*Division of Computing Science and Mathematics*

*Faculty of Natural Sciences*

*University of Stirling*

Social Media and Approval Ratings

Louie Rincon

**Dissertation submitted in partial fulfilment for the degree of   
Master of Science in Big Data**

**August 2020**

Abstract

Twitter is one of the 10 largest social media platforms with over 329 million monthly users[4] and over 145 million active daily users[4]. With an American user base of over 48 million users[9], Twitter was used by Donald Trump prior to and during his tenure as President of the United States. President Trump has used the platform to help keep himself in favour with his supporters who - until the COVID-19 pandemic held him to a fairly consistent approval rating[figure 6].

The aim of this dissertation was to utilize sentiment analysis on Twitter data to investigate any correlative relationship between President Trump’s tweets, sentiment scores and approval ratings prior to and during the COVID-19 pandemic.

To accomplish the goals and objects set forth for this project, python Twitter web scraping packages were used to scape historical tweets made by President Trump from his personal social media account. And the statistical analysis software SPSS was used to calculate the correlation tests.

The objectives were to: Calculate the sentiment scores prior to the pandemic; calculate the sentiment scores during the pandemic; obtain President Trump’s approval rating data prior to and during the pandemic; visualize the resulting sentiment scores and approval ratings; and determine if any correlation between the sentiment data and the approval rating data existed.

The final results were inconclusive since there appeared to be a very low correlation relationship between the Twitter sentiment data, and the approval ratings data. However, additional findings made during the study, such as key characteristics of President Trump supporters, approval rating disparities between white and non-white voters, and the identification of significant dates which influenced President Trump’s approval ratings may suggest that further research in this area is merited.

Attestation

I understand the nature of plagiarism, and I am aware of the University’s policy on this.

I certify that this dissertation reports original work by me during my University project except for the following (*adjust the list below according to the circumstances*):

* The python code used to collect twitter data was taken from examples provided on GetOldtweets3 <https://pypi.org/project/GetOldTweets3/> and modified for the application of this project.

**Signature** *Louie Rincon* **Date** 28/08/2020

Acknowledgements

I would like to thank Dr. David Cairns, my thesis advisor, for his support and guidance this past summer this project, that was filled with a lot of unknow questions and without your direction I would not have been able to complete this project.

Table of Contents

The table of contents below is automatically generated from the paragraphs of style *Heading N* and *Unnumbered N*. To update this after revisions, right-click in the table and choose *Update Field* for the entire table; you may also be prompted about updates when printing.

Abstract 1

Attestation 2

Acknowledgements 3

Table of Contents 4

List of Figures 5

1 Introduction 6

1.1 Scope and Objectives 7

1.2 Achievements 7

1.3 Failures 7

1.4 Overview of Dissertation 7

2 State-of-The-Art 9

3 Project Implementation 11

3.1 Data collection, processing in python, and correlation analysis 11

3.1.1 Selected method for data collection and analysis 11

**3.1.1.1** **Data collection** 12

3.1.2 Interesting Discoveries and limitations of analysis approach 13

3.2 Visualization of Results 16

4 Conclusion 24

4.1 Summary 24

4.2 Evaluation 24

4.3 Future Work 24

References 26

Appendix 1 27

Appendix 2 – User guide 28

Appendix 3 – Installation guide 29

List of Figures

Similarly you can automatically generate a list of figures from paragraphs of style *Figure*. To update this after revisions, right-click in the table and choose *Update Field* for the entire table.

