Academic Year	Module	Assessment Number	Assessment Type
2024	Concepts and Technologies of AI 5CS037	1	Individual

Report Title:

Analysis of the World Happiness Report: Exploring South Asia and Middle East Perspectives.

Student Id   : 2431875

Student Name   : Adarsh Kumar Mandal

Section   : L5CG13

Module Leader   : Siman   Giri

Tutor   : Durga Pokhrel

Submitted on   : 12-20-2024

Contents

Introduction:

The World Happiness Report is an annual survey that ranks countries according to the happiness and well-being of the citizens. It draws on data from various sources, including the Gallup World Poll,
to present the current knowledge on the determinants of happiness and life satisfaction. The report is important, as it considers many aspects of a relationship that influence the happiness of people in the world; these include living conditions, social support, income, generosity, freedom to choose one's life, and the idea of corruption. The report helps demographers and policymakers understand the health of different groups of people and identify areas for improvement; they can then work to develop programs to improve people's health.

Objectives of the Report and Tasks for Each Section:

1. **Data Exploration:**

   - **Objective**:
     The goal of this analysis is to identify the overall trends and patterns in the World Happiness Report data.

   - **Tasks:**
     Identify important variables by analyzing the dataset, compile the happiness score distribution, and investigate the connections between
     the different factors influencing happiness.

2. **South Asia Analysis:**
   - Filter the dataset to focus on South Asian countries.
   - Compute metrics like Composite Score and rank the countries based on this metric.
   - Identify the outliers using the $1.5 \times \text{IQR}$
     rule and discuss the correlations of happiness with other metrics such as Freedom and Generosity.

3. **South Asia vs. Middle East Comparison**:

   - Compare key metrics: GDP per Capita, Social Support, Healthy Life Expectancy, between the two regions.
   - Identify the largest disparities and discuss implications.
   - Perform a gap analysis to comprehend the reasons for the differences in economic performance and happiness levels.

Problem - 1: Getting Started with Data Exploration

1. ## Data Exploration and Understanding:

   In this section, I have gone through basic data overview which implies loading the data and displaying the head portion of the first 10 countries along with their every aspect of happiness analysis and comparisons.

   In the next step we have talked about the number of rows and columns in the dataset. Similarly, data types of every column are discussed. Then we went through the basic statistics like finding the mean, median and the standard deviation of the score column which are 5.275,5.785,1.707 respectively. We printed the country with the highest and lowest happiness scores, checked for missing values, then filtered and sorted them. Then we added new columns called 'Happiness Category'.

|	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

2. ## Data Visualizations

Bar Plot:

   This bar plot is made to display the data visually for the top 10 happiest countries by the perspective of 'Happiness Score'. If we try to interpret the countries as per the bar chart then we can visually stimulate that Finland is the leading country with highest score meanwhile followed by Denmark, Iceland, Sweden and Israel. Netherland is in the sixth position and Norway at the seventh followed by Luxembourg, Switzerland and Australia.

Top 10 Happiest Countries by Score
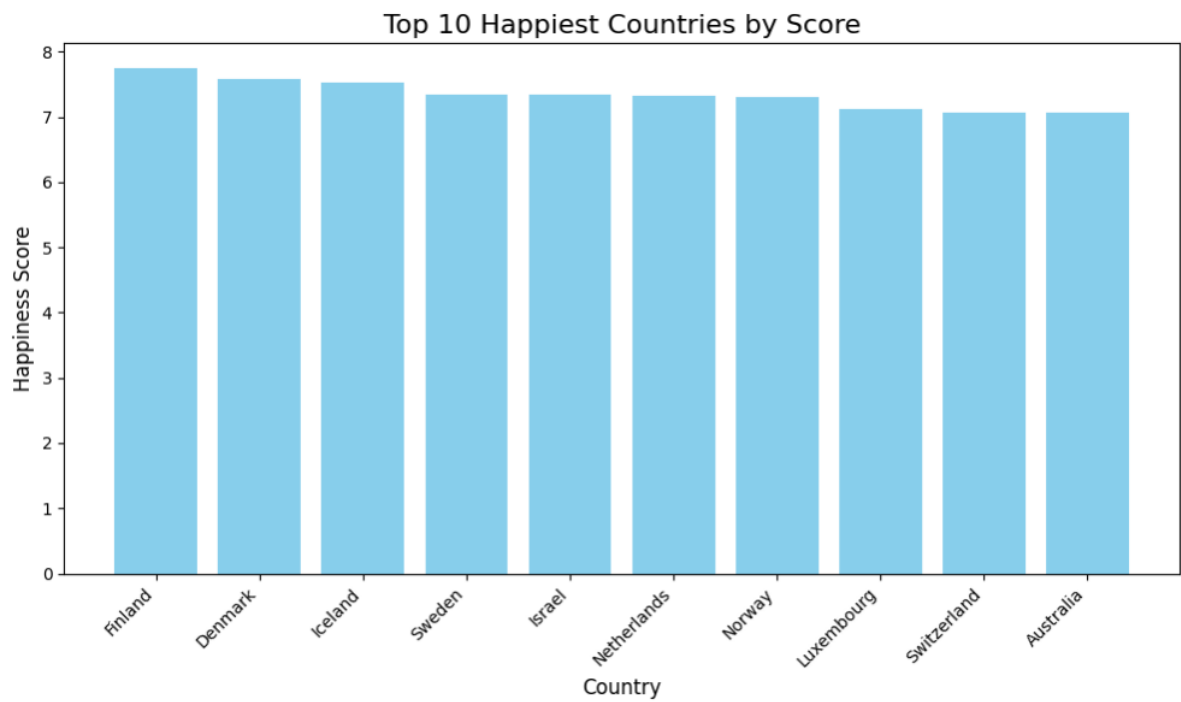
Line Chart:

In this section, we have plotted the line chart to visually represent the data of top 10 unhappiest countries as per their happiness score. For this, we have used matplotlib.pyplot to plot the line chart. Now, if we go through the analysis then the most unhappiness country is Afghanistan with the lowest score followed by Lebanon and Lesotho. The marker represents their happiness score like we have 1.75 for Afghanistan.
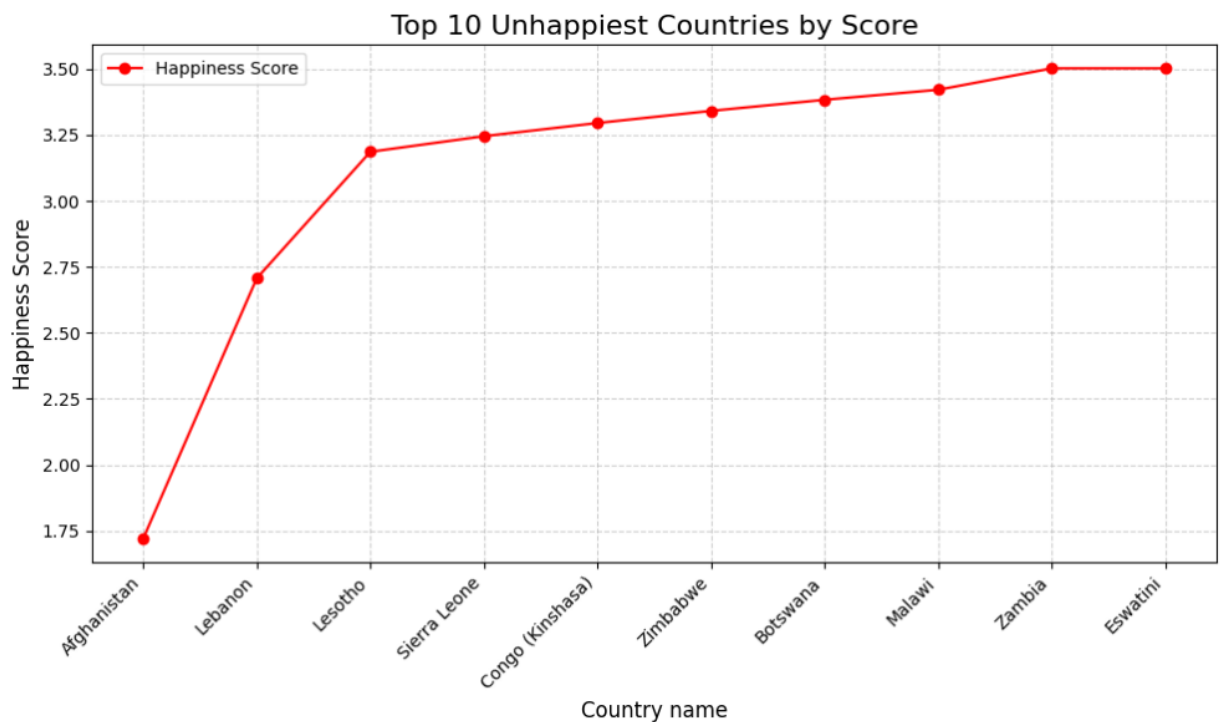
Histogram:

In this we have tried to show the distribution of happiness scores for every country in the dataset then the histogram visualization shows that many countries tend to have scores in the middle range of 4 to 6 which indicates the intermediate level of happiness.

