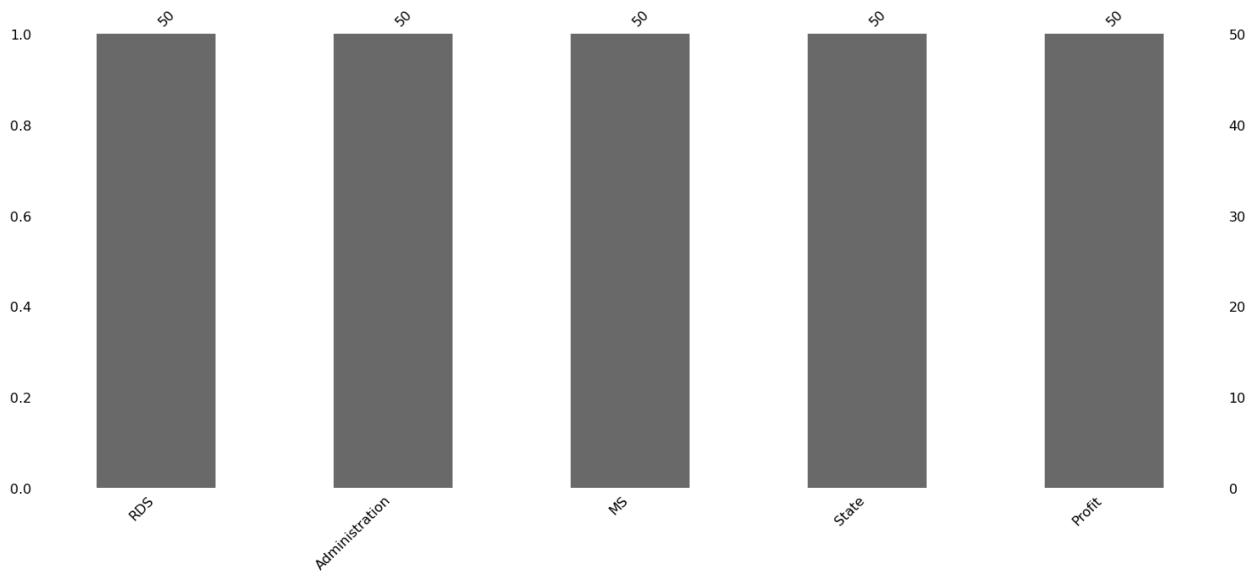
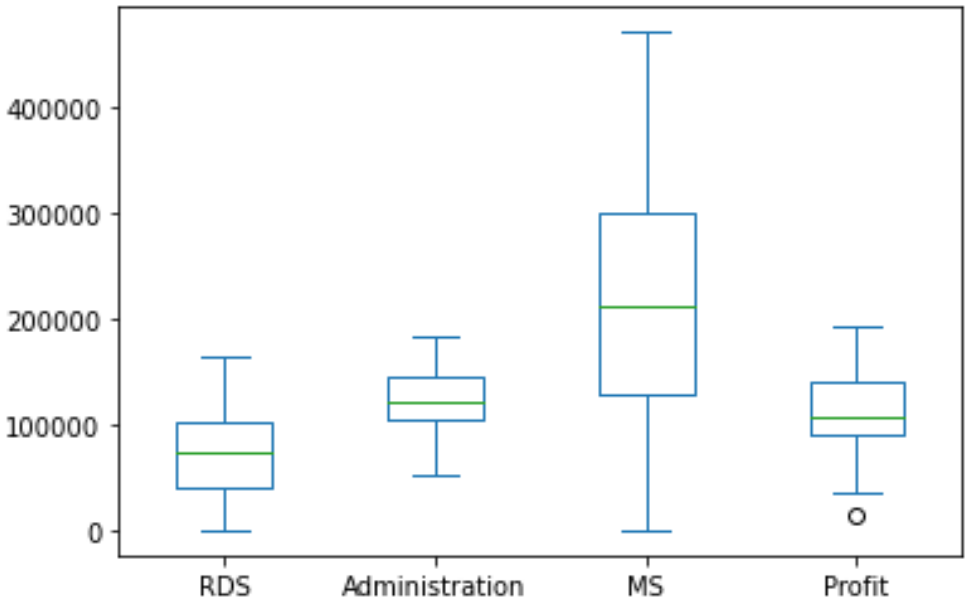
Different Visual Plots for Model building

1) Using bar graph to check missing value:



