Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Linear Regression Plot Section —————————————————————————————————————
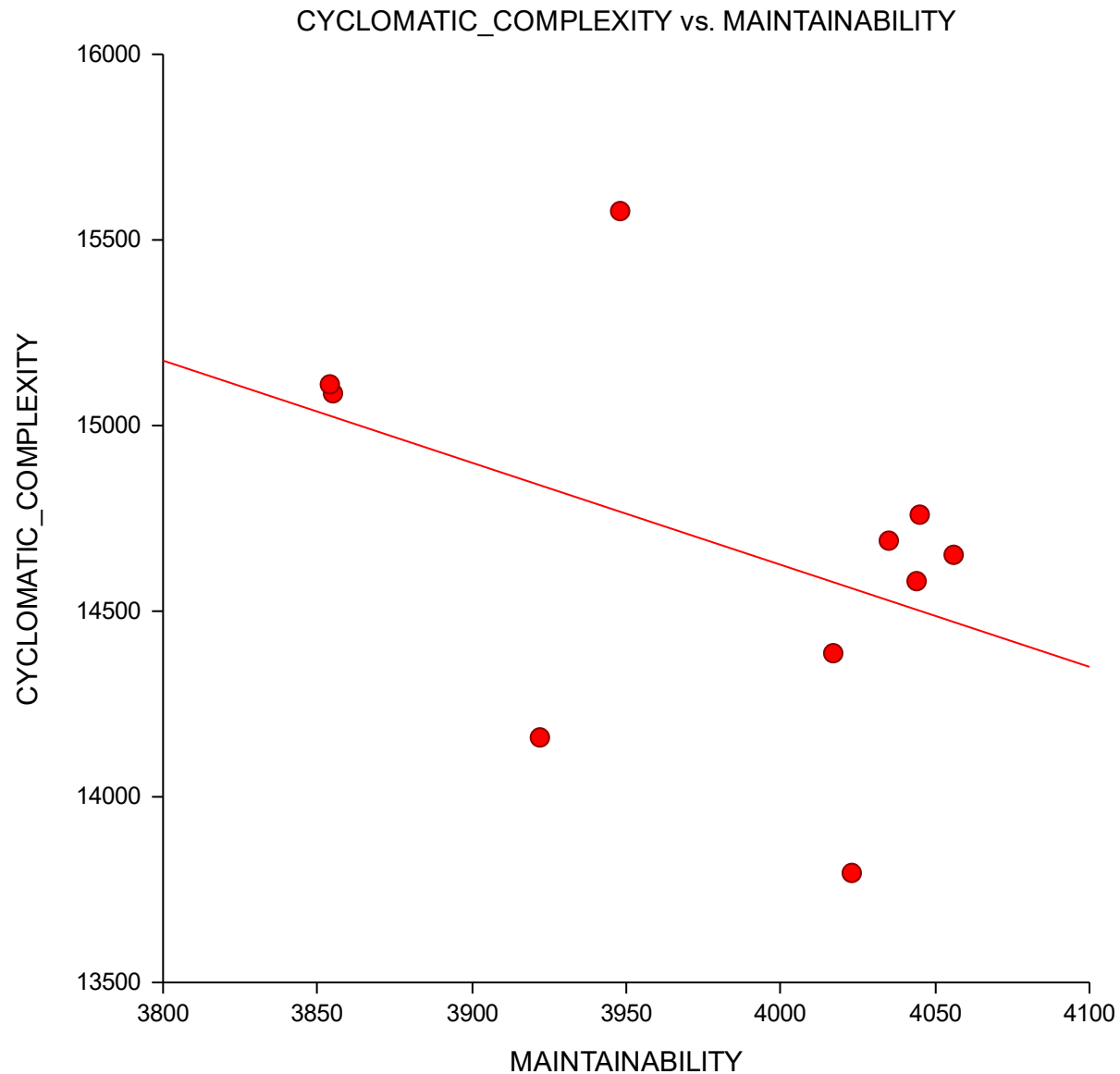
CYCLOMATIC_COMPLEXITY vs. MAINTAINABILITY

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	25677.0202	Rows Prediction Only	0
Slope	-2.7631	Sum of Frequencies	10
R-Squared	0.1855	Sum of Weights	10.0000
Correlation	-0.4307	Coefficient of Variation	0.0331
Mean Square Error	235634.5	Square Root of MSE	485.422

