



# Concepts and Technologies of AI

Statistical Interpretation and Exploratory Data Analysis

Student Id : 2408414  
Student Name : Shashank Pandey  
Section : L5CG20  
Module Leader : Siman Giri  
Tutor : Bibek Khanal  
Submitted on : 12-19-2024

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

The World Happiness Report is a yearly survey done to measure the happiness of residents in different countries of the world, depending on international surveys assessing key factors: GDP per capita, social support, healthy life expectancy, freedom to make life choices, generosity, and perceived corruption. This would therefore be a critical report for both the policymakers and the researchers because it helps in comprehending the contribution factors to happiness while devising ways in which the quality of life on Earth can be improved.

The main goals of the report are summarized into three critical tasks:

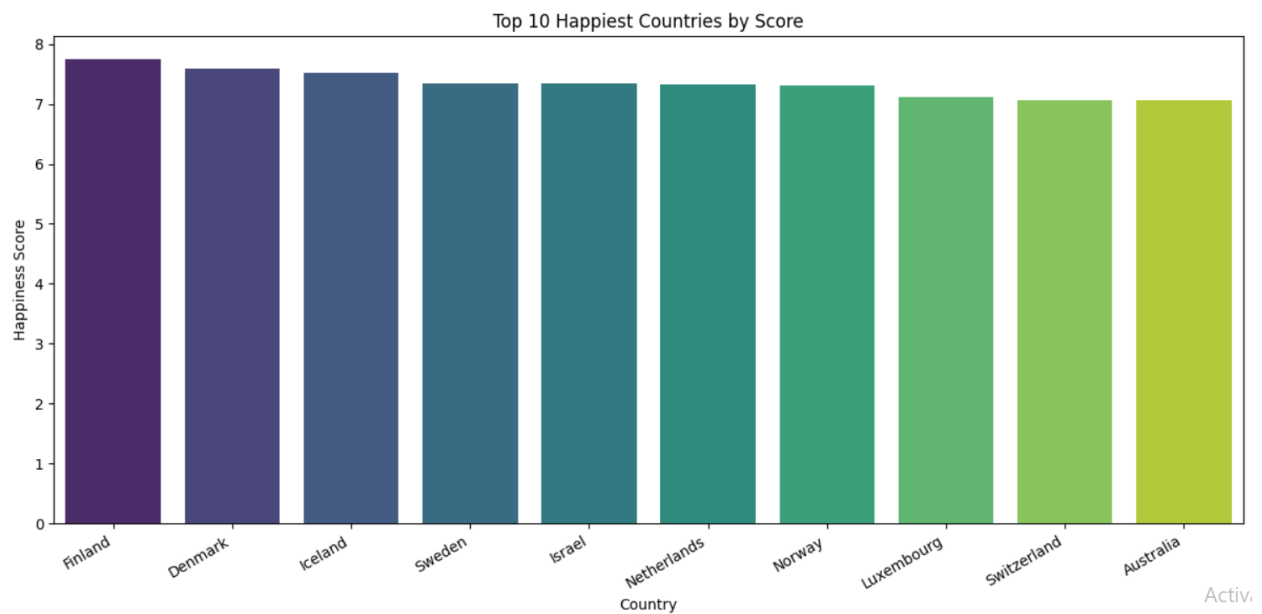
1. Data Exploration: Conducting an analysis of worldwide trends, patterns, and interrelations among significant variables within the dataset.
2. South Asia Analysis: Parsing the Happiness Scores and Crucial Determinants of Well-being in South Asian Countries.
3. South Asia and the Middle East: A comparison of happiness scores and explanatory factors, between South Asia and the Middle East to identify the differences, similarities and regional challenges.

By focusing on these tasks, this report tends to provide updated information about the trend of regional happiness, thus laying the foundation for further research in the development policy by focusing on these tasks.

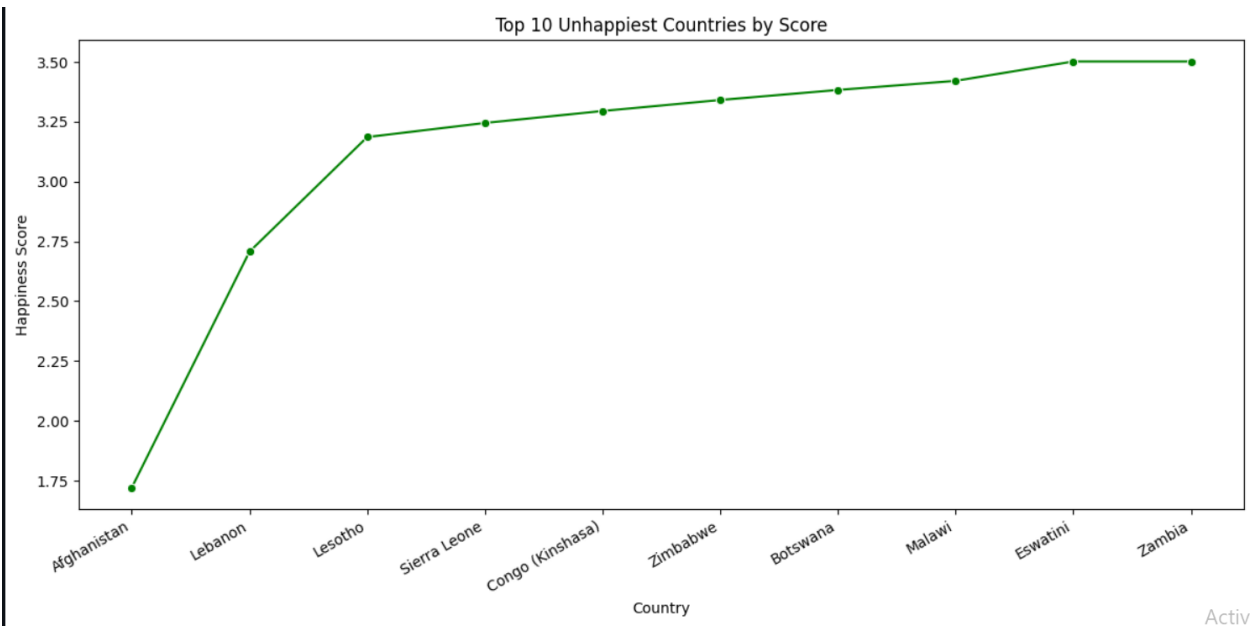
## 1. Data Exploration, Clarification and Visualization.

