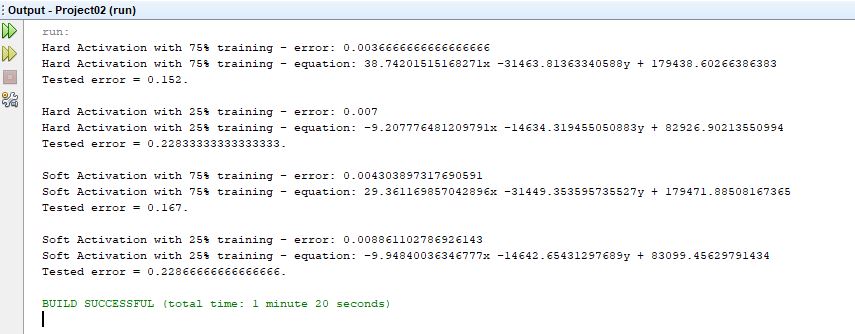
Result display:



Separation line a txt is 75% hard training equation

Separation line b txt is 25% hard training equation

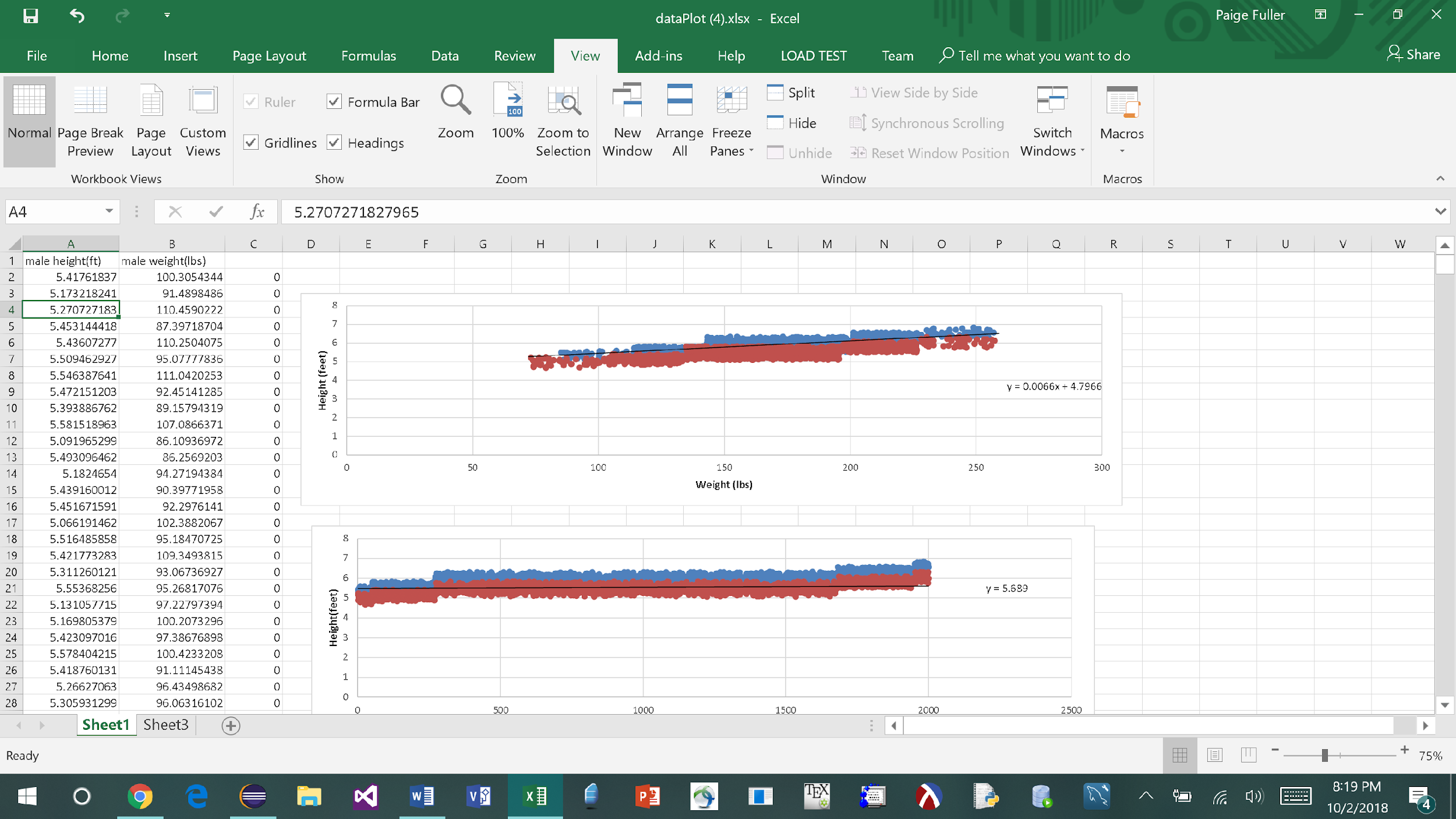
Separation line c txt is 75% soft training equation

