

Q1 - (Price,income) and (price, age)

Q2 - No as (income, age) is the only correlated ~ 0.1 which very small ratio

Q3 - two => **20,30**

Q4 -because the price is categorical so the variable with n levels will be transformed into n-1 variables each with two levels. in our case Levels: 10 20 30 [n= 3] will transform to 2 variables

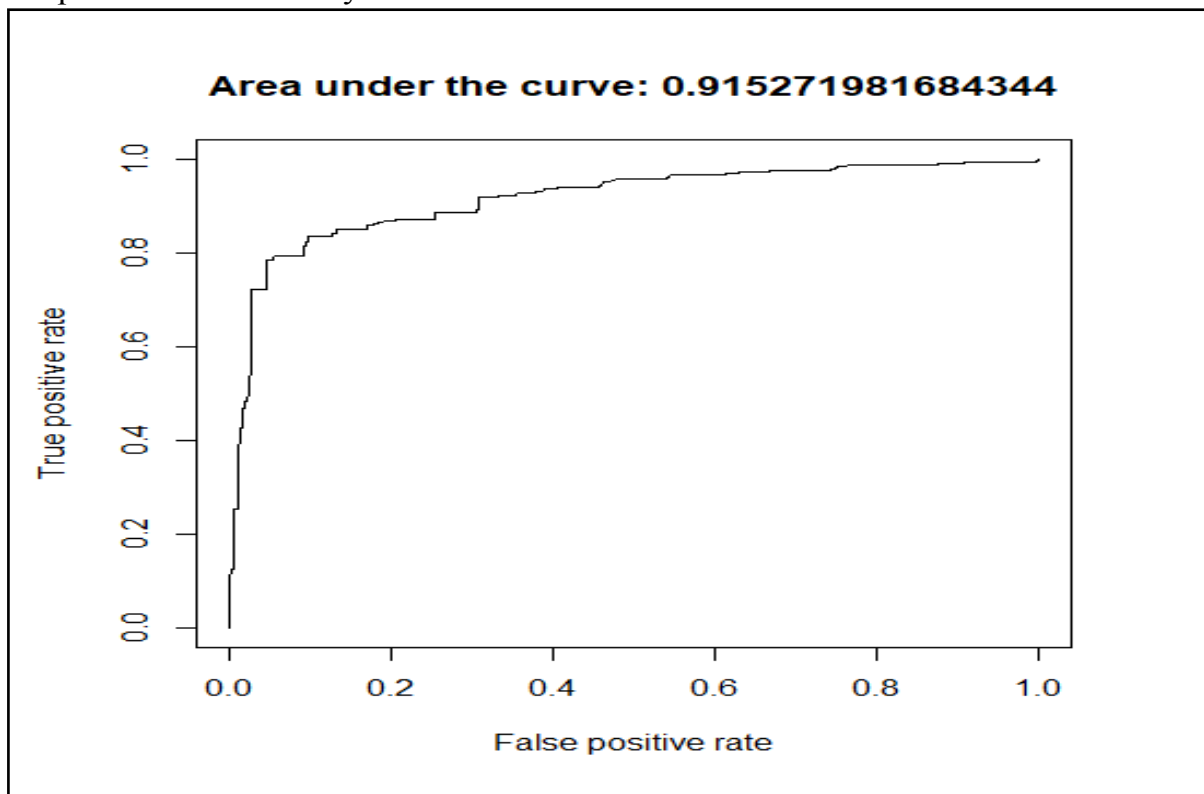
and we can visualizing in R using

```
> contrasts(as.factor(Price))
  20 30
10 0 0
20 1 0
30 0 1
```

Q5 - Write the value of this expression (just the number) ? **0.915272**

What is the maximum value of AUC (ideal case)? **1**

the plot x-axis FPR and y-axis is TPR



Q6 - x-axis is false positive rate (FPR) ,y-axis is true positive rate (TPR) each point indicates **FPR and TPR** for each threshold a higher X-axis value indicates a higher number of False positives than True negatives. While a higher Y-axis value indicates a higher number of True positives than False negatives. So, the choice of the threshold depends on the ability to balance between False positives and False negatives.