Distribution of Happiness Scores
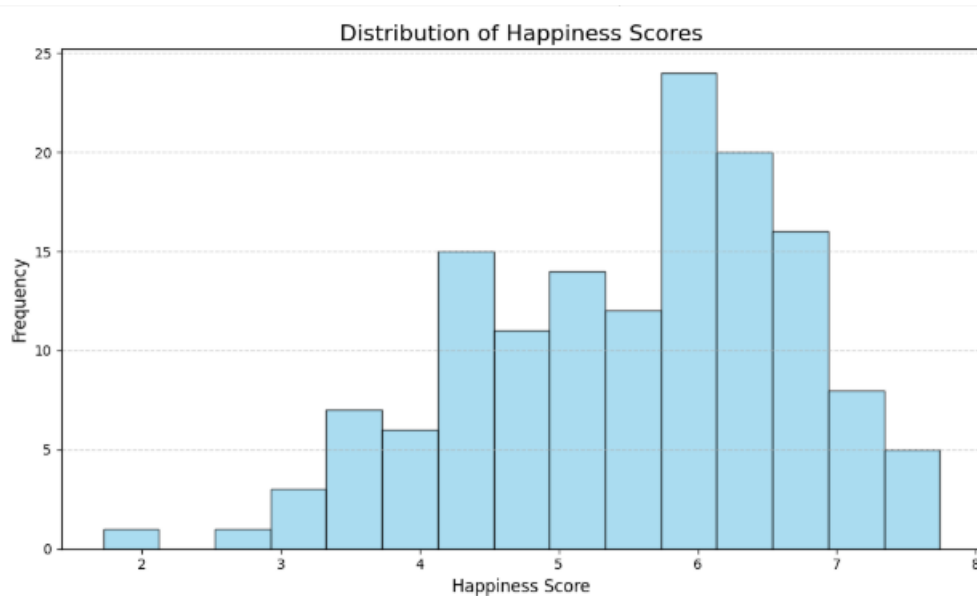
Scatter Plot:

In this visual representation I have tried to visually represent the data of the relationship between GDPs per Capita and Happiness Score. Here a positive trend has been observed which indicates that country with highest GDP per Capita is directly influences the happiness score. Thereby, higher the GDPs, higher the happiness score. Hence, a fluctuation in GDPs significantly affects the happiness score.


Relationship Between GDP per Capita and Happiness Score
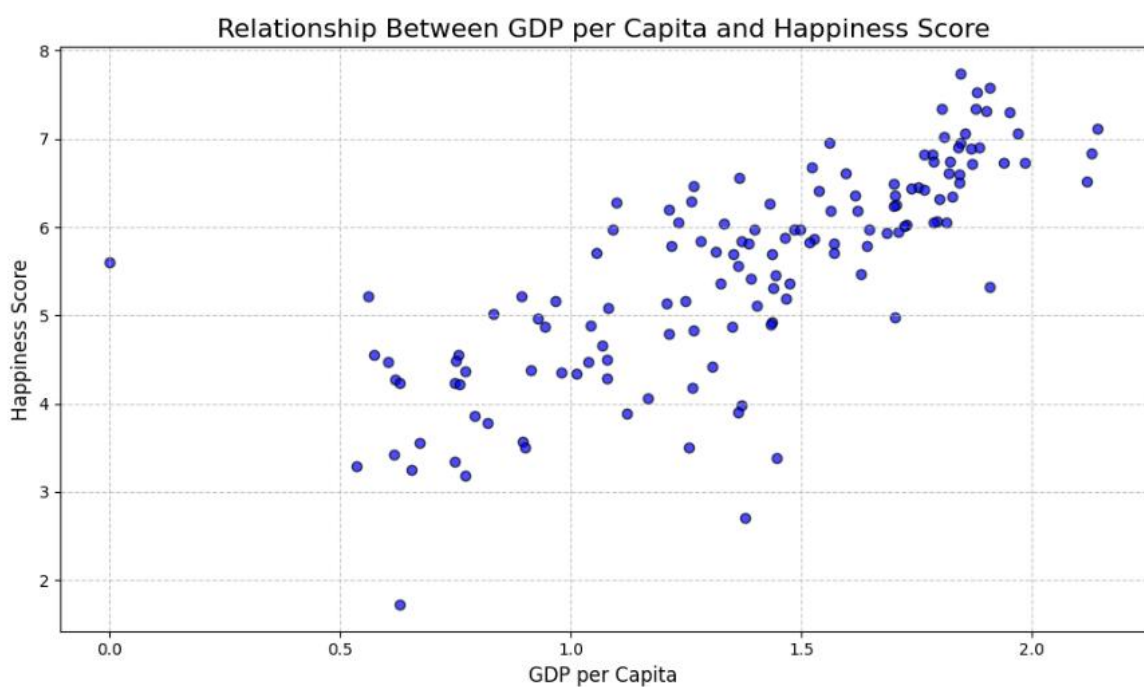
Problem - 2: Some Advance Data Exploration Task

