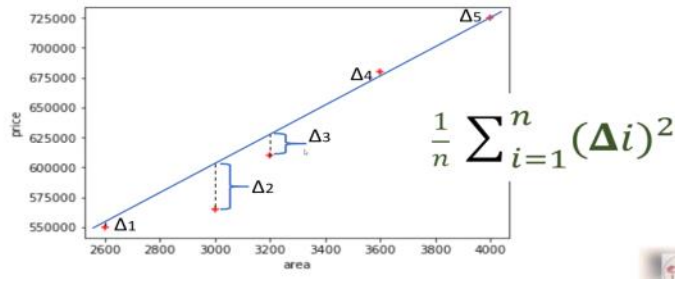
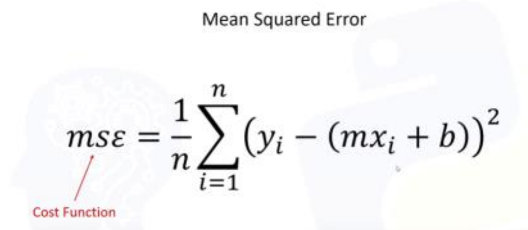
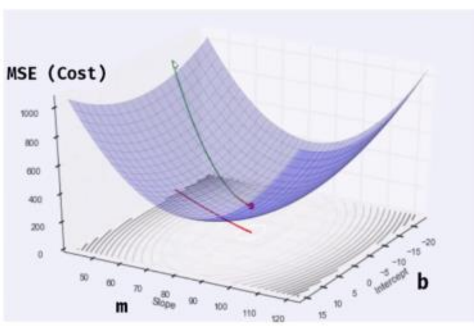
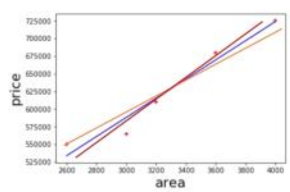
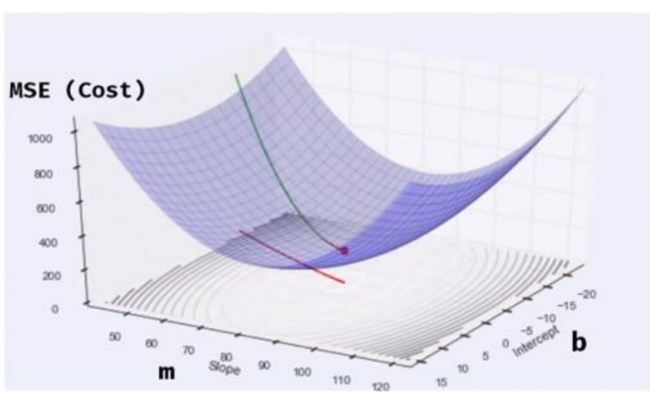


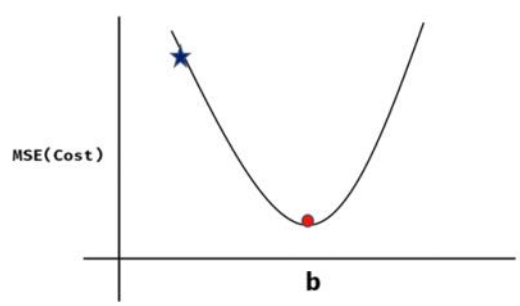


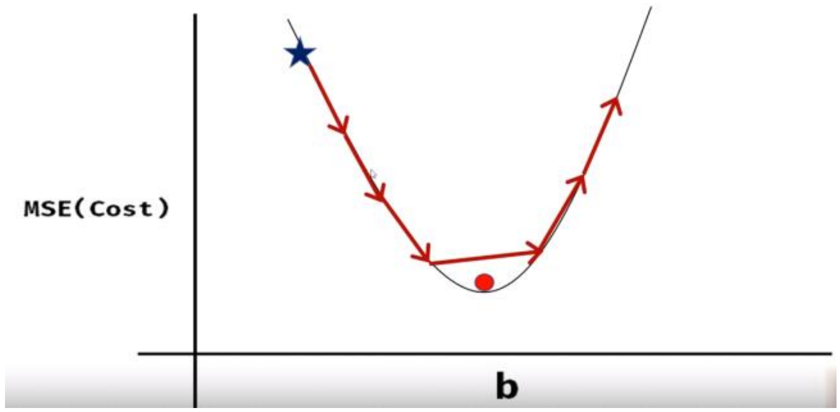
Gradient descent is an algorithm that finds best fit line for given training data set