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? threshold taking for prediction in the binary classification

Q7 - How is the predicted probability affected by changing only price holding all other variables constant?

| Income | Age | Price | PurchaseP |
|--------|--------|-------|-----------|
| 42.492 | 35.976 | 10 | 0.6707408 |
| 42.492 | 35.976 | 20 | 0.4918407 |
| 42.492 | 35.976 | 30 | 0.1826131 |

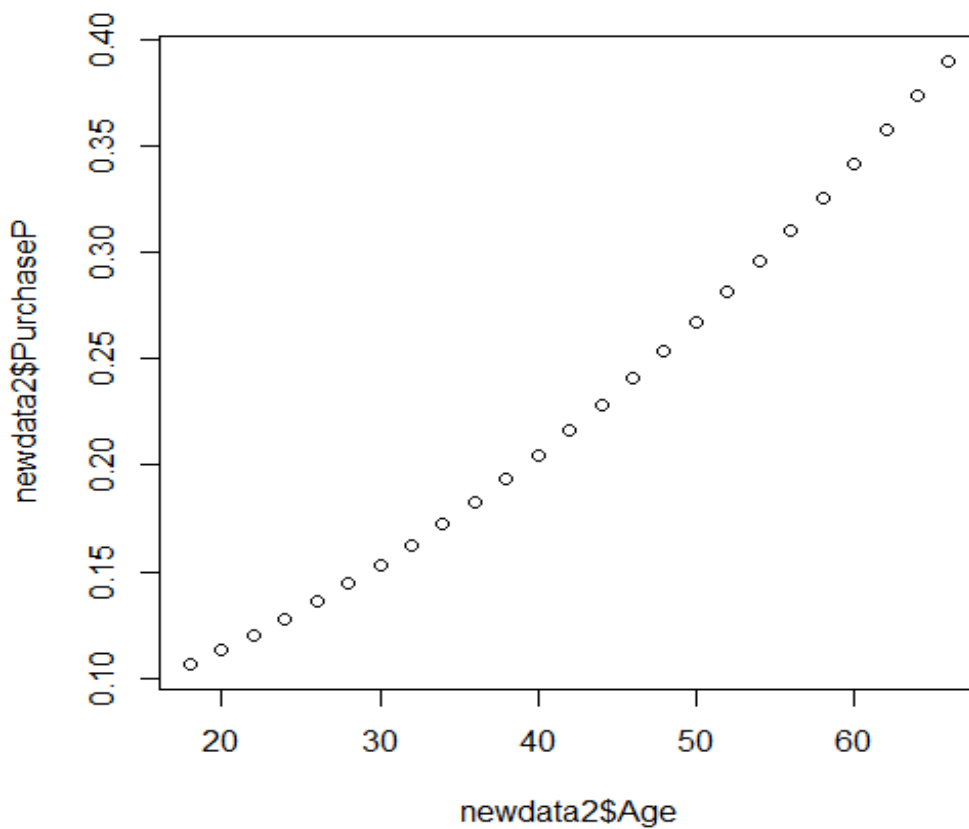
it indicates that the people whose is age ~36 and income ~43 will much likely purchase the lower price 10 rather than higher price 30
so it reflect /represent a specific segment of people on dataset

Q8 -How is the predicted probability affected by changing only age holding all other variables constant?

the plot indicates that people whose income is ~43 and the product price 30 will much more likely to purchase if they are older but with low probabiltly overall.

| Age | Income | Price | PurchaseP |
|-----|--------|-------|-----------|
| 18 | 42.492 | 30 | 0.1063052 |
| 20 | 42.492 | 30 | 0.1131540 |
| 22 | 42.492 | 30 | 0.1203845 |
| 24 | 42.492 | 30 | 0.1280103 |
| 26 | 42.492 | 30 | 0.1360445 |
| 28 | 42.492 | 30 | 0.1444993 |
| 30 | 42.492 | 30 | 0.1533864 |
| 32 | 42.492 | 30 | 0.1627160 |
| 34 | 42.492 | 30 | 0.1724975 |
| 36 | 42.492 | 30 | 0.1827387 |