- ## Setup Task - Preparing the South-Asia Dataset

This allows regional analysis by filtering South Asian countries for a subset of nations with similar socio-economic and cultural contexts.
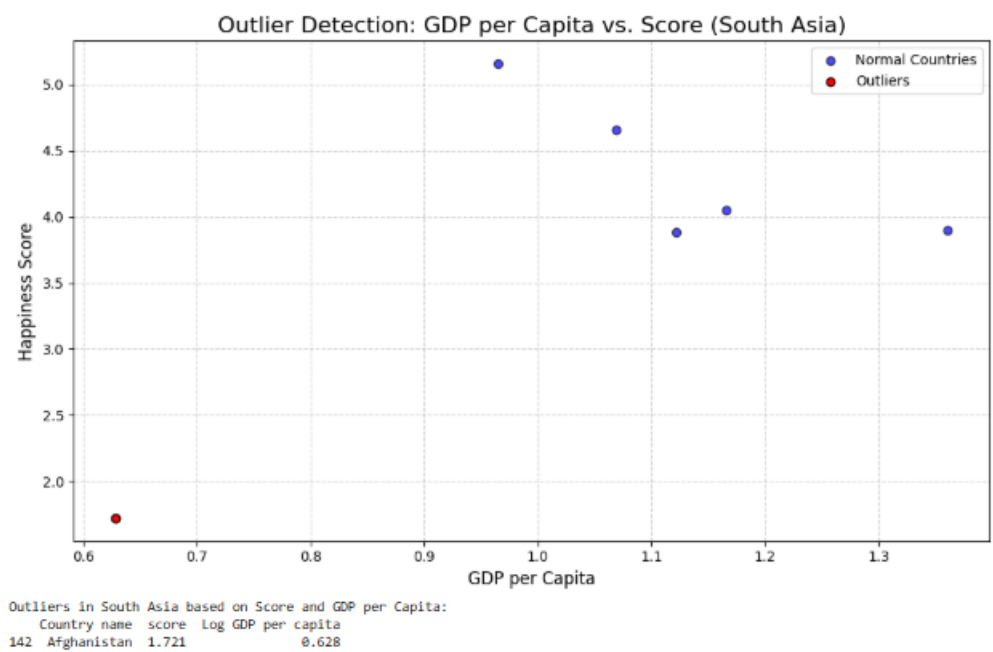
- ## Composite Score Ranking

The Composite Score combines GDP, Social Support, and Healthy Life Expectancy into a balanced ranking system. It highlights the cases of countries performing well in multiple metrics, such as the Maldives, and lagging, such as Afghanistan.
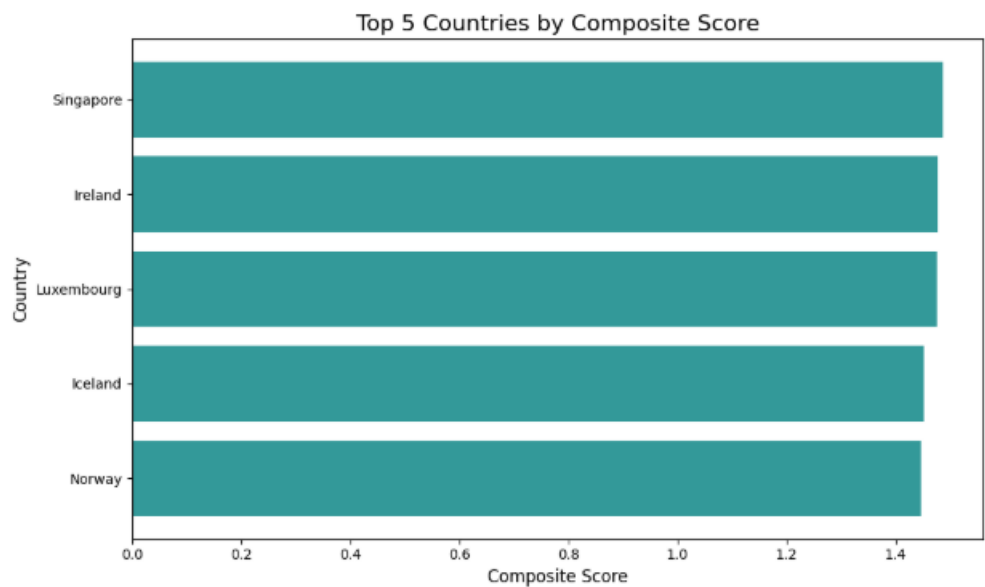

- ## Outlier Detection:

Outliers, such as Afghanistan, with a very low Score and GDP, show the extreme cases. Such outliers distort the results on average and should be considered in isolation for any specific regional analysis.


```
Outliers in South Asia based on Score and GDP per Capita:
        Country name   score   Log GDP per capita
142     Afghanistan    1.721           0.628
```

- ## Exploring Trends Across Metrics

The correlation analysis points to the metrics that most drive happiness: Freedom to Make Life Choices. Poorer relationships, such as in Generosity, point out factors that bear less consequence on well-being in this region.