Summary Statement

The equation of the straight line relating CYCLOMATIC_COMPLEXITY and MAINTAINABILITY is estimated as: CYCLOMATIC_COMPLEXITY = (25677.0202) + (-2.7631) MAINTAINABILITY using the 10 observations in this dataset. The y-intercept, the estimated value of CYCLOMATIC_COMPLEXITY when MAINTAINABILITY is zero, is 25677.0202 with a standard error of 8148.7110. The slope, the estimated change in CYCLOMATIC_COMPLEXITY per unit change in MAINTAINABILITY, is -2.7631 with a standard error of 2.0471. The value of R-Squared, the proportion of the variation in CYCLOMATIC_COMPLEXITY that can be accounted for by variation in MAINTAINABILITY, is 0.1855. The correlation between CYCLOMATIC_COMPLEXITY and MAINTAINABILITY is -0.4307.

A significance test that the slope is zero resulted in a t-value of -1.3498. The significance level of this t-test is 0.2140. Since 0.2140 > 0.0500, the hypothesis that the slope is zero is not rejected.

The estimated slope is -2.7631. The lower limit of the 95% confidence interval for the slope is -7.4837 and the upper limit is 1.9575. The estimated intercept is 25677.0202. The lower limit of the 95% confidence interval for the intercept is 6886.0589 and the upper limit is 44467.9814.

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Regression Estimation Section ──

	Intercept	Slope
Parameter	**B(0)**	**B(1)**
Regression Coefficients	25677.0202	-2.7631
Lower 95% Confidence Limit	6886.0589	-7.4837
Upper 95% Confidence Limit	44467.9814	1.9575
Standard Error	8148.7110	2.0471
Standardized Coefficient	0.0000	-0.4307
T Value	3.1511	-1.3498
Prob Level (T Test)	0.0136	0.2140
Reject H0 (Alpha = 0.0500)	Yes	No
Power (Alpha = 0.0500)	0.7877	0.2220
Regression of Y on X	25677.0202	-2.7631
Inverse Regression from X on Y	73965.3964	-14.8962
Orthogonal Regression of Y and X	72797.5736	-14.6027

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
(25677.0201835632) + (-2.76311469724715) * (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	-0.4307	0.1855	-0.2848
Lower 95% Conf. Limit (r dist'n)	0.2596		
Upper 95% Conf. Limit (r dist'n)	-0.8122		
Lower 95% Conf. Limit (Fisher's z)	-0.8341		-0.7754
Upper 95% Conf. Limit (Fisher's z)	0.2730		0.4201
Adjusted (Rbar)		0.0837	
T-Value for H0: Rho = 0	1.3498	1.3498	0.8405
Prob Level for H0: Rho = 0	0.2140	0.2140	0.4250

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 ──

	X'X	X'X	X'Y	X'X Inverse	X'X Inverse
Index	0	1	2	0	1
0	10	39799	146801	281.7986	-0.07078033
1	39799	1.584523E+08	5.840979E+08	-0.07078033	1.778445E-05
2 (Y'Y)			2.157368E+09		
Determinant		562289			1.778445E-06

