Concepts and Technologies of AI 5CS037

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

Uni ID: 2408485

Submitted by: Aryush Khatri

Lecturer: Ms. Sunita Parajuli

Tutor: Mr. Siman Giri

Submitted on: 2024/12/20

Table of Contents

Introduction:

This World Happiness Report is prepared using different data evaluating happiness levels of different countries across the globe. The data were used from the dataset provided with the assignment. I used different libraries like matplotlib, pandas, NumPy, seaborn etc. to analyze the dataset. This report will thoroughly show how happiness differs from country to country with some calculations and different statical figures.

Problem - 1: Getting Started with Data Exploration

Data Exploration and Understanding:

I have loaded the dataset using read_csv () then printed different attributes of dataset using different methods from their respective libraries that I import.

Some basic statistics like mean, median, standard deviation of happiness score from dataset were calculated. The calculation showed Finland was the happiest country and lowest happy country was Afghanistan. Few null values were seen in some most columns.

Rows were filtered for some higher happiest countries i.e. score>7.5 then that filtered dataset were sorted in descending order using sort_values () method. Then according to question, new column "Happiness Category" was added where values (high, medium, low) were inserted following conditions that were provided.
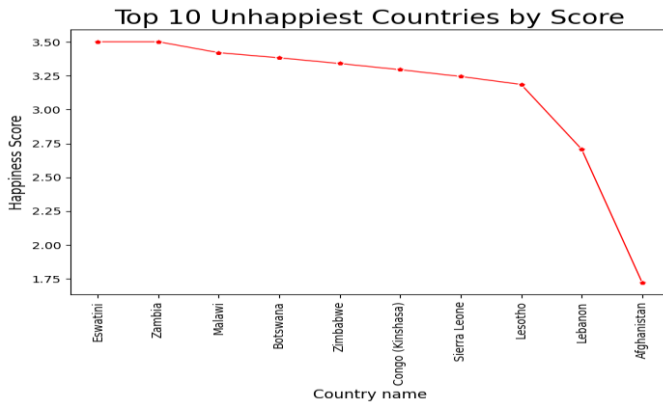
Data Visualizations:

Figure 1.1: Line chart of top 10 Unhappiest Countries

The line chart shows a downward trend with happiness score decreasing as we go to Afghanistan. The unhappiest country is Afghanistan with score 1.721 and least unhappiest countries are Eswatini and Zambia with scores 3.502.

The figure indicates distribution of Happiness score. The histogram is left skewed, and the peak is at happiness score 6 decreasing the frequencies on either side. Most of the frequencies are also around 6.
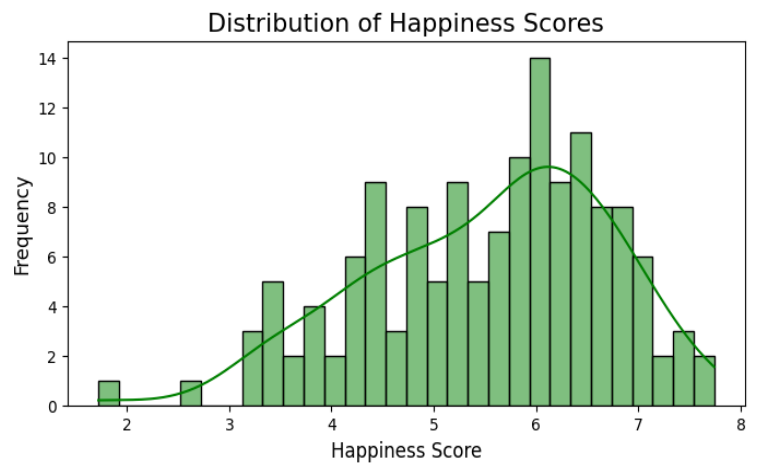

