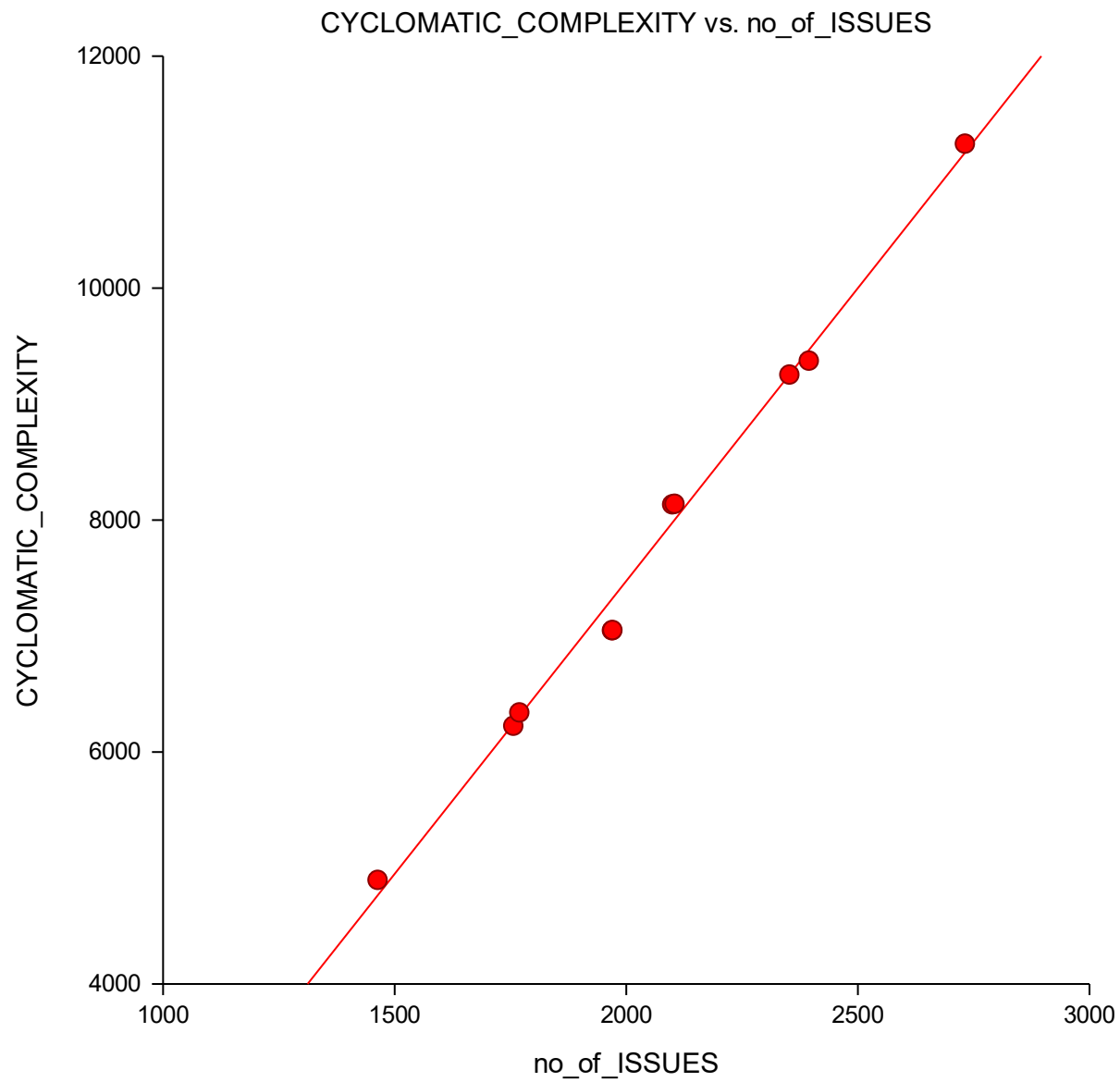


This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Linear Regression Plot Section



