Twitter Mining and Analysis Across Various Twitter Accounts

Author ~ Rinaldo Gagiano -//- Advisor ~ Professor Ricardo Rodriguez

Contents

Introduction

Twitter is an online social media platform for the sharing of entertainment, thoughts, and news through tweets. Tweets are posted by account holders to their profile so their respective followers can see as well as anyone else who ventures onto their accounts.

Tweets generally consist of less than 280 characters but can contain a GIF or video. One may also retweet tweets, which involves attaching the original post to one's personal account. One can also post replies to a tweet, extending the original post with a new tweet that you wish to add.

The question we wanted to answer in this project was 'How does McPherson College's Sporting Twitter account compare with other accounts and what key insights could be drawn from contrasting different approaches to this social medium.' In answering this question, I will draw Twitter API data from various sources, clean and mine the data through the programming tool R, and create visualizations dictating the main aspects that Twitter accounts can be viewed on.

Data Analysis Techniques

Ratio of Replies/Retweets/Organic Tweets

This graph, produced from Twitter data, will show what type of content an account is producing. As mentioned previously an account holder has the option to tweet, retweet, or reply to tweets. Organic Tweets consist of only tweets originally produced by the account holder. This type of insight can show which particular accounts produce what kind of production relating to the purpose of the account.

Tweet Timeline Frequency

This chart depicts a timeline of all tweets obtained from an account, spread over weekly intervals. This type of data is useful as it allows for visual access into tweet frequency averages, tweet dips with potential causation, and time-sensitive trends associated with the particular purpose of the account.

Tweet Publish Locations

Each production of a tweet is logged from a device type. Output device data can give insight into what type of account is being run. A one-man controlled account or a team controlled account will show different trends within the publish locations. This can also be observed through weather the account creates time-sensitive tweets, with up to date productions, or if the account plans to post on a schedule with tweets already constructed and set for a specific date of publication. Note: Tweet Publish Locations does not incorporate data from geographical location.

Word Frequency Analysis

This bar chart involves counting all unique words with the removal of stop words. Stop words are common words such as 'it', 'and', 'the' etc. A chart depicting the usage of certain unique words allows insight into the accounts agenda. When a range of unique words have a higher frequency rate, the account is purposeful and directed with their posts. A lower frequency rate shows the account broadens their post discussions.

Wordcloud Hashtag Frequency Analysis

A hashtag is a word that follows the symbol (#) turning it into a relevant keyword or phrase within a tweet. This helps in categorizing said tweet which aids when doing a Twitter search. Selecting or clicking on a hashtagged word in any tweet will bring you to all tweets that include that hashtag. Hashtags are generally at the end of a tweet but can be included anywhere in a tweet.

A wordcloud is a visual production of hashtags consisting of their frequency rating. The larger the word in the graph indicates the more instances of said word and a difference in color outlines a new tier of increased frequency. Using this type of data visualization outlines the frequency of certain hashtag use and outlines whether the hashtag presence is fundamental to the account or not. Hashtag presence is the standard mechanism that contextualizes tweets among trends. This is useful for when an account establishes its trend to link the account to their followers and the greater Twitter world.

Wordcloud Retweet Account Frequency Analysis

A wordcloud on external user retweet frequency under analysis gives a key insight into the accounts' connection with its followers. If an account is tweeting relatable tweets, more followers will resonate with these posts and will retweet the account, increasing the original account's exposure on the Twitter platform. This, in terms of the wordcloud, will produce a large variety of names if the account is relatable.

Sentiment Analysis of Tweets

Sentiment analysis is based on the scoring of unique words reflecting feelings expressed by the users. These unique words are given a score through a built-in index. Sentiment analysis allows for the insight into what type of language an account is using in terms of tone when tweeting towards the account's followers or the topics of discussion.

Twitter Accounts Analysed

- @BernieSanders ~ Politician
- @elonmusk ~ Influencer
- @chiefs ~ Professional Sporting Team
- @kwucoyotes ~ In-Conference College Sporting Team
- @MAC_Bulldogs ~ Personal College Sporting Team

The reason for choosing these accounts is that each one represents a key insight into the different dealings of the Twitter world.

@BernieSanders - Politician Bernie Sanders

The choosing of Bernie Sanders's account helps to give a view into a politician's world through key insights into how and when they produce tweets. Generally, these types of accounts are run by a professional social media team. For this analysis, it serves as a gauge to judge how an effective Twitter account can be.

@elonmusk - Influencer Elon Musk

Elon Musk, being an influencer and sole contributor to his account, allows for insight into how a singular account holder can have such a large presence in the social media world. The analysis of his tweet interactions should bring key lessons for single account holders.

@chiefs - Major Sporting Team Kansas City Chiefs

The NFL is the largest sporting network on the planet. With this being said, the Kansas City Chiefs are a team within the NFL with a very large Twitter presence. The analysis of this account should translate over to McPherson College Bulldog's smoothly as both accounts share the same goal of promotion and brand awareness. Accessing data such as post timings and consistencies, post types, account hashtags presence, and account retweet presence will give key insights that can be applied to McPherson Account.

@kwucoyotes - College Athletic Twitter Kansas Wesleyan Coyotes

We have chosen Kansas Wesleyan University Sporting Twitter account as a peer institution within the KCAC. Analyzing the Kansas Wesleyan University Coyotes' Twitter account will give key comparables into which the institution's Twitter account better ticks the major boxes of maintaining a solid Twitter presence, and how McPherson can benefit from these.

@MAC_Bulldogs - College Athletic Twitter McPherson College Bulldogs

This is the account for which analysis is happening. The expectation is to give good feedback through a comparison of all other accounts and many of their attributes.

Data Acquisition - How to obtain Twitter API Data

Obtaining Data from Twitter is a straightforward process. All one needs is Twitter API access tokens and the R package 'rtweet'. To obtain access tokens one must first apply for a Twitter developer account. Steps to applying for this can be found at https://developer.Twitter.com/. Once approval has been given and you have access to your developer account you need to create a new app through the 'Create a New App' button located on the site. You will need to give your App a name, a description of what the app does, a Website URL (recommend using your Twitter Profile URL), a Callback URL (use http://127.0.0.1:1410), and a description on how you will personally be using the app. Then simply press the 'create' button. Once that is complete, to find the Twitter access tokens, one must go to the app details of the created app, select 'Keys and Tokens' as one of the headers and all tokens will be present on that section. We can now move to R and install the package 'rtweet'. With that installed we can run the following code (removing all single '#' marks) using those obtained tokens and the app name to create access to Twitter's API data.

```
## store api keys (these are fake example values; replace with your own keys)
#api_key <- "afYS4vbIlPAj096E60c4W1fiK"
#api_secret_key <- "bI91kqnqFoNCrZFbsjAWHD4gJ91LQAhdCJXCj3yscfuULtNkuu"
#access_token <- "9551451262-wK2EmA942kxZYIwa5LMKZoQA4Xc2uyIiEwu2YXL"
#access_token_secret <- "9vpiSGKg1fIPQtxc5d5ESiFlZQpfbknEN1f1m2xe5byw7"

## authenticate via web browser
#token <- create_token(
  #app = "app name",
  #consumer_key = api_key,
  #consumer_secret = api_secret_key,
  #access_token = access_token,
  #access_secret = access_token_secret)
```

Data Manipulation - Converting Raw Data into Usable Data

Installed Packages

(Packages installed within the coding platform R)

- Pacman
- rtweet
- tidyverse
- tidytext
- tm
- wordcloud
- syuzhet
- forestmangr
- rmarkdown
- tinytex

Tweet Mining

The following variables will contain the last 3200 tweets from each Twitter account being analyzed in this project. (3200 tweets is the maximum amount of tweets one can grab with a basic developers account)

Bernie Sanders:

```
bernie_tweets <- get_timeline("@BernieSanders", n=3200)
```

Elon Musk:

```
elon_tweets <- get_timeline("@elonmusk", n=3200)
```

NFL Chiefs:

```
chiefs_tweets <- get_timeline("@chiefs", n=3200)
```

Kansas Wesleyan Coyotes:

```
kwu_tweets <- get_timeline("@kwucoyotes", n=3200)
```

McPherson College Bulldogs:

```
mcpherson_tweets <- get_timeline("@Mac_College", n=3200)
```

Tweet Analysis

Tweet Analysis below is for Bernie Sanders Tweets. The same process is applied for each Twitter Account under analysis.