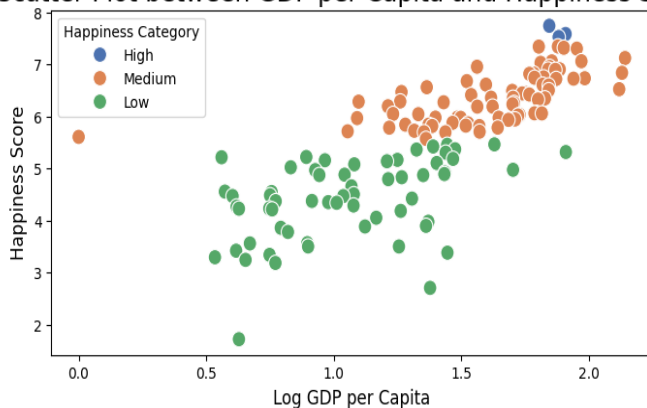*Figure 1.2: Histogram for Happiness Score*

Figure 1.3: Scatter plot between GDP and Score

The figure shows the relationship between GDP per Capita and Happiness Score. They have positive correlation because as one increases another also rises. However, there are few outliers like countries with high score but very low GDP and vice versa.

Problem – 2: South Asia Dataset

Data Exploration and Understanding:

Here I defined new list of South Asian Countries which is filtered accordingly and was saved as csv file. Using the previous dataset we saved, a new column "Composite Score" was created using formula that included GDP per Capita, Social Support, and Healthy Life Expectancy. Now, based on Happiness Score and GDP per Capita I identified outlier using IQR method. According to calculations only Afghanistan is identified as outlier in South Asia. Then I found out correlation of freedom to make life choices and generosity with happiness score which resulted in 0.80 and 0.87 respectively. Then I analyzed the gap between GDP per Capita and Happiness score by subtracting them.

According to calculations only Afghanistan is identified as outlier in South Asia. Correlation shows generosity has slightly more corelated to happiness so fostering it might increase happiness score overall. The analysis shows that Nepal has most negative GDP score gap.
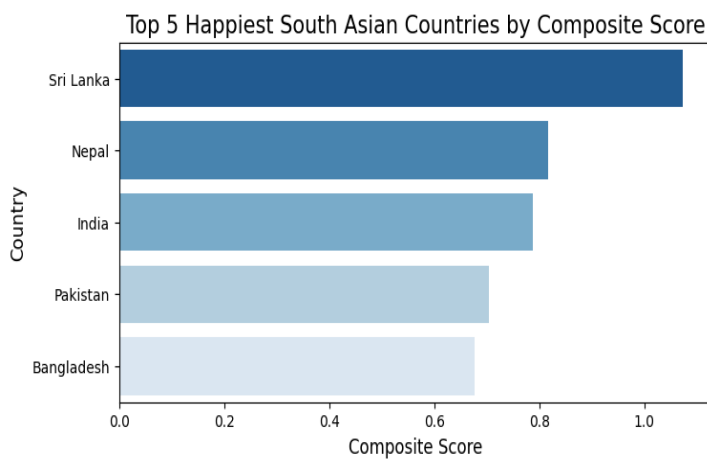
Data Visualizations

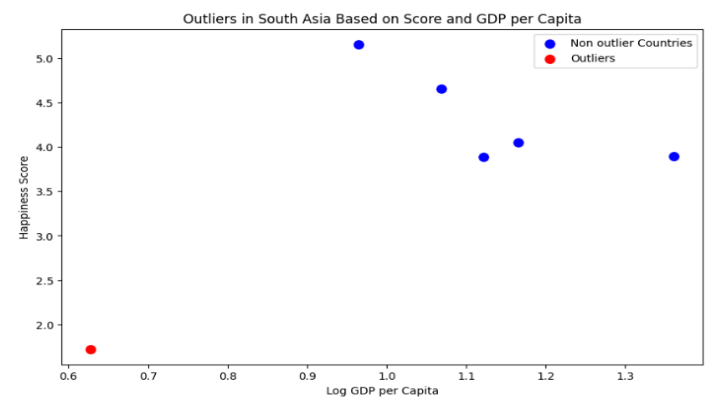Figure 2.1: Top 5 Happiest South Asian Countries

Figure 2.2: Outliers based on Score and GDP per Capita

This horizontal bar graph shows the top 5 Happiest South Asian Countries based on Composite Score. Sri Lanka is the Happiest south Asian country with 1.0739 composite score followed by Nepal and Bangladesh is the least happy country with score 0.6774.

