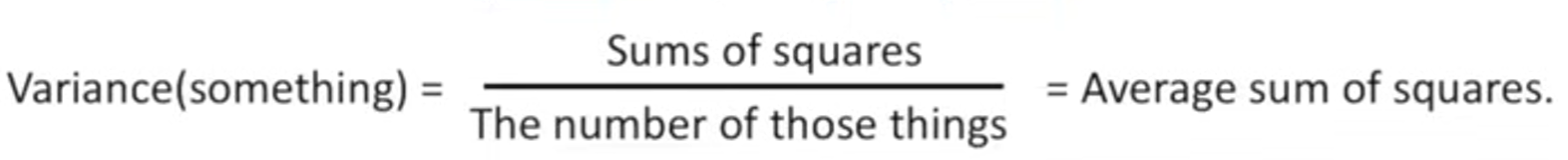
AI ML JOURNEY  
DAY 0

StatQuest – Linear Regression

1. When we draw line in a graph which has some data points, the distance between the line and data point is called “residual”.
2. So what we need to do is we need to keep rotating the line and measuring the sum of the squares of the residual.
3. Then get the least sum of the square. We need to use this line to fit to the data.
4. This line will most probably superimpose the data points and will allow us to predict the y axis with the value of x.
   1. y = intercept + slope
   2. y = constant + constant \* x
5. So now how good will be the guess of the Y is R^2.
6. Now based on the mouse size, we will take mean around the data point and then we will take the sum of squares of the residual. This is called as SS(mean), sum of squares.
7. A screenshot of a computer

   AI-generated content may be incorrect.
8. Now, take the sum of squares of the residual around the line which had least squares fit, this is called as SS(fit), sum of squares around above line.
   1. 
   2. A math equation with black text

      AI-generated content may be incorrect.
   3. 
9. A black text on a white background

   AI-generated content may be incorrect.
10. Using R2, we can check how much variant the y (predictive variable) due to x (dependent variable). R2 should always be under 0 to 1.
11. A diagram of a mathematical equation

    AI-generated content may be incorrect.
12. R2 value becomes insignificant when we only have two data points because a straight line can always connect two points which means SS(fit) is 0 as no residual is present. Now to determine if R2 is significant, we need p-value.
13. A diagram of a graph

    AI-generated content may be incorrect.
14. A white text with black text

    AI-generated content may be incorrect.