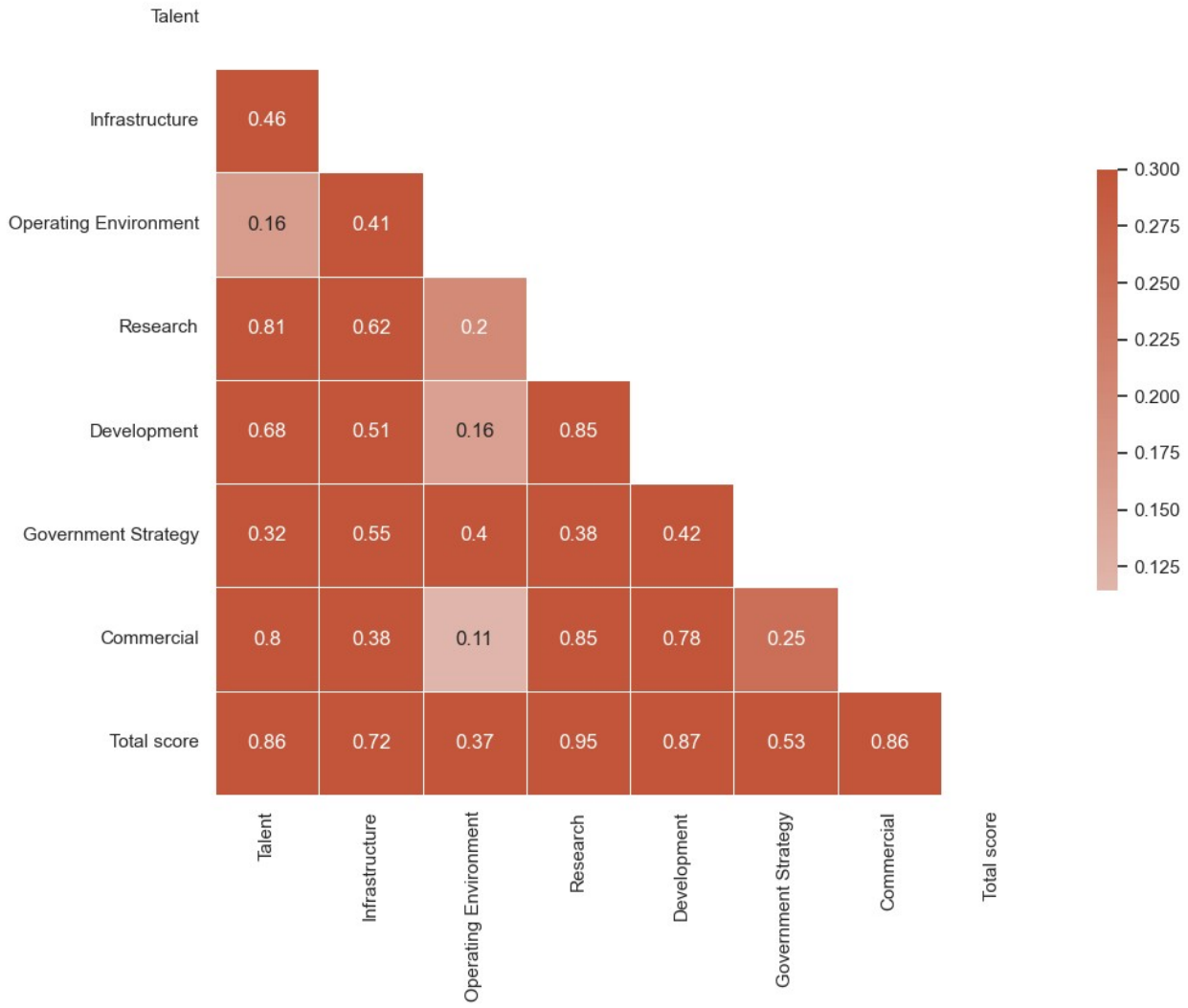
{"anchor":"y4","domain":[0.7875,1]},"xaxis5":{"anchor":"y5","domain":
[0,0.2125]},"xaxis6":{"anchor":"y6","domain":[0.2625,0.475]},"xaxis7":
{"anchor":"y7","domain":[0.525,0.7375]},"xaxis8":
{"anchor":"y8","domain":[0.7875,1]},"yaxis":{"anchor":"x","domain":
[0.625,1]},"yaxis2":{"anchor":"x2","domain":[0.625,1]},"yaxis3":
{"anchor":"x3","domain":[0.625,1]},"yaxis4":{"anchor":"x4","domain":
[0.625,1]},"yaxis5":{"anchor":"x5","domain":[0,0.375]},"yaxis6":
{"anchor":"x6","domain":[0,0.375]},"yaxis7":{"anchor":"x7","domain":
[0,0.375]},"yaxis8":{"anchor":"x8","domain":[0,0.375]}}}

