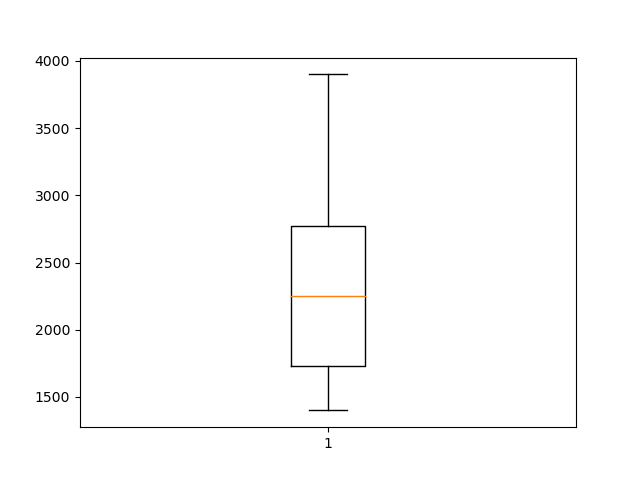
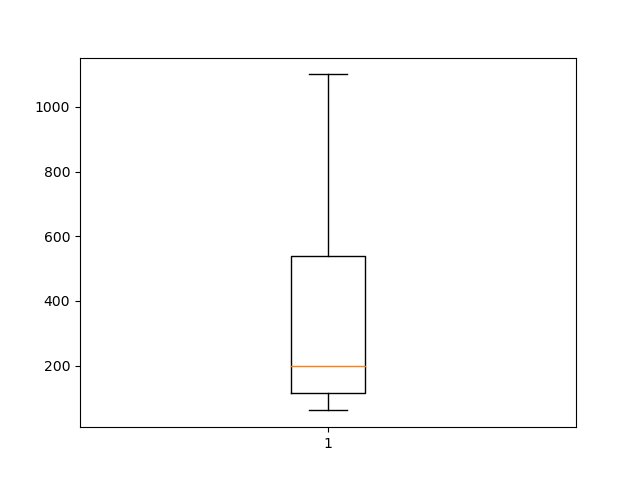
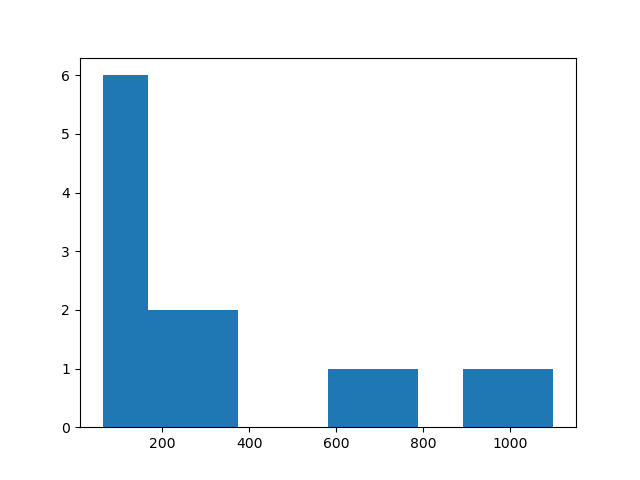
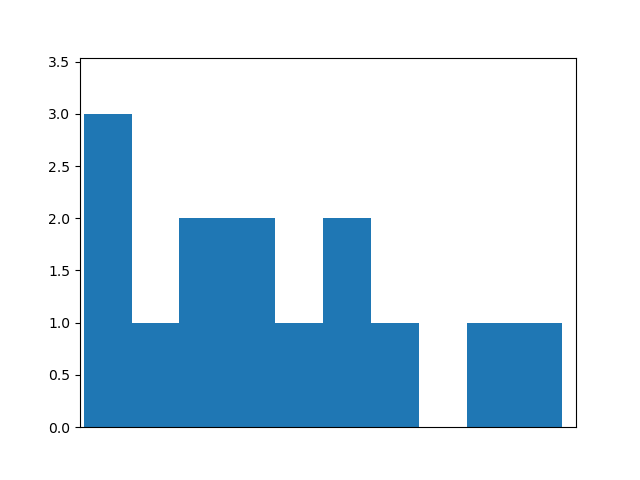
1. Calories consumed predict weight gained using calories consumed



