BIRATNAGAR INTERNATIONAL COLLEGE

University of Wolverhampton

Concepts and Technologies of AI

5CS037

Assignment -- Statistical Interpretation and Exploratory Data Analysis

Name :Sandesh Ghimire
University Id :2505819

1. Introduction

The World Happiness Report is an annual report that estimates the happiness levels of different nations. The people taking part in the survey express their satisfaction with life on a scale that goes from 0 (the least imaginable case) to 10 (the most imaginable case). The 2025 version contains information on 143 countries, such as the happiness ladder score, and the elements determining happiness, which are money, social help, health, personal liberty, donating, and the perception of state honesty.

This report looks at the complete dataset and focuses on comparing South Asia and the Middle East to identify factors that affect happiness differently.

Explanation

The dataset includes the first ten records of each nation along with their respective Happiness score and the determinants which include GDP, Social support, Healthy life Expectancy, Freedom to make life choice, Generosity, Corruption views, and a baseline plus unexplained variance, all these factors being represented in the table below.

	Country name	score	Log GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption	Dystopia + residual
0	Finland	7.741	1.844	1.572	0.695	0.859	0.142	0.546	2.082
1	Denmark	7.583	1.908	1.520	0.699	0.823	0.204	0.548	1.881
2	Iceland	7.525	1.881	1.617	0.718	0.819	0.258	0.182	2.050
3	Sweden	7.344	1.878	1.501	0.724	0.838	0.221	0.524	1.658
4	Israel	7.341	1.803	1.513	0.740	0.641	0.153	0.193	2.298
5	Netherlands	7.319	1.901	1.462	0.706	0.725	0.247	0.372	1.906
6	Norway	7.302	1.952	1.517	0.704	0.835	0.224	0.484	1.586
7	Luxembourg	7.122	2.141	1.355	0.708	0.801	0.146	0.432	1.540
8	Switzerland	7.060	1.970	1.425	0.747	0.759	0.173	0.498	1.488
9	Australia	7.057	1.854	1.461	0.692	0.756	0.225	0.323	1.745

Getting Started with Data Exploration

```
...                                    0
            Country name          object
                   score         float64
        Log GDP per capita       float64
            Social support       float64
    Healthy life expectancy      float64
 Freedom to make life choices    float64
                Generosity       float64
   Perceptions of corruption     float64
         Dystopia + residual     float64
dtype: object
```

It shows the outcome of looking at the DataFrame's details or types. It has eight numerical columns and 'Country name' as a string. It also includes counts of valid entries.

• Basic Statistics:
Mean, Median and Standard deviations:

The mean happiness score is 5.53, median is 5.79, and standard deviation is 1.17 total of 143 countries.

```
Mean: 5.52758041958042

Medain: 5.785

std: 1.1707165099442995
```

Highest and Lowest Score:

7.741 is the highest score of Finland, while 1.721 is the lowest score of Afghanistan.

```
•••    HIGHEST: Finland (7.741)
       LOWEST: Afghanistan (1.721)
```

Missing Value:
Only 3 rows have missing explanatory variables those are (Bahrain, the State of Palestine, and Tajikistan)and Happiness score is completed for all country and datatypes all are numerical expect country name it is string

```
••  The missing value:
                                      0
              Country name            0
                     score            0
          Log GDP per capita          3
             Social support           3
      Healthy life expectancy         3
  Freedom to make life choices        3
                 Generosity           3
    Perceptions of corruption         3
          Dystopia + residual         3

    dtype: int64
```

**• Add New Columns and that categorizes countries into three categories
On the based of score:**

The new column contains the categorical of Happiness as a new column and three categories based on their Score like   High, Medium, and Low added to the dataset. High Countries that score 6.5 or above like Finland and Denmark ,Medium Countries that score between 4.5 and 6.5 are the majority of the countries and Low score Countries that score less than 4.5 are Afghanistan and Lebanon The categorisation is simple and allows anyone to visually see and easily identify very happy, average, and struggling countries without having to check the specific numbers each time.

```
       Country name  score Happiness_Category
0            Finland  7.741               High
1            Denmark  7.583               High
2            Iceland  7.525               High
3             Sweden  7.344               High
4             Israel  7.341               High
..               ...    ...                ...
138  Congo (Kinshasa)  3.295                Low
139      Sierra Leone  3.245                Low
140           Lesotho  3.186                Low
141           Lebanon  2.707                Low
142       Afghanistan  1.721                Low

[143 rows x 3 columns]
```

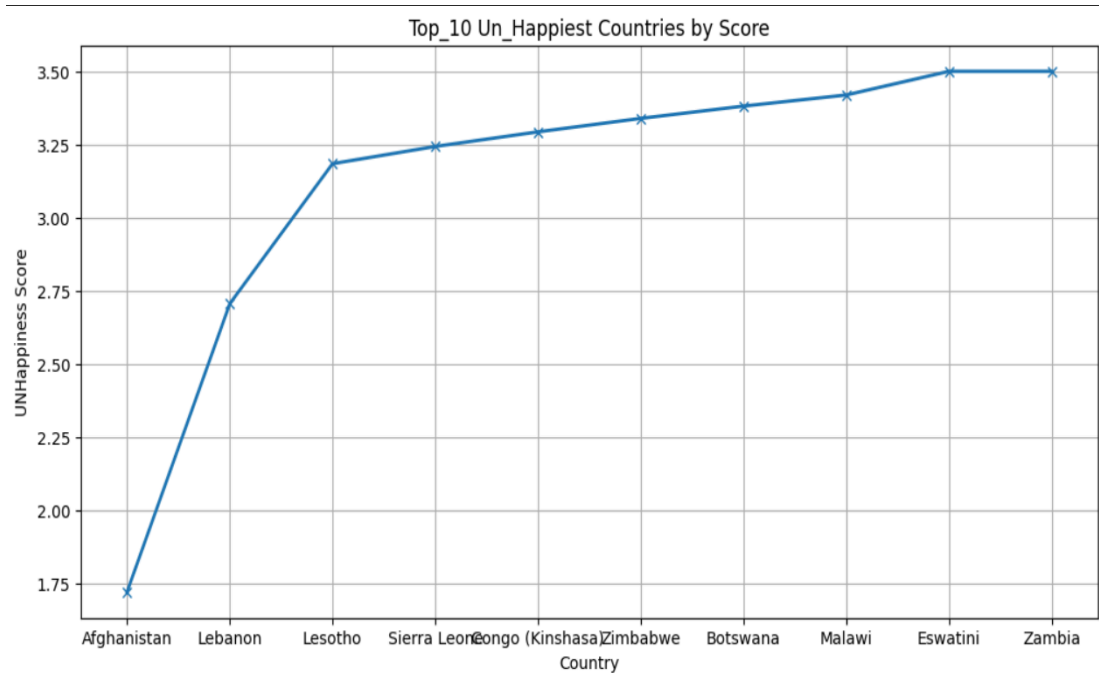|	Country name	score	Log GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption	Dystopia + residual	Happiness_Category
0	Finland	7.741	1.844	1.572	0.695	0.859	0.142	0.546	2.082	High
1	Denmark	7.583	1.908	1.520	0.699	0.823	0.204	0.548	1.881	High
2	Iceland	7.525	1.881	1.617	0.718	0.819	0.258	0.182	2.050	High
3	Sweden	7.344	1.878	1.501	0.724	0.838	0.221	0.524	1.658	High
4	Israel	7.341	1.803	1.513	0.740	0.641	0.153	0.193	2.298	High

