Question 1 (NumPy) – Titanic Dataset Survival Analysis

1. Load the Titanic dataset using Pandas and convert the following numeric columns into a NumPy array:

   o age, fare, and survived.

2. Using **NumPy** (not Pandas), perform the following:

   o a) Remove rows with missing values in any of the selected columns.

   o b) Normalize age and fare columns using Z-score normalization.

   o c) Compute the **mean age and mean fare** for:

      ▪ passengers who **survived**

      ▪ passengers who **did not survive**

   o d) Use np.where() to classify fare as:

      ▪ "Low" if fare < mean,

      ▪ "High" otherwise.

Question 2 (Pandas) – Iris Dataset Summary Report

1. Load the Iris dataset using Pandas.

2. Create a new column called petal_ratio = petal_length / petal_width.

3. Answer the following:

   o a) What is the **average petal_ratio per species**?

   o b) Which species has the **highest standard deviation** in sepal_length?

   o c) Filter the dataset to only include rows where sepal_width is above the overall mean.

   o d) Split the DataFrame into two subsets:

      ▪ One with petal_ratio < 2

      ▪ One with petal_ratio ≥ 2

      ▪ Then stack them vertically using pd.concat.

Submission Format:

• Each answer should include:

   o Code

   o Key outputs (printed or commented)

   o Brief explanation (2–3 lines)