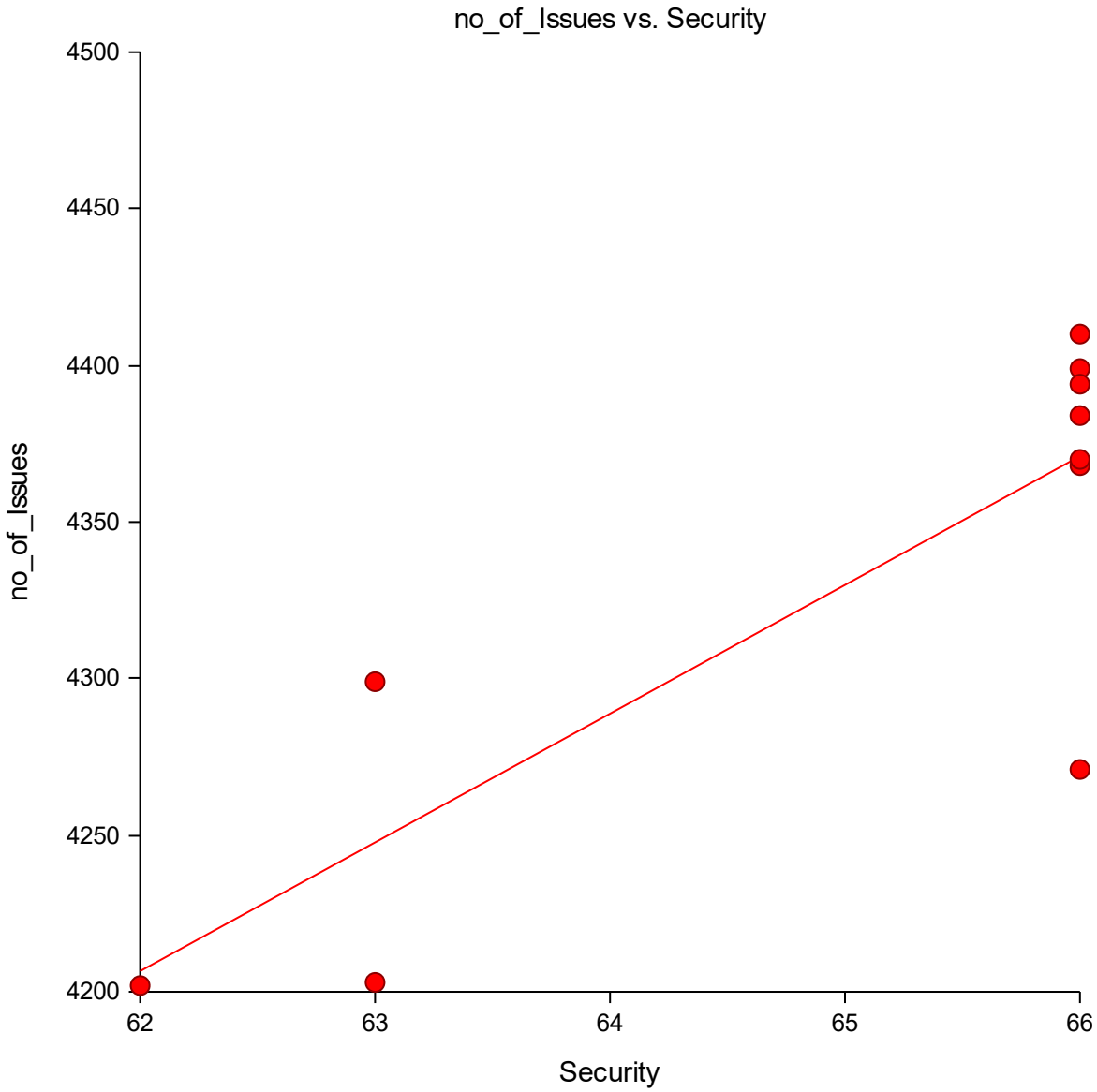
Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security

Linear Regression Plot Section



Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security

Run Summary Section

Parameter	Value	Parameter	Value
Dependent Variable	no_of_Issues	Rows Processed	10
Independent Variable	Security	Rows Used in Estimation	10
Frequency Variable	None	Rows with X Missing	0
Weight Variable	None	Rows with Freq Missing	0
Intercept	1659.5833	Rows Prediction Only	0
Slope	41.0833	Sum of Frequencies	10
R-Squared	0.6964	Sum of Weights	10.0000
Correlation	0.8345	Coefficient of Variation	0.0109
Mean Square Error	2207.979	Square Root of MSE	46.98914

