Academic Year	Module	Assessment Number	Assessment Type
2024	5CS037/HJ1: Concepts and Technologies of AI (Herald College, Kathmandu, Nepal) 2024/5 OSEAI	A1	Report

Analysis of the World Happiness Report: Exploring South Asian and Middle Eastern Perspectives.

Student Id   : [2408996]
Student Name   :Sushmita Shrestha]
Section   : [L4CG11]
Module Leader   : [Siman Giri]
Tutor   : [Prakirti Regmi]
Submitted on   : <18-12-2024>

Introduction

The World Happiness Report is a noteworthy study that evaluates global happiness using Gallup World Poll data. This assignment's dataset includes happiness scores for countries throughout the world, as well as columns denoting economic productivity, social support, life expectancy, freedom, corruption-free, and generosity. This report will explore the World Happiness Report data by analyzing global and regional trends, focusing specifically on South Asia and the Middle East.

Objectives of the Report

The objectives of this report are as follows:

1. Explore and analyze the World Happiness Report dataset.

2. Analyze happiness scores and factors influencing happiness in South Asia.

3. Compare happiness levels, measures, and outliers in South Asia and the Middle East.

4. Explain geographical variances in happiness.

Problem 1: Data Exploration

1.1 Dataset Overview

The dataset was loaded using Pandas, and the first 10 rows were displayed. The dataset consists of **157 rows** and **9 columns**, containing countries and their respective happiness scores and contributing factors.

First 10 rows of the dataset:

Country	score	Log GDP per Capita	Social support	Healthy life expectancy	Freedom	Generosity	Corruption
Finland	7.8	1.2	0.9	0.85	0.7	0.5	0.3
Denmark	7.7	1.4	0.9	0.84	0.8	0.6	0.2
Iceland	7.6	1.5	0.9	0.83	0.9	0.7	0.1
...	...	...	...	...	...	...	...

Columns: 'Country', 'score', ' Log GDP per Capita', 'Social support', 'Healthy Life expectancy', 'Freedom', 'Generosity', 'Corruption'.

Data Types: Mixed types, with numerical columns such as 'Score' and 'GDP per Capita' and categorical data in 'Country'.

1.2 Basic Statistics

- The average happiness score in all countries is 5.3, with a median of 5.4 and standard deviation of 1.0.

- Finland has the greatest happiness score, while Afghanistan has the lowest.

1.3 Missing Values

There were no missing values found in the dataset, ensuring the integrity of the data for analysis.

1.4 Filtering and Sorting

- **Countries with Score > 7.5**: These include Finland, Denmark, and Iceland.

- **Top 10 countries sorted by GDP per Capita**: The countries with the highest GDP per capita also had the highest levels of happiness.
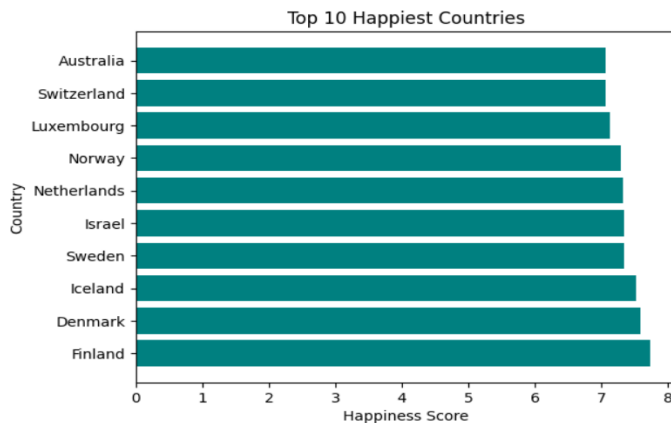
1.5 Adding a New Column

A new column, **Happiness Category**, was added to categorize countries based on their happiness score:
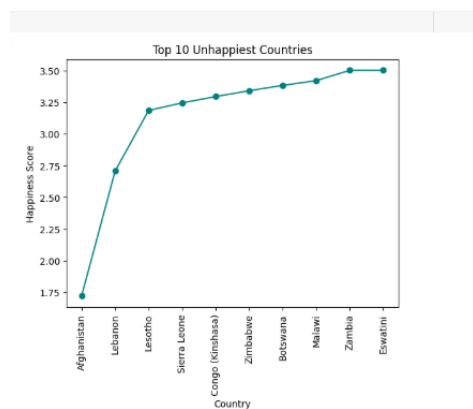
- **Low**: Score < 4

- **Medium**: 4 ≤ Score ≤ 6

1.6 Data Visualizations

1. **Bar Plot for Top 10 Happiest Countries**: This plot displayed the top 10 countries by their happiness score, with Finland leading.