Calories Box Plot Weight Box Plot



Calories Histogram Weight Histogram

wt\_gained cal

count 14.000000 14.000000

mean 357.714286 2340.714286

std 333.692495 752.109488

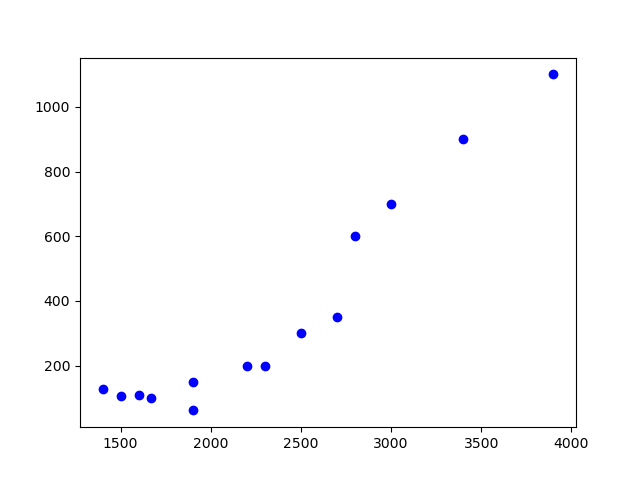
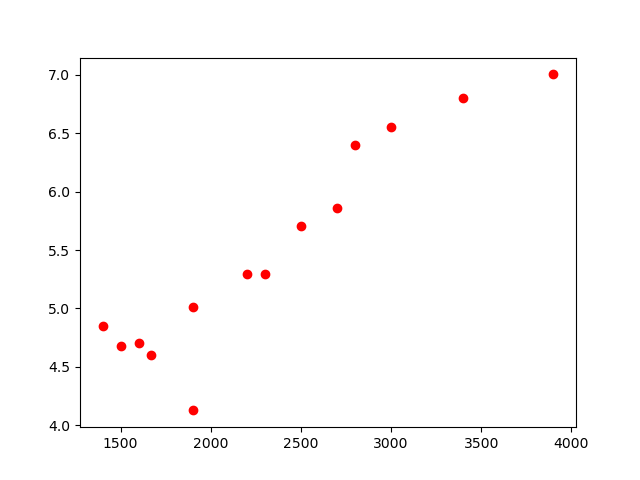
min 62.000000 1400.000000

25% 114.500000 1727.500000

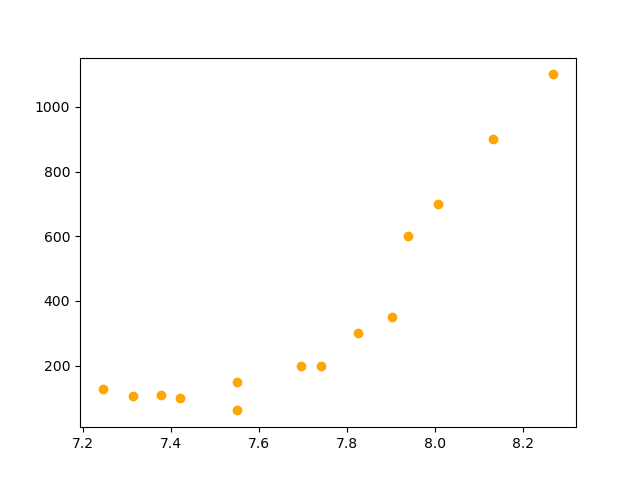
50% 200.000000 2250.000000

75% 537.500000 2775.000000

max 1100.000000 3900.000000



Scatter plot for y(weight) = B + B1(cal) Scatter plot for y(weight) = B + B1log(cal)



Scatter plot for y log(weight) = B + B1(cal)

