Linear Regression Report

Dataset   Untitled
Y = Coupling   X = no_of_Issues

Linear Regression Plot Section ──────────────────────────────────────
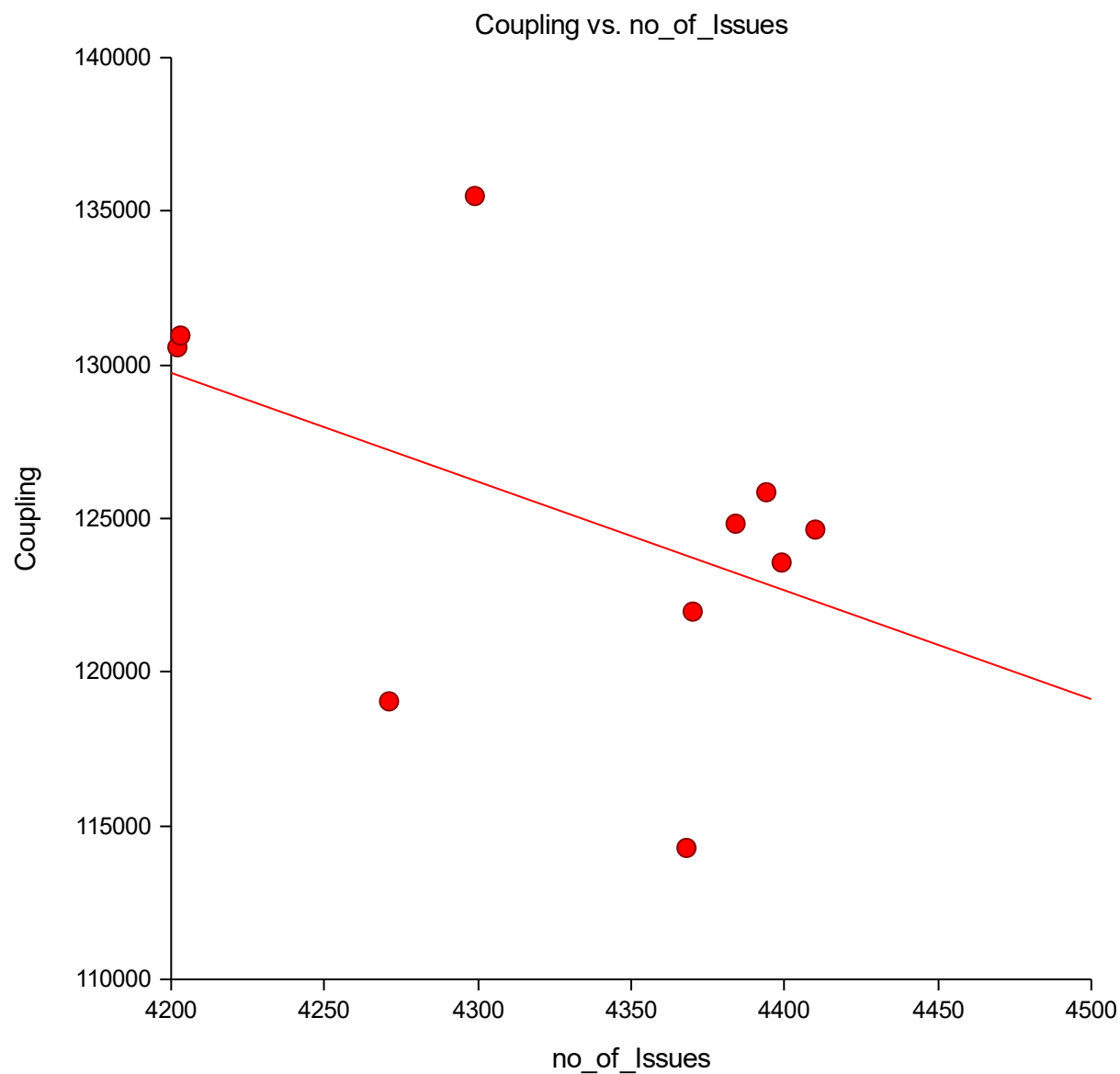
Coupling vs. no_of_Issues

Linear Regression Report

Dataset   Untitled
Y = Coupling   X = no_of_Issues

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

Parameter	Value	Parameter	Value
Dependent Variable	Coupling	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	278077.4590	Rows Prediction Only	0
Slope	-35.3242	Sum of Frequencies	10
R-Squared	0.2141	Sum of Weights	10.0000
Correlation	-0.4627	Coefficient of Variation	0.0461
Mean Square Error	3.330399E+07	Square Root of MSE	5770.961