Separation line d txt is 25% soft training equation

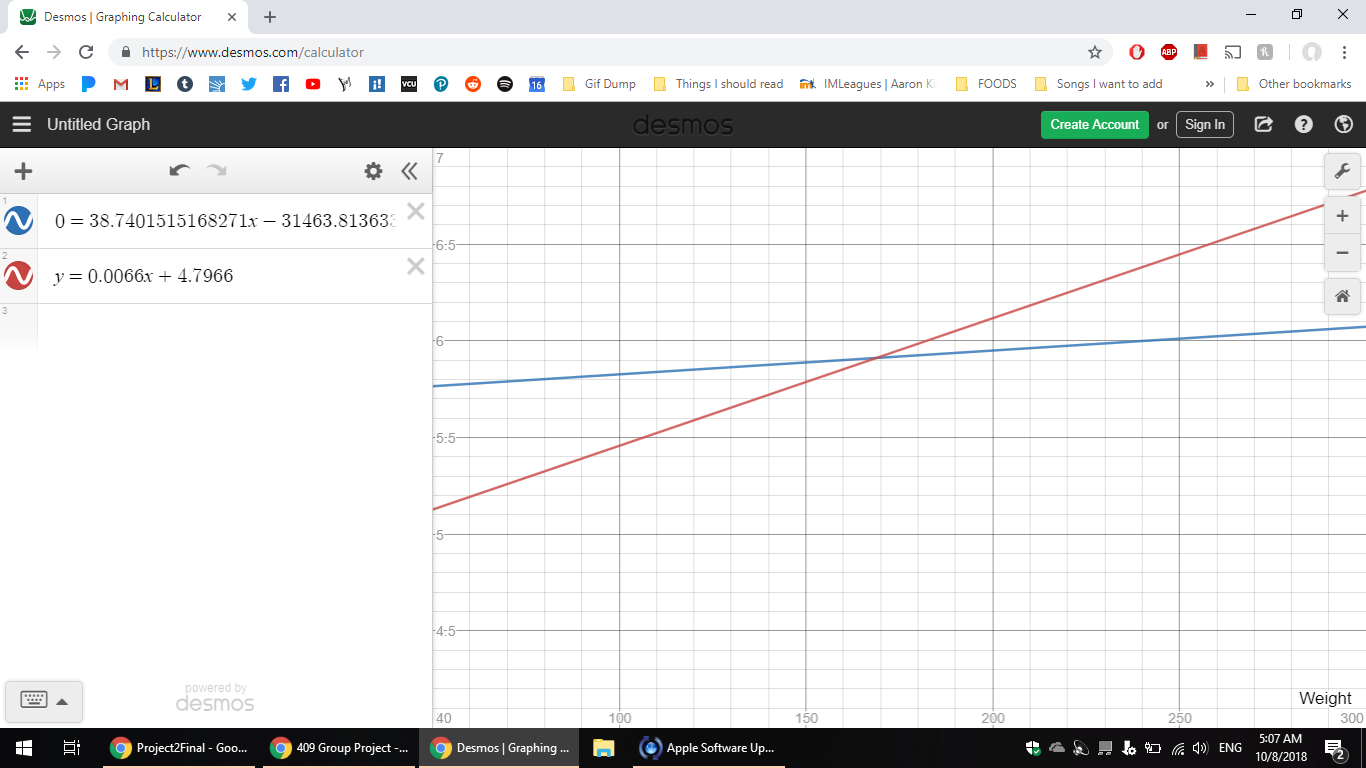
HARD ACTIVATION

**1) Choose 75% of the data for training, and the rest for testing for scenarios. Train and test your neuron. Plot the data and resulting separation line. Specify corresponding errors.**

Below is our old separation line against the data.