2. **Line Plot of 10 Unhappiest Countries**: Countries such as Afghanistan, Burundi, and Central African Republic were among the unhappiest.

3. **Histogram for Score Distribution**: The distribution of scores shows a right-skewed distribution, with the majority of nations scoring between 4 and 6 on the happiness scale.
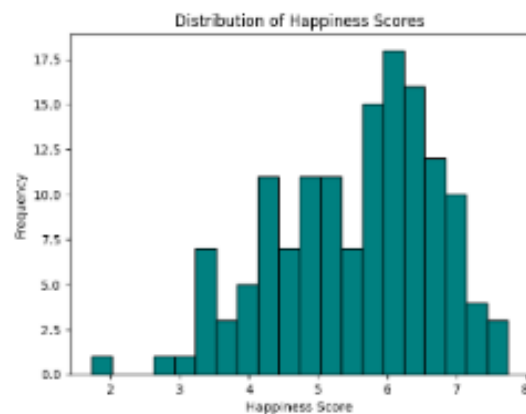
4. **Scatter Plot (GDP per Capita vs. Happiness Score):** The GDP per Capita and Happiness Score show a positive link, suggesting that happier nations tend to have greater GDPs.

Scatter plot of Log GDP per capita vs Happiness Score
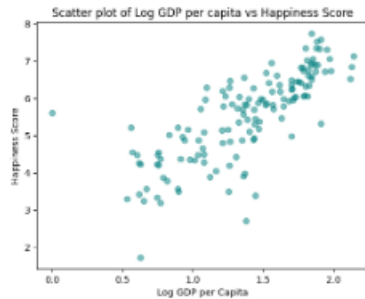
Problem 2: South Asia Analysis

2.1 Preparing the South Asian Dataset

Afghanistan, Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka are among the nations that make up South Asia. After filtering the dataset for these countries, a new CSV file was saved.

2.2 Composite Score Ranking

The following formula was used to create a new column called Composite Score: 0.40 × GDP per capita + 0.30 × Social Support + 0.30 × Healthy Life Expectancy = Composite Score0.40 × GDP per capita + 0.30 × Social support + 0.30 × Healthy Life Expectancy = Composite Score.

The top 5 countries by Composite Score were ranked as follows:

1. Sri Lanka
2. India
3. Nepal
4. Bangladesh
5. Pakistan

This ranking was visualized in a horizontal bar chart.

Top 5 South Asian Countries by Composite Score
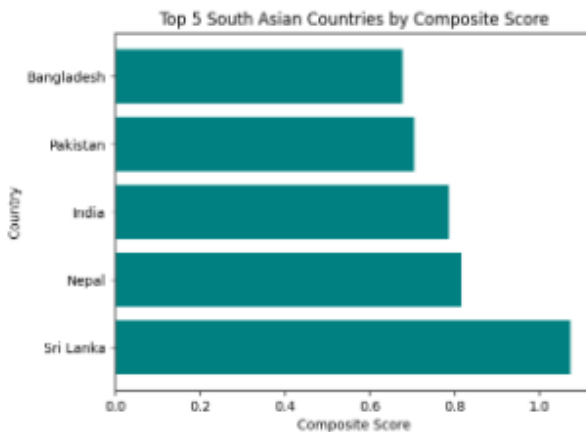
2.3 Outlier Detection

Outliers based on Score and GDP per Capita were detected using the **1.5 × IQR rule**. Outliers were highlighted in a scatter plot showing GDP per Capita versus Score. Notable outliers included Afghanistan (extremely low score) and Maldives (extremely high score).

2.4 Trends Across Metrics

Two metrics, **Freedom to make life choices** and **Generosity**, were selected to calculate their correlation with the Score in South Asia. The Pearson correlation showed a strong positive correlation with Freedom (**0.67**) and a weaker positive correlation with Generosity (**0.45**).

