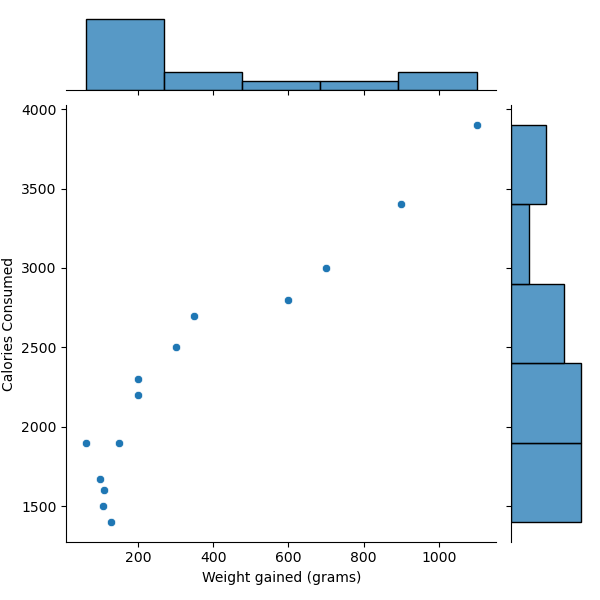
Single Linear regression

1. )Cal Consumed

Cals is used to predict Weight

There is a heft correlation between X and y .

The Scatter plot with linear regression line appears with less RMSE if we use normal SLR with no transformation technique

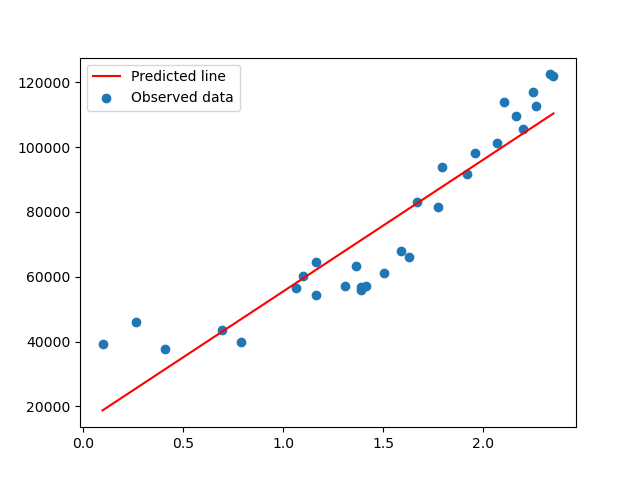
MODEL RMSE

0 SLR 103.302502

1 Log model 141.005382

2 Exp model 118.045157

3 Poly model 117.414500



2.)Delivery Time predicted by Sorting time

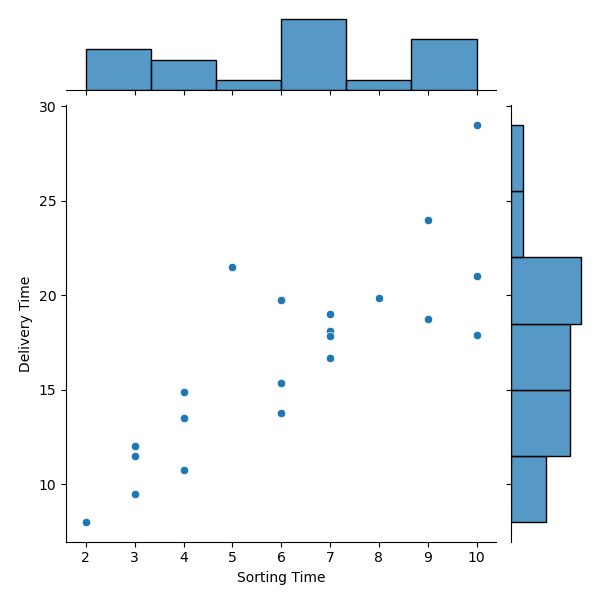
Corr btw both X ( Sorting time ) and y (Del time) is 0,85

MODEL RMSE

0 SLR 2.79165

1 Log model 2.79165

We can use Log technique to achieve low RMSE and check for SSE for Train and test to see which performed better.