Below here is the old separation line from Project 1 (RED) and the new line for hard activation using 75% of the data for testing (BlUE). Line equation is shown on the top result picture.

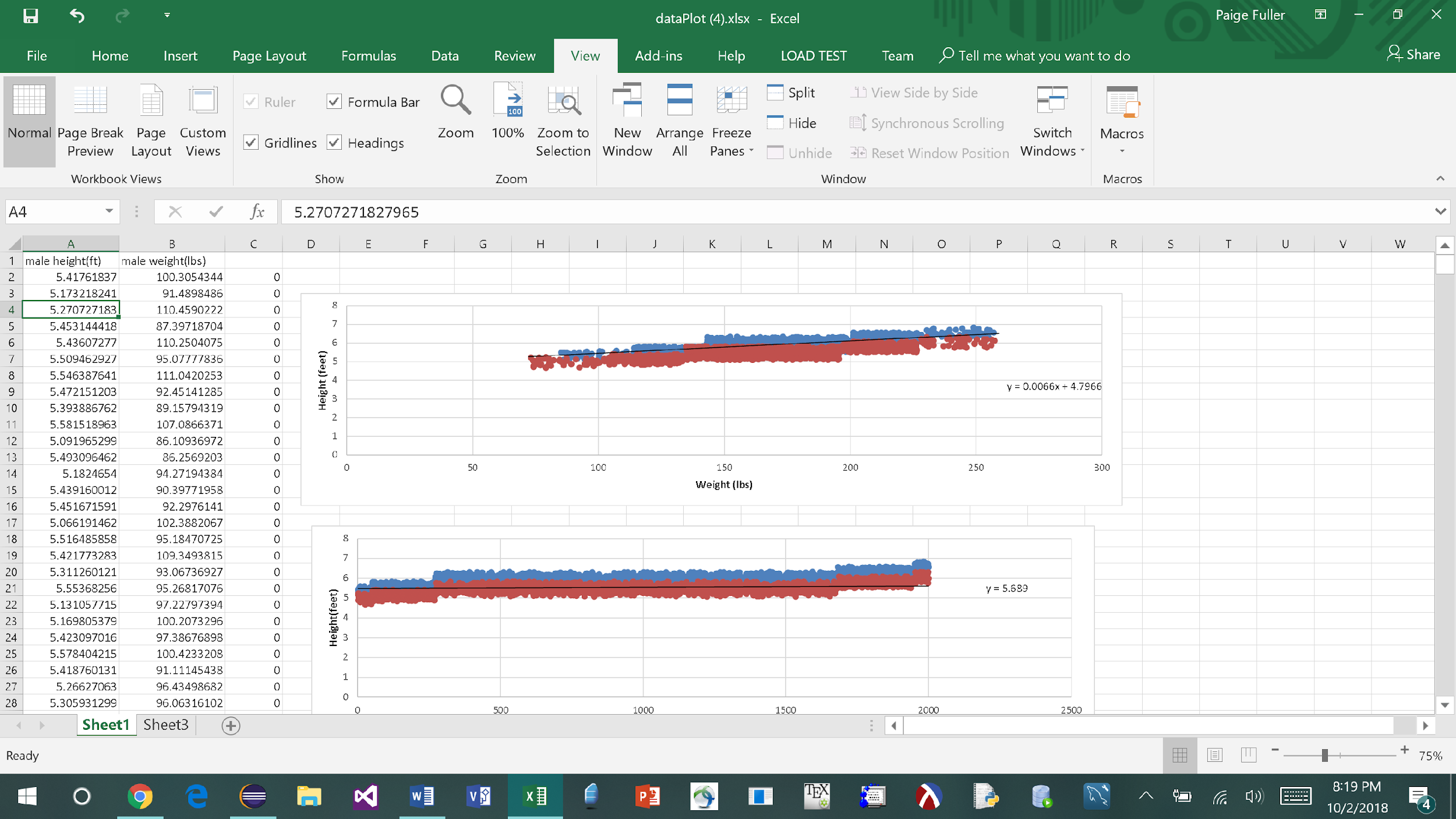


The error for the hard activation with 75% of the data used for training is:  
0.003666666666666666

