HERALD COLLEGE
KATHMANDU

5CS037 - Concepts and Technologies of AI

Individual coursework

Assessment – Statisitical Interpretation and Exploratory Data Analysis

Report Title: Analysis of the World Happiness

Report: Exploring South Asia and Middle East

Perspectives.

Autumn2024

Student Name: Mohammad Rashid Siddiqui

Student ID: 2413748

Module Leader: Mr. Siman Giri

Tutor: Ms. Durga Pokharel

Submission Date: 20th Dec, Friday, 2024
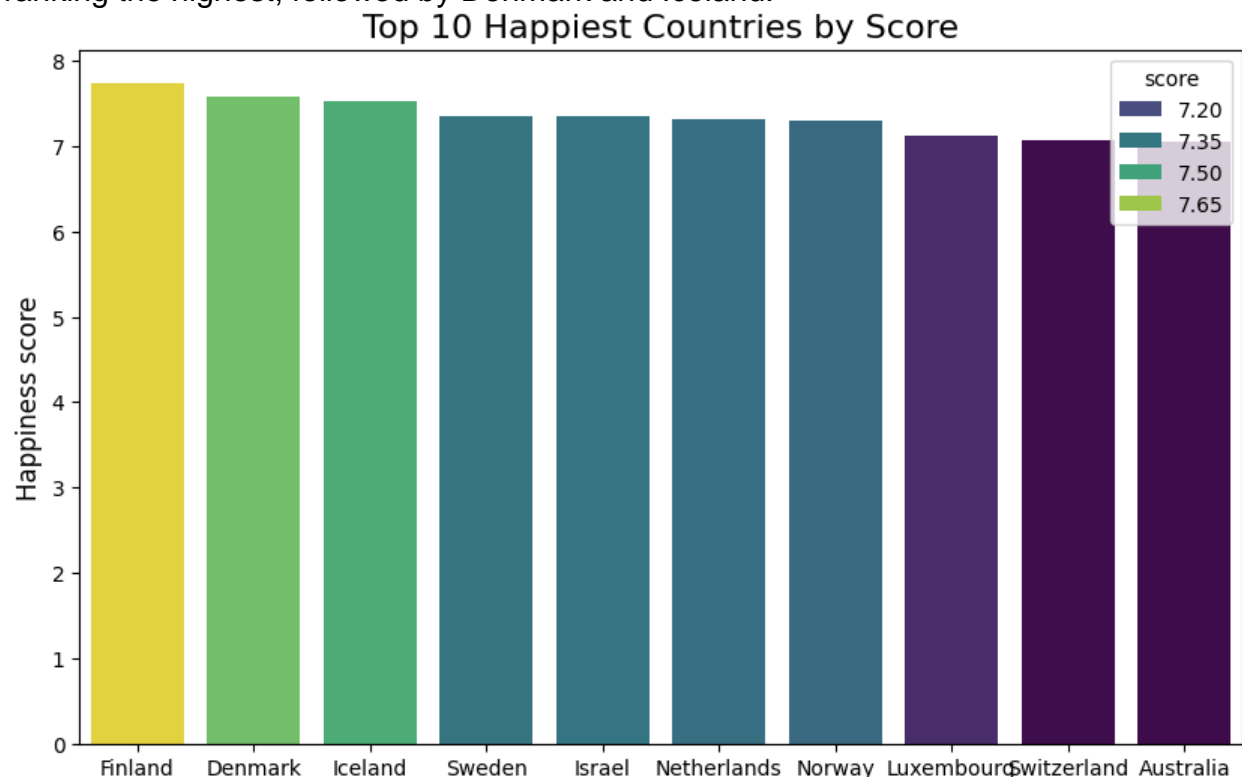
Acknowledgements