Remove retweets: Removal of all retweets from base 3200 tweets

```
bernie_tweets_organic <- bernie_tweets[bernie_tweets$is_retweet==FALSE, ]
```

Remove replies: Removal of all replies from base 3200 tweets

```
bernie_tweets_organic <- subset(bernie_tweets_organic, is.na(bernie_tweets_organic$reply_to_status_id))
```

(Organic Tweets are all tweets within the base 3200 with the removal of all replies and retweets. Essentially, organic tweets are all tweets originally posted by an account.)

Organic Tweet with most Favourites :

```
bernie_tweets_favorite <- bernie_tweets_organic %>% arrange(-favorite_count)
```

Organic Tweet with most Retweets

```
bernie_tweets_retweet <- bernie_tweets_organic %>% arrange(-retweet_count)
```

Keeping only the retweets

```
bernie_retweets <- bernie_tweets[bernie_tweets$is_retweet==TRUE,]
```

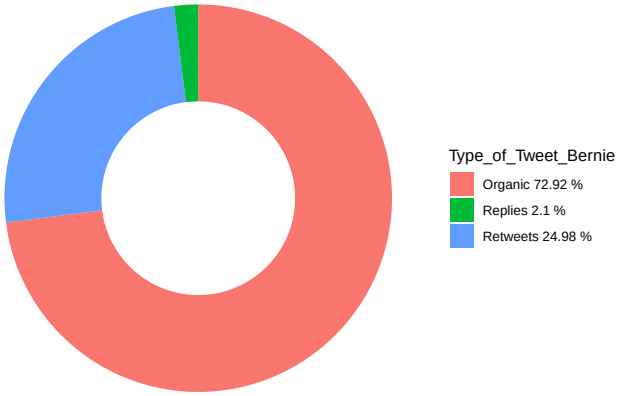
Keeping only the replies

```
bernie_replies <- subset(bernie_tweets, !is.na(bernie_tweets$reply_to_status_id))
```

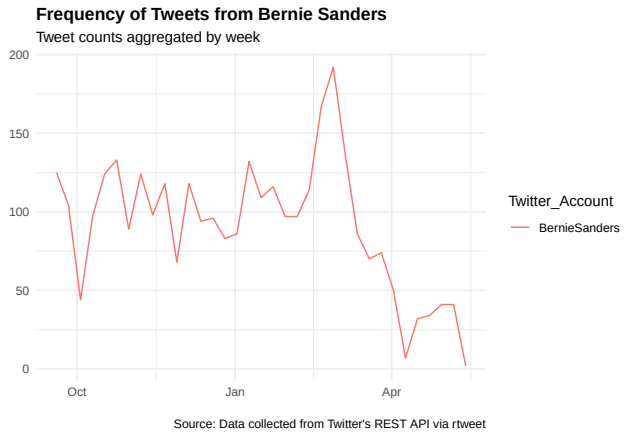
Data Analysis

Bernie Sanders

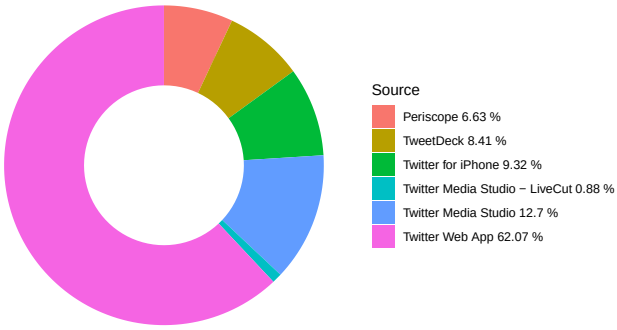
Ratio of Replies/Retweets/Organic Tweets