Score vs. Freedom to Make Life Choices (South Asia)

Score vs. Generosity (South Asia)
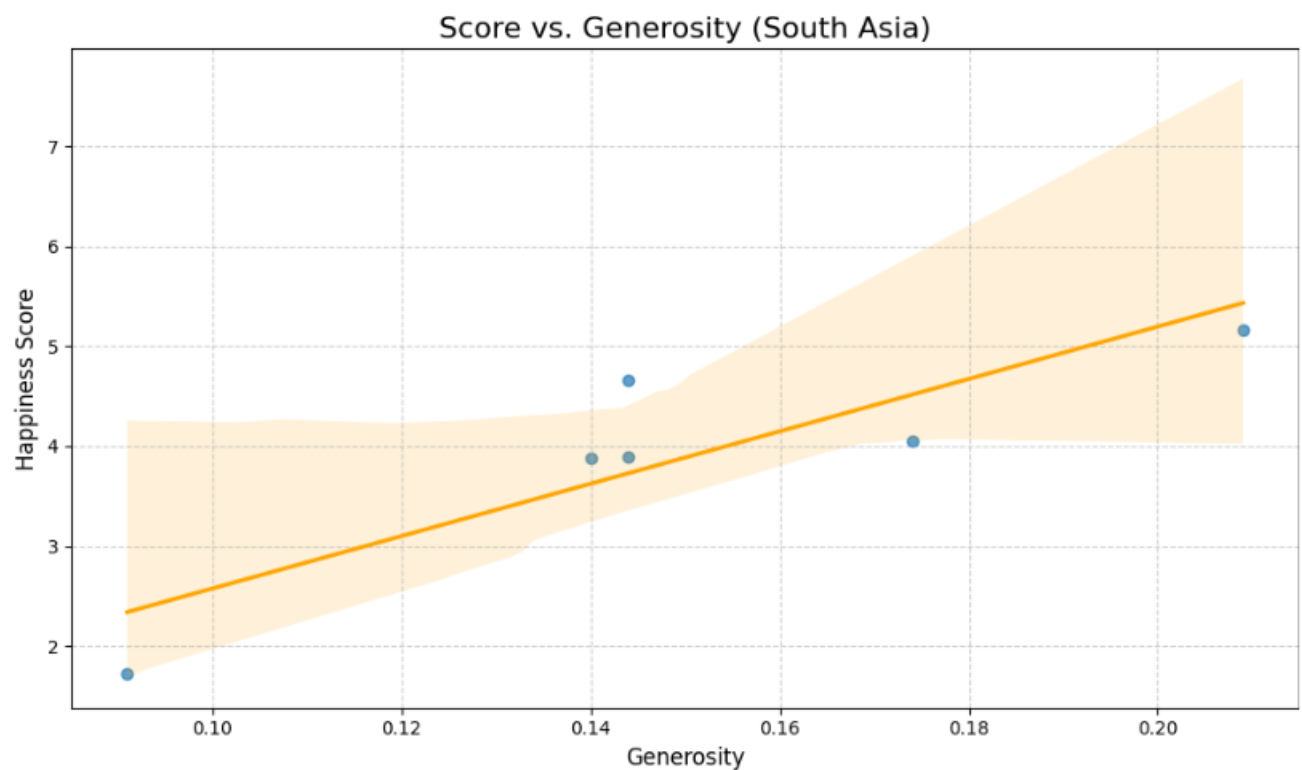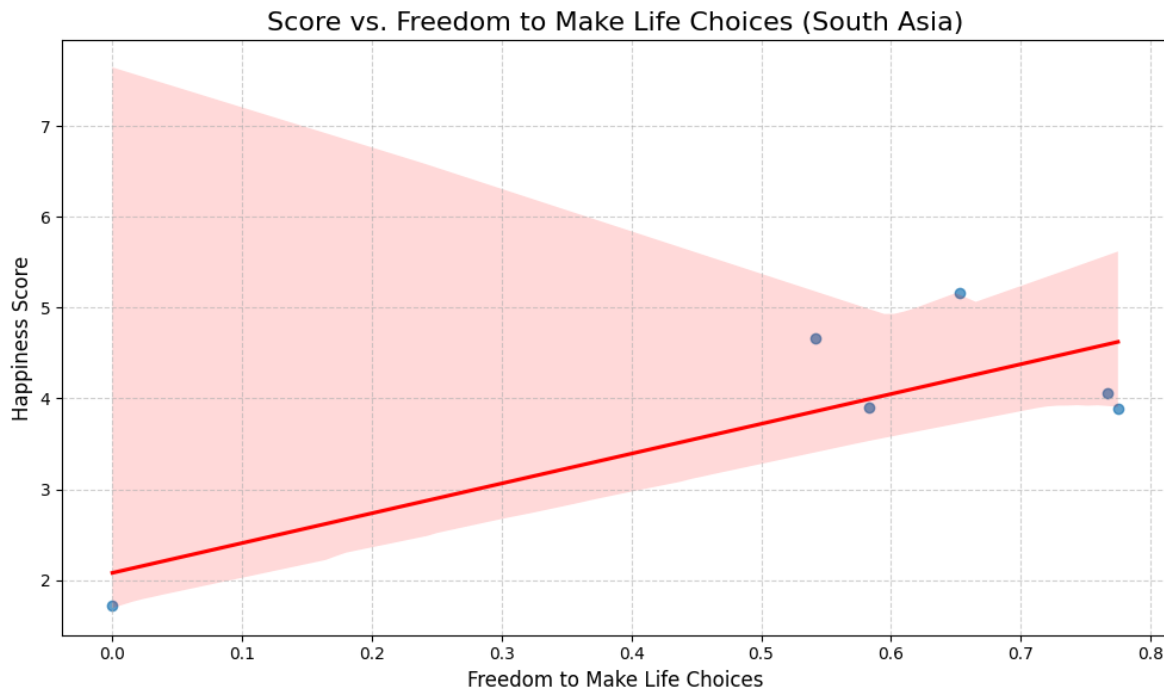
- ## Gap Analysis