This figure shows the outlier in the south Asian countries which was found using IQR method and the outlier country is Afghanistan as it is far from other countries according to score and GDP.

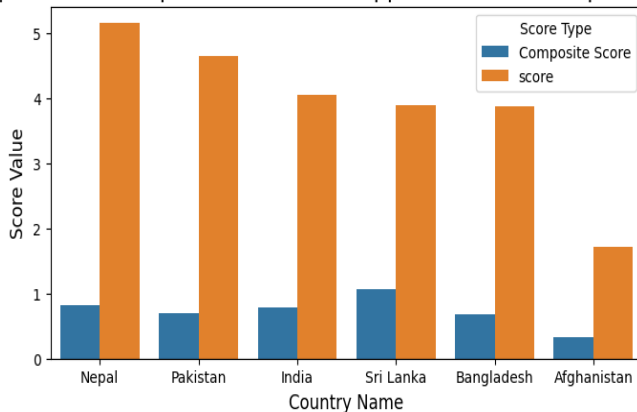Comparison of Composite Score and Happiness Score for Top 10 Countries

Figure 2.3: Composite score vs Happiness Score

The bar graph shows a comparison of composite score and happiness score. The happiness score is much higher than the composite score in every country. Composite score is consistent as it depends upon many factors. The difference between composite and happiness score is largest in Nepal.
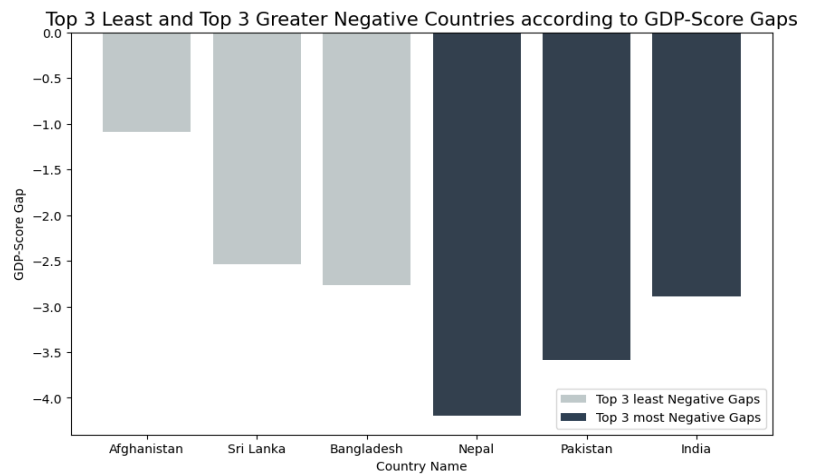
The bar chart shows the top 3 countries with the least Negative and most negative GDP and Score gaps. Some countries have less difference while some countries have significantly higher gaps. Nepal seems to have most negative gaps of -4.193 followed by Pakistan and Afghanistan seems to have least

Figure 2.4: Countries with GDP Score Gap

Problem 3: South Asia vs Middle East Dataset

Data Exploration and Understanding:

In this part I defined a DataFrame for middle east countries and filtered it. Then calculated mean and standard deviation of their happiness score and compared it with South Asian countries. Similarly, I then calculated the top 3 and bottom 3 countries from each region and compared their scores. Coefficient of Variation of both regions is calculated using formula (standard deviation/mean). Calculated correlation of score with other metrics like freedom to make life Choices and Generosity within each region. After that, for identifying outlier in both regions based on Score and GDP per capita I used IQR method.