• Bar Plot:
This bar chart illustrates the ten happiest countries ranked by scores, with Finland in first place at 7.7, and Denmark, Iceland, and the other countries being in the range of 7.0 to 7.7 out of 8.
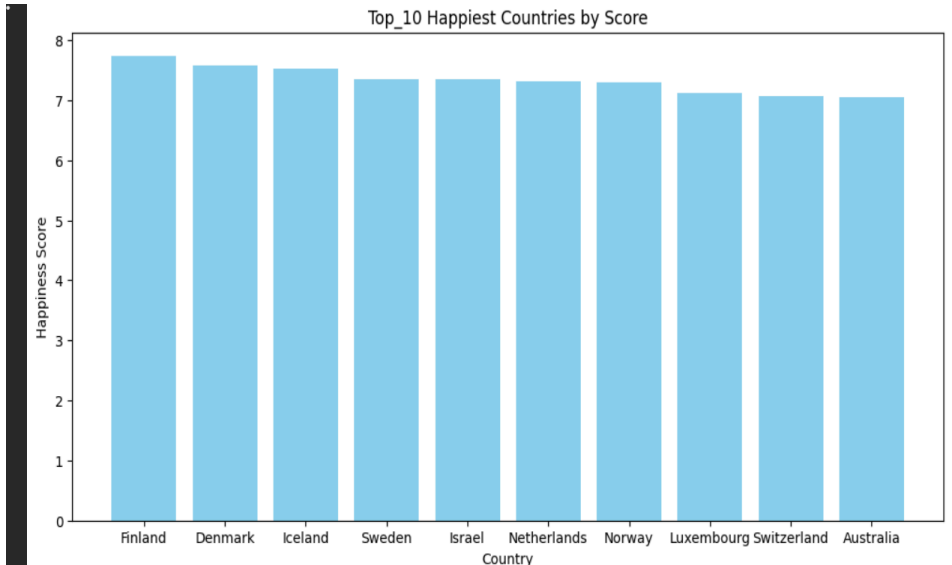

• Line Plot:

The line plot indicates that Afghanistan, with the lowest score, is the most unhappy country while Lebanon, going through the same crisis, is no happier than the other nations that score below 3.5, which is usually regarded as a low score compared to the average happiest country.


Top_10 Un_Happiest Countries by Score

Plot a histogram:

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

This **histogram shows the graph of distribution of average scores on national happiness on the scale 0-10.The datasheet, having rated about 140-150 nations with the scale Ladder scale,disclosed a skewed pattern of the right with most (more than 80%) in the 4-7 range.maximum frequency in the range of 5-6, and minimum scores below 4 and extremely high scores more than 7 are very few,indicating inequality in life satisfaction across the world due to such factors as economic stability, social**

Distribution of Happiness Scores
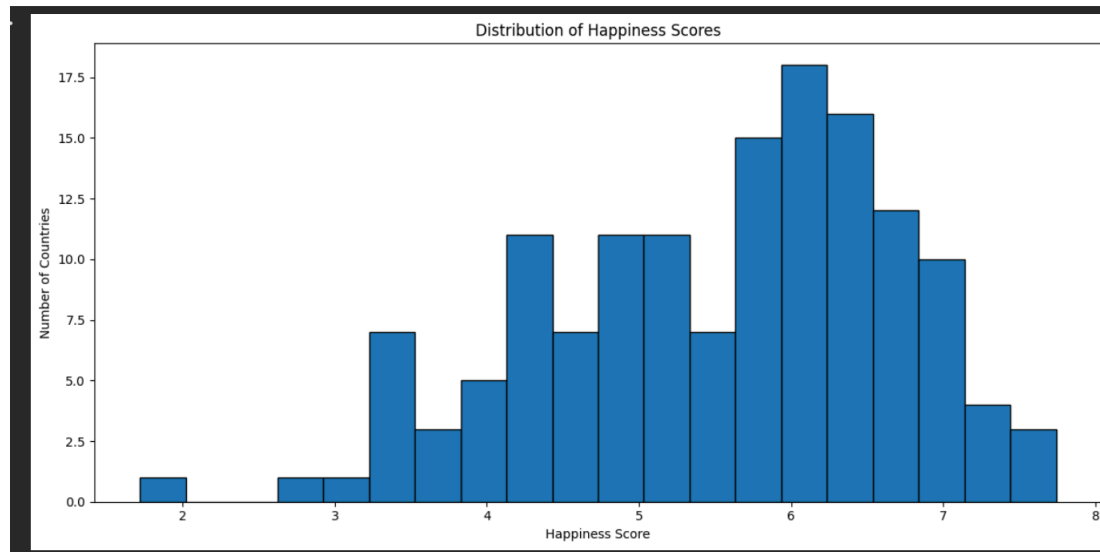
• **Scatter Plot:**

The graph displays the relationship between the economic prosperity of a nation (illustrated by GDP per capita, shown in a log scale at the bottom) and its people's contentment (represented by happiness score on the left, 2-8).

Happiness Score vs GDP per Capita
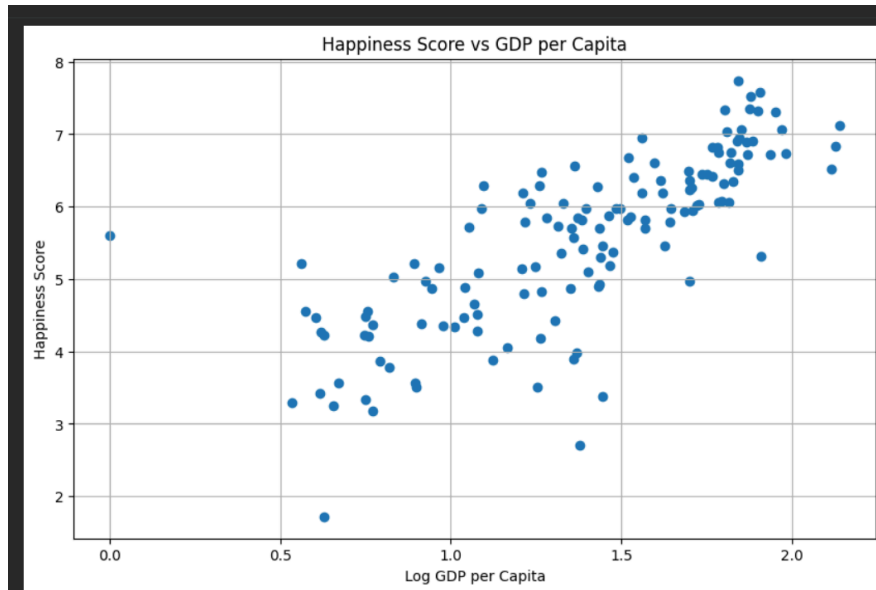
Task - 2 - Composite Score Ranking:

In the table displayed, South Asian nations are arranged according to their composite score on a scale of 0-1, with Nepal attaining the first position at 0.816 raw score, again giving Nepal first place at 5.16 which maybe reflects human development or some other criteria.

```
...      Country name  Composite_score
92             Nepal           0.8159
107         Pakistan           0.7039
125            India           0.7874
127        Sri Lanka           1.0739
128       Bangladesh           0.6774
142      Afghanistan           0.3238
         Healthy life expectancy  Composite_score
92                         0.443           0.8159
107                        0.321           0.7039
125                        0.417           0.7874
127                        0.586           1.0739
128                        0.513           0.6774
142                        0.242           0.3238
         Social support  Composite_score
92                0.990           0.8159
107               0.600           0.7039
125               0.653           0.7874
127               1.179           1.0739
128               0.249           0.6774
142               0.000           0.3238
```

```
         Country name  Composite_score
127         Sri Lanka           1.0739
92              Nepal           0.8159
125             India           0.7874
107          Pakistan           0.7039
128        Bangladesh           0.6774
142       Afghanistan           0.3238
```
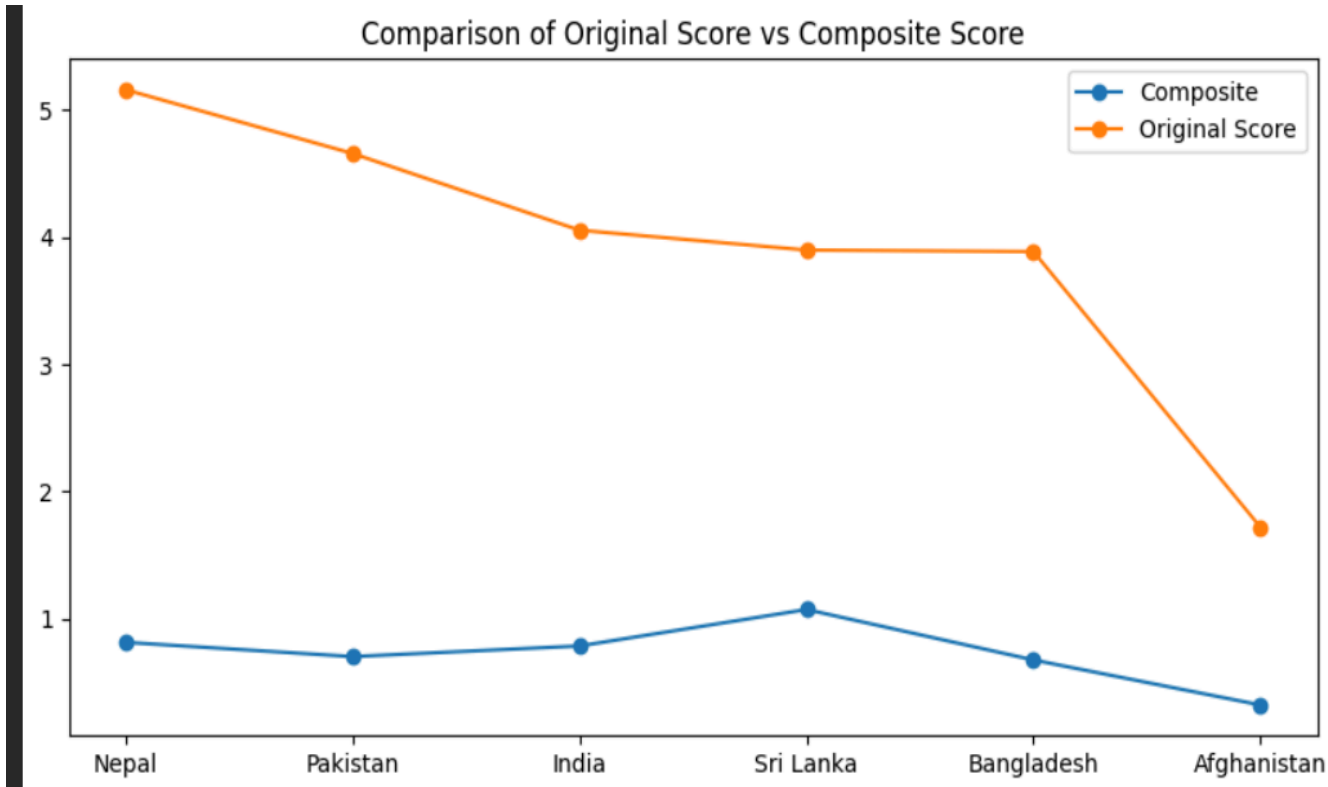
Horizontal bar chart showing the Composite Score:

The depicted bar graph corroborated that Nepal came first with a composite score of about 0.82, while India was right behind with approximately 0.79, followed by Pakistan with 0.70, Bangladesh with 0.68, and Afghanistan with a distant score of 0.32 among the five South Asian neighbors.

Top 5 South Asian Countries by Composite Score
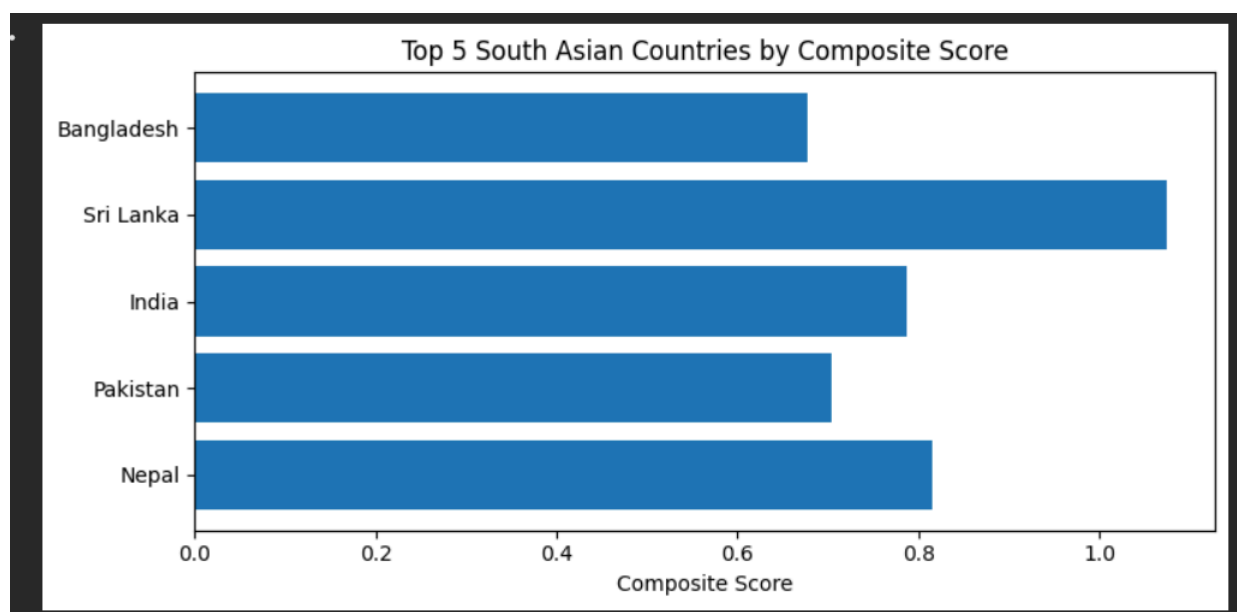
2. Composite Score on rankings based on the
 with some visualization plot.

The scatter plot shows a strong linear correlation between the Original Happiness Score (0-10)
and the two variables, with Nepal at the top and Afghanistan at the bottom.