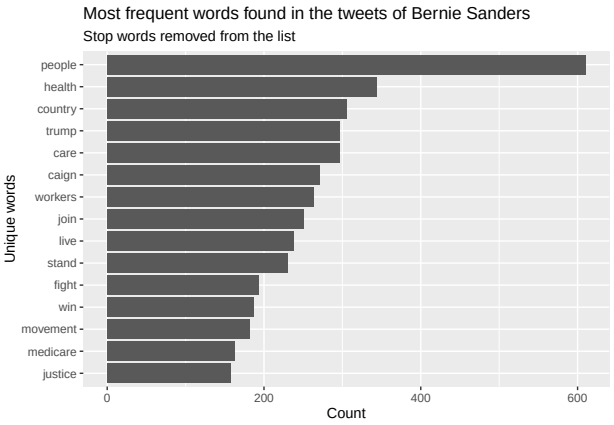
Tweet Timeline Frequency



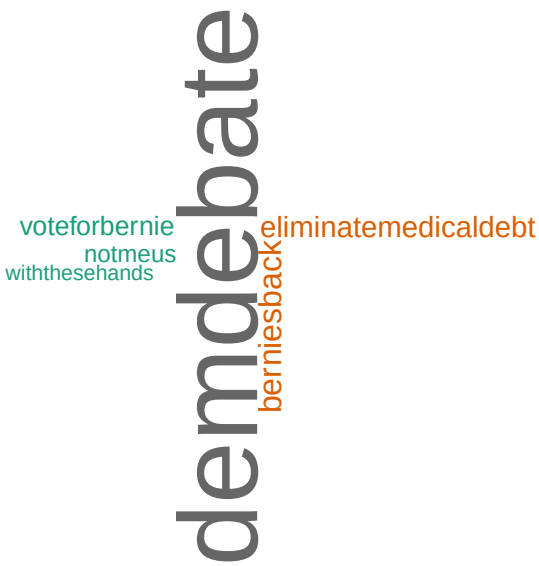
Tweet Publish Locations



Word Frequency Analysis

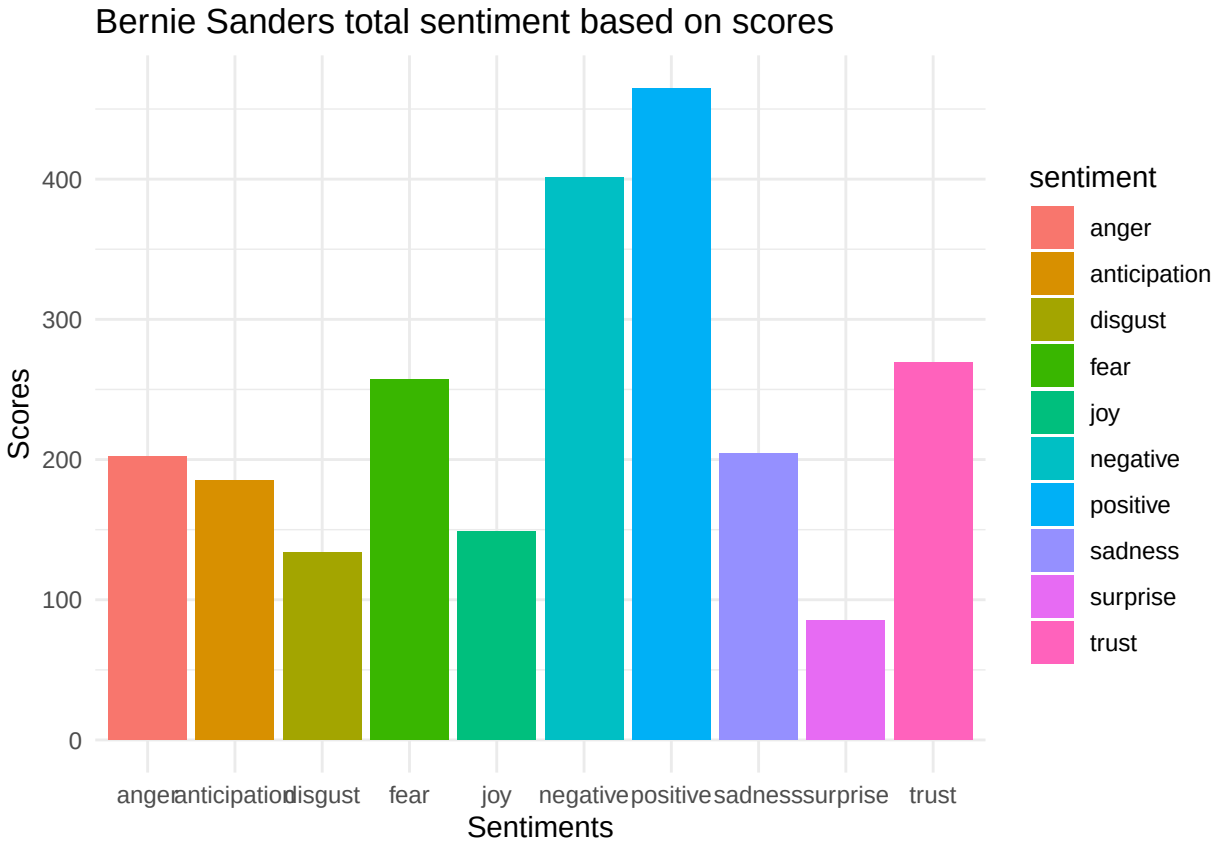


Wordcloud Hashtag Frequency Analysis



Wordcloud Retweet Account Frequency Analysis

Sentiment Analysis of Tweets



Ratio of Replies/Retweets/Organic Tweets

Bernie Sanders' ratio of replies, retweets, and tweets reveals that a major majority of his Twitter account production comes from organic original tweets. This is logical as a politician talks to his people through his or her voice rather than retweeting people or replying to comments.

Tweet Timeline Frequency

The tweet timeline frequency of Bernie Sanders's tweets consists of a timeframe spanning September 2019 to May 2020. This period of 8 months gives us a total of 3200 tweets, indicating that Bernie Sanders' Twitter account is very active. This timeline frequency also shows that within those 8 months, the account produces a range of 50 tweets to 150 tweets every week with an estimated average of 100 tweets per month.

Tweet Publish Locations

The Tweet Publish Locations graph gives us a general idea of where most of the account's tweets are being produced. There isn't a large spread of tweet locations for this account as most are being produced on a computer or similar device. This is standard for an account with high original content that is produced at a very up to date fashion rather than a third party application on a time scheduled bases.

Word Frequency Analysis

The word frequency analysis produces a nice even spread of frequently used words, with the keyword 'people' being the most used the most. This gives us major insight into the general target of the politicians' Twitter agenda. The general spread of unique words, negating 'people', involves words being used between 200 to 350 times which shows moderate focus in Twitter posts.

Wordcloud Hashtag Frequency Analysis

The wordcloud shows us that the account has a very small variety in hashtag presence. Combining this with 'demdebate' being the largest hashtag by far within the hashtag presence shows that Bernie Sanders' hashtag presence is very specified, ideal for a political Twitter account.

Wordcloud Retweet Account Frequency Analysis

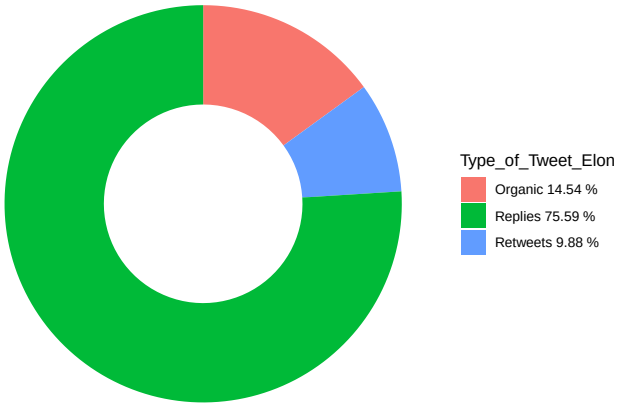
This wordcloud gives us a good insight into the accounts retweet structure. All the words associated in the wordcloud are evenly sized therefore not one account or a few accounts are the main contributors for retweeting the main account's tweets. Also, there is a large variety of different accounts retweeting the main accounts tweets which should be expected as a politician is generally speaking to people and acting as their voice. The large variety of retweeters highlights that the Twitter account resonates within a lot of followers.

Sentiment Analysis of Tweets

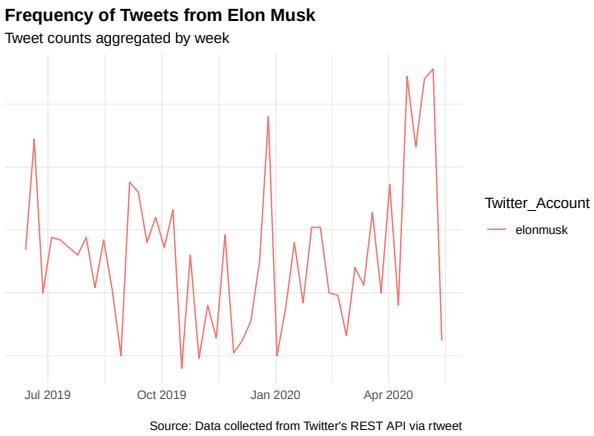
The sentiment analysis of tweets takes in all unique words from the collected tweets and gives each one a rating of all present categories ranging from 'Anger' to 'Trust'. Bernie Sanders Twitter account gives us a large collection of 'Positive' and 'Negative' words, as the account must promote Bernie as well as bring down others that he opposes. There is also a secondary standout between 'Fear' and 'Trust'. Typical of an account that speaks to the reasoning of people, as fear and trust are key factors that politicians use to sway their audience.

Elon Musk

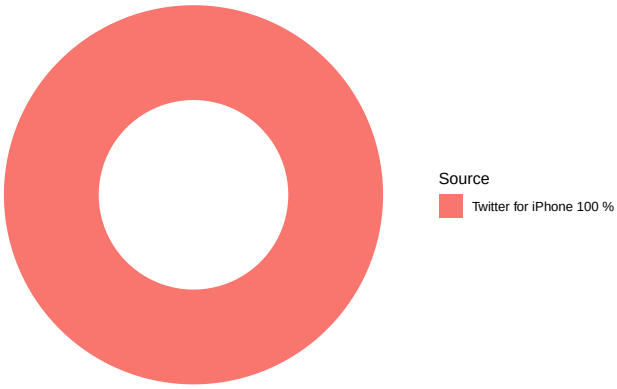
Ratio of Replies/Retweets/Organic Tweets



