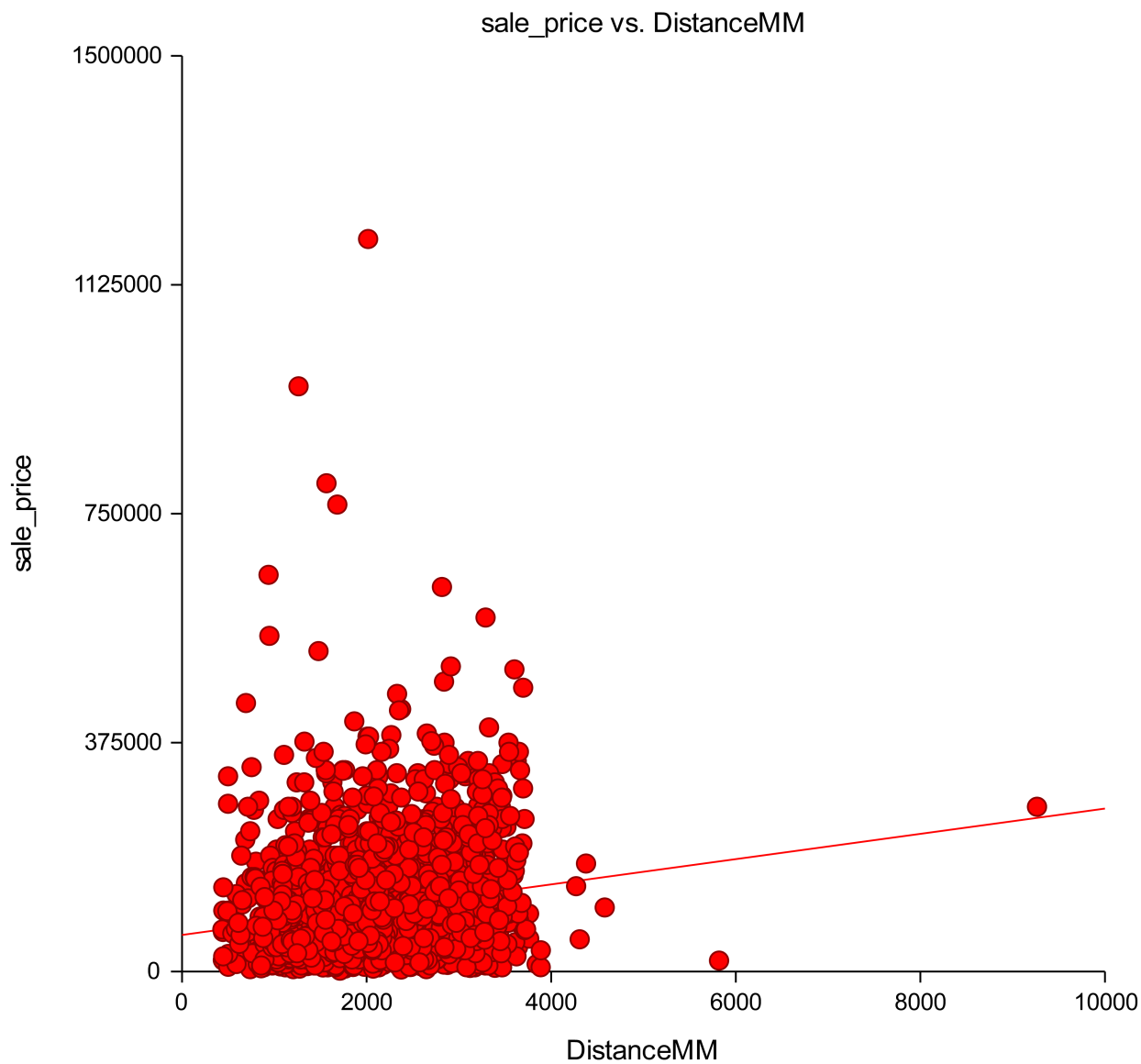


Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
Y = sale_price X = DistanceMM

