Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Linear Regression Plot Section
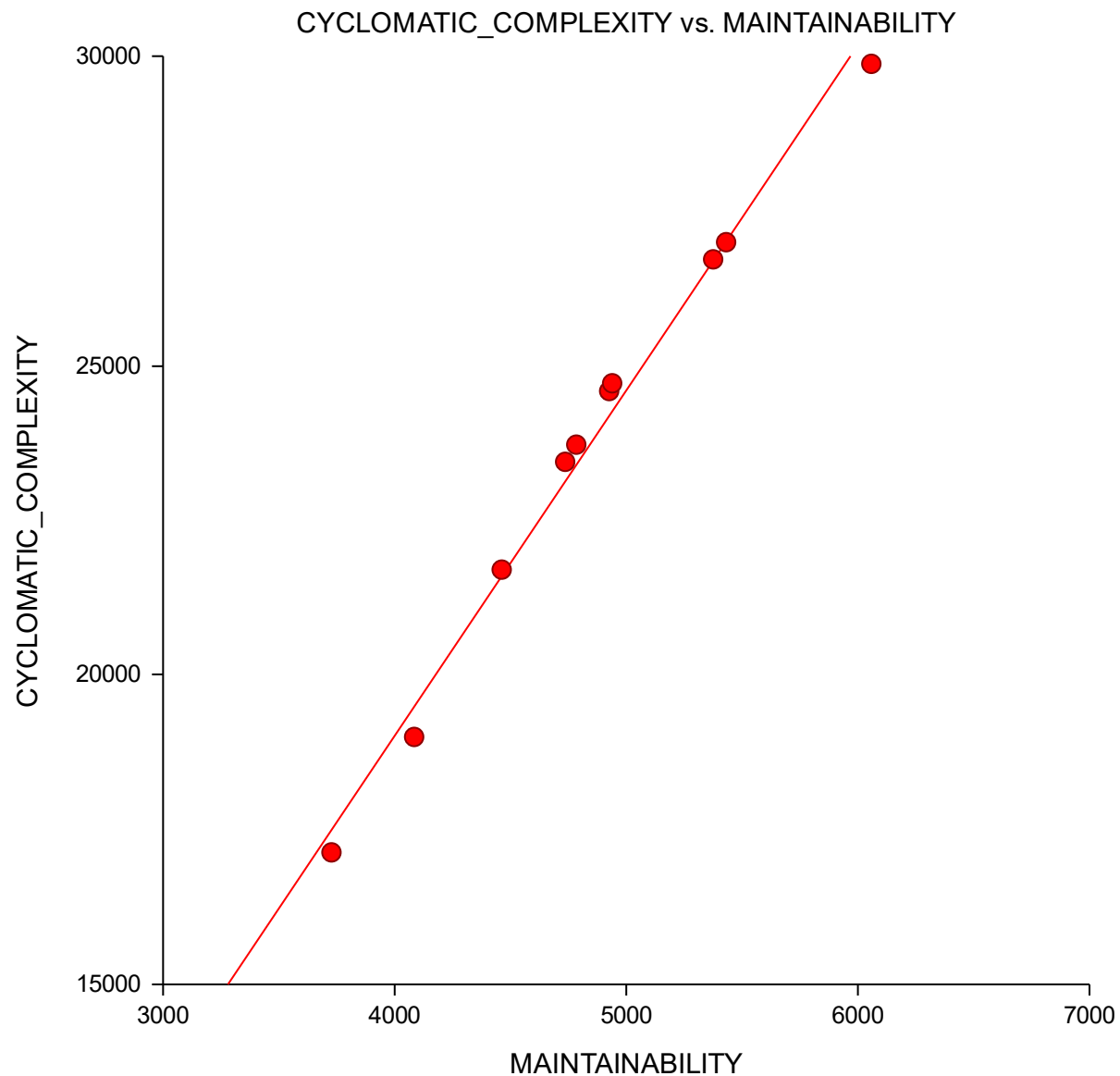
Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Run Summary Section ——————————————————————————

Parameter	Value	Parameter	Value
Dependent Variable	CYCLOMATIC_COMPLEXITY	Rows Processed	10
Independent Variable	MAINTAINABILITY	Rows Used in Estimation	10
Frequency Variable	None	Rows with X Missing	0
Weight Variable	None	Rows with Freq Missing	0
Intercept	-3289.6125	Rows Prediction Only	0
Slope	5.5810	Sum of Frequencies	10
R-Squared	0.9895	Sum of Weights	10.0000
Correlation	0.9948	Coefficient of Variation	0.0173
Mean Square Error	168753.2	Square Root of MSE	410.7958

