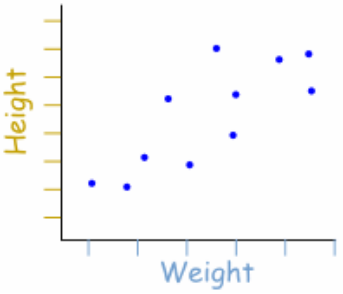
Scatter Plots



A graph of plotted points that show the relationship between two sets of data.

In this example, each dot represents one person's weight versus their height.

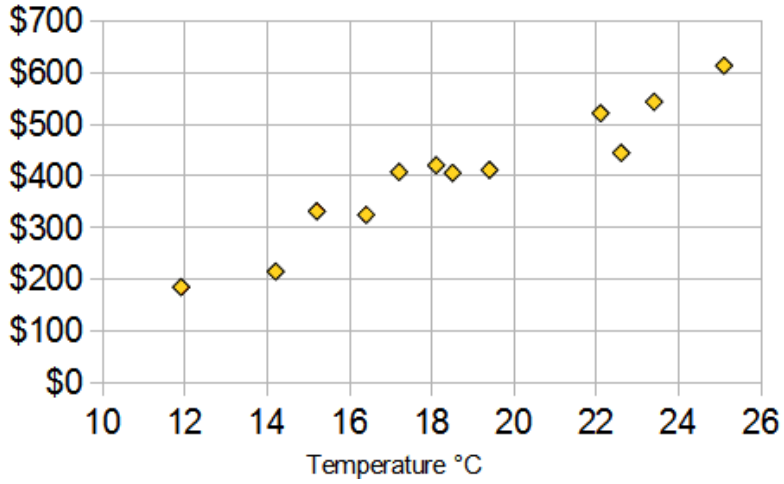
(The data is plotted on the graph as "Cartesian (x,y) Coordinates")

Example:

The local ice cream shop keeps track of how much ice cream they sell versus the noon temperature on that day. Here are their figures for the last 12 days:

Ice Cream Sales vs Temperature	
Temperature °C	**Ice Cream Sales**
14.2°	$215
16.4°	$325
11.9°	$185
15.2°	$332
18.5°	$406
22.1°	$522
19.4°	$412
25.1°	$614
23.4°	$544
18.1°	$421
22.6°	$445
17.2°	$408

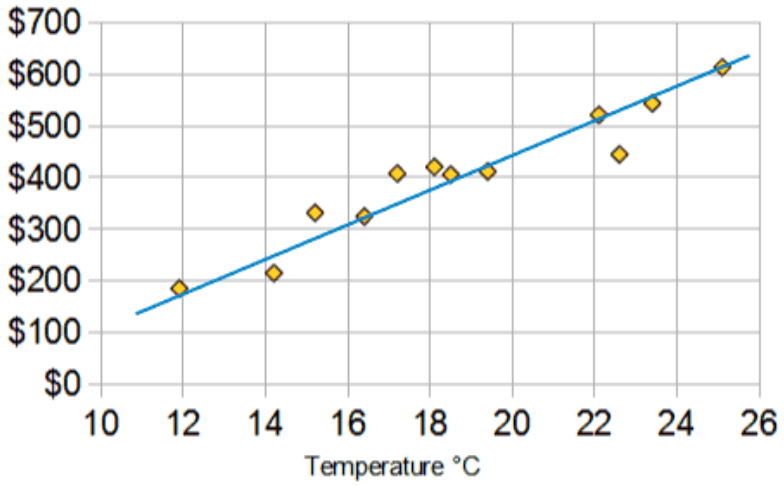
And here is the same data as a Scatter Plot:



It is now easy to see that **warmer weather leads to more sales**, but the relationship is not perfect.