[Figure 1 Example sentiment analysis 7](#_Toc49941912)

[Figure 2 President Trump tweet to support a border wall 10](#_Toc49941913)

[Figure 3 Code used to retrieve Twitter data 12](#_Toc49941914)

[Figure 4 Code used to calculate sentiment scores 12](#_Toc49941915)

[Figure 5 Sentiment scores 12](#_Toc49941916)

[Figure 6 President Trump approval data 13](#_Toc49941917)

[Figure 7 President Obama approval ratings 13](#_Toc49941918)

[Figure 8 White voter approval ratings 2019-2020 14](#_Toc49941919)

[Figure 9 Non-white voter approval ratings 2019-2020 15](#_Toc49941920)

[Figure 10 President Trump 2019 approval ratings 17](#_Toc49941921)

[Figure 11 2019 Sentiment Scores 18](#_Toc49941922)

[Figure 12 Approval ratings and sentiment scores combined 19](#_Toc49941923)

[Figure 13 2020 approval ratings 20](#_Toc49941924)

[Figure 14 2020 sentiment scores 21](#_Toc49941925)

[Figure 15 Approval ratings and sentiment scores combined 22](#_Toc49941926)

[Figure 16 2019 Correlation test results 23](#_Toc49941927)

[Figure 17 2020 correlation results 23](#_Toc49941928)

# Introduction

Twitter is a social media platform that was first utilized by President Obama which led to him being dubbed the “first social-media president, President Obama was indeed the first presidency to make use of services like Twitter, Facebook, Snapchat, and Instagram.”[1] and as a result of him and his administration adoption of social media it helped get him elected as the 44th president of the United States. Him having done so created a formula for other political leaders to follow, which gave them an opportunity to interact digitally with their supporters and voter bases which historically was only done in person before the rise and utilization of social media.

In the recently aired BBC documentary “Trump in Tweets” the documentary walks you through “Trump’s first innocuous tweet to his early celebrity beefs, from his feuds with politicians to the extraordinary exchanges with North Korea’s leader and his continuing war on anyone who opposes him, Trump uses Twitter to tell the world what he is thinking and doing at all times.”[2]

“The film explores how his tweets provide an insight into his psyche, as well as maintaining his cult-like popularity amongst his followers. And in light of the recent events and protests across the US (and the rest of the world) the programme asks whether his tweets are a threat to democracy - or the mark of a president who has revolutionised the way leaders communicate with their people.”[2]

With over 85.5 million followers on his personal Twitter account[3] President Trumps reaches a very wide global audience, and with such a large Twitter following it led to questions if his Twitter followers particularly American citizens approval rating of President Trumps is at all influenced or impacted by the tweets he makes from his personal Twitter account.

**Background and Context**

Twitter has over 329 Million users on its platform[4], President Trump is one of those users, he also has over 85.5 million followers on Twitter[3]. He uses his personal Twitter account quite frequently as it is not controlled like the presidential Twitter account is, so he is quick to go to his personal account as he is free to say and so anything he likes since it is unmoderated.

With all of his Twitter use comes tweets that can be analysed and looked at to produce a sentiment analysis and ultimately a sentiment score. This will be addressed in further detail in section 2 State-of-the-Art.

A close up of a logo

Description automatically generated

Figure Example sentiment analysis

## Scope and Objectives

The scope of this project was to look at historical, and current Twitter data from President Trumps personal Twitter account and collect his approval ratings during his presidency. The goal was to conduct a sentiment analysis of his tweets before the corona virus pandemic, and during the pandemic while producing visualizations showing both his sentiment scores, and his approval ratings during the set time periods that I have been chosen. The chosen timeframe dates were as follows January of 2019 to December of 2019 for the pre-pandemic Twitter data, and January of 2020 to July of 2020 for the Twitter data during the pandemic.

The following were key objectives with this dissertation:

* Calculate the sentiment scores prior to the pandemic
* Calculate the sentiment scores during the pandemic
* Obtain his approval ratings prior to the pandemic
* Obtain his approval ratings during the pandemic
* Visualize his sentiment scores and approval ratings
* Try to identify any correlation between the two sets of data

## Achievements

This section provides the major achievements in relation to the objectives of this dissertation.