This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
 Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Run Summary Section

| Parameter | Value | Parameter | Value |
|----------------------|-----------------------|--------------------------|----------------|
| Dependent Variable | CYCLOMATIC_COMPLEXITY | | Rows Processed |
| 10 | | | |
| Independent Variable | no_of_ISSUES | Rows Used in Estimation | 10 |
| Frequency Variable | None | Rows with X Missing | 0 |
| Weight Variable | None | Rows with Freq Missing | 0 |
| Intercept | -2634.3214 | Rows Prediction Only | 0 |
| Slope | 5.0500 | Sum of Frequencies | 10 |
| R-Squared | 0.9927 | Sum of Weights | 10.0000 |
| Correlation | 0.9964 | Coefficient of Variation | 0.0216 |
| Mean Square Error | 28073.99 | Square Root of MSE | 167.5529 |

Summary Statement

The equation of the straight line relating CYCLOMATIC_COMPLEXITY and no_of_ISSUES is estimated as: $\text{CYCLOMATIC_COMPLEXITY} = (-2634.3214) + (5.0500) \text{ no_of_ISSUES}$ using the 10 observations in this dataset. The y-intercept, the estimated value of CYCLOMATIC_COMPLEXITY when no_of_ISSUES is zero, is -2634.3214 with a standard error of 319.5634. The slope, the estimated change in CYCLOMATIC_COMPLEXITY per unit change in no_of_ISSUES, is 5.0500 with a standard error of 0.1529. The value of R-Squared, the proportion of the variation in CYCLOMATIC_COMPLEXITY that can be accounted for by variation in no_of_ISSUES, is 0.9927. The correlation between CYCLOMATIC_COMPLEXITY and no_of_ISSUES is 0.9964.

A significance test that the slope is zero resulted in a t-value of 33.0219. The significance level of this t-test is 0.0000. Since $0.0000 < 0.0500$, the hypothesis that the slope is zero is rejected.

The estimated slope is 5.0500. The lower limit of the 95% confidence interval for the slope is 4.6973 and the upper limit is 5.4026. The estimated intercept is -2634.3214. The lower limit of the 95% confidence interval for the intercept is -3371.2359 and the upper limit is -1897.4070.

This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Regression Estimation Section

| Parameter | Intercept B(0) | Slope B(1) |
|----------------------------------|-------------------|---------------|
| Regression Coefficients | -2634.3214 | 5.0500 |
| Lower 95% Confidence Limit | -3371.2359 | 4.6973 |
| Upper 95% Confidence Limit | -1897.4070 | 5.4026 |
| Standard Error | 319.5634 | 0.1529 |
| Standardized Coefficient | 0.0000 | 0.9964 |
| T Value | -8.2435 | 33.0219 |
| Prob Level (T Test) | 0.0000 | 0.0000 |
| Reject H0 (Alpha = 0.0500) | Yes | Yes |
| Power (Alpha = 0.0500) | 1.0000 | 1.0000 |
| Regression of Y on X | -2634.3214 | 5.0500 |
| Inverse Regression from X on Y | -2710.6684 | 5.0870 |
| Orthogonal Regression of Y and X | -2707.8071 | 5.0857 |

Notes:

The above report shows the least-squares estimates of the intercept and slope followed by the corresponding standard errors, confidence intervals, and hypothesis tests. Note that these results are based on several assumptions that should be validated before they are used.

Estimated Model

$(-2634.32142966334) + (5.04999341469568) * (\text{no_of_ISSUES})$

This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
 Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Correlation and R-Squared Section

| Parameter | Pearson Correlation Coefficient | R-Squared | Spearman Rank Correlation Coefficient |
|------------------------------------|---------------------------------------|-----------|--|
| Estimated Value | 0.9964 | 0.9927 | 1.0000 |
| Lower 95% Conf. Limit (r dist'n) | 0.9823 | | |
| Upper 95% Conf. Limit (r dist'n) | 0.9990 | | |
| Lower 95% Conf. Limit (Fisher's z) | 0.9840 | | 1.0000 |
| Upper 95% Conf. Limit (Fisher's z) | 0.9992 | | 1.0000 |
| Adjusted (Rbar) | | 0.9918 | |
| T-Value for H0: Rho = 0 | 33.0219 | 33.0219 | |
| Prob Level for H0: Rho = 0 | 0.0000 | 0.0000 | 0.0000 |

Notes:

The confidence interval for the Pearson correlation assumes that X and Y follow the bivariate normal distribution. This is a different assumption from linear regression which assumes that X is fixed and Y is normally distributed.