Line of Best Fit

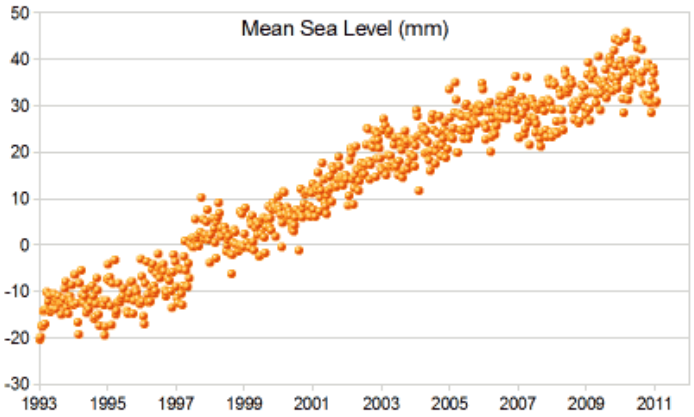
You can also draw a "Line of Best Fit" (also called a "Trend Line") on your scatter plot:



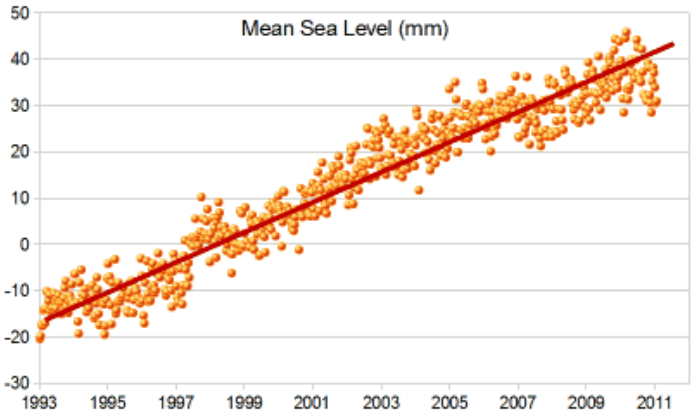
Try to have the line as close as possible to all points, and **as many points above the line as below**.

Example: Sea Level Rise

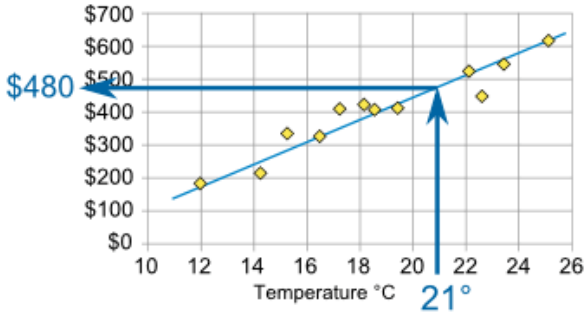
A Scatter Plot of Sea Level Rise:



And here I have drawn on a "Line of Best Fit".

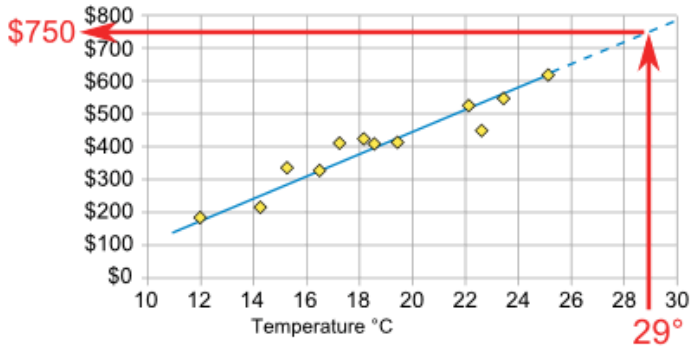


Interpolation and Extrapolation



Interpolation is where we find a value **inside** our set of data points.

Here we use **linear interpolation** to estimate the sales at 21 °C.



Extrapolation is where we find a value **outside** our set of data points.

Here we use **linear extrapolation** to estimate the sales at 29 °C (which is higher than any value we have).

Careful: **Extrapolation** can give misleading results because we are in "uncharted territory".