The GDP-Score Gap illustrates disparities between economic output and happiness levels. Positive gaps (high GDP, low happiness) suggest inefficiencies in translating wealth into well-being, while negative gaps (low GDP, high happiness) reflect cultural resilience and strong social support systems.
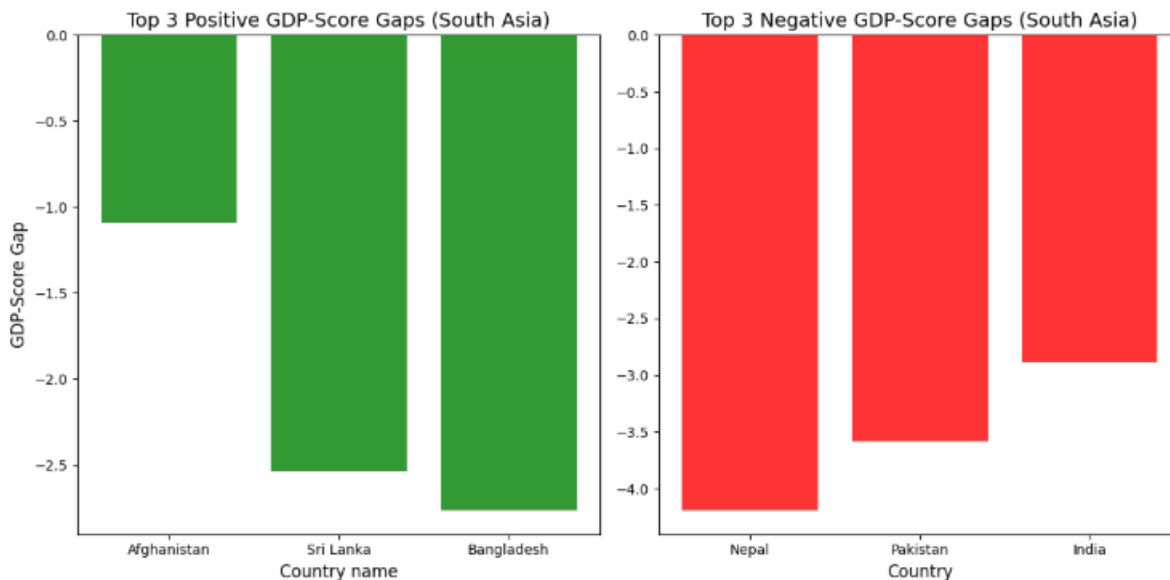

Problem – 3: Comparative Analysis

Preparing the Middle Eastern Dataset: Tasks:

- ## Descriptive Statistics

South Asia represents lower happiness averages compared to the Middle East. While the latter enjoys higher GDP and health infrastructure, South Asia grapples with systemic socio-economic challenges.

