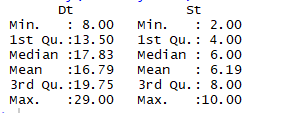
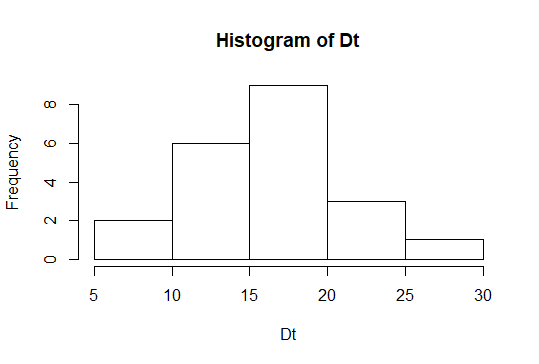
# To Predict delivery Time using Sorting time:

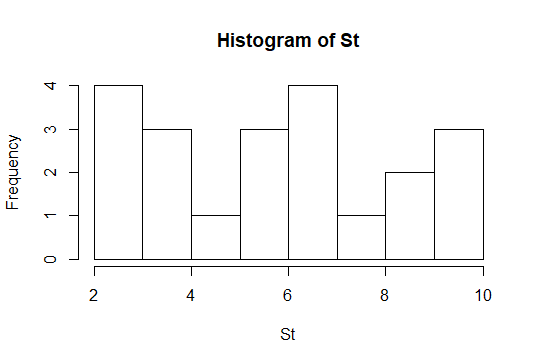
1. Data Collection: delivery\_time.csv
2. EDA:

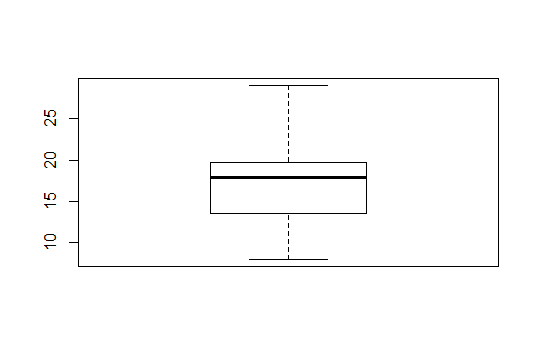


Standard Deviation(Dt): 5.07

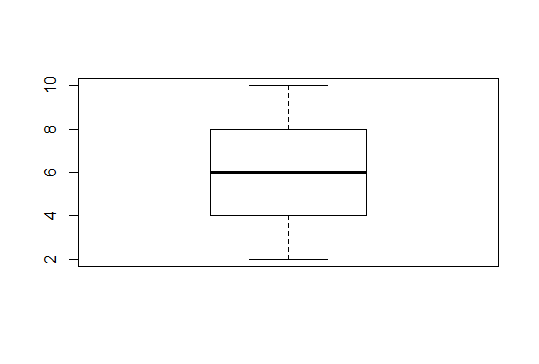
Standard Deviation(St): 2.54



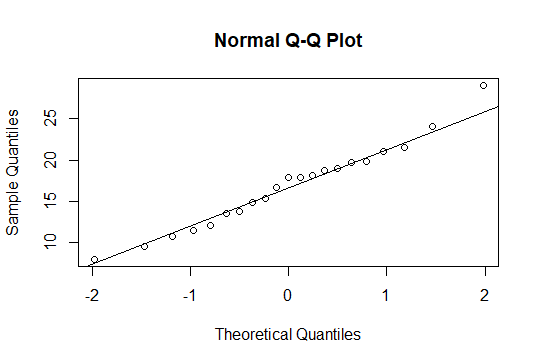




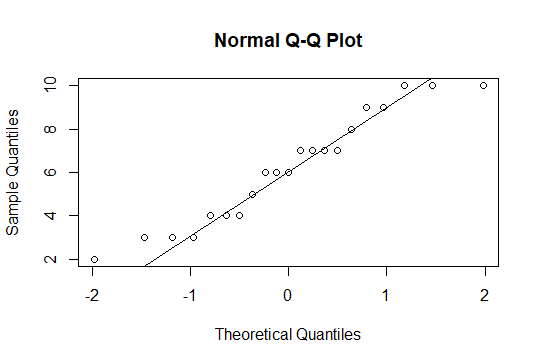
Box Plot for Delivery Time(No Outliers)



Box Plot for Sorting Time(No Outliers)



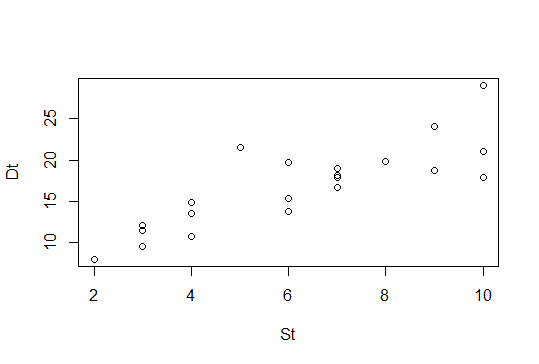
Q-Q Plot for Delivery Time(Normal)



Q-Q Plot for Sorting Time(Not Normal)

Since Sorting Time is not Normal, so we have to transform our model accordingly.

