SPSS LAB 1
Lab: Working with NumPy, Pandas, Seaborn, and Matplotlib

Question:

You are given a dataset named "data.csv" containing information about individuals, including their name, age, gender, income, and category. You need to perform a series of data analysis and visualization tasks using Python libraries such as NumPy, Pandas, Seaborn, and Matplotlib.

1. Data Loading and Initial Exploration:
    a. Use Pandas to load the data from "data.csv" into a DataFrame called df. Display the first 5 rows of the DataFrame.
    b. Perform a preliminary data exploration using Pandas functions:
        i. Print basic information about the DataFrame, including data types.
        ii. Provide summary statistics for numeric columns.
        iii. Check for missing values and display the count of missing values in each column.
        iv. Show the value counts for the "Category" column.
2. Data Cleaning and Transformation:
    a. Handle missing values by filling them with the mean value for numeric columns (e.g., "Age" and "Income").
    b. Perform a data transformation by adding 1 to the "Age" column.
3. Data Visualization:
    a. Create a box plot to visualize the distribution of "Income" across different categories ("Category" column).
    b. Generate a histogram of "Age" with 10 bins and a kernel density estimate (KDE).
    c. Create a pair plot to explore relationships between numeric columns, with data points color-coded by the "Category" column.
4. NumPy Operations:
    a. Use NumPy to create an array called income_values containing the "Income" values from the DataFrame.
    b. Calculate the mean income value using NumPy and print the result.