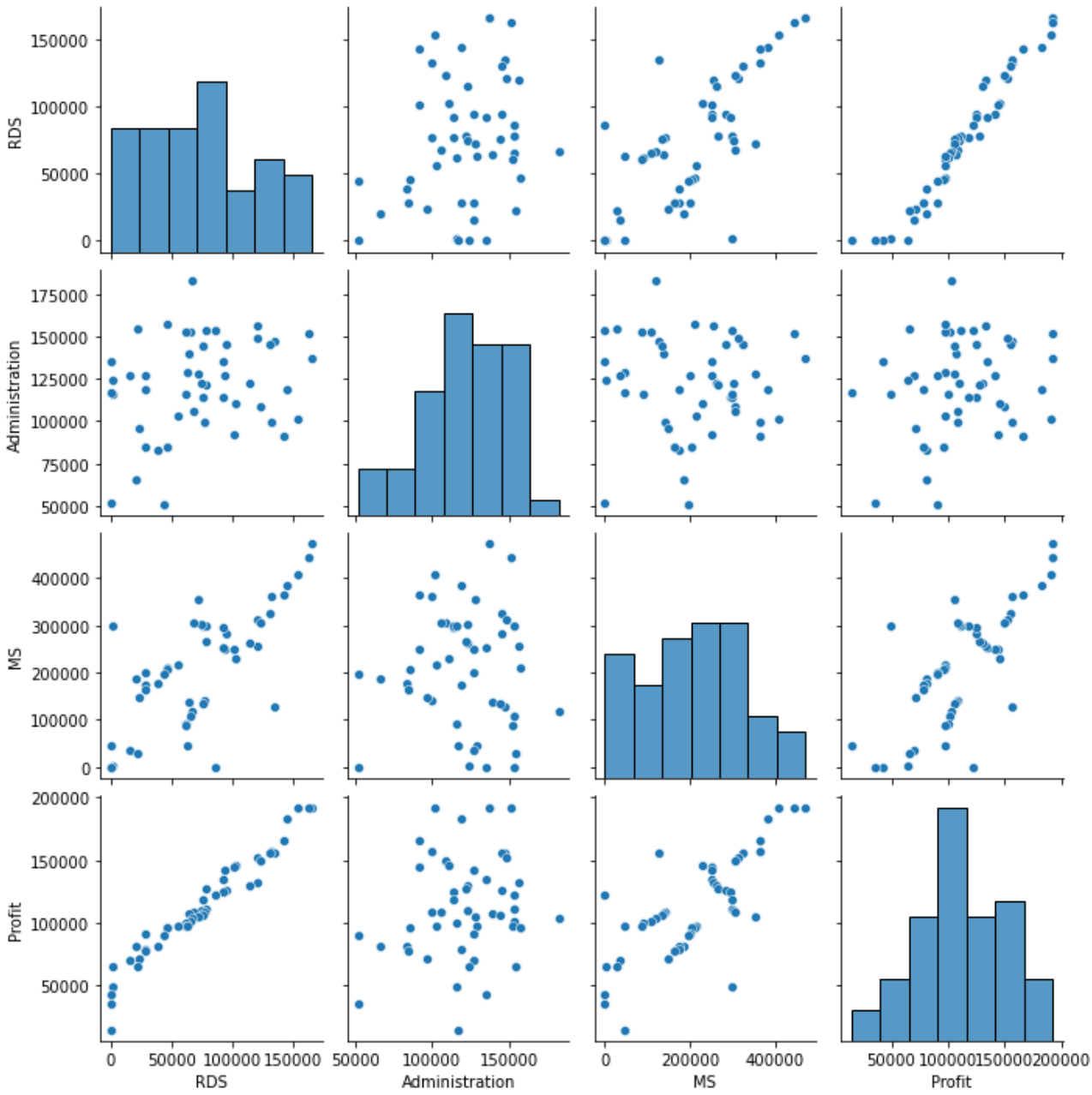
Conclusion – There is no Null Values

2) Box Plot for Outlier Detection:



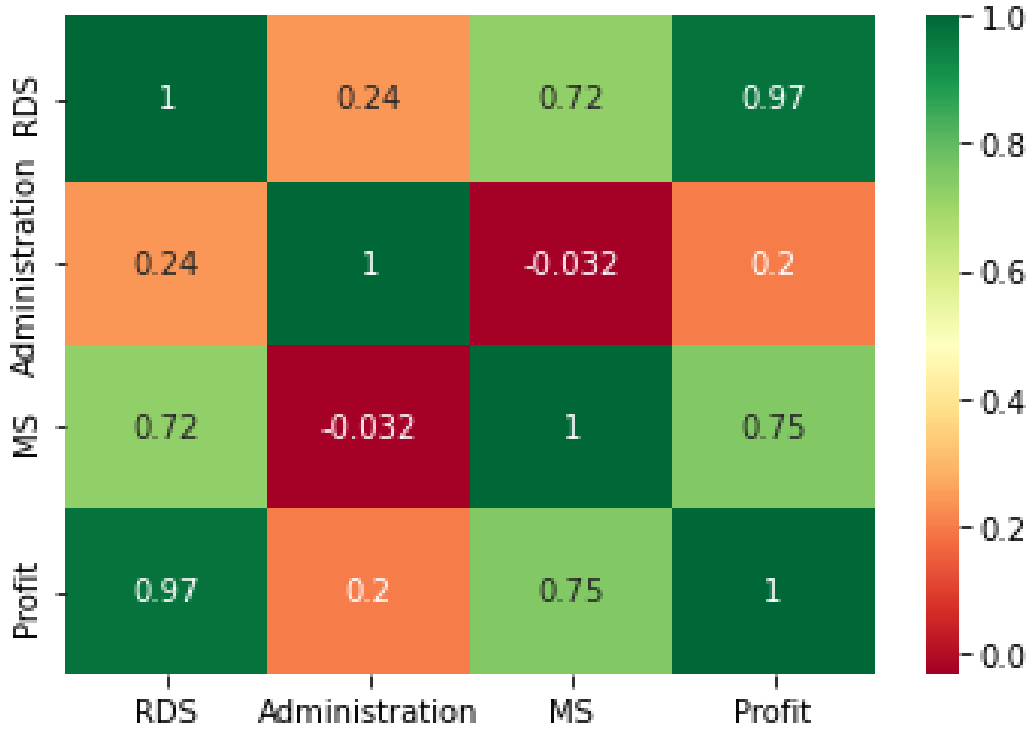
Conclusion – Only one outlier in profit column

3) For checking Correlation used Pair plot :



Conclusion – Profit and RDS have correlation

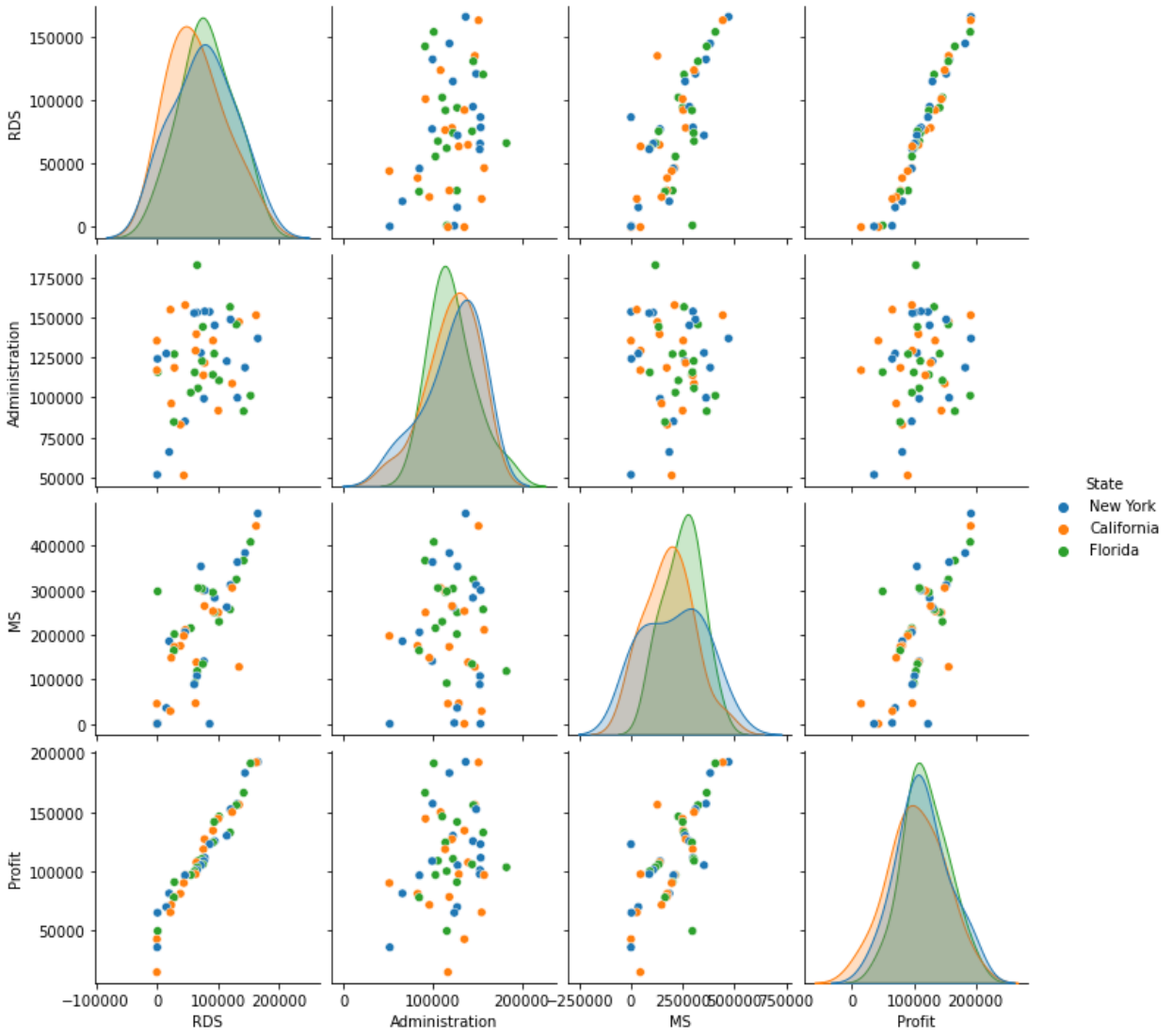
4) Heat map for correlation:



Conclusion – Indicates Profit and RDS have strong positive correlation which is approx. 0.97

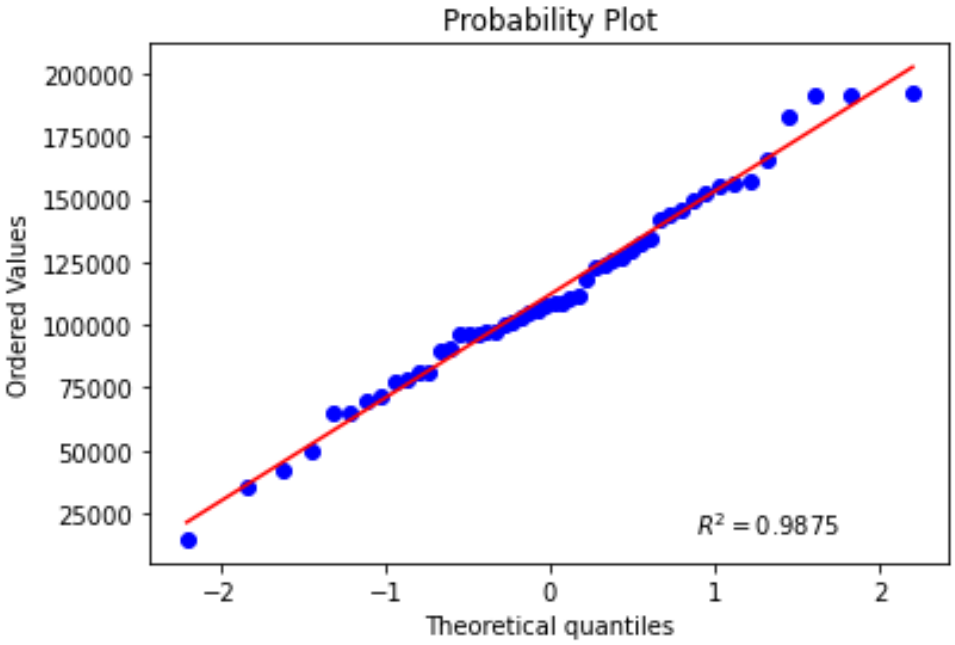
MS and Administration have negative correlation which approx. -0.032

5) Pair plot with hue on state :



