Question 2:

You are provided with a dataset. Your task is to build a machine learning model to detect its Y variable using various Data Mining techniques within a 1.5-hour time frame.

1. Data Exploration and Visualization:

Load the dataset and explore its structure using Pandas.   
Visualize key features to gain insights into the data.

2. Data Preprocessing:   
Handle any missing values and outliers in the dataset.   
Perform feature scaling and transformation if necessary.

3. Model Building and Evaluation:   
Split the dataset into training and testing sets (e.g., 70% training, 30% testing).   
Build and train a classification model using any 2 of the following algorithms:

* KNN
* NaiveBayes
* SVM

Evaluate the model's performance using metrics like accuracy, precision, recall, and F1- score on the test set.  
Visualize the confusion matrix and ROC curve for model evaluation.   
Which model classification accuracy is better ?   
  
Note: You can use Python libraries like Pandas, NumPy, and Matplotlib for data manipulation, model building, and visualization.