Scatterplot

- Be able to use your calculator to graph the points. Make sure that your scatterplot has all of the components: title, axes scaled (does not have to start at zero), axes labeled, points plotted well (not individually labeled), graph centered around the points.
- Outliers are in the y-direction,
- Influential points (a special type of outlier) are in the x-direction.
- If data is far from zero, make sure the axes start appropriate values to display the data in a visually helpful way.
- Direction, Form, Strength, Outliers. Include r-value.

Correlation: r

- Strength (include r in your paragraph), form (linear or non-linear), direction (positive or negative), and outliers. Include context in your paragraph. DO NOT DISCUSS THE SLOPE OR THE LSRL HERE.
- See page 156 from textbook for more details about r.
- REMEMBER: CORRELATION DOES NOT EQUAL CAUSATION!

Residual Plot and Coefficient of Determination: r^2

- r^2 is used to determine the fit of the LSRL for the scatterplot. If r^2 is close to 1 (or 100%), then the LSRL could be a good fit, but the Residual Plot should not have a parabolic pattern. Even if r^2 is close to 1, a Residual Plot with a pattern indicates that the LSRL is not a good fit.
- Residual Plots should look scattered, like snow, to indicate the LSRL is a good fit.
- r^2 measures how much prediction error is eliminated when using the LSRL instead of the mean of y-data. Link to a helpful Khan Academy article.
- When asked to interpret r^2 , you would say "the percent variation of y that can be explained by x."
- Residual value = Actual value Predicted by the LSRL (think A-P)
- YOU ARE REQUIRED TO DRAW AND INTERPRET THE RESIDUAL PLOT.

Calculator

- Create a scatterplot
- Enter data into lists
- Create residual plot
- Find LSRL, r, and r² values
- Use 1-var stats

Least Squares Regression Line (LSRL)

- $\hat{y} = a + bx$
- When writing an LSRL always define the variables (i.e. state what y-hat is and x is)
- a: y-intercept. When x is zero units, the predicted y-value is _____ units.
- b: slope. As x increases by 1 unit, the predicted y-value increases/decreases by _____ units.
- x: explanatory variable
- \hat{y} : predicted response variable
- To find b by hand use: $b = r \frac{Sy}{Sr}$
- To find a by hand use: $a = \overline{y} b\overline{x}$
- Be able to graph the LSRL on your scatterplot, if necessary.

Computer Output (Minitab)

- Know how to write the LSRL, interpret the standard deviation of the residuals (s), find r^2 , find r (positive or negative).
- The standard deviation of the residuals is "the average distance the points are away from the LSRL in the y-direction," or "the average length of the residuals."
- Remember the constant is the y-intercept in the Coef column. The slope is the number below that.
- **s** is the standard deviation of the residuals. The standard deviation of the residuals, when using a least-squares regression line, is the typical vertical distance from a data point to a point on the regression line.

Predictor	Coef	SE Coef	Т	P
Constant	44.540	1.813	24.57	0.000
Revision time	0.55528	0.05501	10.09	0.000

S = 6.35447 R-Sq = 72.8% R-Sq(adj) = 72.1%