Vocabulary: Least-Squares Best Fit Lines



**Vocabulary**



* Least-squares best fit line – a line that estimates the linear relationship between two variables in a scatter plot.
	+ The least-squares best fit line is also called the *least-squares regression line*.
	+ The least-squares fit line is the line that minimizes the total sum of the squares of the *residuals*, as shown by the purple squares in the image to the right.
* Outlier – a value that is significantly different from the other values in a data set.
	+ In a scatter plot with a trend line, the outliers are usually the points farthest from the trend line.
* Residual – the difference between an observed value in a data set and its predicted or theoretical value.
	+ In a scatter plot with a trend line, the residual of a point is the vertical difference from the data point to the trend line (the difference between *y*-values).
* Scatter plot – a graph of (*x*, *y*) points that shows the general relationship between two variables, *x* and *y*.
	+ Generally, the variable on the horizontal (*x*) axis is the *independent variable*, and the variable on the vertical (*y*) axis is the *dependent variable*.
* Trend line – a line that fits the points in a scatter plot well.

**trend line**

* The slope of the trend line indicates the type of correlation the variables have.
* A positive slope indicates a positive correlation, a negative slope indicates a negative correlation, and a slope of zero indicates no correlation.
* For example, the trend line shown to the right suggests a negative correlation between the variables.