## Commentary on the Assignment Given

### Solution 1

Table 1

| | Time | Age | Year | Thickness | Sex | Status | Ulcer |
|---|---|---|---|---|---|---|---|
| **Min** | 10 | 4. | 1962 | 0.100 | 0.000 | 1.000 | 0.000 |
| **1st Q** | 1525 | 42 | 1968 | 0.970 | 0.000 | 1.000 | 0.000 |
| **Median** | 2005 | 54 | 1970 | 1.940 | 0.000 | 2.000 | 0.000 |
| **Mean** | 2153 | 52.46 | 1970 | 2.920 | 0.385 | 1.790 | 0.439 |
| **3rd Q** | 3042 | 65 | 1972 | 3.560 | 1.000 | 2.000 | 1.000 |
| **Max** | 5565 | 95 | 1977 | 17.42 | 1.000 | 3.000 | 1.000 |
| **S. D** | 1122 | 16.67 | 2.575 | 2.959 | 0.487 | 0.551 | 0.497 |

From Table 1 it can been seen that 38% of the patient in the data are male, about 44% of the tumor was ulcerated, more than 50% of the patient survived after the operation. The average age of the patient that took the operation is 52

### Solution 2

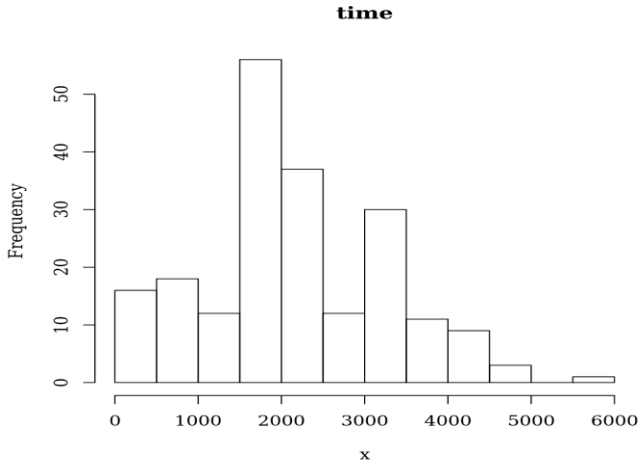