Two confidence intervals are given. The first is based on the exact distribution of Pearson's correlation. The second is based on Fisher's z transformation which approximates the exact distribution using the normal distribution. Why are both provided? Because most books only mention Fisher's approximate method, it will often be needed to do homework. However, the exact methods should be used whenever possible.

The confidence limits can be used to test hypotheses about the correlation. To test the hypothesis that rho is a specific value, say r_0 , check to see if r_0 is between the confidence limits. If it is, the null hypothesis that $\rho = r_0$ is not rejected. If r_0 is outside the limits, the null hypothesis is rejected.

Spearman's Rank correlation is calculated by replacing the original data with their ranks. This correlation is used when some of the assumptions may be invalid.

Summary Matrices

| Index | X'X 0 | X'X 1 | X'Y 2 | X'X Inverse 0 | X'X Inverse 1 |
|-------------|----------|--------------|--------------|------------------|------------------|
| 0 | 10 | 20607 | 77722 | 3.637558 | -0.001716678 |
| 1 | 20607 | 4.366524E+07 | 1.662237E+08 | -0.001716678 | 8.330556E-07 |
| 2 (Y'Y) | | | 6.349087E+08 | | |
| Determinant | | 1.2004E+07 | | | 8.330556E-08 |

Variance - Covariance Matrix of Regression Coefficients

| Index | VC(b) 0 | VC(b) 1 |
|-------|------------|------------|
| 0 | 102120.7 | -48.19398 |
| 1 | -48.19398 | 0.02338719 |

This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
 Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Tests of Assumptions Section

| Assumption/Test | Test Value | Prob Level | Is the Assumption Reasonable at the 0.2000 Level of Significance? |
|--|------------|------------|---|
| Residuals follow Normal Distribution? | | | |
| Shapiro Wilk | 0.8772 | 0.121258 | No |
| Anderson Darling | 0.5270 | 0.178951 | No |
| D'Agostino Skewness | -1.2439 | 0.213537 | Yes |
| D'Agostino Kurtosis | -0.1784 | 0.858438 | Yes |
| D'Agostino Omnibus | 1.5791 | 0.454050 | Yes |
| Constant Residual Variance? | | | |
| Modified Levene Test | 1.1869 | 0.307687 | Yes |
| Relationship is a Straight Line? | | | |
| Lack of Linear Fit F(0, 0) Test | 0.0000 | 0.000000 | No |

No Serial Correlation?

Evaluate the Serial-Correlation report and the Durbin-Watson test if you have equal-spaced, time series data.

Notes:

A 'Yes' means there is not enough evidence to make this assumption seem unreasonable. This lack of evidence may be because the sample size is too small, the assumptions of the test itself are not met, or the assumption is valid.

A 'No' means the that the assumption is not reasonable. However, since these tests are related to sample size, you should assess the role of sample size in the tests by also evaluating the appropriate plots and graphs. A large dataset (say $N > 500$) will often fail at least one of the normality tests because it is hard to find a large dataset that is perfectly normal.

Normality and Constant Residual Variance:

Possible remedies for the failure of these assumptions include using a transformation of Y such as the log or square root, correcting data-recording errors found by looking into outliers, adding additional independent variables, using robust regression, or using bootstrap methods.

Straight-Line:

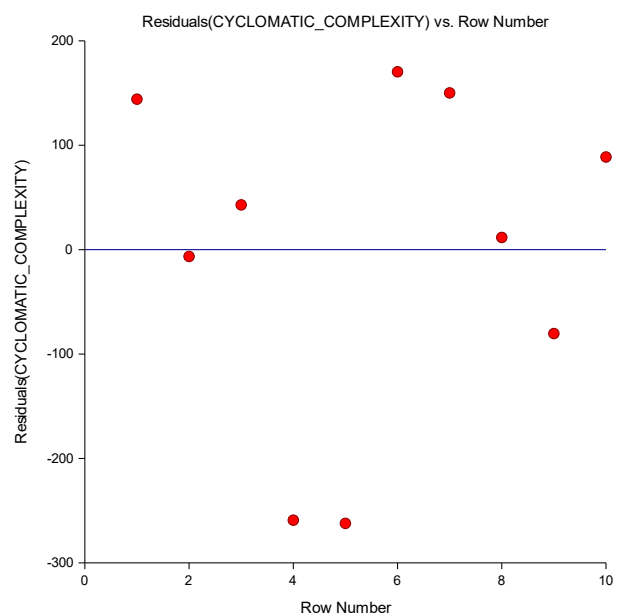
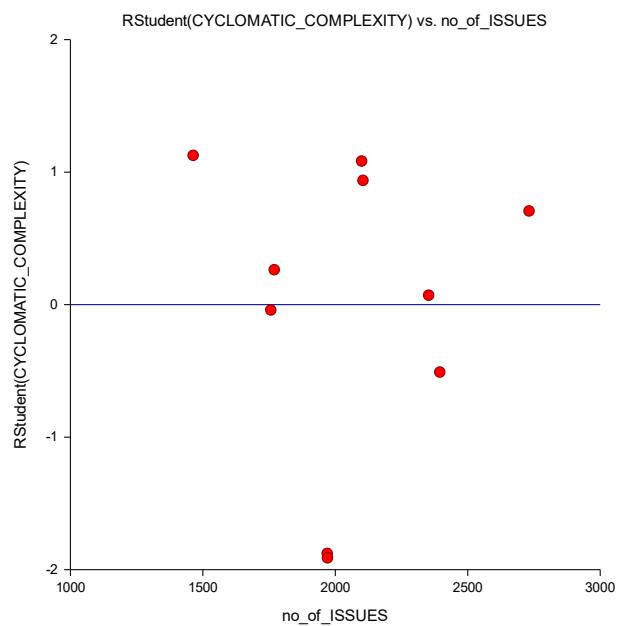
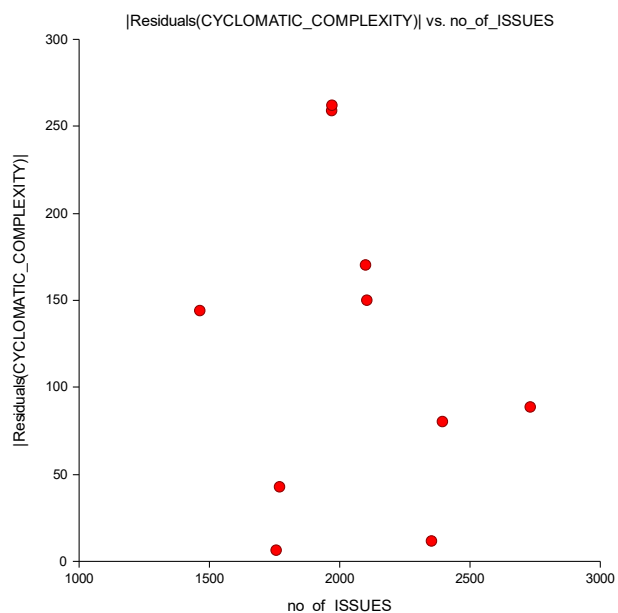
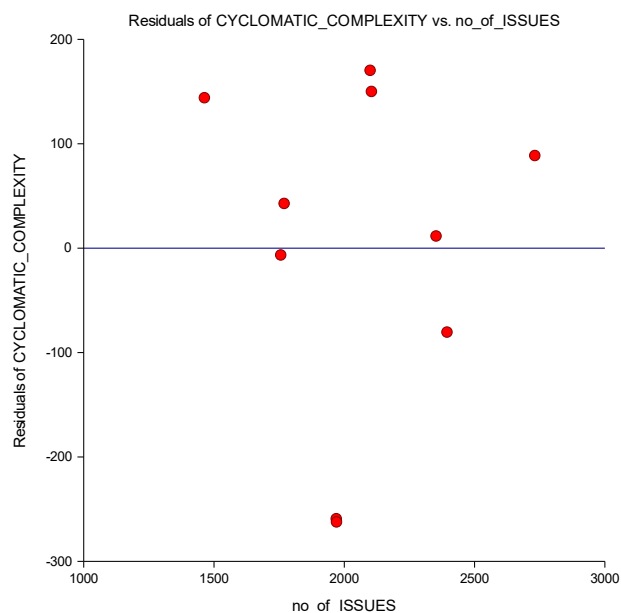
Possible remedies for the failure of this assumption include using nonlinear regression or polynomial regression.