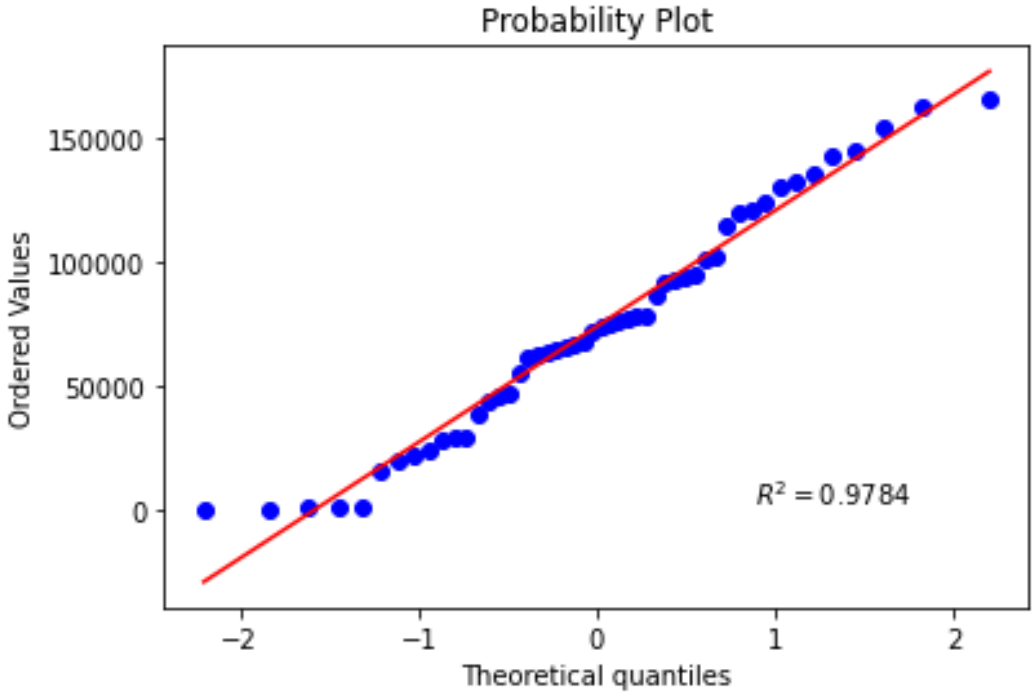
Conclusion – Profit and RDS have correlation

6) Probability Plot for Profit



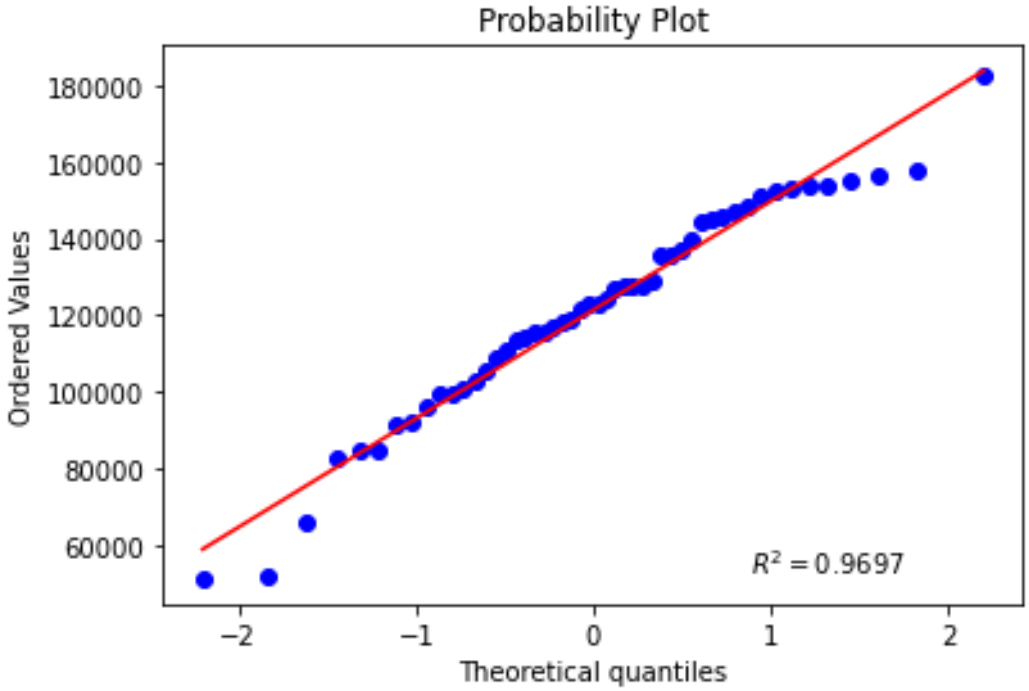
Conclusion – R square value for Profit is 0.9875