MODEL RMSE

0 SLR 103.302502

1 Log model 141.005382

2 Exp model 118.045157

3 Poly model 117.414500

OLS Regression Results

===========================================================================

Dep. Variable: wt\_gained R-squared: 0.897

Model: OLS Adj. R-squared: 0.888

Method: Least Squares F-statistic: 104.3

Date: Fri, 15 Jan 2021 Prob (F-statistic): 2.86e-07

Time: 13:37:34 Log-Likelihood: -84.792

No. Observations: 14 AIC: 173.6

Df Residuals: 12 BIC: 174.9

Df Model: 1

Covariance Type: nonrobust

==========================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

Intercept -625.7524 100.823 -6.206 0.000 -845.427 -406.078

cal 0.4202 0.041 10.211 0.000 0.331 0.510

===========================================================================

Omnibus: 3.394 Durbin-Watson: 2.537

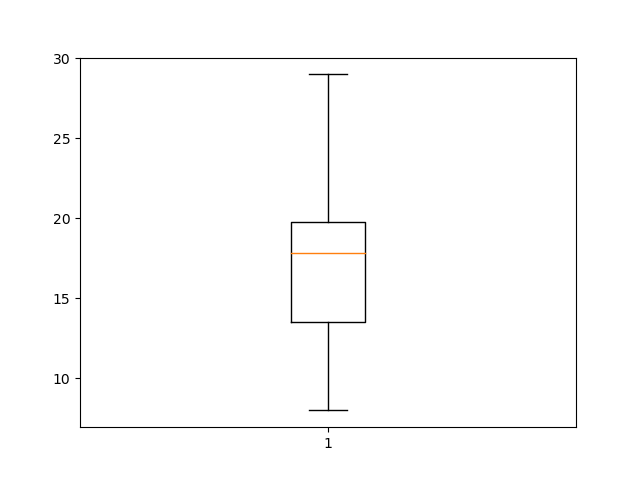
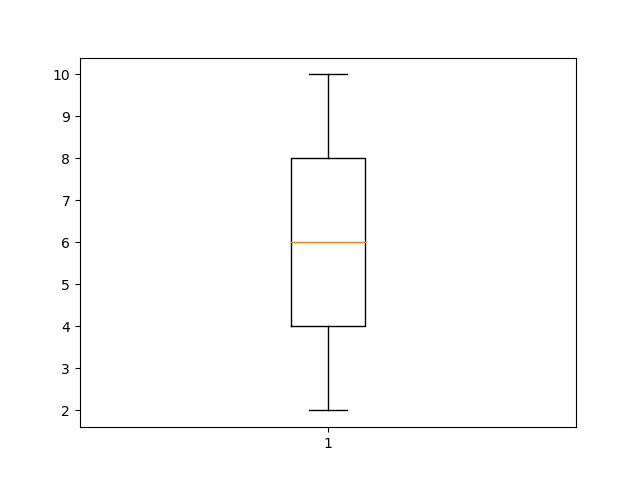
Prob(Omnibus): 0.183 Jarque-Bera (JB): 1.227