Introduction

The World Happiness Report is a key global study that measures happiness levels in different countries, providing valuable insights into well-being. It helps policymakers, researchers, and the public understand the factors that influence happiness. The report's dataset includes happiness scores and key factors like social support, life expectancy, economic performance, freedom, lack of corruption, and generosity. This study focuses on three main goals: first, to identify global trends and patterns in happiness; second, to explore what affects happiness and the scores in South Asia; and third, to compare happiness levels and differences between South Asia and the Middle East, highlighting cultural and regional contrasts. This analysis aims to deepen our understanding of happiness from both regional and comparative perspectives.
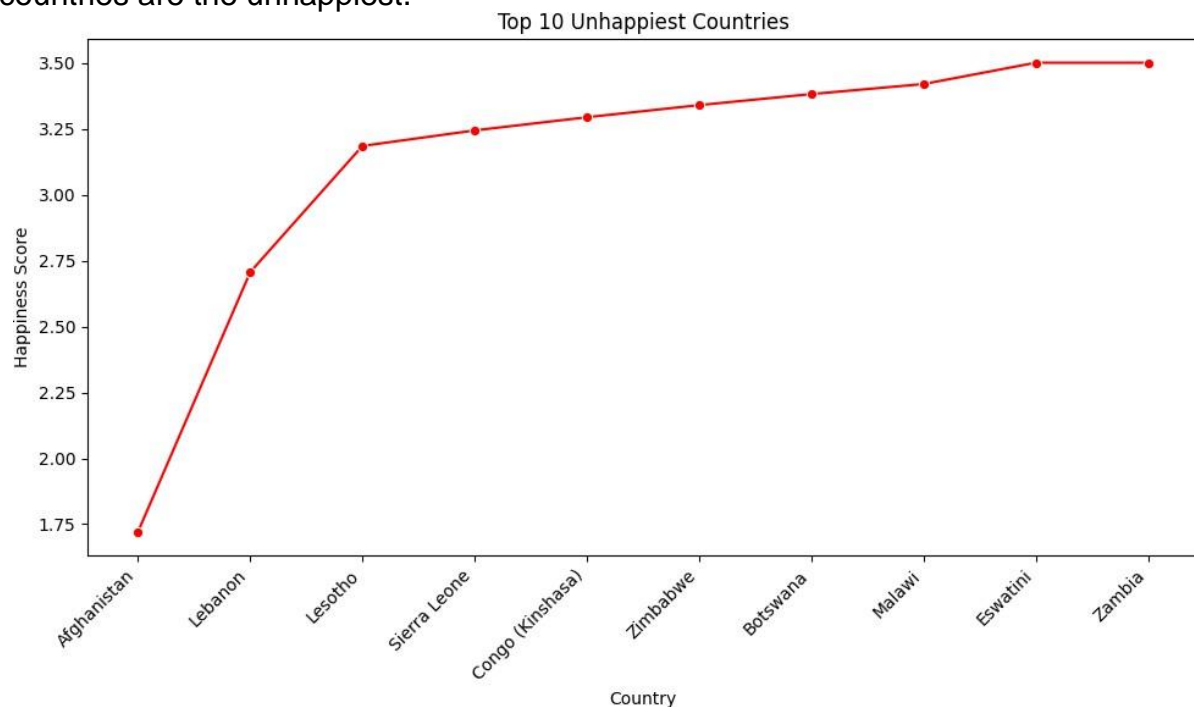
Objectives

The goals of this report are:

1. To study the World Happiness Report dataset and identify key patterns and trends.
2. To examine happiness scores and the factors influencing happiness in South Asia.
3. To compare happiness levels, metrics, and unique differences between South Asia and the Middle East.
4. To highlight the geographical differences in happiness between these two regions.

The analysis of basic statistics provides important insights into happiness scores. The average happiness score is 5.53, with a median of 5.79 and a standard deviation of 1.17. Finland has the highest happiness score, while Afghanistan has the lowest, according to the World Health Organization. The dataset has no missing values. When sorting the data, the top 10 happiest countries based on their happiness scores include Finland, Denmark, and Iceland. To enhance the classification of countries, a new "Happiness Category" was introduced. Countries with scores below 4 were labeled "Low," those with scores between 4 and 6 were categorized as "Moderate," and countries with higher scores were classified as "High." A bar chart was created to display the top 10 happiest countries based on their happiness scores, with Finland ranking the highest, followed by Denmark and Iceland.

Top 10 Happiest Countries by Score

The countries with the lowest happiness scores were listed. Afghanistan was the least happy, followed by Lebanon and Lesotho. This chart makes it easy to see which countries are the unhappiest.
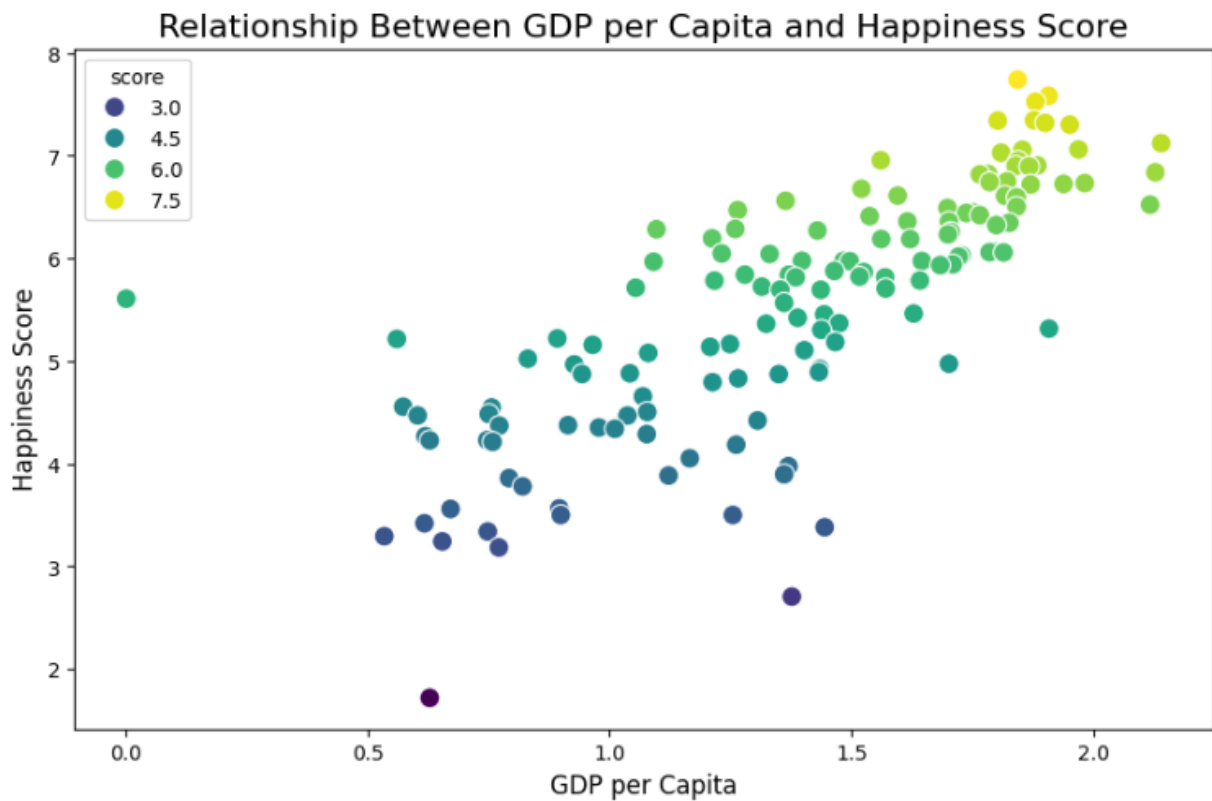