7) Probability Plot for RDS



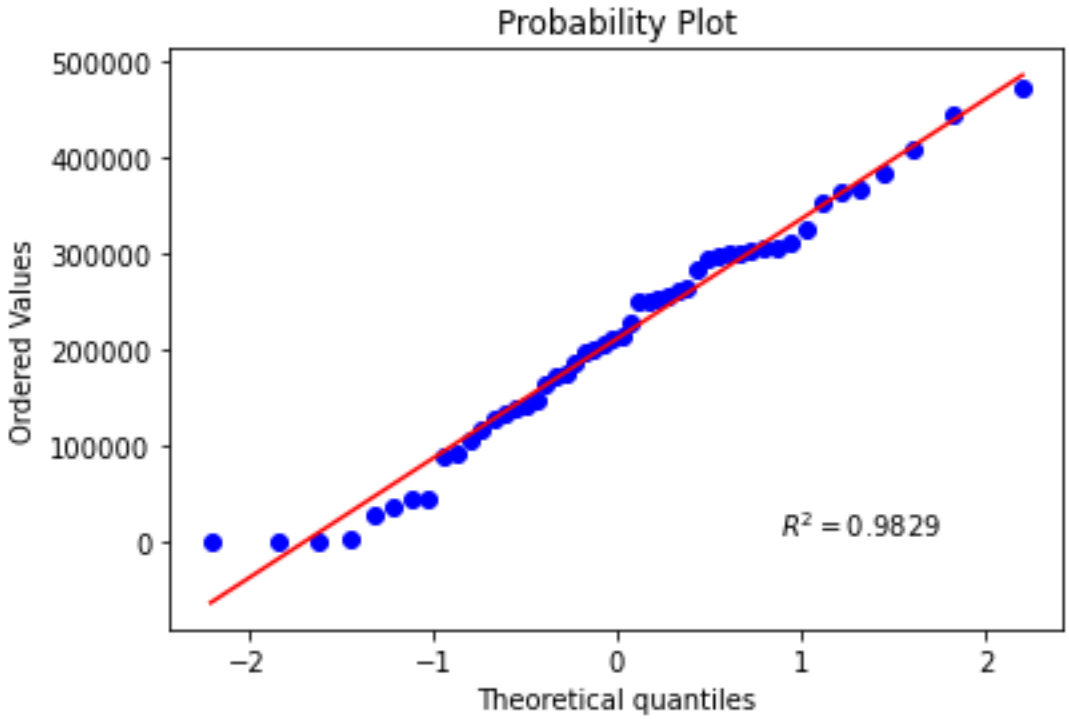
Conclusion – R square value for RDS is 0.9784

8) Probability Plot for Administration



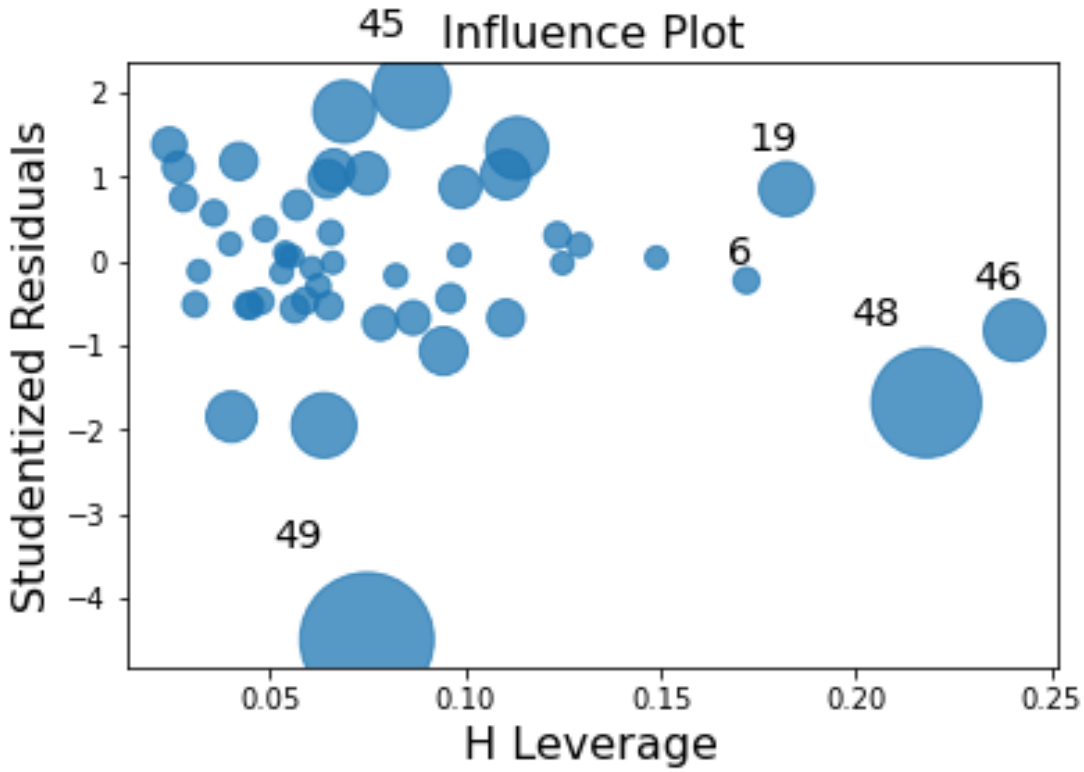
Conclusion – R square value for Administration is 0.9697

