Lab 5

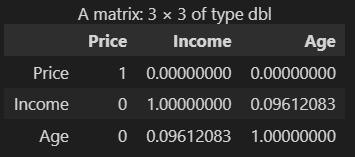
Predictive Analysis ||

Karim Mahmoud Kamal Sec: 2 BN: 12

Mustafa Mahmoud Hamada Sec: 2 BN: 25

Eng. Omar Samir

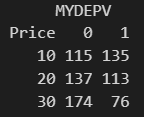
## Q1) Write the variable pairs that are not correlated at all to each other.

(Price, Age)

(Price, Income)

## Q2) Are there any highly correlated variables in this dataset?

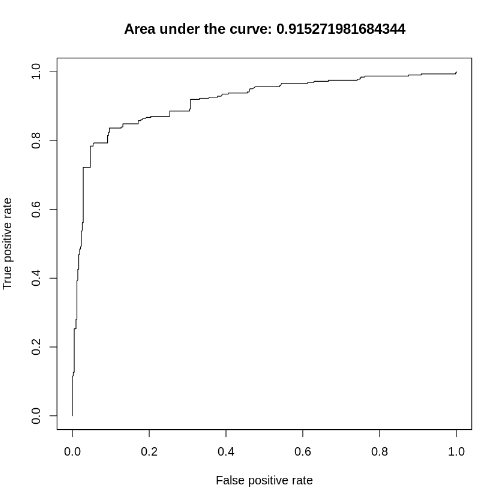
No, as all pair have a low correlation value

Q3) How many categories are there for the Price variable? There are 3 categories for price (10, 20, 30)

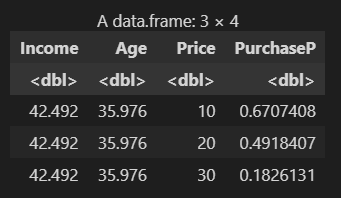
Q4) Why it is divided into two entries only in model? Because if we have n categories (which is 3 in this example) so it is divided into n-1 variables (which is 2).

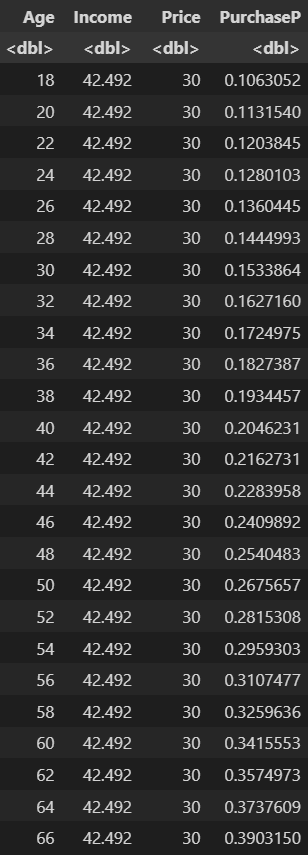
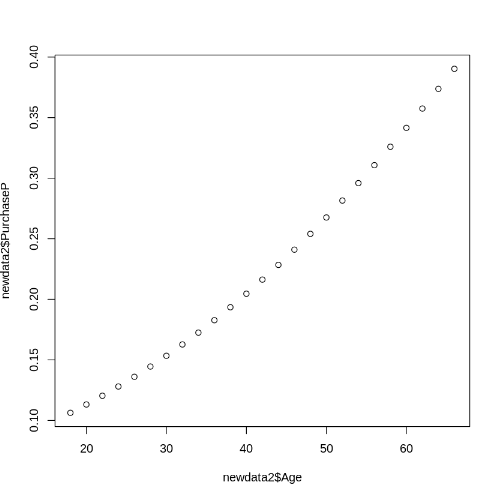
Q5.1) Write the value of AUC. Area under the curve = 0.915271981684344

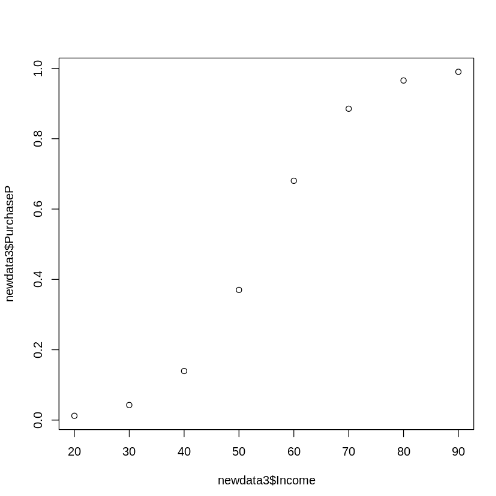
Q5.2) What is the maximum value of AUC (ideal case)? The maximum AUC is 1 so it maximizes the TPR and minimizes FPR



Q6) What does each point in the ROC graph represent? 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? Each point on ROC graph represents a different classification threshold. The classification threshold is the point at which a predicted probability or score is converted into a binary classification (positive or negative). A good model is characterized by a situation where, as we move from left to right along the threshold, the TPR approaches values near 1 with only a small change in the FPR.

Q7) How is the predicted probability affected by changing only Price holding all other variables constant? The probability of purchasing something more expensive decreases as the price increases.

Q8) How is the predicted probability affected by changing only Age holding all other variables constant? The predicted probability tends to increase as older clients are considered.

Q9) How is the predicted probability affected by changing only Income holding all other variables constant? The predicted probability increases as the income increases. This is logical because when income increases, clients are more likely to have more income and hence purchase with a higher probability.