Comparison of Original Score vs Composite Score

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

Task - 3 - Outlier Detection:

Afghanistan is clearly a statistical outlier as it exhibits both Happiness Score 1.721 which is much less than the lower limit of about 2.73 and log GDP per capita 0.628, which is considerably less than about 0.729.

```
•••    UnHappiness Score by outliner country :
          Country name   score
      142  Afghanistan   1.721

       Gdp by outliner country :
          Country name   Log GDP per capita
      142  Afghanistan                 0.628
```

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

3.3 Create a scatter plot with GDP per Capita on the x-axis and Score on the y-axis, highlighting outliers in a different color.
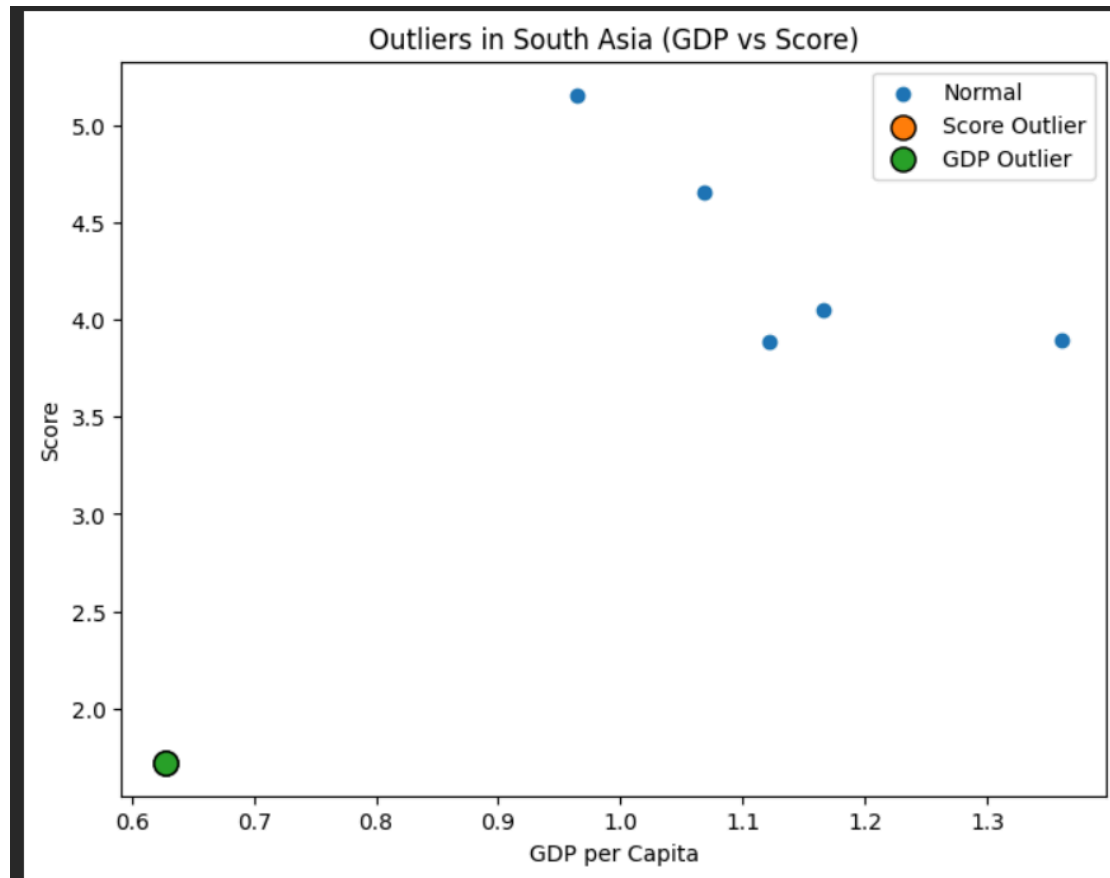
The scatter plot displayed below shows the connection between GDP per capita and the happiness scores of South Asian countries.The greater part of the countries are indicated in blue, meaning they fall within the normal range of GDP and happiness scores. The GDP outlier with very low income and low score depicted by the green dot represents economic disparity.

Task - 4 - Exploring Trends Across Metrics:

```
··  Pearson correlation: (score vs Freedom to make life choices): 0.801

    Pearson correlation: (score vs Generosity): 0.875
```

The outcome indicates that in South Asia happiness levels exhibit a positive correlation not only with Generosity but also with Freedom to choose one's own way of living.