9) Probability Plot for MS



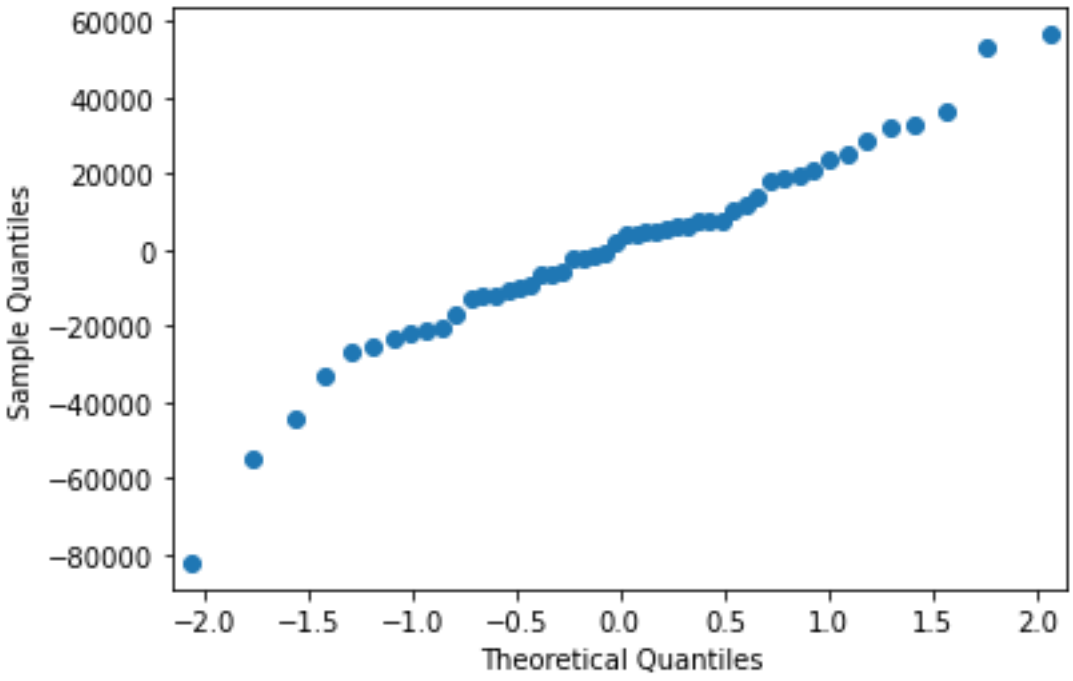
Conclusion – R square value for MS is 0.9829

10) Influence Plot



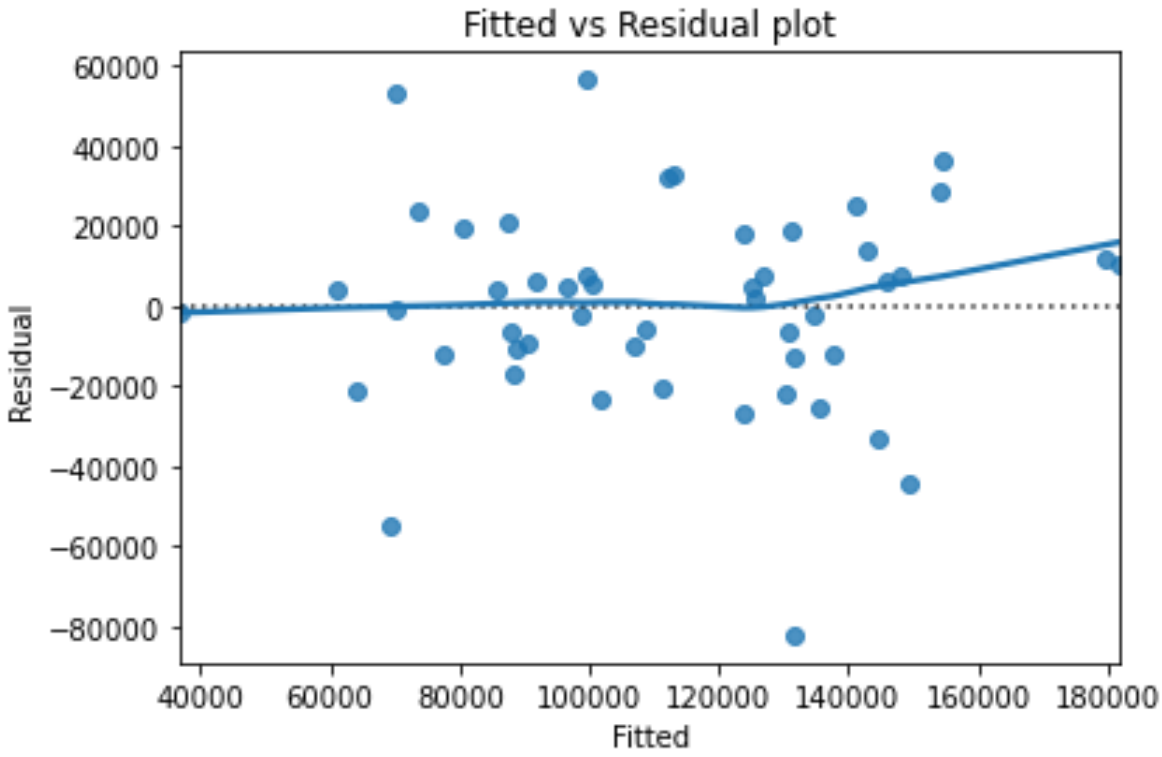
Conclusion – Row 49 shows high influence on data

