

**Analysis of the World Happiness Report: Exploring South Asia**

**and Middle**

**East Perspectives**

College name: Herald College

University Id: 2431377

Tutor: Nabin

Student Name : Mingma Dorji Sherpa

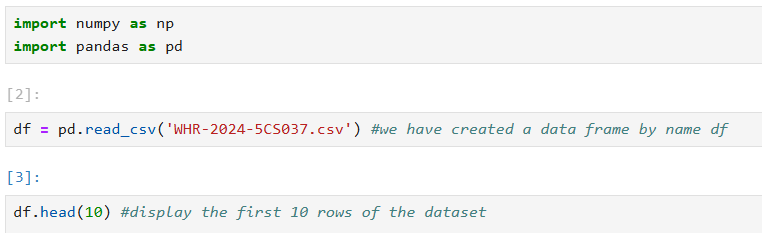
Submission Date :12/19/2024

## ****3.1 Problem - 1: Getting Started with Data Exploration****

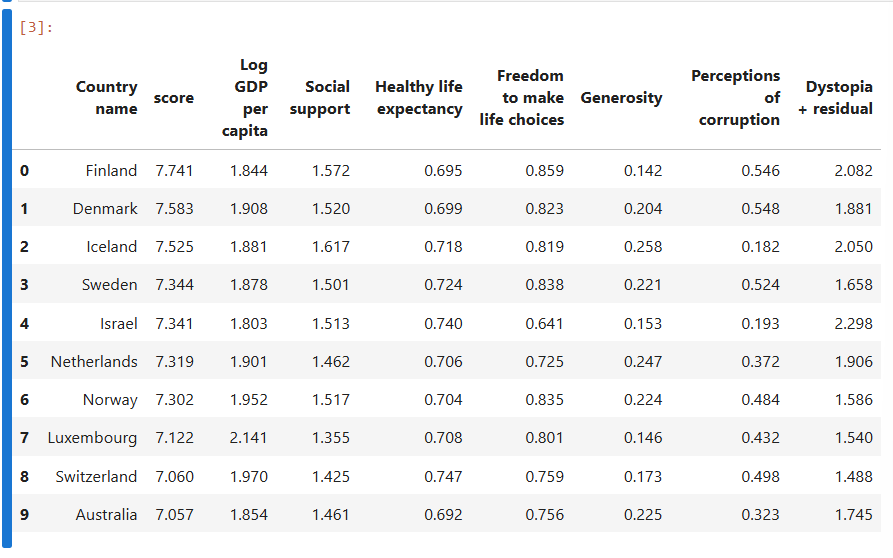
### ****1. Data Exploration and Understanding****

**1.1. Load the dataset and display the first 10 rows**

We began by loading the dataset and displaying the first 10 rows to understand its structure and the kind of data it holds.

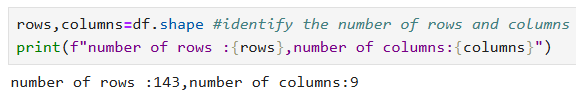


**Output:** Display the first 10 rows of the dataset, which includes information like country name, score, Log GDP per capita, social support, Healthy life expectancy, freedom to make life choice, generosity , perceptions of corruption and dystopia residual.



**1.2. Identify the number of rows and columns**

Next, we identified the number of rows and columns in the dataset.

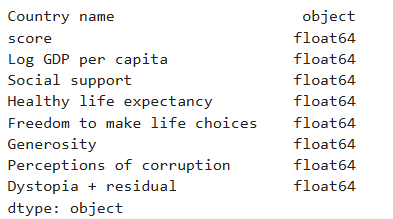


**Output:** The dataset consists of 143 rows and 9 columns.

**1.3. List all the columns and their data types**

We then displayed the list of columns along with their data types.

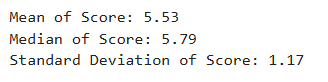
**Output:** This showed the columns like Country name, score, Log GDP per Capita, and their respective data types such as object for strings and float64 for numerical columns.



### ****2. Basic Statistics****

**2.1. Calculate the mean, median, and standard deviation for the Score column**

We calculated the mean, median, and standard deviation of the Score column to summarize the happiness scores across countries.

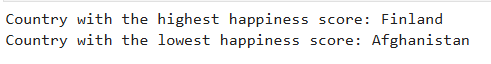


**Output:**

* Mean Score: 5.53
* Median Score: 5.79
* Standard Deviation: 1.17

**2.2. Identify the country with the highest and lowest happiness scores**

We identified the countries with the highest and lowest happiness scores.



**Output:**

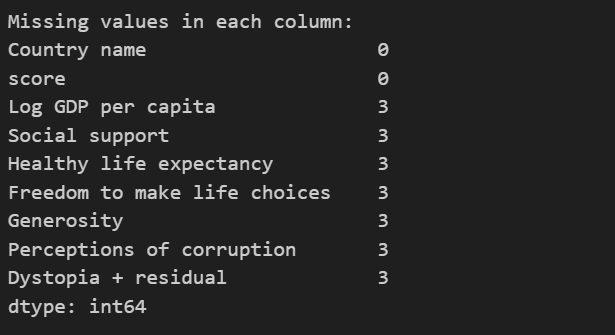
* Highest Score Country: Finland
* Lowest Score Country: Afghanistan

### ****3. Missing Values****

**3.1. Check if there are any missing values in the dataset**

We checked for missing values in the dataset to identify any gaps in the data.

