Analysis of the World Happiness

Report: Exploring South Asia and Middle East Perspectives.

Student Id: 230164
Student Name: Sujal Shrestha
Section: L4CG9
Module Leader: Siman Giri
Tutor: Bibek Khanal
Submitted on: 2024/12/19

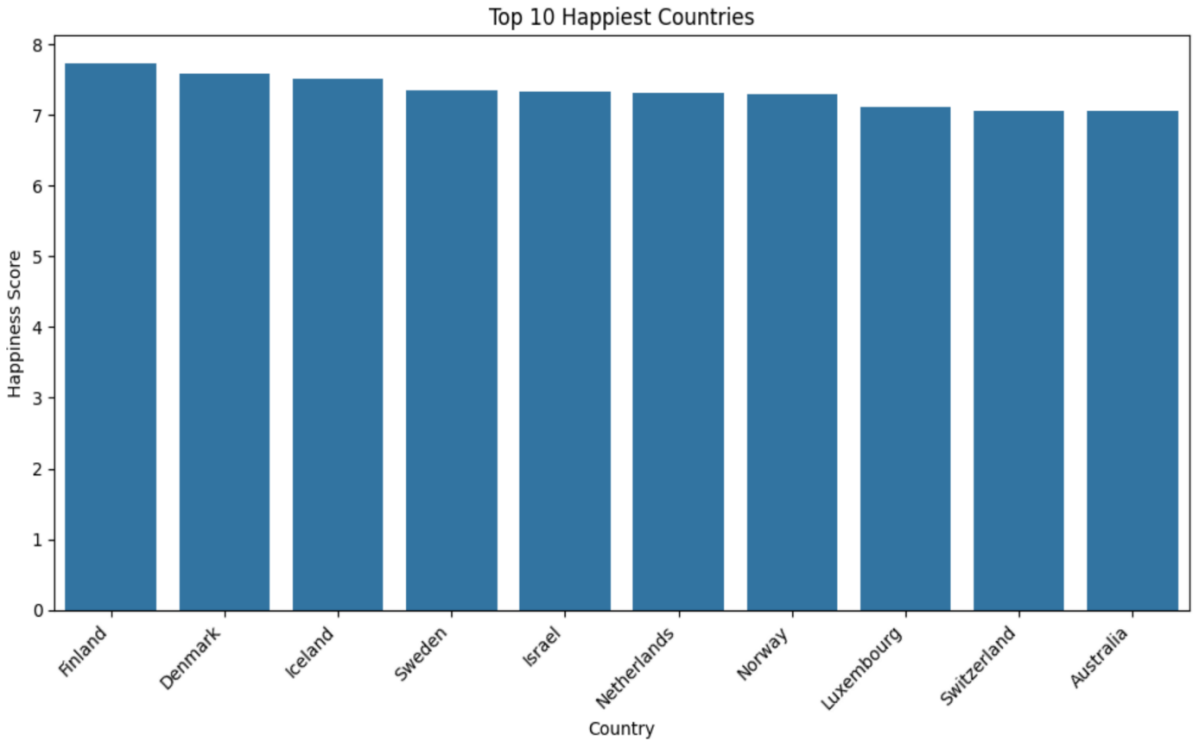
Title

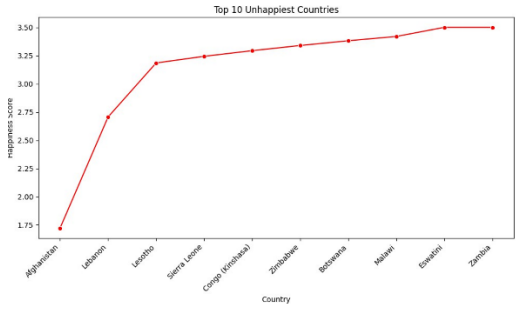
- Analysis of the World Happiness Report : A Data-Driven Exploration of Global and Regional Trends.