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

The equation of the straight line relating Coupling and no_of_Issues is estimated as: Coupling = (278077.4590) + (-35.3242) no_of_Issues using the 10 observations in this dataset. The y-intercept, the estimated value of Coupling when no_of_Issues is zero, is 278077.4590 with a standard error of 103620.6665. The slope, the estimated change in Coupling per unit change in no_of_Issues, is -35.3242 with a standard error of 23.9272. The value of R-Squared, the proportion of the variation in Coupling that can be accounted for by variation in no_of_Issues, is 0.2141. The correlation between Coupling and no_of_Issues is -0.4627.

A significance test that the slope is zero resulted in a t-value of -1.4763. The significance level of this t-test is 0.1781. Since 0.1781 > 0.0500, the hypothesis that the slope is zero is not rejected.

The estimated slope is -35.3242. The lower limit of the 95% confidence interval for the slope is -90.5003 and the upper limit is 19.8519. The estimated intercept is 278077.4590. The lower limit of the 95% confidence interval for the intercept is 39127.7736 and the upper limit is 517027.1445.

Linear Regression Report

Dataset   Untitled
Y = Coupling   X = no_of_Issues

Regression Estimation Section ───────────────────────────────────────

Parameter	Intercept B(0)	Slope B(1)
Regression Coefficients	278077.4590	-35.3242
Lower 95% Confidence Limit	39127.7736	-90.5003
Upper 95% Confidence Limit	517027.1445	19.8519
Standard Error	103620.6665	23.9272
Standardized Coefficient	0.0000	-0.4627
T Value	2.6836	-1.4763
Prob Level (T Test)	0.0278	0.1781
Reject H0 (Alpha = 0.0500)	Yes	No
Power (Alpha = 0.0500)	0.6534	0.2561
Regression of Y on X	278077.4590	-35.3242
Inverse Regression from X on Y	839497.4434	-164.9824
Orthogonal Regression of Y and X	839401.1134	-164.9601

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
(278077.459024966) + (-35.3241937701969) * (no_of_Issues)

Linear Regression Report

Dataset   Untitled
Y = Coupling   X = no_of_Issues

Correlation and R-Squared Section ─────────────────────────────────

Parameter	Pearson Correlation Coefficient	R-Squared	Spearman Rank Correlation Coefficient
Estimated Value	-0.4627	0.2141	-0.2848
Lower 95% Conf. Limit (r dist'n)	0.2244		
Upper 95% Conf. Limit (r dist'n)	-0.8249		
Lower 95% Conf. Limit (Fisher's z)	-0.8459		-0.7754
Upper 95% Conf. Limit (Fisher's z)	0.2355		0.4201
Adjusted (Rbar)		0.1159	
T-Value for H0: Rho = 0	1.4763	1.4763	0.8405
Prob Level for H0: Rho = 0	0.1781	0.1781	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 0	X'X 1	X'Y 2	X'X Inverse 0	X'X Inverse 1
Index	**0**	**1**	**2**	**0**	**1**
0	10	43300	1251237	322.4011	-0.07443444
1	43300	1.875472E+08	5.415801E+09	-0.07443444	1.71904E-05
2 (Y'Y)			1.568984E+11		
Determinant		581720			1.71904E-06

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

	VC(b) 0	VC(b) 1
Index	**0**	**1**
0	1.073724E+10	-2478964
1	-2478964	572.5089

Linear Regression Report

Dataset   Untitled
Y = Coupling   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	0.8929	0.182945	No
Anderson Darling	0.6934	0.070234	No
D'Agostino Skewness	-0.7215	0.470593	Yes
D'Agostino Kurtosis	0.6841	0.493933	Yes
D'Agostino Omnibus	0.9885	0.610019	Yes
Constant Residual Variance?			
Modified Levene Test	3.6216	0.093529	No
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 = Coupling   X = no_of_Issues

