R script code

```
# Vihari Reddy Tummuru and Max Goswitz
# MIS 545 Section 01
# Lab08Group13GoswitzTummuru.R
# Description: In this lab, we use the Naive Bayes method to predict what type
# of dwelling (Condo, Home, Apartment) an individual lives in based on other
# factors listed in the dataset like age, if they are a female, if they live
# alone, and if they have a paid salary.

# Environment Setup ------------------------------------------------------
# Install the tidyverse & e1071 packages
# install.packages("tidyverse")
# install.packages("e1071")

# Install the tidyverse & e1071 libraries
library(tidyverse)
library(e1071)

# Set the working directory to your Lab08 Folder
setwd("C:/Users/ual-laptop/Downloads/Lab08")

# Read DwellingType.csv into a tibble called dwellingType
dwellingType <- read_csv(file = "DwellingType.csv",
                         col_types = "filll",
                         col_names = TRUE )

# Display dwellingType in the console
print(dwellingType)

# Display the dwellingType structure
str(dwellingType)

# Display the dwellingType summary
summary(dwellingType)

# Set seed using 154 as random seed
set.seed(154)

# Split the dataset into dwellingTypeTraining (75% of records) and
# dwellingTypeTesting (25% of records)
sampleSet <- sample(nrow(dwellingType),
                    round(nrow(dwellingType) * 0.75),
                    replace = FALSE)
```

```r
# Set dwellingTypeTraining (75% of records)
dwellingTypeTraining <- dwellingType[sampleSet, ]

# Set dwellingTypeTesting (25% of records)
dwellingTypeTesting <- dwellingType[-sampleSet, ]

# Generate the Naive Bayes Model to predict DwellingType based on other
# variables in the dataset
dwellingModel <- naiveBayes(formula = DwellingType ~ . ,
                            data = dwellingTypeTraining,
                            laplace = 1)

# Build properties for each record in the testing dataset and store
# them in dwellingTypeProbability
dwellingTypeProbability <- predict(dwellingModel,
                                   dwellingTypeTesting,
                                   type = "raw")

# Display dwellingTypeProbability on the console
print(dwellingTypeProbability)

# Predict the classes for each record in the testing dataset & store them
# in dwellingTypePrediction
dwellingTypePrediction <- predict(dwellingModel,
                                   dwellingTypeTesting,
                                   type = "class")

# Display dwellingTypePrediction on the console
print(dwellingTypePrediction)

# Evaluate the model by forming a confusion matrix
dwellingConfusionMatrix <- table(dwellingTypeTesting$DwellingType,
                                 dwellingTypePrediction)

# Display the confusion matrix on the console
print(dwellingConfusionMatrix)

# Calculate the model predictive accuracy & store it into a variable
# called predictiveAccuracy
predictiveAccuracy <- sum(diag(dwellingConfusionMatrix)) /
  nrow(dwellingTypeTesting)

# Display the predictiveAccuracy on the console
print(predictiveAccuracy)
```

Questions:

2. Provide a possible explanation as to why the model was better able to predict home and apartment dwellers than condo dwellers.

Naïve Bayes models assume all independent variables are separate from one another and have equal importance on predicting the dependent variable. However, PaySalaried could have a higher impact on determining the individual's dwellingType, yet equal emphasis could have been placed on the LiveAlone & Female variables, both of which could not be as important for making predictions and thus made it harder for our model to correctly classify condos.

3. How could this model be used by a direct mailing marketing company to optimally target their mailing?

This model could be used by a direct mailing marketing company to optimally target their mailing by sending mail about certain products and deals to consumers' homes based on factors like gender and if they are a salary worker. If the marketing company wants to target high-income females within a certain age range, they can use this model to identify which dwelling type this customer demographic tends to live in so the mailing company can send them mail or deals.