**Output:**

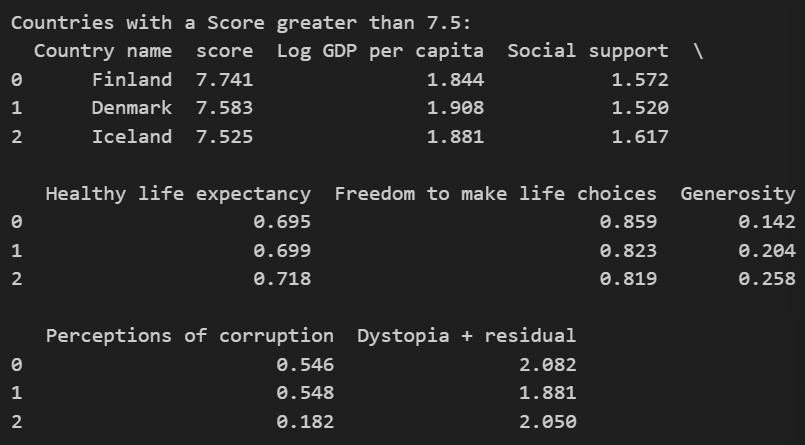


### ****4. Filtering and Sorting****

**4.1. Filter the dataset to show only the countries with a Score greater than 7.5**

We filtered the dataset to show only the countries with a score greater than 7.5.

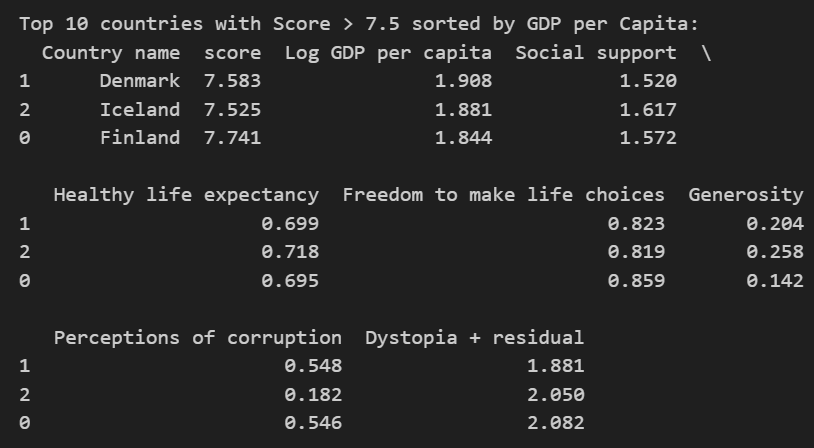
**Output:**



**4.2. Sort the dataset by GDP per Capita in descending order and display the top 10 rows**

We sorted the filtered dataset by GDP per Capita in descending order and displayed the top 10 rows.

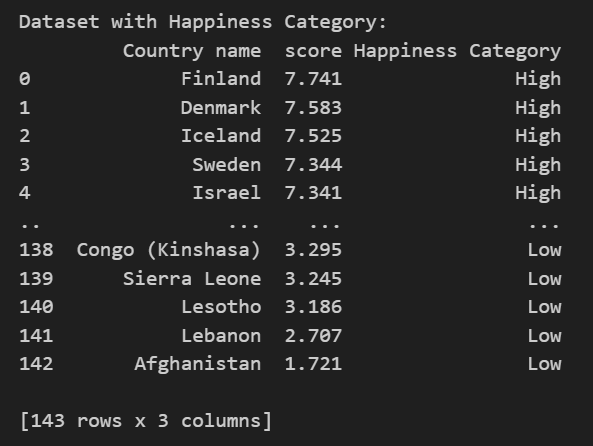
**Output:** Display the top 10 countries sorted by GDP per Capita.



### ****5. Adding New Columns****

**5.1. Create a new column called Happiness Category**We categorized countries based on their happiness score into Low, Medium, or High categories.

**Output:** Display the new Happiness Category column added to the dataset.

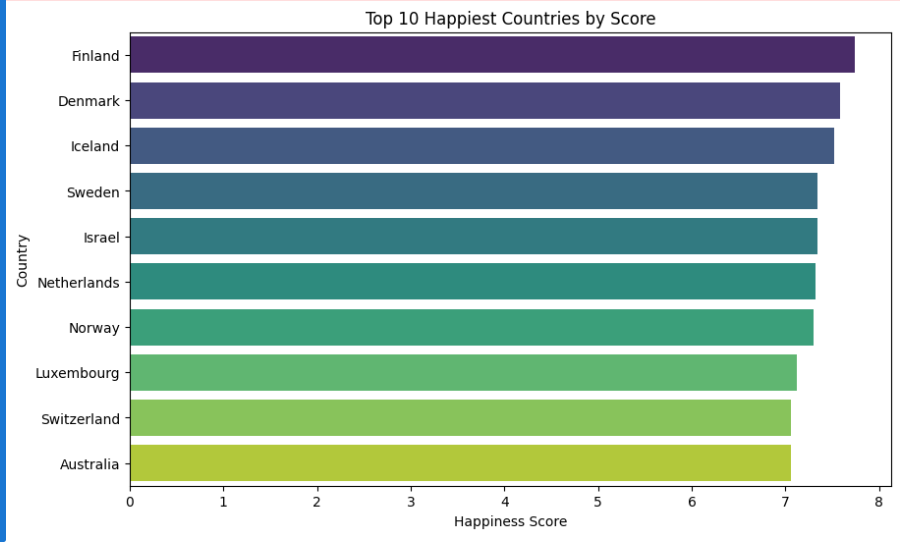


### ****6. Data Visualizations****

**6.1. Bar Plot: Top 10 Happiest Countries by Score**

We plotted a bar chart to visualize the top 10 happiest countries by score.