Top 10 Unhappiest Countries

The histogram shows that the highest happiness score is 17.5, and the lowest is 0.2. This visualization clearly displays the range of happiness scores in the dataset.

Distribution of Happiness Scores

Scatter Plot:

This scatter plot displays the relationship between Happiness Score and GDP per capita.

Problem - 2 - Some Advance Data Exploration Task:

Task-1

1. Preparing the South Asian Dataset

Nepal, India, Sri Lanka, Pakistan, Afghanistan, Bhutan, Bangladesh and Maldives was saved
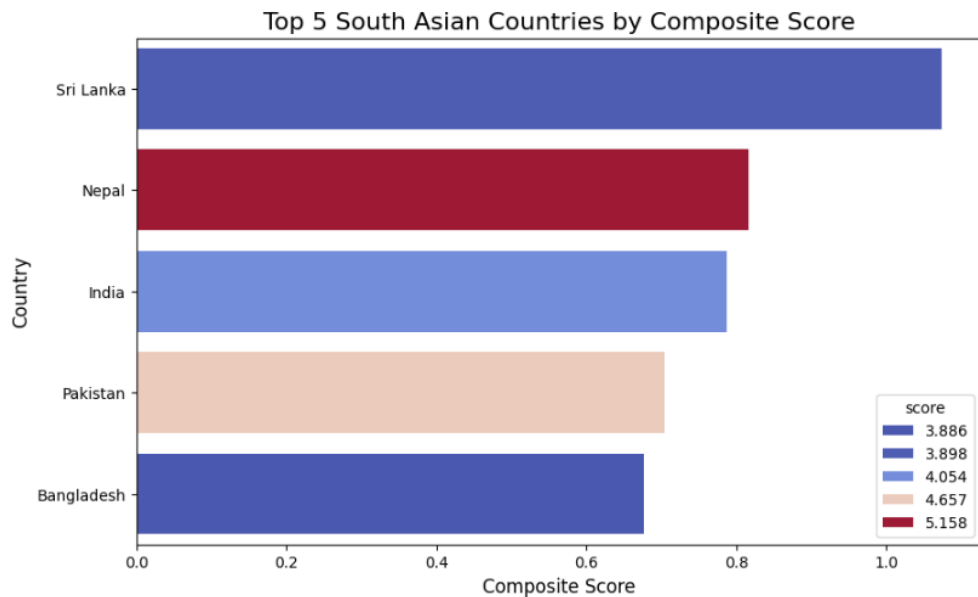
Task-2

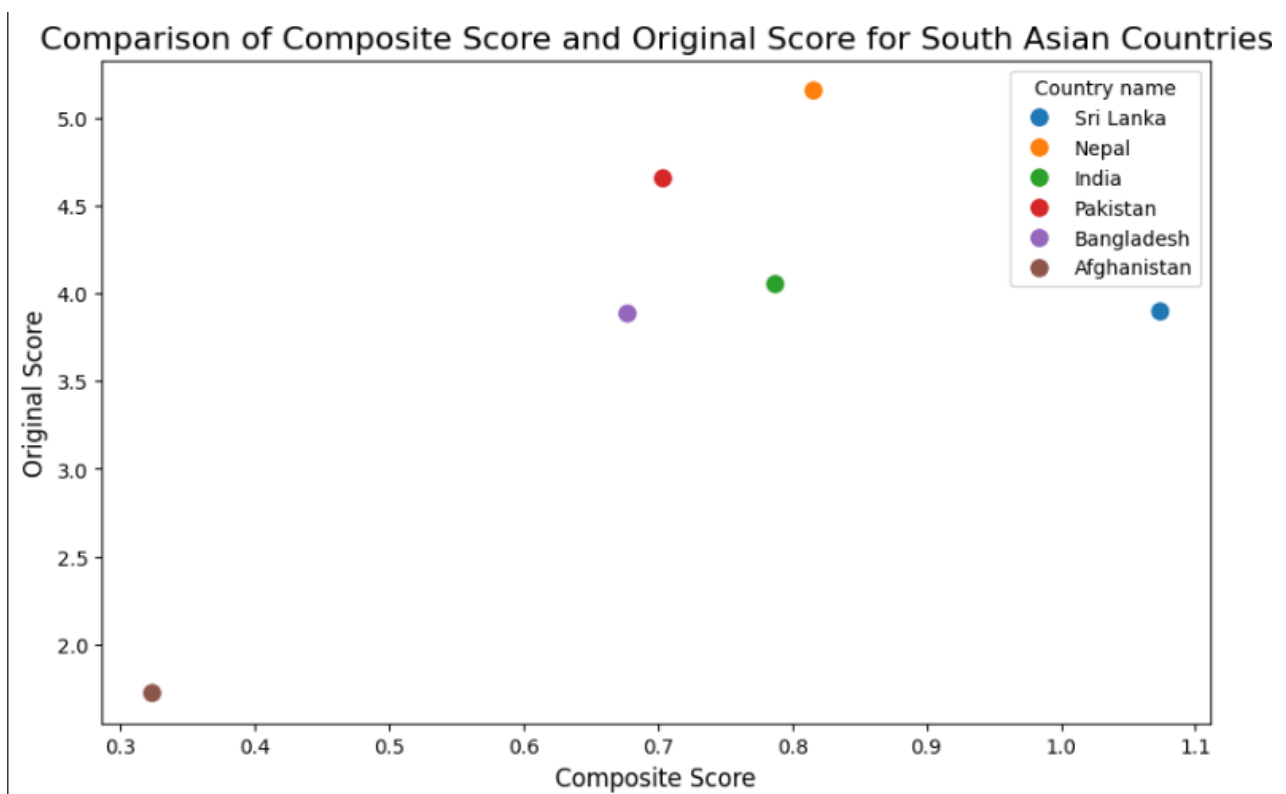1. New column was created named composite score by using below formula