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

4.2 Create scatter plots
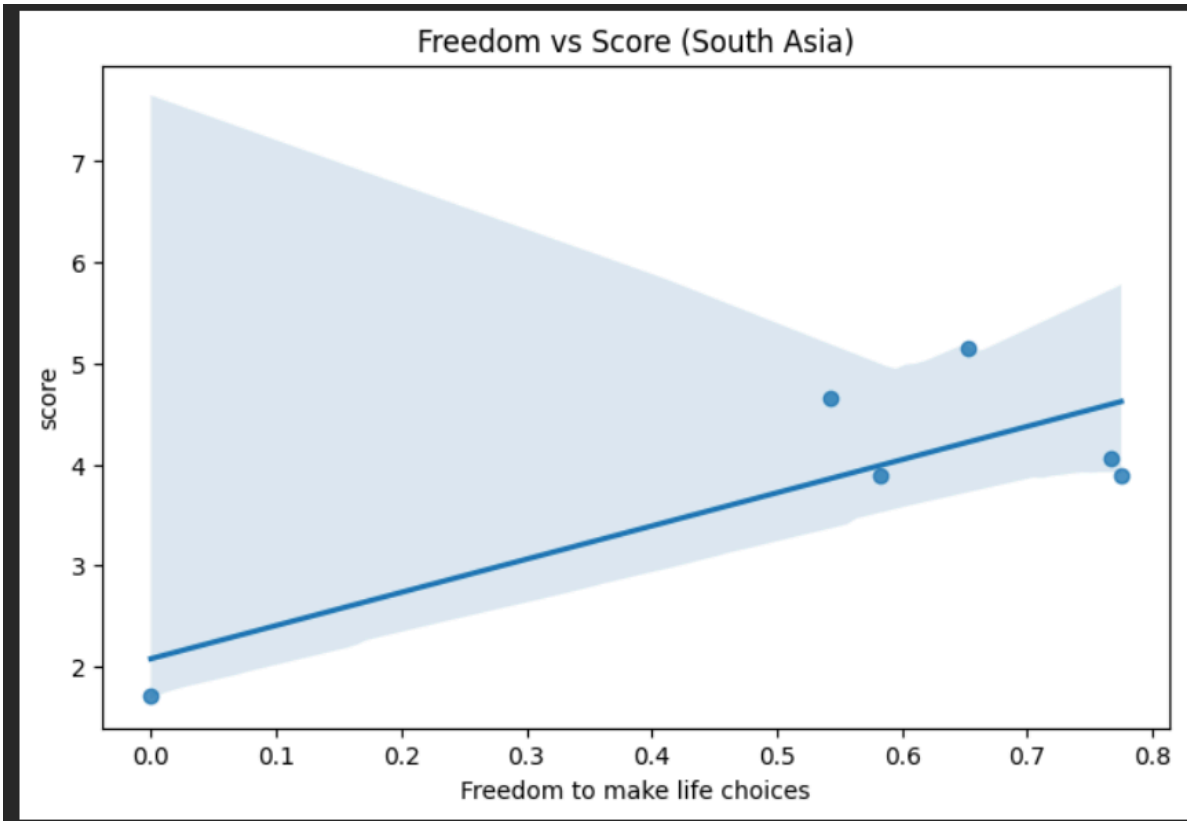
Freedom vs Score (South Asia)

Generosity vs Score (South Asia)
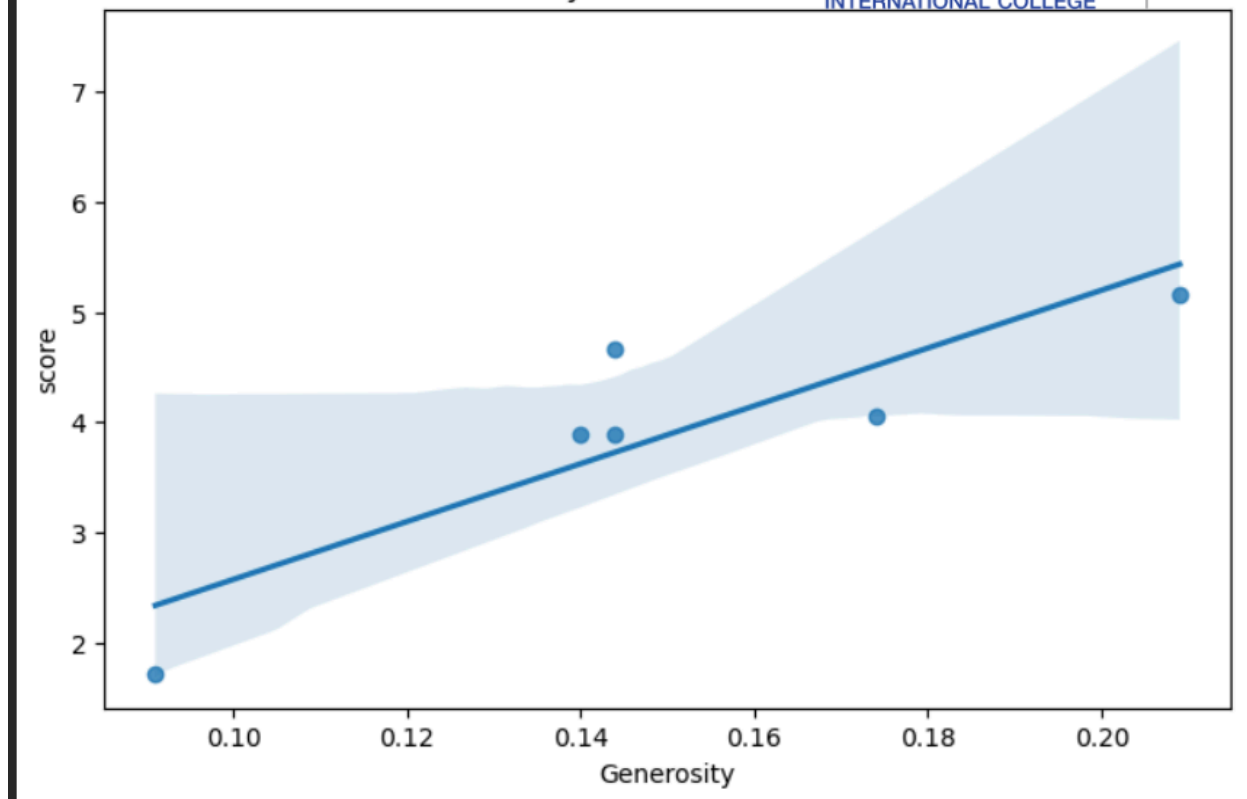
These results indicate that happiness levels in South Asia are positively related not only to Generosity but also to the Freedom to make decisions about one's lifestyle.

4.4

```
Correlation (Freedom vs Score): 0.8005185224163315
Correlation (Generosity vs Score): 0.874512371253192
```

The most powerful association within South Asian countries belongs to the Generosity aspect, while there is a very slight connection of Freedom to make life choices among these factors that represent the weakest link.South Asian nations the strongest relationship is between Generosity where there is slightly a weakest link with Freedom to make life choices .

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

Task - 5 - Gap Analysis:

5.1

```
        Country name   GDP_Score_Gap
92           Nepal          -4.193
107       Pakistan          -3.588
125          India          -2.888
127      Sri Lanka          -2.537
128     Bangladesh          -2.764
142    Afghanistan          -1.093
```
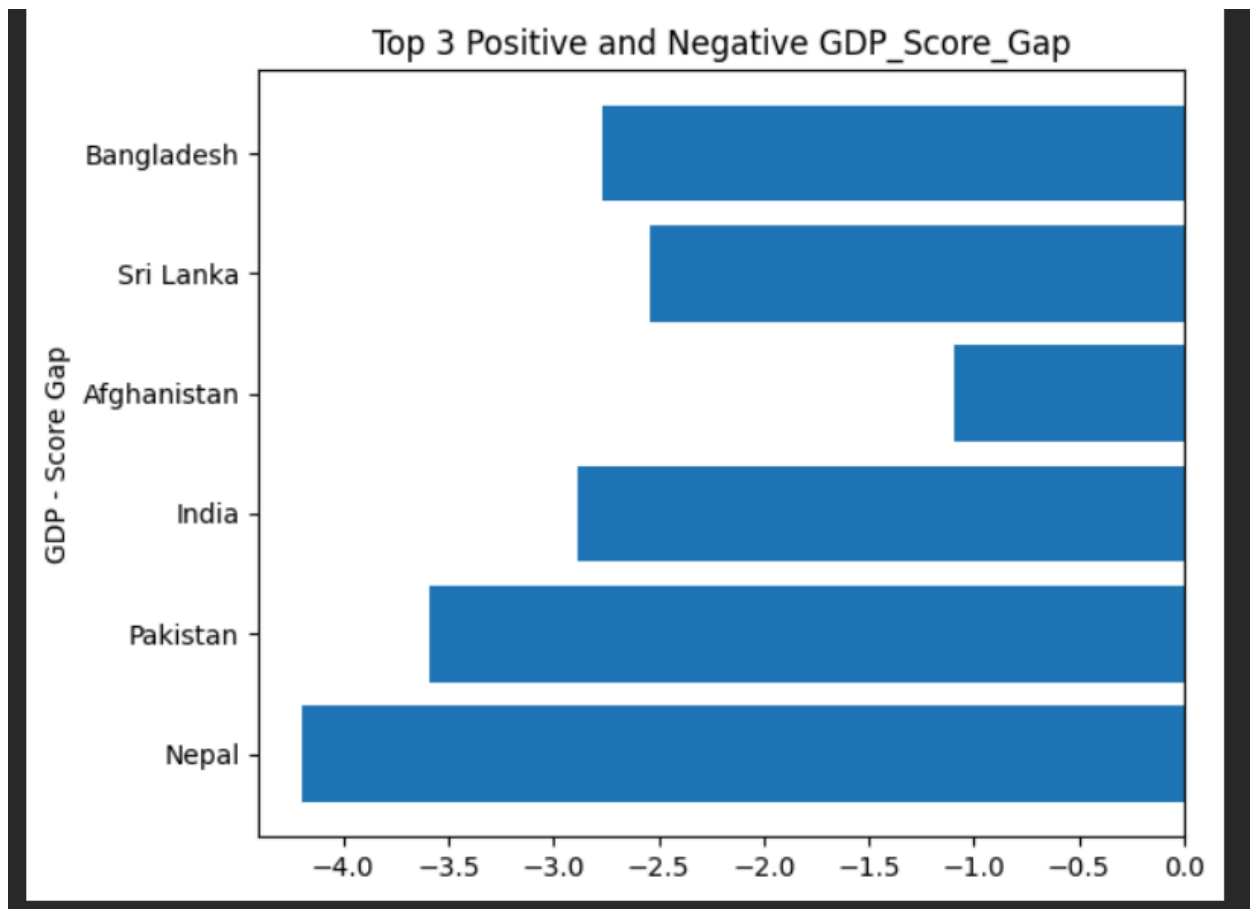
```
Ascending order:

        Country name   GDP_Score_Gap
142    Afghanistan          -1.093
127      Sri Lanka          -2.537
128     Bangladesh          -2.764

Descending order:

        Country name   GDP_Score_Gap
92           Nepal          -4.193
107       Pakistan          -3.588
125          India          -2.888
```

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

5.2 bar graph

Top 3 Positive and Negative GDP_Score_Gap

Both the table and the graph reveal that the happiness scores of South Asian countries are less than their GDPs, with Nepal and Pakistan at the negative GDP gaps. It indicates that the majority of the population's unhappiness is in relative terms with the wealth deficits, while Afghanistan has the case of the smallest deficit owing to its income and the extremely low happiness prevailing there.