Summary Statement

The equation of the straight line relating no_of_Issues and Security is estimated as: no_of_Issues = (1659.5833) + (41.0833) Security using the 10 observations in this dataset. The y-intercept, the estimated value of no_of_Issues when Security is zero, is 1659.5833 with a standard error of 623.6322. The slope, the estimated change in no_of_Issues per unit change in Security, is 41.0833 with a standard error of 9.5916. The value of R-Squared, the proportion of the variation in no_of_Issues that can be accounted for by variation in Security, is 0.6964. The correlation between no_of_Issues and Security is 0.8345.

A significance test that the slope is zero resulted in a t-value of 4.2833. The significance level of this t-test is 0.0027. Since 0.0027 < 0.0500, the hypothesis that the slope is zero is rejected.

The estimated slope is 41.0833. The lower limit of the 95% confidence interval for the slope is 18.9650 and the upper limit is 63.2016. The estimated intercept is 1659.5833. The lower limit of the 95% confidence interval for the intercept is 221.4849 and the upper limit is 3097.6818.

Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security

Regression Estimation Section

Parameter	Intercept B(0)	Slope B(1)
Regression Coefficients	1659.5833	41.0833
Lower 95% Confidence Limit	221.4849	18.9650
Upper 95% Confidence Limit	3097.6818	63.2016
Standard Error	623.6322	9.5916
Standardized Coefficient	0.0000	0.8345
T Value	2.6612	4.2833
Prob Level (T Test)	0.0288	0.0027
Reject H0 (Alpha = 0.0500)	Yes	Yes
Power (Alpha = 0.0500)	0.6462	0.9616
Regression of Y on X	1659.5833	41.0833
Inverse Regression from X on Y	495.1318	58.9980
Orthogonal Regression of Y and X	495.6121	58.9906

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
(1659.58333333302) + (41.0833333333139) * (Security)

Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security

Correlation and R-Squared Section

Parameter	Pearson Correlation Coefficient	R-Squared	Spearman Rank Correlation Coefficient
Estimated Value	0.8345	0.6964	0.7416
Lower 95% Conf. Limit (r dist'n)	0.4092		
Upper 95% Conf. Limit (r dist'n)	0.9524		
Lower 95% Conf. Limit (Fisher's z)	0.4316		0.2101
Upper 95% Conf. Limit (Fisher's z)	0.9598		0.9348
Adjusted (Rbar)		0.6584	
T-Value for H0: Rho = 0	4.2833	4.2833	3.1269
Prob Level for H0: Rho = 0	0.0027	0.0027	0.0141

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	650	43300	176.1417	-2.708333
1	650	42274	2815486	-2.708333	0.04166667
2 (Y'Y)			1.875472E+08		
Determinant		240			0.004166667

Variance - Covariance Matrix of Regression Coefficients

Index	VC(b) 0	VC(b) 1
0	388917.1	-5979.943
1	-5979.943	91.99913

Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security

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	0.8857	0.151721	No
Anderson Darling	0.5483	0.158328	No
D'Agostino Skewness	-1.9876	0.046855	No
D'Agostino Kurtosis	1.4931	0.135414	No
D'Agostino Omnibus	6.1799	0.045504	No
Constant Residual Variance?			
Modified Levene Test	0.1494	0.709226	Yes
Relationship is a Straight Line?			
Lack of Linear Fit F(1, 7) Test	0.0171	0.899718	Yes

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.

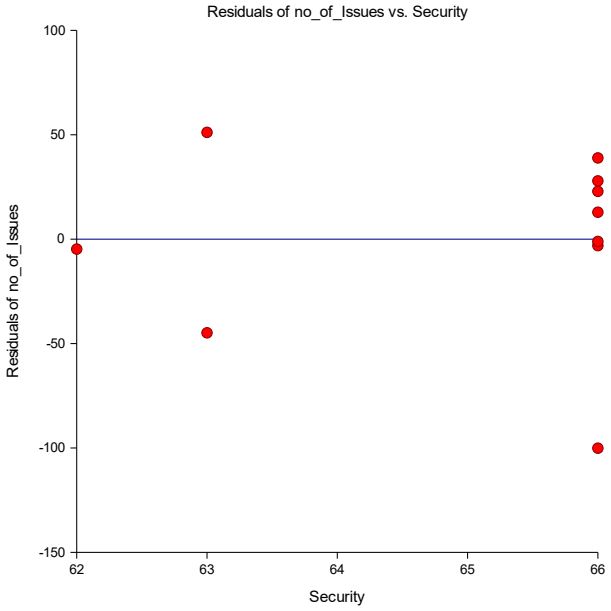
Normality and Constant Residual Variance:
Possible remedies for the failure of these assumptions include using a transformation of Y such as the log or square root, correcting data-recording errors found by looking into outliers, adding additional independent variables, using robust regression, or using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