Composite Score = 0.40 ×GDP per Capita + 0.30 ×Social Support + 0.30 ×Healthy Life Expectancy

2. The countries were sorted by their composite score, from highest to lowest.

3. A graph with horizontal bars is used to compare the composite scores of five countries..

4. The graph shows that the composite score and the original score don't match up.
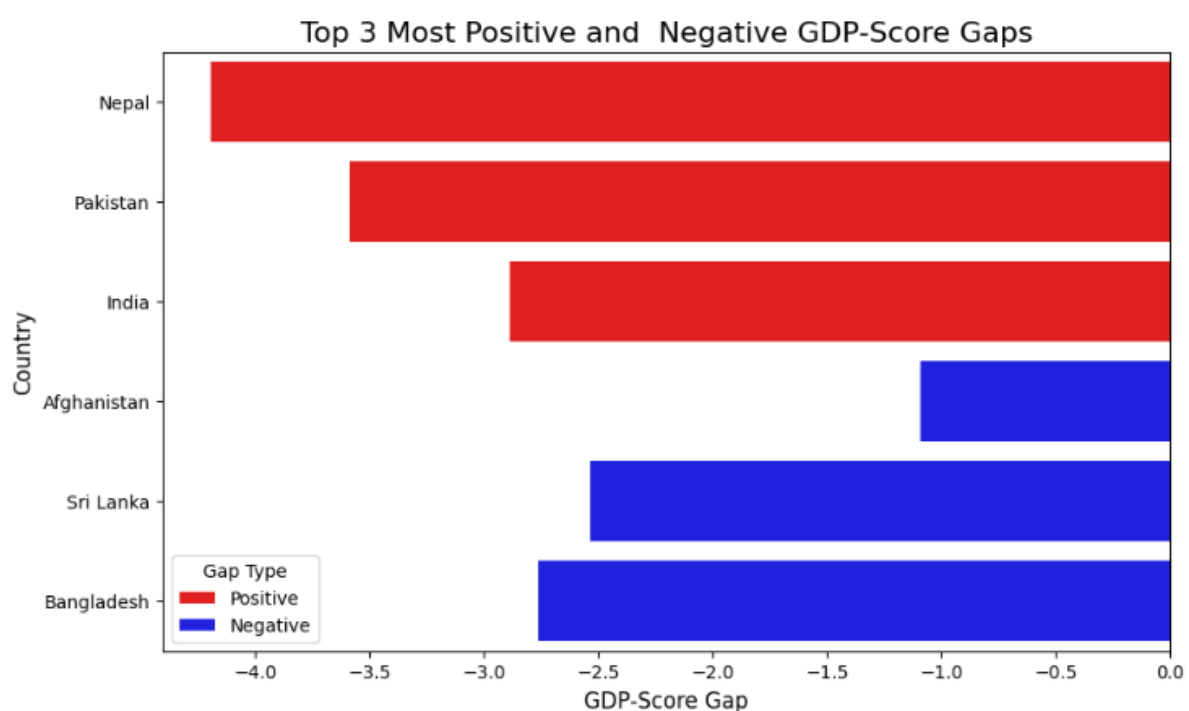
Task-3

Afghanistan is very different from other South Asian countries when it comes to happiness and wealth. It's considered an outlier because its happiness score and GDP per capita are way outside the normal range for the region. Outliers are data points that are much higher or lower than most of the other data points. They can happen for many reasons, like mistakes in measurement, rare events, or just really unusual situations.