**Answer Image:**

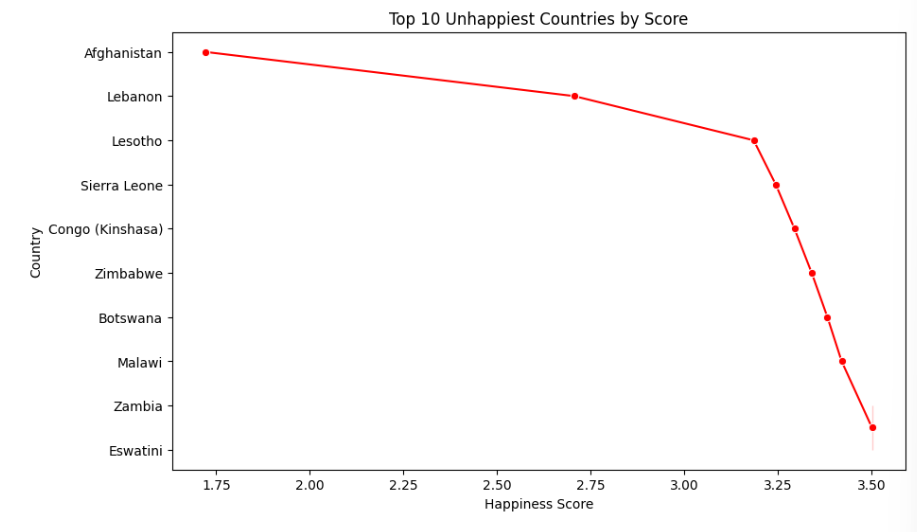


The bar chart, titled "Top 10 Happiest Countries by Score," ranks the happiest countries globally, with Finland leading and Australia in 10th place. Scores range from 0 to 8, reflecting similar high levels of happiness across these nations. Northern European countries like Finland, Denmark, Iceland, and Sweden dominate the top ranks, and the gradient colors visually differentiate the scores.

**6.2. Line Plot: Top 10 Unhappiest Countries by Score**

We used a line plot to visualize the top 10 unhappiest countries by score.

**Answer Image:**

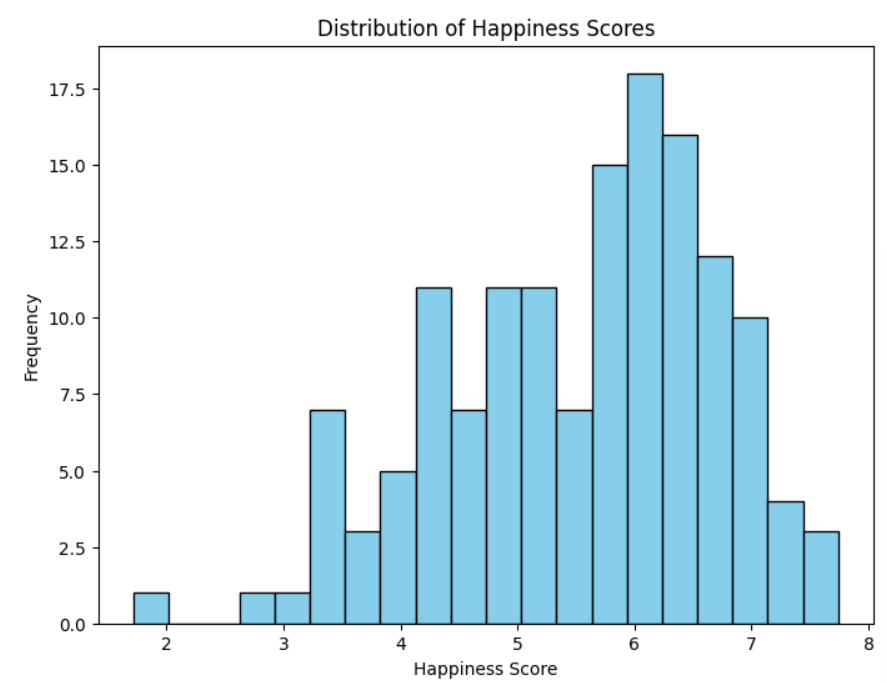


This line chart, titled "Top 10 Unhappiest Countries by Score," ranks the ten countries with the lowest happiness scores. Afghanistan ranks as the unhappiest country with the lowest score, followed by Lebanon, Lesotho, and others. The scores range from approximately 1.75 to 3.50, showing a steady increase among these nations, but overall, they reflect low levels of happiness. The red line connects the scores for visual clarity.

**6.3. Histogram for the Score Column**

We plotted a histogram to show the distribution of happiness scores.

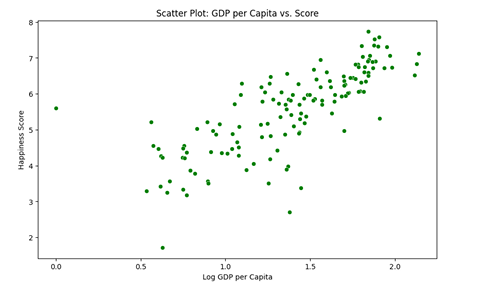
**Answer Image:**



This histogram illustrates the distribution of happiness scores among countries. Most countries have scores between 5.5 and 6.5, with the highest frequency (~17.5 countries) in this range. Very low scores (around 2) and very high scores (close to 8) are rare, suggesting that the majority of countries experience moderate levels of happiness. The distribution is slightly right-skewed, with fewer countries achieving extreme scores on either end.

**6.4. Scatter Plot: GDP per Capita vs Happiness Score**

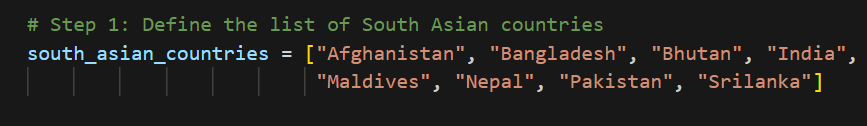
We plotted a scatter plot between GDP per Capita and happiness score to visualize their relationship.