Summary Statement ——————————————————————————

The equation of the straight line relating CYCLOMATIC_COMPLEXITY and MAINTAINABILITY is estimated as: CYCLOMATIC_COMPLEXITY = (-3289.6125) + (5.5810) MAINTAINABILITY using the 10 observations in this dataset. The y-intercept, the estimated value of CYCLOMATIC_COMPLEXITY when MAINTAINABILITY is zero, is -3289.6125 with a standard error of 992.7133. The slope, the estimated change in CYCLOMATIC_COMPLEXITY per unit change in MAINTAINABILITY, is 5.5810 with a standard error of 0.2028. The value of R-Squared, the proportion of the variation in CYCLOMATIC_COMPLEXITY that can be accounted for by variation in MAINTAINABILITY, is 0.9895. The correlation between CYCLOMATIC_COMPLEXITY and MAINTAINABILITY is 0.9948.

A significance test that the slope is zero resulted in a t-value of 27.5154. 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 5.5810. The lower limit of the 95% confidence interval for the slope is 5.1132 and the upper limit is 6.0487. The estimated intercept is -3289.6125. The lower limit of the 95% confidence interval for the intercept is -5578.8135 and the upper limit is -1000.4114.

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Regression Estimation Section ————————————————————————

Parameter	Intercept B(0)	Slope B(1)
Regression Coefficients	-3289.6125	5.5810
Lower 95% Confidence Limit	-5578.8135	5.1132
Upper 95% Confidence Limit	-1000.4114	6.0487
Standard Error	992.7133	0.2028
Standardized Coefficient	0.0000	0.9948
T Value	-3.3138	27.5154
Prob Level (T Test)	0.0106	0.0000
Reject H0 (Alpha = 0.0500)	Yes	Yes
Power (Alpha = 0.0500)	0.8261	1.0000
Regression of Y on X	-3289.6125	5.5810
Inverse Regression from X on Y	-3575.7584	5.6399
Orthogonal Regression of Y and X	-3566.9448	5.6381

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
(-3289.61245400761) + (5.58097614566748) * (MAINTAINABILITY)

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Correlation and R-Squared Section

Parameter	Pearson Correlation Coefficient	R-Squared	Spearman Rank Correlation Coefficient
Estimated Value	0.9948	0.9895	1.0000
Lower 95% Conf. Limit (r dist'n)	0.9746		
Upper 95% Conf. Limit (r dist'n)	0.9986		
Lower 95% Conf. Limit (Fisher's z)	0.9771		1.0000
Upper 95% Conf. Limit (Fisher's z)	0.9988		1.0000
Adjusted (Rbar)		0.9882	
T-Value for H0: Rho = 0	27.5154	27.5154	
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 r0, check to see if r0 is between the
confidence limits. If it is, the null hypothesis that rho = r0 is not rejected. If r0 is
outside the limits, the null hypothesis is rejected.

Spearman's Rank correlation is calculated by replacing the orginal 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	48522	237904	5.83977	-0.001182921
1	48522	2.395403E+08	1.17725E+09	-0.001182921	2.437907E-07
2 (Y'Y)			5.788944E+09		
Determinant		4.10188E+07			2.437907E-08

Variance - Covariance Matrix of Regression Coefficients

Index	VC(b) 0	VC(b) 1
0	985479.8	-199.6217
1	-199.6217	0.04114046

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

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	0.9048	0.247168	Yes
Anderson Darling	0.4476	0.279716	Yes
D'Agostino Skewness	-0.8645	0.387336	Yes
D'Agostino Kurtosis	-0.8880	0.374530	Yes
D'Agostino Omnibus	1.5359	0.463971	Yes
Constant Residual Variance?			
Modified Levene Test	0.0000	0.998467	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.