Linear Regression Plot Section



Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

Run Summary Section

Parameter	Value	Parameter	Value
Dependent Variable	sale_price	Rows Processed	1826
Independent Variable	DistanceMM	Rows Used in Estimation	1815
Frequency Variable	None	Rows with X Missing	11
Weight Variable	None	Rows with Freq Missing	0
Intercept	60244.2560	Rows Prediction Only	0
Slope	20.7318	Sum of Frequencies	1815
R-Squared	0.0278	Sum of Weights	1815.0000
Correlation	0.1666	Coefficient of Variation	0.8998
Mean Square Error	8702106748.10817	Square Root of MSE	93285.083202558

Summary Statement

The equation of the straight line relating sale_price and DistanceMM is estimated as: $\text{sale_price} = (60244.2560) + (20.7318) \text{DistanceMM}$ using the 1815 observations in this dataset. The y-intercept, the estimated value of sale_price when DistanceMM is zero, is 60244.2560 with a standard error of 6422.5714. The slope, the estimated change in sale_price per unit change in DistanceMM, is 20.7318 with a standard error of 2.8820. The value of R-Squared, the proportion of the variation in sale_price that can be accounted for by variation in DistanceMM, is 0.0278. The correlation between sale_price and DistanceMM is 0.1666.

A significance test that the slope is zero resulted in a t-value of 7.1935. 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 20.7318. The lower limit of the 95% confidence interval for the slope is 15.0832 and the upper limit is 26.3804. The estimated intercept is 60244.2560. The lower limit of the 95% confidence interval for the intercept is 47656.2473 and the upper limit is 72832.2647.

Descriptive Statistics Section

Parameter	Dependent	Independent
Variable	sale_price	DistanceMM
Count	1815	1815
Mean	103677.2138	2094.9910
Standard Deviation	94580.9111	759.9726
Minimum	1500.0000	436.3154
Maximum	1200000.0000	9263.5934

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

Regression Estimation Section

Parameter	Intercept B(0)	Slope B(1)
Regression Coefficients	60244.2560	20.7318
Lower 95% Confidence Limit	47656.2473	15.0832
Upper 95% Confidence Limit	72832.2647	26.3804
Standard Error	6422.5714	2.8820
Standardized Coefficient	0.0000	0.1666
T Value	9.3801	7.1935
Prob Level (T Test)	0.0000	0.0000
Prob Level (Randomization Test*)		0.0010
Reject H0 (Alpha = 0.0500)	Yes	Yes
Power (Alpha = 0.0500)	1.0000	1.0000
Regression of Y on X	60244.2560	20.7318
Inverse Regression from X on Y	-1461473.1434	747.0917
Orthogonal Regression of Y and X	-1461374.8958	747.0448

Estimated Model

$(60244.2559851201) + (20.7318113055527) * (\text{DistanceMM})$

* Number of Monte Carlo Samples = 1000, Computer-Generated Random Seed = 5912797.

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.

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

Bootstrap Section

Estimation Results		Bootstrap Confidence Limits		
Parameter	Estimate	Conf. Level	Lower	Upper
Intercept				
Original Value	60244.2560	0.9000	49392.3947	71014.0943
Bootstrap Mean	60211.4147	0.9500	47461.8399	73357.7508
Bias (BM - OV)	-32.8413	0.9900	42461.6673	78490.5867
Bias Corrected	60277.0973			
Standard Error	6599.8170			
Slope				
Original Value	20.7318	0.9000	15.7629	25.6049
Bootstrap Mean	20.7240	0.9500	14.7137	26.5600
Bias (BM - OV)	-0.0078	0.9900	12.7349	28.8966
Bias Corrected	20.7396			
Standard Error	3.0103			
Correlation				
Original Value	0.1666	0.9000	0.1225	0.2096
Bootstrap Mean	0.1672	0.9500	0.1125	0.2181
Bias (BM - OV)	0.0006	0.9900	0.0895	0.2381
Bias Corrected	0.1660			
Standard Error	0.0269			
R-Squared				
Original Value	0.0278	0.9000	0.0111	0.0402
Bootstrap Mean	0.0287	0.9500	0.0068	0.0423
Bias (BM - OV)	0.0009	0.9900	0.0000	0.0465
Bias Corrected	0.0268			
Standard Error	0.0091			
Standard Error of Estimate				
Original Value	93285.0832	0.9000	84587.4508	101605.4523
Bootstrap Mean	93032.0145	0.9500	82207.4710	103026.4878
Bias (BM - OV)	-253.0687	0.9900	78513.8706	105103.7171
Bias Corrected	93538.1519			
Standard Error	5138.1093			
Orthogonal Intercept				
Original Value	-1461374.8958	0.9000	-1862183.2755	-794603.6843
Bootstrap Mean	-1506012.0688	0.9500	-1924115.2622	-606630.7655
Bias (BM - OV)	-44637.1730	0.9900	-2036997.9272	-126080.1108
Bias Corrected	-1416737.7228			
Standard Error	336200.4092			

