**Section A**

**Exercise 1:**

The correlation coefficient is equal to **0.7917762**

**Exercise 2:**

The average income is **47682.1**

**Exercise 3:**

536.6094

**Exercise 4:**

68.75%

**Exercise 5:**

44

**Exercise 6:**

34.00504%

**Exercise 7:**

34734.6347380095 USD

**Exercise 8:**

554.286363

**Exercise 9:**

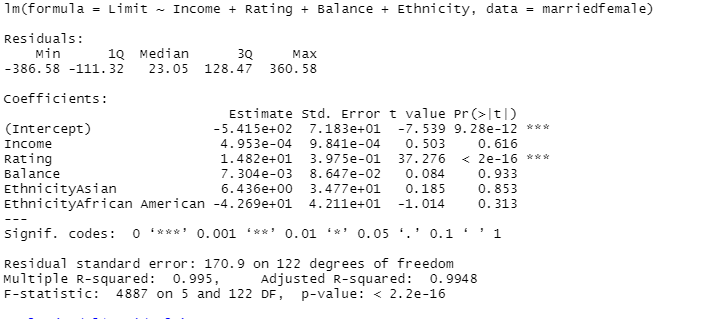
47

**Exercise 10:**

3480.5274060401

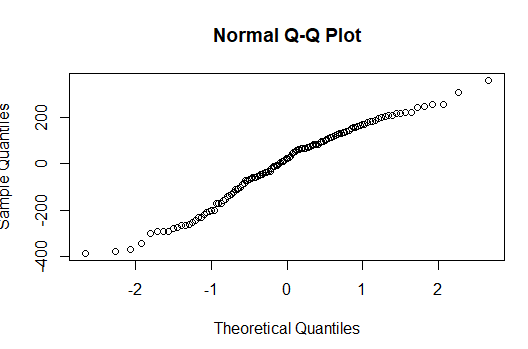
**Section B**

**Exercise 2:**

****

1. Limit = beta0 + beta1\*Income + beta2\*Rating + beta3\*Balance + beta4\*Ethnicity   
   Limit (y hat) = -541.5413 + 0.0004952687 \* Income + 14.81807 \* Rating + 0.007304326 \* Balance + 6.43607 \* EthnicityAsian - 42.68517 \* EthnicityAfrican American
2. Income: All things being constant a rise in income of one dollar will increase the limit by 0.0004952687  
   Rating: Increasing rating by one unit will increase the limit by 14.81807  
   Balance: When balance goes up by one dollar, the credit limit will go up by 0.007304326  
   EthnicityAsian: If the ethnicity of a person is a asian the credit limit will go up by approximately $6.43607  
   EthnicityAfrican American: When a person is an African American his/her limit is less by approximately $42.68517
3. …
4. Hypothetic testing.  
   Ho: beta1 = beta2 = beta3 = beta4 = 0  
   Ha: at least one of betas is not equal to 0  
   Rejection region:   
   We reject Ho if the p-value is less than alpha  
   P – value = 0.000000000….022  
   α = 0.05  
   Decision: Since p-value is far less than alpha, or to be more precise it is almost 0, we reject Ho and make a conclusion that our model is significant.
5. I have the following numerical values for my module: Income, Rating and Balance.  
   The following are the hypothesis for them  
   For income: For rating: For balance:  
   Ho: beta1 = 0 Ho: beta2 = 0 Ho: beta3 = 0  
   Ha: beta1 not equal to 0 Ha: beta2 not equal to 0 Ha: beta3 not equal to 0  
   Decision: Reject Ho if,   
    p-value < α   
   P values for income and balance are greater than alpha. Their values are 0.616 and 0.933 respectively. The alpha is 0.05. Hence, we fail to reject Ho and conclude that these factors are insignificant for my model. Regarding the rating, its p-value is almost equal to zero. Thus, we reject Ho and make a conclusion that its value is significant for the model.

It can be seen that distribution is normal



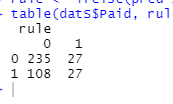
No Multicollinearity:

To check this assumption we need correlation coefficients.

* Between income and rating: 0.8034959
* Between income and balance: 0. 4539398
* Between rating and balance: 0. 8481232

There are violations. Coefficient for rating and balance, and income and rating are high, they ARE HIGHER THAN 80%

**Exercise 5:**



The confusion matrix.

Accuracy: (235+27)/(108+27+27+235) = 0.65994962The model is not highly accurate

Sensitivity: 27/ (27+108) = 0.2 The proportion of all 1s that are correctly classified as 1

Specificity: 235/(235+27) = 0.89694656 The proportion of all 0s that are correctly as 0s

Precision: 27/(27+27) = 0.5 The proportion of predicted 1s that are actually 1s