Task-4

I found that generosity and freedom to make life choices are strongly linked to happiness. Generosity has the strongest connection, as shown by a correlation coefficient of 0.8745. This means that as generosity increases, happiness tends to increase as well. Freedom to make life choices also has a positive correlation, though slightly weaker, with a coefficient of 0.8005. When we looked at the data visually, we saw that both generosity and freedom to make life choices have a positive linear relationship with happiness. However, generosity's relationship is even stronger, as it follows the trend line more closely.
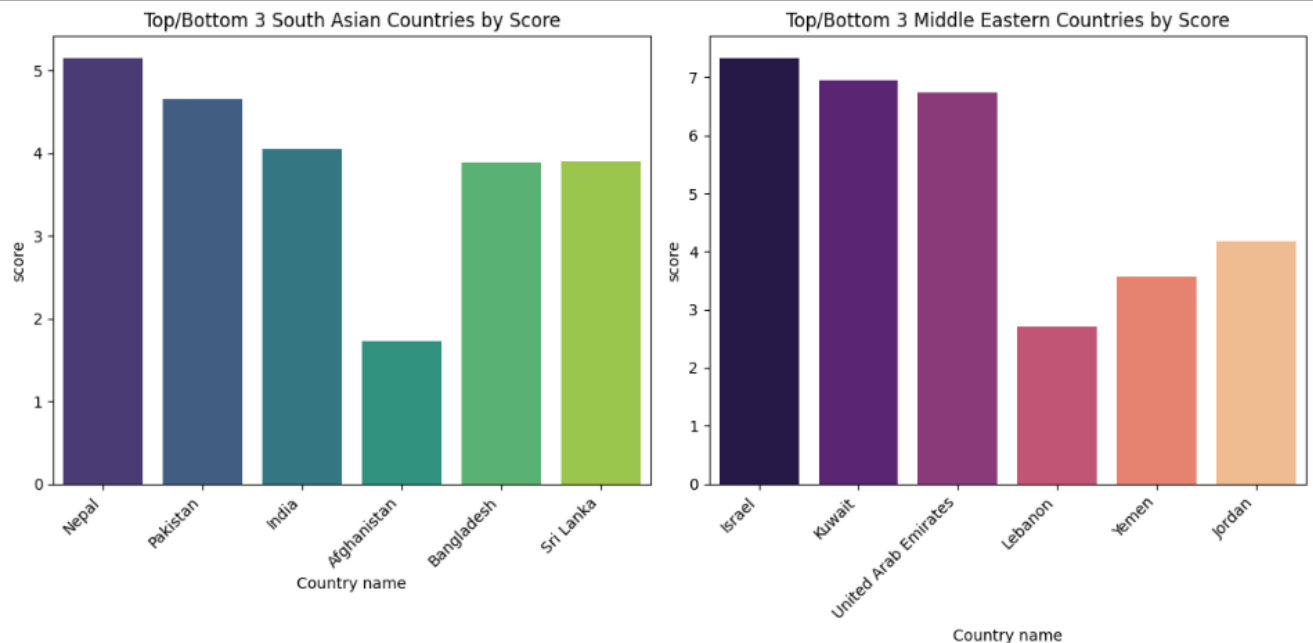
Task-5

I added a new column called "GDP-Score Gap" to show the difference between a country's GDP per capita and its happiness score. I then ranked the South Asian countries by this gap, both from smallest to largest and largest to smallest. I made a graph to show the countries with the biggest positive and negative gaps, so we can see these differences more clearly.

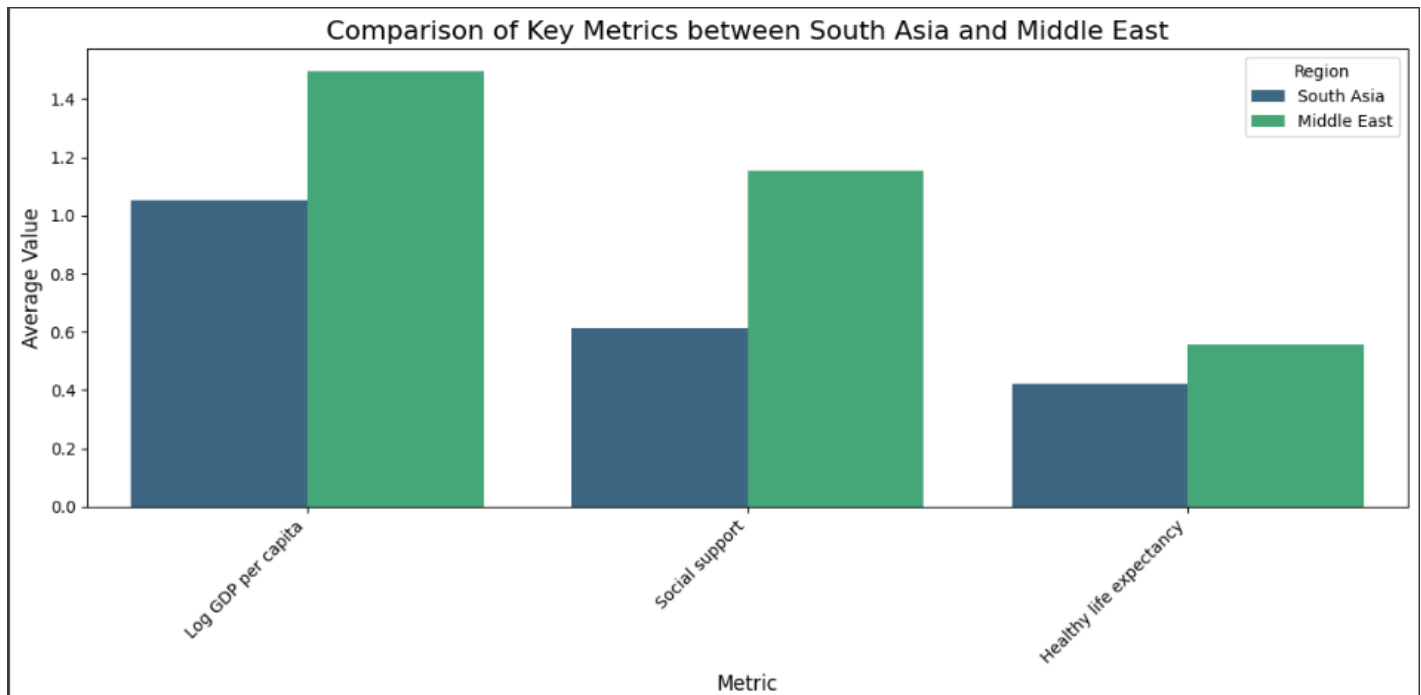Top 3 Most Positive and Negative GDP-Score Gaps