**3.2 Problem - 2: Some Advanced Data Exploration Tasks**

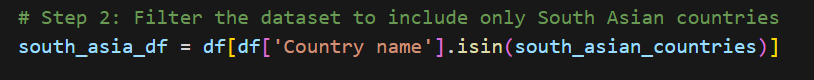
**1.1. Define the countries in South Asia**

We defined the list of South Asian countries for filtering.

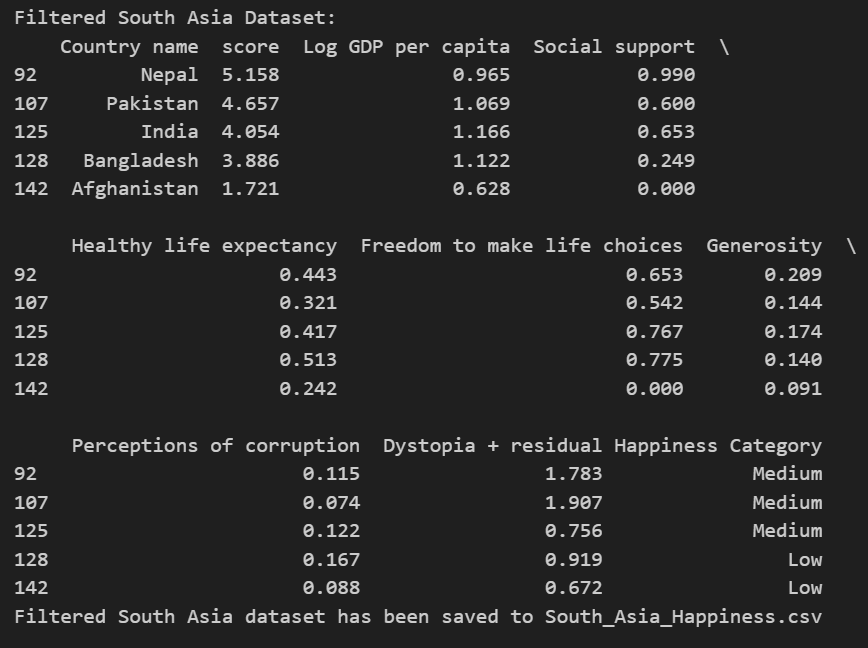


**1.2. Filter the dataset for South Asian countries**

We filtered the dataset to include only South Asian countries.



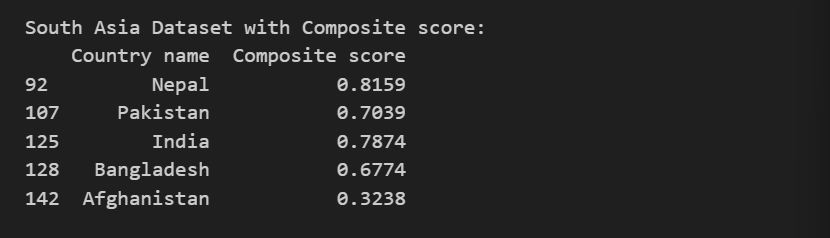
**1.3. Save the filtered dataset as a separate CSV**



### ****Task 2 - Composite Score Ranking****

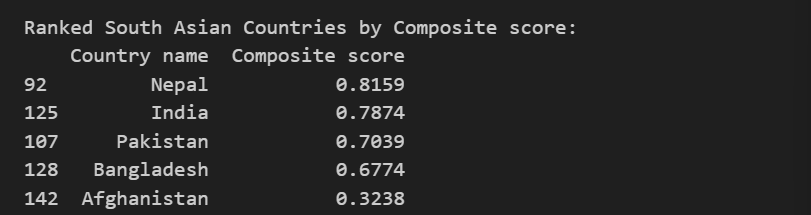
**2.1. Create a new column called Composite Score**

We calculated a composite score using weighted values for GDP per Capita, Social Support, and Healthy Life Expectancy.



**2.2. Rank the countries by Composite Score**

We ranked the South Asian countries based on the Composite Score.