Skew: -0.203 Prob(JB): 0.541

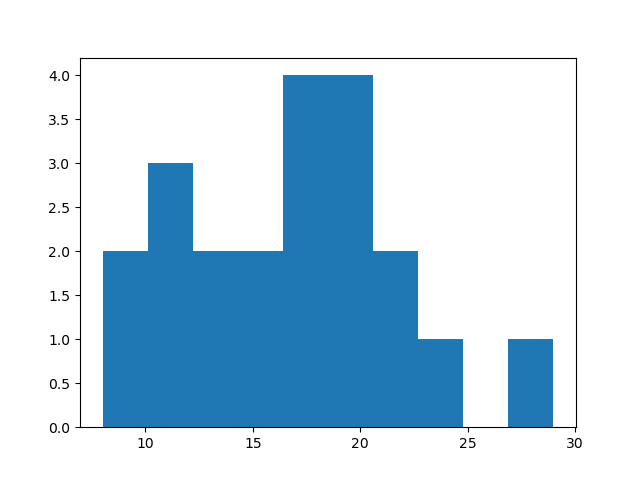
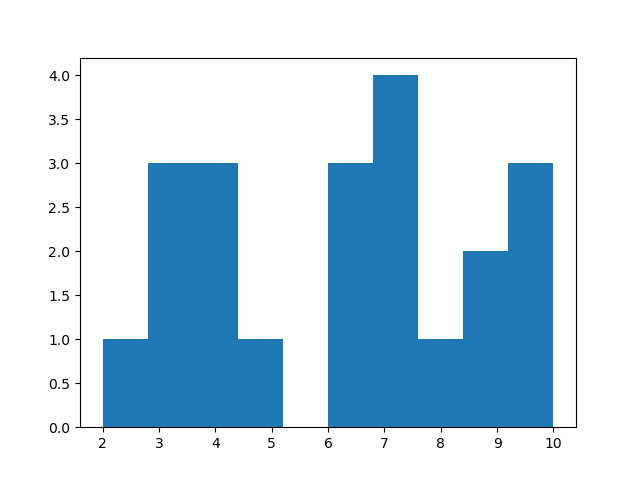
Kurtosis: 1.608 Cond. No. 8.28e+03