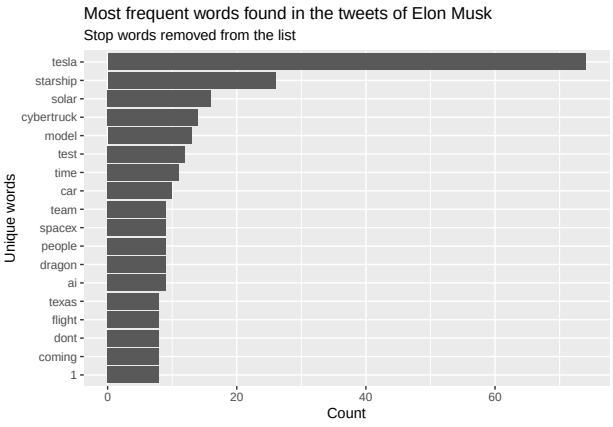
Tweet Timeline Frequency



Tweet Publish Locations



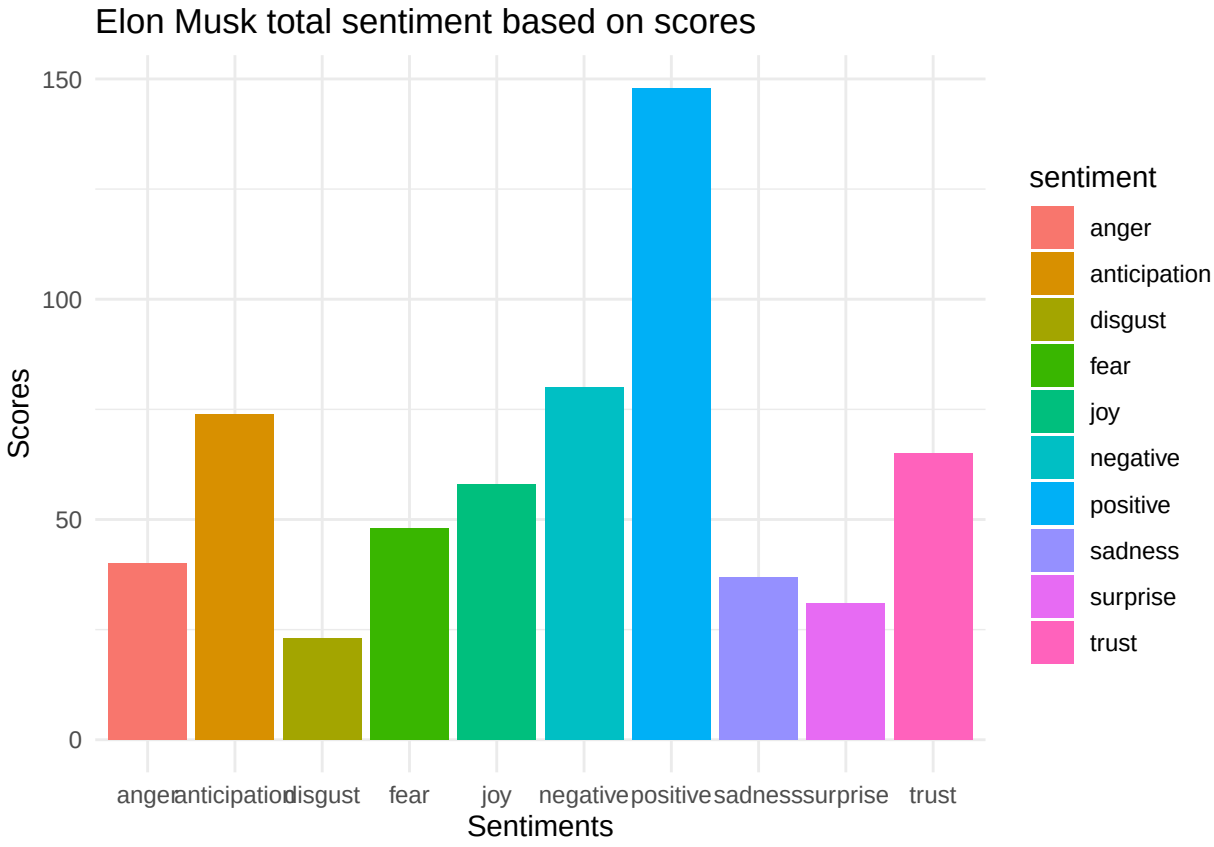
Word Frequency Analysis



Wordcloud Hashtag Frequency Analysis

Wordcloud Retweet Account Frequency Analysis

Sentiment Analysis of Tweets



Ratio of Replies/Retweets/Organic Tweets

Elon Musk's ratio of replies, retweets, and tweets reveals that most of his account content is produced through replies to various tweets. As an influencer, Elon comments on a wide variety of posts directed at him or general posts that he deems reply worthy. A small portion of his content is organic tweets and retweets as his Twitter account is one factor that affects his business, so branding is very important.

Tweet Timeline Frequency

The tweet timeline frequency of Elon Musk's tweets consists of a timeframe spanning June 2019 to May 2020. This period of 11 months gives us a total of 3200 tweets, indicating that Elon Musk's Twitter account is quite active yet not as fiercely active as Bernie Sanders. The timeline frequency shows that within those 11 months, the account produces a range of 0 tweets to 100 tweets every week with an estimated average of 75 tweets per week Elon Musk's tweet frequency is quite varying with large dips and spikes, indicating he tweets when he deems needed rather than tweeting on a schedule.

Tweet Publish Locations

Obtaining the relevant data for this graph was unsuccessful. Either Elon's Twitter account as a proxy on his output selection, resulting in the default output of 'iPhone app', or Elon Musk only tweets from his phone 100% of the time. No insight can be drawn.

Word Frequency Analysis

From this frequency analysis produced, we gain insight that Elon doesn't use too many keywords too often with the exception of 'Tesla' with over 70 mentions. This is not high comparing to Bernie Sander Twitter account which indicates that Elon Musk's Twitter account has a large spread of overall word repetition. This indicates that the discussion topics in which Elon posts or replies too, have a very broad range of points rather than a specified discussion as seen in Bernie Sanders' account.

Wordcloud Hashtag Frequency Analysis

Elon Musk only produces one main hashtag within his tweeting records, 'armageddon69'. This could be an indication that Elon has a single agenda in mind when presenting hashtags but as his account consists of typically replies and his word frequencies are quite low, he may not have a particular interest in hashtag use.

Wordcloud Retweet Account Frequency Analysis

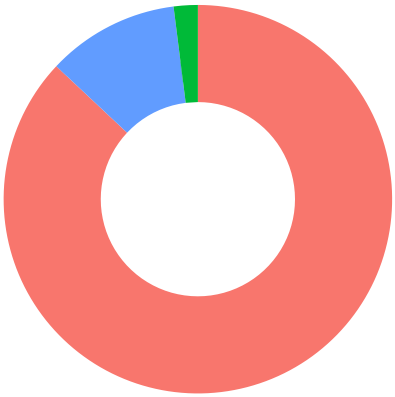
The Retweet Account wordcloud leans heavily on his company 'SpaceX'. This wordcloud is not as large and diverse at Bernie Sanders wordcloud, indicating that his tweets aren't highly retweetable, yet it still contains numerous accounts that retweet Elon's tweets. This reveals his Twitter recognition is present yet not a total connection to his followers' beliefs.

Sentiment Analysis of Tweets

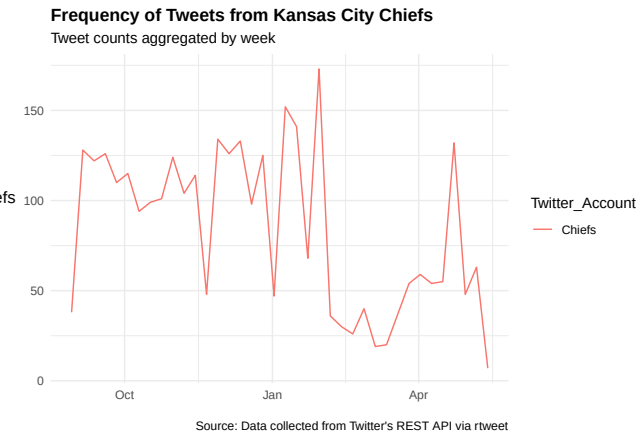
Elon Musk's Twitter account gives us a minimal spread with most wording being graded as positive. As the company's CEO, his tweets would have to gear in a 'Positive' sentiment as each tweet affects stocks. A secondary sentiment that is large is 'Anticipation'. This can be accounted for as this Twitter account is generally used as a promotion and advertisement of the companies Elon controls.