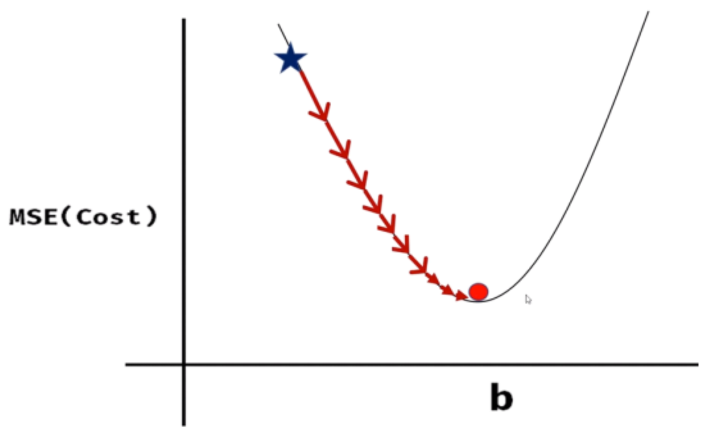


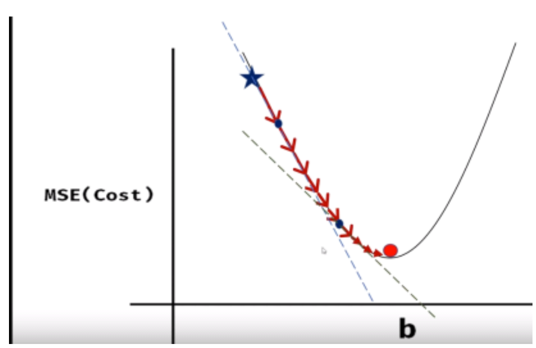






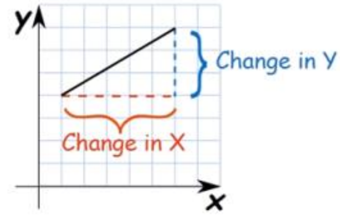
Fix values of y then we can miss the Minimum value of curvature.



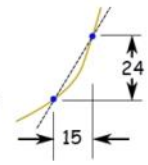


It is all about slope!

$$\text{Slope} = \frac{\text{Change in Y}}{\text{Change in X}}$$



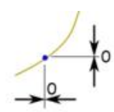
We can find an **average** slope between two points.



average slope = $\frac{24}{15}$

But how do we find the slope **at a point**?

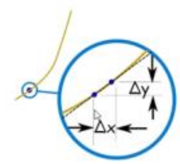
There is nothing to measure!



slope = $\frac{0}{0}$ = ???

But with derivatives we use a small difference ...

... then have it **shrink towards zero**.



We write **dx** instead of **"Δx heads towards 0"**.

And "the derivative of" is commonly written $\frac{d}{dx}$:

$$\frac{d}{dx}x^2 = 2x$$

"The derivative of $\boldsymbol{x^2}$ *equals* ***2x****"*

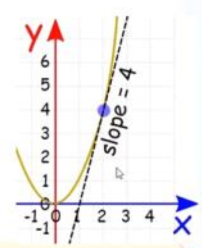
or simply *"d dx of* $\boldsymbol{x^2}$ *equals* ***2x****"*

What does $\frac{d}{dx}x^2 = 2x$ mean?

It means that, for the function x^2, the slope or "rate of change" at any point is **2x**.

So when **x=2** the slope is **2x = 4**, as shown here:

Or when **x=5** the slope is **2x = 10**, and so on.



Note: sometimes $f'(x)$ is also used for "the derivative of":

$$f'(x) = 2x$$

Partial derivates

$$f(x,y) = x^3 + y^2$$

$$\partial f/_{\partial x} = 3x^2 + 0 = 3x^2$$

$$\partial f/_{\partial y} = 0 + 2y = 2y$$

$$ms\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - (mx_i + b)\right)^2$$

To solve our problem
We need to find the partial derivative of m and b
Partial derivatives gives the slope

$$\partial/_{\partial m} = \frac{2}{\mathrm{n}}\sum_{i=1}^{n} -x_i\left(y_i - (mx_i + b)\right)$$

$$\partial/_{\partial b} = \frac{2}{\mathrm{n}}\sum_{i=1}^{n} -\left(y_i - (mx_i + b)\right)$$

$$\mathrm{m} = \mathrm{m}\ - \text{learning rate} * \partial/_{\partial m}$$

$$\mathrm{b} = \mathrm{b}\ - \text{learning rate} * \partial/_{\partial b}$$