- ## Top and Bottom Performers

Top countries like the UAE and Bhutan prove that effective governance and economic development have a great potential impact. The bottom performers, including Yemen and Afghanistan, really point out the devastating impact of conflict and instability on happiness.

Top 3 Performers in Middle East — Bottom 3 Performers in Middle East
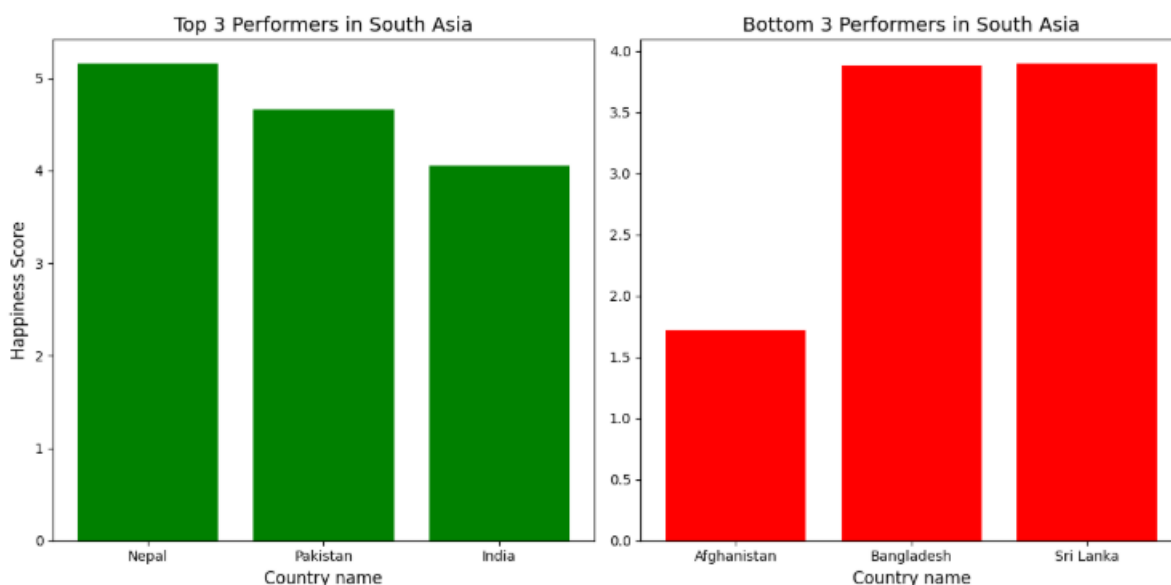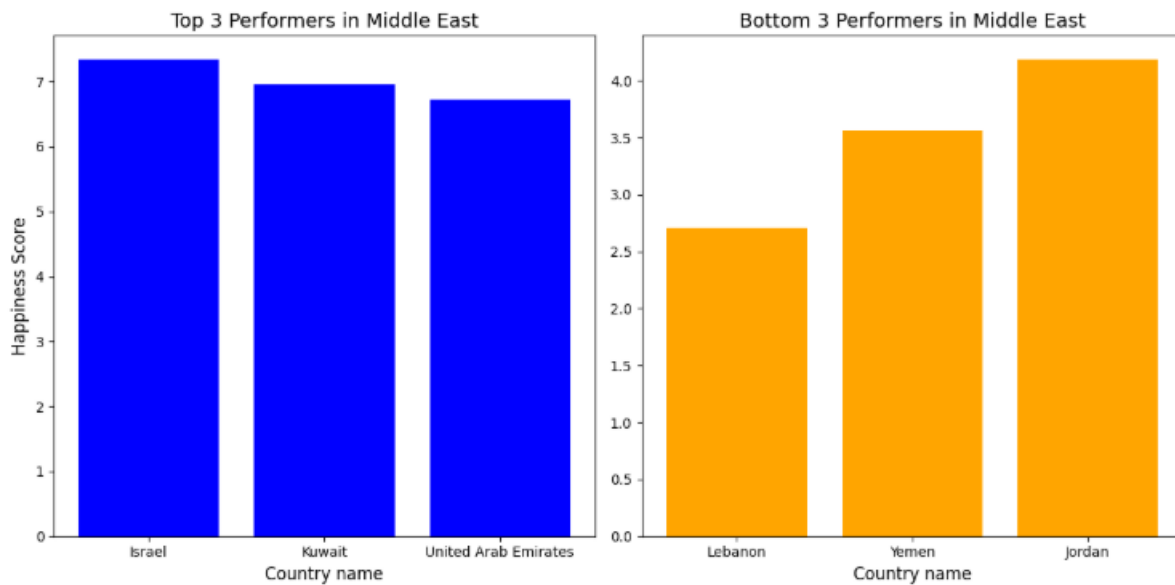
- **Metric Comparisons**

GDP per Capita has the largest range, underlining the Middle East's wealth advantage. Social Support and Healthy Life Expectancy also have an advantage in the Middle East but to a lesser
extent, thus showing more balanced access to those factors.