NFL Chiefs

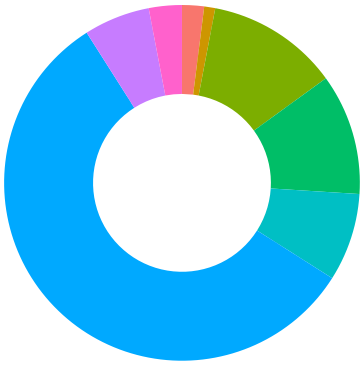
Ratio of Replies/Retweets/Organic Tweets



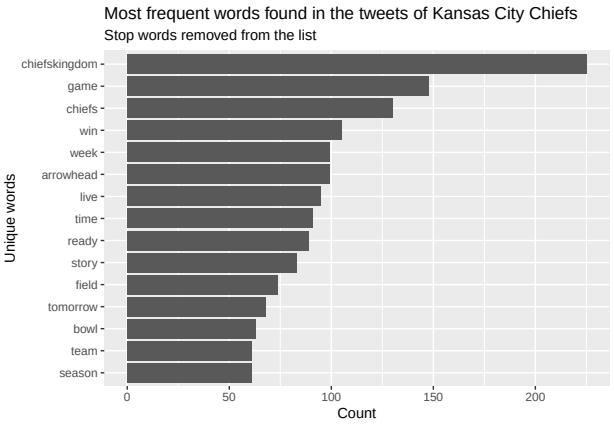
Tweet Timeline Frequency



Tweet Publish Locations



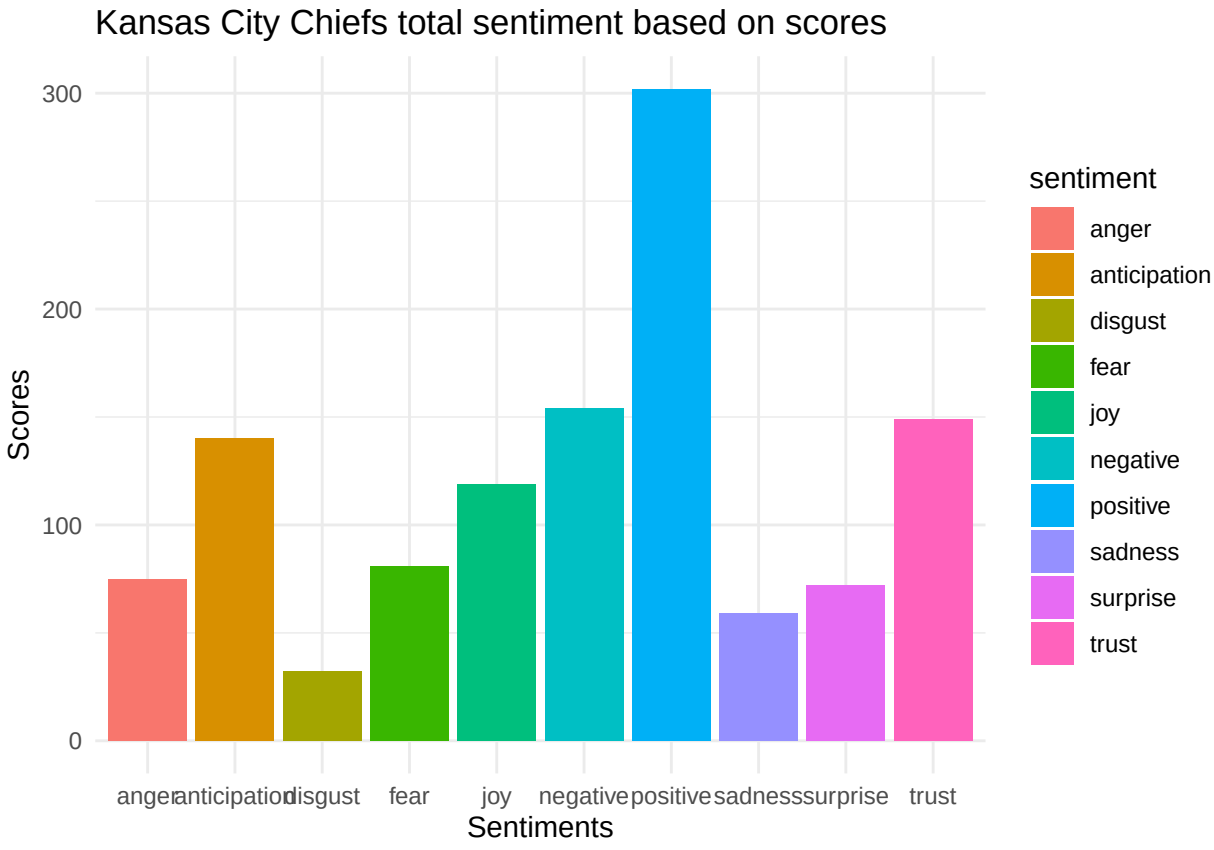
Word Frequency Analysis



Wordcloud Hashtag Frequency Analysis

Wordcloud Retweet Account Frequency Analysis

Sentiment Analysis of Tweets



Ratio of Replies/Retweets/Organic Tweets

The Kansas City Chiefs NFL sporting team ratio of replies, retweets, and tweets reveal that a major majority of tweets posted are organically posted tweets. As a sporting team, organic content is very important. Very little retweets and replies are produced as the team's Twitter is heavily associated with the face of the brand. This is an important insight when understanding what makes a sporting accounts' Twitter a successful one.

Tweet Timeline Frequency

The tweet timeline frequency of Chiefs' tweets spans from September 2019 to May 2020. This period of 8 months gives us a total of 3200 tweets, indicating that Chiefs' is a incredibly highly active Twitter account with an even larger tweet production then that to Bernie Sanders' account. The timeline frequency shows that within those 8 months, the account produces a range of 50 tweets to 200 tweets every week with an estimated average of 100 tweets per week. The Chiefs' Tweet frequency is quite stable with high tweet production consisting of only three drastic dips during peak holidays within this timeframe, then a complete upward spike in February as the team went on to win the Super Bowl of 2020.

Tweet Publish Locations

The tweet publish locations graph for the Chiefs indicate a wide variety of forms for submitting tweets, indicating that not one sole tweeter is tweeting away on the account, rather a collection or team of people posting carefully planned tweets at one time. A key component in the spread is 'Twitter Media Studio'. This application is a very used and useful tool for companies that run a Twitter Account. Twitter Media Studio allows users to manage, measure, and monetize videos, images, and GIFs on Twitter. A visual presence for a sporting team is important as a limit of 280 characters can be limiting for certain promotions and advertisements.

Word Frequency Analysis

Chiefs' frequency analysis graph outlines that producers of the tweets for the Chiefs use keywords highly repetitively. The graph indicates somewhat of an even spread, 'chiefskingdom' being the front runner with over 200 repetitions, and most words being repeated from 80 to 150 times. The even yet high spread of keyword frequencies indicates the Chiefs' Twitter account has a specific agenda when communicating with its fans.

