# **DEFINITION OF DONE: R CODE OVERVIEW:**

**NOTE: This isn’t a report, but could be a helpful pointer in your report.**

1. **Data Preprocessing and Exploration:**
   * Reads the dataset and performs some exploratory data analysis (EDA) like displaying the structure, head, and variable names of the dataset.
   * Converts relevant variables to factors.
2. **Feature Selection Using Cramer's V:**
   * Calculates Cramer's V values for each predictor variable with the target variable.
   * Selects predictor variables with Cramer's V values above a specified threshold for feature selection.
3. **Data Partitioning:**
   * Partitions the dataset into training and test sets.
4. **Naive Bayes Modeling and Prediction:**
   * Fits a Naive Bayes model on the training data using the **naiveBayes** function from the **e1071** package.
   * Makes predictions on the test data.
5. **Confusion Matrix and Evaluation Metrics:**
   * Constructs a confusion matrix using the **confusionMatrix** function from the **caret** package.
   * Extracts accuracy, specificity, and precision from the confusion matrix.
   * Calculates the F1-score using precision and recall.
6. **Visualization:**
   * Visualizes the confusion matrix as a heatmap using **ggplot2**.
   * Creates a density plot of class probabilities.
7. **Performing Data mining Techniques After Feature Selection:**
   * Repeats the Naive Bayes modeling, prediction, and evaluation steps after feature selection.