Objective 1: Calculating the sentiment scores of president Tumps tweets prior to the pandemic was accomplished

Objective 2: calculating the sentiment scores of president trumps scores during the pandemic

Objective 3: obtain President Trumps approval ratings data prior to the pandemic

Objective 4: obtain his approval ratings during the pandemic

Objective 5: create visualizations of the sentiment scores, and his approval ratings

## Failures

Objectives that were unable to be accomplished:

Objective 6: establishing a correlative relationship between president Trumps tweets sentiment scores and his approval ratings. This will be discussed further in the chapter 3 section

## Overview of Dissertation

The dissertation will outline the steps taken to reach the objectives that were set out for the project. Below are the descriptions of what each chapter covers:

Chapter 1 – The background and context of the project will be outlined, objectives set, and

achievements listed.

Chapter 2 – state-of-the-art will cover previous studies conducted on sentiment analysis and its applications to Twitter data.

Chapter 3 – The project implementation chapters will cover, the various technologies used to collect the Twitter data, process the data, visualize the of the data, and discuss the results of the project.

Chapter 4 – The Evaluation, and summary of the project will be provided, as well as any conclusions that have been drawn from the project; as well as any future project work will be suggested.

# State-of-The-Art

What is sentiment analysis? “Sentiment Analysis is the process of determining whether a piece of writing is positive, negative or neutral. A sentiment analysis system for text analysis combines natural language processing (NLP) and machine learning techniques to assign weighted sentiment scores to the entities, topics, themes and categories within a sentence or phrase.”[5]

Sentiment analysis approach is broken down into 4 steps:

* Step one: Break each text document down into its component parts (sentences, phrases, tokens and parts of speech)[5]
* Step two: identify each sentiment-bearing phrase and component[5]
* Step three: Assign a sentiment score to each phrase and component (-1 to +1)[5]
* Step four (optional): Combine scores for multi-layered sentiment analysis[5]

When conducting sentiment analysis, two scores are generated in the results of the analysis; Polarity and Subjectivity. For polarity the score range is -1 to 1 meaning the closer to -1 is the more negative it is and the closer to 1 the score is the more positive it is. Subjectivity the score range is 0 to 1 meaning the closer to 0 a score is the more objective the statement is meaning it is more factual. The closer to 1 a score is the more subjective it is meaning is it a personal opinion.

External research of sentiment analysis revealed 5 key points about sentiment analysis and classification. Key point 1: Social Media is the main resource for sentiment analysis. Key point 2: before starting sentiment analysis, pre-processing is needed to remove numbers, stemming, speech tagging, punctuation, lowercase and stop words. Key point 3: sentiment classification is broken into three approaches Machine learning, lexicon-bases, and hybrid. Key point 4: sentiment analysis uses the evaluation metrics of precision , recall, f-score, and accuracy. Key point 5: visualization of the results, graphs, histograms, and confusion matrices are utilized.[14]

The journal publication “The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers” looks at research topics in which sentiment analysis has been used. They concluded:

“We have seen a massive increase in the number of papers focusing on sentiment analysis and opinion mining during the recent years. According to our data, nearly 7,000 papers of this topic have been published and, more interestingly, 99% of the papers have appeared after 2004 making sentiment analysis one of the fastest growing research areas. Although the present paper focuses on the research articles of sentiment analysis, we can see that the topic is getting attention in the general public”[15]

In researching for similar dissertations that were focused on analysing sentiment analysis scores and attempting to find a correlation between a government officials approval rating. Two discoveries were made:

The first discovery is that a majority of Sentiment analysis research projects tend to be focused on the various methods of sentiment analysis, machine learning, and their applications to Twitter data. Which is further supported by “The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers” publication. In this dissertation example “A Comparative Study of Twitter Sentiment Analysis Methods for Live Applications” the abstract of the dissertation is “Sentiment Analysis is growing exponentially due to the importance of the automation in mining, extracting and processing information in order to determine the general opinion of a person. The problem that this thesis proposes to address is to determine what methods are more suitable to extract subjective impressions in real time from Twitter. For live applications, since the opinions collected from Twitter are limited to certain amount of characters and it will happen in a real-time environment, this provides an interesting scenario; we will test using both the Machine Learning Approach and the Lexicon-based Approach, and then combine them in an effort to increase the accuracy. In order to test the real-time factor, I will implement a web service with the purpose of collecting real-time feedback from Twitter in real-time, which will be later processed and analysed for accuracy and real time performance.”[6]