Comparative Analysis

Descriptive Statistics:

```
••   Mean score of southasia counter : 3.895666666666667

     Standard Devations of south asia country:  1.1770690152521504

     Mean score of middle east country : 5.412100000000001

     Standard Devations of middle east country:  1.5662011684327144
```

```
Middle East has higher happiness score on average.
```

Though the variation in happiness scores was quite similar in both regions, the Arab nations still managed to maintain a higher average score of 5.412 as compared to the South Asian countries which only had 3.895.

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

Top and Bottom Performers:

```
Top Three :      Country name   score
92           Nepal  5.158
107       Pakistan  4.657
125          India  4.054

 bottom 3:       Country name   score
142  Afghanistan  1.721
128    Bangladesh  3.886
127     Sri Lanka  3.898

Top Three :               Country name   score
4                  Israel  7.341
12                 Kuwait  6.951
21   United Arab Emirates  6.733

bottom_3 :
    Country name  score
141       Lebanon  2.707
132        Yemen  3.561
124        Jordan  4.186
```

In South Asia the nations Nepal, India and Pakistan have the greatest number of scores while Afghanistan, Bangladesh, Sri Lanka, Lebanon, Yemen and Jordan have the lowest number of scores but Uae, Israel and Kuwait get the highest rating.

Plot Graph :

Top 3 Countries: Score Comparison
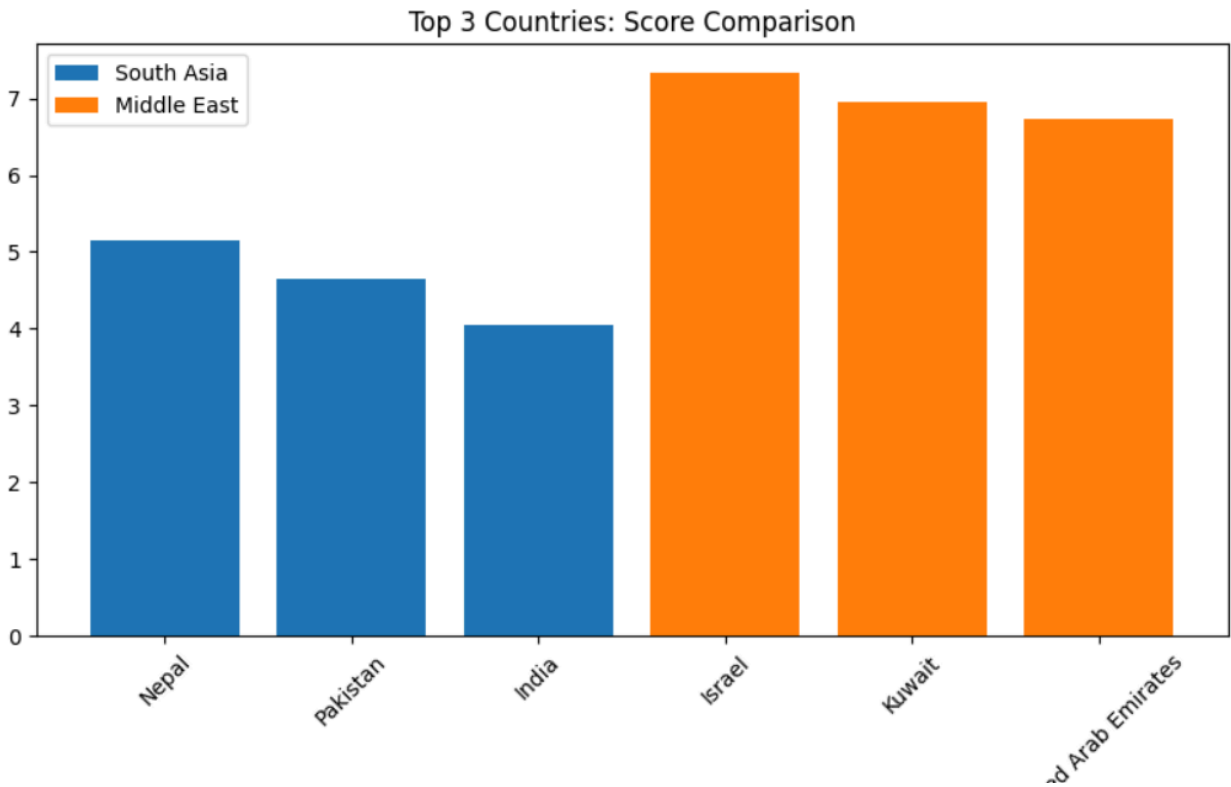
This bar graph show comparison score between two region but as shows graph middle east country have highest number of Score as compare to South Asian nation like Nepal,India and Pakistan

Metric Comparisons:

Metric Comparison (South asia vs Middle east)
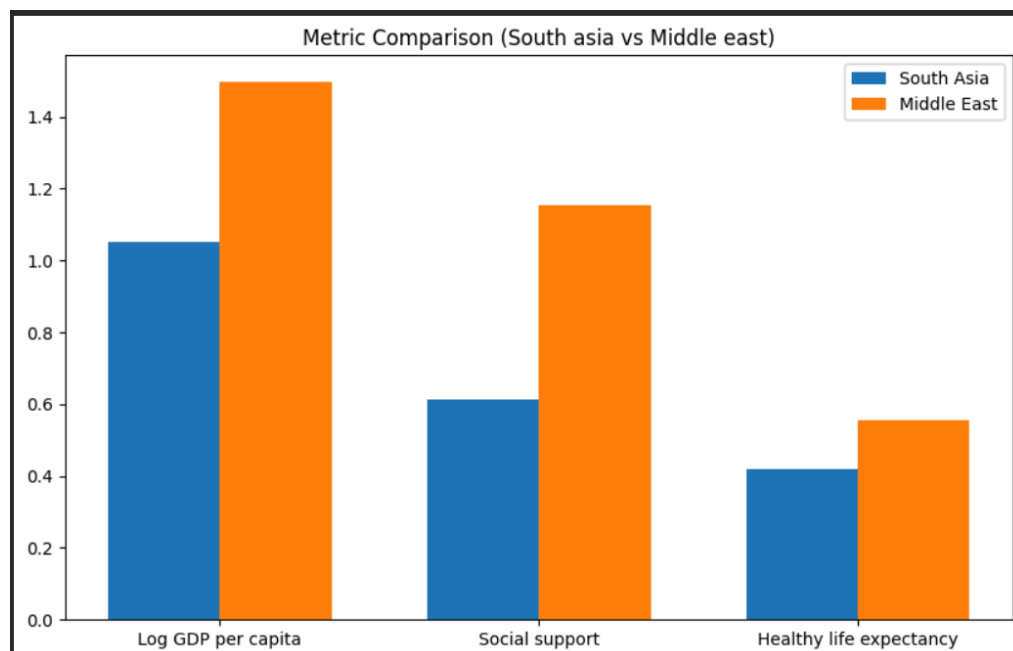
Usually, those countries in the Middle East are in the higher range of all these factors including GDP, Social Support and Healthy Life Expectancy, and according to the chart.