===========================================================================

1. Delivery time :Predict delivery time using sorting time



Delivery time box plot Sorting time Box Plot



Delivery Time histogram Sorting time Histogram

del\_time sort\_time

count 21.000000 21.000000

mean 16.790952 6.190476

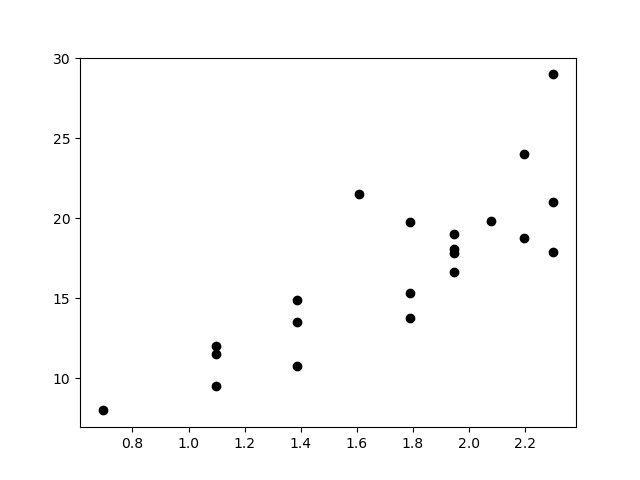
std 5.074901 2.542028

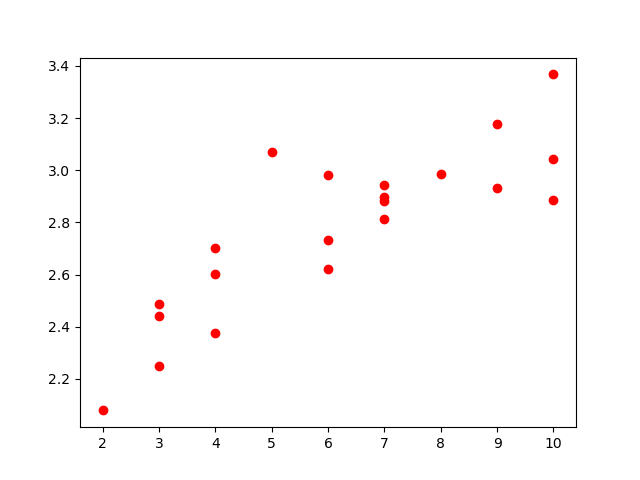
min 8.000000 2.000000

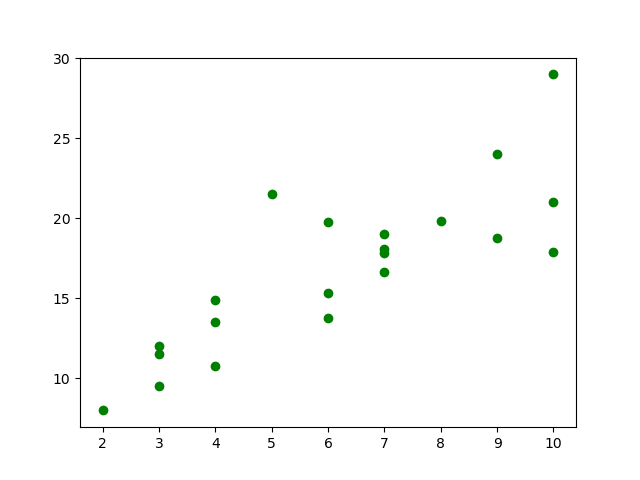
25% 13.500000 4.000000

50% 17.830000 6.000000

75% 19.750000 8.000000

max 29.000000 10.000000

 Scatter plot for y(delivery time) = B + B1log(sorting time)

 Scatter plot for y Log(delivery time) = B + B1(sorting time)

Scatter plot for y(delivery time) = B + B1(sorting time)

MODEL RMSE

0 SLR 3.997528

1 Log model 3.786004

2 Exp model 3.541549

3 Poly model 1.326790

=============================================================================

OLS Regression Results

==============================================================================

Dep. Variable: np.log(del\_time) R-squared: 0.711

Model: OLS Adj. R-squared: 0.696

Method: Least Squares F-statistic: 46.73

Date: Fri, 15 Jan 2021 Prob (F-statistic): 1.59e-06

Time: 22:18:32 Log-Likelihood: 7.7920

No. Observations: 21 AIC: -11.58

Df Residuals: 19 BIC: -9.495

Df Model: 1

Covariance Type: nonrobust

==============================================================================

coef std err t P>|t| [0.025 0.975]

-------------------------------------------------------------------------------------------------------------------------------

Intercept 2.1214 0.103 20.601 0.000 1.906 2.337

sort\_time 0.1056 0.015 6.836 0.000 0.073 0.138

==============================================================================

Omnibus: 1.238 Durbin-Watson: 1.325

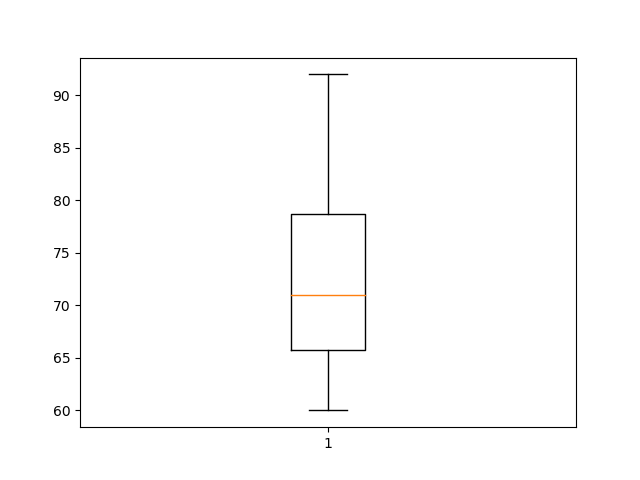
Prob(Omnibus): 0.538 Jarque-Bera (JB): 0.544

