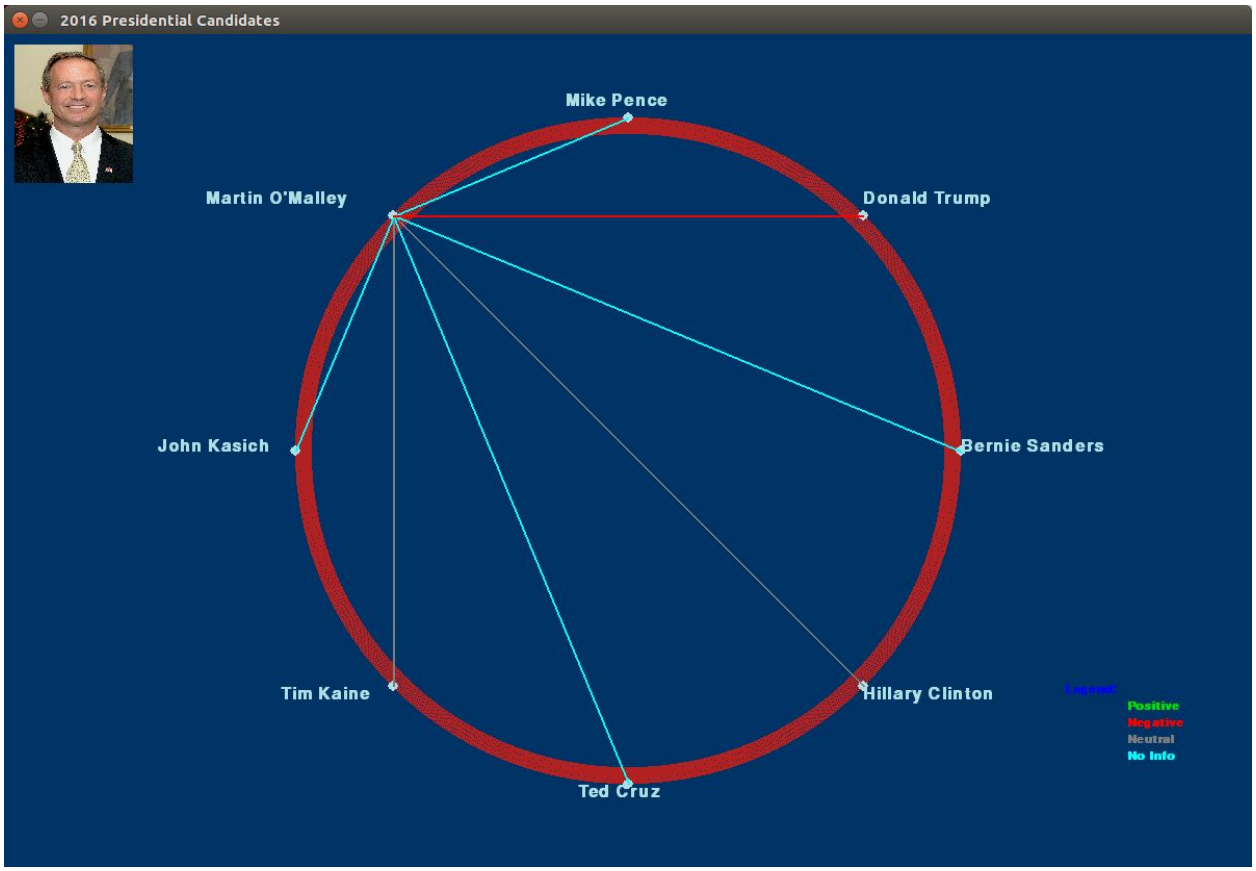
Project Partners
Prava Dhulipalla
Minju Kang
Alex Core

Project Overview
This interactive data visualization project mined tweets from various candidates from the 2016 Presidential Election. Sentiment analysis was performed on the tweets to determine if it was mostly negative, positive, or neutral. The main component of the project was the final data visualization, which included a ring with the candidate's names on it. If one hovers over the candidate's name, an information box pops up and lines pop, color-coded red, green, and grey, based on whether their tweets about another candidate was largely neutral, positive, or negative (line per tweet about another candidate). Clicking on a candidate's name causes the lines to stay.

