Ephrem Glushchenko

Electric Motor Batch Size Forecast Analysis Report

Research Objective:

Harry Daniels is a quality control engineer for the Specific Electric Corporation. Specific manufactures electric motors. One of the steps in the manufacturing process involves the use of an automatic milling machine to produce slots in the shafts of the motors. Each batch of shafts is tested, and all shafts that do not meet required dimensional tolerances are discarded. The milling machine must be readjusted at the beginning of each new batch because its cutter head wears slightly during production. Harry is assigned the job of forecasting how the size of a batch affects the number of defective shafts in the batch so that he can select the best batch size. He collects the data for the average batch size of 13 batches shown in Table P-13 and assigns you to analyze it.

a. Plot the data as a scatter diagram.

b. Fit a simple linear regression model.

c. Test for the significance of the slope coefficient.

d. Examine the residuals.

e. Develop a curvilinear model by fitting a simple linear regression model to some transformation of the independent variable.

f. Test for the significance of the slope coefficient of the transformed variable.

g. Examine the residuals.

h. Forecast the number of defectives for a batch size of 300.

i. Which of the models in parts b and e do you prefer?

j. Write Harry a memo summarizing your results.

Problem Definition: Forecast how the size of a batch affects the number of defective shafts within the batch.

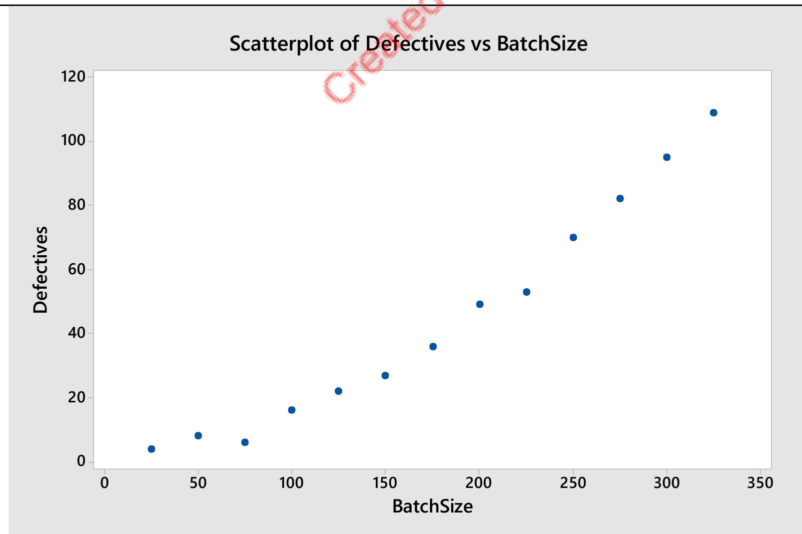
Hypothesis:
H_0: p = 0
H_1: p ≠ 0

Decision Rule: If the p-value less than the alpha, reject null hypothesis.

Conclusion: The p-value of 0.000 is less than the alpha level of 0.05, reject the null hypothesis. This shows that there is a linear relationship. We also reject the null hypothesis for Anova since the p-value is also less than the alpha level.

Interpretation: The tests show significance but no normality or homoscedasticity. Transformation need to be done to that all assumptions are met and show normality.

The points in the normal probability plot of do not seem to follow the line very closely. This behavior suggests a not very good fit between the data and normal distribution. The normal probability plot suggests there is some reason to doubt the normal assumption. The second plot in the first row looks bowed, this behavior would suggest a nonlinear relationship between BatchSize and Defectives. This plot does not fit the normal assumption. The histogram in the bottom row seems to be negatively skewed, not bell-shaped and not centered at zero. This histogram does not fit the normal assumption. The versus order plot also does not fit the normal assumption because it does not cross the zero line between points like it should.

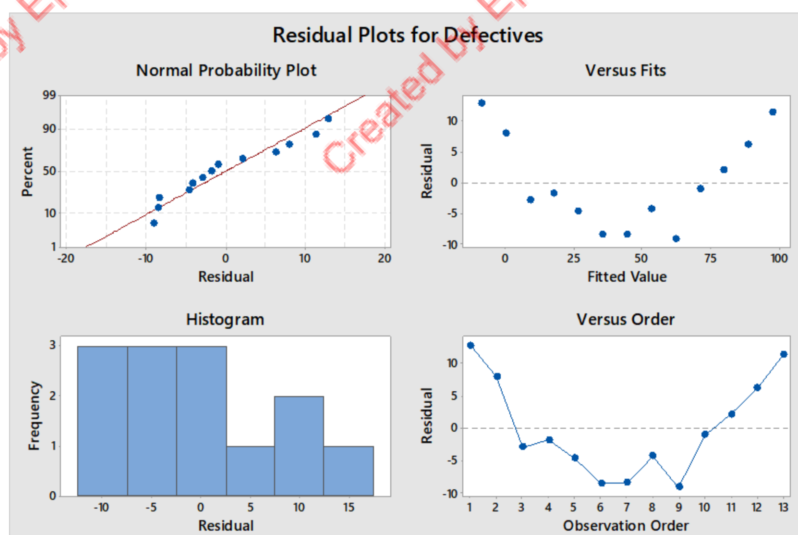


After reviewing the scatterplot, the points seem to be following an upward trend that as the BatchSize increases so does the number of Defectives. This is non stationary and indicates a relationship but must be tested further.