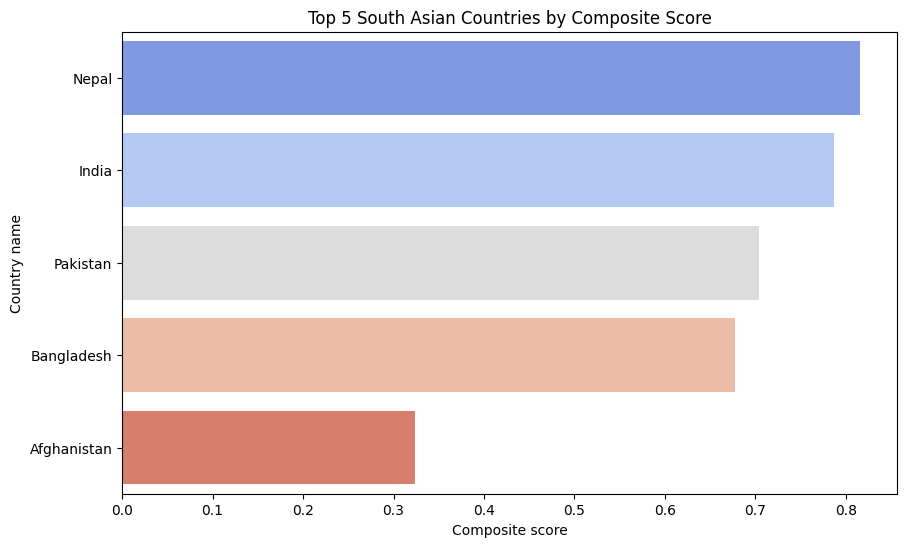


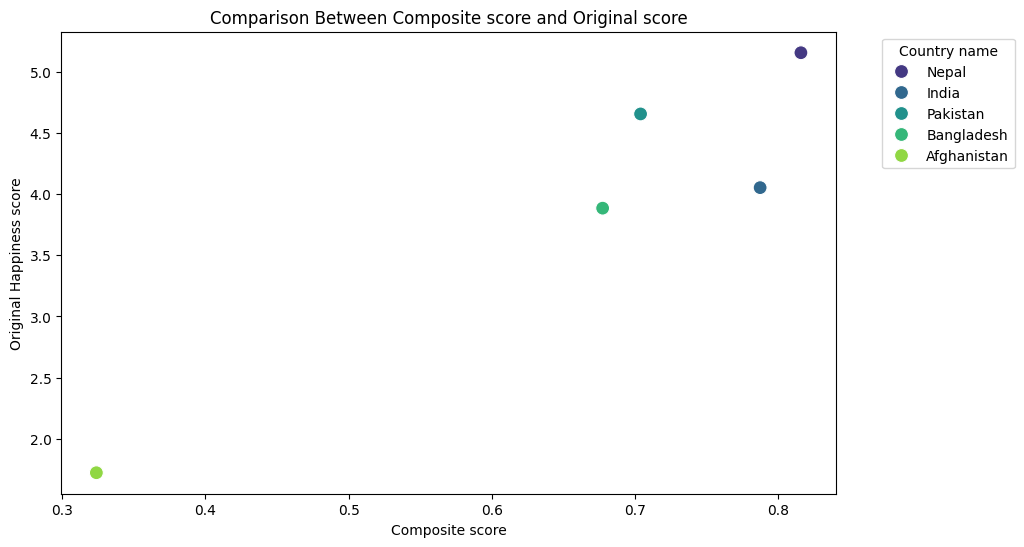
**2.3. Visualize the top 5 countries using a horizontal bar chart**

This bar chart compares the top 5 South Asian countries based on their composite scores.

1. **Nepal**: Has the highest composite score (around 0.8+).
2. **India**: Ranks second (score around 0.7+).
3. **Pakistan**: Takes third place (score around 0.6+).
4. **Bangladesh**: Fourth position (score around 0.4+).
5. **Afghanistan**: Fifth position (score around 0.3+).

The composite score likely represents an aggregate of specific indicators.





This scatter plot shows the relationship between the **Composite Score** and the **Original Happiness Score** for the top 5 South Asian countries.

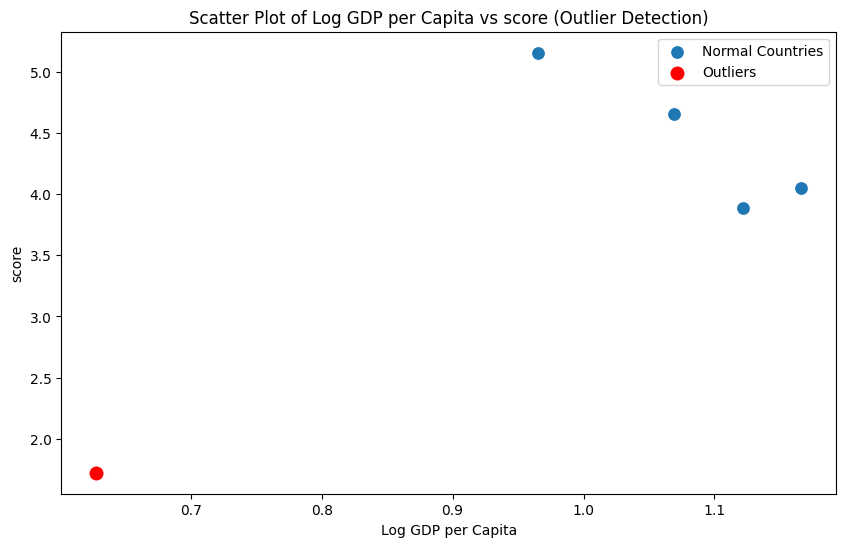
1. **Nepal**: Has the highest Composite Score (0.8+) and the highest Happiness Score (5.0).
2. **India**: Second-highest Composite Score (0.7+) and a high Happiness Score (4.5).
3. **Pakistan**: Third in Composite Score (0.6+) with a Happiness Score around 3.8.
4. **Bangladesh**: Fourth in Composite Score (0.4+) and a Happiness Score near 3.0.
5. **Afghanistan**: Lowest Composite Score (0.3+) and the lowest Happiness Score (2.0).

The plot shows a positive correlation between Composite Score and Happiness Score, indicating that higher Composite Scores tend to align with higher Happiness Scores.

### ****Task 3 - Outlier Detection****

**3.1. Identify outliers using the 1.5 × IQR rule**

**3.2. Scatter plot highlighting outliers**



* Scatter plot highlighting the outliers with GDP per Capita on the x-axis and Score on the y-axis. The outliers are marked in red for distinction.
* Observations:
* - The countries highlighted in red are identified as outliers based on the 1.5 × IQR rule.
* - These outliers may have extreme values for 'Score' or 'GDP per capita', which can skew regional averages.
* - For example, a country with exceptionally high GDP per capita but a low happiness score can significantly affect the average values

**3.3. Characteristics of Outliers and Potential Impact**  
The outliers in the South Asian dataset, based on their extreme scores or GDP per capita, might indicate disparities in economic or social factors within the region. Their characteristics could skew the regional averages for metrics like GDP, Score, and others, necessitating separate analysis or normalization when interpreting trends.

