Dylan Vanelli
EDUC295S Final Project
April 30, 2018

Research Reflection Memo

The primary reason I was interested in conducting social media research is because it is a topic that I have never perused academically, despite the fact that I use social media almost everyday. As a computer science major, I knew that I wanted to center my final research on something quantitative. I really believe that data can tell a powerful story if presented in a digestible way. After sitting in on Brian Norberg's seminar on social media analysis with Tableau, I knew immediately that I would focus on the same type of research for my final project. After Case Study 3, I also knew that researching a contemporary issue is the most engaging to me. Therefore, I decided to harvest data from Twitter related to the #MarchForOurLives D.C. event, which took place right around the time we were given details on our final projects.

Conducting data harvesting on Twitter is a difficult task and needs to be well-organized. Brian and I ran into a few bumps along the way. First, we ran into a tricky software compatibility issue with Tableau. With Tableau's most recent version, the data Brian harvested for me was corrupted. I fixed this by installing an older version of Tableau and it worked like a charm. The other challenge to conducting this sort of research is timing. At first, I was not certain which time intervals of the protest would contain the richest data. As a consequence, I did not plan out which time intervals I wanted beforehand, so I ended up pestering Brian a few times in order to get data that covered the most important parts of the event. Because data collection is computationally expensive, after each request I would have to wait a few days to receive data (no fault to Brian. Data collection just takes a long time without a supercomputer).

This kind of research poses one more unique challenge. Because the dataset is original, the analysis must also be original. On the web, the data of many social media events has been analyzed, but since #MarchForOurLives D.C. was so recent, examples to base my research on were scarce. As a result, all of my subject matter was made through experimentation. Using Tableau, I compared many different data variables in order to discover trends (or a lack thereof). In the end, this process made me more proficient with Tableau, which culminated in me being able to produce sophisticated visualizations by the end of the project.

I highly encourage students to peruse this kind of research. Delving into the unknown and discovering new trends is incredibly interesting. However, I do stress that as the dataset grows, it becomes increasingly difficult to remain organized. In addition, software issues will always be prominent as software is constantly updated. Lastly, Brian Norberg is one of the best assets at Duke when conducting this kind of research. With his knowledge of social media and data analysis tools, this kind of project is possible.