When GDP is higher than happiness, it suggests that things like inequality, corruption, and environmental damage can hurt people's well-being, even if the economy is doing well. Materialism and a lack of trust in others can also make people less happy. On the other hand, when happiness is higher than GDP, it shows that strong social connections, resilience, and cultural values are important for well-being. Welfare programs can also improve people's lives, even in poorer countries. These findings suggest that we need to tackle inequality and corruption, while also supporting social welfare and community to create lasting happiness.

Problem- 3

1. The graph shows a comparison between Southeast Asian and Middle Eastern countries. In Southeast Asia, Nepal, Pakistan, and India are the top three, while in the Middle East, Israel, Kuwait, and the UAE lead. On the other hand, Sri Lanka, Bangladesh, and Afghanistan are the bottom three in Southeast Asia, and Jordan, Yemen, and Lebanon are the lowest-ran in the Middle East.
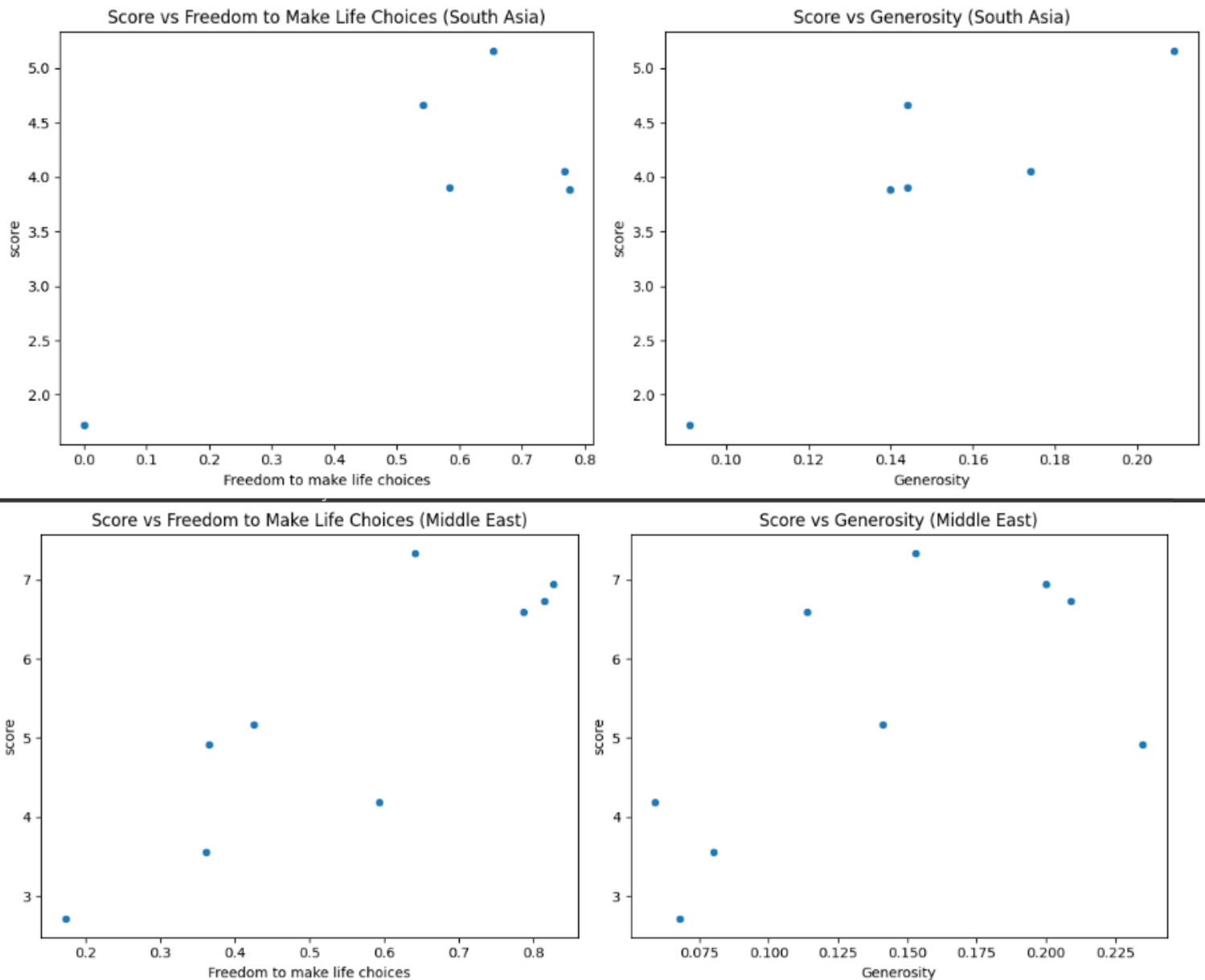
2. The graph compares South Asian and Middle Eastern countries based on GDP per capita, social support, and health life expectancy. It's interesting to note that there's a big difference in social support between the two regions. South Asia averages around 0.4, while the Middle East averages 1.4.
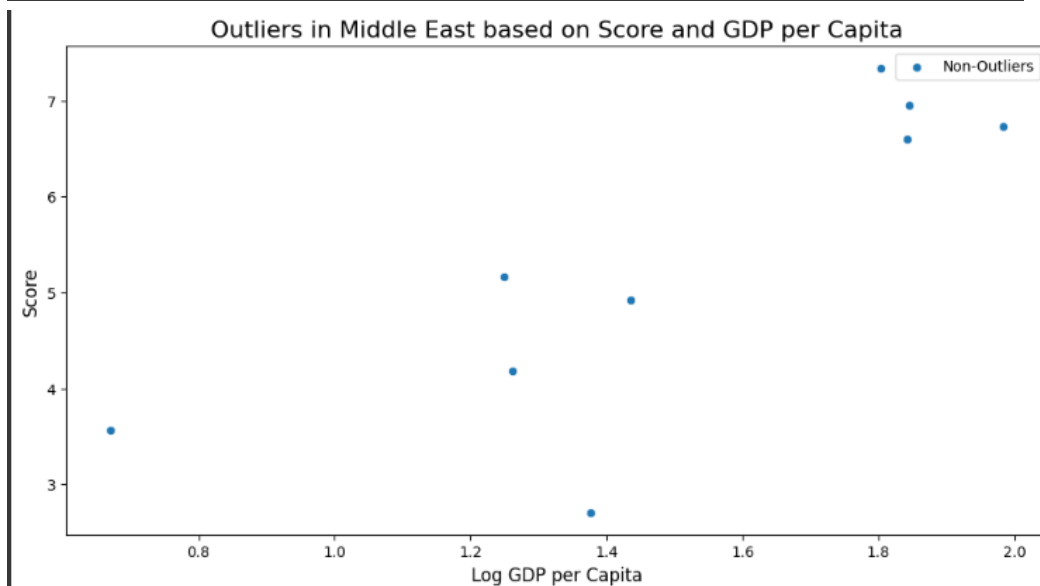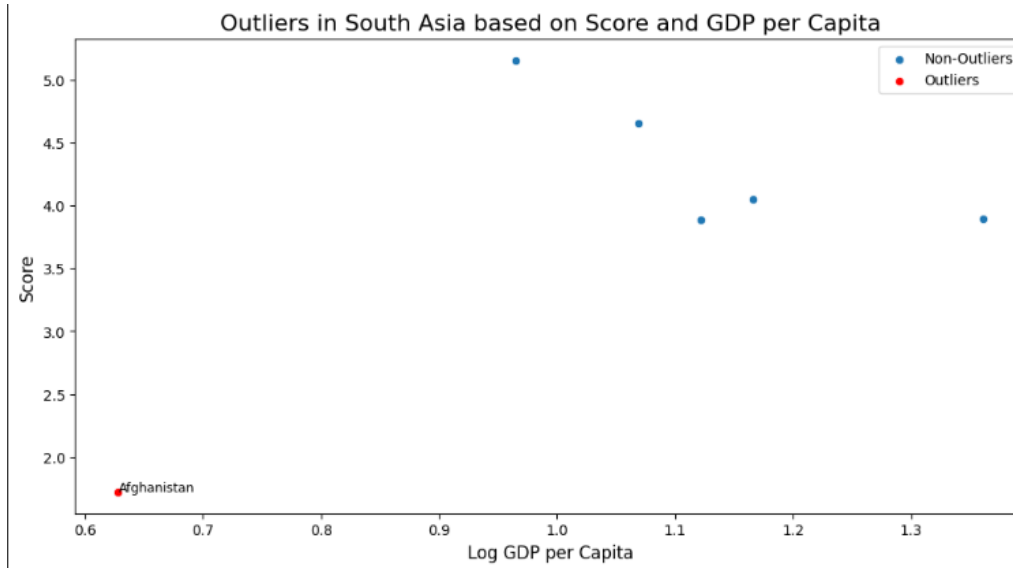

4. The range of happiness scores in South Asia is 3.437, with a coefficient of variation of 30.21%. In the Middle East, the range is 4.634 and the coefficient of variation is 28.94%. This means that happiness levels vary more in South Asia.