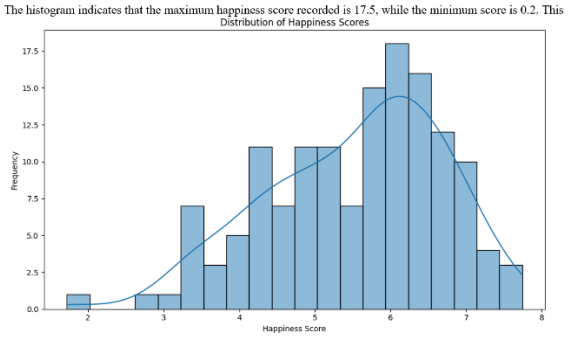
B. Introduction

- "The World Happiness Report serves as a crucial resource for understanding the well-being of countries globally. It emphasizes various factors contributing to happiness, such as economic stability, social support, and health. This report aims to analyze happiness scores in South Asia and the Middle East, exploring descriptive statistics, performance comparisons, and correlation analyses to derive insights about regional happiness."

Problem 1: Preparing the Middle Eastern Dataset







Problem - 2 - Some Advance Data Exploration Task:

Task - 1 - Setup Task - Preparing the South-Asia Dataset:

Steps:

1. Define the countries in South Asia with a list for example:

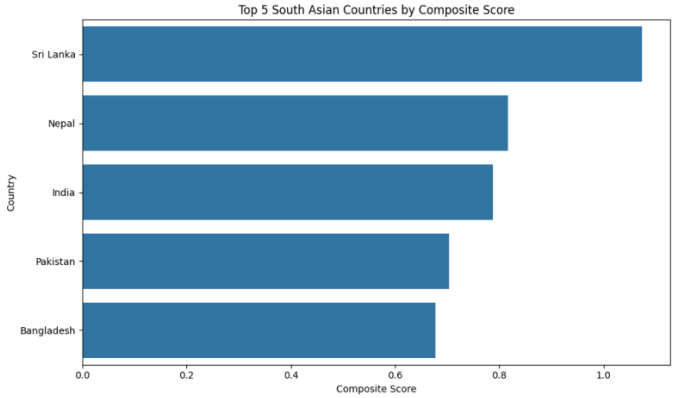
south asian countries = ["Afghanistan", "Bangladesh", "Bhutan", "India",

"Maldives", "Nepal", "Pakistan", "Srilanka"]

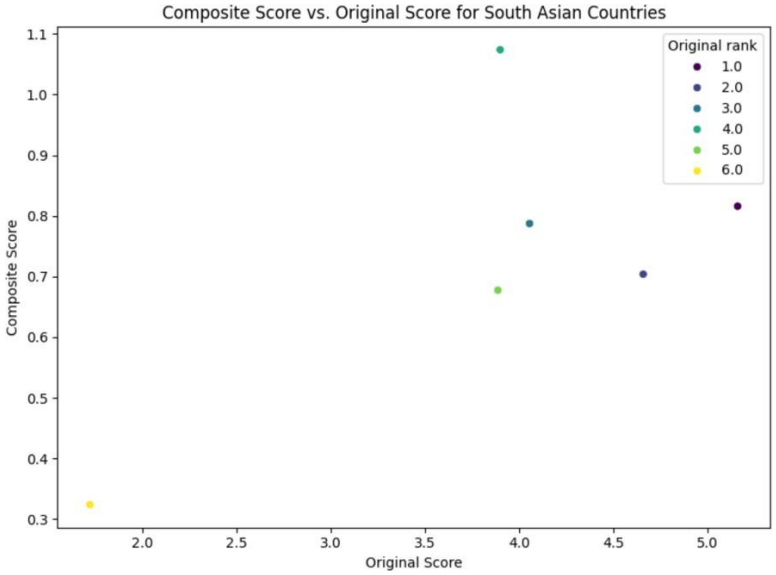
2. Use the list from step - 1 to filtered the dataset {i.e. filtered out matching dataset from list.}
3. Save the filtered dataframe as separate CSV files for future use.

Key Findings

1. **Top 5 South Asian Countries by Composite Score**:
    - The bar plot shows the top 5 South Asian countries by Composite Score: Sri Lanka, Nepal, India, Pakistan, and Bangladesh.
    - Sri Lanka has the highest Composite Score at 1.0739, followed by Nepal (0.8159), India (0.7874), Pakistan (0.7039), and Bangladesh (0.6774).



