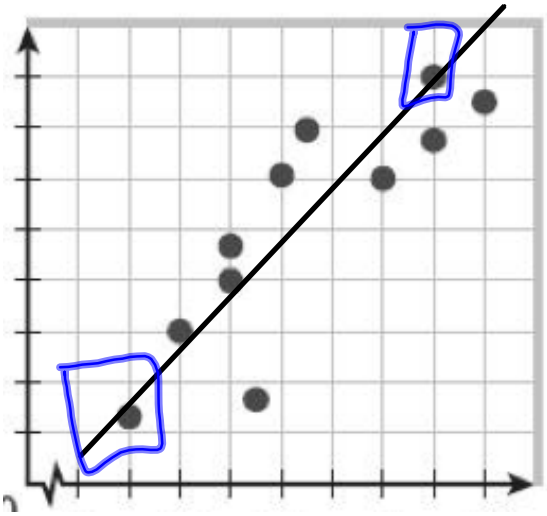
5.4/5.7 Best Fitting Lines

Usually there is NO single line that passes through all of the data points, so we try to find the line that best fits the data. This is called the best-fit Line. We can do this by drawing a line on a graph.

Example: On the scatter plot, draw a best-fit line.



- Choose 2 points on the line. $(\overset{x_1}{2},\overset{y_1}{1})$ $(\overset{x_2}{8},\overset{y_2}{8})$
- Find the slope between the 2 points. $\frac{y_2-y_1}{x_2-x_1}=\frac{8-1}{8-2}=\frac{7}{6}$
- Use slope-intercept or point-slope form to find the equation of the line.

$y=mx+b$

$y-y_1=m(x-x_1)$

m=slope
b=y-intercept

$y-1=\frac{7}{6}(x-2)$

$y-1=\frac{7}{6}x-\frac{14}{6}+\frac{6}{6}$

$+1 \qquad +1$

$y=\frac{7}{6}x-\frac{8}{6}$

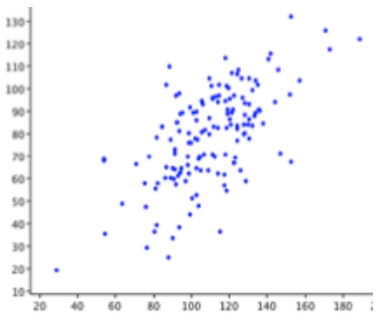
$$\boxed{y=\frac{7}{6}x-\frac{4}{3}}$$

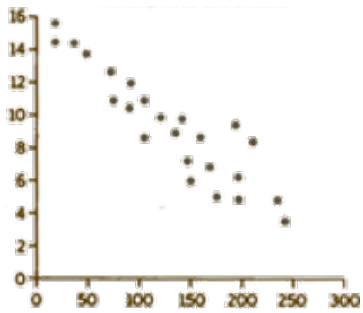
The number *r*, which indicates how well a set of data can be estimated by a straight line, is correlation.

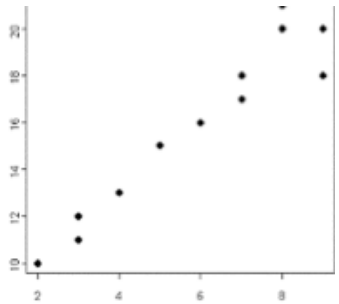
Correlation *r* is always between -1 & 1.

When the points on a scatter plot can be approximated by a line with positive slope, x and y have a positive correlation. When the points can be approximated by a line with negative slope, x and y have a negative correlation. When points CANNOT be approximated by a straight line, there is no Linear correlation.

Example: Label each graph with its correlation. (positive, negative, no corr.)







none

non linear

negative

Linear

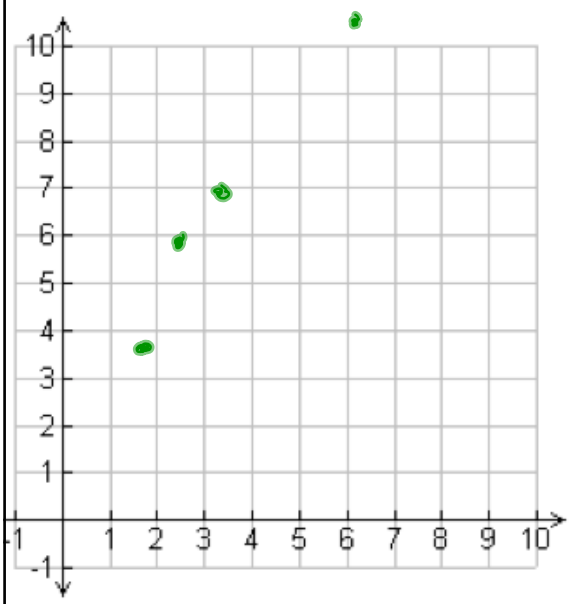
positive

Linear

We can decide when a Linear model can be used to represent real-life data.

We do this by creating a scatter plot of the data.

X	Y
1	4
3	6
4	7
7	10
9	12
10	13



On the graph, are the points almost in a straight line?

yes

If the points look like they're almost in a straight line on the graph, then the data can be modeled with a Linear equation.