Figure 1

Distribution of survival time in days since the operation

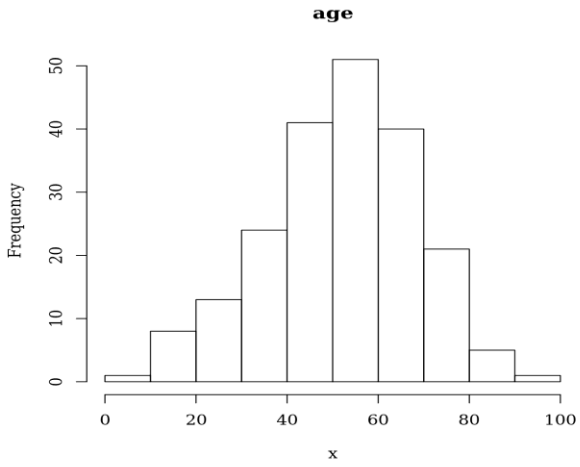


Figure 2

Distribution of age in years at the time of the operation

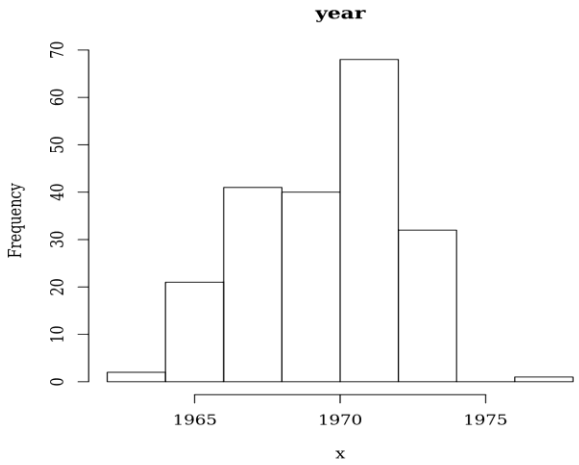


Figure 3

Distribution of year of operation

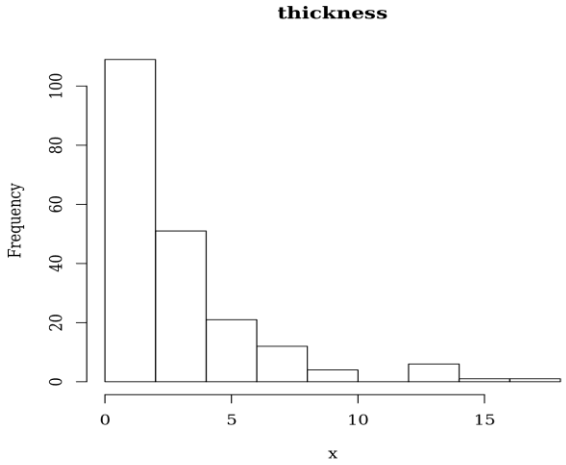


Figure 3

Distribution of tumour thickness in mm

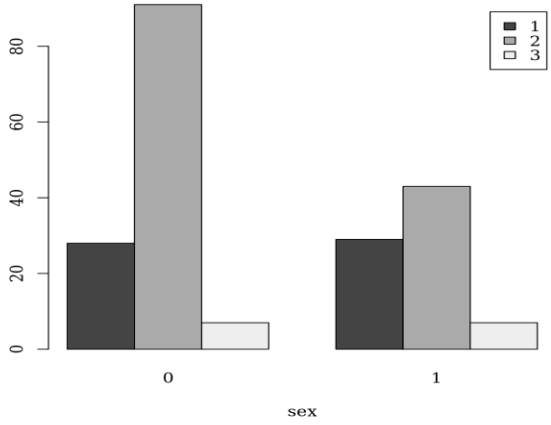


Figure 4

Bar-chart of sex showing the survival status of each gender

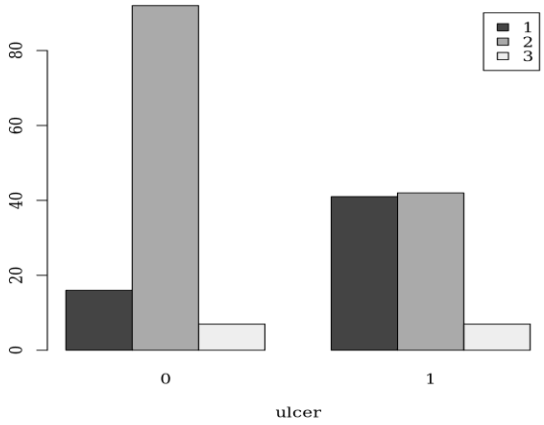


Figure 5

Bar-chart of patient whose tumour was ulcerated or not, showing their survival status.

**Solution 3 and 4**

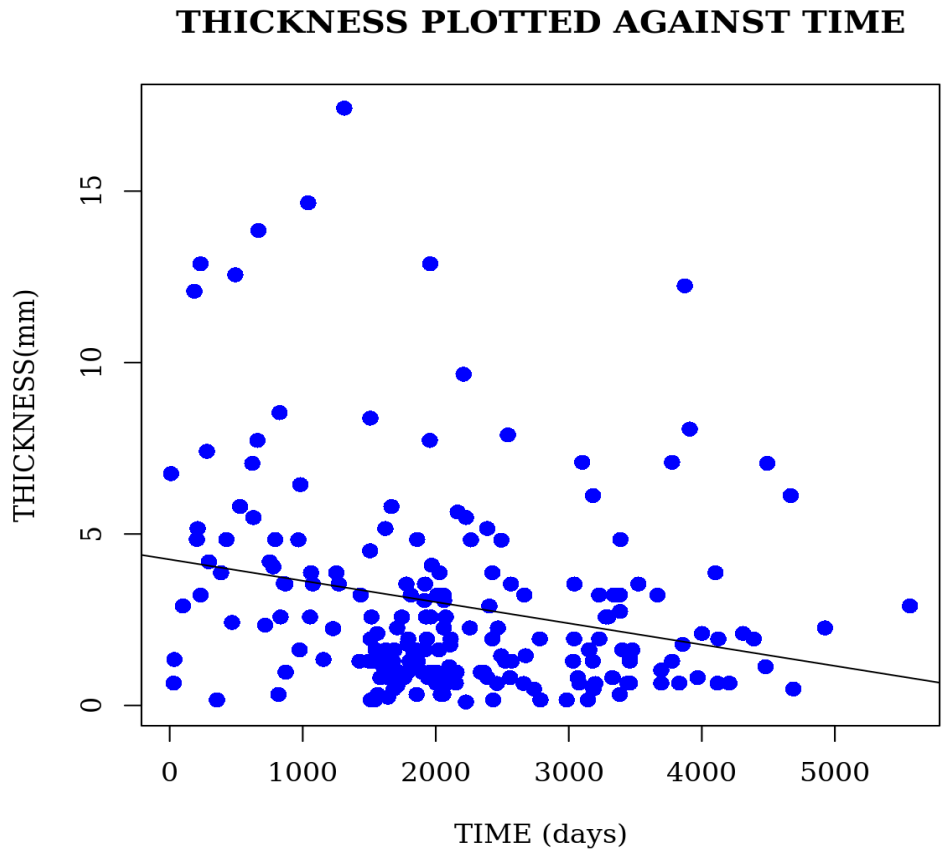


Figure 6

From this figure it can be seen that tumour thickness (mm) has a **Negative Relationship** with survival time in days since the operation with R-square value of 0.055 and Pearson Correlation of -0.235

