World Happiness Report Dataset (Minor project)

Task List for World Happiness Report Dataset

1. Import Libraries
   - Import `pandas`, `numpy`, `seaborn`, and `matplotlib` for data analysis and visualization.

2. Load the Dataset
   - Load the World Happiness Report dataset into a DataFrame.

3. Initial Data Exploration
   - Display the dataset's first and last few rows using `.head()` and `.tail()`.
   - Check the shape of the dataset using `.shape()`.

4. Summary of Data
   - Display the column names, data types, and non-null counts using `.info()`.
   - Provide a summary of the numeric columns using `.describe()`.

5. Check for Missing Values
   - Identify the count of missing values in each column using `.isnull().sum()`.

6. Handle Missing Data
   - Fill missing values using appropriate methods (e.g., mean, median, or drop rows).

7. Check for Duplicates
   - Check for duplicate rows using `. duplicated().sum()`.
   - Remove duplicates if any are present.

8. Rename Columns (if needed)
   - Rename columns for better readability using `.rename()`.

9. Identify Outliers
   - Detect outliers using visual methods such as box plots for numeric columns.

10. Data Transformation
   - Apply transformations on skewed columns (e.g., log transformation).

11. Data Visualization
   - Plot histograms or bar charts to visualize the distribution of important variables (e.g., Happiness Score).
12. Correlation Analysis
   - Compute and visualize the correlation matrix between numeric columns using a heatmap.

13. Pair Plot
   - Use Seaborn's `pair plot` to visualize relationships between multiple variables.

14. Grouping and Aggregation
   - Group data by a categorical column (e.g., region) and calculate the mean of the Happiness Score.

15. Draw Insights
   - Based on the analysis, write down insights and conclusions from the data.
   - Discuss factors influencing Happiness Score.