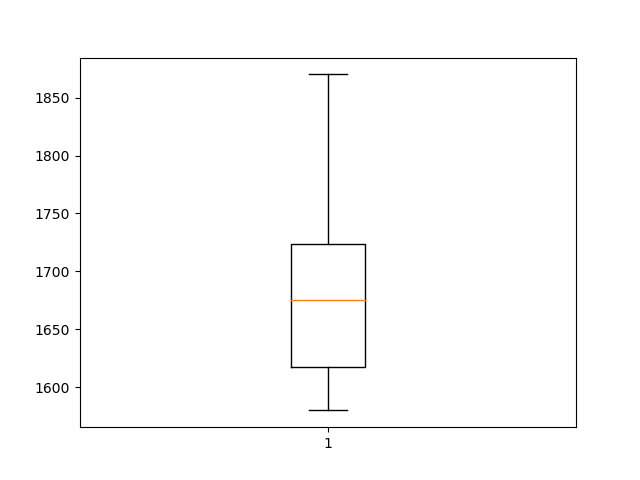
Skew: 0.393 Prob(JB): 0.762

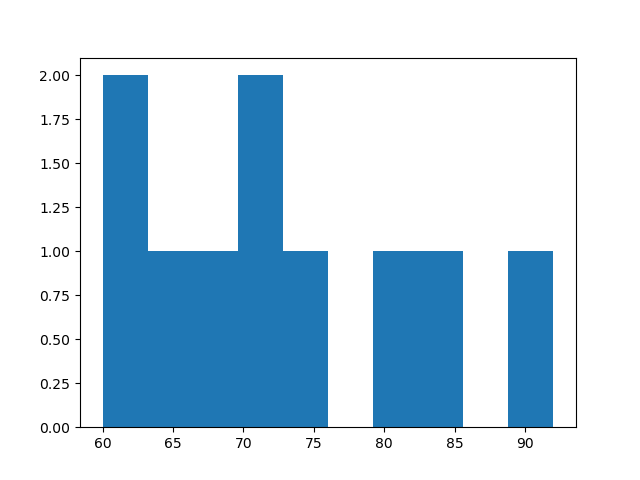
Kurtosis: 3.067 Cond. No. 18.3

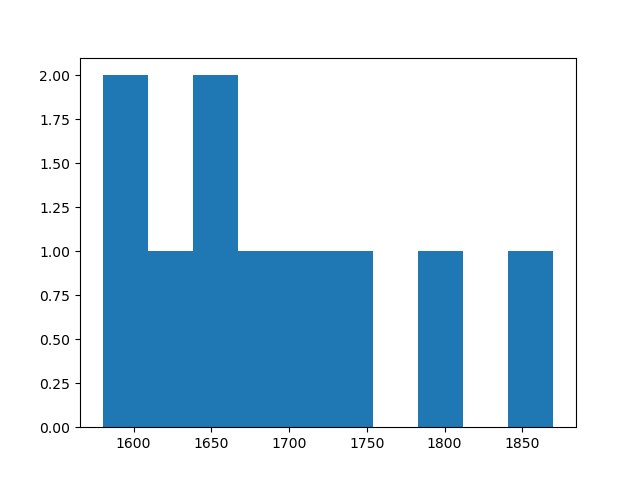
==============================================================================