This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

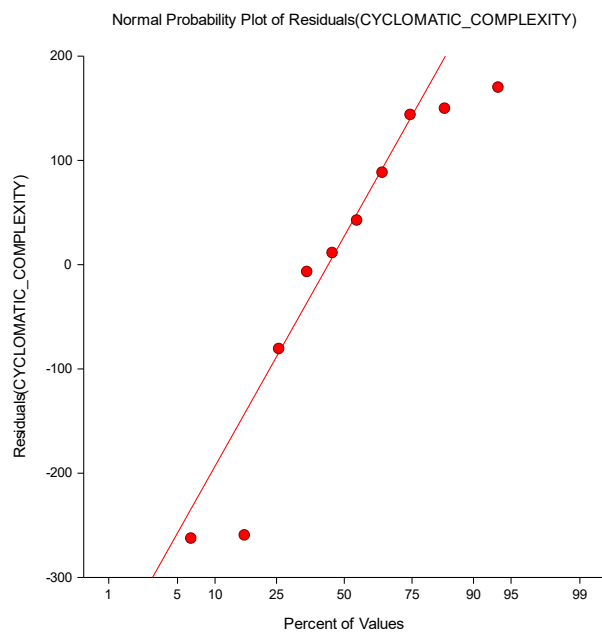
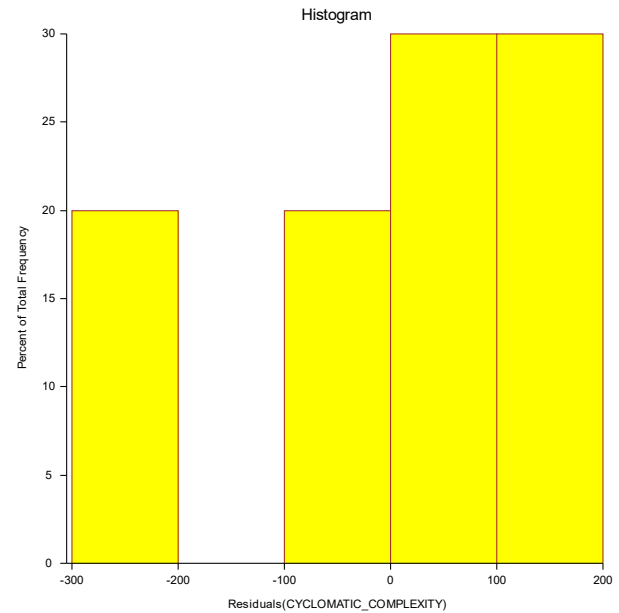
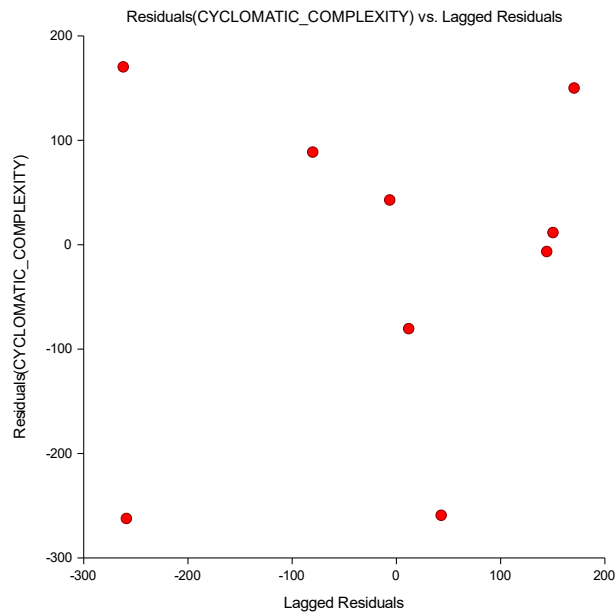
Residual Plots Section



This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES



This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
 Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Procedure Input Settings

Autosaved Template File

C:\Users\KASATLA\Documents\NCSS 12\Procedure Templates\Autosave\Linear Regression and Correlation - Autosaved 2018_3_30-0_44_46.t153

