Big Data

Logistic Regression

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Sec** | **BN** | **ID** |
| Sara Gamal Gerges | 1 | 20 | 9210455 |
| Eman Ibrahim | 1 | 14 | 9210265 |

## Q1)

A white screen with black text

AI-generated content may be incorrect.

|  |
| --- |
| A screenshot of a computer program  AI-generated content may be incorrect.  [1] "Price and Income are not correlated"  [1] "Price and Age are not correlated" |
|  |
| |  | | --- | |  | |

## Q2)

No, there is no highly correlated variables because no correlation value is close to 1 or -1.

## Q3) A close-up of a white background AI-generated content may be incorrect.

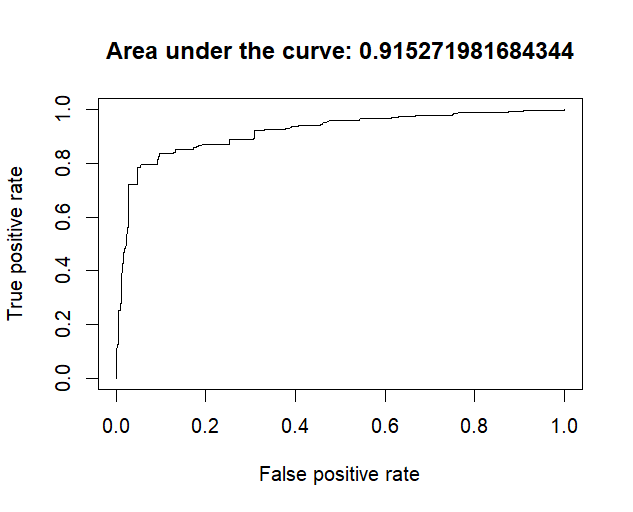
## Q4)

## Q5)

1. Write AUC value 0.915272
2. Maximum value of AUC is 1 (ideal case)

**AUC ranges from 0 to 1, where:**

* **1** → Perfect classification (ideal case).
* **0.5**→ Random guessing (no predictive power).
* **< 0.5** → Worse than random (model is making wrong predictions more often).



## Q6)

Each point in the ROC curve represents a different threshold used by the classifier.

**What changes from one point to another?**

The value that changes and drives both **True Positive Rate (TPR)** and **False Positive Rate (FPR)** is the **classification threshold**.

**How does the threshold affect TPR & FPR?**

* **Lower threshold → More positives predicted**
  + **Higher TPR**
  + **Higher FPR**
* **Higher threshold → More negatives predicted**
  + **Lower TPR**
  + **Lower FPR**

## Q7)

A screenshot of a computer

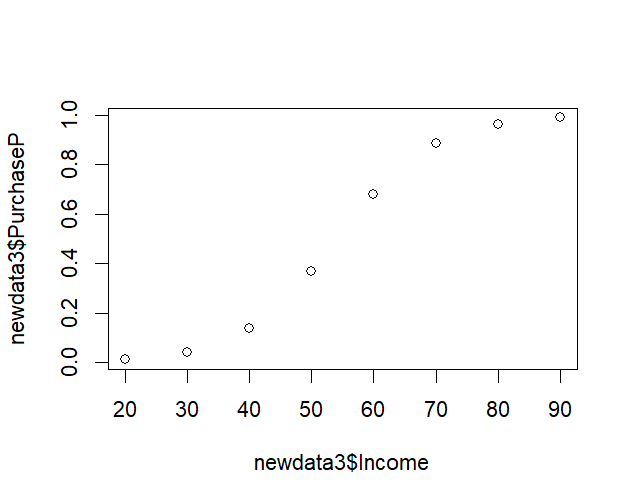
AI-generated content may be incorrect.

As the price increases the predicted probability of purchase decreases.

## A screenshot of a computer code AI-generated content may be incorrect.Q8)

It ‘s clear that as the age increases the probability of purchase increases.

## Q9)

A screenshot of a computer code

AI-generated content may be incorrect.

It ‘s clear that as the income increases the probability of purchase increases.