Logistic Regression Report

Name	Section	B.N	ID
Peter Atef	1	18	9202395
Beshoy Morad	1	19	9202405

1. Write the variable pairs that are not correlated at all to each other.

```r
if (FALSE) {
  "
          Price      Income          Age
  Price      1 0.00000000 0.00000000
  Income     0 1.00000000 0.09612083
  Age        0 0.09612083 1.00000000
  "

  # Variables that are not correlated at all to
each other are:
  # Price and Income
  # Price and Age
}
```

2. Are there any highly correlated variables in this dataset?

```r
# There are no highly correlated variables in this
dataset because the correlation values are not
close to 1 or -1.
```

3. How many categories are there for the Price variable?

```
# There are 3 categories for the Price variable:
10, 20, and 30.
```

4. Why is it divided into two entries only in the model?

```
• # The Price variable is divided into two entries
  because it is a categorical variable.
• # for n = 3 categories, you need n-1 dummy
  variables to represent the categories.
• # so we have 2 dummy variables for the Price
  variable.
```
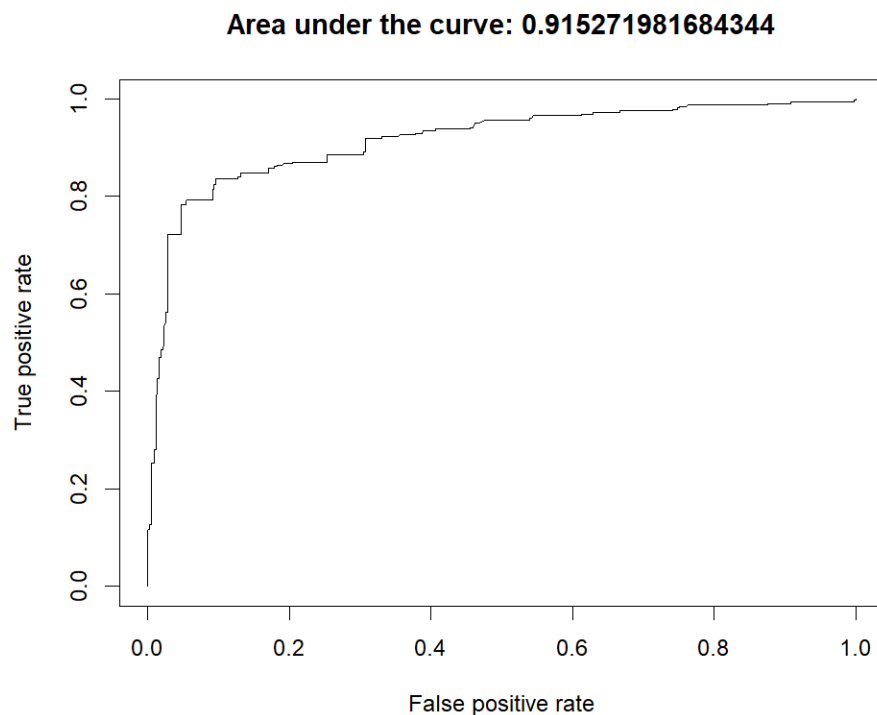
5. Q5: (Note: For this part, you may need to search and read about the ROC curve.)

   a. Write the value of AUC.

```
i. # 0.915272
```

   b. What is the maximum value of AUC (ideal case)?

```
i. # one
```

Area under the curve: 0.915271981684344

6. What does each point in the ROC graph represent?
   a. In other words, what is the value that changes and drives TPR
      and FPR to change too from one point to another in the graph?

```
● # The value that affects the TPR and FPR is the
  threshold value.
● # The threshold value is the value that the model
  uses to classify the data into classes.
● # We draw the ROC curve by changing the threshold
  value and calculating the TPR and FPR for each
  threshold value.
● # Then we choose the best threshold value that
  gives the largest AUC value.
```
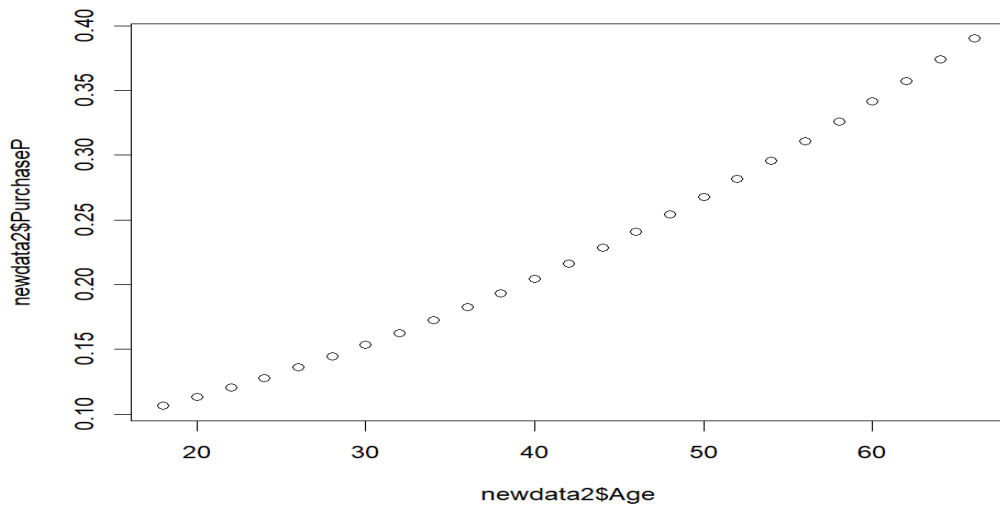
7. How is the predicted probability affected by changing only the Price
   holding all other variables constant?

```
●   "
●    Income    Age Price PurchaseP
● 1 42.492 35.976    10 0.6707408
● 2 42.492 35.976    20 0.4918407
● 3 42.492 35.976    30 0.1826131
●    "
●   #  As the price increases, the predicted
  probability of purchase decreases.
●   # Holding income and age constant, a higher price
  leads to a lower likelihood of purchase.
```

8. How is the predicted probability affected by changing only Age
   holding all other variables constant?

```
● # As the age increases, the predicted
  probability of purchase increases.
```

- # Holding income and price constant, a higher age leads to a higher likelihood of purchase.

newdata2$Age

9. How is the predicted probability affected by changing only Income holding all other variables constant?

- # As the income increases, the predicted probability of purchase increases.
- # Holding age and price constant, a higher income leads to a higher likelihood of purchase.

newdata3$Income