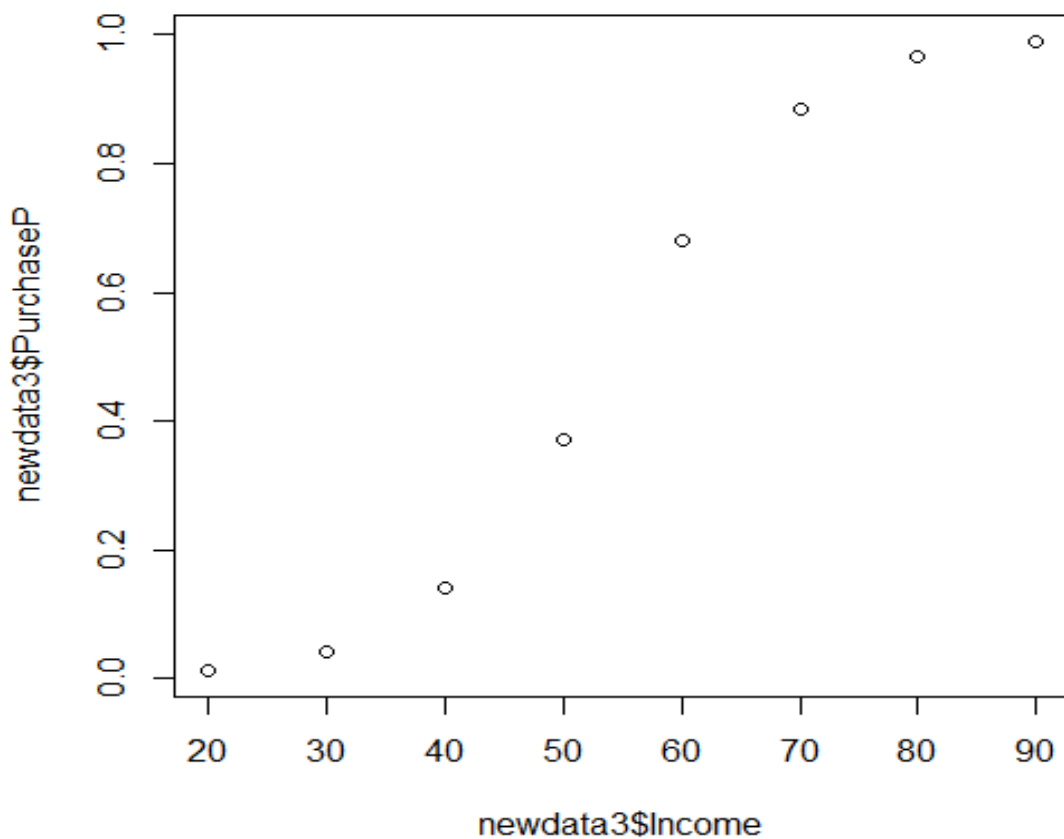
| | | | |
|----|--------|----|-----------|
| 38 | 42.492 | 30 | 0.1934457 |
| 40 | 42.492 | 30 | 0.2046231 |
| 42 | 42.492 | 30 | 0.2162731 |
| 44 | 42.492 | 30 | 0.2283958 |
| 46 | 42.492 | 30 | 0.2409892 |
| 48 | 42.492 | 30 | 0.2540483 |
| 50 | 42.492 | 30 | 0.2675657 |
| 52 | 42.492 | 30 | 0.2815308 |
| 54 | 42.492 | 30 | 0.2959303 |
| 56 | 42.492 | 30 | 0.3107477 |
| 58 | 42.492 | 30 | 0.3259636 |
| 60 | 42.492 | 30 | 0.3415553 |
| 62 | 42.492 | 30 | 0.3574973 |
| 64 | 42.492 | 30 | 0.3737609 |
| 66 | 42.492 | 30 | 0.3903150 |



Q9 -How is the predicted probability affected by changing only income holding all other variables constant?

that the people whose age ~36 with product's price 30 will be more likely to purchase if their income is higher or equal of 60\$ and 60 \$ is depending on threshold 0.5 meaning that higher income ≥ 60 will give high probability of purchase

| Income | Purchase |
|--------|------------|
| 20 | 0.01219091 |
| 30 | 0.04281102 |
| 40 | 0.13948050 |
| 50 | 0.37004640 |
| 60 | 0.68039246 |
| 70 | 0.88525564 |
| 80 | 0.96546923 |
| 90 | 0.99022745 |



and that is final prediction if purchase depend on all variables price , age , income

| Age | Income | Price | Prob |
|-----|--------|-------|-----------|
| 51 | 47 | 20 | 0.7454717 |
| 62 | 18 | 20 | 0.0933234 |
| 39 | 96 | 30 | 0.9959175 |
| 39 | 61 | 30 | 0.7291606 |
| 50 | 45 | 30 | 0.3353513 |
| 50 | 62 | 20 | 0.9512435 |
| 45 | 39 | 20 | 0.4586317 |
| 54 | 44 | 20 | 0.6885863 |
| 54 | 82 | 10 | 0.9983910 |
| 24 | 82 | 10 | 0.9954071 |