Here is another example of a dissertation that is focused on sentiment analysis: “Sentiment Analysis and Opinion Mining has become a research hot-spot with the rapid development of social network websites. Twitter is a typical social network application with millions of users expressing their sentiment every day. In this work, we explored comprehensively the methodologies applied in sentiment classification over Twitter data: lexicon-based, rule-based and machine learning-based methods.”[7]

The second discovery is that this dissertation appears to be the first of its kind. A project that is attempting to utilize sentiment analysis scores, to explore if there is a correlative relationship between the sentiment scores of the tweets that have been made by President Trump, and his public approval ratings. The initial hypothesis is that there may be a correlation between the data. President Trumps presidential campaign relied on racist resentment and anti-Immigrant sentiments[8] which he tweeted about these topics during his presidential run in order to stir up outrage and drum up support for his candidacy. Which is why the hypothesis is that if he used Twitter to gain followers and support for his presidential run, in theory his continued use of Twitter should continue to influence his supporters or non-supporters opinions of him.

A screenshot of a cell phone screen with text

Description automatically generated

Figure President Trump tweet to support a border wall

# Project Implementation

This section of chapters will cover the technology, tools, and methods used to conduct sentiment analysis. This section will also cover the tools, and technology used in this project. In order to obtain, visualize the data, and run the statistical analysis. This section will also cover discoveries that were made while the project was underway, the resulting visualizations, and overview of the results.

## Data collection, processing in python, and correlation analysis

When conducting a sentiment analysis: The python coding language, and a series of python packages that consist of machine learning, analytics, and visualization tools are typically used to extract, analyse, and visualize the data.

In order to conduct a Sentiment analysis on language you need to utilize the Natural Language Toolkit (NLTK) python package this python package allows you to process the unstructured text that is located within a Twitter tweet. Other python packages that are typically used to process the text data, and visualize it are typically pandas, numpy, and matplotlib.pyplot.

For this project data was going to analysed from pre-pandemic dates and during the pandemic so the chosen timeframe was going to be January of 2019 to December 2019 for pre pandemic dates, and during the pandemic the dates chose were January 2020 to July 2020. In order to do that an alternative needed to be chosen in order to accesses the amount of Twitter data needed in order to conduct the correlation analysis of sentiment analysis scores and President Trumps approval ratings.

Web scraping was used to access all of the necessary twitter data. Web scraping has no rate limits and does not required the individual to sign up for a developer account with Twitter. Since web scraping is not using Twitters API tool the tweet data is pulled directly from the Twitter website, as long as the Twitter users whose tweets are being collected profile is not private, and the tweets are public the information can be collected.

In order to conduct the correlation analysis SPSS software was used. “SPSS software platform offers advanced statistical analysis, a vast library of machine learning algorithms, text analysis, open source extensibility, integration with big data and seamless deployment into applications.”[10]

### Selected method for data collection and analysis

The method selected to collect the Twitter data was via web spacing method utilizing a python package called Get Old Tweets, which allows you to input the desired Twitter account into the python code and pull the desired amount of data with relative ease and no rate limits. The NLTK package that was chosen to conduct the sentiment analysis and scoring is a python package named TextBlob. “TextBlob is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.”[11]

The decision to use TextBlob for NLP analysis was the result of looking at other sentiment analysis projects and discovering that TextBlob was a commonly used python package for the NLP analysis.

#### **Data collection**

After deciding on the method to be used for data collection, and the NLTK package to be used to conduct the sentiment analysis. Then the execution of pulling the data into a python notebook and analysing the data via TextBlob. In figure 3. you can see the code that was used to in order to pull the Twitter data in the python notebook. In figure 4. you can see that code that was used in order to generate the sentiment scores for the tweets. In figure 5. you can see the generated sentiment scores, after conducting the sentiment analysis of President Trumps Twitter data.