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

Happiness Disparity:

The happiness range in south asia is 3.437 points , while Middle East have wider range of 4.63 points when we look at CV south asia's countries have higher score and middle east nations as compared both region in CV south asia`s take lead but in range middle east far away from another region.

```
South Asia Range: 3.4370000000000003
Middle East Range: 4.634

South Asia CV: 0.3021482883337427
Middle East CV: 0.28938880812119405
```

Correlation Analysis:

The region of South Asia is characterized by two metrics, generosity and freedom to make life choices, which exhibited a very close correlation with happiness score. The civilizations that are more liberal and more open report greater satisfaction with life. Hence, it can be concluded that there is a weaker relationship with generosity, however, freedom still has a much wider connection in the Middle Eastern nation. The main difference factor of both regions is 'total freedom and pleasure are both personal matters, but giving is still more important in the east than in the west.'

```
ME Correlation (Freedom vs Score): 0.8632202433827543
ME Correlation (Generosity vs Score): 0.6275236536964182
```

Scatter Plot and compare them :

Middle East: Freedom vs Score

Middle East: Generosity vs Score

Freedom vs Score (South Asia)


Generosity vs Score (South Asia)
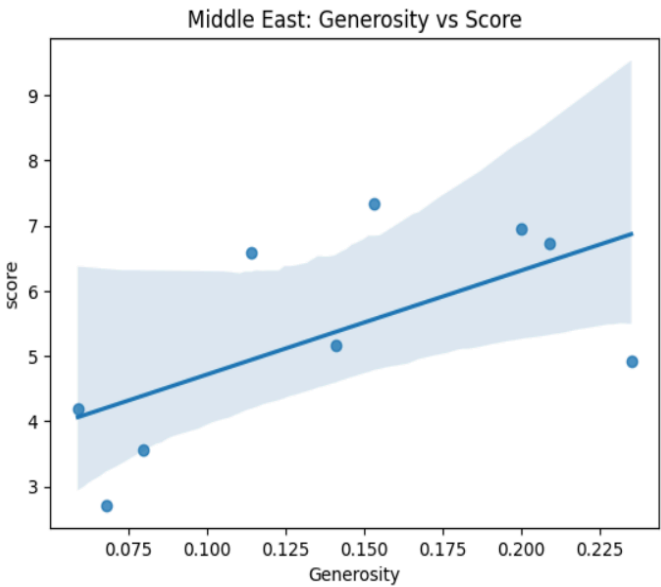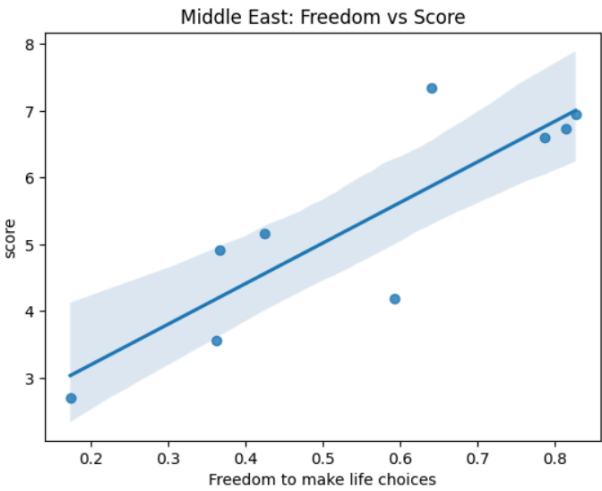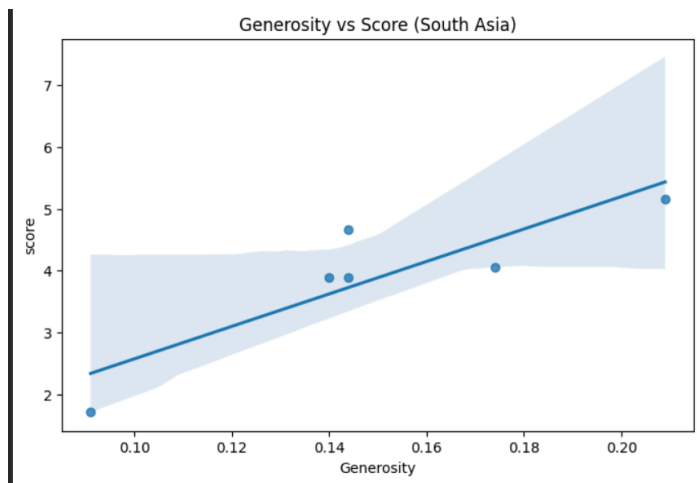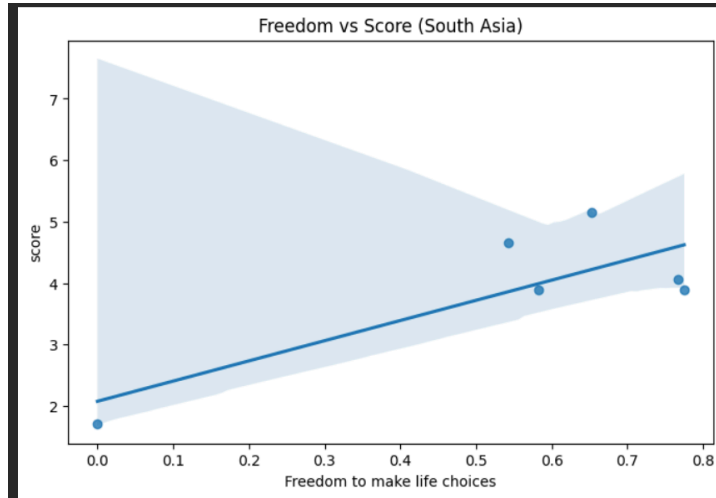
Scatter plots with trend lines confirm stronger links, and they are the clearest and most intelligible with higher upward slopes for South Asia (especially for generosity). Using the 1.5 × IQR method, Afghanistan stands out as the only clear outlier in South Asia – its score (1.721) is extremely low andGDP per capita (0.628) is very low as well.

Outlier Detection:
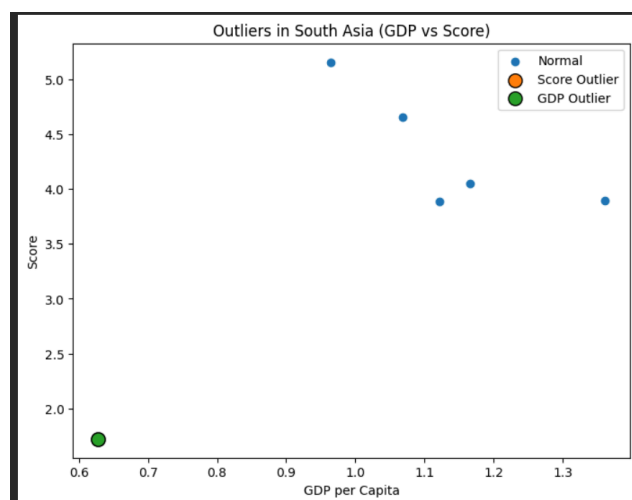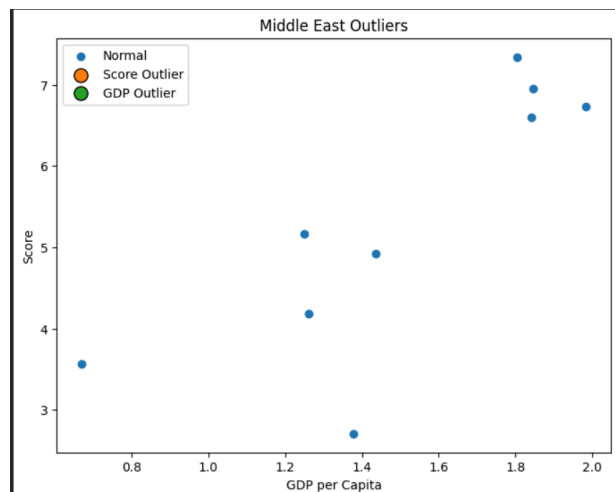
No Middle Eastern country meets the criteria of a statistical outlier, but Lebanon and Yemen are among the lowest scorers despite some areas having a good GDP.Negative impact of Afghanistan on south asia is more than that of any single nation in the Middle East Asia.