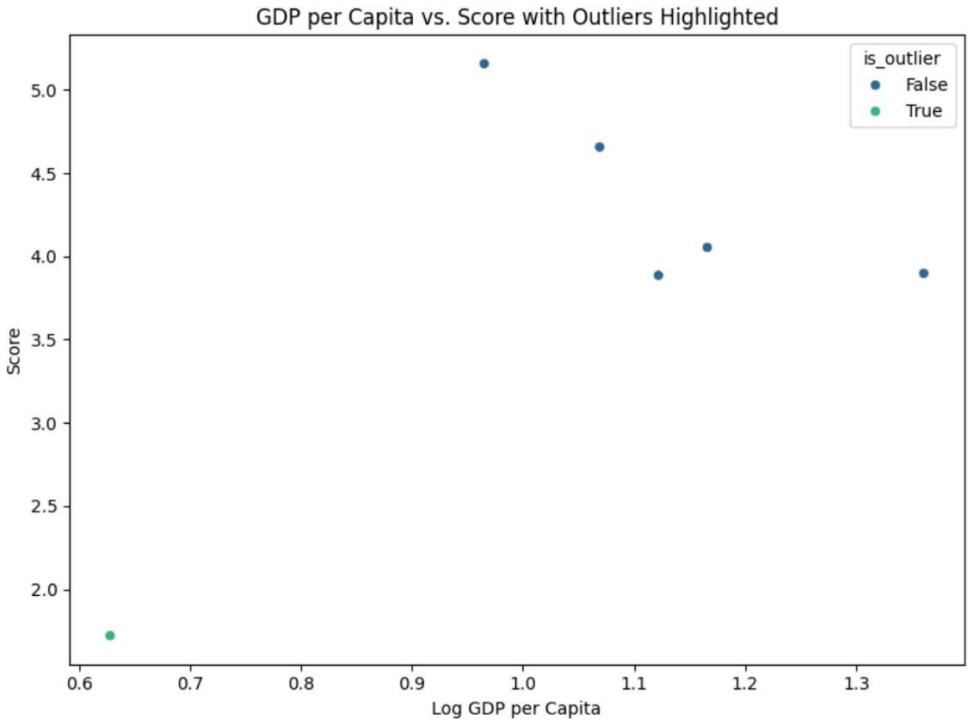
1. **Relationship between Composite Score and Original Score**:
    - The scatter plot shows the relationship between the Composite Score and the Original Score for the South Asian countries.
    - The plot reveals a positive correlation between the two scores, indicating that the Composite Score is generally aligned with the Original Score.
    - However, there are some deviations, as seen by the varying distances of the data points from the diagonal line, suggesting that the Composite Score provides additional insights beyond the Original Score.


Composite Score vs. Original Score for South Asian Countries
Original rank
1.0
2.0
3.0
4.0
5.0
6.0
Composite Score
Original Score
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
2.0
2.5
3.0
3.5
4.0
4.5
5.0

Task - 3 - Outlier Detection: Tasks:

1. Identify outlier countries in South Asia based on their Score and GDP per Capita.

The analysis identified the top five South Asian countries by Composite Score, with Sri Lanka leading at 1.0739, followed by Nepal (0.8159), India (0.7874), Pakistan (0.7039), and Bangladesh (0.6774). A scatter plot illustrated the relationship between the Composite Score and Original Score, revealing a positive correlation that indicates alignment between the two, though some data points deviated from the diagonal line, suggesting that the Composite Score offers additional insights. Outlier detection highlighted Afghanistan as a significant outlier, exhibiting a notably lower happiness score and GDP per capita, which could distort regional averages and perceptions of overall well-being. The findings suggest that while rankings based on Composite Scores generally reflect Original Scores, the Composite Score provides a more nuanced view of well-being; for instance, Sri Lanka's higher Composite Score indicates strong performance across the evaluated factors, whereas Bangladesh's lower score points to potential challenges in economic and social support metrics.



Task - 4 - Exploring Trends Across Metrics:

Tasks:

1. Choose two metrics (e.g., Freedom to Make Life Choices and Generosity) and calculate their correlation

{pearson correlation} with the Score for South Asian countries.