Scatter Plot :

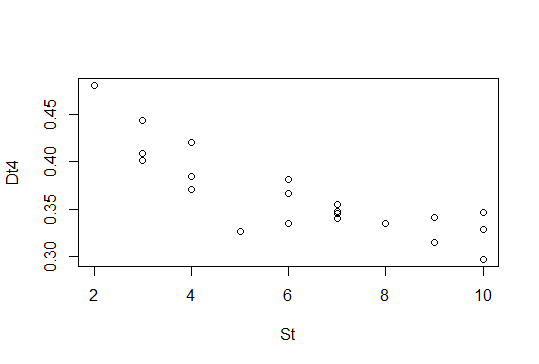


The Scatter plot Suggests:

* Direction is positive
* Heteroscedastic
* Linear

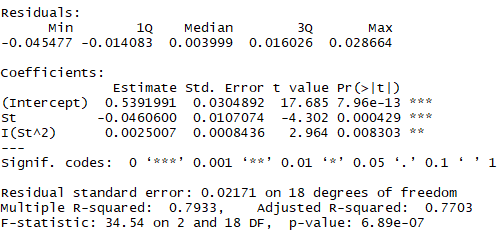
1. Data Mining:

After Transformation our Scatterplot is:



The scatter plot suggests:

* Non linear
* Negative correlation
* Homoscedastic



B0 = 0.54 is significant

B1 = -0.05 is significant

B2 = 0.003 is significant

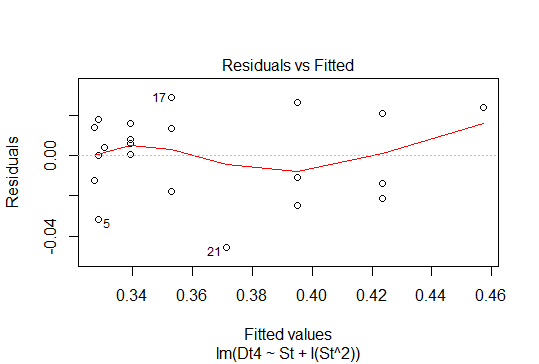
R2 value = 0.79

F-statistic and p- value suggests our model is good.

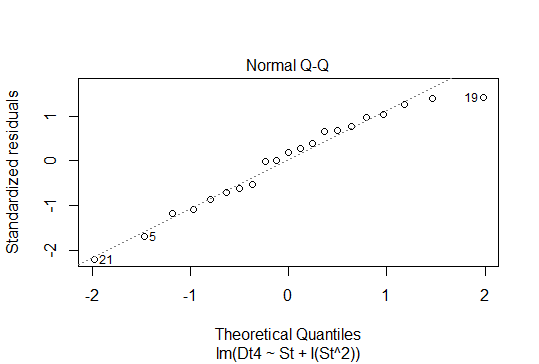
From our R code , we can come to a final model as:

1/log(Delivery Time) = 0.54 – 0.05\*Sorting time+ 0.003\* (Sorting Time)^2

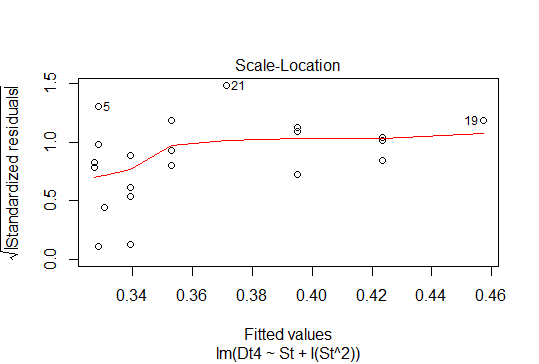
1. Line Assumptions:



Residuals Vs Fitted Plot Shows that residuals are Independently Distributed



The above Plot shows that errors Are Normally Distributed



The above Plot suggests that there is no Heteroscedasticity Problem

|  |  |
| --- | --- |
| Model | R2 Value |
| M1 | 0.68 |
| M2 | 0.71 |
| M3 | 0.70 |
| M4 | 0.70 |
| M5 | 0.78 |
| M6 | 0.79 |