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

Bootstrap Section (Continued)

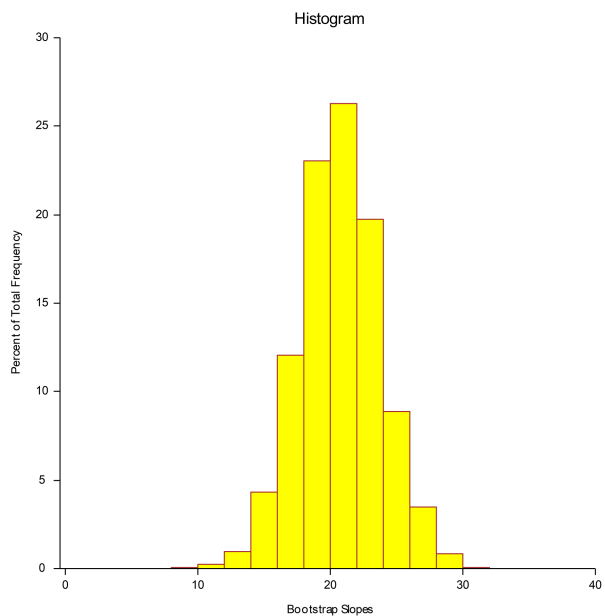
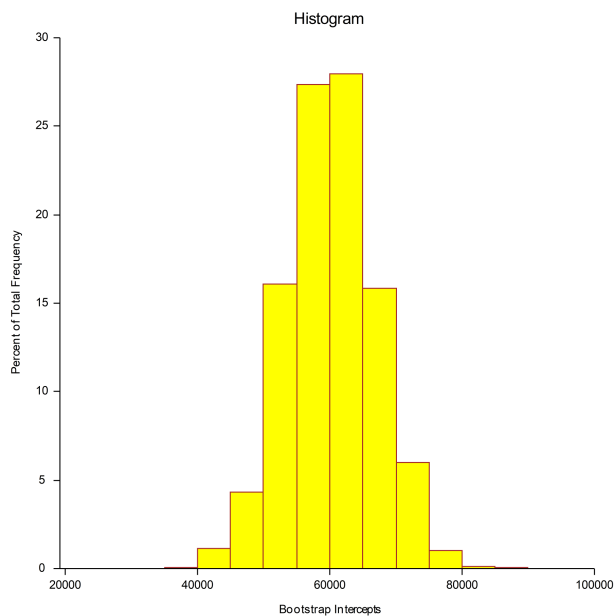
Estimation Results		Bootstrap Confidence Limits		
Parameter	Estimate	Conf. Level	Lower	Upper
Orthogonal Slope				
Original Value	747.0448	0.9000	427.9366	939.0933
Bootstrap Mean	768.6035	0.9500	334.0236	969.2325
Bias (BM - OV)	21.5587	0.9900	100.6510	1025.2509
Bias Corrected	725.4861			
Standard Error	161.6213			

Sampling Method = Observation, Confidence Limit Type = Reflection, Number of Samples = 3000,
 Computer-Generated Random Seed = 5912797.