we knows that

Infrastructure and Operating Environment are skewened to the right a little Other numeric variables are significantly left-skewed The values of Talent, Research, Development, Commercial and Total score have clear outliers to the right side of the distribution (it is related to two countries in Power players cluster)

```
sns.set_theme(style="white")

d = data1.copy()

corr = d.corr()


mask = np.triu(np.ones_like(corr, dtype=bool))


f, ax = plt.subplots(figsize=(11, 9))


cmap = sns.diverging_palette(230, 20, as_cmap=True)

sns.heatmap(corr, mask=mask, cmap=cmap, vmax=.3, center=0, annot=True,
            square=True, linewidths=.5, cbar_kws={"shrink": .5})
```

```
C:\Users\SAMEERM. CHANNE\AppData\Local\Temp\
ipykernel_760\2920554842.py:5: FutureWarning:

The default value of numeric_only in DataFrame.corr is deprecated. In
a future version, it will default to False. Select only valid columns
or specify the value of numeric_only to silence this warning.


<Axes: >
```



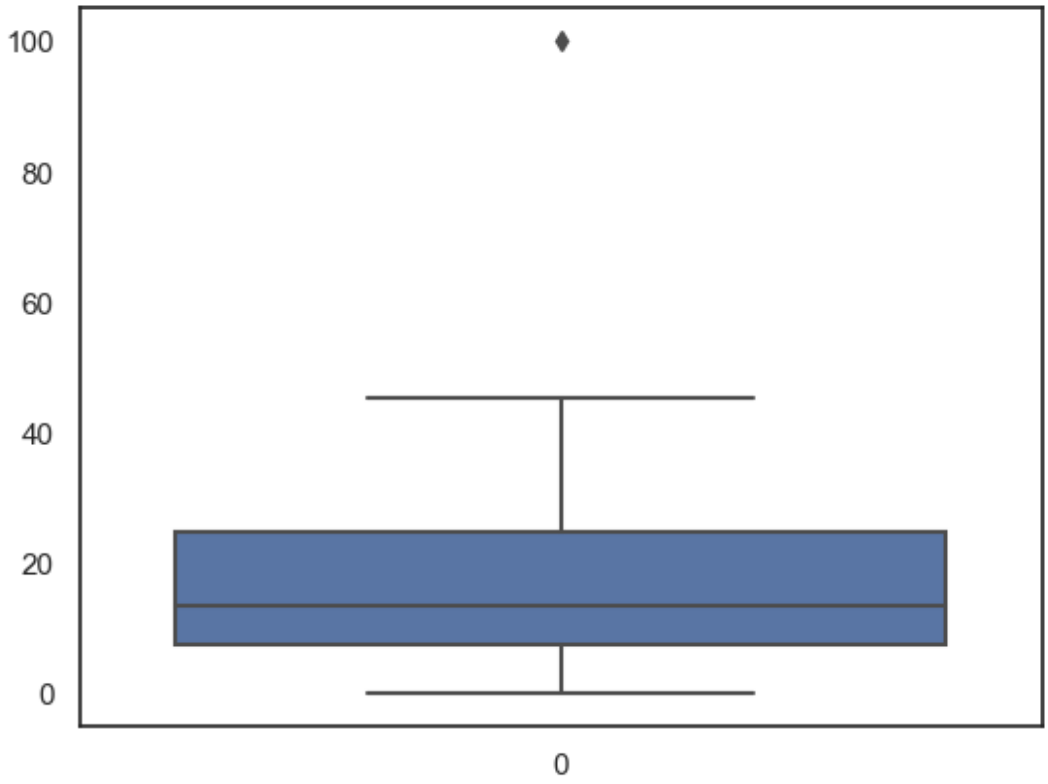
We knows that

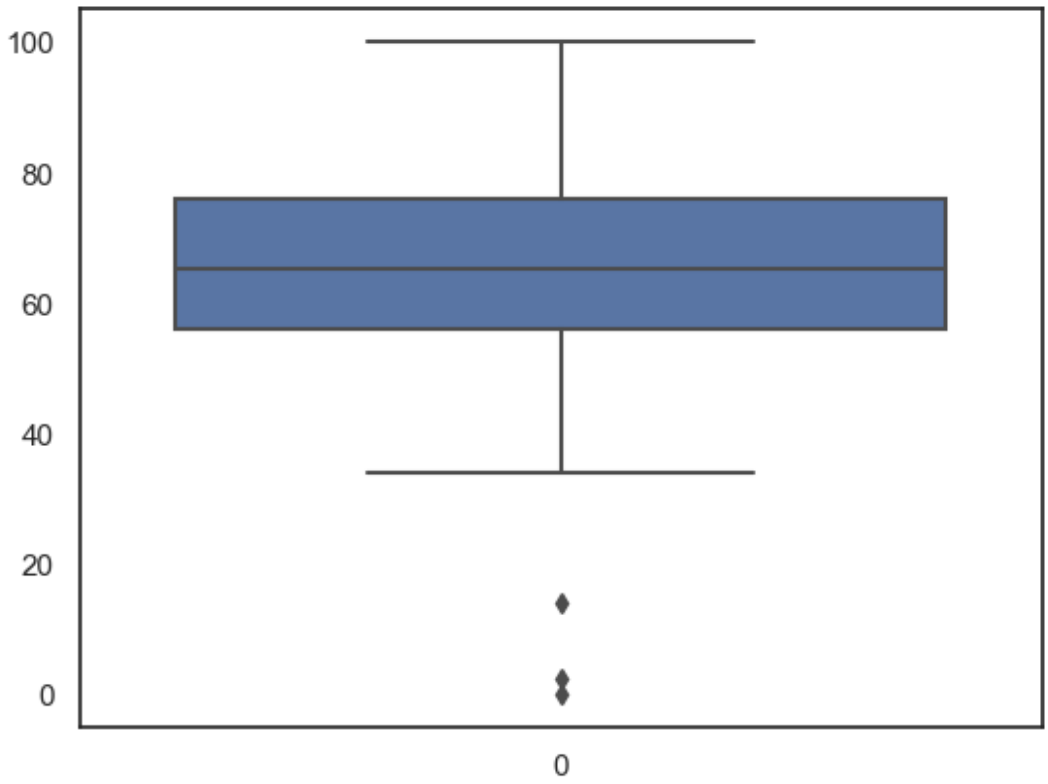
Total score is highly correlated with Research (0.95), Development (0.87), Talent (0.86), and Commercial (0.86) variables Total score displays mediuum correlations with Government Strategy and Operating Environment variables (that is, the contribution of such variables to the Total score is less then the highly correlated variables)

