Assignment 1: KNN Classification on the Wine Quality Dataset

Question:

You are tasked with building a KNN classifier to predict the quality of wine based on various chemical properties in the **Wine Quality dataset**. Follow these steps to complete the assignment:

1. Download the **Wine Quality dataset** from the link provided below.
2. Preprocess the dataset:
    a. Handle any missing values if present.
    b. Standardize the features using `StandardScaler`.
3. Split the data into training and testing sets (80% training, 20% testing).
4. Build a KNN classifier with k=5 and train it on the training data.
5. Evaluate the model's performance on the test data using accuracy, confusion matrix, and classification report.
6. Visualize the confusion matrix as a heatmap.
7. Experiment with different values of k (from 1 to 10) and plot the test accuracy for each value.

Expected Output:

- Classification accuracy and confusion matrix.
- A plot showing how test accuracy changes as k increases.
- Classification report including precision, recall, and F1-score.

Assignment 2: KNN Classification on the Titanic Dataset

Question:

- In this assignment, you will use the **Titanic dataset** to predict whether a passenger survived the sinking of the Titanic based on various features. Complete the following tasks:
- Download the **Titanic dataset** from the link provided below.
- Preprocess the data:
    - Handle any missing values, especially in the `Age` and `Embarked` columns.
    - Convert categorical variables (e.g., `Sex`, `Embarked`) into numerical values using one-hot encoding.
    - Standardize the features using `StandardScaler`.
- Split the data into training and testing sets (70% training, 30% testing).
- Build a KNN classifier with k=5 and train it on the training data.
- Evaluate the model's performance on the test data using accuracy, confusion matrix, and classification report.
- Test the model by predicting the survival outcome for a new passenger with the following features:
    - `Pclass = 1` (First class)
    - `Sex = 'female'`
    - `Age = 29`
    - `SibSp = 0` (No siblings/spouses aboard)
    - `Parch = 0` (No parents/children aboard)
    - `Fare = 50`
    - `Embarked = 'C'` (Cherbourg)
- Experiment with different values of k (from 1 to 15) and plot the test accuracy for each value.

Expected Output:

- Accuracy, confusion matrix, and classification report for the model.
- Predicted survival outcome for the new passenger.
- A plot showing the accuracy vs. different values of k