Notes:

The main purpose of this report is to present the bootstrap confidence intervals of various parameters. All gross outliers should have been removed. The sample size should be at least 50 and the sample should be 'representative' of the population it was drawn from.

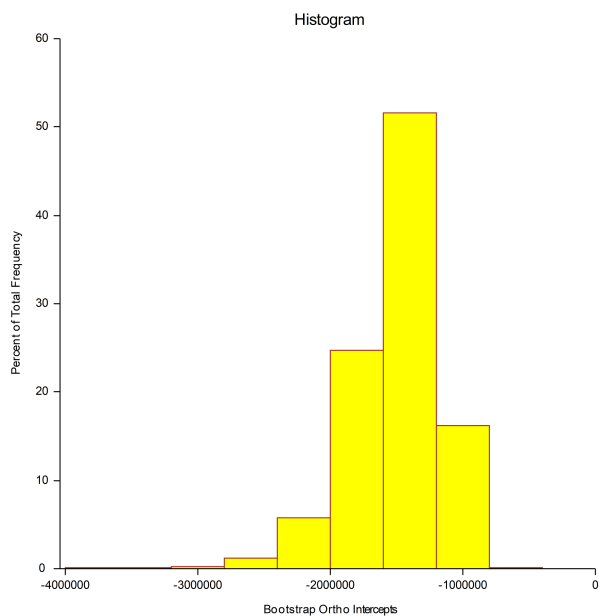
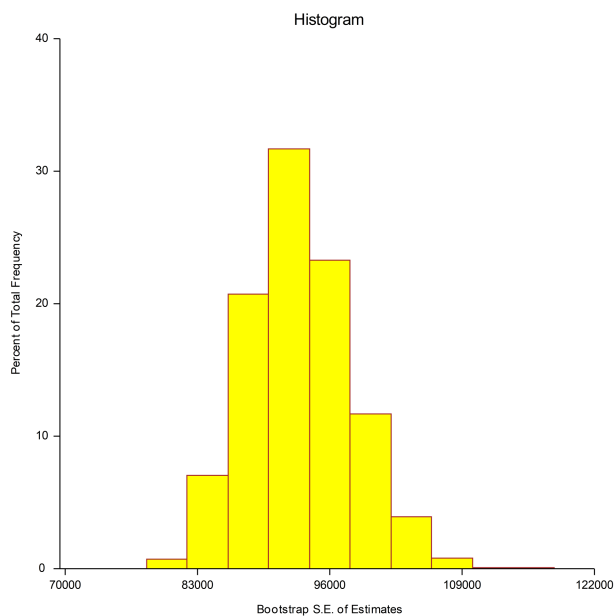
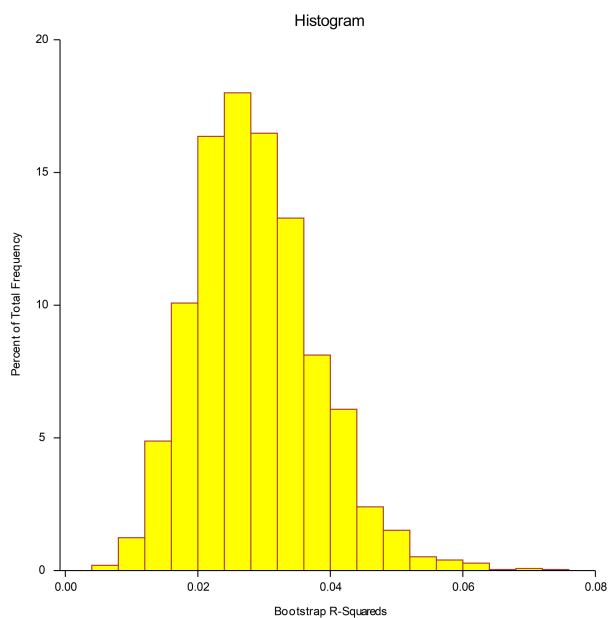
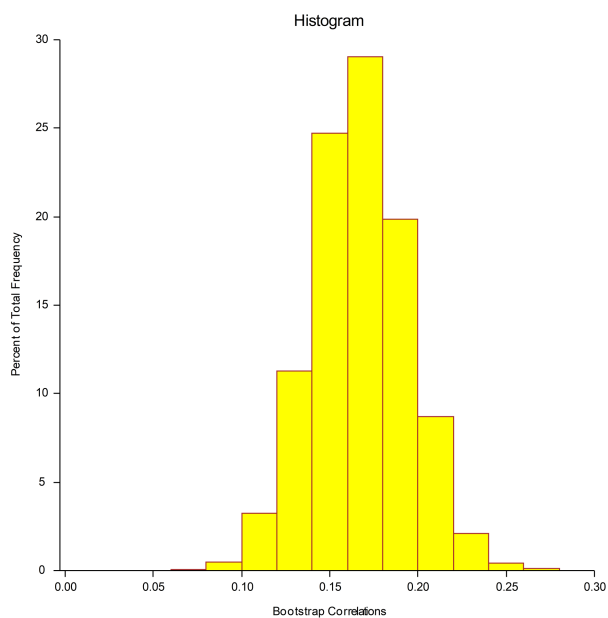
Bootstrap Histograms Section



Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
Y = sale_price X = DistanceMM

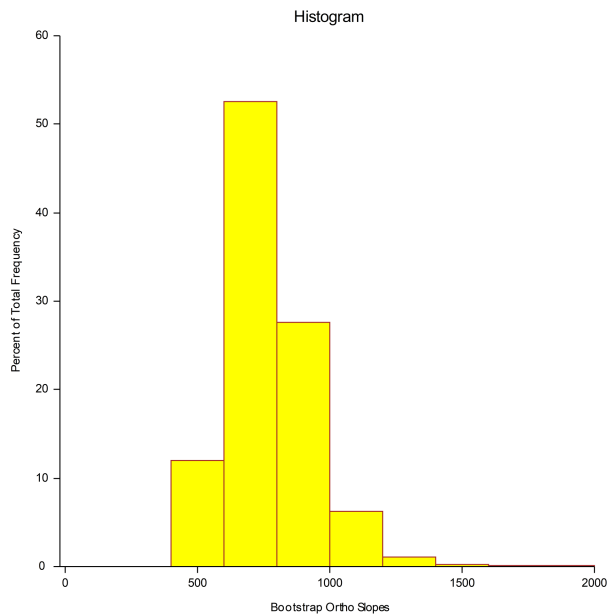
Bootstrap Histograms Section (Continued)



Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

Bootstrap Histograms Section (Continued)



Analysis of Variance Section

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (5%)
Intercept	1	19509370850569.9	19509370850569.9			
Slope	1	450305871800.945	450305871800.945	51.7468	0.0000	1.0000
Error	1813	15776919534320.1	8702106748.10817			
Lack of Fit	1636	14609033789193.9	8929727255.00852	1.3534	0.0051	
Pure Error	177	1167885745126.17	6598224548.7354			
Adj. Total	1814	16227225406121.1	8945548735.45813			
Total	1815	35736596256691				

Standard Deviation of Residuals

$s = \text{Square Root}(8702106748.10817) = 93285.083202558$

Notes:

The above report shows the F-Ratio for testing whether the slope is zero, the degrees of freedom, and the mean square error. The mean square error, which estimates the variance of the residuals, is used extensively in the calculation of hypothesis tests and confidence intervals.

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

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	0.7790	0.000000	No
Anderson Darling	82.1619	0.000000	No
D'Agostino Skewness	29.0897	0.000000	No
D'Agostino Kurtosis	21.4430	0.000000	No
D'Agostino Omnibus	1306.0129	0.000000	No
Constant Residual Variance?			
Modified Levene Test	6.7577	0.009410	No
Relationship is a Straight Line?			
Lack of Linear Fit F(1636, 177) Test	1.3534	0.005148	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.

