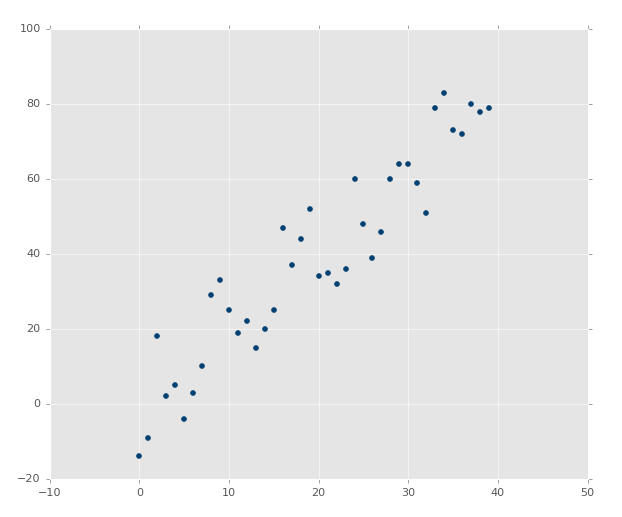
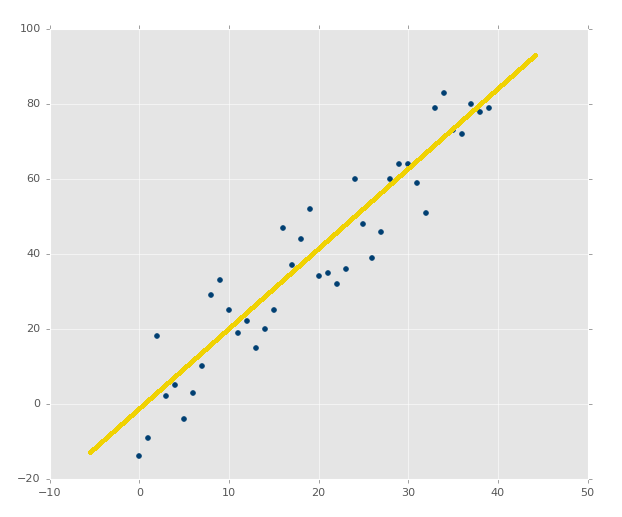
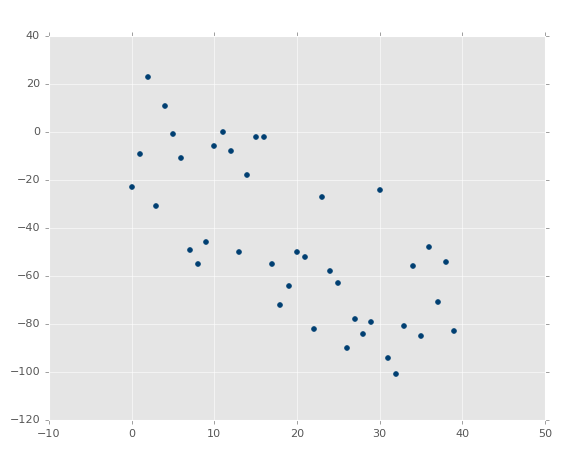
Correlation:



The above image clearly has a nice correlation. If you were asked to draw a best-fit line by estimation, you should be able to easily do something like this:

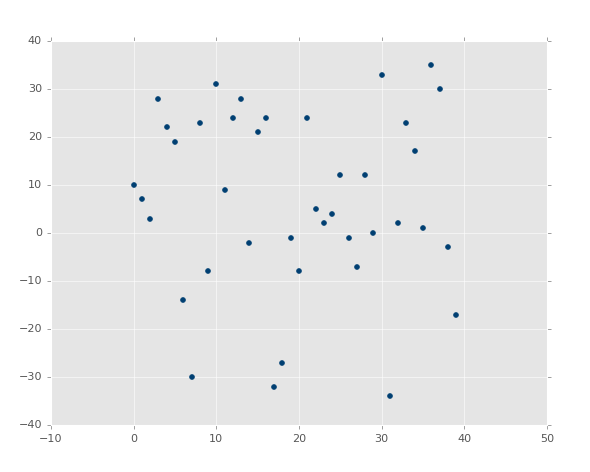


What about an image like:



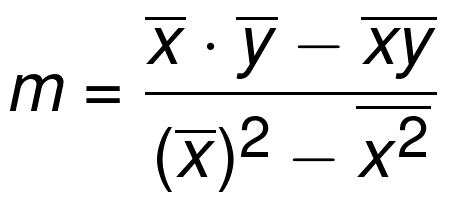
It is not AS obvious, but it is clearly a negative correlation. You could probably draw a best-fit line, but you're much more likely to not actually have the single best fit line in this case.

Finally, what about:

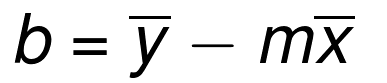


Well?! There is a best-fit line, but good luck figuring it out!

The slope, m, of the best-fit line is defined as:



The bar over the letters signifies a mean/average. If two letters are sitting next to eachother, you multiply them. The xs and ys are from all of the existing coordinates. So we now know the m (y intercept) of the best-fit line's definition in y=mx+b, now we just need the b (the slope). There's a formula for that too:



Okay great! ...except the whole part where this isn't a math tutorial, it's a programming tutorial. That means we actually need to build a program to do this for us. In the next tutorial, we're going to do just that, as well as explain why I am going to have us programming it rather than using a module!