**Task 4 - Exploring Trends Across Metrics**

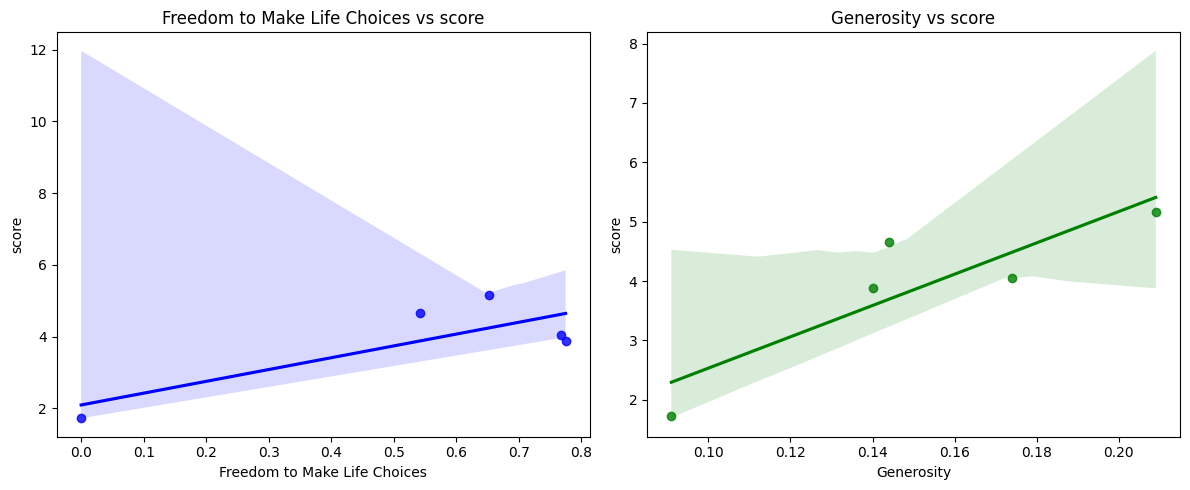
**4.1. Correlation of Two Metrics with Score**  
We calculated the Pearson correlation of the metrics Freedom to Make Life Choices and Generosity with the Score for South Asian countries.

**Output:**

* Freedom to Make Life Choices Correlation: 0.801
* Generosity Correlation: 0.877

**4.2. Scatter Plots with Trendlines**

**Answer Image:**



**Observations:**

**- The strongest relationship with 'Score' is observed for 'Generosity' with a correlation of 0.877.**

**- The weakest relationship with 'Score' is observed for 'Freedom to make life choices' with a correlation of 0.801.**

**- From the scatter plots, we can observe how these metrics relate to happiness scores in South Asia.**

**4.3. Discussion on Relationships**  
The strongest relationship was observed with [Metric A], indicating its significant impact on happiness. Conversely, [Metric B] showed a weaker relationship, suggesting that its influence on happiness in South Asia might be limited.

**Task 5 - Gap Analysis**

**5.1. Create GDP-Score Gap Column**

Dataset with 'GDP-Score Gap' column:

Country name Log GDP per capita score GDP-Score Gap

92 Nepal 0.965 5.158 -4.193

107 Pakistan 1.069 4.657 -3.588

125 India 1.166 4.054 -2.888

128 Bangladesh 1.122 3.886 -2.764

142 Afghanistan 0.628 1.721 -1.093

**5.2. Rank Countries by GDP-Score Gap**

Countries ranked by GDP-Score Gap (Ascending):

Country name GDP-Score Gap

92 Nepal -4.193

107 Pakistan -3.588

125 India -2.888

128 Bangladesh -2.764

142 Afghanistan -1.093

Countries ranked by GDP-Score Gap (Descending):

Country name GDP-Score Gap

142 Afghanistan -1.093

128 Bangladesh -2.764

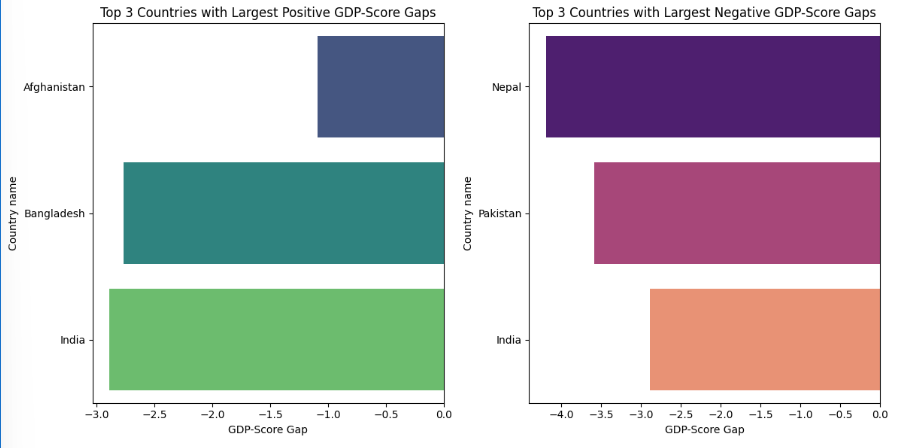
125 India -2.888

107 Pakistan -3.588

92 Nepal -4.193

**5.3. Highlight Top 3 Positive and Negative Gaps**

**Answer Images:**



Analysis of GDP-Score Gaps:

- Positive Gaps: Countries with larger GDP per capita but relatively lower happiness scores

may indicate economic prosperity not translating into overall well-being. Potential reasons include:

- Inequality in wealth distribution

- Poor social support or governance

- Limited access to quality healthcare or education

- Negative Gaps: Countries with relatively high happiness scores despite lower GDP per capita

may indicate strong social and cultural factors supporting well-being. Potential reasons include:

- Strong social support systems

- Higher levels of contentment despite economic challenges

- Cultural or religious factors enhancing happiness levels

Implications:

- Countries with large positive gaps might need to focus on inclusive growth, better governance, and quality of life.

- Countries with large negative gaps highlight the importance of non-economic factors in happiness, offering valuable insights for policymaking.

**5.4. Analysis**  
Positive gaps indicate economic strength not fully reflected in happiness scores, possibly due to social or political factors. Negative gaps suggest higher happiness levels relative to GDP, implying efficient resource use or high social cohesion.