1. Emp\_data : Build a prediction model for Churn out\_rate





 Churn Out Rate Box plot Salary Hike Box Plot



Churn out rate hHistogram Salary Hike Histogram

salary churn

count 10.000000 10.000000

mean 1688.600000 72.900000

std 92.096809 10.257247

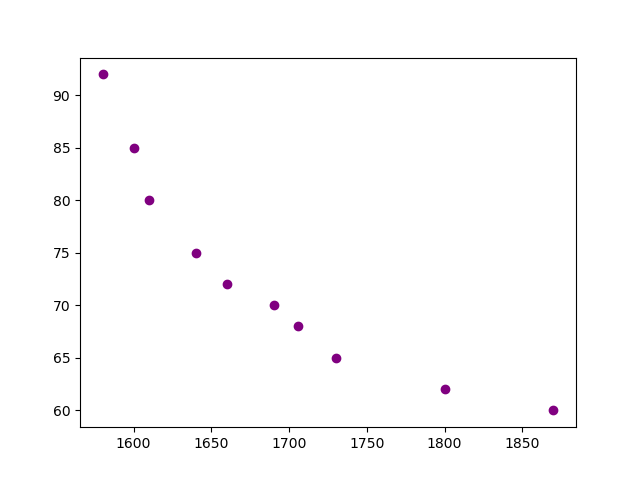
min 1580.000000 60.000000

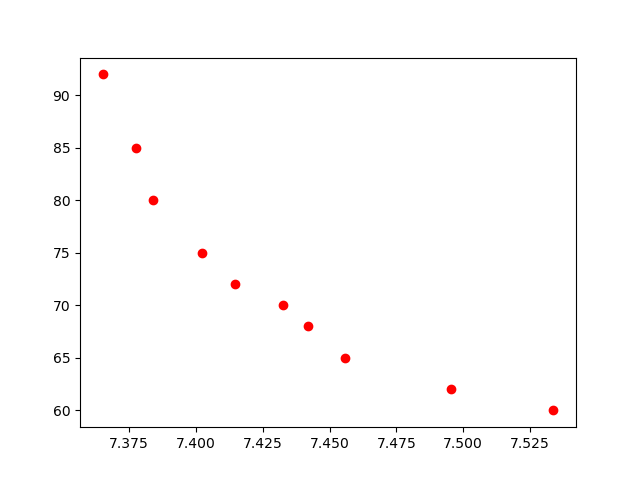
25% 1617.500000 65.750000

50% 1675.000000 71.000000

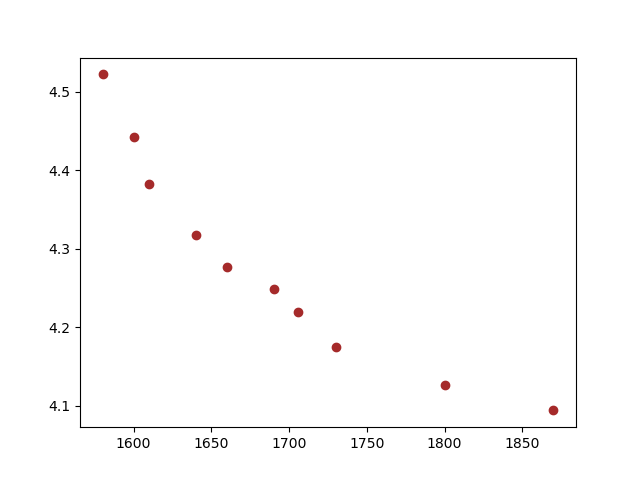
75% 1724.000000 78.750000

max 1870.000000 92.000000



Scatter plot for y(churn out rate) = B + B1(salary hike)

Scatter plot for y(churn out rate) = B + B1Log(salary hike)



Scatter plot for y Log(churn out rate) = B + B1(salary hike)

MODEL RMSE

0 SLR 3.997528

1 Log model 3.786004

2 Exp model 3.541549

3 Poly model 1.326790

OLS Regression Results

===========================================================================

Dep. Variable: np.log(churn) R-squared: 0.984

Model: OLS Adj. R-squared: 0.979

Method: Least Squares F-statistic: 210.1

Date: Fri, 15 Jan 2021 Prob (F-statistic): 5.63e-07

Time: 23:28:53 Log-Likelihood: 26.728

No. Observations: 10 AIC: -47.46

Df Residuals: 7 BIC: -46.55

Df Model: 2

Covariance Type: nonrobust

===========================================================================

coef std err t P>|t| [0.025 0.975]

--------------------------------------------------------------------------------------

Intercept 23.1762 2.415 9.597 0.000 17.466 28.887

salary -0.0207 0.003 -7.351 0.000 -0.027 -0.014

I(salary \* salary) 5.605e-06 8.17e-07 6.857 0.000 3.67e-06 7.54e-06

===========================================================================

Omnibus: 0.520 Durbin-Watson: 1.272

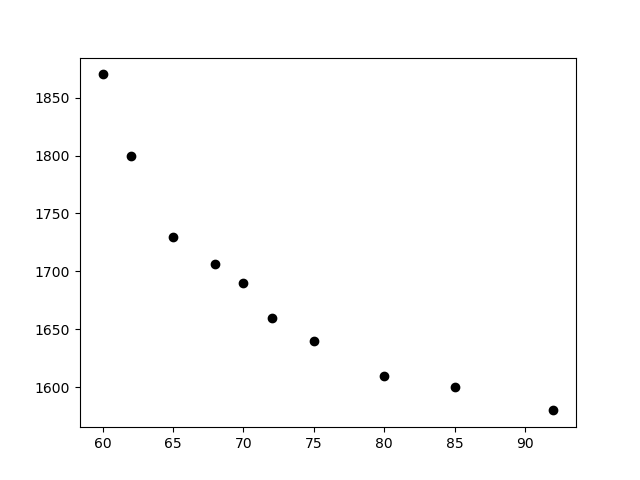
Prob(Omnibus): 0.771 Jarque-Bera (JB): 0.506