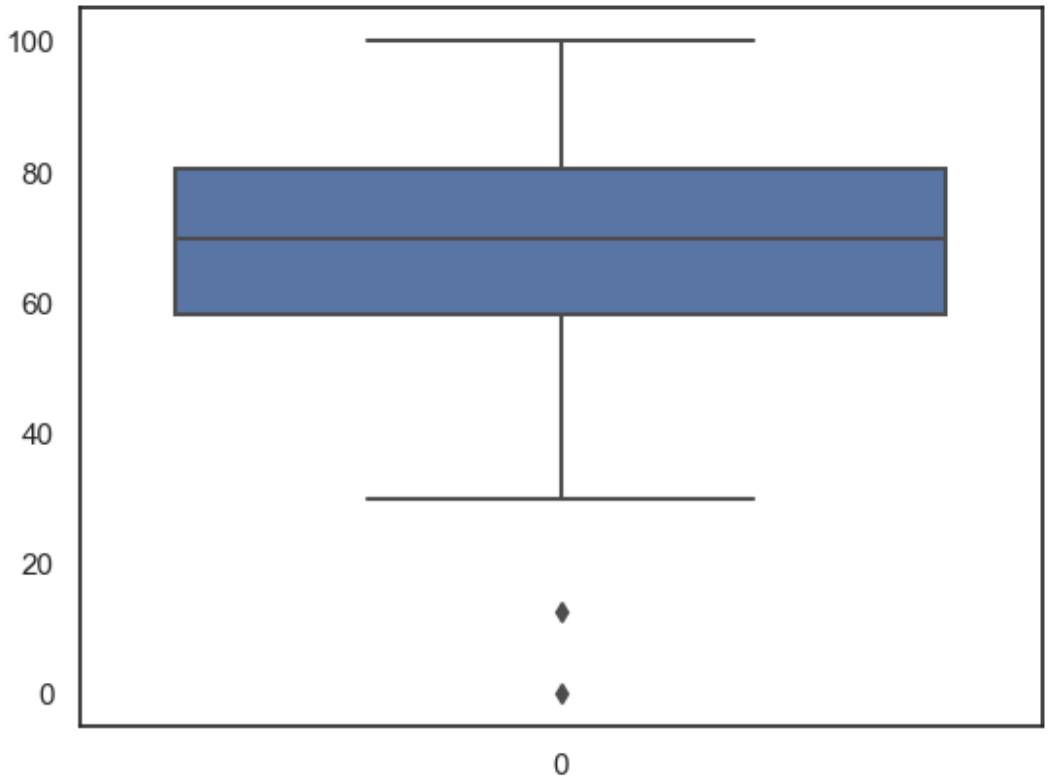
```
data1.columns
```

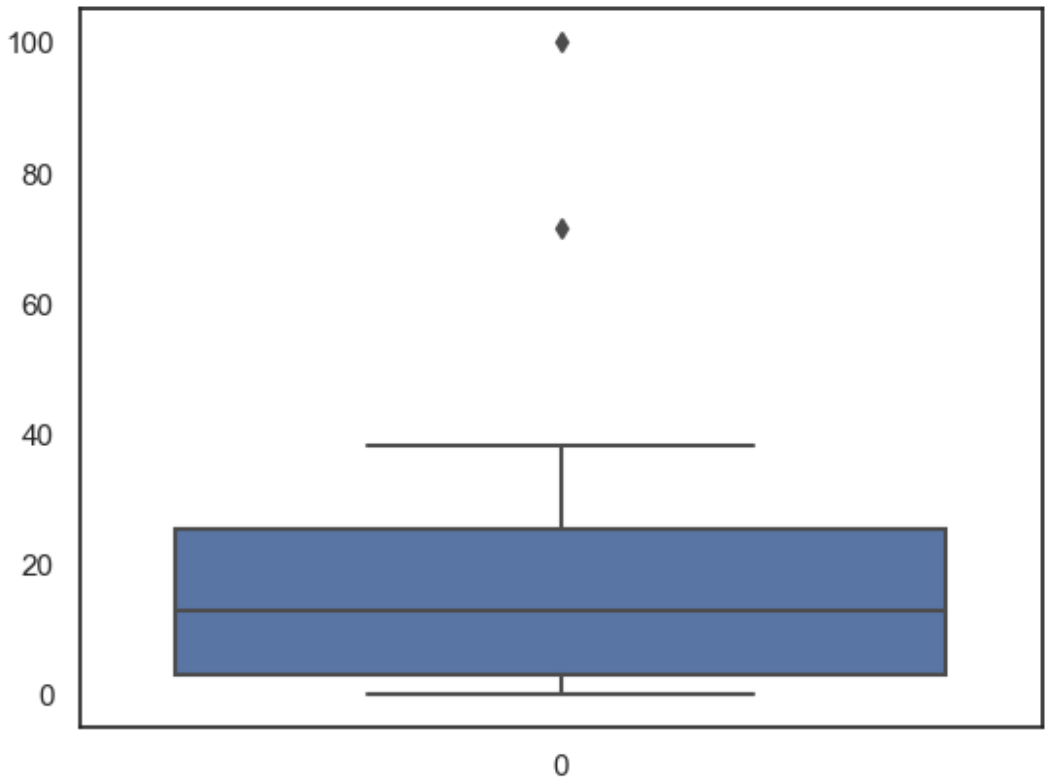
```
Index(['Country', 'Talent', 'Infrastructure', 'Operating Environment',
       'Research', 'Development', 'Government Strategy', 'Commercial',
       'Total score', 'Region', 'Cluster', 'Income group', 'Political
regime'],
      dtype='object')
```