Residual Plots Section
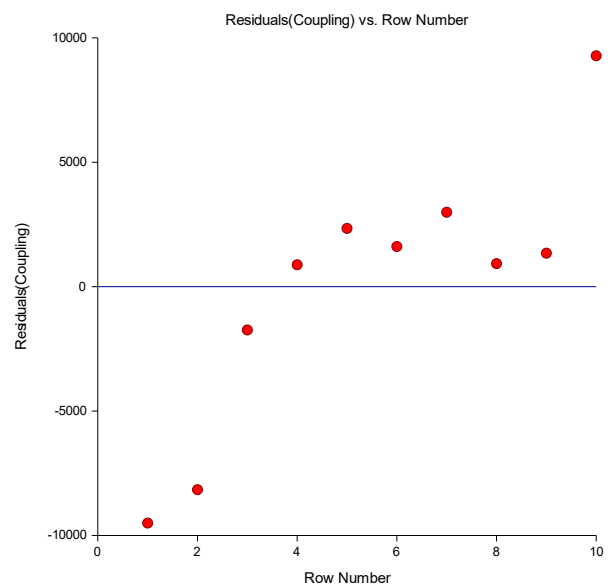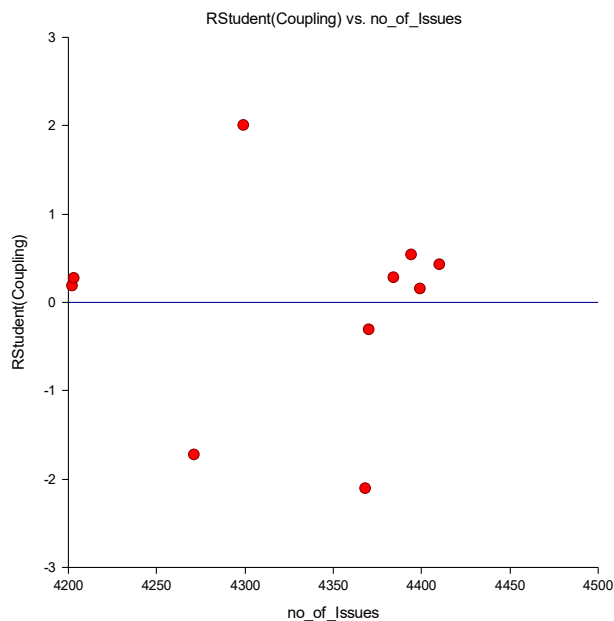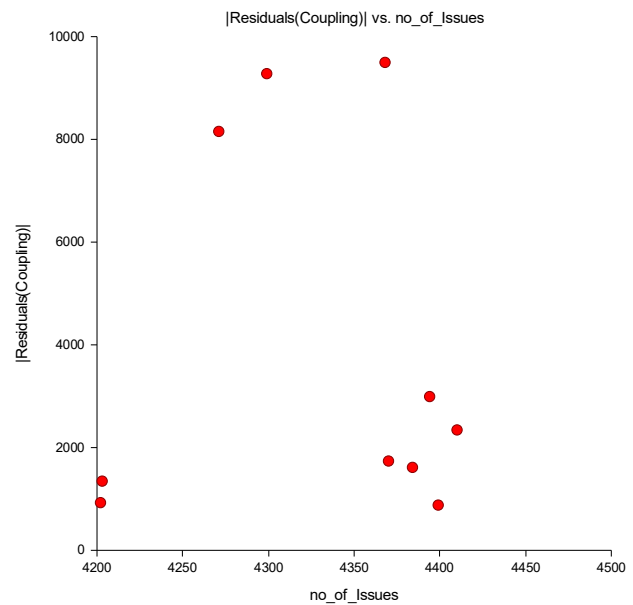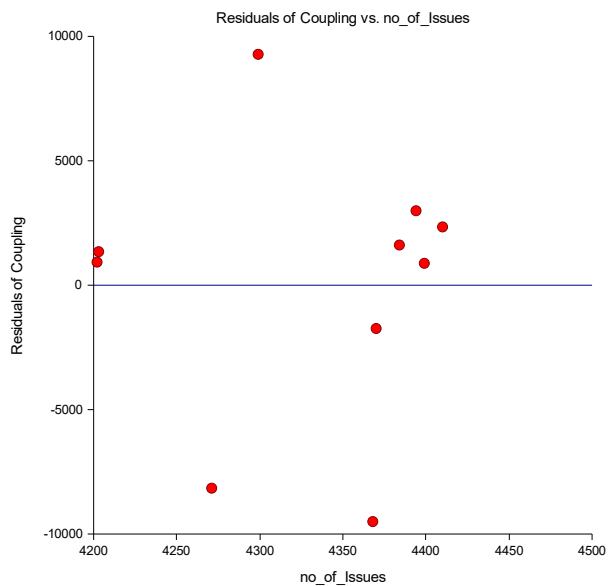
Linear Regression Report