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Residual Plots Section
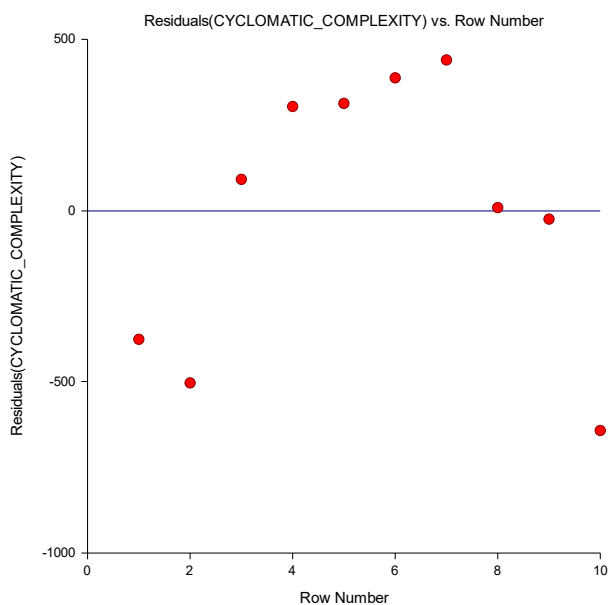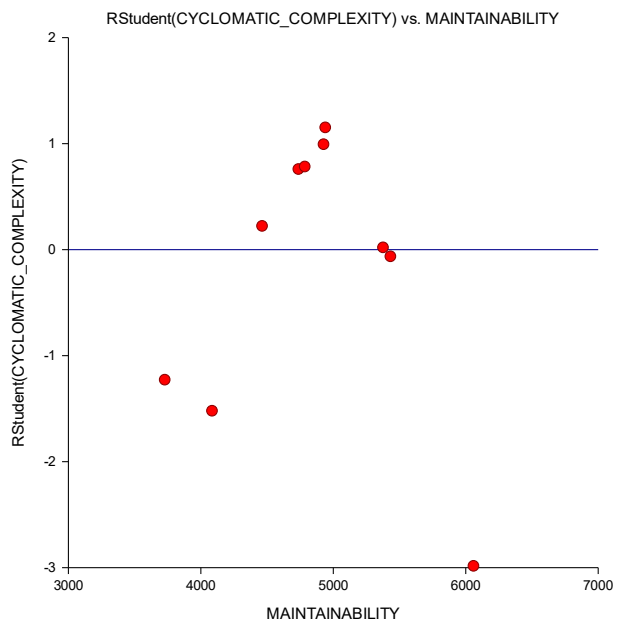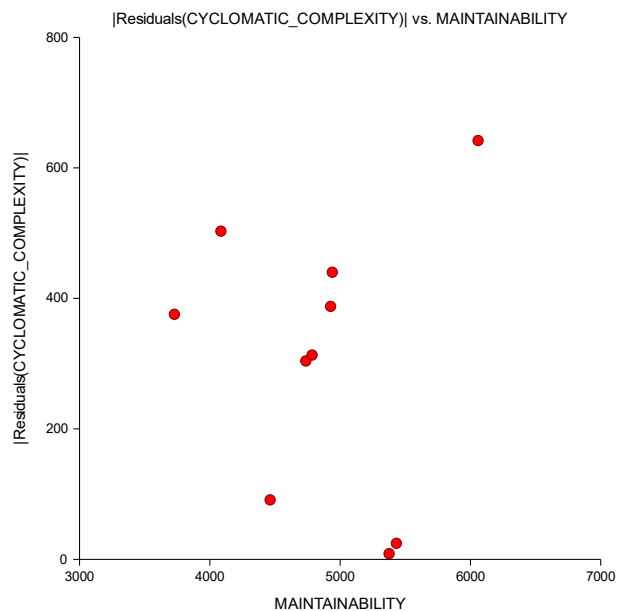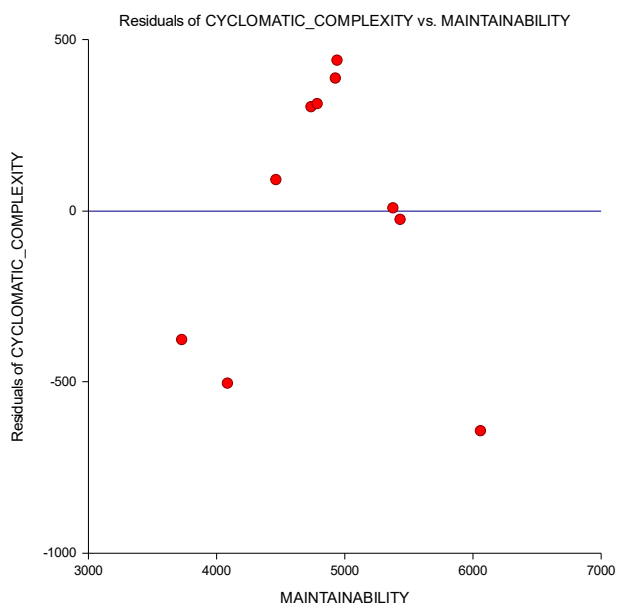
Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Residuals(CYCLOMATIC_COMPLEXITY) vs. Lagged Residuals