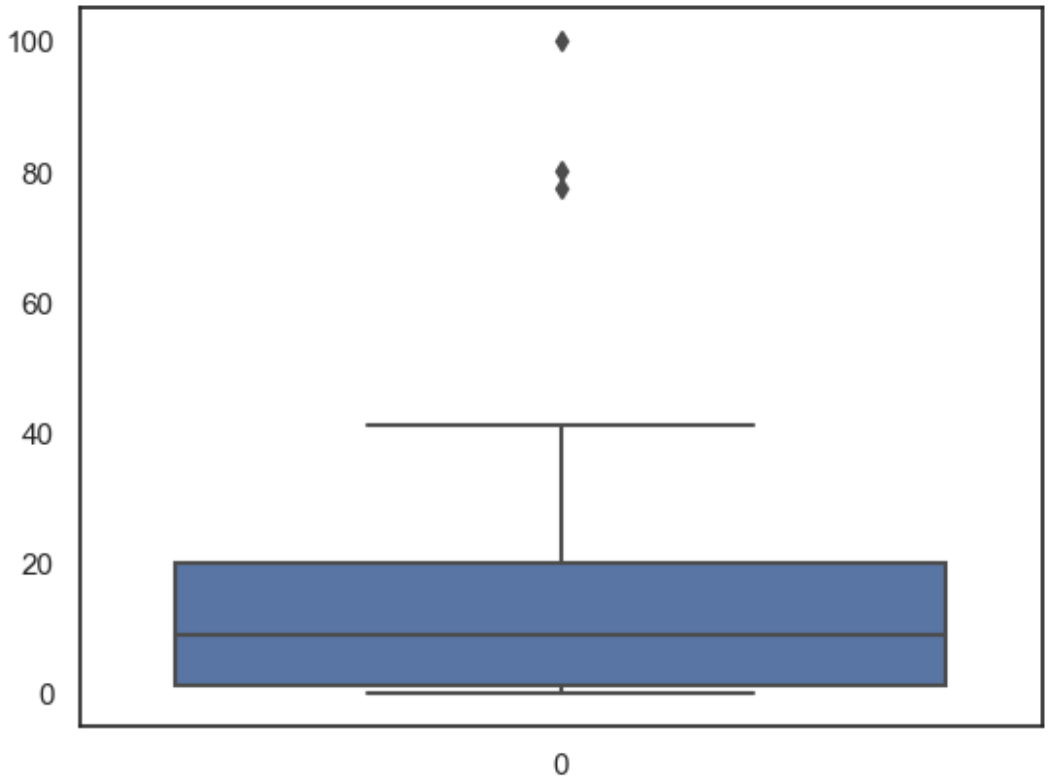
```
variables = ['Talent', 'Infrastructure', 'Operating Environment',
       'Research', 'Development', 'Government Strategy', 'Commercial',
       'Total score',]

i = 1
for var in variables:
    plt.figure(i)
    sns.boxplot(data1[var])
    i += 1
```