Dataset   Untitled
Y = Coupling   X = no_of_Issues

Residuals(Coupling) vs. Lagged Residuals

Histogram

Normal Probability Plot of Residuals(Coupling)
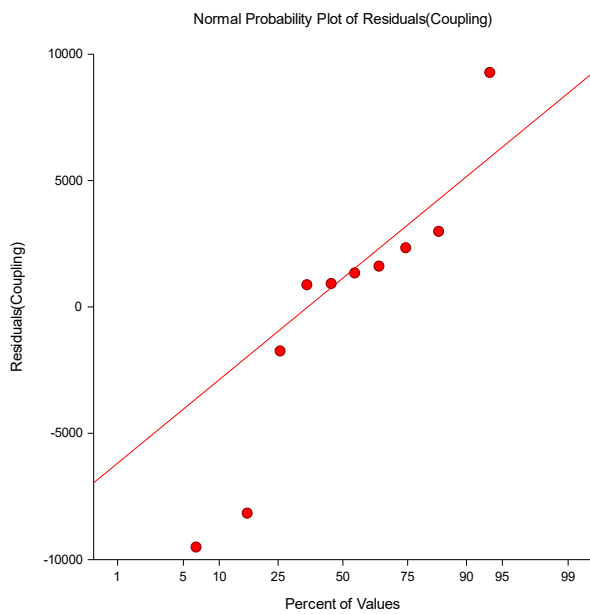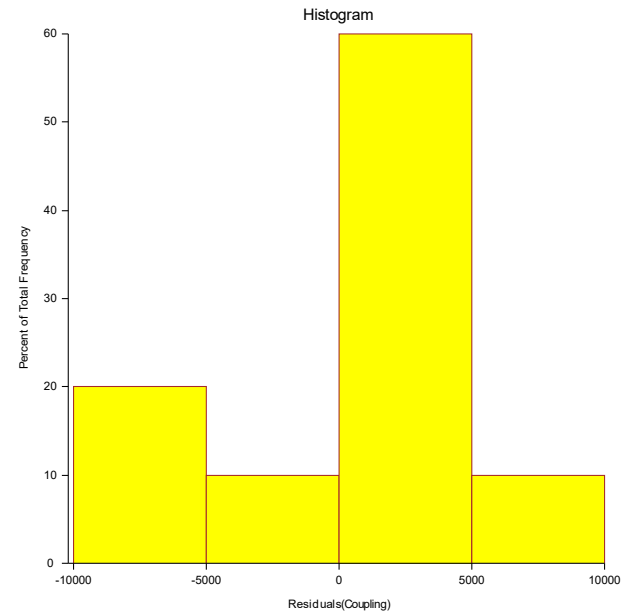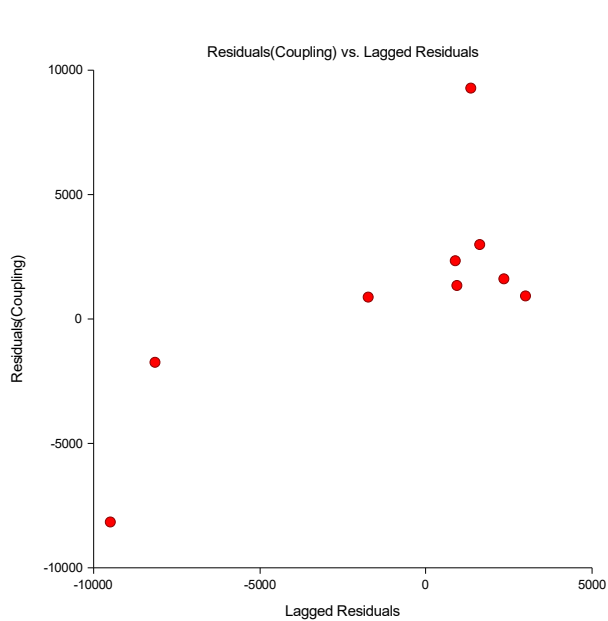
Linear Regression Report

Dataset   Untitled
Y = Coupling   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_29-23_11_31.t153

Variables Tab
-- Variables ---
Y: Dependent Variable(s):   Coupling
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

<div align="center">

Linear Regression Report

</div>

Dataset   Untitled
Y = Coupling   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

Linear Regression Report

Dataset   Untitled
Y = Coupling   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