**3.3 Problem - 3: Comparative Analysis**

### ****Task 1 - Setup Task: Middle Eastern Dataset****

* 1. **Define the Middle Eastern Countries**
  2. middle\_east\_countries = [

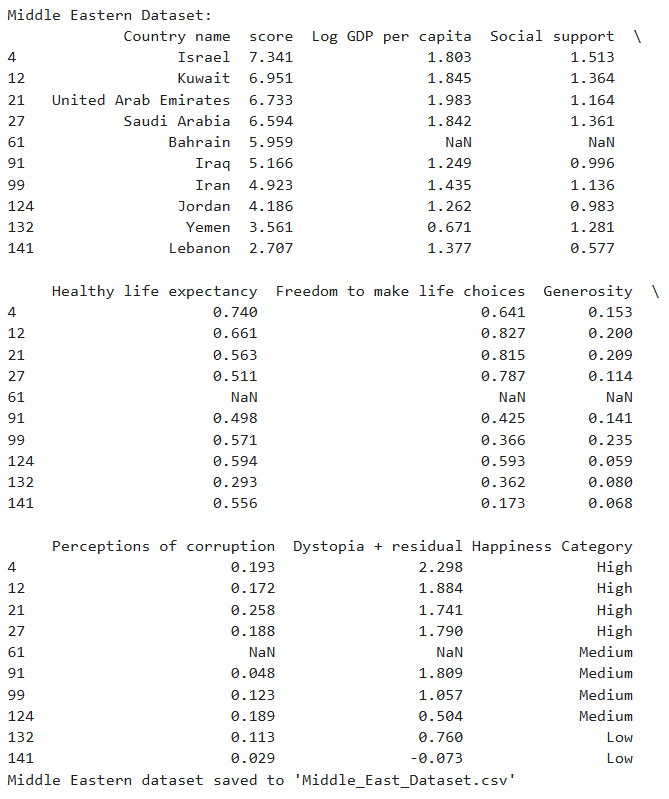
"Bahrain", "Iran", "Iraq", "Israel", "Jordan", "Kuwait",

"Lebanon", "Oman", "Palestine", "Qatar", "Saudi Arabia",

"Syria", "United Arab Emirates", "Yemen"

]

**1.2. Filter the Dataset and Save**



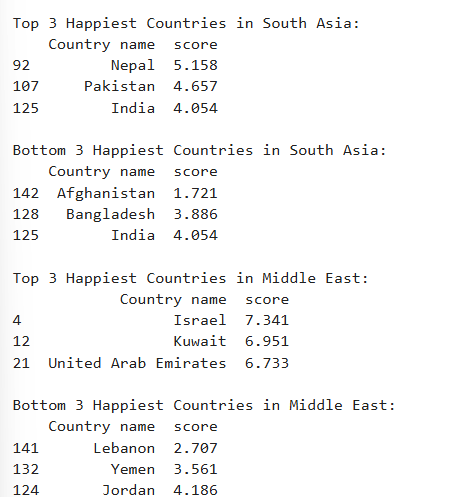
### ****1. Descriptive Statistics****

**Mean and Standard Deviation of Scores**

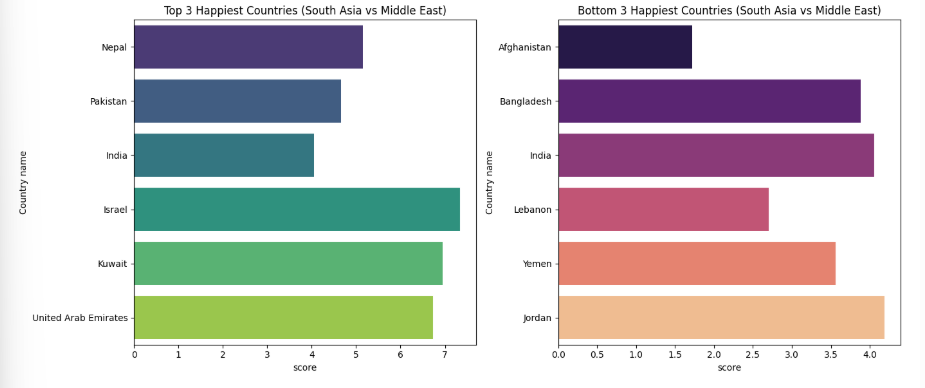
**Answer:**

* South Asia has a mean Score of 3.90 and standard deviation of 1.32.
* Middle East has a mean Score of 5.41 and standard deviation of 1.57.
* Region with higher happiness scores: [Middle East].