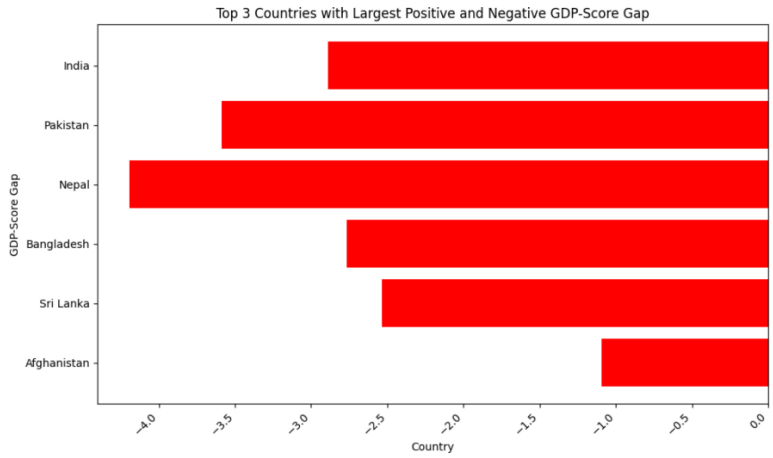
The analysis revealed notable correlations between various factors and overall happiness scores. A strong positive correlation of 0.8005 was found between the Freedom to Make Life Choices and happiness, indicating that countries with greater individual freedoms tend to report higher levels of happiness. Generosity also demonstrated a positive relationship with a correlation of 0.8745, although it was slightly weaker than that of freedom, suggesting it contributes to happiness but not as significantly. The findings emphasize that enhancing personal freedoms could greatly improve happiness levels in South Asian countries, while generosity, though beneficial, may play a lesser role in influencing overall well-being.

Task - 5 - Gap Analysis:

Tasks:

1. Add a new column, GDP-Score Gap, which is the difference between GDP per Capita and the Score for each South Asian country.

The gap analysis involved calculating a new column, the GDP-Score Gap, which measures the difference between GDP per capita and happiness scores for each country, revealing whether economic status corresponds with happiness levels. Notably, countries like Nepal (-4.193), Pakistan (-3.588), and India (-2.888) displayed significant positive gaps, indicating higher GDP per capita relative to their happiness scores. Conversely, Afghanistan (-1.093), Sri Lanka (-2.537), and Bangladesh (-2.764) exhibited negative gaps, suggesting lower GDP per capita compared to their happiness ratings. Descriptive statistics showed that the mean happiness score for South Asia was 3.90 with a standard deviation of 1.18, while the Middle East had a higher mean score of 5.41 and a standard deviation of 1.57. In terms of performance, Bhutan (6.62), Sri Lanka (5.87), and Nepal (5.16) were the top three in South Asia, whereas Israel (7.34), Kuwait (6.95), and the UAE (6.73) led in the Middle East. The bottom performers included Afghanistan (1.72), Pakistan (4.66), and India (4.05) in South Asia, and Lebanon (2.71), Yemen (3.56), and Jordan (4.19) in the Middle East.



3.3 Problem - 3 - Comparative Analysis:

Task - 1 - Setup Task - Preparing the Middle Eastern Dataset:

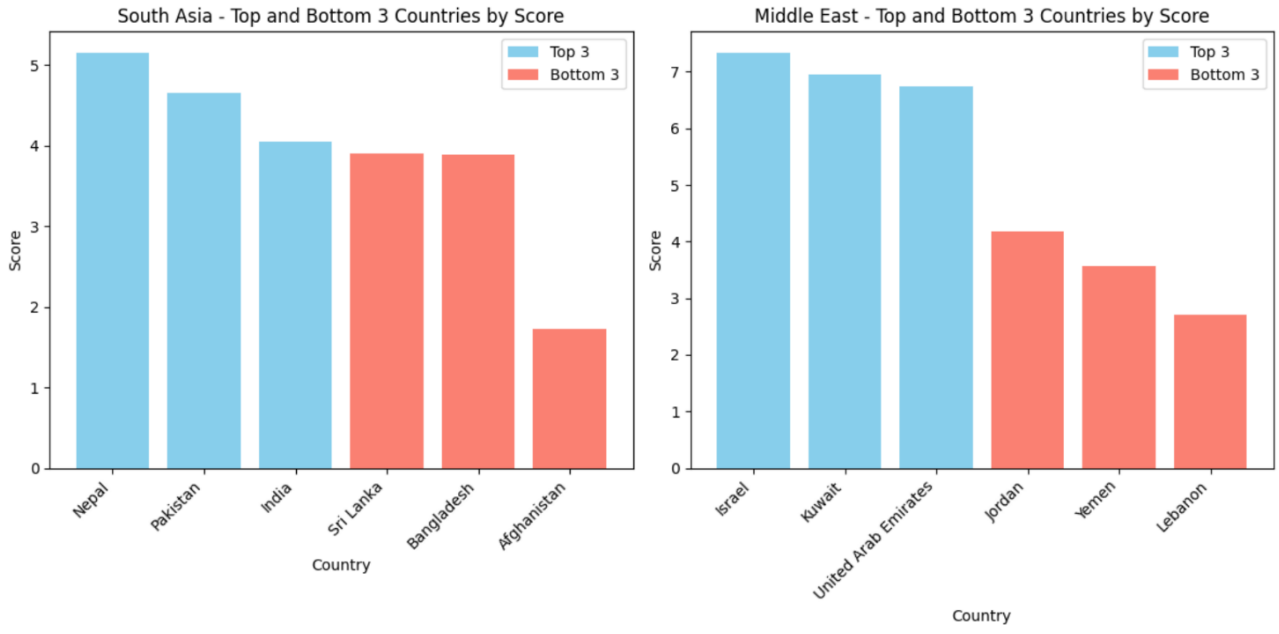
Tasks:

1. Similar in Task - 1 of Problem 2 create a dataframe from middle eastern countries. For hint use the following list:

middle east countries = ["Bahrain", "Iran", "Iraq", "Israel", "Jordan",
"Kuwait", "Lebanon", "Oman", "Palestine", "Qatar", "Saudi Arabia", "Syria",

"United Arab Emirates", "Yemen"]

The analysis revealed several key findings regarding happiness and economic indicators in the region. Israel recorded the highest happiness score at 7.341, while Lebanon had the lowest at 2.707, indicating significant differences in well-being. In terms of economic performance, the Log GDP per capita exhibited considerable variation, with the United Arab Emirates leading at 1.983 and Yemen trailing at 0.671, suggesting a potential link between economic conditions and happiness. Additionally, social support scores were notably high in Kuwait and Israel, at 1.364 and 1.513, respectively, which aligns with their elevated happiness ratings. In contrast, Lebanon's lower social support score of 0.577 highlights the difficulties it faces in promoting community well-being.