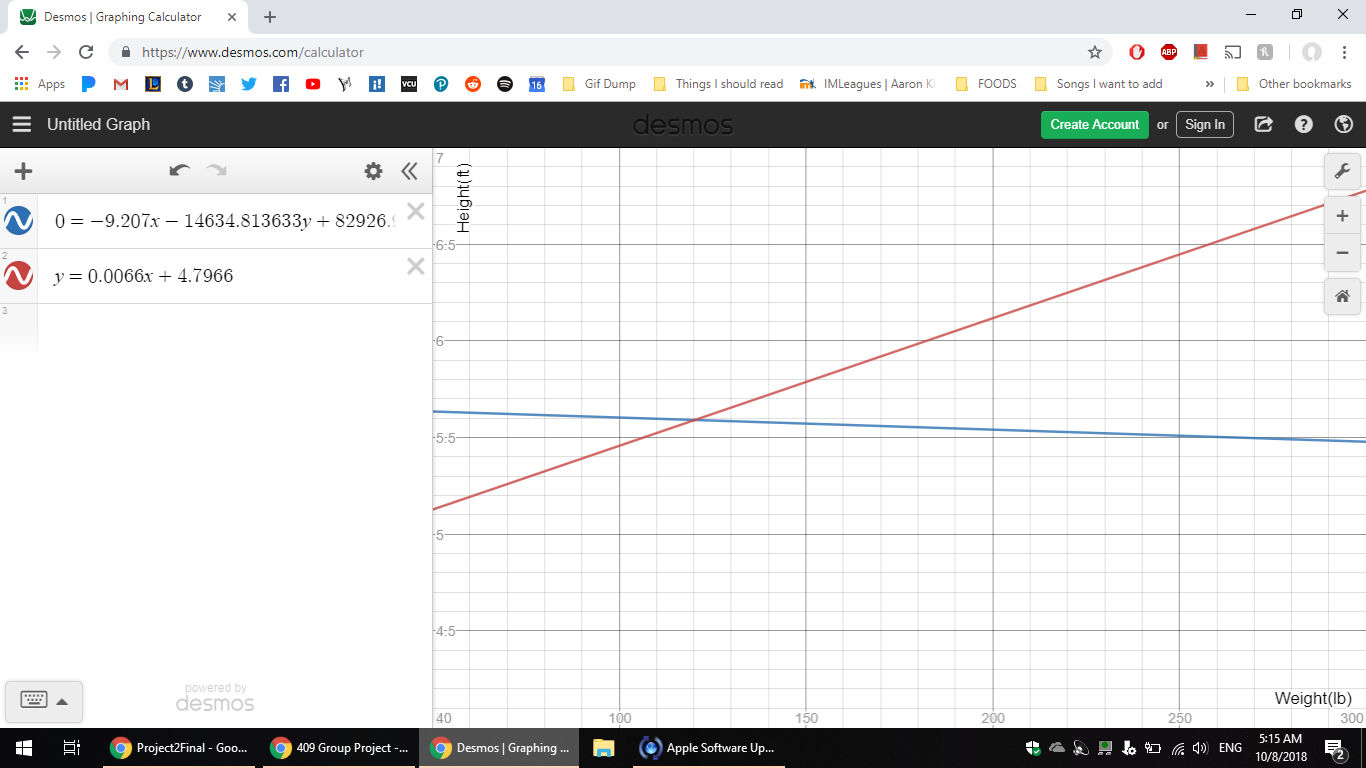
The error for the hard activation with 25% of the data used for testing is: 0.152

**2) Choose 25% of the data for training, and the rest for testing for scenarios. Train and test your neuron. Plot the data and resulting separation line. Specify corresponding errors.**

Below is our old separation line against the data.



Below here is the old separation line from Project 1 (RED) and the new line for hard activation using 25% of the data for testing (BLUE). Line equation is shown on the top result picture.



