1. From R:
    1. All code from your R script (Code should be presented single-spaced in a fixed-width font. Adjust the font size so that no lines of code extend to the next line in the document)

```
# YUEH-TING WU
# MIS 545 Section 02
# Lab08WuY.R
# In this R programming, importing a csv file and build a Naive
Bayes models to
# make a prediction.

# Install the tidyverse and e1071 packages
# install.packages("tidyverse")
# install.packages("e1071")

# Load the tidyverse and e1071 libraries
library(tidyverse)
library(e1071)

# Set the working directory to Lab08 folder
setwd("~/MIS 545/Lab08")

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

# Display dwellingType in the console
print(dwellingType)

# Display the structure of dwellingType in the console
str(dwellingType)

# Display the summary of dwellingType in the console
summary(dwellingType)

# Ramdonly split the dataset into dwellingTypeTraining (75% of
records) and
# dwellingTypeTesting (25% of records) using 154 as the random
seed
set.seed(154)
SimpleSet <- sample(nrow(dwellingType),
                    round(nrow(dwellingType) * 0.75),
                    replace = FALSE)
dwellingTypeTraining <- dwellingType[SimpleSet, ]
dwellingTypeTesting <- dwellingType[-SimpleSet, ]
```

```
# Generate the Naive Bayes model to predict DwellingType based
on the other
# variables in the dataset.

dwellingTypeModel <- naiveBayes(formula = DwellingType ~ .,
                                data = dwellingTypeTraining,
                                laplace = 1)

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

# Display dwellingTypeProbability on the console
print(dwellingTypeProbability)

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

# Display dwellingTypePrediction on the console
print(dwellingTypePrediction)

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

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

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

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

2. Answer the following question in a sentence: Provide a possible explanation as to why the model was better able to predict home and apartment dwellers than condo dwellers?

This model shows that it has 52% of times we will be able to accurately predict in the testing dataset, thus it may fail when it made a prediction about the condo dwellers.

3. Answer the following question in a sentence: How could this model be used by a direct mailing marketing company to optimally target their mailings?

A direct mailing marketing company can use this model to predict what type of products the customer will be interested in and mail them a related one.