Wordcloud Hashtag Frequency Analysis

Hashtag frequency wordcloud for Chiefs contains multiple hashtags used within the account's tweets. The main hashtag is also 'chiefskingdom', seen in the word frequency analysis, indicating that a higher number of tweets contain this hashtag. This larger variation of one main hashtag amongst several others indicates that the Chiefs' Twitter account also has one main hashtag presence they employ in creating a name for their presence on Twitter, as seen with the Bernie Sanders' account. A large number of other hashtags specifies that the Chiefs Twitter account is also very active with many different trends allowing for a large audience reach.

Wordcloud Retweet Account Frequency Analysis

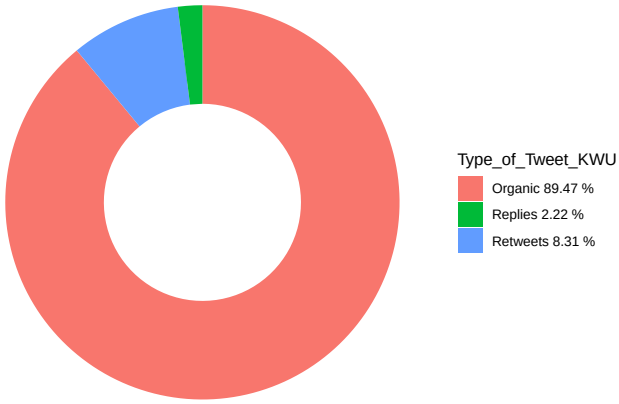
The Retweet Account wordcloud contains multiple major retweeters. This indicates that retweets are a larger contributor to the Twitter account than its hashtag presence, as the Chiefs Twitter account is a very retweeted account. This sign combined with its heavy organic tweet makeup indicates that the Twitter account for the Chief's is very fan-based and for the people. This would be the case for a major sporting team as its Twitter platform is an advertisement and promotional voice for the team and its fans.

Sentiment Analysis of Tweets

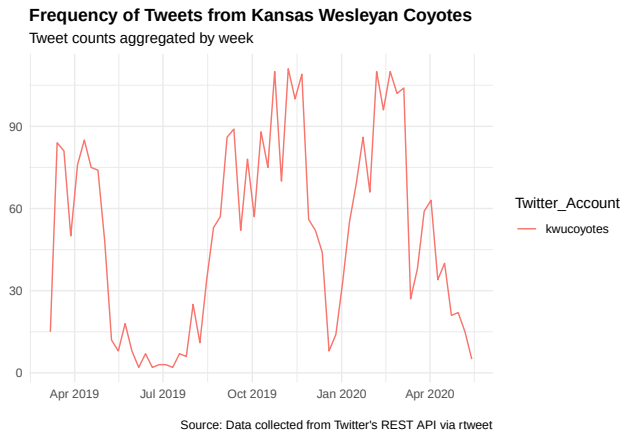
Chiefs account gives us an even spread with most wording being graded as positive. Secondary sentiments are 'Anticipation' and 'Trust'. This is quite a key component of sporting team notion as a sporting account not only speaks on the current events of the sporting team but as well as the future of the program as to where it's heading and what plans it has for the future. This is important because revealing to the fans that the future is good and bright is what is need to ensure fans remain fans of that particular organization.

Kansas Wesleyan University Coyotes

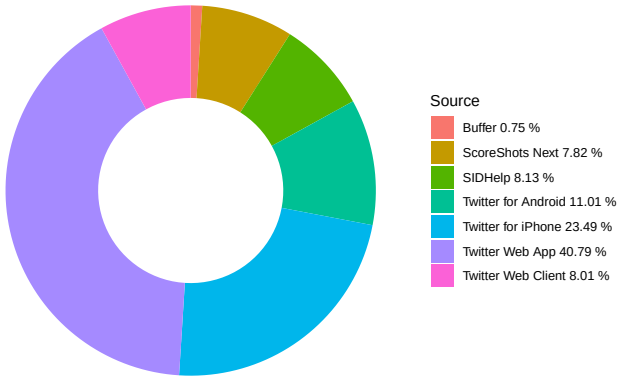
Ratio of Replies/Retweets/Organic Tweets