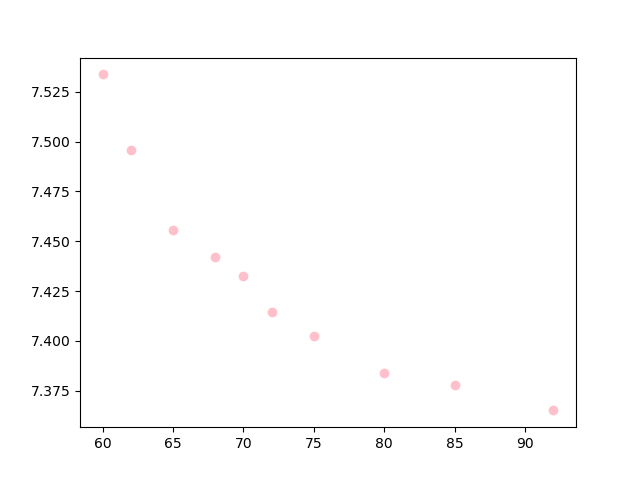
Skew: -0.088 Prob(JB): 0.777

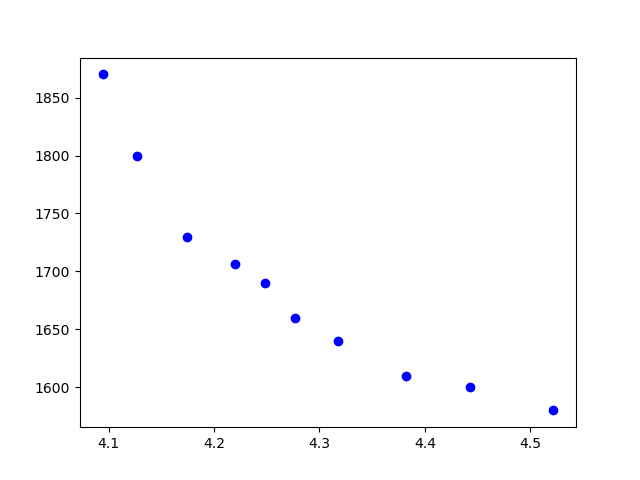
Kurtosis: 1.913 Cond. No. 1.10e+09

===========================================================================

1. Salary\_hike : Build a prediction model for Salary\_hike



 Scatter plot for y(salary Hike) = B + B1(churn out rate)

 Scatter plot for y Log(salary hike) = B + B1(churn out)

Scatter plot for y Log(salary hike) = B + B1(churn out rate)

MODEL RMSE

0 SLR 35.892635

1 Log model 31.069521

2 Exp model 34.268550

3 Poly model 13.025272

==========================================================================

OLS Regression Results

===========================================================================

Dep. Variable: np.log(salary) R-squared: 0.979

Model: OLS Adj. R-squared: 0.972

Method: Least Squares F-statistic: 160.0

Date: Fri, 15 Jan 2021 Prob (F-statistic): 1.44e-06

Time: 23:55:26 Log-Likelihood: 34.809

No. Observations: 10 AIC: -63.62

Df Residuals: 7 BIC: -62.71

Df Model: 2

Covariance Type: nonrobust

==========================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------------------------------------------------

Intercept 8.8843 0.170 52.109 0.000 8.481 9.287

churn -0.0345 0.005 -7.560 0.000 -0.045 -0.024

I(churn \* churn) 0.0002 3.02e-05 6.519 0.000 0.000 0.000

===========================================================================

Omnibus: 0.341 Durbin-Watson: 1.516

Prob(Omnibus): 0.843 Jarque-Bera (JB): 0.031

Skew: -0.049 Prob(JB): 0.984

Kurtosis: 2.744 Cond. No. 3.40e+05

==========================================================================