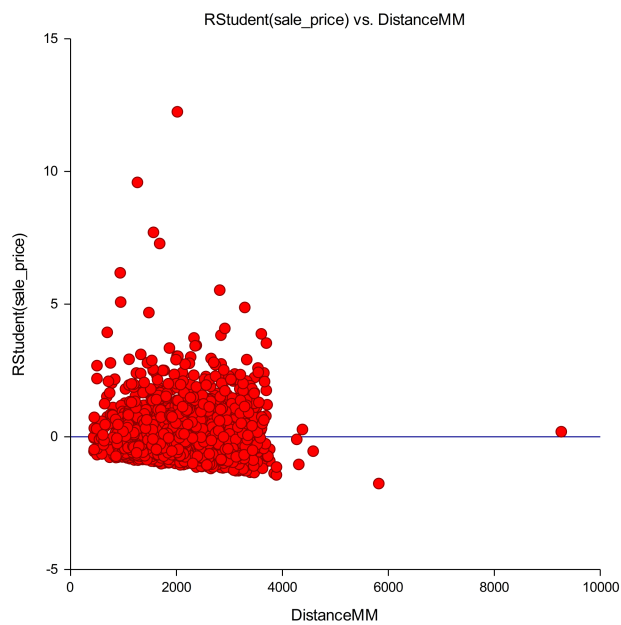
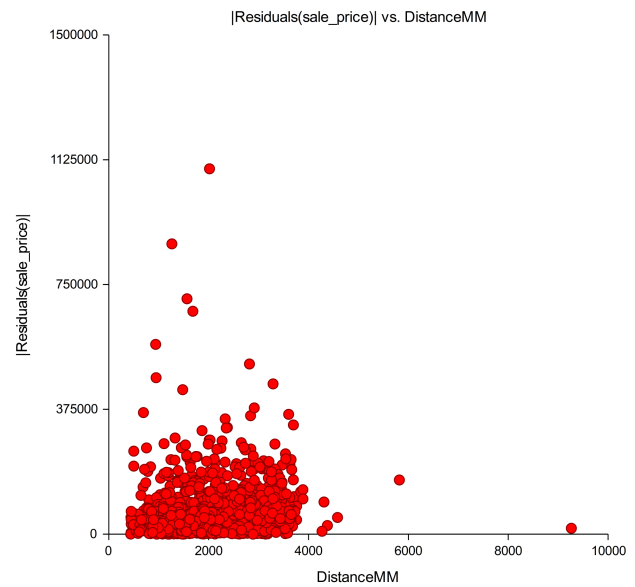
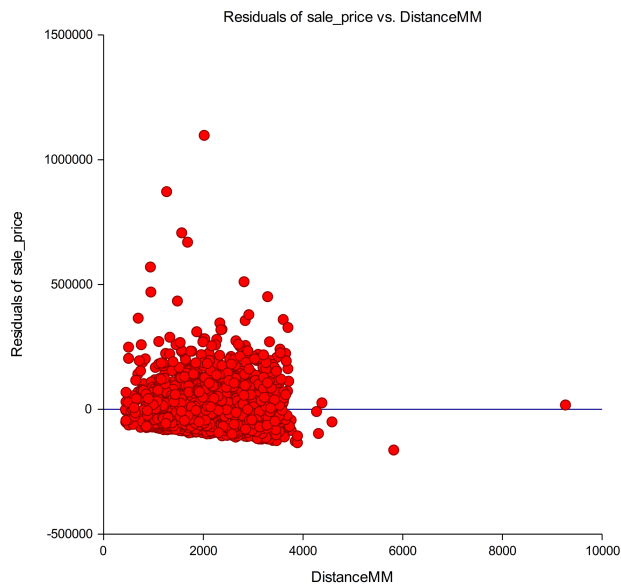
Straight-Line:

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

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
Y = sale_price X = DistanceMM

