World Happiness Index and Inflation Dataset Analysis

This project explores the relationship between economic stability and overall well-being by analyzing the *World Happiness Index and Inflation Dataset*. The dataset, which is available on [Kaggle](#), includes data from 148 countries spanning the years 2015 to 2023. It combines the World Happiness Index with key economic indicators such as GDP per capita, inflation and social factors like freedom, social support, and generosity. In this project, the term **Score** is used as a proxy for a country's happiness level.

The analysis focuses on answering five primary questions:

1. Distribution of Scores

Objective:
The goal here is to understand the overall spread of the **Score** values in the dataset and to identify any potential outliers. This involves generating descriptive statistics and creating visual representations such as a histogram (with a density estimate) and a boxplot.

Approach Summary:

- Load the dataset and verify the column names to ensure that the **Score** column is correctly identified.

- Generate descriptive statistics (e.g., count, mean, standard deviation) for the **Score**.

- Create visualizations: a histogram helps to visualize the distribution shape and a boxplot highlights any outliers.

2. GDP per Capita and Score Correlation

Objective:
This section examines whether there is a relationship between a country's economic wealth (measured by GDP per capita) and its **Score**. The analysis involves both visualization and statistical computation.

Approach Summary:

- Create a scatter plot to visually inspect the relationship between GDP per capita and the **Score**, using color coding (by continent) to add another layer of insight.

- Calculate the Pearson correlation coefficient to quantitatively assess the strength and direction of the relationship.

3. Inflation Trends by Continent and Their Relation to Score

Objective:
This analysis investigates how inflation rates have changed over time across different continents and examines how these trends compare with the changes in **Score** values.

Approach Summary:

- Aggregate the data by grouping it according to both continent and year to calculate the average inflation rate and average **Score**.

- Use line plots to display the trend of inflation rates over time for each continent.

- Similarly, create a line plot to show how the average **Score** changes over time across continents.

4. Countries with Significant Changes in Scores

Objective:
This part of the analysis identifies which countries have experienced the most significant improvements or declines in **Score** over the period of study (from 2015 to 2023).

Approach Summary:

- Reshape the dataset using a pivot table where each row represents a country and columns represent the different years.

- Calculate the change in **Score** by subtracting the value in 2015 from that in 2023 for each country.

- Identify and list the top 10 countries with the greatest positive change (improvements) and the top 10 with the most negative change (declines).

5. Impact of Social Factors on Score

Objective:
This analysis examines how social factors specifically freedom, social support, and generosity influence the overall **Score**. This is performed through correlation analysis and pairwise visualizations.

Approach Summary:

- Compute the Pearson correlation coefficients between each of the social factors and the **Score** to assess their relationships.

- Create a pairplot to visually explore the pairwise relationships among the social factors and the **Score**, which helps in identifying patterns or potential nonlinear interactions.

Overall Recommendations

- **Error Handling and Data Quality:**
  Before conducting any analysis, add steps to check for missing values, outliers, or incorrect data entries. This improves the robustness of your analysis.

- **Documentation and Clarity:**
  Each section of the analysis should include clear comments and explanations to help others understand the logic behind your approach. This is especially useful if someone unfamiliar with the project needs to get up to speed quickly.

- **Modular Code and Reusability:**
  Consider creating reusable functions for common tasks such as data cleaning, plotting, and calculating statistics. This makes your code cleaner and easier to maintain.

- **Extended Analysis:**
  Based on your audience and project goals, you might add further analyses such as advanced modeling, additional visualizations, or deeper exploration of specific indicators. Always include the rationale behind any extra analysis steps.

Reference

Data Source:
World Happiness Index and Inflation Dataset