In the problem, the given EDA has been conducted on the World Happiness Report dataset; it summarizes the data by calculating mean, median, and standard deviation to bring countries that have the highest and lowest scores. It checks for missing values in the data, filters those scores above 7.5, and sorts according to the GDP per capita. A new “Happiness Category” column is created, categorizing scores as Low, Medium or High. The Visualizations include a bar plot for the happiest countries, a line plot for the unhappiest, a histogram for score distribution, and a scatter plot showing GDP vs Happiness.

### I. Bar plot for the happiest countries:



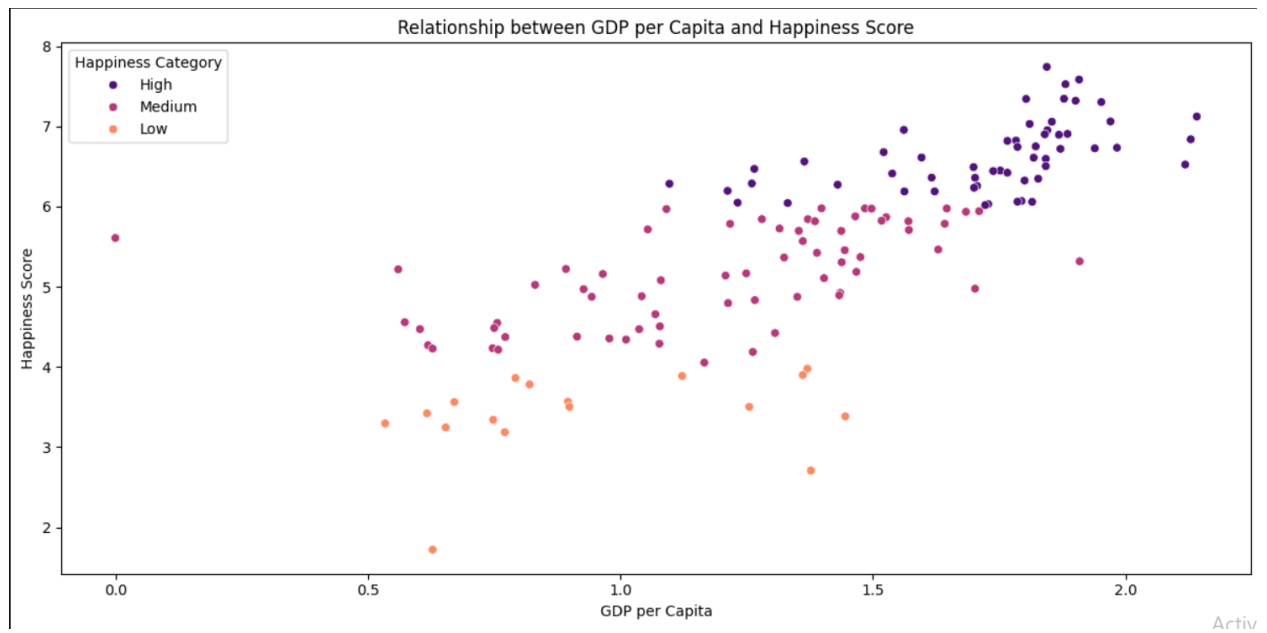
## II. Line plot for the unhappiest countries:



## III. Histogram plot for score distribution:



#### IV. Scatter plot showing GDP vs Happiness:

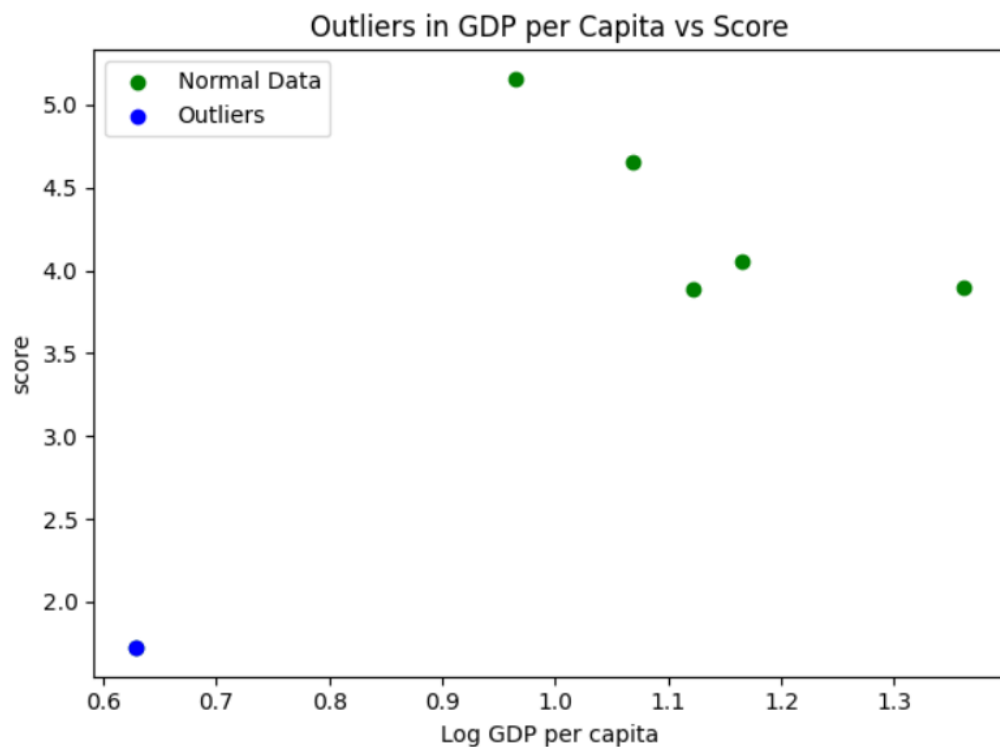


From the given visualization plot, we conclude that the higher the GDP per capita the higher will be the happiness scores.

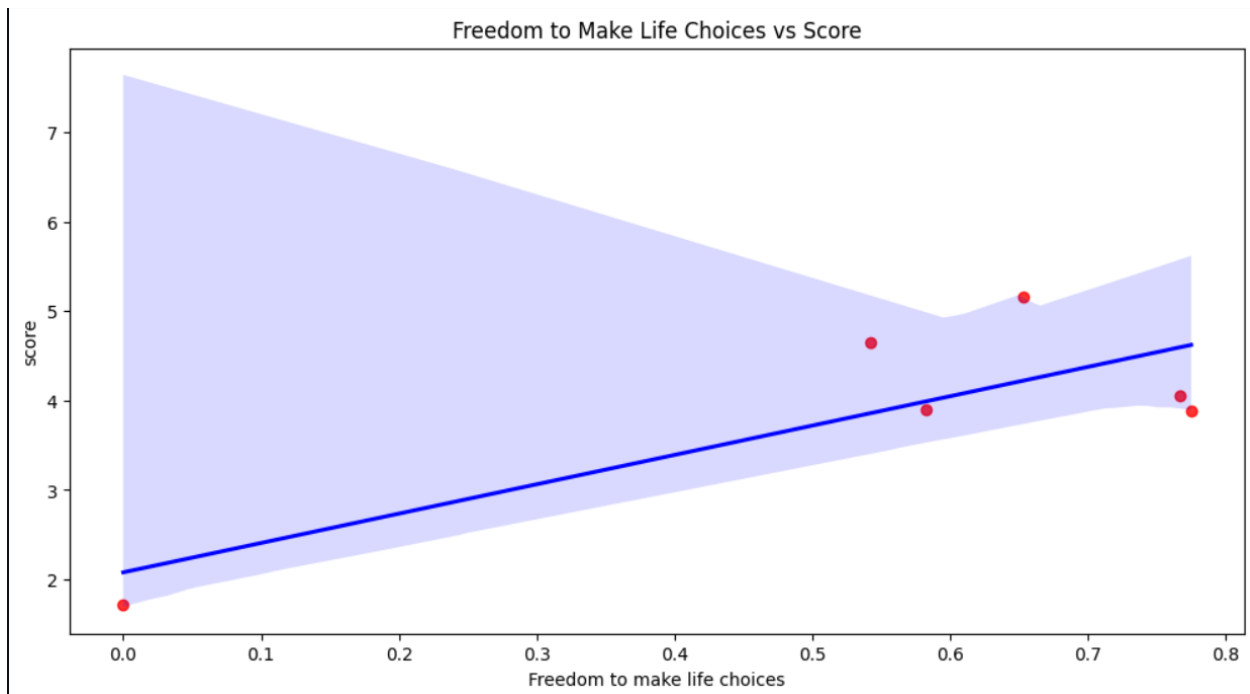
## 2. Advance Data Exploration Task:

The following study carries out some preliminary exploratory data investigation and analysis by exploring the dataset on the World Happiness Report, focusing on South Asian countries. First, filtering of data by country for South Asian nations is done, followed by the computation of a composite score calculated by multiplying Log GDP per capita, social support, and healthy life expectancy by their respective weight factors. Finally, countries are ranked according to the composite score and the top 5 are visualized as a bar plot. A scatter plot compares the original happiness scores with the composite scores. Using the  $1.5 * \text{IQR}$  rule for outlier detection and visualizing outliers of GDP per capita vs. Happiness Score: Pearson's correlation coefficients are computed between freedom to make life choices, generosity, and happiness scores, supported by corresponding visualizations. Gap analysis was performed for the countries featuring the largest positive and negative GDP-Score gaps, visualized in a bar chart.