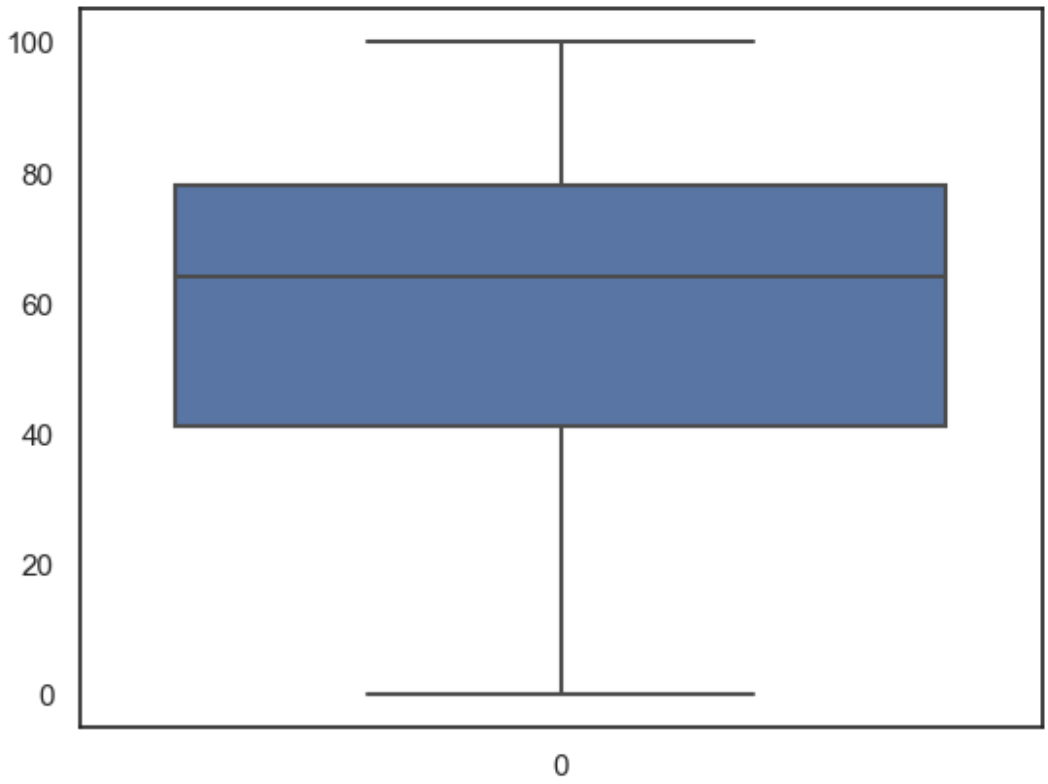


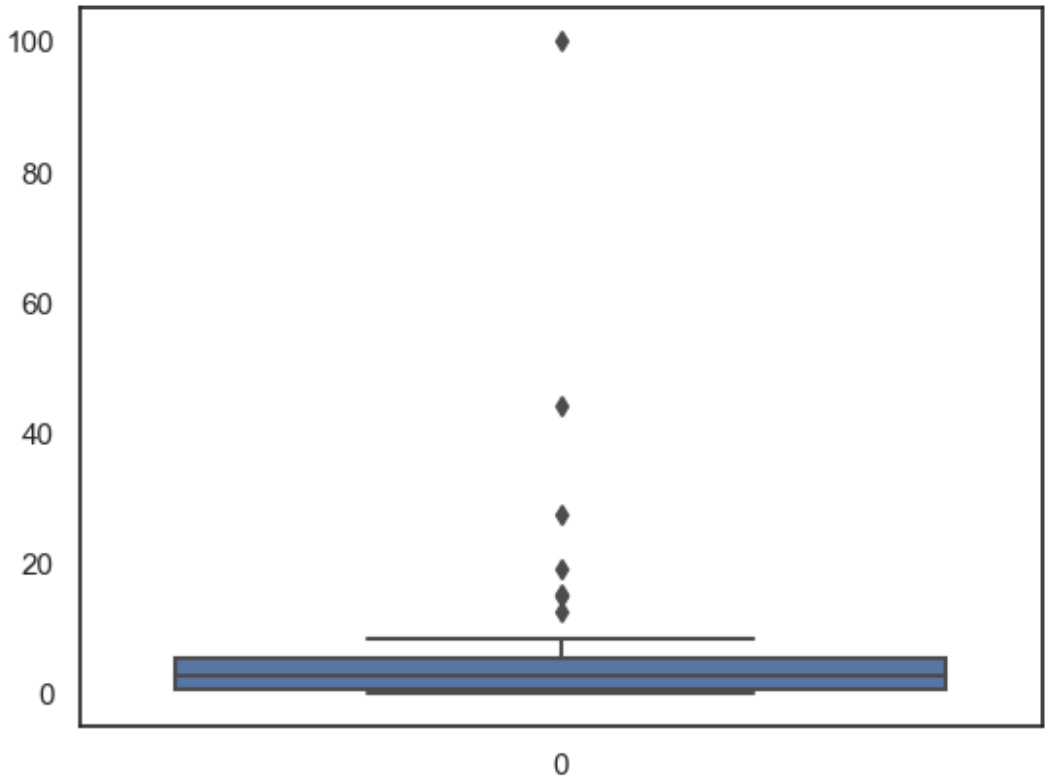

100
80
60
40
20
0
0

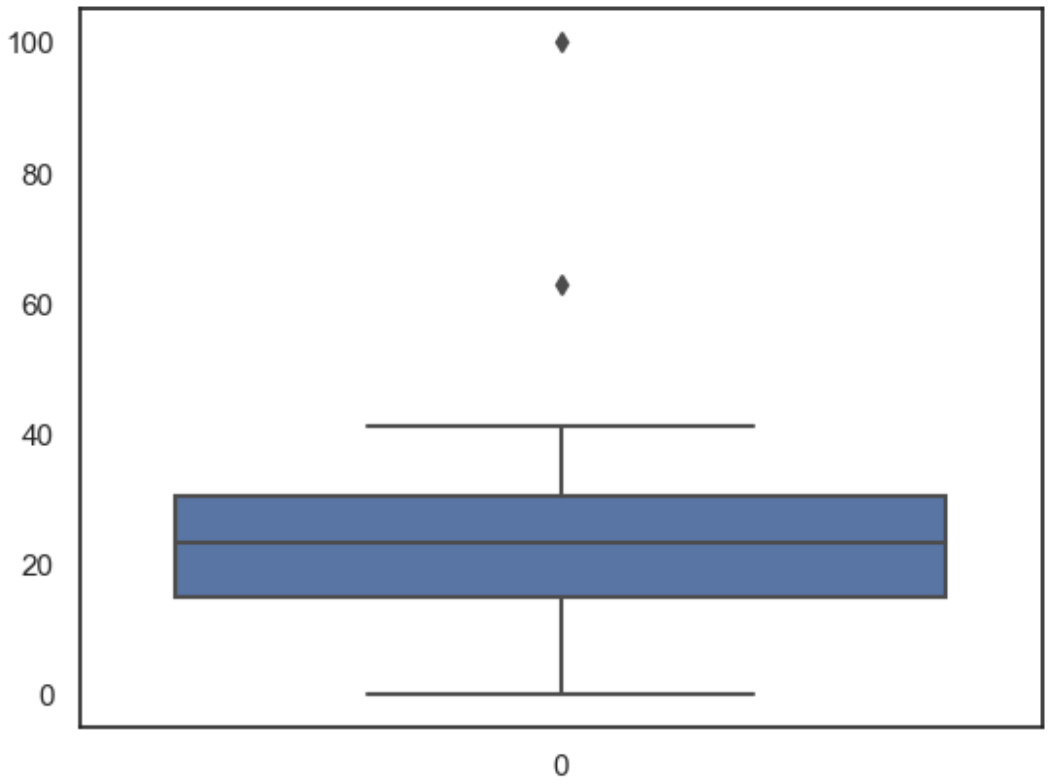

100
80
60
40
20
0
0


100
80
60
40
20
0
0


100
80
60
40
20
0
0


100
80
60
40
20
0
0


100
80
60
40
20
0
0



We find that

the biggest group of the countries in the dataset are European countries with High income level, fitting into Waking Up cluster countries with High income level, fitting into Waking Up cluster, are the most numerous group in Aasia-Pacific and Middle East regions, too the most 'crowdy' group in Americas is Waking Up, with Upper middle income level Traditional champions and Raising star countries are all with High level of incomes One of the Power players (USA) is with the High level of incomes, and another Power player (China) in Uppper middle income group the biggest groups in Europe and Asia-Pacific regions represent the countries with High income and either Liberal or Electoral democracy the biggest group of countries in Americas contains the countries with High income level and Liberal democracy as a political regime the biggest group of countries in Middle East contains the countries with High income level and Closed autocraty as a political regime

Conclusion : In the chapters below, we are going look deeper at the business-level and public administration insights drawn from the data in AI dataset.

First of all, we are going to look at the relations between Total score and other variables in the dataset (see Total Score Insights chapter).Then we are going to outline the holistic insights on

AI Implementation AI Research and Development AI Governance and Commercial contexts In the series of the final chapters at the end of the notebook, we are going to perform bivariative and multivariate analysis for the rest of the numeric features in the dataset (besides Total score).