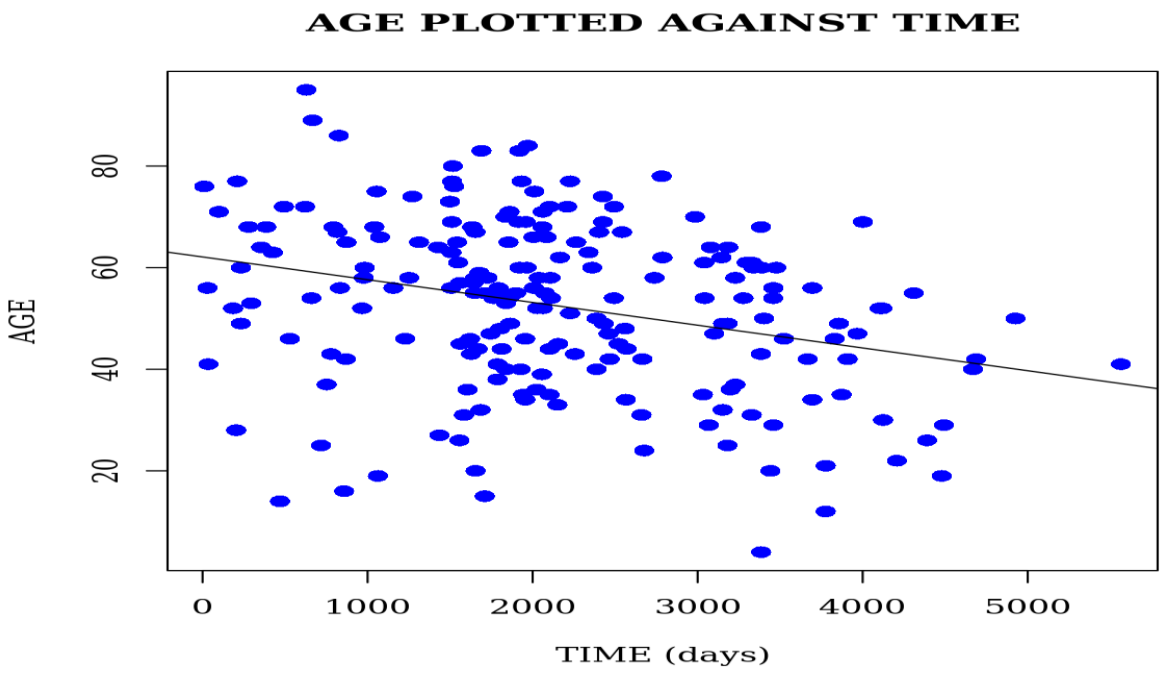


Figure 7

From this figure it can be seen that age in years at the time of the operation has a **Negative Relationship** with survival time in days since the operation with R-square value of 0.09 and Pearson Correlation of -0.302