According to mean, it was observed Middle East has higher happiness scores on average. Nepal is the happiest country with score 5.158 and Afghanistan is least happy country with score 1.721 in South Asia. Similarly, Israel is the happiest country with score 7.341 and Lebanon is least happy country with score 2.707 in Middle East. Based on Coefficient of Variation South Asia has slightly greater variability with cv 0.30. No outliers were observed in Middle East while Afghanistan is the only outlier in South Asia
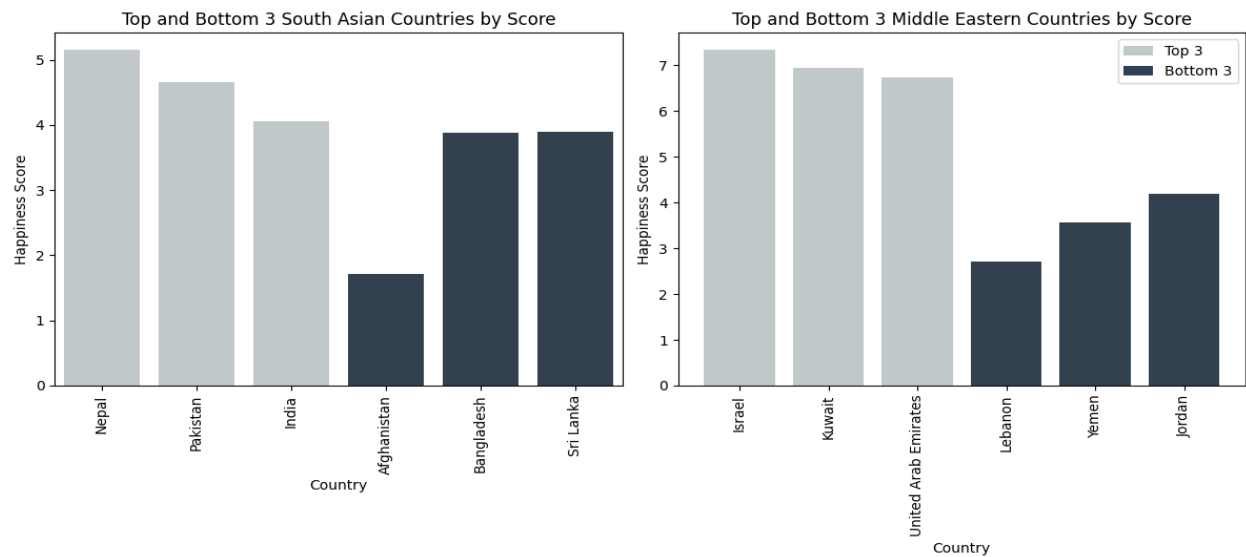
Data Visualization

Figure 3.1: South Asian vs Middle East Countries Happiness score

The first bar chart depicts the top 3 and bottom 3 happy countries in South Asia and the second one shows of Middle East. In overall comparison, these bar chart shows that the Middle East countries are much happier than South Asian countries.
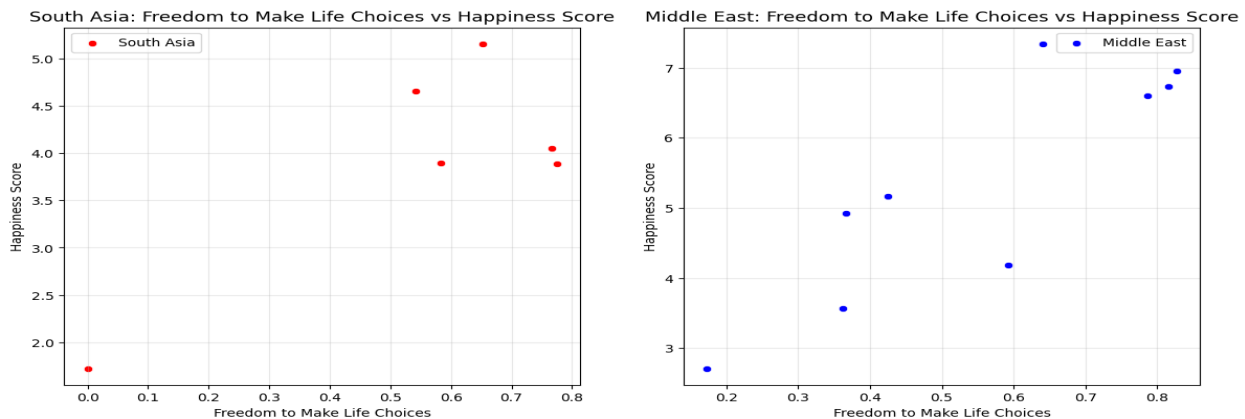