Variance - Covariance Matrix of Regression Coefficients ─────────────

	VC(b)	VC(b)
Index	0	1
0	6.640149E+07	-16678.29
1	-16678.29	4.19063

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	0.9043	0.244258	Yes
Anderson Darling	0.6213	0.105829	No
D'Agostino Skewness	-0.4600	0.645552	Yes
D'Agostino Kurtosis	0.6478	0.517142	Yes
D'Agostino Omnibus	0.6311	0.729371	Yes
Constant Residual Variance?			
Modified Levene Test	0.2732	0.615348	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.
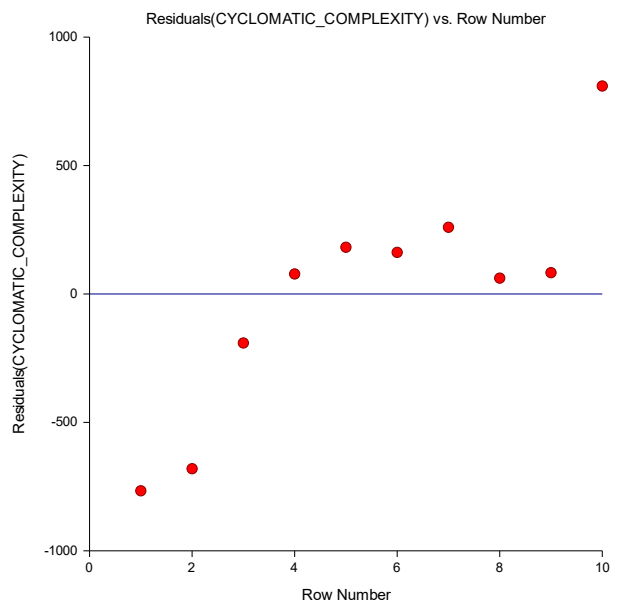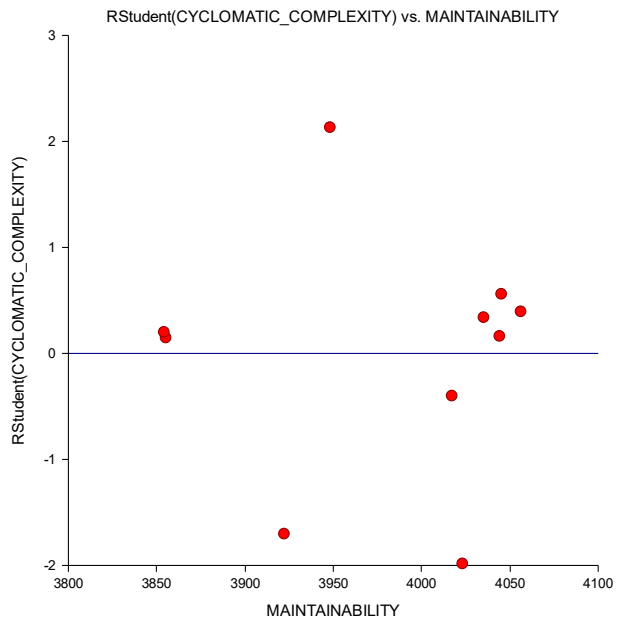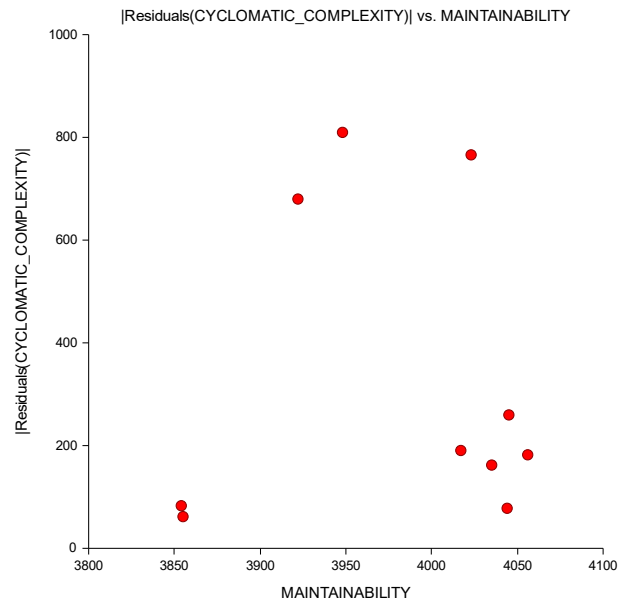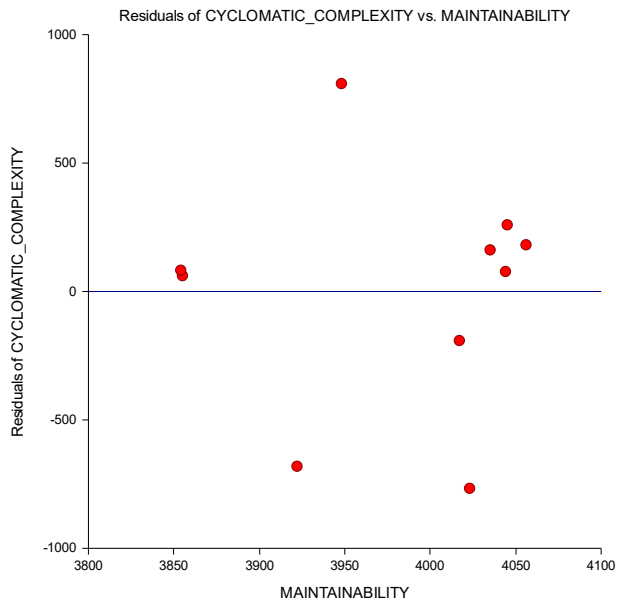
Straight-Line:
Possible remedies for the failure of this assumption include using nonlinear regression or
polynomial regression.

Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

Residual Plots Section
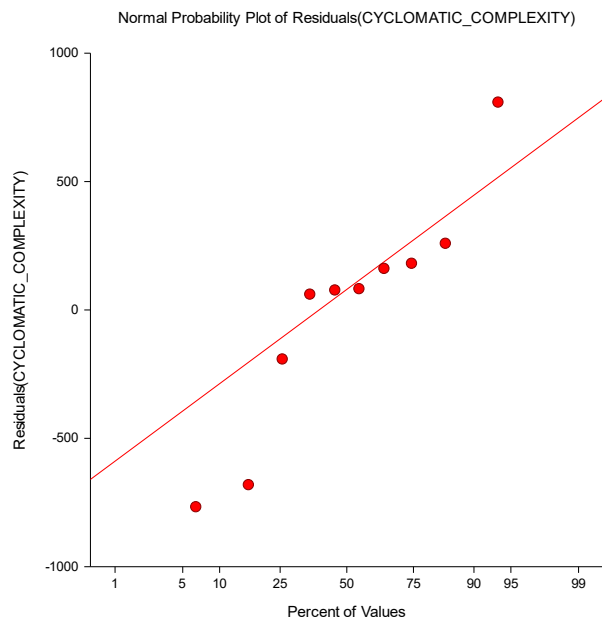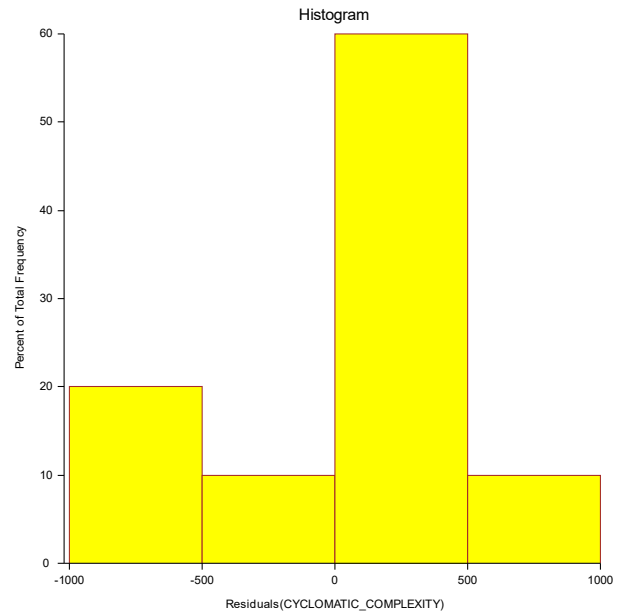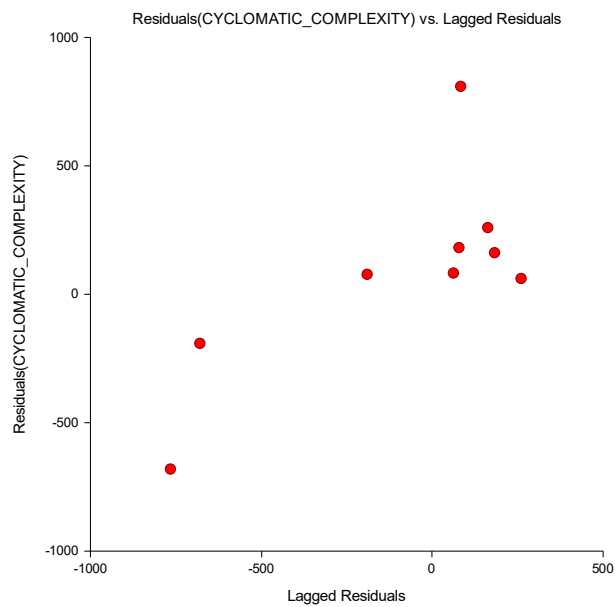
Linear Regression Report

Dataset   Untitled
Y = CYCLOMATIC_COMPLEXITY   X = MAINTAINABILITY

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_2_6.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