**Your name(s): Pierre Arquillano**

**Executive Summary/Abstract**

This exploratory analysis focuses on Facebook statuses posted by Trump from a set period of time. The aim of this analysis is to gain insights into the reactions, sentiment of Facebook users towards statuses posted by Trump.

The dataset used for this analysis consists of Facebook statuses posted by Donald Trump from 2009-2016. The Data imported was scraped from Facebook and saved as TrumpStats.csv.

The analysis begins with a preliminary examination of the data, including basic statistics such as the distribution of types of statuses, the progression of number of reactions, and the op reactions expressed by the audience. The results reveal that Trumps Audience are prone to like his Statuses and are generally happy with them.

To further understand the interests and personal lives of Facebook users, topic modeling is performed on the statuses. The analysis identifies the most common themes and topics that Trump discusses in his posts, which include Trump, America, Great, Hillary, and Jobs. The results suggest that Trump is advocating for making America Great, talking about jobs and Hillary Clinton.

Overall, this exploratory analysis provides valuable insights into the behavior of Facebook users towards Trump and highlights the importance of social media as a platform for understanding an audience’s perception about specific people. The results can be used to inform future research on the impact of social media on presidential campaigns and what issues candidates should focus on to attract a larger audience.

**Introduction/Background**

The Project is a exploratory analysis on Facebook statuses posted by Donald Trump. The project uses machine learning methods and visualizations to gain insight from statuses posted by Donald Trump.

The Dataset in use is data scraped from Facebook. This data includes statuses posted by Donald trump, the date and time the status was posted and the numbers of reactions, likes, loves, sads and angrys each post received from Trumps audience.

**Methods and Results**

In this part, answer the questions asked in "Your Task" section of the assignment.

Briefly explain your methods and share your results and visualizations.

Chart, bar chart

Description automatically generated Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated Chart, bar chart, histogram

Description automatically generated Chart, bar chart

Description automatically generated

1. Use charts (histogram/bar graph) to display the distribution of types of statuses (photo, video, link, and status).

The visualization shows the distribution of the types of statuses posted by Trump. From the bar graph we can see that the top two types of statuses posted by trump are Photo and Status

2. Plot a chart showing the progression of likes, loves, and sads, and angrys over time.

The next two visualizations show the progression of reactions from Trumps audience towards his posts. We can see from the first of the two visualization that the likes reactions are trending upwards and have a steep increase as the beginning of 2016 approaches.

The second of the two visualizations is a progression chart which displays the newer reactions that were introduced in early 2016. This visualization subset the data in posts only from 2016 and onwards. As well as only tracking loves sads and angrys, leaving out likes because of the vast difference in counts.

3. What are the top reactions expressed by Trump’s Facebook audience?

The next visualization displays the top reactions by trumps audiences. The first of the two visualizations show that most expressed reaction by Trumps audience is that they like his posts.

The second visualization, removes like from the chart and focuses on the new reactions implemented to get a better idea of the audiences reactions. While a large majority of the audience Love Trumps reactions. It is followed by Angry, which shows that some of trumps posts angered his audience.

4. Analyze the content in the status\_message to find the top 10 topics (Topic modeling LDA can be used). Can you tell if fans are particularly happy or upset about the topics? Chart

Description automatically generated

From this visualization we can see most fans are happy with the topics. We can also see a large number of his upset audience are upset with topic 0

5. What are the themes (topics or top words) that generated a lot of shares? Chart, bar chart, histogram

Description automatically generated

Topic 0 generated the greatest number of shares by a large margin. Topics 1 and 2 generated the second and third most number of shares.

6. What are the themes (topics or top words) that generated a lot of num\_reactions?

Chart, histogram

Description automatically generated

This chart displays that Topic 0 has the most shares, Followed by Topic 2 and Topic 1

**Conclusions**

After analyzing this dataset, we can see that topics Trump focuses on are topics that his audience agree with. With a majority of the topics, his audience gave a positive reaction to. We can interpret this as Trump being able to cater to his audience and deliver on the topics that they will like. Through this analysis we are able to see how the audience feels about a certain topic and whether they are happy about the stance Trump is making about that topic. This insight shows that social media is a great source for determining what your audience feels towards you. It is a great tool to determine which topics generate the most positive reactions and with that information one can focus more on those topics to gain better favour with one’s audience.

Challenges faced when completing this project is that many posts contained hashtags. These hashtags contained combined words. This was dealt with by using regular expression operations to separate them by finding a pattern. Most hashtags capitalized the start of each word, therefore by searching for that pattern I was able to isolate each word from the hashtag.

**Future Research Considerations**

Other information that can be extracted in future work is creating a model to predict the reaction of Trumps audience on future status posts. By utilizing the status posts and their corresponding emotion, we can use this dataset as the training model for prediction of future status posts.