Results



Although our original code (the unfinished code) was interactive, all it did was display the candidate's information and bios when hovering over the name. However, it didn't fit in with our MVP plan, nor did it utilize to previously existing code files (the class is called Software Design,

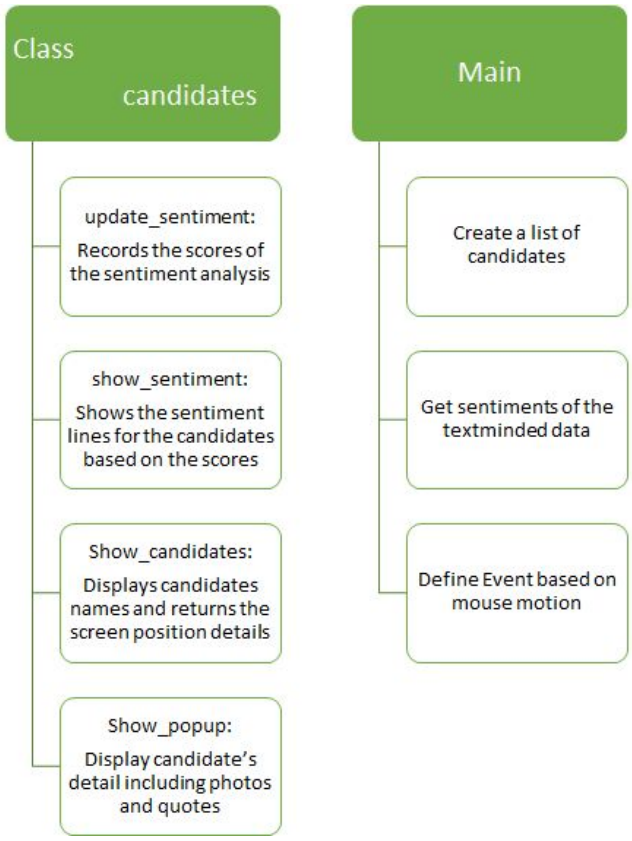
after all). Now, the code uses tweets, analyzes them, and displays lines going towards other candidates. If the line is red, it means the tweets were mostly negative. Green, positive. Grey, neutral. And finally, blue, positive. Because this is found based on a sentiment analyzer that was created 'from scratch,' it might be slightly off but overall, it seems fine. One thing that would be good for future steps is including more data so that there are less 'No info' lines. The lines are only displayed when one hovers over the candidate's names, which is apparent when one reads the README.md file, but not otherwise (another place to change). However, it was finished, it fit our MVP, and we all survived.

Implementation

First, to conduct the Sentimental Analysis on the twitter data, we pickled the twitter data. Pickling the twitter data was simply proceeded by using pickle library of python. By using retrieve_tweets function and plugging in the November 18th's twitter API number with user's id (ex: for Bernie, retreive_tweets('@SenSanders', 'bernietwitters.pickle', '796785451713593344' (API number for November 18th's twitter)), twitter data are saved as .pickle form. By using if statement, the code loads the saved data if there is any therefore it does not mine the text every single time whenever it runs.

Second, the phrase extractor was the portion of the code that split up the tweets into phrases/sentences. Then, it performs sentiment analysis on them based on the negative and positive word list. This is done on each of the candidate's tweets. Here is where we had to make a decision whether to make our own sentiment analyzer or to use the preexisting one. Ultimately, creating our own one fit in with learning goals/personal interests more. The main file displays the graphic interface and takes care of all the 'mouse hovering' and line drawings. main.py is the file that is run when you want to program to run. Although the functions of these are very basic, they were relatively difficult to implement (as somebody who hasn't done this type of thing before).

Although we wanted to do a lot more with the project, ultimately, we decided against it. For MP5, Prava finished the code, and even that took more than the max hours needed for MP5. If we wanted to pass the MVP, that would have required a lot more time. Given that we had a result that was satisfactory, we left the code at that.



Reflection

Up until the few last days of the project, we thought our project is under our control. However, our teamwork and teammates didn't get along very well therefore we couldn't finish the project in the time. Even though we didn't finish the project in time, each of our teammate achieved the personal goal. Minju learned how to pickle the twitter data which she didn't know for the Mini Project 3 and general knowledge about pygame while Prava learned how to implement pygame and interactive programming generally. Alex also learned how to implement pygame and start developing interactive programming.

Prava is finishing this project for her Mini Project 5 and finally, we will complete the Mini Project 4. Since we achieved our personal learning goal, this project could have been successful if teamwork went well and if we had better communication within us. Not finishing the project is unfortunate but still we learned a lot. We should have taken the advice to finish the write-up and submission with only half of the work done, but we really thought that we would be able to get it to finish, due to some false promises, so we held off - this was also a mistake.

Further steps would be mining more data, either from Twitter or other sources, adding more candidates, and improving the graphical interface.