The error for the hard activation with 25% of the data used for training is: 0.007

The error for the hard activation with 75% of the data used for testing is: 0.228333333333333333

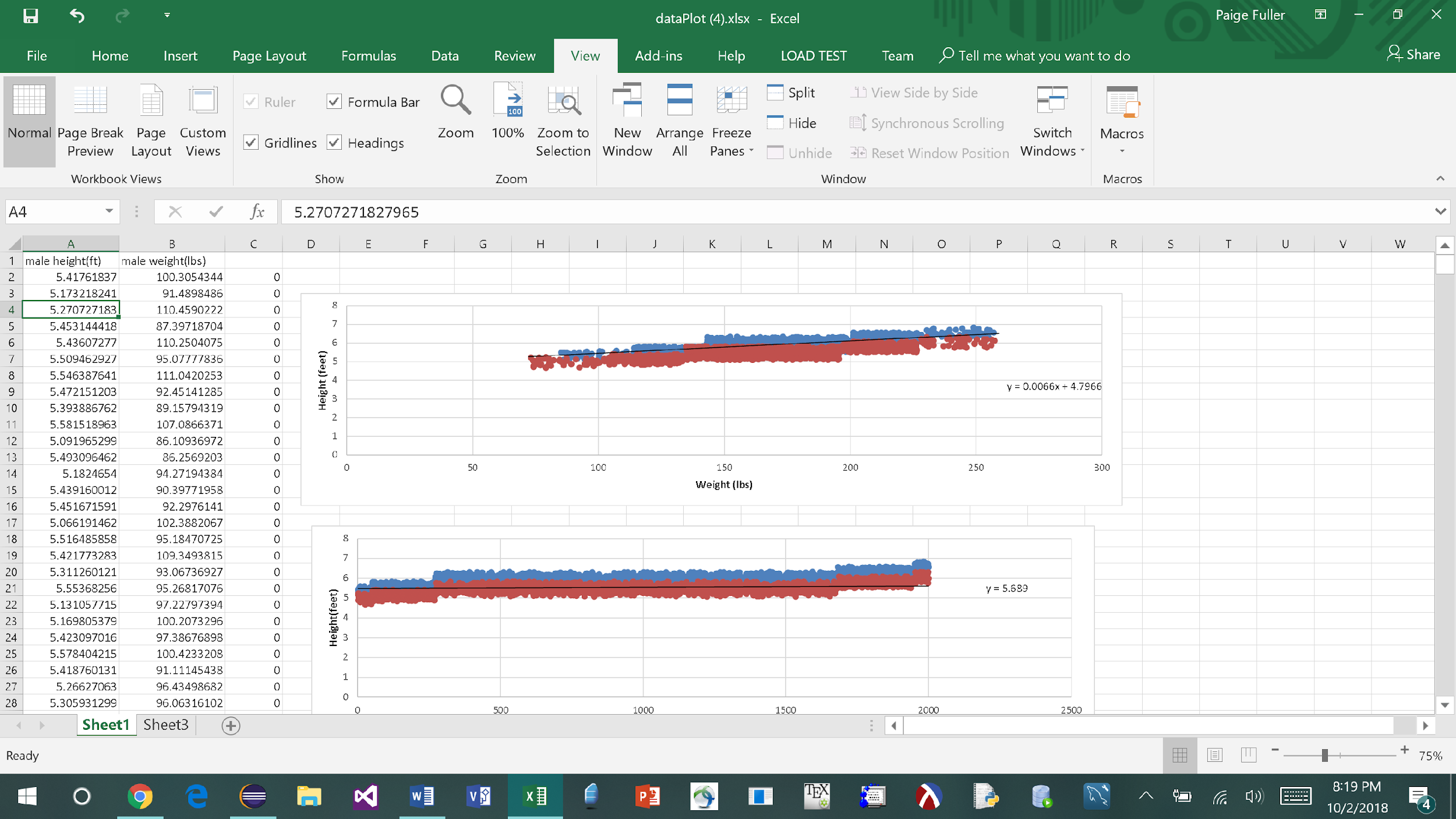
**3) Compare 1. and 2. and discuss.**

Both training’s coefficient is 0.3. The 75% training did about 7% better than the one with 25% training. This would make be expected as our data seems to follow a linear positive trend which fits better with the 75% training’s positive slope, whereas the 25% training has a negative slope. It would also be reasonable for the 75% training to have less error for the rest of data because there are less to test than the 25% training and more data points to train from. Despite having less error, the difference is hardly large.