Visualization:
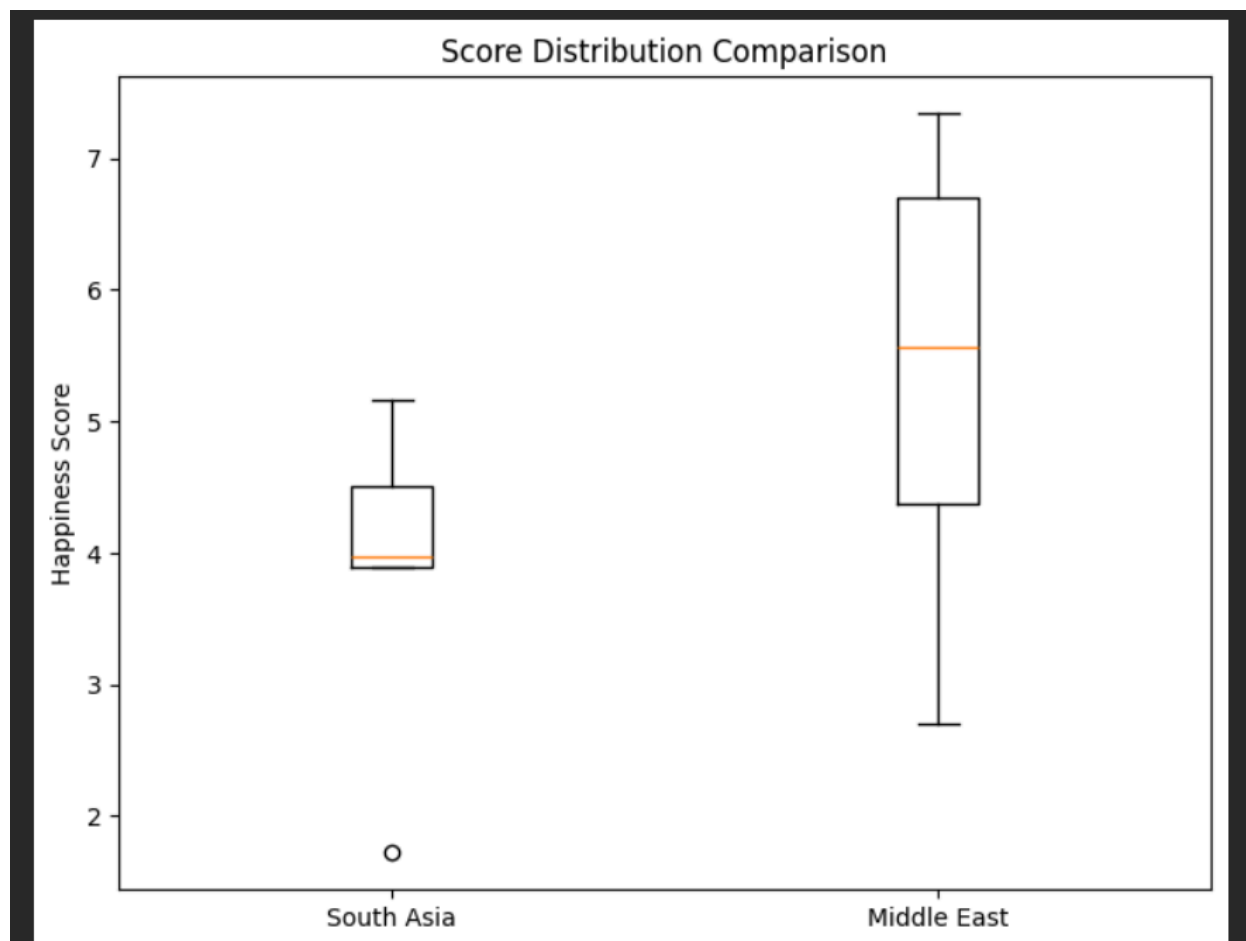
The Box plot shows that the Middle Eastern nations scored on an average significantly better than the South Asian ones while the scores from the Middle East were not that diverse in terms of their dispersion.

2. Global Picture of Happiness

With a score of 7.74, Finland was the happiest country in the world and the title holder for 2024, while Denmark, Iceland, Sweden, and Israel were the closest competitors. The top 10 places still primarily consist of the Nordic countries along with Israel, Netherlands, Luxembourg, Switzerland, Australia, and Norway. Such nations possess the features of vast wealth, highly efficient social security systems, long-life expectancy, and high government trust. The ten least happy countries are situated at the opposite extreme of the scale, their places being mostly in sub-Saharan Africa and war-torn areas: with their scores all below 3.5, the list includes countries like Afghanistan, Lebanon, Lesotho, etc. The worldwide ranking of all 143 nations shows that the majority of the ratings fall within the 4.5 to 6.5 range, thus creating a roughly normal (bell-shaped) distribution with a slight leftward shift. This implies that there are still some extremely unhappy people, but the vast majority of the world population thinks the same way, which means having a life satisfaction level that is considered average.

UNIVERSITY PARTNER

UNIVERSITY OF
WOLVERHAMPTON

BIRATNAGAR
INTERNATIONAL COLLEGE

ing

3. Relationship between Wealth and Happiness

Log GDP per capita and happiness score have a great positive correlation (r = 0.77) when we plot them which usually means that richer countries are happier. But still, a little money adds to happiness only when the GDP per capita is between log =1.8 -- 2.0 this is the classic Easterlin paradox. This is the reason why some Middle East Asia scored high but still not at the very top. Money can help a lot but only up to a point. After that, social factors prevail.

4. Comparison between two nation South Asia vs Middle East

The average happiness score for the Middle East stands at 5.35, a figure that is considerably above South Asia's score of 3.90, and this mainly due to the income and social support in the region being the highest. Social support is the factor with the greatest difference: the Middle East gets almost twice the score of the South Asia region's average. Afghanistan and Lebanon are placed lower than several less wealthy but peaceful African countries because of the conflict of happiness. In South Asia, people's generosity and freedom are the main sources of happiness, whereas in the Middle East, wealth and social connections are the primary factors.

5. Key Insights and Interpretation

The average happiness score in the Middle East (5.35) was much higher than that in South Asia (3.90), and the main reasons for it were the higher GDP and stronger social support. The difference in social support is the greatest, as the Middle East gives almost twice the average score of South Asia. Conflict affects happiness more than poverty does—this is evidenced by Afghanistan and Lebanon, which have lower happiness levels than several poorer but conflict-free African countries. The people of South Asia value sharing and liberty the most, while those of the Middle East count money and personal connections as their top priorities. Middle East nations such as (United Arab Emirates, Kuwait, and Saudi Arabia) are examples for the rest of the world that very fast economic growth accompanied by stability and welfare can very quickly increase national happiness.

6. Conclusion:

The World Happiness Report for 2025 suggests, to a certain extent, that the factors of happiness are not limited to the financial aspect; it is a matter of health, social connections, freedom, and trust as well. The Nordic countries keep on being the best places where wealth is associated with social trust and equality. The Middle East region, especially the Gulf countries, is much better off than South Asia and this is further aggravated by the situation in Afghanistan. One of the major contributions of this research to the revelation of global well-being patterns is the application of simple yet powerful Python tools like pandas and seaborn for data analysis.

References

All images and tables were created by me utilizing the official dataset (CSV) given by the instructor and executed in the manner of (Python/pandas/matplotlib/seaborn).

Insert link of github:https://github.com/Sundesh280/Math-course-work