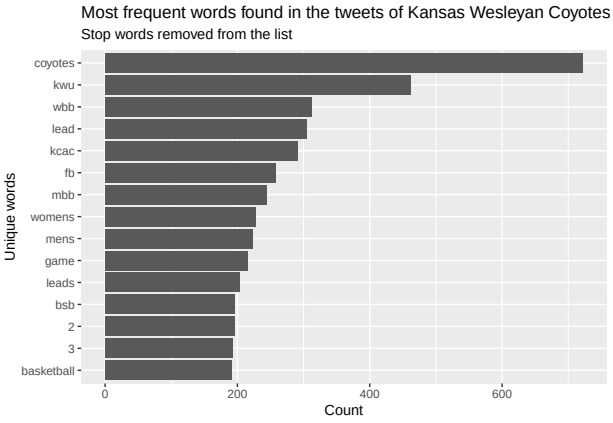
Tweet Timeline Frequency



Tweet Publish Locations



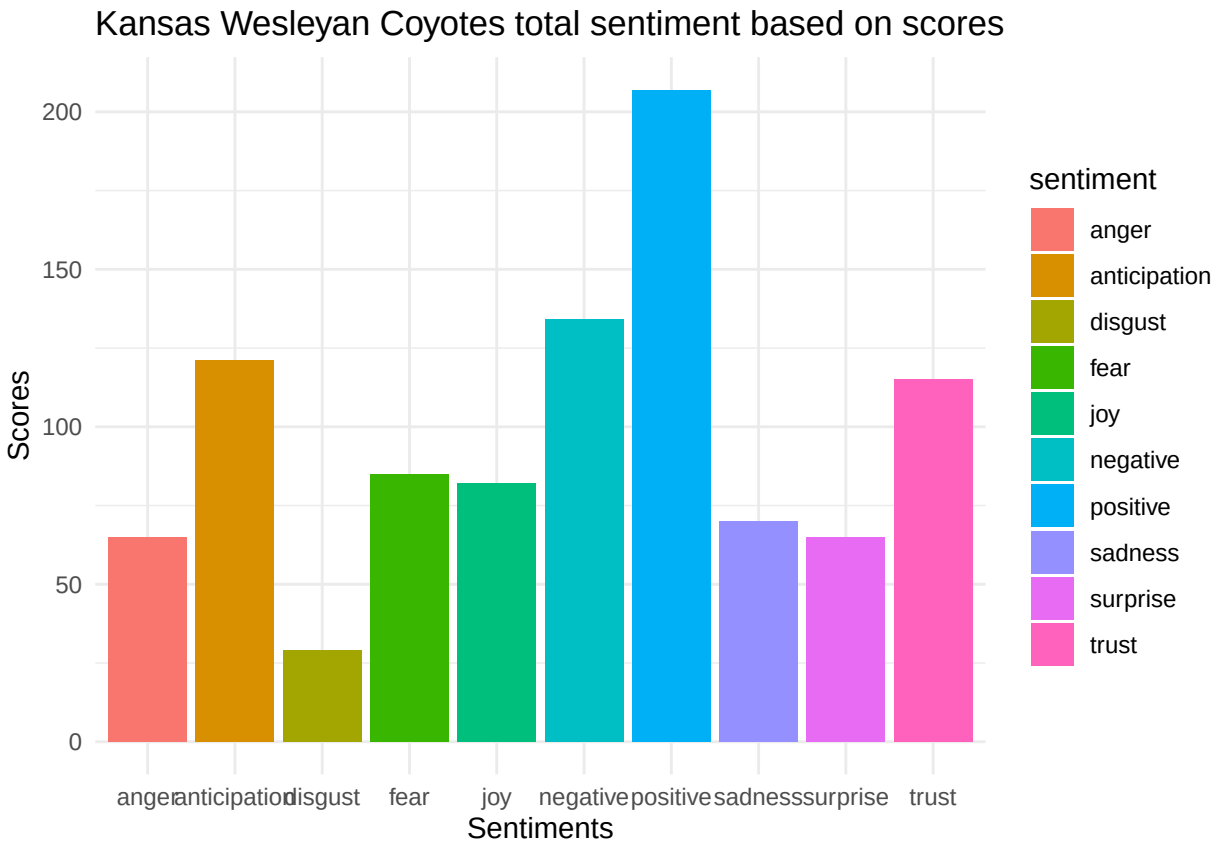
Word Frequency Analysis



Wordcloud Hashtag Frequency Analysis

Wordcloud Retweet Account Frequency Analysis

Sentiment Analysis of Tweets



Ratio of Replies/Retweets/Organic Tweets

The Kansas Wesleyan University Coyotes ratio of replies, retweets, and tweets reveal that a major majority of tweets posted are organically posted tweets. As a sporting team, organic content is very important. Very little retweets and a bare supply of replies are produced as the team's Twitter is heavily associated with the face of the brand. This is in relatable connection to the Chiefs' sporting account.

Tweet Timeline Frequency

The tweet timeline frequency of KWU tweets spans across March 2019 to May 2020. This period of 14 months gives us a total of 3200 tweets, indicating that the KWU Coyotes is a moderately active Twitter account but nothing to the scale of major Twitter accounts analyzed above. The timeline frequency shows that within those 14 months, the account produces a range of 0 tweets to 120 tweets every week with an estimated average of 55 tweets per week. The KWU Coyote Twitter produces high tweet count during college semesters and low tweet counts during breaks. Similar to the Chiefs' account for dips during holidays and a high tweet production during in-season play.

Tweet Publish Locations

The locations graph for the KWU Coyotes indicates a minimal variety of forms for submitting tweets, indicating that a singular person or small group of people man the account. The College attempted to use a third-party app 'Scoreshots Next' but resulted in a very small number of tweets contributed by this application.

Word Frequency Analysis

KWU Coyotes frequency analysis graph outlines that producers of the tweets for the KWU Coyotes use keywords highly repetitively. A major keyword used is the team's mascot 'coyotes' as it is repeated over 700 times within all 3200 tweets. The next highest is 'KWU' repeating close to 500 times. After those two keywords, the graph indicates an even spread of words ranging from 200 to 300 hundred repetitions, which is a high count for average word repetition. This indicates that KWU is very specific with its wording choice in its tweet productions. This was done even more so than compared to Chiefs' Twitter. The large repetition of words might be a hindrance, as being so specif in the accounts agenda and repetition may lead to detached and bored followers.

Wordcloud Hashtag Frequency Analysis

Hashtag frequency wordcloud for KWU Coyotes is very minimal. One major hashtag presence is the hashtag 'rollyotes', being the main contributor to the wordcloud. Other than that only five other hashtags are present within all 3200 tweets with little repetition. This is ideal when pushing for a singular hashtag presence especially within an account that is leading a fan base, as seen with Chiefs and Bernie Sanders.

Wordcloud Retweet Account Frequency Analysis

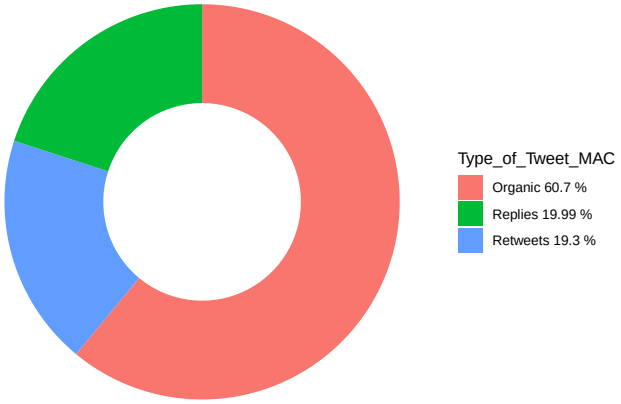
The Retweet Account wordcloud contains a decent amount for a small college that KWU is. This wordcloud is still smaller than that compared to Bernie Sanders or Chief's respective Twitter accounts. Main retweet accounts are comprised of small Twitter accounts owned or ran by the college indicating that their Twitter presence does resognate with certain followers but may not contain enough connection wihtin its followers to deem a retweet.

Sentiment Analysis of Tweets

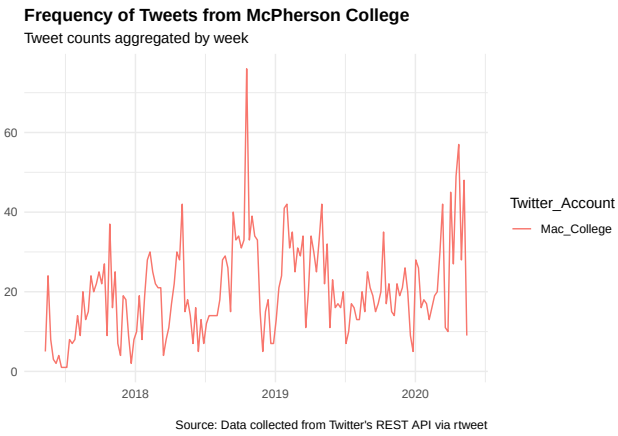
Sentiment analysis of KWU Coyotes Twitter accounts is similar to that of the chiefs sporting account. A high presence of 'Positive' words with secondary sentiments being 'Anticipation', 'Negative', and 'Trust'. Typical of an account that spreads positivity within the account at present while also speculating positively about prospects.

McPherson College Bulldogs

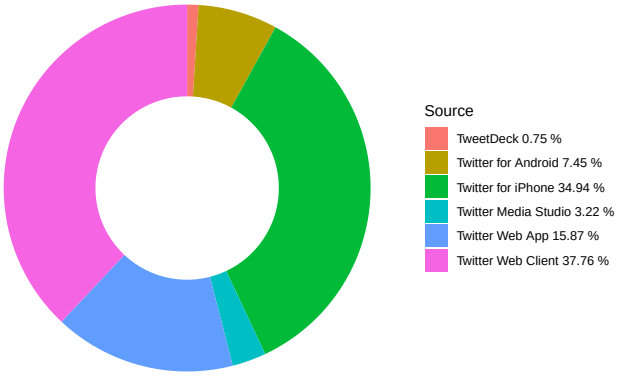
Ratio of Replies/Retweets/Organic Tweets



