NM2207

Session 09 Challenges

Before attempting the challenges, you are expected to have watched and coded along with the Lecture videos. A tutorial is meant to practice the skills presented in the video lecture, and show you more applications of it. Tutors will explain the challenges and answer the questions you may have**.**

**The challenges are due to be completed at the end of class each week for full credit which is also attendance. Submitting Saturday night of the same week accounts for half the credit.**

# Overview of what we will do today:

* Reading data from an online file (<https://2207-resources.s3.ap-southeast-1.amazonaws.com/senticnet_sg.csv>)
* Programmatically splitting data
* Visualizing data

**Part 1 – Reading Data**

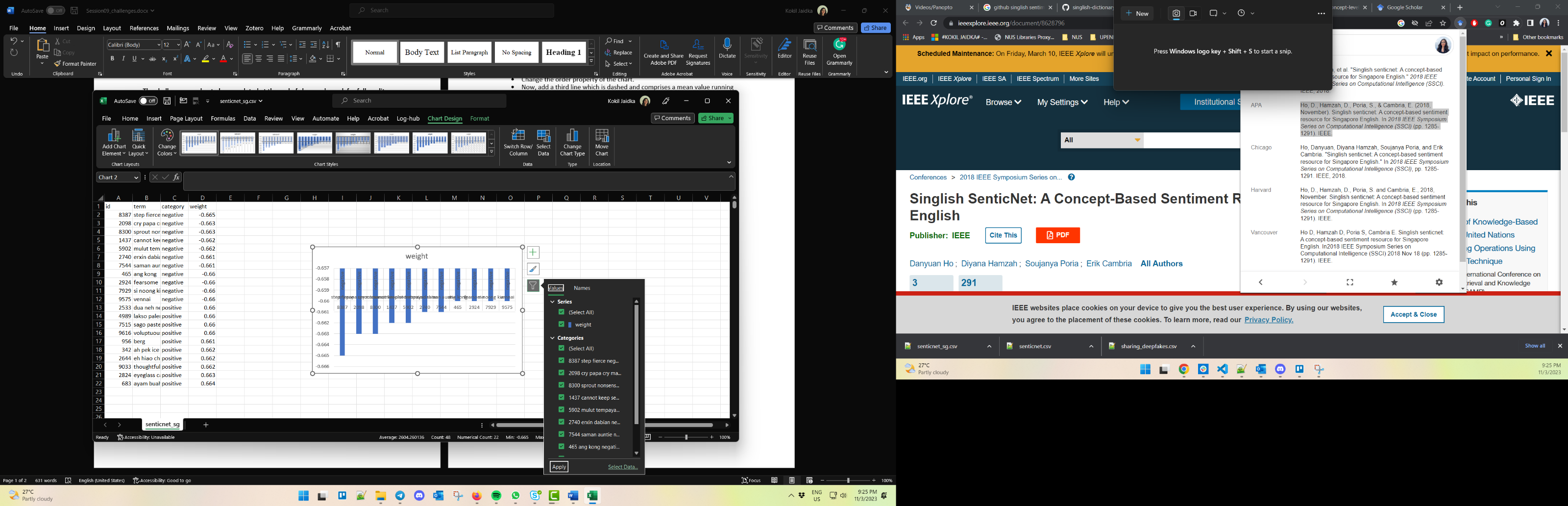
Following the steps similar to the codealong, read in the file (linked above).

The file is the research output of the following paper and describes the Singaporean/Singlish vocabulary and their associated sentiment values.

Ho, D., Hamzah, D., Poria, S., & Cambria, E. (2018, November). Singlish senticnet: A concept-based sentiment resource for Singapore English. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1285-1291). IEEE.

First, simply explore the file. Use your skills from Week 8 to:

1. Sort the data according to the weight column
2. Filter rows so that you can reduce the number of rows to a number that would look nice and understandable in a chart. Maybe the words with the top ten weights (most positive words) and bottom ten words (most negative words), for two charts?
3. Use Insert->Bar charts in Excel to explore what a possible chart could look like. First “Select data” to select only the axis labels that you want to keep.



****

**Part 2 Splitting Data**

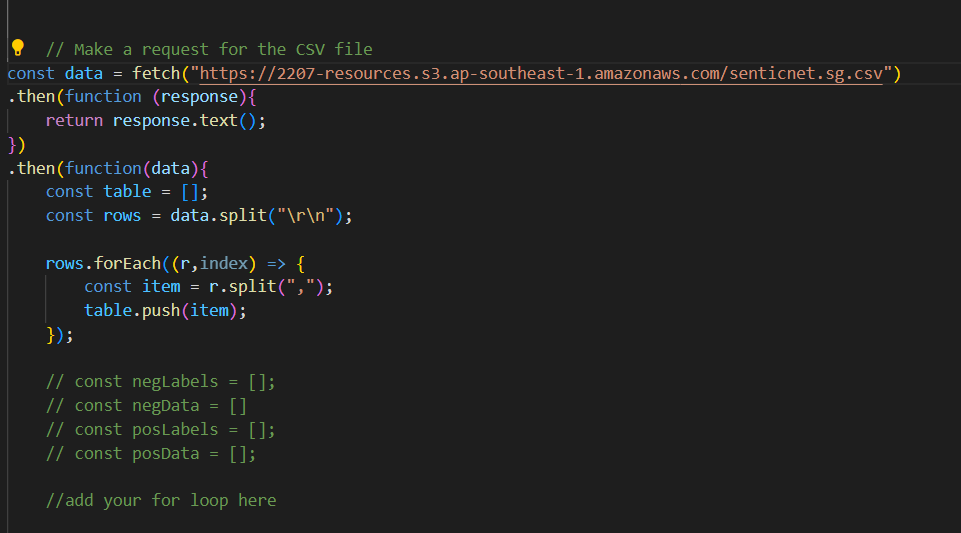
Now, follow the steps in the codealong to read in your data. Remember that it will read in the entire file into the table, not simply the rows and columns you want.

Next, use a for loop (for i =0; i<table.length;i++) and if statements within the for loop

This is simply an example to create a bar chart of the ten most negative and most positive words, but you can use a different idea.

if table[i][3] <-0.66 (ie if you want that in order to be on your plot, the 4th column in the ith row should be less than -0.66) to:

1. “push” your labels into a negLabel array (negLabel.push(table[i][1])),
2. and data into a negdata array (negData.push(table[i][3])
3. Similarly, for most positive words, in a different labels and data aray



Very importantly, within the for loop check using console.log if table[i][3] is indeed the value you want.

**Part 3 Visualize**

Now, modify your bar chart so that it’s able to use your labels and data to show the top ten most negative words in the Singlish sentiment lexicon. You will need to change the dataObj’s properties, such as the labels, the data, the chart title.

Does it match what you expected based on your Excel plot? Can you change the fill colors?

Can you make the chart horizontal so that the labels are easier to read?

Can you create a second bar chart for the positive words?