Twitter Visualization Activity (3155):

Task 1:

1. Explain the above code and visualization in your own words?

The code that was provided to us is setting up how the three csv files we downloaded will be displayed. We then added some extra lines to interpret the files into the chart display. The visualization of the files was displaying the frequency that various first and last names were being used in the fake twitter accounts that the csv files contained. By looking at the bar diagram we can identify what the fake accounts frequently used.

1. Can you improve this above bar graph? If yes, how will you try to improve it and explain the process in steps.

I would improve this bar graph by making it easier to interpret. The vertical axis is obvious that these are the names that were being generated. However, the horizontal axis could use a label that tells the viewer the bars are indicating the frequency of use in the files that were provided. A more advanced functionality that I think could be added – would be to add a hover feature allowing the user to view each tweet that the account had posted. The steps for doing this would be to add code that breaks each iteration of the name’s appearance into its own segment. To specify, from 0 to 1 occurrence the bar would stop then from 1 to 2 the bar would continue with a small couple pixel break in between the bars. When the user clicks on the segment of the bar that the name had occurred in, a small pop up box would appear with the tweets body.

Task 2:

1. What do we mean by sentiment analysis?

I think that sentiment analysis means to pick apart the tweet by certain words that are used and using that to deduct what the topic of the tweet could potentially be about.

1. After implementation, write your observation about the visualization of user profile in detail.

This is a much better visual representation for the csv’s. The bar graph left out too much essential information and I think that the way it is being presented in this way does a much better job at gauging the accuracy of the information to each topic.

1. Would you like to change or add anything in this visualization? Why? And How?

I think this visual representation is good as it is. It adds the core of the functionality that I thought the other one should implement while also adding a way to categorize accuracy.