

StatCrunch Lab 4 for Statistics 301

Topics: scatter plots, linear regression, prediction, residual, diagnostics

Scatter plot, linear regression, prediction, residual. Consider grain yield (tons) versus distance from water (feet) data.

dist, x	0	10	20	30	45	50	70	80	100	120	140	160	170	190
yield, y	500	590	410	470	450	480	510	450	360	400	300	410	280	350

1. *Scatter plot, regression.* The least-squares line is $\hat{y} = 515.45 - 1.06x$.

StatCrunch: Relabel var1 as distance and var2 as grain yield. Type data into two columns. Stat, Regression, Simple Linear, X-Variable: distance, Y-Variable: grain yield, Next, Next, check Save Residuals, Next, check plot the fitted line, Calculate. Linear regression is given on first page of output, click Next to see scatter plot and regression on this plot.

2. *Prediction.*

At distance $x = 100$, $\hat{y} = -1.06x + 515.45 = -1.06(100) + 515.45 \approx 409.5$.

StatCrunch: Click Options (in Simple Linear Regression box!), Edit, Back, check Predict Y for X = 100, Calculate.

3. *Residual.*

At distance $x = 100$, $y - \hat{y} \approx 360 - 409.5 = -49.5$

StatCrunch: Look in Residuals column in data, beside distance $x = 100$ and $x = 140$.

Diagnostics: residual plot, boxplot of residuals. Consider grain yield (tons) versus distance from water (feet) data.

dist, x	0	10	20	30	45	50	70	80	100	120	140	160	170	190
yield, y	500	590	410	470	450	480	510	450	360	400	300	410	280	350

StatCrunch: Stat, Scatter Plot, X-Variable: distance, Y-Variable: grain yield, Next, Next, check Save Residuals, Next (three times!), check Plot the fitted line, check Residuals vs. X-values, Calculate gives Simple linear regression results, Next gives plotted regression, Next gives Residual plot. Graphics, Boxplot, Residuals, Next, check Use fences to identify outliers, Create Graph! gives boxplot.