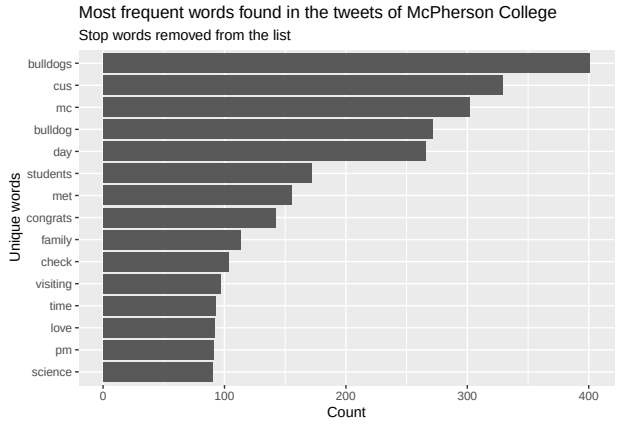
Tweet Timeline Frequency



Tweet Publish Locations



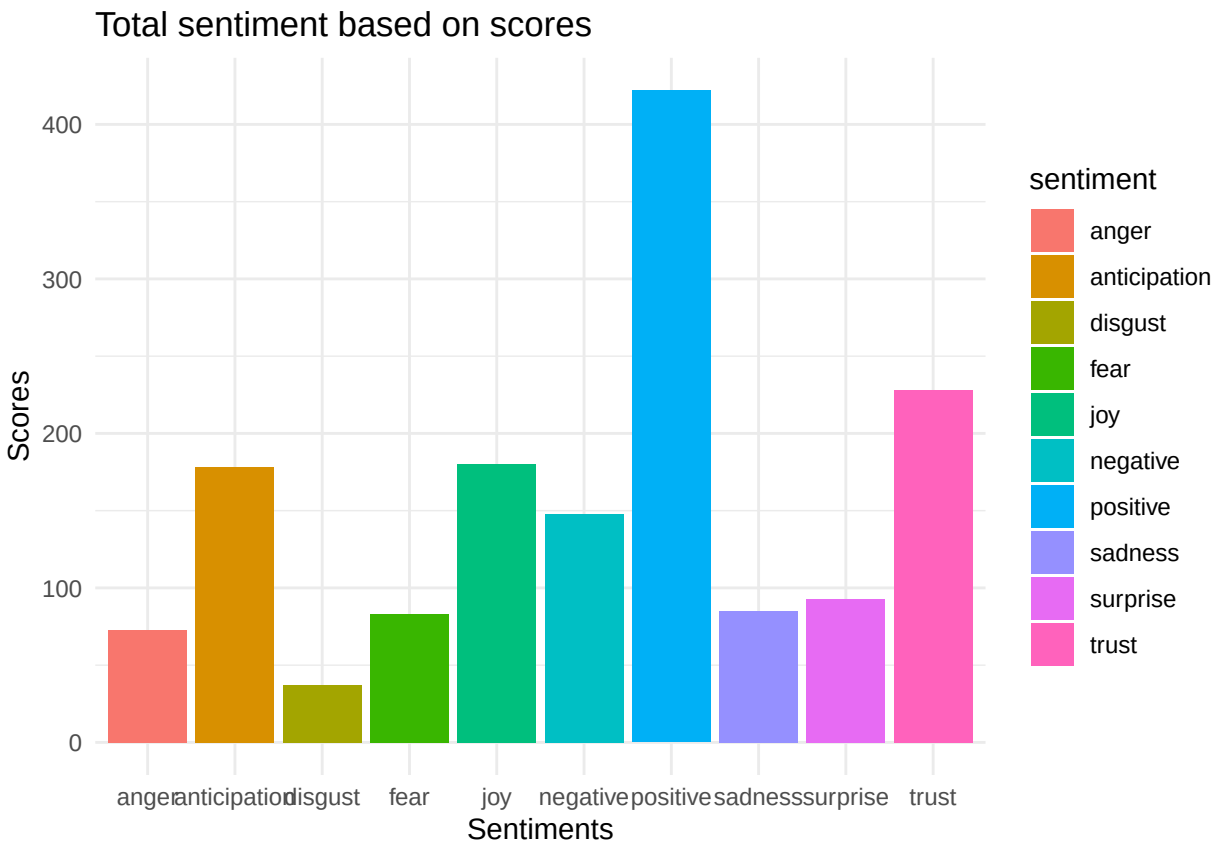
Word Frequency Analysis



Wordcloud Hashtag Frequency Analysis

Wordcloud Retweet Account Frequency Analysis

Sentiment Analysis of Tweets



Ratio of Replies/Retweets/Organic Tweets

The McPherson College Bulldogs ratio of replies, retweets, and tweets reveal a moderate makeup of tweets posted are organically posted tweets yet not as large as others compared above. As a sporting team, organic content is very important and as seen in larger colleges and sporting teams' organic content should be by far the major make up their sporting Twitter account.

Tweet Timeline Frequency

The tweet timeline frequency of McPherson Bulldogs' tweets spans across April 2017 to May 2020. This period of 3 years gives us a total of 3200 tweets, indicating that the McPherson Bulldogs is a very dormant Twitter account. The timeline frequency shows that within those 3 years, the account produces a range of 0 tweets to 50 tweets every week with an estimated average of 20 tweets per week. The McPherson Bulldogs' Twitter account has a very minimal tweet production.

Tweet Publish Locations

The locations graph for the McPherson Bulldogs indicates a minimal variety of forms for submitting tweets, indicating that a singular person or small group of people man the account. 'Twitter Media Studio' is present in the graph but only holds a small sliver of total publish location percentage.

Word Frequency Analysis

McPherson Bulldogs' Twitter account has a good overall word frequency spread. Most of the major words are over 250 repetitions. Larger repetition of 'Bulldogs' and 'Bulldog' is a good sign the accounts agenda is in line with promoting the college's sporting brand.

Wordcloud Hashtag Frequency Analysis

Hashtag frequency wordcloud for McPherson Bulldogs is very minimal, similar to that of the KWU Coyotes. One major hashtag presence is the hashtag 'bulldogpride', being the main contributor to the wordcloud. Other than that, only seven other hashtags are present within all 3200 tweets with little repetition. The hashtag presence for McPherson College Bulldogs' Twitter account does align with other accounts and their general identification with a singular hashtag.

Wordcloud Retweet Account Frequency Analysis

The Retweet Account wordcloud of McPherson Bulldogs is large and numerous. Many retweeters range from within the college's own accounts to personal accounts of people linked to the school. This is an indication of the dedication people have toward McPherson college and their intrigue to Twitter posts.

Sentiment Analysis of Tweets

Sentiment analysis of McPherson College Bulldogs' Twitter account reveals that the Twitter account has a very high positive word count. The sentiment analysis also brings forward the category of 'joy' which is not present in other accounts. This is important because it sets McPherson College aside in a positive manner.

Conclusions

Through social media analysis, more specifically Twitter data analysis, one can find many factors that contribute to a Twitter account. It is with these factors that analyzing and comparing accounts, that one can draw conclusions about the effectiveness and reach of specific accounts and, with mathematical data, outline where and how an account can be improved.

McPherson College Bulldog's Twitter account has been examined through careful analysis. With simple key steps, turning the account into a Twitter account with a large Twitter presence should be simple.

- Tweet Production: We suggest a focus on more original content, promoting and advertising the College's sporting teams. This incorporates an increase in the frequency of tweets, ranging around 55 tweets per week, and special consideration when retweeting and replying to other users.
- Third-party applications: McPherson College Bulldogs' Twitter account could implement stronger use of Twitter Media Studio as well as potential third-party applications for the production of its tweets. This will ensure production of on a wide variety of content as well as the analytical management of video engagement and monetizing of the account to increase the college's sporting revenue. The use of third-party applications will also allow for tweet scheduling as future planning and 'hype' of sporting programs are seen to be an important part of a sporting Twitter account.
- Hashtag presence: McPherson College Bulldogs' Twitter account could potentially update its Hashtag presence within its Twitter account. This could involve establishing a new main hashtag to pursue heavily while adopting smaller hashtags to broaden the Twitter presence of McPherson College.
- Tweet Wording Choice: The Bulldogs' Twitter account should continue its positive wording presence on Twitter while also encompassing joyful words to set apart McPherson College from the rest. The Twitter account can also start to incorporate 'anticipation' wording choice within the tweets regarding the outlook of the sporting program for a general enticement of its followers.

Bibliography

Obtaining and using access tokens
- https://rtweet.info/articles/auth.html

A Guide to Mining and Analysing Tweets with R
- https://towardsdatascience.com/a-guide-to-mining-and-analysing-tweets-with-r-2f56818fdd16

R Markdown from R Studio
- https://rmarkdown.rstudio.com/lesson-1.html

R Markdown Cookbook
- https://bookdown.org/yihui/rmarkdown-cookbook/