Figure 3.2: Freedom to make choices vs happiness score in South Asia and Middle East

The above figures show the relationship between freedom to make life choices and Happiness Score in South Asia and Middle East. In both regions the scattered plots between them show positive correlation as when freedom to make life choices increases Happiness Score also increases and vice versa. South Asian countries are generally happier even they have less freedom to make life choices than Middle East

This scatter plot detects outliers in South Asia and Middle East for the relation between Score and GDP per Capita. This figure shows that countries in south Asia have very low GDP per Capita and moderate happiness score with one outlier having very low GDP and score. However, middle east has no outlier and have widely spread happiness score and GDP per Capita
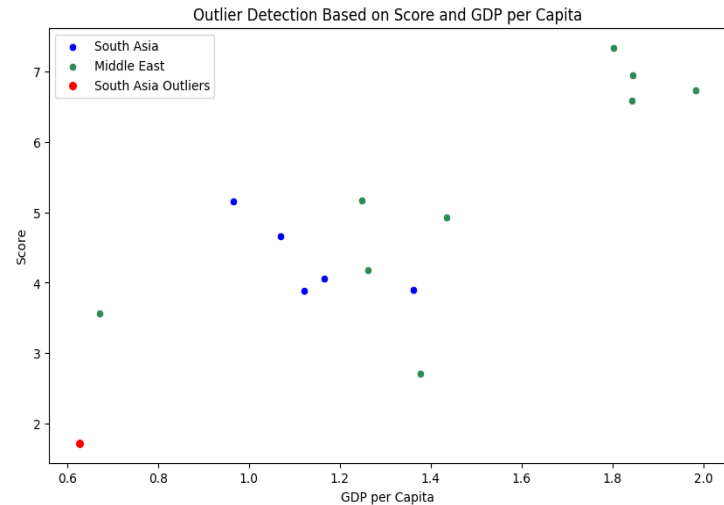

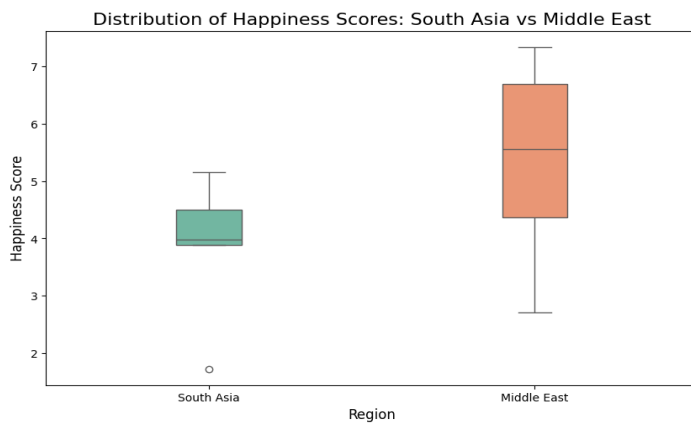*Figure 3.3: Outlier detection based on score and GDP*

Figure 3.4: Distribution of happiness score in both region

The boxplot shows the distribution of Happiness Score in South Asia and Middle East. The median happiness score for South Asia is around 4. The South Asian plot shows scores are tightly clustered with smaller range with a outlier below 2 shows a country with very low happiness score. The Middle East plot shows a widely spread happiness score with a median around 5 and has no outlier.

Conclusion:

This report analyzed happiness across countries, focusing on South Asia and the Middle East. Finland was the happiest country globally, while Afghanistan ranked the lowest and was identified as an outlier in South Asia due to its low GDP and score. South Asia showed moderate happiness levels with Sri Lanka leading, while the Middle East had higher average scores and more variability. Generosity (correlation: 0.87) had a stronger impact on happiness in South Asia than freedom of choice (0.80). Boxplots revealed tightly clustered scores in South Asia with one outlier, whereas the Middle East had a wider range and no outliers. Overall, the Middle East is happier on average, but fostering generosity could improve happiness in South Asia.