R Script:

```
# Aadhithya Dinesh
# MIS 545 Section 02
# Lab08DineshA.R
# Import a dataset of people and generate a Naive Bayes model to predict
# a person's dwelling type based on demographic data. We will be assigning
# data types, building a model and testing for model fit.

# install.packages("tidyverse")
# install.packages("e1071")

library("tidyverse")
library("e1071")

# set the working directory
setwd("~/MIS/Classes/MIS545/Assignments/Lab08")

dwellingType <- read_csv(file = "DwellingType.csv",
                         col_types = "filll",
                         col_names = TRUE)

# print the dwellingType tibble
print(dwellingType)

# print the structure of dwellingType
print(str(dwellingType))

# print the summary of dwellingType
print(summary(dwellingType))

# set the seed to 154
set.seed(154)
sampleSet <- sample(nrow(dwellingType),
                    round(nrow(dwellingType)*0.75),
                    replace = FALSE)
# loading 75% of the training dataset
dwellingTypeTraining <- dwellingType[sampleSet, ]

# loading the remaining 25% of the dataset for testing
dwellingTypeTesting <- dwellingType[-sampleSet, ]

# generating the naiveBayes model for finding the dwelling type
dwellingTypeModel <- naiveBayes(formula = DwellingType ~ .,
                                data = dwellingTypeTraining,
                                laplace = 1)

dwellingTypeProbability <- predict(dwellingTypeModel,
                                   dwellingTypeTesting,
                                   type = "raw")
```

```
# printing the probability for each record in the testing dataset
print(dwellingTypeProbability)

dwellingTypePrediction <- predict(dwellingTypeModel,
                                  dwellingTypeTesting,
                                  type = "class")
# printing the prediction of dwellingType based on the model
print(dwellingTypePrediction)

# displaying the confusion matrix
dwellingTypeConfusionMatrix <- table(dwellingTypeTesting$DwellingType,
                                     dwellingTypePrediction)

print(dwellingTypeConfusionMatrix)
# displaying the predictive accuracy of the naive bayes model
predictiveAccuracy <- sum(diag(dwellingTypeConfusionMatrix)) /
  nrow(dwellingTypeTesting)

print(predictiveAccuracy)
```

Answers:

1. Naïve Bayes usually does not do a great job with lesser number of independent features. Here we have just 4 of them, which is not enough to get in a good accurate model.
2. This model has a predictive accuracy of 52.33%, which is low. The first course of action would be to improve the accuracy and then the mailing marketing company could use the model to predict if a customer is most likely to book a condo, an apartment or a home. Based on the prediction there could be targeted advertisements sent to the relevant customers.