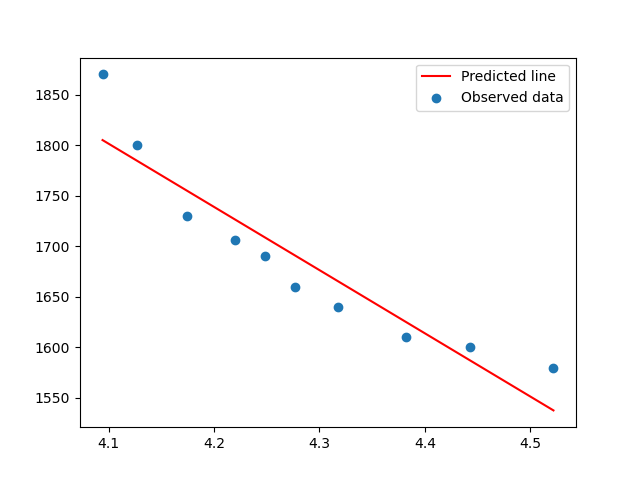


--> 12002150006.1002 for Test Data

9603211653.488464 --> A lot of RMSE in Train data but better at Test data

3.)Emp Data

Downward negative correlation observed between x and y



The predicted line by LOG transformation technique seems to have given good R2 value for this best fit line and low RMSE.

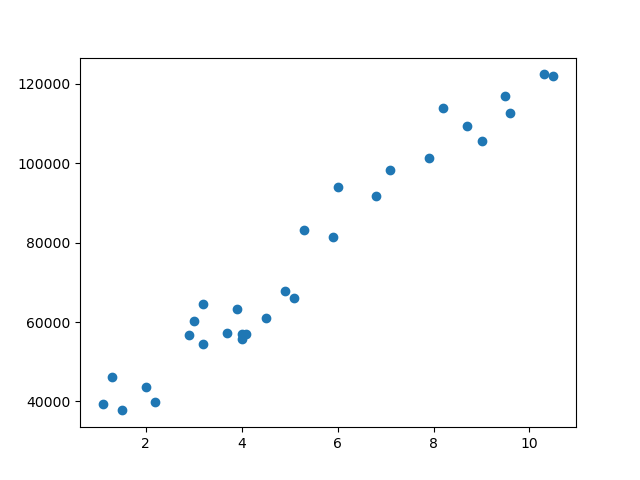
MODEL RMSE

0 SLR 35.892635

1 Log model 31.069521

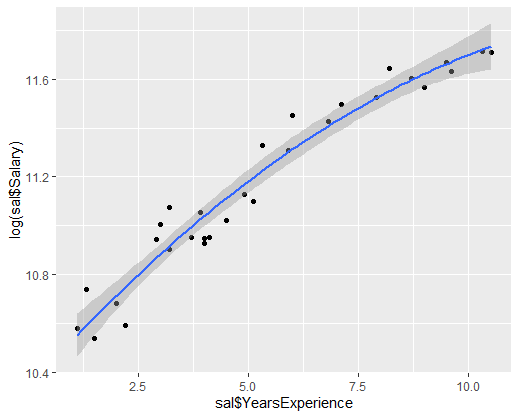
2 Exp model 34.268550

4.) Salary Hike



X and y have high amount or correlation between each other .

Y is dependent on X to prove its variability.



MODEL RMSE

0 SLR 5592.043609

1 Log model 10302.893706

2 Exp model 7213.235077

3 Poly model 5391.081583

Low RMSE observed in Polynomial Transformation so we will take that model for best fit line.

5849.0417615844335 --> RMSE fro Test Data

5235.602483490605 --> less training error as low RMSE in Train data . A case of slight Overfitting