- **Happiness Disparity**

The higher variability in the Middle East stems from the sharp contrasts between rich countries like UAE and poorer ones like Yemen and Syria. The lower variability of South Asia reflects
more homogeneously low happiness scores within the region.

- **Correlation Analysis**

Metrics like Freedom to Make Life Choices are strongly correlated with happiness, which suggests
that autonomy is an important factor. The weaker correlations with Generosity suggest cultural differences
in perception and impact of this metric.

- **Outlier Detection**

Countries like Afghanistan and Yemen are outliers that show how regional happiness can be severely
influenced by socio-economic and political instability.

- Visualization:
  - Distribution Shapes: South Asia's scores are very closely bunched at the low end, while for the Middle East, this distribution is wider due to disparities in wealth.
  - Medians: The Middle East has a higher median, reflecting overall better happiness levels compared to South Asia.
  - Outliers include Afghanistan, which is an extreme low outlier for the South Asian region, Yemen and Syria for the Middle East, while UAE is a high outlier because of its high economic performance.

Score Distribution by Region
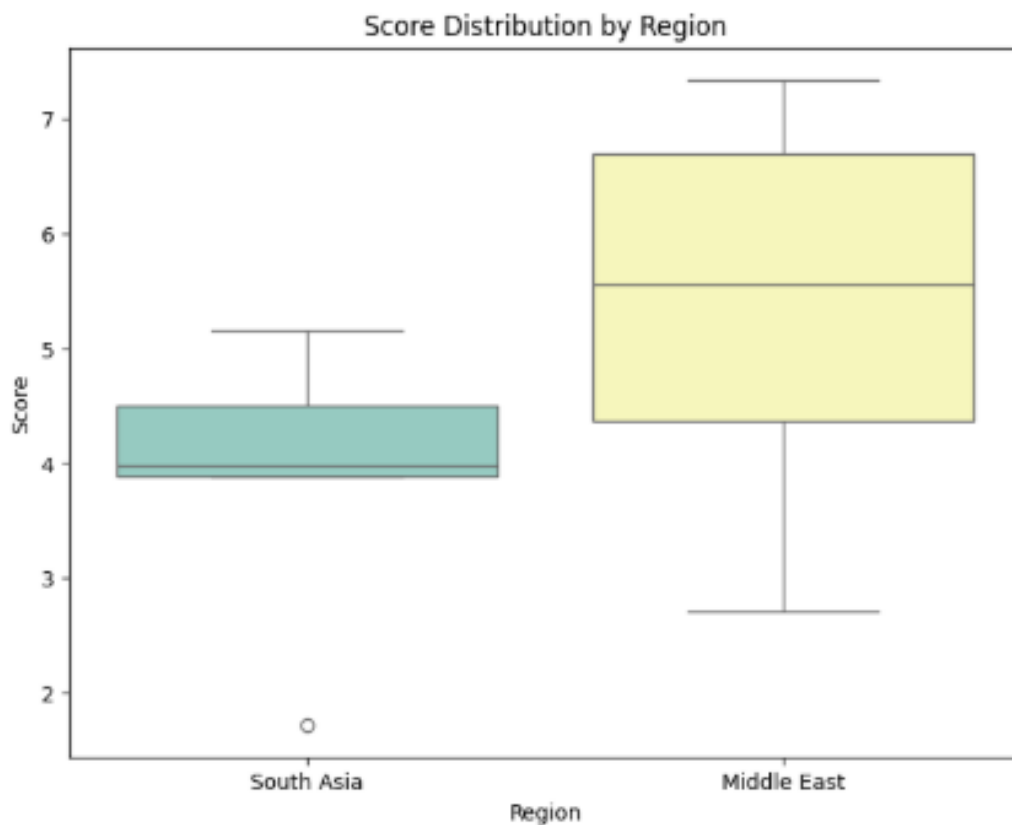
Conclusion

The report sums up the following:

- Regional inequalities, especially in GDP per Capita, between South Asia and the Middle East.
- The importance of autonomy, social support, and life expectancy in driving happiness.
- The impact of systemic inefficiencies and cultural factors on regional happiness scores.
- This increased understanding gives deeper insight into the tasks and what they are meant to portray.