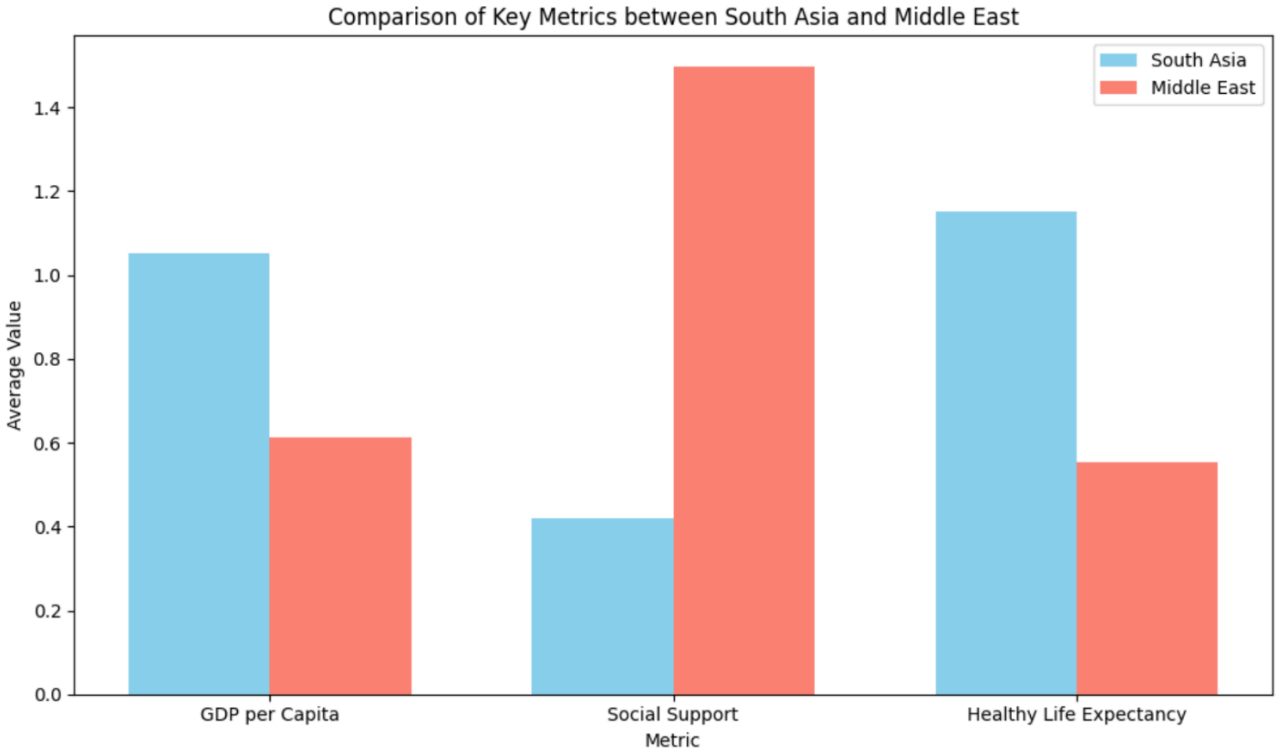


3. Metric Comparisons:

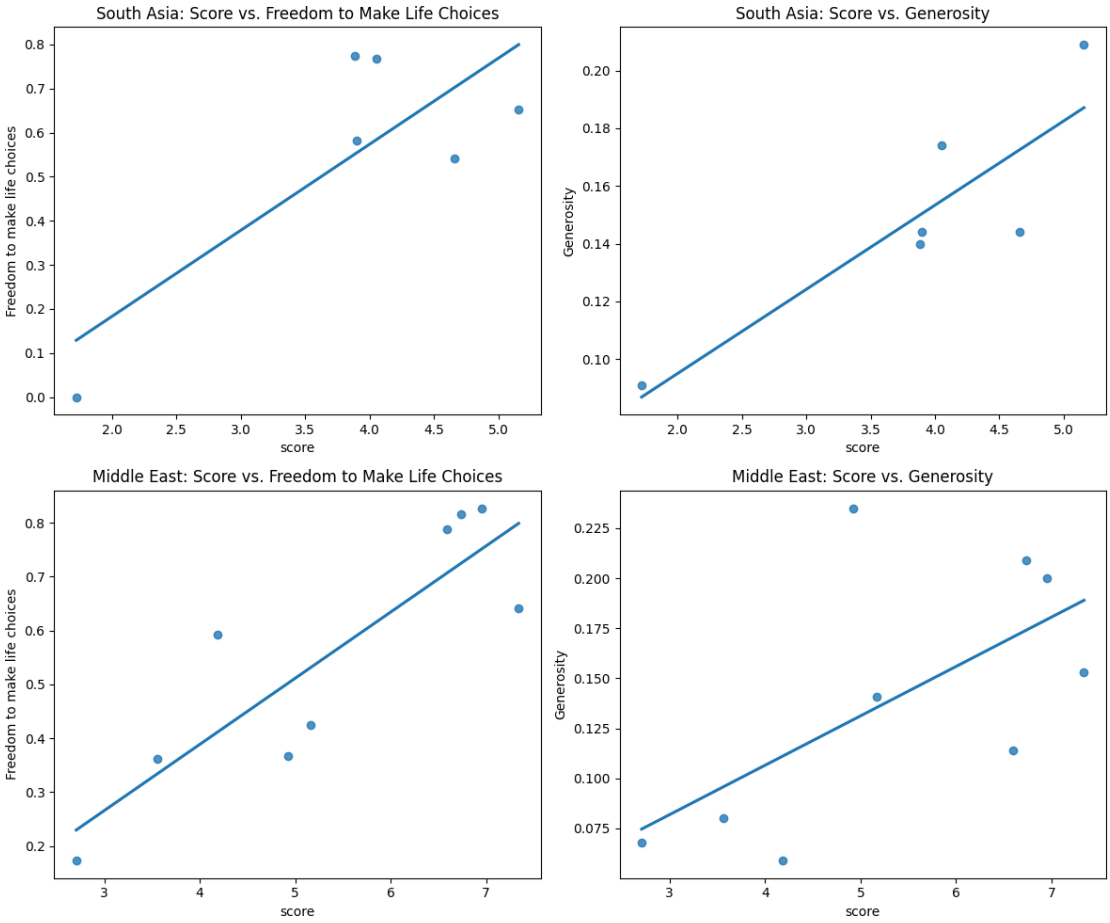
• Compare key metrics like GDP per Capita, Social Support, and Healthy Life Expectancy between the regions using grouped bar charts.