A picture containing bird

Description automatically generated

Figure Code used to retrieve Twitter data

**A screenshot of a cell phone

Description automatically generated**

Figure Code used to calculate sentiment scores

A picture containing knife

Description automatically generated

Figure Sentiment scores

Once all of the tiwtter data was collected, and analysed; then the process of obtaining the most reliable data sournce for presendtial approval ratings was undertaken. That led to the discovery of a data set provided by the University of California Santa Baraba which has compiled the presidential approval ratings via the Gallup Poll[12] which is a american analytics and advsoroy company that amongst other things conducts presedential surverys aross the united states.

A screenshot of a cell phone

Description automatically generated

Figure President Trump approval data

### Interesting Discoveries and limitations of analysis approach

Some key discoveries and limitations made during analysis and research processes

* Twitter API call limited access to tweet data
* President Trump was a poor candidate for analysis
* Attempts to generate word clouds and other visualizations were not needed
* White trump supporters approval ratings are very flat until May of 2020 had a 9% drop in approval
* Trump supporters have a distinct personality characteristic

Obtaining the necessary Twitter data: When researching options for gather Twitter data, the two methods that are available to use are either using Twitters API service, which is free to use after you have applied for and have been granted accesses to a Twitter developer account. But after lots of research, a key discovery was made with the Twitter API. When collecting data via Twitters API call, users are rate limited to collecting 200 tweets per API call and are allotted access to a maximum of 3,200 tweets. When using Twitters API to pull tweet data the data is only accessible to upwards of 7 days in the past. This major limitation in data access with Twitters API call meant that the API service for the data collection was not a viable option as it limited the scope of the project drastically and limited how much data could be accessed at any given time.

During the analysis of approval ratings, it was discovered that President Trumps approval ratings have been quite flat during his presidency as you can see in figure 6. There is one notable exception and that is in May/June of 2020 where you can see a noticeable drop in his approval ratings. President Obama approval ratings during his two presidential terms are quite dynamic. As you can see in figure 7. below you can see very clear and very sizeable fluctuations in President Obamas approval ratings during his Presidency.

A screenshot of a cell phone

Description automatically generated

Figure President Obama approval ratings

When initially starting on this project there were multiple attempts to generate word clouds and other visualizations. But after some reflecting it was realized that these were a wasted effort on this project. As the goal was to look at the numerical scores given to the Twitter data and attempt to correlate that information to the approval ratings data. A word cloud would be a beneficial piece to a similar project in the future; or if future work on this project was to happen. I will cover the benefits of a word cloud in the future work section 4.3.

Another discovery when undertaking this project was that white voters approval rating of president trump was extremely flat in comparison to non-white voters as you can see in the figure 8 and figure 9 images.

A close up of a map

Description automatically generated

Figure White voter approval ratings 2019-2020

A close up of text on a white background

Description automatically generated

Figure Non-white voter approval ratings 2019-2020

Trump supports distinct personality traits. While conducting this project an article in psychology today listed the 5 main personality characteristics of a supporter of President Trump:

1. Authoritarian Personality Syndrome

Authoritarianism refers to the advocacy or enforcement of strict obedience to authority at the expense of personal freedom and is commonly associated with a lack of concern for the opinions or needs of others. Authoritarian personality syndrome—a well-studied and globally-prevalent condition—is a state of mind that is characterized by belief in total and complete obedience to one’s authority. Those with the syndrome often display aggression toward outgroup members, submissiveness to authority, resistance to new experiences, and a rigid hierarchical view of society. The syndrome is often triggered by fear, making it easy for leaders who exaggerate threat or fear monger to gain their allegiance.[13]