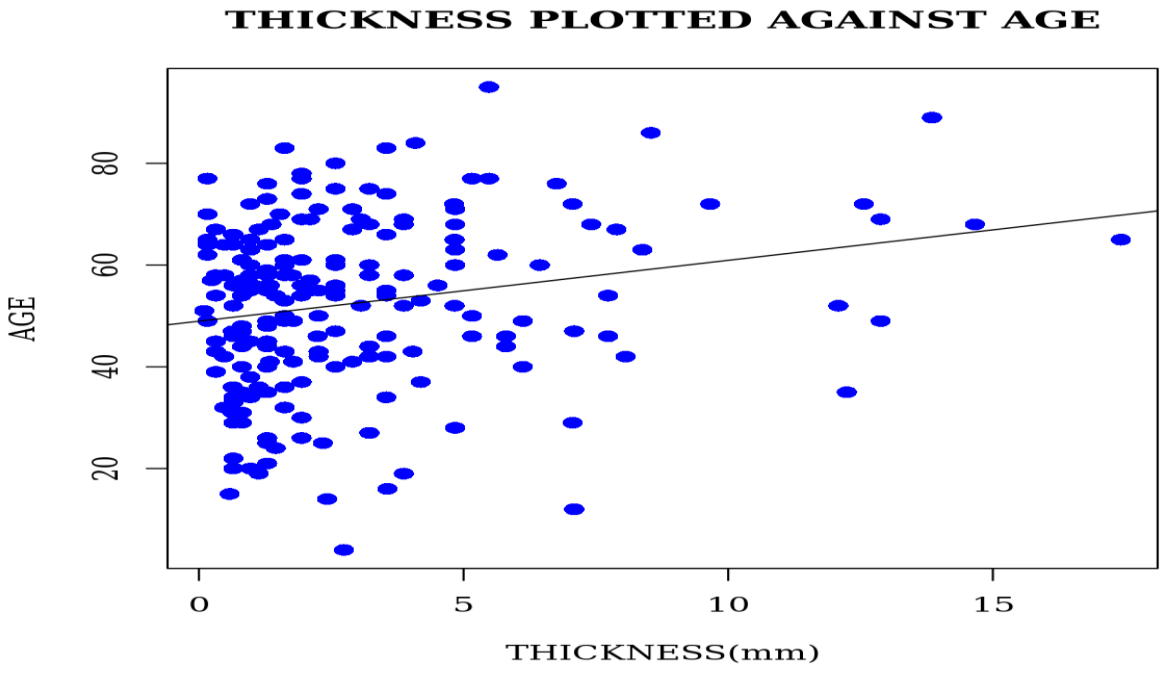


Figure 8

From this figure it can be seen that tumour thickness (mm) has a **Positive Relation** with the age in years at the time of the operation with R-square value of 0.05 and Pearson Correlation of 0.212

## Solution 5

After grouping the data into two sample set based on the patient gender, T-test was carried out. T-tests are handy hypothesis tests in statistics when we want to compare means. We can compare a sample mean to a hypothesized or target value using a one-sample t-test.

From the result of the t-test carried out there was a high similarity between the gender group.

## Solution 6

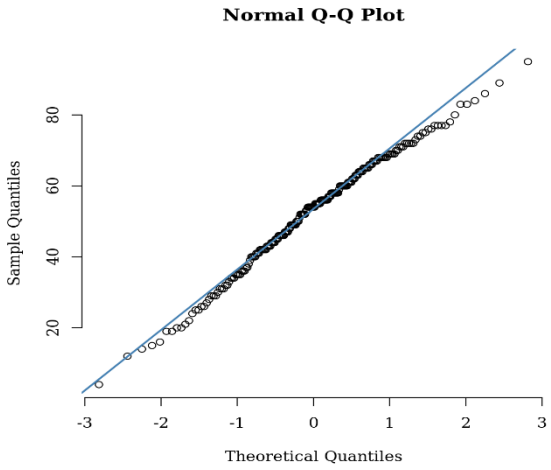


Figure 9

Q-Q Plot for age, this graph indicates that age in years at the time of the operation is **Normally Distributed**

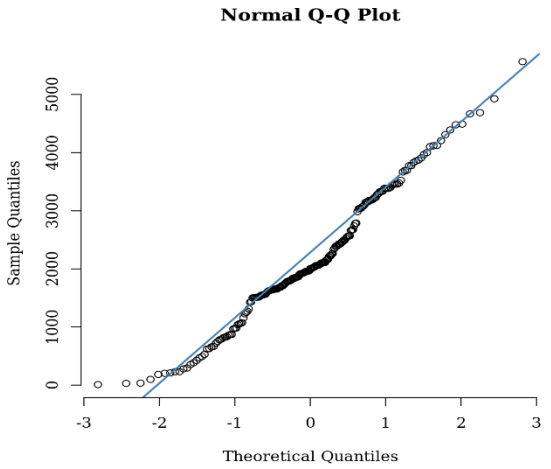


Figure 10

Q-Q Plot for time, this graph indicates that the survival time in days since the operation is **Normally Distributed**

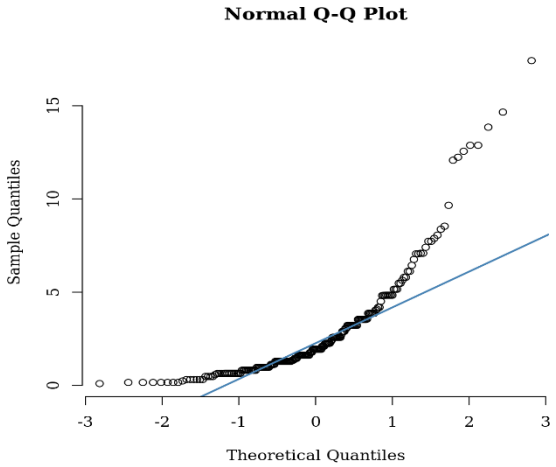


Figure 11

Q-Q Plot for tumour thickness, this graph indicates that the tumour thickness in mm. is **Exponentially Distributed**

## Question 7

In conclusion, it appears that patients with ulcerated tumors were older, more likely to be male, and had thicker stage tumors. Include these variables in the model .