Instructions: Answer each question as thoroughly as possible. Round answers to 4 decimal places as needed. Exact answers are best when possible. Be sure to answer all parts of each question.

1. In the file **325quiz5data.xlsx**, there is data related customers to have or have not tried a particular brand of lasagna. The file contains data on Age, Weight, Income, Pay Type, Car Value, Credit Card Debt, Gender, whether they live alone, Dwelling Type, number of Mall Trips, Neighborhood and whether they have tried the lasagna or not. Have Tried column has been recoded as Have_Tried_Dummy. Load this data into R, and create a logistic model using Income as the explanatory variable. You will need to convert Have_Tried_Dummy to a factor. Then use glm(Have_Tried_Dummy ~Income, data=mydata, family="binomial") (updating your data filename as needed). Use summary() to display the results.

   a. Describe your hypothesis test of the model. Is the coefficient for Income significant? ☑
   $H_0: \beta_1 = 0, \beta_1 \neq 0: H_a$ p-value = 0.00068 reject nulle   yes

   b. Create a plot of the data Income and Have_Tried_Dummy and plot the model on the graph (you should be able to do this inside ggplot).

   c. Interpret the coefficients of Income in the model in terms of odds of whether they have tried the lasagna or not.   income increases the odds of having tried the lasagna

   d. Redo the model with Age, Weight, Income, Car Value and Credit Card Debt in your model. Are all the variables significant? If not which ones?

   Income p-value 0.022   reject
   Age   " 2×10^{-16} reject
   Weight   "   0.00671   reject
   Car Value "   0.09930   fail to reject
   CC Debt   " 4×10^{-14} reject

   intercept "   0.85449   fail to reject

   Intercept and Car Value do appear to significantly affect the odds of trying the lasagna.