2. Social dominance orientation

Social dominance orientation (SDO)—which is distinct but related to authoritarian personality syndrome—refers to people who have a preference for the societal hierarchy of groups, specifically with a structure in which the high-status groups have dominance over the low-status ones. Those with SDO are typically dominant, tough-minded, and driven by self-interest.[13]

3. Prejudice

It would be grossly unfair and inaccurate to say that every one of Trump’s supporters have prejudice against ethnic and religious minorities, but it would be equally inaccurate to say that many do not. It is a well-known fact that the Republican party, going at least as far back to Richard Nixon’s “southern strategy,” used strategies that appealed to bigotry, such as lacing speeches with “dog whistles”—code words that signalled prejudice toward minorities that were designed to be heard by racists but no one else.[13]

4. Intergroup contact

Intergroup contact refers to contact with members of groups that are outside one’s own, which has been experimentally shown to reduce prejudice. As such, it’s important to note that there is growing evidence that Trump’s white supporters have experienced significantly less contact with minorities than other Americans. [13]

5. Relative deprivation

Relative deprivation refers to the experience of being deprived of something to which one believes they are entitled. It is the discontent felt when one compares their position in life to others who they feel are equal or inferior but have unfairly had more success than them.[13]

This psychological information offered some key insights into a potential explanation as to why President Trumps approval ratings have stayed so consistent throughout the duration of his presidency amongst white voters. This information also offers up a potential way to reshape this projects or similar projects approaches in future work, this will be expanded on further in section 4.3.

## Visualization of Results

Two scores are generated for sentiment analysis. The first of these – known as Polarity ranges from -1 to 1, meaning the closer the score is to -1 is the more negative the sentence is and the closer to 1 the score is the more positive the sentence is. The second score – Subjectivity - ranges from 0 to 1. The closer the Subjectivity score is to 0 the more objective (or factual) the statement is. A score closer to 1 is indicative of a more subjective statement i.e. it is based on personal opinion.

For approval rating, two sets of data were visualised. The first set of data was the approval rating which looks at the individuals that have a positive view of President Trump. The second set is disapproval data which was for individuals who have a negative view of President Trump.

Six visualizations were created from the collected Twitter and approval ratings data. Three figures are from 2019 (pre-pandemic) and present 1) sentiment scores[figure 10], 2) approval ratings[figure 11] and 3) a combination of both sentiment scores and approval ratings[figure 12]. The remaining three figures[figure 13,14,15] present sentiment scores, approval ratings and a combination of both data sets during the pandemic in 2020.

Two images present the correlation between the 2019 pre pandemic data (which data), and approval ratings. The other results are for the correlation test of the 2020 pandemic sentiment data, and approval ratings. Different methods were attempted in order to identify the best way of presenting the resulting data. The data was displayed by year, quarter, and month. But as previous mentioned in section 3.1.2 the approval rating data is so flat that it didn’t yield any discernible differences, so it was decided to display the data by month as it best displayed the relative flatness of the data.

As you can see from the visualized data that it is extremely flat across the board. President Trumps approval ratings have had very little variations in the rating data. The key significant change is president trumps approval rating happened in May of this year when his approval rating dropped by almost 10 points. Even with this sentiment analysis scores they somewhat reflect his approval ratings with two sizeable variations in 2019 of his sentiment scores when the average went from about .02 to .25 around may and then dropped from 0.10 in October to -0.10 in November of 2019.

For the correlation analysis testing was conducted in SPSS. The Pearson correlation coefficient was chosen as the method to conduct the correlation testing, as the two data sets were both continuous sets of data. A looking at the results revealed that there were not any significant correlations between approval ratings and sentiment analysis scores. As you can see from the results of both correlation tests there was not a high level of significance in the relationship of the data.