Correlation: BatchSize, Defectives

Correlations

Pearson correlation	0.977
P-value	0.000



Problem Definition: Develop and evaluate the graphs for the four transformations and make the determination of which transformation is best.

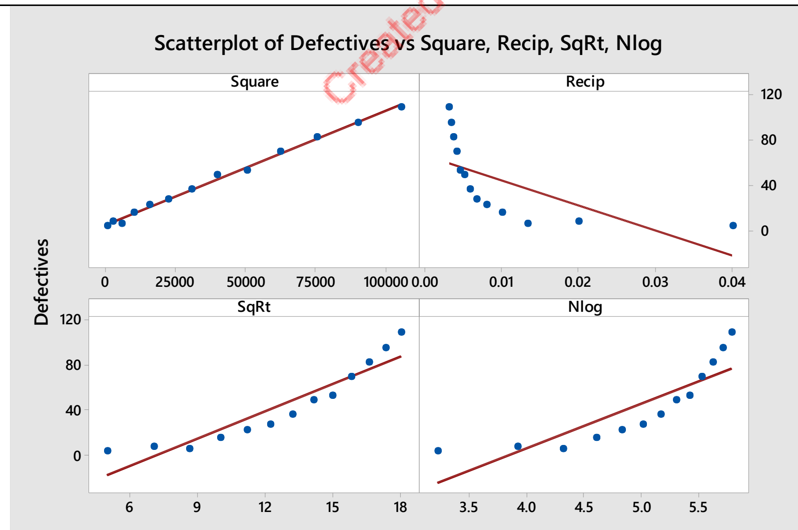
Hypothesis:
H_0: p = 0
H_1: p ≠ 0

Decision Rule: If the p-value less than the alpha, reject null hypothesis.

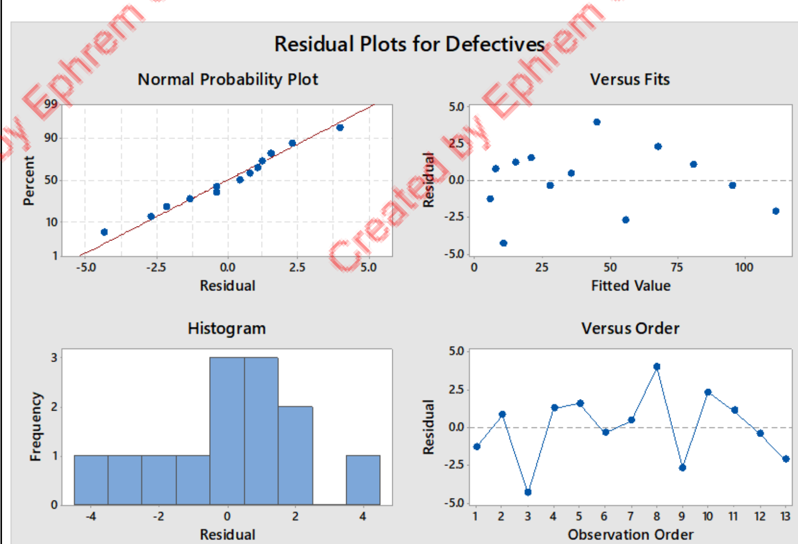
Conclusion: The p-value of 0.000 is less than the alpha level of 0.05, reject the null hypothesis. This shows that there is a linear relationship.

Interpretation: The test show significance and now since that data has been transformed using the square transformation all assumptions have been met and show normality.

The residuals plots for defectives now shoes all graphs meeting all assumptions. The normal probability plot shows the data closely following the line, the versus fits is showing no bow and no pattern, the histogram is centered at zero and has a bell shaped like curve, and the versus ordered crosses the zero line and follows the up and down pattern which shows normality.



After developing four transformations, we can now evaluate each transformation and the effect it has on the data. The square transformation shows the points closely following the line while the recip transformation shows the data exponentially decreasing and flattening out. SquareRoot and Nlog transformations are similar with the data showing a bow and not following the line at all. After evaluating all the transformations, the Square transformation is the best the use.



Regression Equation

Defectives = 4.70 + 0.001008 Square

Prediction

Fit	SE Fit	95% CI	95% PI
95.4111	1.17325	(92.8288, 97.9934)	(89.6468, 101.175)

Ephrem Glushchenko

Memo

To: Senior Management

Re: Transformations

To Management Personal:

As per your request we have completed the requested transformations analysis and prediction for the BatchSize and Defectives data. When predicting the number of defectives for a batch of 300, we got a fit value of 95.4111. The fit value in our case was slightly over the actual number of defectives for a batch size of 300, which was 95. This is still very acceptable and only off by .4111. The average number of defectives for a batch of 300 using the prediction intervals is between 92.8288 and 97.993, while the prediction interval for the number of defectives for a randomly selected batch of 300 is between 89.6468 and 101.175. Some linearity issues were found in the data, the versus fits graph showed the data in a pattern of a bow and the versus ordered graph did not cross the zero-horizontal line. Other assumptions of normality were not met and therefore had to be transformed. We found that the square transformation was the best due to its normal probability plot showing the points closely following the line and that all the assumptions of normality were met. Through our findings we found that the model was statistically significant.