5. After analyzing how happiness scores relate to factors like freedom of choice and generosity within each region, we created scatter plots to visualize these relationships

Score vs Freedom to Make Life Choices (South Asia)

Score vs Generosity (South Asia)

Score vs Freedom to Make Life Choices (Middle East)

Score vs Generosity (Middle East)

6. I have identified countries in both regions that are outliers based on their happiness scores and GDP per capita. These outlier countries usually have unique characteristics
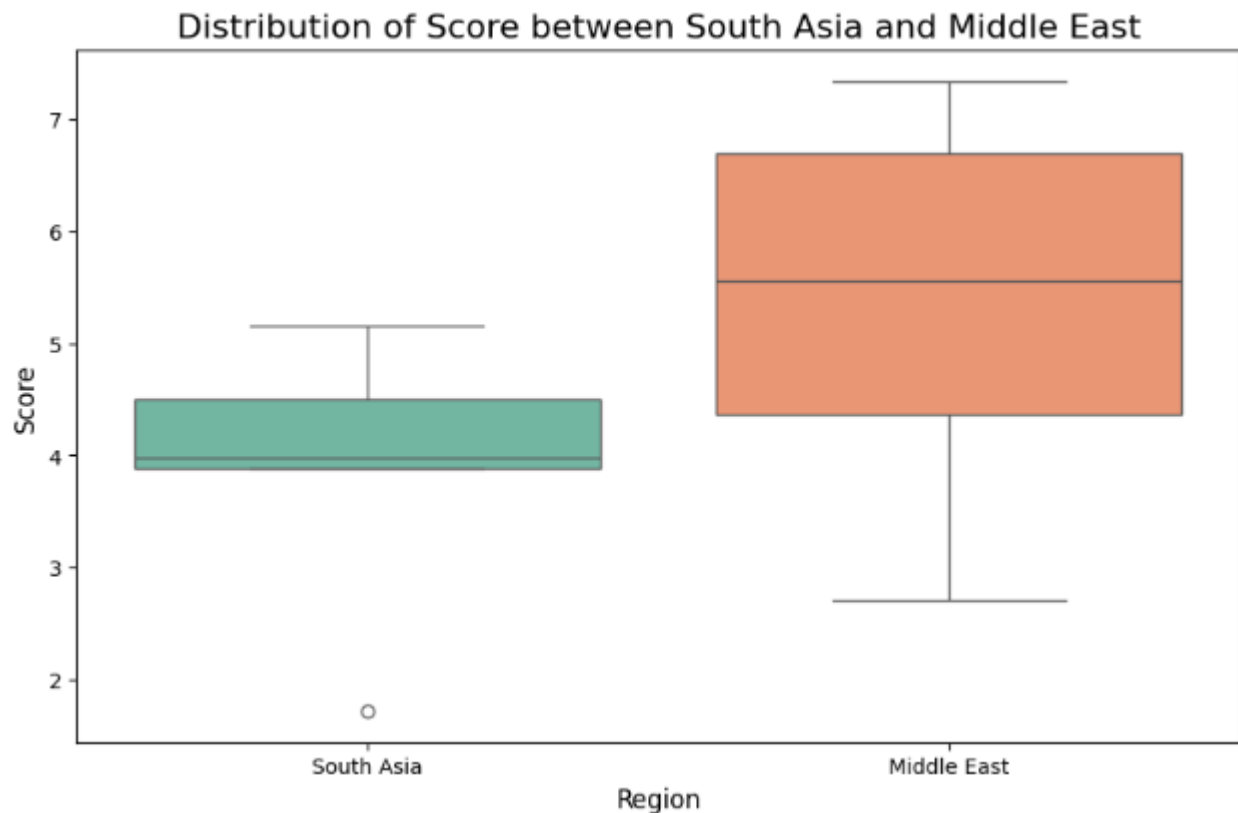

Outliers in South Asia based on Score and GDP per Capita

Outliers in Middle East based on Score and GDP per Capita

7. We created box plots to compare how happiness scores are distributed in South Asia and the Middle East.

Distribution Shapes: The box plots show us the overall shape of how happiness scores are distributed in each region.

Medians: The line inside each box shows the median happiness score. A higher median means that region is generally happier.

Outliers: The points outside the whiskers of the box plots are outlier countries with unusually high or low happiness scores.

Distribution of Score between South Asia and Middle East

Conclusion:

This report compared happiness levels in South Asia and the Middle East. We found that happiness levels vary more in South Asia, and that South Asia generally has lower social support and GDP per capita compared to the Middle East. Generosity was the strongest factor linked to happiness, while freedom to make life choices had a slightly weaker connection. When GDP is higher than happiness, it suggests issues like inequality and materialism. When happiness is higher than GDP, it suggests strong social connections and resilience. Countries like Afghanistan and Lebanon, which are outliers, significantly impact their regions' average happiness levels. Overall, social support and healthcare are crucial for happiness in both regions. To improve happiness, it's important to address inequality and promote strong communities.

GitHub - https://github.com/Rashidsid/Concept-and-technology-of-ai/commit/6b671651e30801b916563e05de2640dbd68000b2#diff-c563088098d3379af3255e2bac251dc6ebb6d049dc0455ef9453f59b5048cb11R1