Question 1: Lists and Loops

Write a Python program that performs the following:

1. Define the list:
   numbers = [15, 8, 22, 7, 31, 4, 17]
2. Print all even numbers from the list.
3. Create a new list containing the square of each odd number.
4. Print the new list.

Question 2: Strings and Dictionaries

Write a program to count how many times each word appears in a sentence.

1. Given the sentence:
   "The book was interesting because the book covered many topics and the topics discussed in the book were engaging"
2. Convert the sentence to lowercase and split it into words.
3. Create a dictionary where:
   - Keys are unique words
   - Values are the word counts
4. Print the dictionary.

Question 3: Functions and Conditional Logic

Define a function called is_prime(n) that checks if a number is prime.

1. Given the list:
   nums = [2, 4, 5, 10, 13, 17, 20, 23]
2. Use the function to check each number in the list.
3. Store and print only the prime numbers in a new list.

Question 4: NumPy Array Operations and Statistics

Use NumPy to simulate and analyze student test scores.

1. Generate a 2D NumPy array scores of shape (10, 5) representing scores of 10 students in 5 subjects, with random integers between 50 and 100.
2. Perform the following:
   - Compute the average score per student.

- Compute the maximum score per subject.
- Add a new column with each student's average score.
- Print the final array.

Question 5: Saving Data to CSV

Using the array from Question 4:

1. Create a CSV file named student_scores.csv with the following header: "Subject1,Subject2,Subject3,Subject4,Subject5,Average"
2. Save the data (including the average column) to the file.
3. Ensure the file contains:
   - One row per student
   - Comma-separated values
   - The header row at the top

Question 6: Creating and Exploring a DataFrame

Create a Pandas DataFrame with the following structure:

- Columns: Product (string), Price (float), Quantity (integer)
- Add at least 8 rows of sample data of your choice.

Perform the following operations:

1. Add a column TotalValue calculated as Price × Quantity.
2. Sort the DataFrame in descending order by TotalValue.
3. Group the data by Price and compute the average Quantity for each price group.

Question 7: Analysis of a Real-World Dataset (Titanic)

Use the Titanic dataset:
🔗 https://www.kaggle.com/c/titanic/data (Download train.csv)

1. Load the dataset using Pandas.
2. Display the total number of missing values in each column.
3. Filter and display:
   - Passengers who are under 18 years old
   - Female passengers who survived

4. Calculate the average fare paid by passengers in each passenger class (pclass).

Question 8: Data Cleaning and Aggregation (Heart Disease Dataset)

Use the UCI Heart Disease dataset:
🔗 https://archive.ics.uci.edu/dataset/45/heart+disease
(File: processed.cleveland.data)

1. Assign these column names when loading the file:
   ['Age', 'Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS', 'RestingECG', 'MaxHR', 'ExerciseAngina', 'Oldpeak', 'ST_Slope', 'NumVesselsFluoro', 'Thalassemia', 'HeartDisease']
2. Replace '?' with np.nan and convert relevant columns to numeric.
3. Perform the following:
   - Display the count of missing values in each column.
   - Replace missing values in NumVesselsFluoro and Thalassemia using the median or mode, whichever is appropriate.
   - Drop any row where more than two columns have missing values.
   - Group the data by HeartDisease and compute the average cholesterol level for each group.

Question 9: Visualizing Student Scores

Given this data representing student performance:

Name	Math	Science	English	History
Alice	85	90	78	84
Bob	78	82	85	80
Charlie	92	89	88	91
David	88	94	80	79
Eva	76	75	82	77

Tasks:

1. Create a bar chart showing each student's Math score.
2. Create a line plot comparing all 4 subjects for each student (one line per subject).
3. Add proper title, axis labels, and a legend to the plot.

Question 10: Visualizing Random Data Distribution

Use NumPy and Matplotlib to generate and visualize data.

1. Generate 1000 random numbers from a normal distribution with mean = 50 and std = 10.
2. Plot a histogram of the data.
3. Customize the plot with:
    - 20 bins
    - A vertical line indicating the mean
    - Title, axis labels, and grid lines