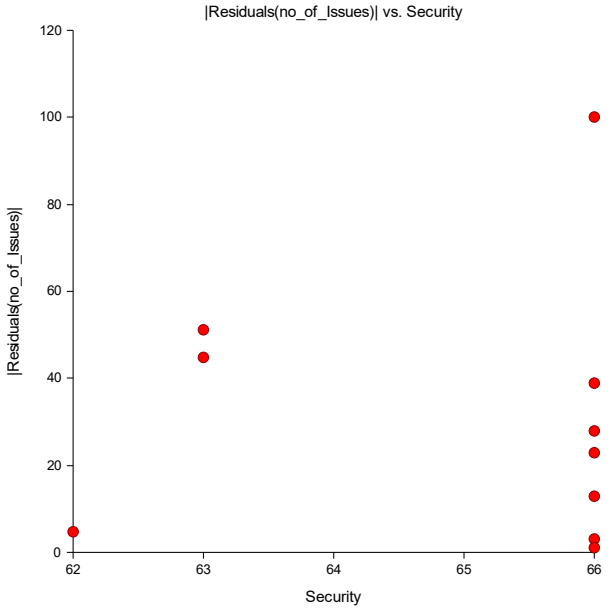
Straight-Line:
Possible remedies for the failure of this assumption include using nonlinear regression or polynomial regression.

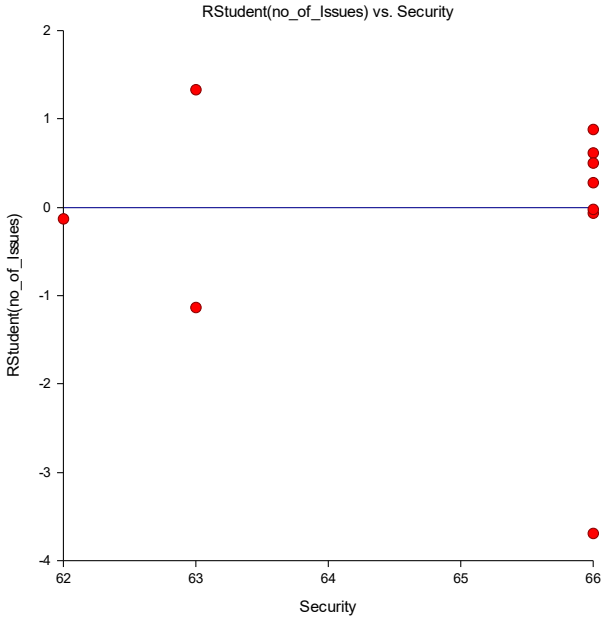
Linear Regression Report

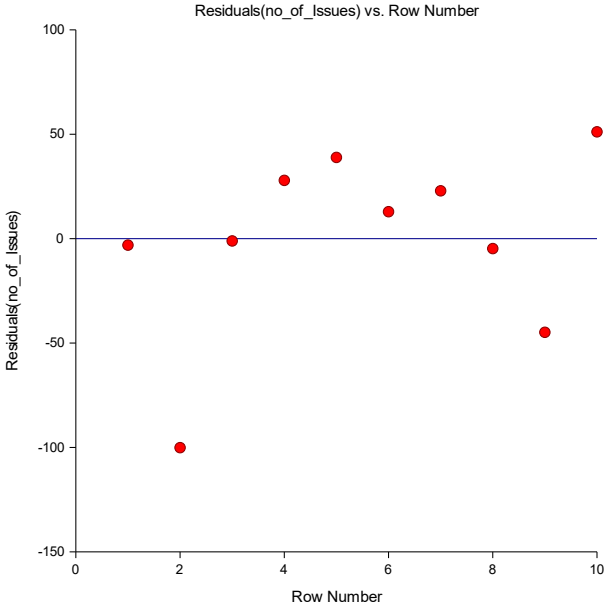
Dataset Untitled
Y = no_of_Issues X = Security

