Assignment 1

Objective:

Develop Python programming and problem-solving skills by tackling tasks relevant to both social science and general computational logic. This assignment will help you improve your data analysis, programming, and visualization skills.

Instructions:

1. Discuss ideas with your classmates if you need, but you must submit your own work.
2. Provide your solutions as a Jupiter Notebook with code, outputs, and brief comments explaining your logic. Use markdown cells for additional context.
3. Attempt all questions, progressing from basic to advanced.

Basic Problems:

1. **Simple Survey Analysis**
   Write a program that collects responses to a simple survey.

- Ask the user a question, e.g., "How satisfied are you with our services on a scale from 1 to 5?"
- Collect 5 responses.
- Calculate and display the **average satisfaction score** at the end.

2. **Age Group Classification**
   Write a program to classify individuals into age groups based on their age:

- Input: A list of ages, e.g., [15, 23, 35, 42, 67, 80].
- Output: A list of age group labels, e.g., [Teen, Young Adult, Adult, Adult, Senior, Senior].
    - Teen: 13–19
    - Young Adult: 20–35
    - Adult: 36–59
    - Senior: 60+

3. **Basic Text Analysis**
   Ask the user to input a sentence. Write a program to:

- Count the total number of words.
- Identify how many times each word occurs.
- Output the counts in a simple table format.

4. **Gender Distribution in a Dataset**
   Imagine you have survey data about people's gender identities stored as a list of strings, e.g., ["Male", "Female", "Female", "Male"]
   Write a program to:

- Count the number of people in each gender category.
- Display the counts as percentages.
- Visualize the distribution as a **pie chart** using Matplotlib.

5. **Simple Trend Analysis**
   You're given a small dataset of monthly unemployment rates, e.g., [5.1, 4.8, 4.6, 4.7, 5.0, 4.9].

- Write a program to calculate the **average unemployment rate**.
- In order to identify whether the unemployment rate is **increasing** or **decreasing** overall, plot the trend using a **line graph** with proper labels and titles.

6. **Population Statistics (NumPy)**
   Using NumPy, analyze the following population age dataset:
   ages = [15, 23, 35, 42, 67, 80, 29, 38, 54, 21, 19, 33, 45].

- Calculate the **mean**, **median**, and **standard deviation** of the ages.
- Determine the **percentage of the population** in each age group (Teen, Young Adult, Adult, Senior).
- Visualize the age distribution as a **bar chart**.

7. **Income Inequality Analysis**
   Given an income dataset (in thousands):
   incomes = [20, 22, 25, 30, 35, 35, 40, 45, 50, 90, 120, 200].

- Calculate the **mean** and **median** income.
- Plot the income data using a **histogram** to visualize the distribution.
- Discuss whether the data indicates income inequality.

8. **Data Simulation and Visualization**
   Simulate survey data using NumPy. For example:

- Create a dataset of **100 random ages** between 10 and 80.
- Create another dataset of **random monthly incomes** between $500 and $10,000.
- Analyze and visualize:
    - The age distribution as a **bar chart**.
    - The relationship between age and income using a **scatter plot**.

9. **Number Guessing Game**
   Write a program for a simple **number guessing game**:

- The program selects a random number between 1 and 50.
- The user tries to guess the number, and the program provides feedback:
    - If the guess is too high or too low.
    - If the guess is correct, congratulate the user and display the total number of attempts.

Submission: A Jupiter notebook containing the code, outputs, and brief comments explaining each step of the analysis, if necessary. For your comments and headers, please use the *markdown* option.