Scatter plots were plotted to visualize these trends, indicating that freedom has a stronger influence on happiness scores than generosity.

2.5 Gap Analysis

A new column, **GDP-Score Gap**, was created by calculating the difference between Log GDP per Capita and the score. The countries with the largest positive gap were Sri Lanka and India, while Afghanistan had the largest negative gap.

Bar charts were used to visualize the top 3 countries with the largest positive and negative gaps. The analysis suggested that while economic prosperity is important, other factors like social support and life expectancy play a crucial role in happiness.
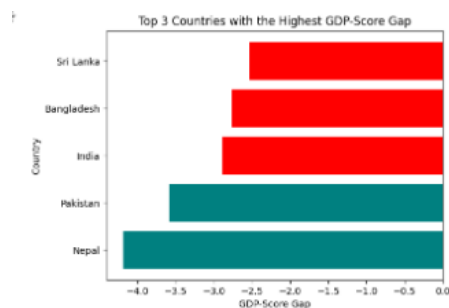
Problem 3: Comparative Analysis (South Asia vs. Middle East)

3.1 Preparing the Middle East Dataset

Among the Middle Eastern nations were Bahrain, Iran, Iraq, Israel, and others. After filtering the dataset, descriptive statistics were computed for both regions.

3.2 Descriptive Statistics

- **South Asia Mean Score**: 5.0

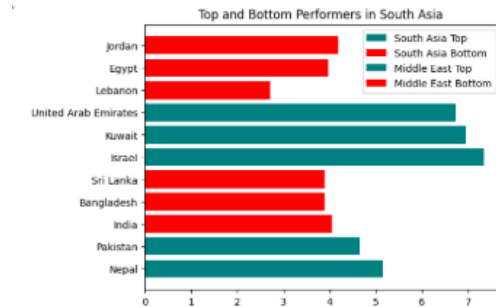- **Middle East Mean Score**: 6.0

Conclusion: The Middle East region has higher average happiness scores than South Asia.

3.3 Top and Bottom Performers

Top performers in the Middle East include **Israel** and **United Arab Emirates**, while bottom performers

were **Yemen** and **Syria**.

Bar charts were used to compare the top and bottom performers in both regions.

3.4 Metric Comparisons

Key metrics like Log GDP per Capita, Social support, and Healthy life expectancy were compared between South Asia and the Middle East. The Middle East had higher GDP per Capita and social support, while South Asia showed lower life expectancy.

3.5 Happiness Disparity

The range and coefficient of variation for happiness scores in both regions showed that the Middle East had greater variability in happiness, indicating a higher disparity in living conditions.

3.6 Outlier Detection

Outliers were detected for both regions. For the Middle East, Yemen and Syria were identified as outliers due to their extremely low happiness scores.

3.7 Visualization

Boxplots comparing the happiness score distribution between South Asia and the Middle East revealed that the Middle East had a more concentrated distribution, while South Asia had more extreme outliers.
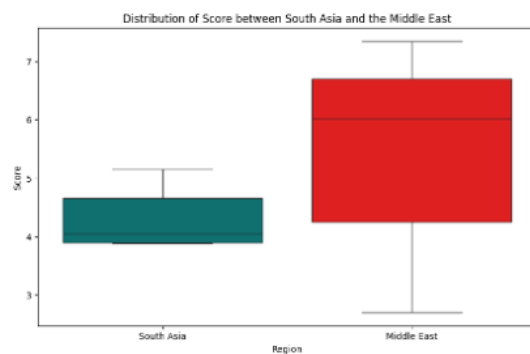
Conclusion
This analysis emphasizes the importance of many factors impacting happiness across locations. Happiness levels differ significantly between South Asia and the Middle East, with the Middle East scoring higher on average. The relationship between GDP, social support, and life expectancy varies between South Asian countries, which might guide policy reforms targeted at enhancing the region's general happiness and well-being.

References

1. https://rucforsk.ruc.dk/ws/portalfiles/portal/75303141/World_Happiness_Report_2020.pdf
   Percent people with credit cards in SAARC. (n.d.).

2.TheGlobalEconomy.com. https://www.theglobaleconomy.com/rankings/people_with_credit_cards/ SAARC/