Histogram

Normal Probability Plot of Residuals(CYCLOMATIC_COMPLEXITY)
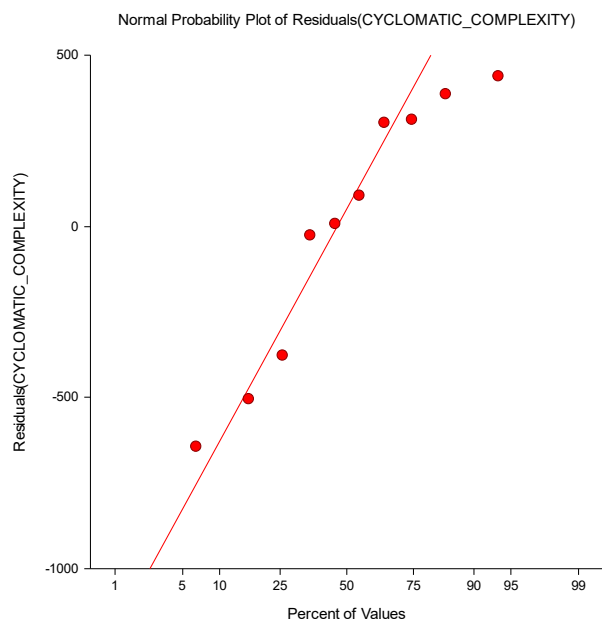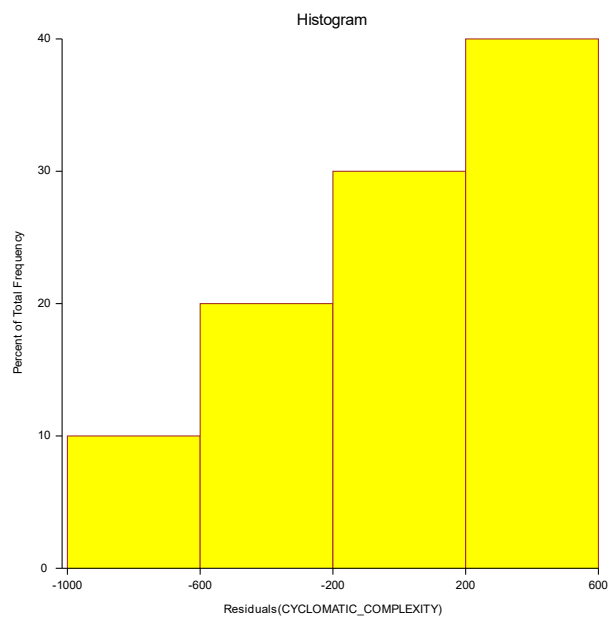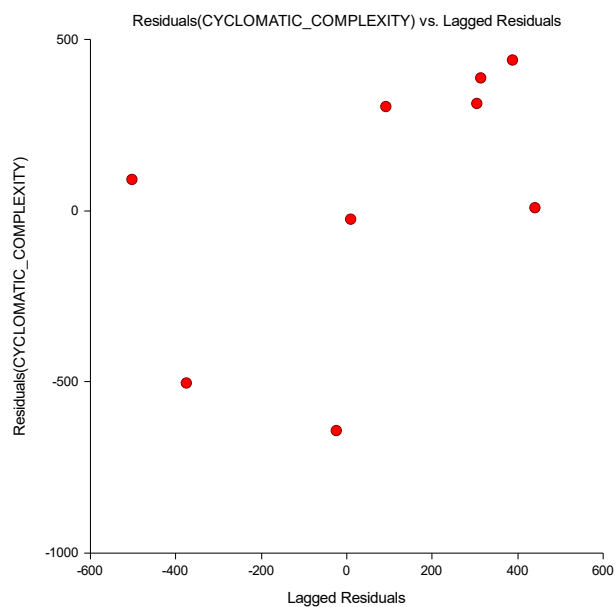
Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Procedure Input Settings ──

Autosaved Template File
C:\Users\KASATLA\Documents\NCSS 12\Procedure Templates\Autosave\Linear Regression and Correlation -
Autosaved 2018_3_30-1_4_30.t153

Variables Tab
-- Variables --
Y: Dependent Variable(s):   CYCLOMATIC_COMPLEXITY
X: Independent Variable: MAINTAINABILITY
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

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

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

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

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