Variables Tab

```
-- Variables -----
Y: Dependent Variable(s):      CYCLOMATIC_COMPLEXITY
X: Independent Variable:       no_of_ISSUES
Frequency Variable:            <Empty>
Weight Variable:               <Empty>

-- Model Specification -----
Remove Intercept                Unchecked

-- Resampling (Increases computation time) -----
Calculate Bootstrap C.I.'s      Unchecked
Run Randomization Tests         Unchecked

-- Alpha Levels -----
Alpha for C.I.'s and Tests:     0.050
Alpha for Assumptions:         0.20
```

Reports Tab

```
-- Select Report / Plot Group -----
Select a Group of Reports and Plots:  Display only those items that are CHECKED BELOW
Show Notes                           Checked
Show All Rows                        Checked

-- Select Reports -----
.. Summaries .....
Run Summary                          Checked
Summary Statement                    Checked
Descriptive Statistics                Unchecked
Correlation and R-Squared             Checked
Summary Matrices                     Checked

.. Estimation .....
Regression Estimation                 Checked

.. ANOVA .....
ANOVA                               Unchecked

.. Assumptions .....
Assumptions                          Checked
  Levene Groups:                      2
Durbin-Watson                        Unchecked
PRESS                                Unchecked
```


This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
 Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Procedure Input Settings (Continued)

Reports Tab (Continued)

-- Prediction -----

| | |
|------------------------------|-----------|
| Predict Y at these X values: | <Empty> |
| Predicted Y - C.L. | Unchecked |
| Predicted Y - P.L. | Unchecked |

-- Row-by-Row Lists -----

| | |
|-------------------------|-----------|
| Original Data | Unchecked |
| Predicted Y Means | Unchecked |
| Predicted Y Individuals | Unchecked |
| Simultaneous Bands | Unchecked |
| Predicted X Means | Unchecked |
| Predicted X Individuals | Unchecked |

-- Regression Diagnostics -----

| | |
|---------------------------|-----------|
| Residuals | Unchecked |
| Residual Diagnostics | Unchecked |
| Leave One Row Out | Unchecked |
| Outlier Detection Chart | Unchecked |
| Influence Detection Chart | Unchecked |
| Outlier-Influence Chart | Unchecked |

Report Options Tab

-- Report Options -----

| | |
|-----------------|--------|
| Precision: | Single |
| Variable Names: | Names |

-- Decimal Places -----

| | |
|----------------------|-----|
| Probability: | 4 |
| Beta (Coefficients): | 4 |
| SE: | 4 |
| T: | 4 |
| R2: | 4 |
| X: | 4 |
| Y: | 4 |
| Residuals: | 4 |
| Std Residuals: | 4 |
| Sum Squares: | All |
| Matrix: | All |

This report is for evaluation purposes only.
There are 23 days remaining in your free trial (Expires on 4/20/2018).

Linear Regression Report

Dataset Untitled
 Y = CYCLOMATIC_COMPLEXITY X = no_of_ISSUES

Procedure Input Settings (Continued)

Plots Tab

| | |
|-----------------------|---------|
| -- Select Plots ----- | |
| Y vs X | Checked |
| RStudent vs X | Checked |
| Histogram | Checked |
| Residuals vs X | Checked |
| Residuals vs Row | Checked |
| Probability Plot | Checked |
| Residuals vs X | Checked |
| Serial Correlation | Checked |
| -- Plot Options ----- | |
| Y vs X Plot Size: | Medium |
| All Other Plot Sizes: | Small |

Resampling Tab

| | |
|--|----------------|
| -- Bootstrap Calculation Options ----- | |
| .. Sampling | |
| Samples (N): | 3000 |
| Sampling Method: | Observations |
| Retries: | 50 |
| .. Estimation | |
| Percentile Type: | Ave X(p[n+1]) |
| C.I. Method: | Reflection |
| Bootstrap Confidence Coefficients: | 0.90 0.95 0.99 |
| -- Randomization Test Options ----- | |
| Monte Carlo Samples: | 1000 |

Storage Tab

| | |
|-------------------------------|-------------------|
| -- Data Storage Options ----- | |
| Storage Option: | Do not store data |