Residual Plots Section



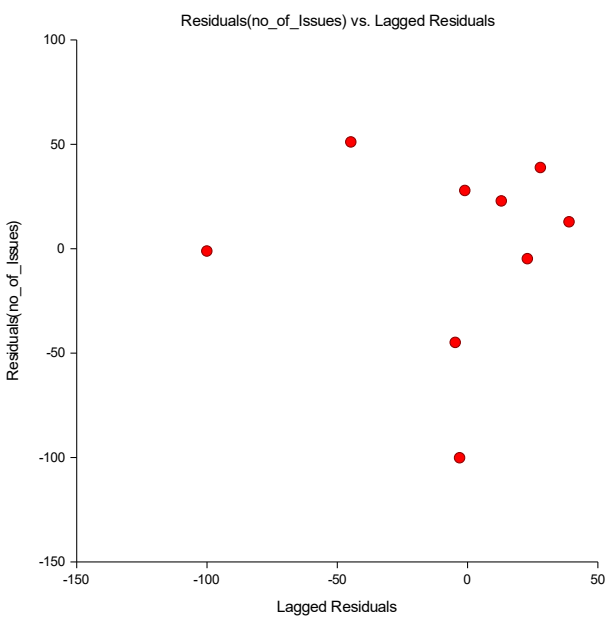


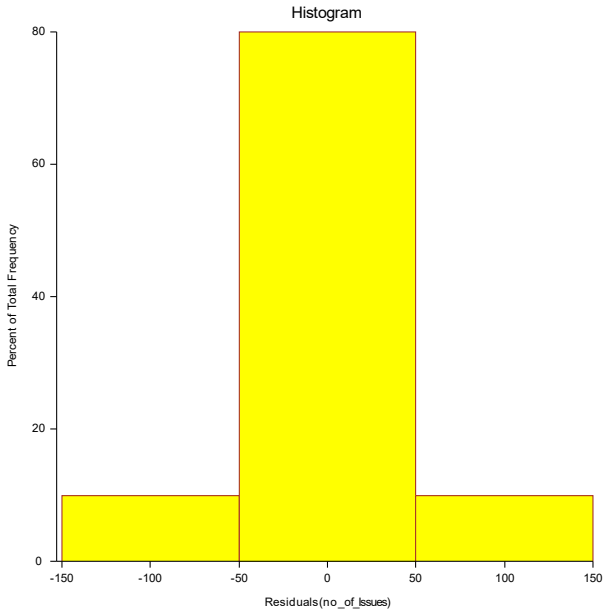


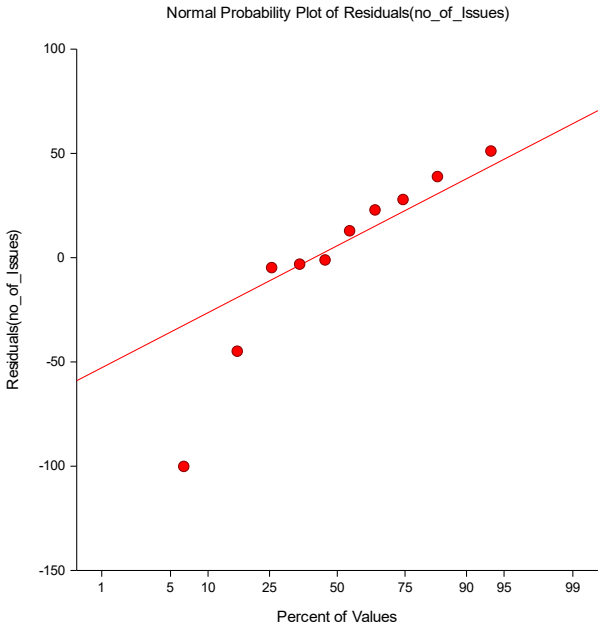


Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security







Linear Regression Report

Dataset Untitled
Y = no_of_Issues X = Security

Procedure Input Settings

Autosaved Template File
C:\Users\KASATLA\Documents\NCSS 12\Procedure Templates\Autosave\Linear Regression and Correlation - Autosaved 2018_3_30-0_10_25.t153

Variables Tab

-- Variables	
Y: Dependent Variable(s):	no_of_Issues
X: Independent Variable:	Security
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 = no_of_Issues X = Security

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 = no_of_Issues X = Security

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