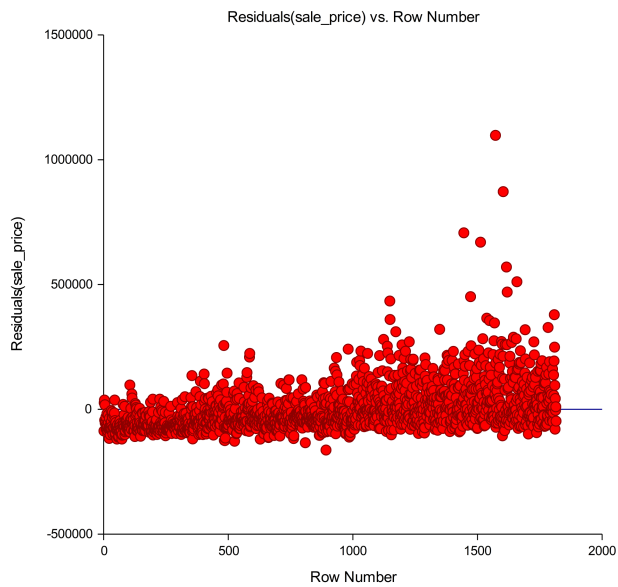
Residuals vs X Plots



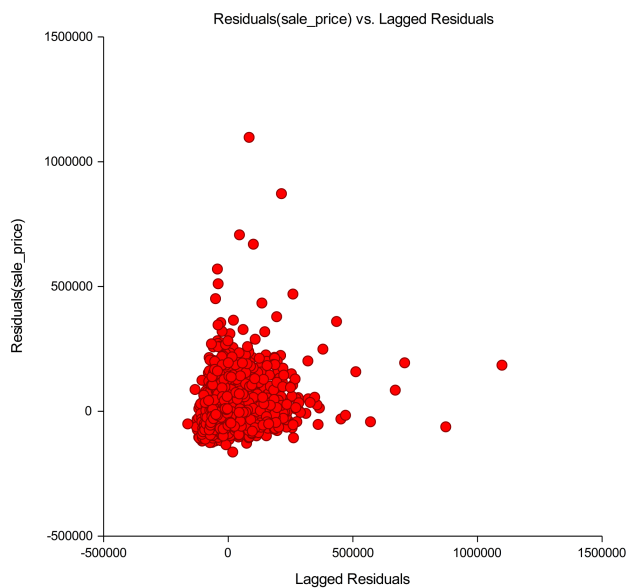
Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
Y = sale_price X = DistanceMM

Sequence Plot: Residuals vs Row



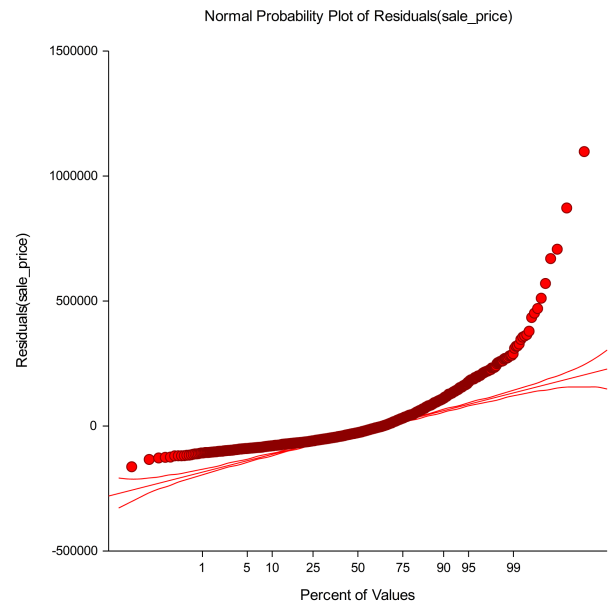
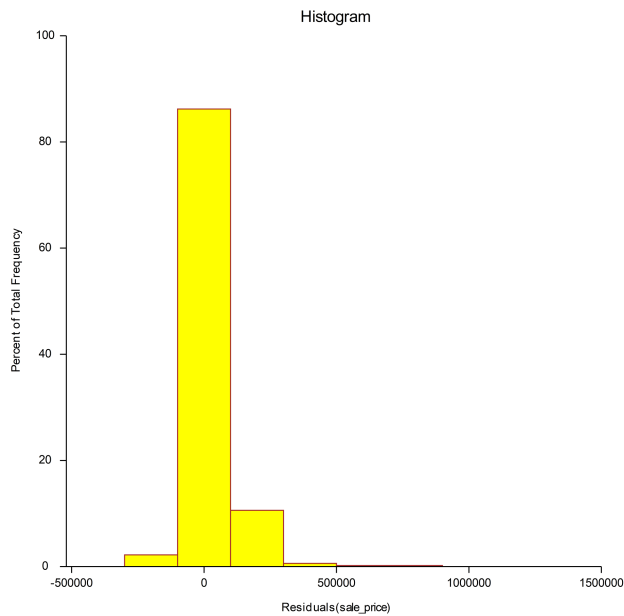
Serial Correlation Plot: Residuals vs Lagged Residuals



Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
Y = sale_price X = DistanceMM

Distributional Plots of Residuals



Storage Report

Stored Item	Column Number	Column Name
Predicted Y	57	Predicted_sale_price
Predicted X	58	Predicted_DistanceMM
Residuals	59	Residualsx
RStudent	60	RStudentx

This report lists the columns in the dataset into which the various items were stored.

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS

Y = sale_price X = DistanceMM

Procedure Input Settings

Autosaved Template File

C:\Users\Dr_Ia\Documents\NCSS 2022\Procedure Templates\Autosave\Linear Regression and Correlation - Autosaved 2022_8_9-16_25_58.t153

Variables Tab

-- Variables -----	
Y: Dependent Variable(s):	sale_price
X: Independent Variable:	DistanceMM
Frequency Variable:	<Empty>
Weight Variable:	<Empty>

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

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

-- 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	Checked
Correlation and R-Squared	Unchecked
Summary Matrices	Unchecked

.. Estimation	
Regression Estimation	Checked

.. ANOVA	
ANOVA	Checked

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

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS

Y = sale_price X = DistanceMM

Procedure Input Settings (Continued)

Reports Tab (Continued)

-- Prediction	
Predict Y at these X values:	<Empty>
Predicted Y - C.L.	Checked
Predicted Y - P.L.	Checked
-- 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:	Double
-- 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

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

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
 Y = sale_price X = DistanceMM

Procedure Input Settings (Continued)**Plots Tab (Continued)**

-- 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
 -- Random Numbers for Bootstrap Confidence Intervals and the Randomization Test -----
 Random Seed: Random

Storage Tab

-- Select Items to Store with the Dataset -----

Predicted Y	Checked
Standard Error of the Predicted Y Mean	Unchecked
Lower C.L. of the Predicted Y Mean	Unchecked
Upper C.L. of the Predicted Y Mean	Unchecked
Standard Error of a Predicted Y Individual	Unchecked
Lower C.L. of a Predicted Y Individual	Unchecked
Upper C.L. of a Predicted Y Individual	Unchecked
Predicted X	Checked
Lower C.L. of the Predicted X Mean	Unchecked
Upper C.L. of the Predicted X Mean	Unchecked
Lower C.L. of a Predicted X Individual	Unchecked
Upper C.L. of a Predicted X Individual	Unchecked
Residuals	Checked
RStudent	Checked
Hat Diagonals	Unchecked
CovRatio	Unchecked
DFFITS	Unchecked
DFBETAS	Unchecked
MSE(i)	Unchecked
Cook's D	Unchecked
Standard Error of the Predicted Band	Unchecked
Lower Confidence Band of the Predicted Y Mean	Unchecked
Upper Confidence Band of the Predicted Y Mean	Unchecked
LOESS Values	Unchecked

Linear Regression Report

Dataset C:\Users\Dr_Ia\Desktop\BHTotalData.NCSS
Y = sale_price X = DistanceMM

Procedure Input Settings (Continued)

Storage Tab (Continued)

.. Storage Options	
Storage Location:	In empty column(s) after the last column with data
Skip one or more columns before item storage	Unchecked
Rename the storage column(s) using the item name(s)	Checked
Add a Note about each stored item in the Column Info table	Checked
Overlay storage data when the Group By System is active	Unchecked