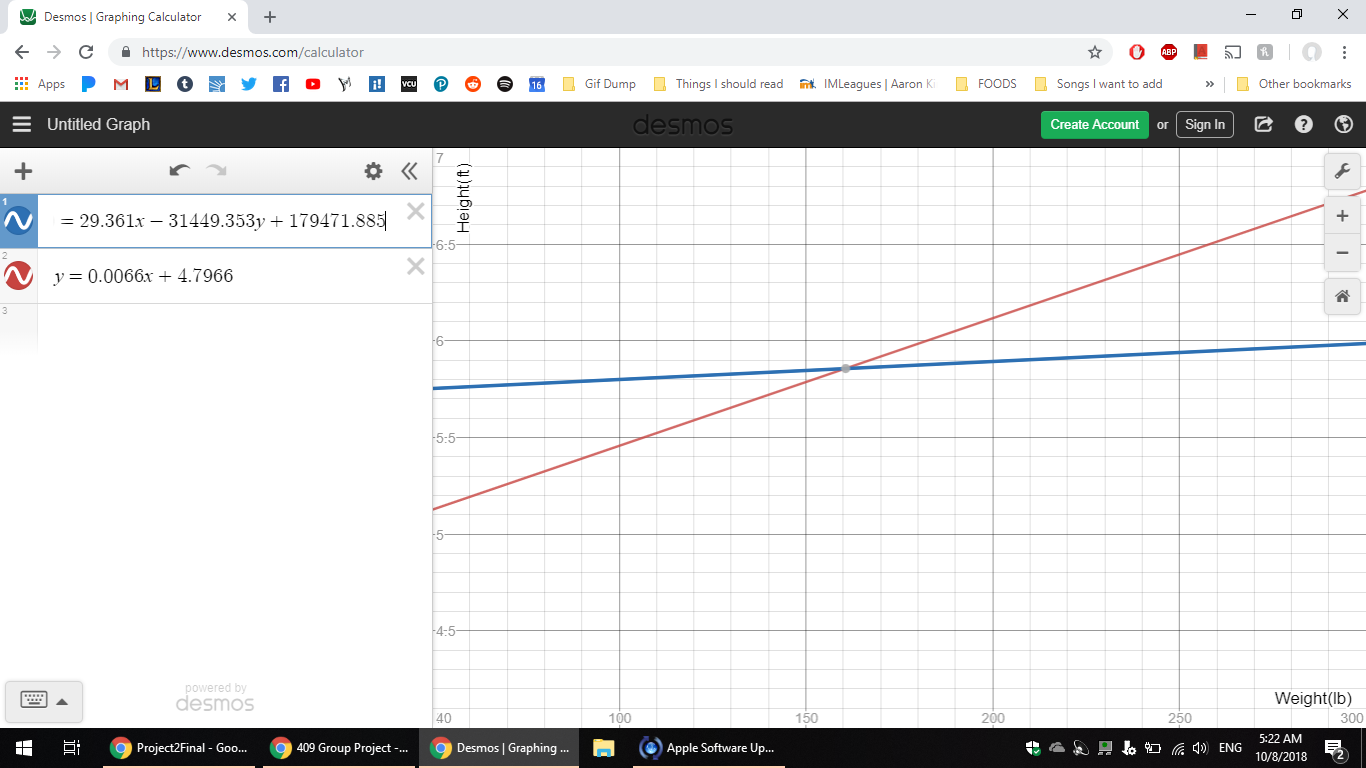
SOFT ACTIVATION

**1) Choose 75% of the data for training, and the rest for testing for scenarios. Train and test your neuron. Plot the data and resulting separation line. Specify corresponding errors.**

Below is our old separation line against the data.



Below here is the old separation line from Project 1 (RED) and the new line for soft activation using 75% of the data for testing (BLUE). Line equation is shown on the top result picture.

