Maggie Jakus
2/25/15
Software Design
Mini Project 3

Project Overview

For my project, I used Twitter and Bing. I pulled trending topics from Twitter and then searched Bing for each of the topics. I then calculated the average polarity and subjectivity of each trending topic. From this project, I hoped to be able to compare the sentiment of popular topics of discussion at any given time.

Implementation

I used Pattern to access Twitter and Bing data. Twitter returns trending topics as a list of strings, and I put this list of strings into Bing to be analyzed. I stored the topics and their corresponding Bing search results as a dictionary, so that I could quickly access the results. My goal was to analyze these Bing search results so I could see what people were talking about at any given time and how they feel more generally about these subjects. I took my dictionary of topics and Bing searches and analyzed each result from the Bing search for a particular topic for its polarity and subjectivity. I calculated the average subjectivity and polarity for each topic and returned a dictionary containing topics as keys and their corresponding polarity and subjectivity as tuples.

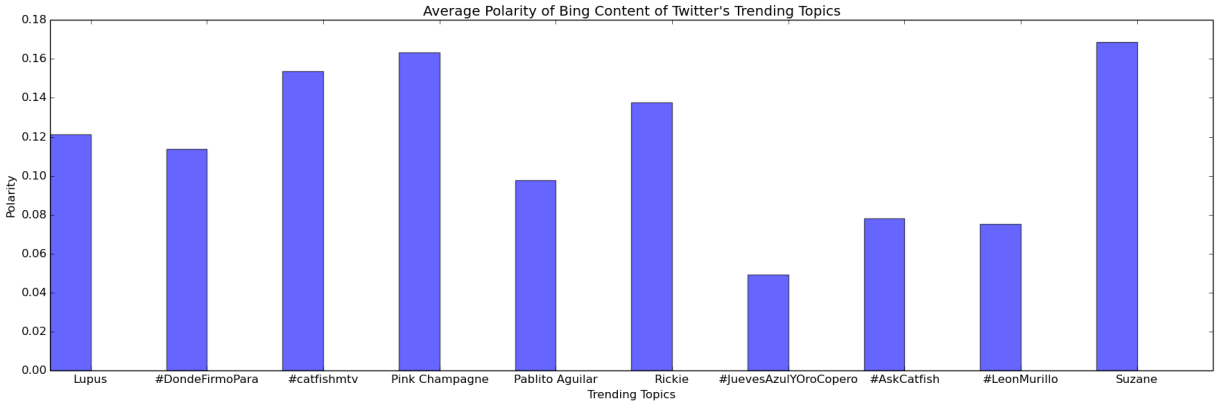
One decision I had to make was whether to combine my two bing functions into one. This would have kept everything in the same place, but would have made keeping track of dictionaries and values much more difficult. It also would have made debugging much harder, as a lot of my issues had to do with the type of data I was analyzing. For example, it took me a while to realize that when I return a dictionary of keys (trending topics from Twitter), the value is not a list of strings, but a list of lists of strings. Had I kept my two functions as one, this error would have been much more difficult to detect.

Results

It was very interesting to see which things returned positive results and which did not. The TV show “Catfish” is on MTV from 8 am to 11 pm on Wednesdays, when these data were collected. As can be seen from the data, Catfish was a popular trending topic, and the data on Bing must also be somewhat positive. What was surprising to me was that “Lupus” was a trending topic, and that it had such a high polarity.

No trending topic had an overall negative average polarity, and the average polarities were all within a relatively small range. How is language used? What content makes it onto Bing's search radar? Who primarily uses Twitter, and how does a trending topic become a top 10 topic? It would

be interesting to know the answers to these questions in order to better understand the data.



Reflection

Overall, I think this project went well for me. This was the first project I did where I figured out the structure of my project almost entirely on my own. I had a lot of trouble with running out of searches I could use. I needed several Bing API licenses to get through this project, which was mostly due to the fact that I didn't plan ahead. I wish I had created a better visual representation of my data. What I'm trying to show is very cool, but I wasn't able to present it well.