Since the results of the pearson correlation coefficient were above 0 that means that there is some sort of correlation between the data; but it is not significant enough to factually say that there is in fact a statistical relationship between the two sets of data. The results of the correlation coefficient scores ranged from .1 to .4 which based on the Pearson scale that means that there was a small to medium relationships between the sentiment analysis data, and President Trumps approval ratings.

|  |  |
| --- | --- |
| Figure President Trump 2019 approval ratings |  |
| Figure 2019 Sentiment Scores |  |
| Figure Approval ratings and sentiment scores combined |  |

|  |
| --- |
| Figure 2020 approval ratings |
| Figure 2020 sentiment scores |
| Figure Approval ratings and sentiment scores combined |

A screenshot of a cell phone

Description automatically generated

Figure 2019 Correlation test results

**A screenshot of a cell phone

Description automatically generated**

Figure 2020 correlation results

# Conclusion

## Summary

The goal of this project was an attempt to use the sentiment scores of President Trumps tweets; and establish a correlation between the sentiment scores, and his approval ratings. This project was successful in its implementation, and analysis. What was a failure/remains inconclusive was the attempt to try to establish a correlation between the two sets of data was not achieved; there was not enough empirical data, to support a concrete relationship between the two sets of data.

As mentioned in the State-of-the-Art section, this project appeared to be the first of its kind, which left the approach of this project to be quite broad in scope, which is why it is felt that the correlation tests were not successes since the analyses was conducted on large and unrefined data set. But as progress was made on this project key discoveries were made about President Trumps supporters, key points where there were significant drops in his approval ratings.

The final result of this project has set up foundation for further studies, and investigations into establishing an empirical relationship between social media use, and its potential influence on presidential approval ratings. This will be expanded on in section 4.3 future work.

## Evaluation

In section 1.2 a set of goals and objectives for this project were established in order to ensure this projects success. Of the six objectives that were created for this project 5 of them were successfully completed.

The completed objectives:

Objective 1: The sentiment scores of president Tumps tweets prior to the pandemic was accomplished

Objective 2: The sentiment scores of president trumps scores during the pandemic

Objective 3: obtain his approval rating data prior to the pandemic

Objective 4: obtain his approval ratings during the pandemic

Objective 5: create visualizations of his sentiment scores and his approval ratings

Incomplete:

Objective 6: establish a correlative relationship between President Trumps tweets sentiment scores, and his approval ratings.

Through research and evaluation, it was determined what would be the best technology and methods to use in order to get this project completed. That technology was python, GetOldTweets, Textblob, and SPSS. These 4 tools were imperative in collecting, conduction the necessary sentiment analysis on President Trumps tweets, and generating the sentiment scores needed in order to perform the correlation analysis between the sentiment scores and President Trumps approval ratings. As you can see in figure 13., and 14. the results yielded very low correlation scores amongst the data. Which was not very surprising given the interesting information that was learned about voter profiles, and characteristics of President Trumps supporters as further research, and information was collected while undertaking this project.

## Future Work

The limitations in the results of this study were uncovered as the projected progressed further, the discovery of the flatness of President Trumps approval ratings meant he was probably a poor test subject as there was not a lot of changes over time with his approval ratings during his presidency, whereas President Obama has a lot of fluctuation in his approval ratings during his presidency. With more fluctuations in presential approval ratings there would have been more opportunities to target key dates to conduct a correlation analysis. Having a bigger data \ approval ratings data set. It appeared that typically approval rating surveys were conducted bi-monthly so there was not a lot of obtained data in regard to presidential approval ratings. Filtering the tweets is another way that could possibly improve the correlation test results as it would have removed certain sentiment scores that would not have been needed to be analysed if they did not meet the tweet filter criteria.

How this project could be developed further is by having a more targeted analysis, this project was quite broad in scope which makes no surprise that the results turned out as they did. As mentioned in section 3.1.2 some of the biggest discoveries were made while this project was underway. Learning about President Trumps supporters 5 key traits, the fact that white voters had a very consistent approval rating of President Trump before the significant dip in May of 2020. This project could be developed further to look at key points in spikes or drops in President Trumps approval ratings.