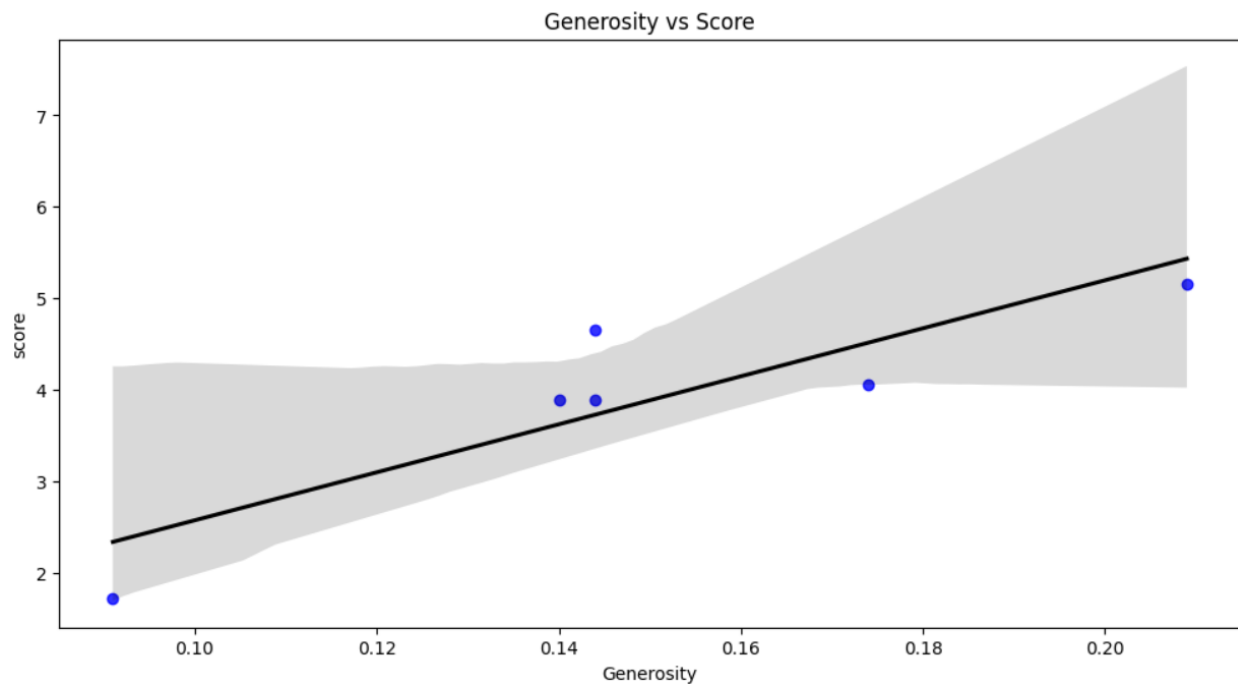
- i. A scatter plot comparing the outliers in GDP per capita with the score.



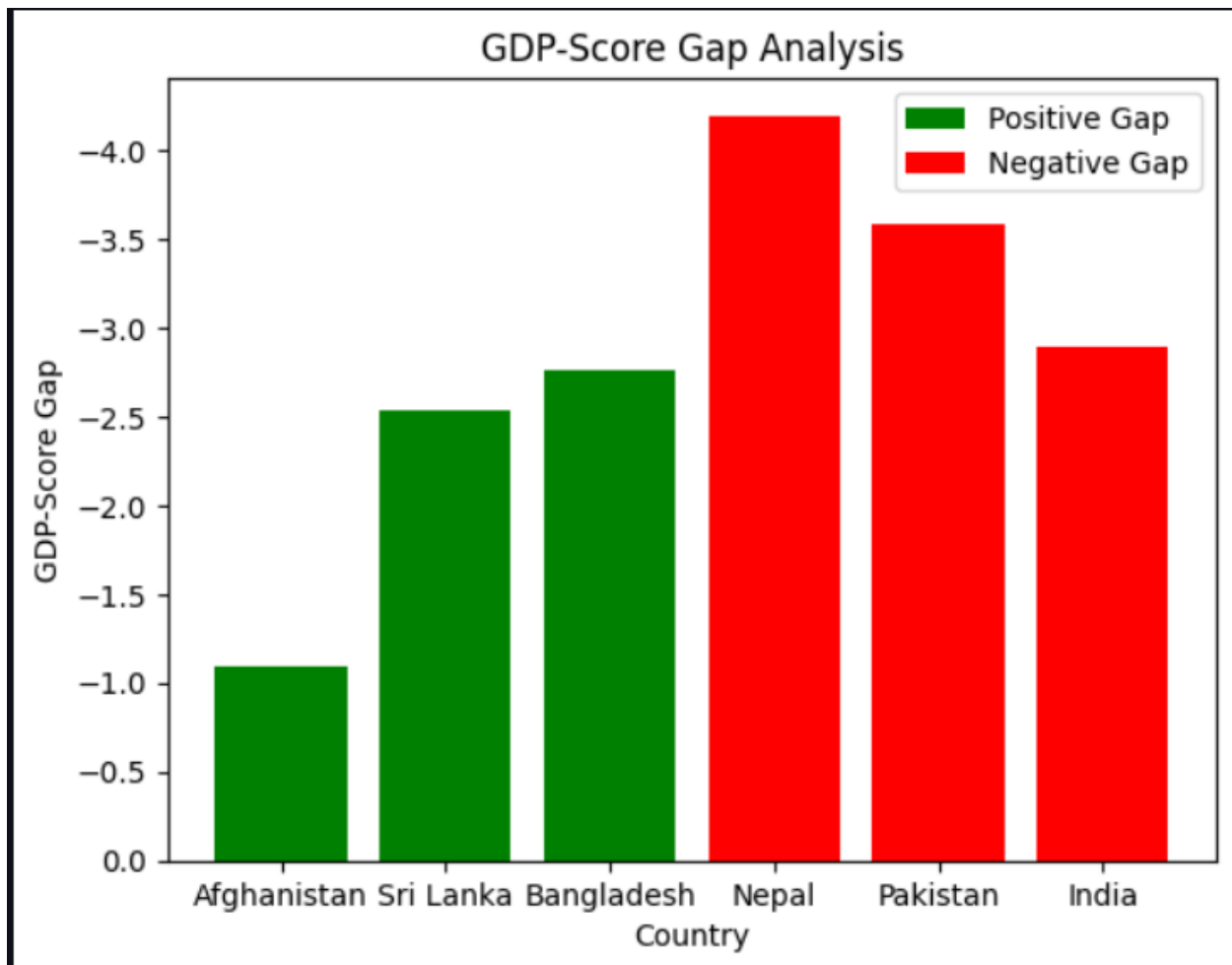
ii. Visualizing plot comparing Freedom to Make Life Choices vs Scores:



iii. Visualizing plot comparing Generosity vs Scores:



iv. Bar chart featuring the largest positive and negative GDP-Score:

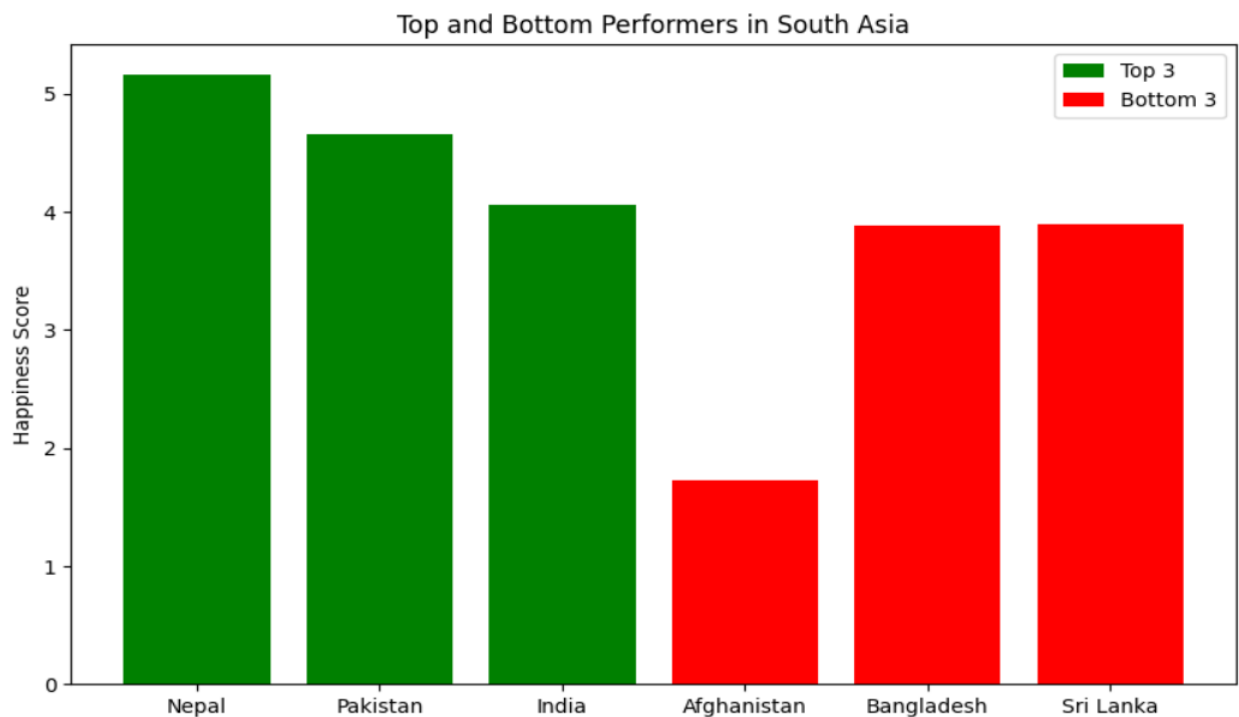




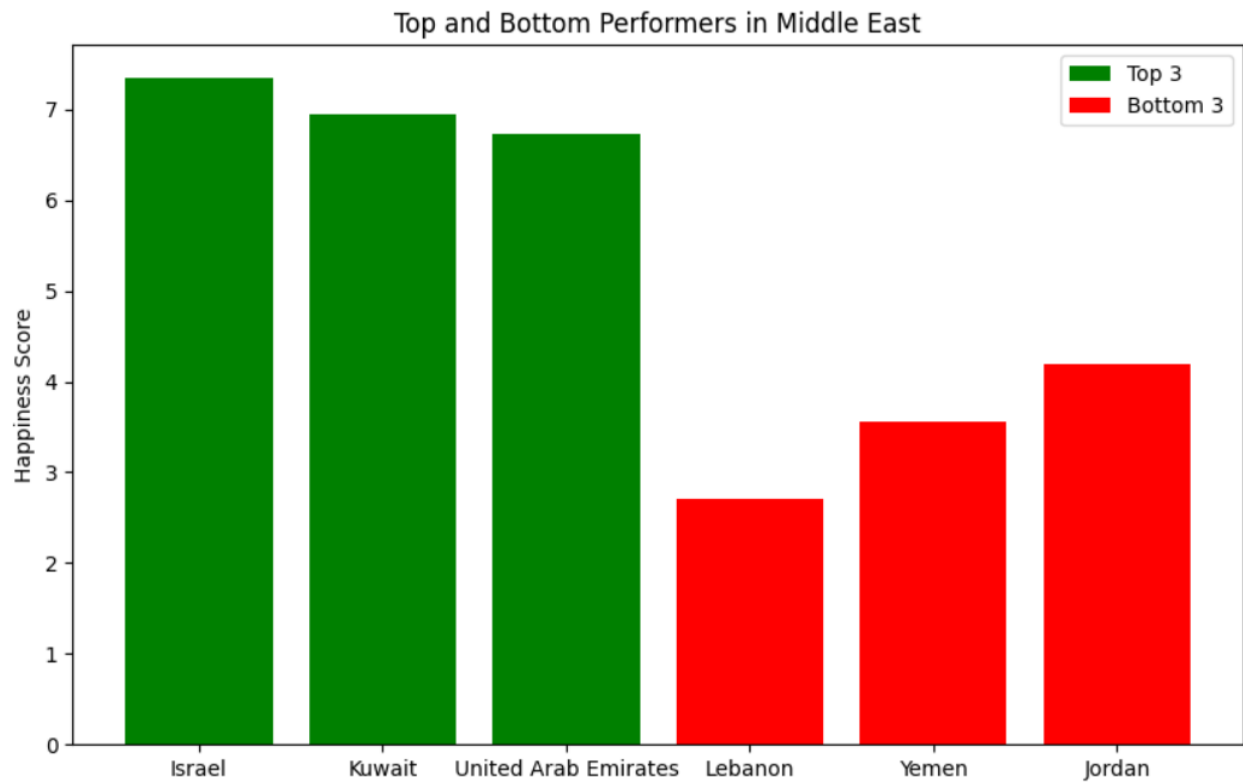
### 3. Comparative Analysis between South Asian and Middle Eastern Countries:

In this problem, a comparison between happiness and other metrics in South Asia and the Middle East is made using the World Happiness Report dataset. Data of each region was extracted, focusing on countries of South Asia and the Middle East. Descriptive statistics show that the average happiness score for the Middle East was 5.67 compared to 4.84 for South Asia. Further analysis revealed the top and lowest rankers in both areas. In South Asia, Sri Lanka, Maldives, and Nepal top the graphs, while at the Middle East, Israel, UAE, and Qatar are on the higher end. Log GDP per capita, Social Support, and Healthy Life Expectancy are the key metrics. GDP-wise, the Middle East was good; whereas social support and life expectancy are better in South Asia. Outlier detection showed very high differences in happiness scores from Log GDP per capita; hence, both regions have significant outliers. In correlation analysis, it could be seen that happiness is positively related to such parameters as freedom and generosity, but with different strengths.

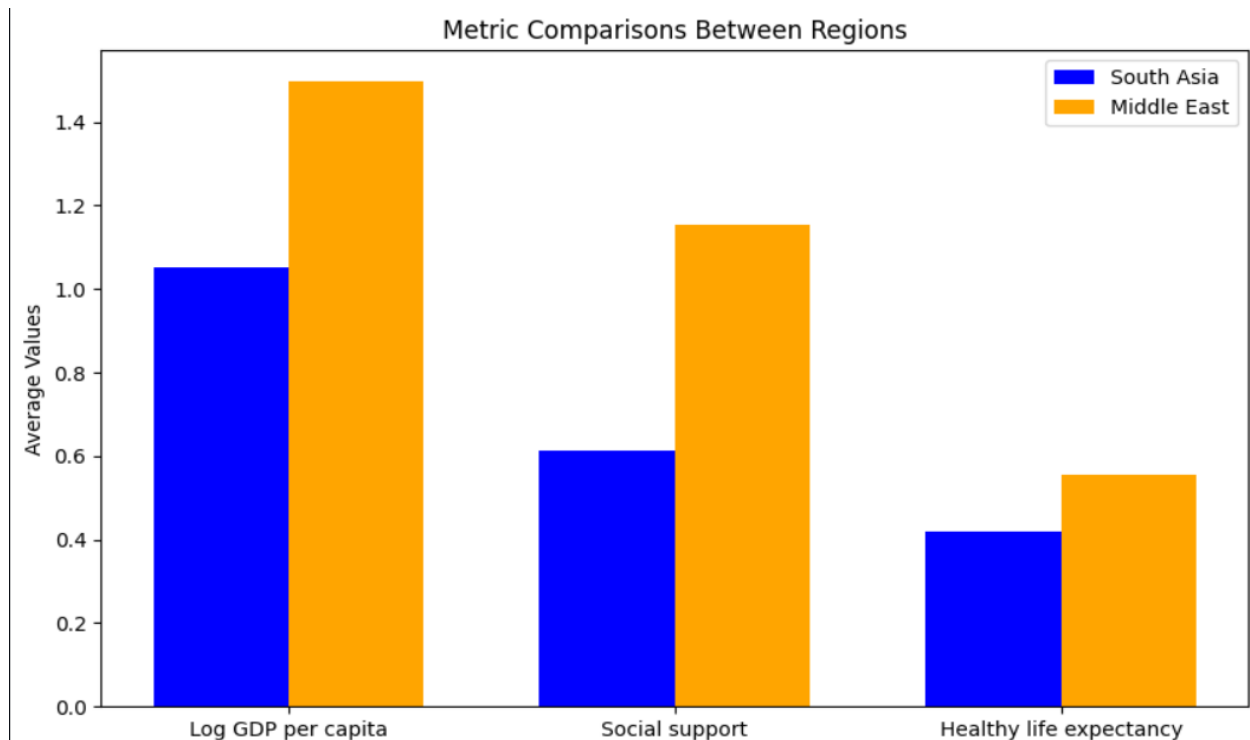
i. Bar chart which shows Top and Bottom Performers in South Asia:



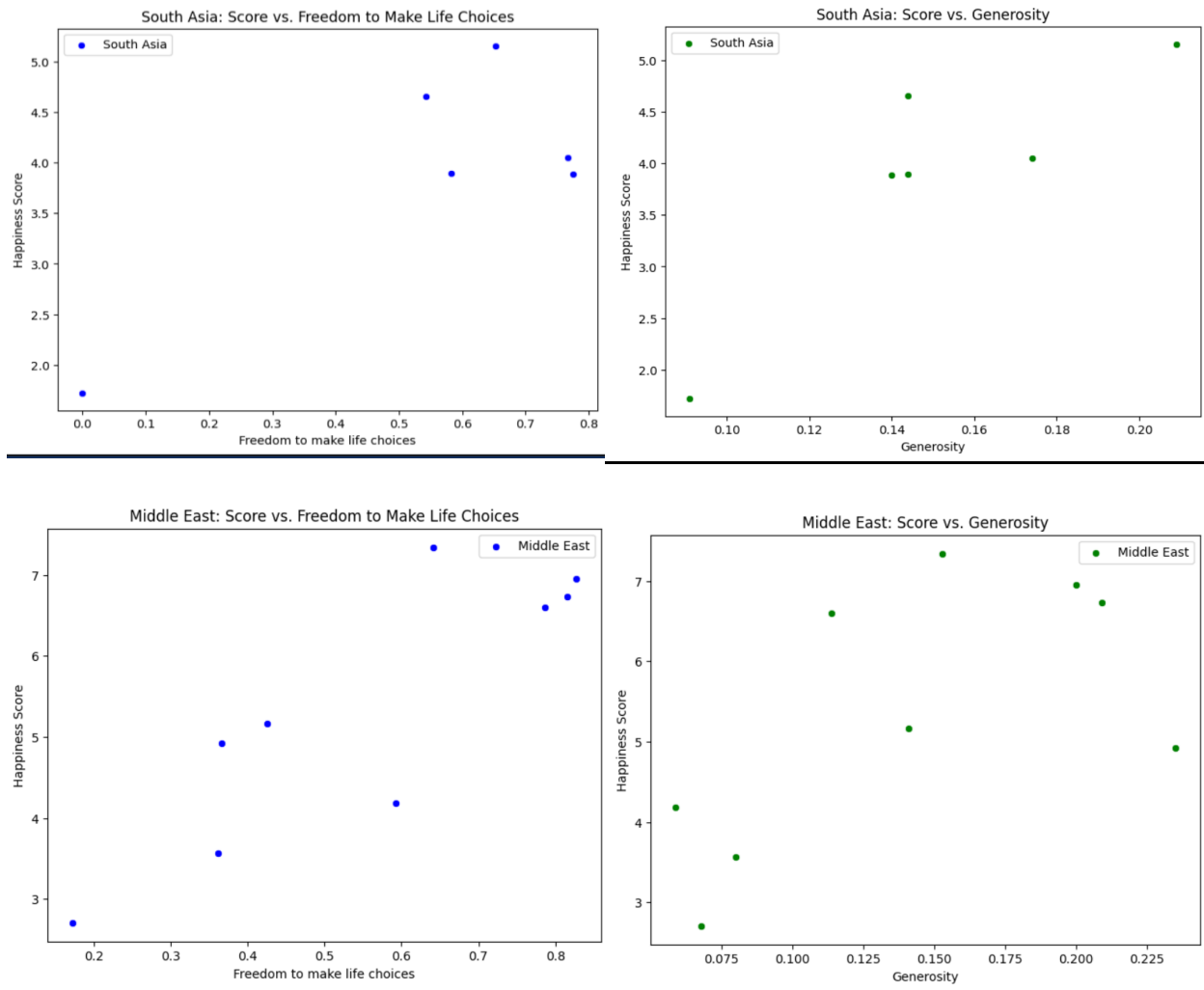
ii. Bar chart which shows Top and Bottom Performers in Middle East:



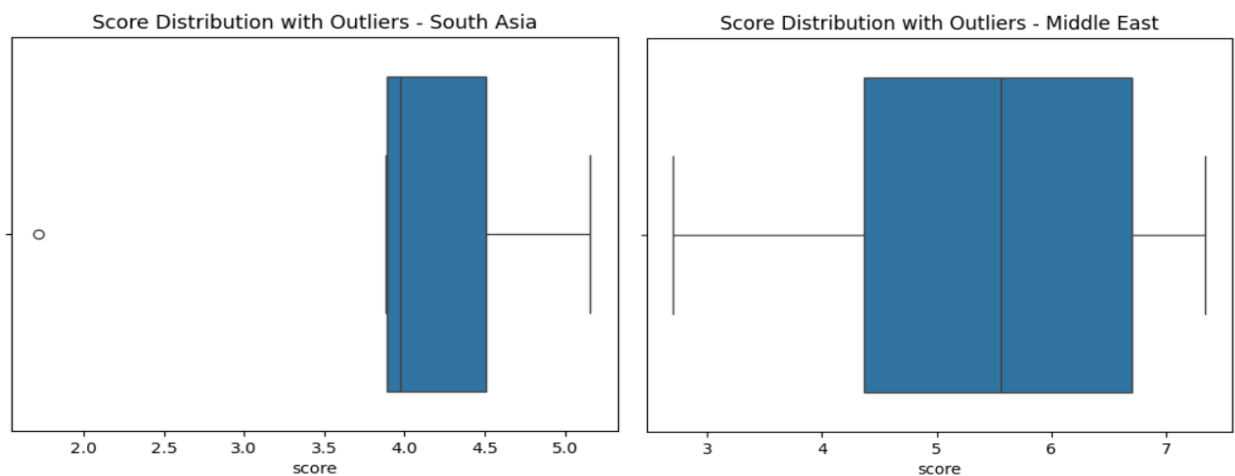
iii. Bar chart which shows metric comparisons between two regions:

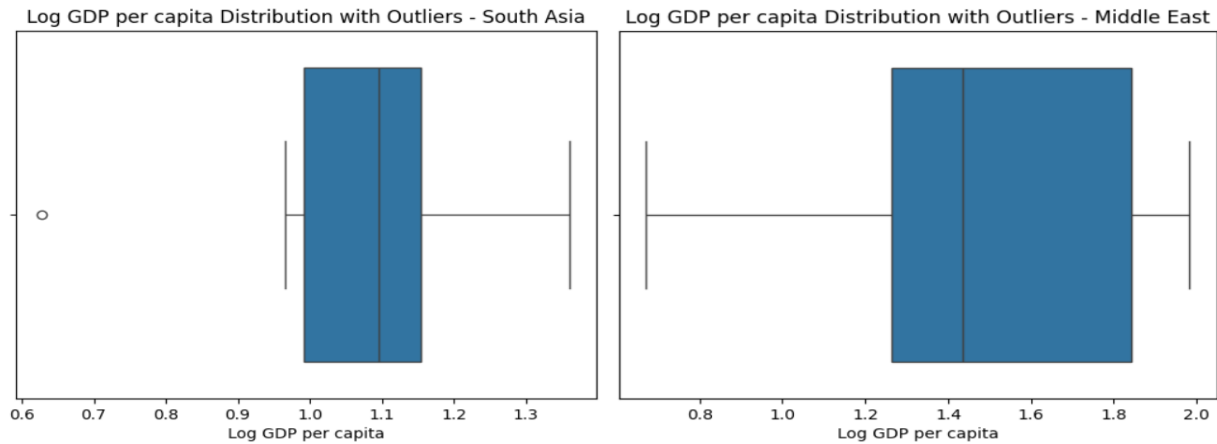


iv. Scatter plots for South Asia and Middle East:

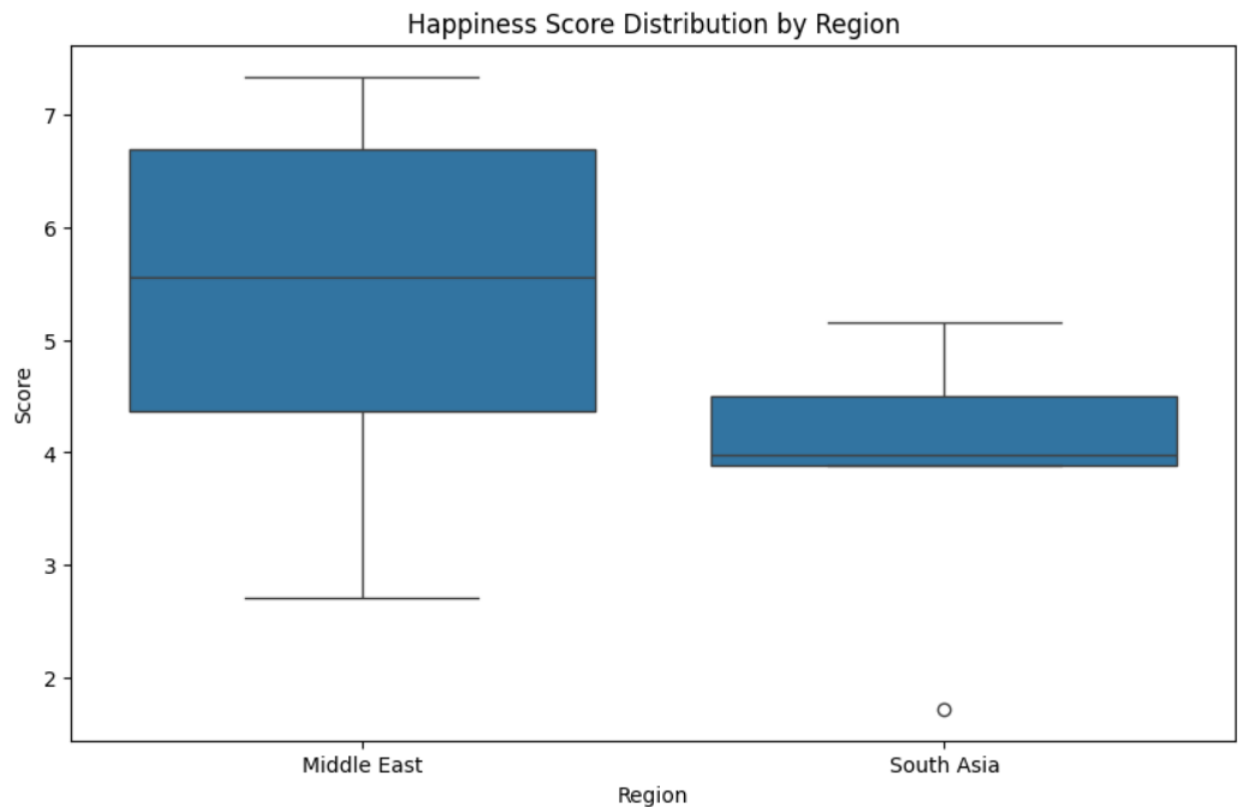


v. Outliers detection of South Asia and Middle East regions:





vi. Happiness Score Distribution by Regions:



So, in a way, this means that although the Middle East poses a higher average happiness score, South Asia had much more variation, hence much more inequality in happiness across the region.

## Conclusion:

The test run on the dataset generated from the World Happiness Report provides some interesting conclusions regarding happiness indices on a global and regional scale. Scores in happiness have been evaluated concerning GDP per capita, social support, and life expectancy, shedding light on significant trends and differences in the importance of these variables. Generally, countries with higher scores boast big GDP, social networks, and healthy populations as can be seen in the rankings of Israel, UAE, and Qatar in the Middle East and Sri Lanka, Maldives, and Nepal in South Asia.

Outlier detection highlights that happiness differences may be huge and usually linked to GDP and social inequalities. Analysis of correlation lends credence to freedom and generosity, impacting happiness positively but not consistently, indicating that culture and region play a significant role in defining these relations.

That is, social support and life expectancy, for instance, where the South Asia part has a more qualified rating than the Middle East shows the multidimensional nature of happiness. The study successfully finds differences between GDP scores and happiness, reflecting the deeper complexity of welfare beyond economic measurement.

This detailed analysis, which is supported by various visualizations, creates a clear understanding of the happiness determinants and serves as a basis for further specific improvements in the happiness initiatives that are being developed around the globe and regionally.