• Which metrics show the largest disparity between the two regions?

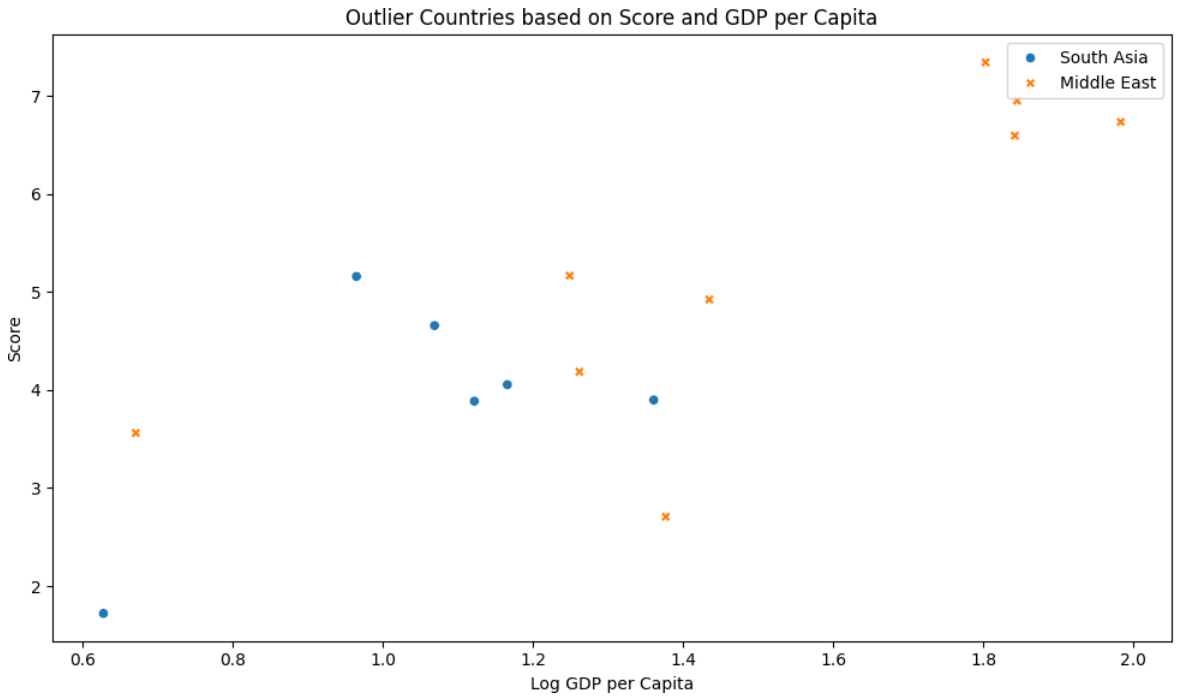
In this study, average values for key metrics were computed for both South Asia and the Middle East, and a DataFrame was constructed to aid in visualizing these metrics. A grouped bar chart was created to depict the average values of GDP per Capita, Social Support, and Healthy Life Expectancy, clearly illustrating the differences between the two regions. Additionally, the analysis revealed the metric with the greatest disparity, underscoring the significant variations in well-being between South Asia and the Middle East.




```
The largest disparity is observed in: Social Support
```



Outlier countries in both regions have been identified based on their happiness scores and GDP per capita. These outliers typically possess distinct characteristics that significantly impact their happiness levels and economic performance. Their presence can distort regional averages and comparisons, emphasizing the necessity for a nuanced interpretation of the data when analyzing regional trends.


Outlier Countries based on Score and GDP per Capita
South Asia
Middle East
Score
Log GDP per Capita
2
3
4
5
6
7
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0