11) Q-Q Plot



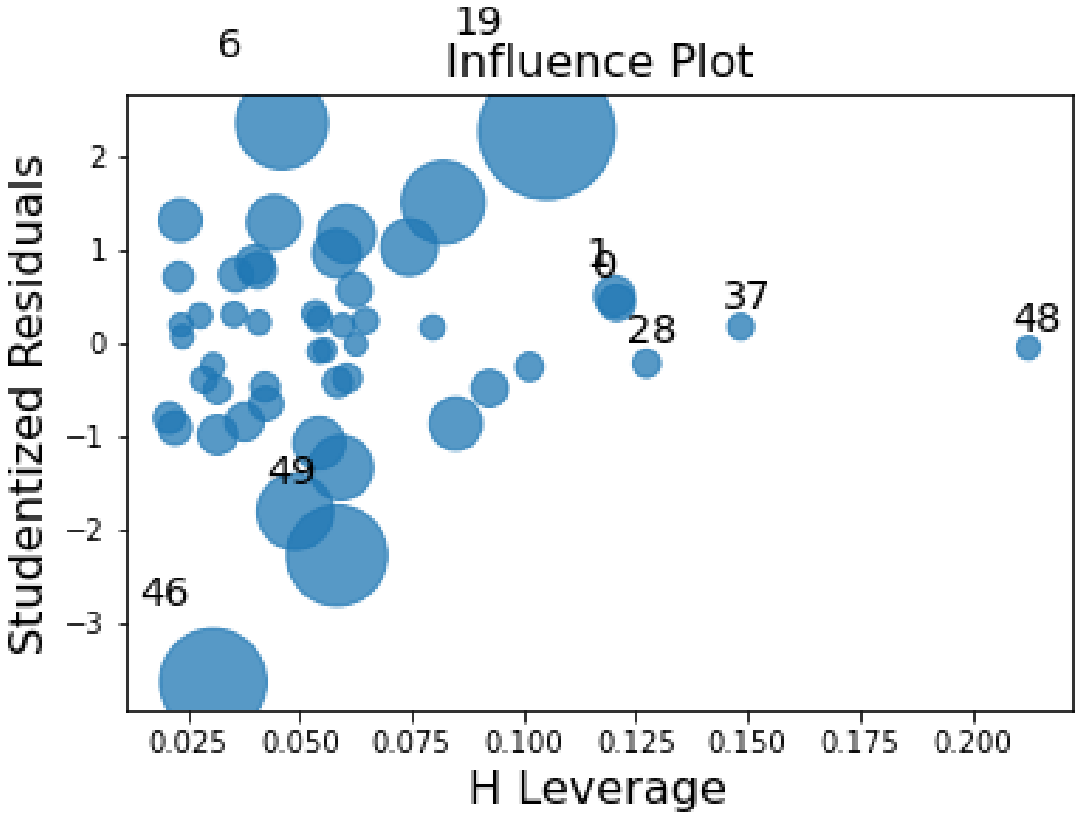
Conclusion – Each of our distributions has the same quantiles

12) Fitted Plot vs Residual



Conclusion - When a linear regression model is suitable for a data set, then the residuals are more or less randomly distributed around the 0 line.

13) Influence Plot for Final Model



Conclusion – Data well fitted for model H leverage corresponded with residual.