As well as using a graph (like above) we could come up with a formula to help us

Example:

We can estimate a straight line equation from two points from the graph above

Let's estimate two points on the line near actual values: **(12°, $180)** and **(25°, $610)**

$$\text{slope "m"} = \frac{\text{change in y}}{\text{change in x}} = \frac{\$610 - \$180}{25° - 12°} = \frac{\$430}{13°} = 33 \text{ (rounded)}$$

Put that slope and the point **(12°, $180)** into the "point-slope" formula:

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

$$y - 180 = 33(x - 12)$$

$$y = 33(x - 12) + 180$$

$$y = 33x - 396 + 180$$

$$y = 33x - 216$$

Now we can use that equation to **interpolate** a sales value at 21°:

$$y = 33 \times 21 - 216 = \$477$$

And to **extrapolate** a sales value at 29°:

$$y = 33 \times 29 - 216 = \$741$$

The values are close to what we got on the graph. But that doesn't mean they are more (or less) accurate. They are all just estimates.

Don't use extrapolation too far! What sales would you expect at 0° ?

$$y = 33 \times 0 - 216 = -\$216$$

Hmmm... **Minus $216**? We have extrapolated too far!

Note: we have been using **linear** (based on a **line**) interpolation and extrapolation, but there are many other types, for example we could use polynomials to make curvy lines, etc.

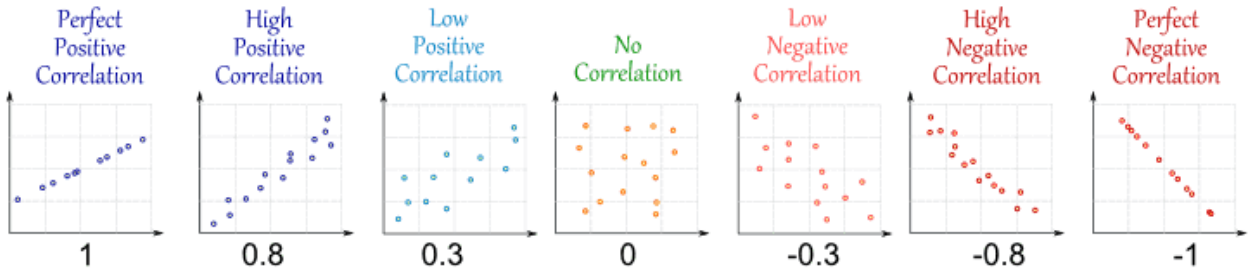
Correlation

When the two sets of data are strongly linked together we say they have a **High Correlation**.

The word Correlation is made of **Co-** (meaning "together"), and **Relation**

- Correlation is **Positive** when the values **increase** together, and
- Correlation is **Negative** when one value **decreases** as the other increases

Like this:



(Learn More About Correlation)

Negative Correlation

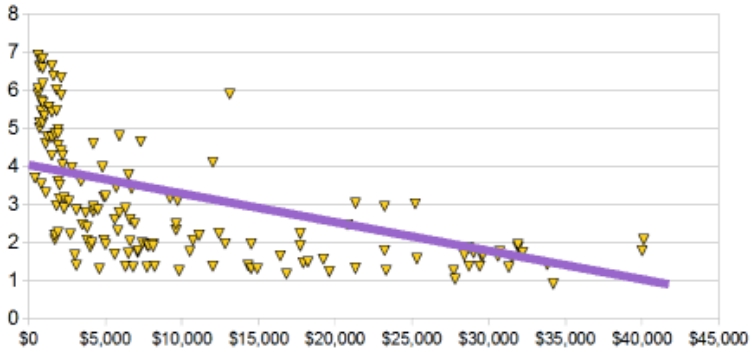
Correlations can be negative, which means there **is** a correlation but one value goes down as the other value increases.

Example : Birth Rate vs Income

The birth rate tends to be **lower** in richer countries.

Below is a scatter plot for about 100 different countries.

Country	Yearly Production per Person	Birth Rate
Madagascar	$800	5.70
India	$3,100	2.85
Mexico	$9,600	2.49
Taiwan	$25,300	1.57
Norway	$40,000	1.78



It has a **negative correlation** (the line slopes down)

Note: I tried to fit a straight line to the data, but maybe a **curve** would work better, what do you think?

Question 1 Question 2 Question 3 Question 4 Question 5
Question 6 Question 7 Question 8 Question 9 Question 10
Question 11 Question 12

Search :: Index :: About :: Contact :: Contribute :: Cite This Page :: Privacy