How this could be approached is by filtering his tweets during key dates, and using a filter to look for, and analyse those tweets in those specific timeframes. Another way to further improve further a future project would be to filter for tweets that fit his supporters characterise profiles, look for tweets that meet one of the 5 criteria that his supporters meet Authoritarian Personality Syndrome, Social dominance orientation, Prejudice, Intergroup contact, and Relative deprivation, e.g. Figure 2. And by doing so the analysis will be much more targeted; but also reflective of his supporters, and non-supporters which should help improve the results of a correlation test since the audience will be segmented into supporters and non-supporters.

This project has definitely opened up potential new ways to apply, and use sentiment analysis outside of the standard approaches.

References

[1] Ian Bogost, “Obama Was Too Good at Social Media” TheAtlantic.com, 06-01-2017, https://www.theatlantic.com/technology/archive/2017/01/did-america-need-a-social-media-president/512405/

[2] “Trump in Tweet” BBC.co.uk,06-07-2020, https://www.bbc.co.uk/mediacentre/proginfo/2020/27/trumps-in-tweets

[3] @realDonaldTrump, Twitter.com, 08-01-2020, https://twitter.com/realDonaldTrump

[4] Esteban Ortiz-Ospina, “The rise of social media” OutWorldinData.com, 18-09-2019, https://ourworldindata.org/rise-of-social-media

[5] “Sentiment analysis explained” lexalytics.com, 08-01-2020, https://www.lexalytics.com/technology/sentiment-analysis

[6] Angel Cambero “A Comparative Study of Twitter Sentiment Analysis Methods for Live Applications” B. Thomas Golisano College of Computing and Information Sciences Rochester Institute of Technology Rochester, New York 01-08-2016, https://scholarworks.rit.edu/cgi/viewcontent.cgi?article=10367&context=theses

[7] Bo Yuan “SENTIMENT ANALYSIS OF TWITTER DATA” Rensselaer Polytechnic Institute Troy, New York March 2016, https://www.cs.rpi.edu/~szymansk/theses/bo.ms.16.pdf

[8] Marc Hooghe, University of Leuven Ruth Dassonneville, Université de Montréal “Explaining the Trump Vote: The Effect of Racist Resentment and Anti-Immigrant Sentiments” 01-07-2018, https://www.cambridge.org/core/services/aop-cambridge-core/content/view/537A8ABA46783791BFF4E2E36B90C0BE/S1049096518000367a.pdf/explaining\_the\_trump\_vote\_the\_effect\_of\_racist\_resentment\_and\_antiimmigrant\_sentiments.pdf

[9] Salman Aslan, “Twitter by the Numbers: Stats, Demographics & Fun Facts” omnicoreagency.com,

 Last updated: Feb 10, 2020, https://www.omnicoreagency.com/twitter-statistics/#:~:text=There%20are%2048.35%20million%20monthly,monetizable%20daily%20active%20Twitter%20users.

[10] ibm.com, https://www.ibm.com/uk-en/products/spss-statistics

[11] “TextBlob: Simplified Text Processing” TextBlob, https://textblob.readthedocs.io/en/dev/

[12] Presidential Job Approval, https://www.presidency.ucsb.edu/statistics/data/presidential-job-approval

[13] Bobby Azarian Ph.D. An Analysis of Trump Supporters Has Identified 5 Key Traits, psychology today.com, 31-12-2017, <https://www.psychologytoday.com/gb/blog/mind-in-the-machine/201712/analysis-trump-supporters-has-identified-5-key-traits>

[14] Symeon Symeonidis, Democritus University of Thrace, kdnuggets.com, 2018 <https://www.kdnuggets.com/2018/03/5-things-sentiment-analysis-classification.html>

[15] Mika V. Mäntylä Daniel Graziotin, Miikka Kuutila, University of Oulu, Institute of Software Technology, University of Stuttgart, The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers, http://arxiv-export-lb.library.cornell.edu/pdf/1612.01556

Appendix 1

Appendix 2 – User guide

Appendix 3 – Installation guide

Python3

SPSS

Python package GetOldTweets3

TextBlob