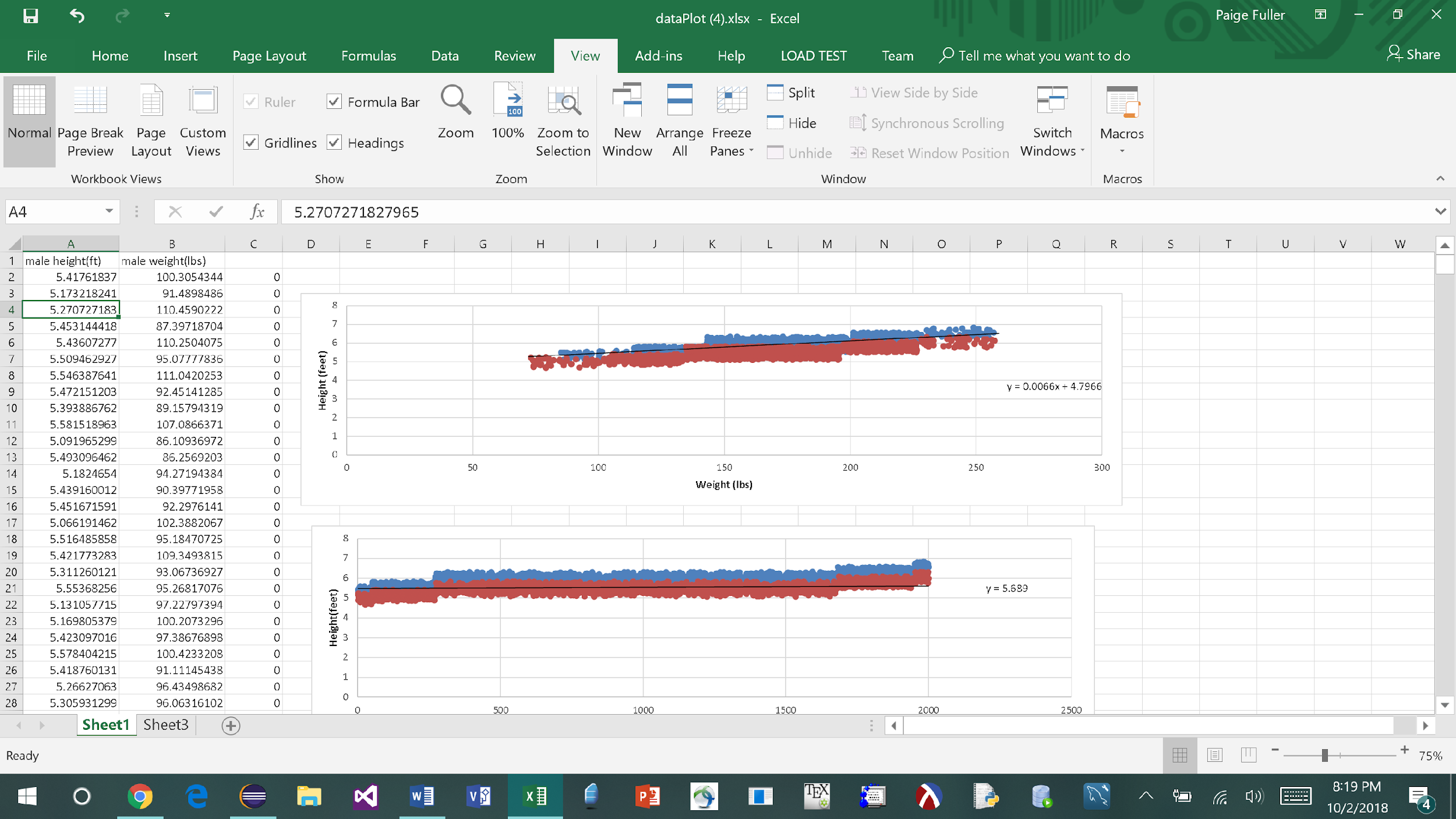


The error for the soft activation with 75% of the data used for training is: 0.0043038973117690591

