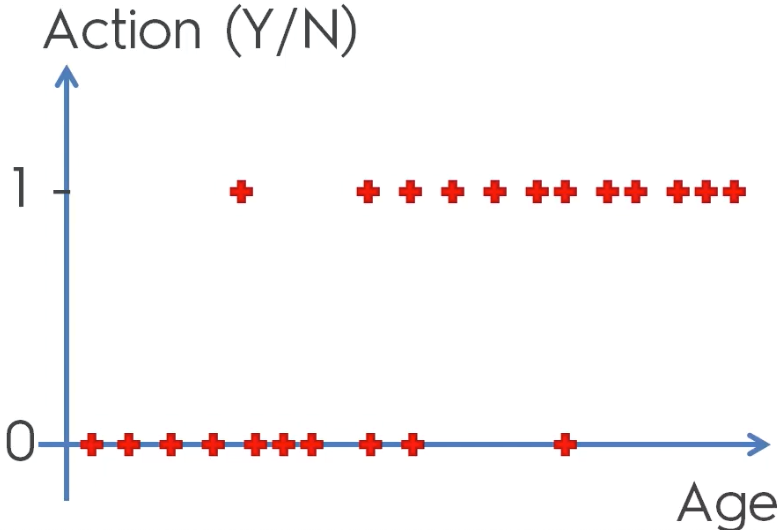
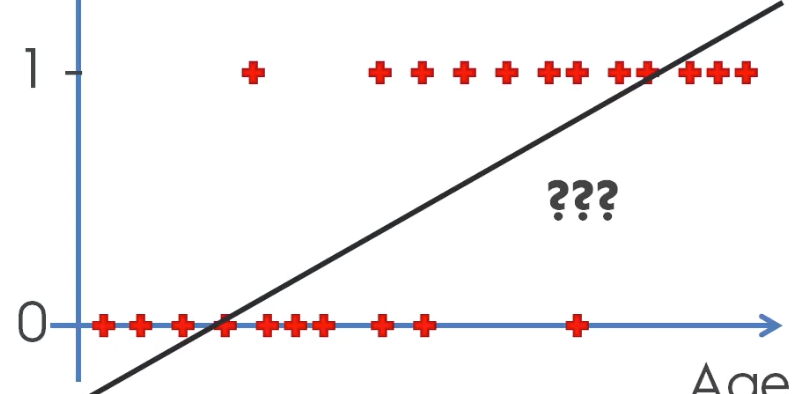
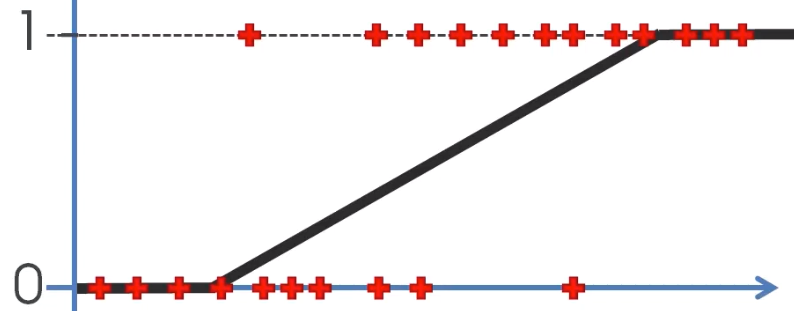
***Logistic Regression***

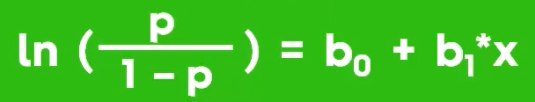


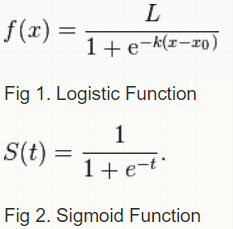
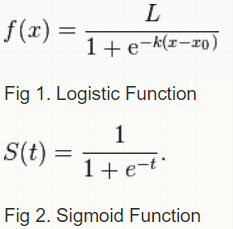
* Can initially see *some* sort of correlation (old = more likely to take an action based on *this* offer)
* Linear regression won’t work



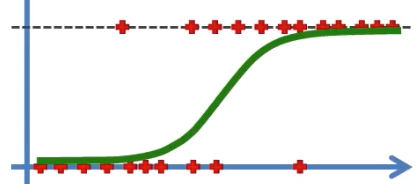
* Instead of predicting *exactly* if someone will take action, we predict the probability/likelihood of them taking an action
* Action goes from 0-1, as do probabilities
* Middle of line makes sense 🡪 probability goes up as age goes up
* But ends do not, as probability cannot be > 1 or < 0
* Can interpret as people above that crossing of the horizontal 1 line and the regression line = 100% of taking action, and crossing the horizontal 0 line = 0% chance



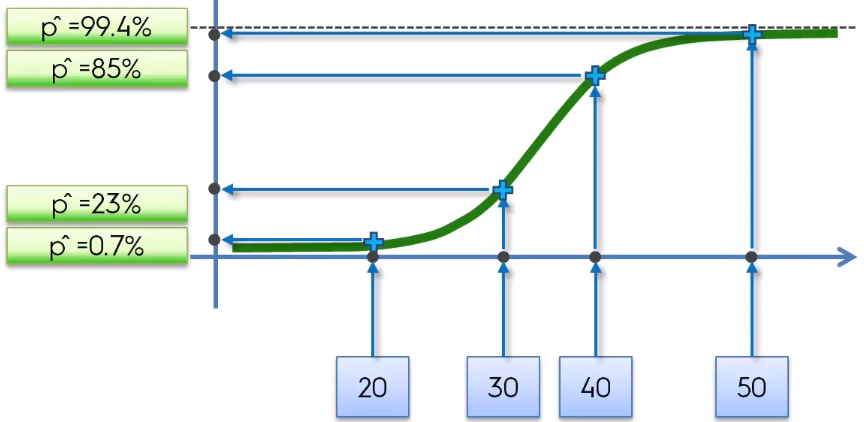
* This is the **sigmoid function**
*  **🡺**  🡺 solve for y + plug in 🡺
* Green = This is the formula for logistic regression



* **sigmoid function** = a special case of the Logistic function when L = 1, k = 1, x0 = 0
* L = maximum value the function can take.
* e−k(x−x0) is always >= 0, so the maximum point is achieved when it = 0, + this is at L/1.
* x0 controls where on the x-axis the growth should be, because if we put x0 in for x, x0−x0 cancel out + e0 = 1, so you end up with f(x0)=L/2 = the midpoint of the growth.
* The parameter k controls how steep the change from the min to the max value is



* *Same thing as linear regression* = find the best line to fit the observations, it’s just not straight
* We use that line to predict probabilities, **p^**
* Take 4 random inputs + predict p^ by projecting these values onto our best-fit logistic regression curve = **fitted values** 🡪 then project to left for probabilities



* We use this probability as a **score**
* If we want an actual prediction for y^, we need to set a threshold for the p^ to be greater than for us to predict 1 and if not, predict 0
* Ex: threshold = .5 (50%)