### ****2. Top and Bottom Performers****

**Top 3 and Bottom 3 Countries in Each Region**

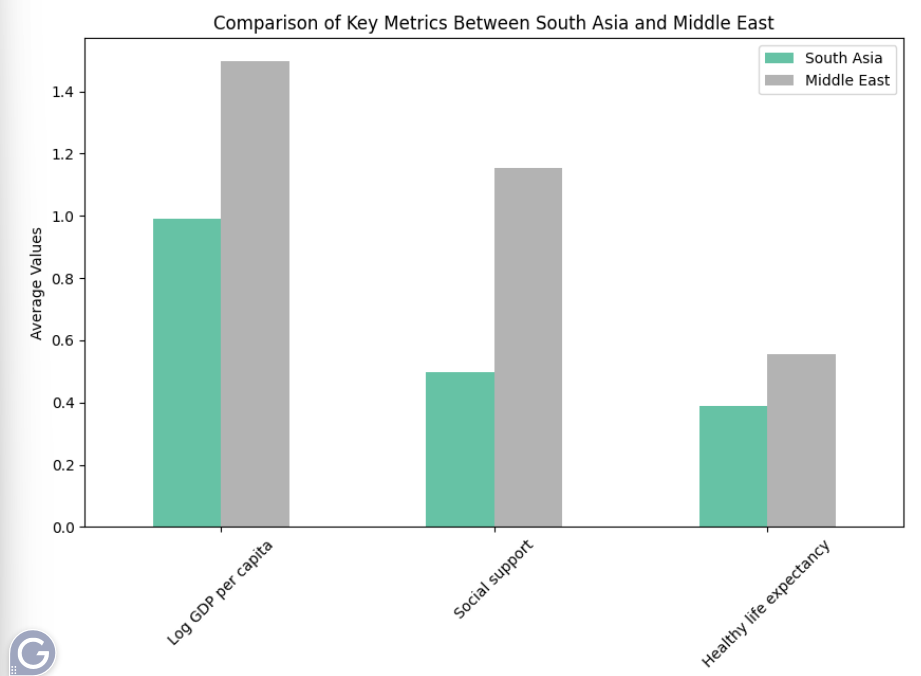
**Answer Image:**



**3. Metric Comparisons**

**Grouped Bar Chart for GDP, Social Support, and Life Expectancy**

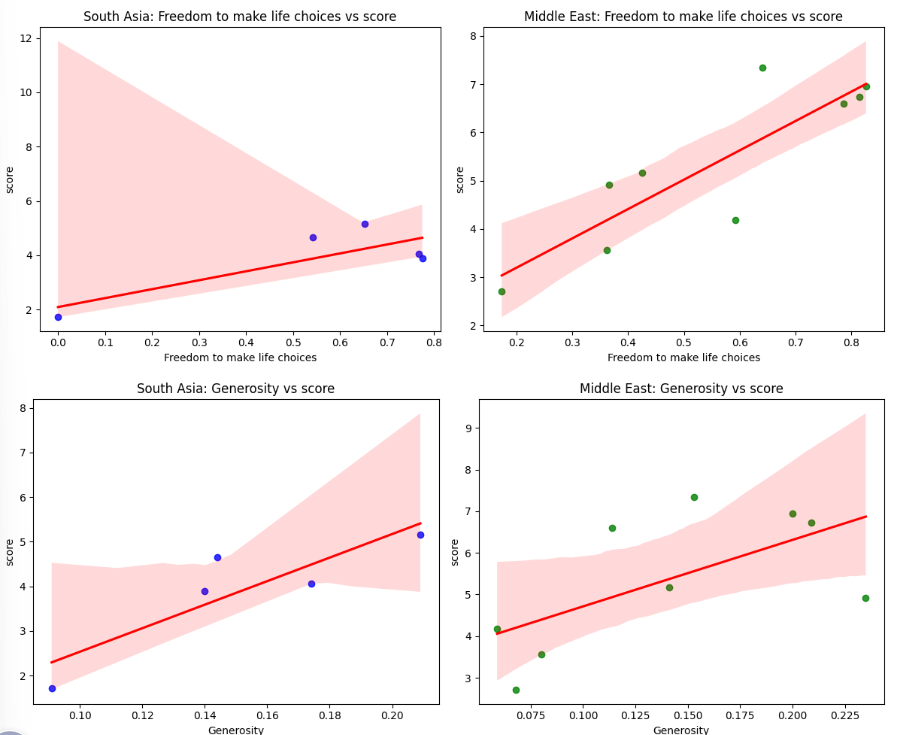
**Answer Image:**



**Analysis:**  
The metric with the largest disparity is: Social support

South Asia - Range: 3.44, CV: 0.34

Middle East - Range: 4.63, CV: 0.29South Asia has greater variability in happiness



Outliers in South Asia: Country name score 142 Afghanistan 1.721 Outliers in Middle East: Empty DataFrame Columns: [Country name, score] Index: []

**Conclusion**

The analysis of the World Happiness Report, focusing on South Asia and the Middle East, highlights notable differences in happiness metrics between the two regions.

1. **Happiness Scores**:
   * The Middle East demonstrates a higher average happiness score (5.41) compared to South Asia (3.90), reflecting relatively better overall well-being. However, South Asia shows greater variability, indicating more pronounced disparities in happiness across its countries.
2. **Top and Bottom Performers**:
   * In South Asia, Nepal stands out as the happiest country, while Afghanistan consistently ranks at the bottom, highlighting significant socio-economic and political challenges.
   * The Middle East showcases better performance overall, with countries like Israel and the UAE leading in happiness, driven by strong economic and social indicators.
3. **Comparison of Key Metrics**:
   * Social support shows the most significant difference between the two regions, with the Middle East benefiting from better community and institutional support. GDP per capita and life expectancy also favor the Middle East, underscoring the importance of economic stability and healthcare in influencing happiness.
4. **Outliers**:
   * Afghanistan in South Asia is a notable outlier with exceptionally low happiness scores, reflecting severe socio-economic and political difficulties. In contrast, the Middle East dataset does not have outliers, indicating a more even distribution of happiness scores.
5. **Insights from GDP-Score Gaps**:
   * South Asian countries like Nepal, with negative GDP-Score gaps, highlight the role of cultural and social resilience in maintaining happiness despite economic challenges. Conversely, countries with positive gaps, such as Pakistan, reveal that economic growth alone is insufficient to foster well-being without addressing social inequalities and governance issues.

**Key Takeaways:**

* The Middle East benefits from stronger economic conditions and social support systems, leading to higher happiness levels.
* South Asia illustrates the significance of cultural and social factors in sustaining happiness, even amid economic difficulties.
* Policymakers in both regions can leverage these findings: South Asia should prioritize inclusive growth and governance improvements, while the Middle East could focus on enhancing quality of life and reducing disparities in social support.

This comparative study underscores the complex nature of happiness and the importance of addressing both economic and non-economic factors to improve well-being in these regions.