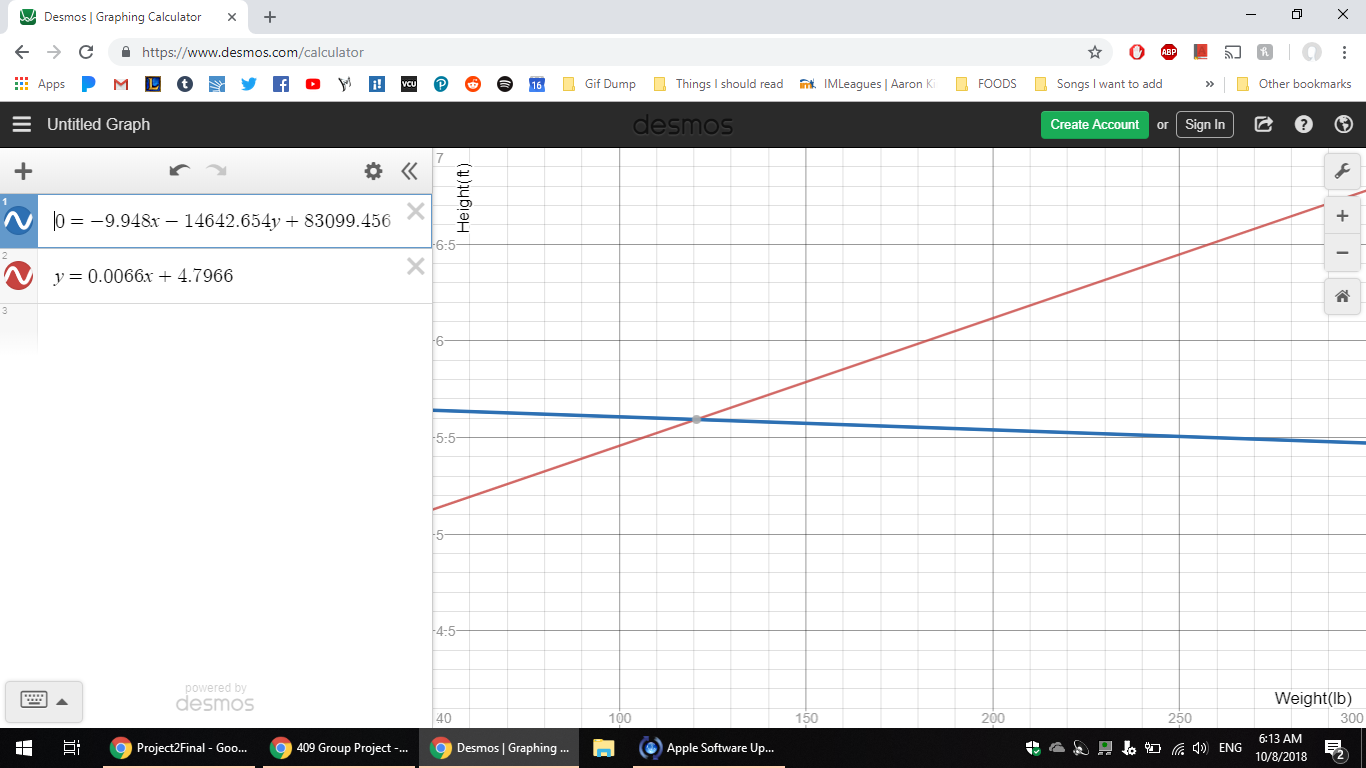
The error for the soft activation with 25% of the data used for testing is: 0.167

**2) Choose 25% of the data for training, and the rest for testing for scenarios. Train and test your neuron. Plot the data and resulting separation line. Specify corresponding errors.**

Below is our old separation line against the data.



Below here is the old separation line from Project 1 (RED) and the new line for soft activation using 25% of the data for testing (BLUE). Line equation is shown on the top result picture.



The error for the soft activation with 25% of the data used for training is: 0.008861102786926143

The error for the soft activation with 75% of the data used for testing is: 0.228666666666666666

**3) Compare 1. and 2. and discuss.**

For both training, K value is 1.0. Like the hard activation, the 75% training did better in terms of error, but in this case with less of a difference of about 6%. This would make be expected as our data seems to seems to follow a positive linear trend which is similar to that of the slope of the 75% training, whereas like in the hard activation, the 25% training has a negative slope. Also in terms of the training error, the 75% training did better by a factor of 2. The reason being 75% training having less error for the rest of data because there are less to test than the 25% training and more data points to train from.

**Pr. 2.1 Extra credit (soft vs. hard activation function) Compare and discuss results form using hard activation and soft activation solutions for 1. Do the same for 2. and 3.**

Comparison for training with 75% of the data and having the same amount of training cycles. The training error is about the same for hard vs soft activation, but the testing error is a little different by a very slight margin 1.5% with the Hard Activation being the better one.

Comparison for training with 25% of the data also having the same number of training cycles. Both the training error and especially the testing error are incredibly similar with no significant difference amongst the two activation.

At least in our cases, there seemed to be little difference between hard and soft activation. Even in the comparison in each between the 75% training and the 25% training remain the same as well. Both activations have the 75% training be the less error prone one, but with a same difference of about 6-7%. Also in both activations, the training error of 75% is double that of the error of the 25% training.

Workload distribution:

We work as a team, we peer coded and worked together on this report via google doc.