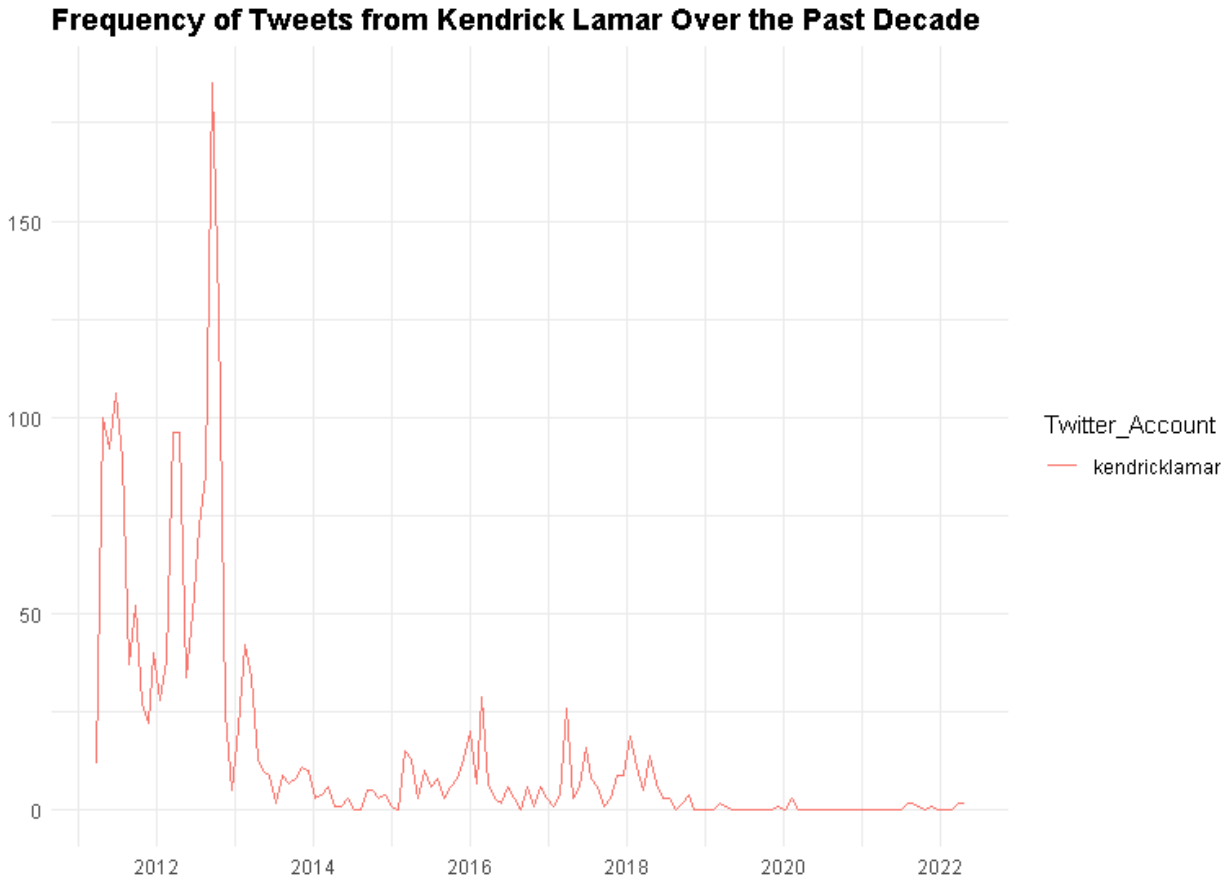
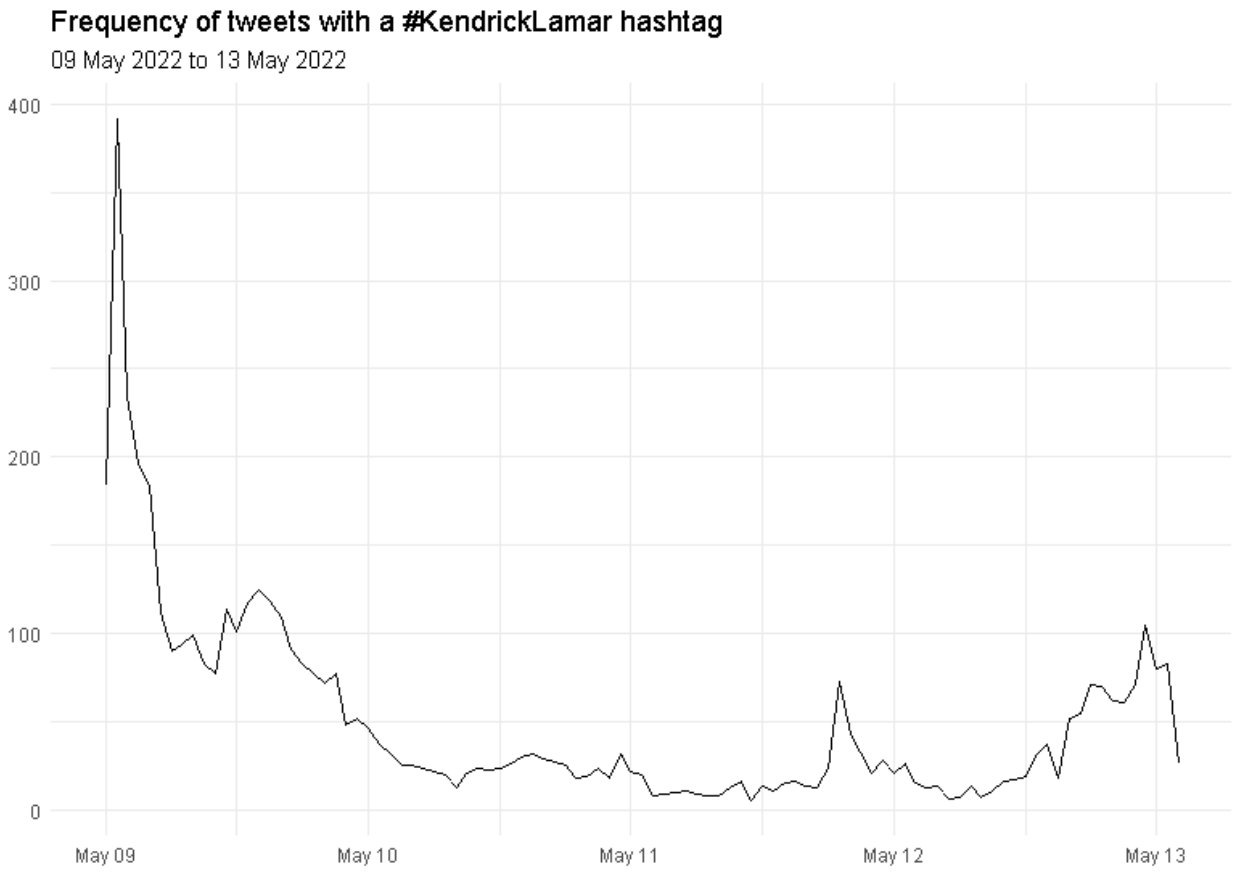
Giorgi Samushia
CMPT363
Public Perception of Kendrick Lamar's Return to Music

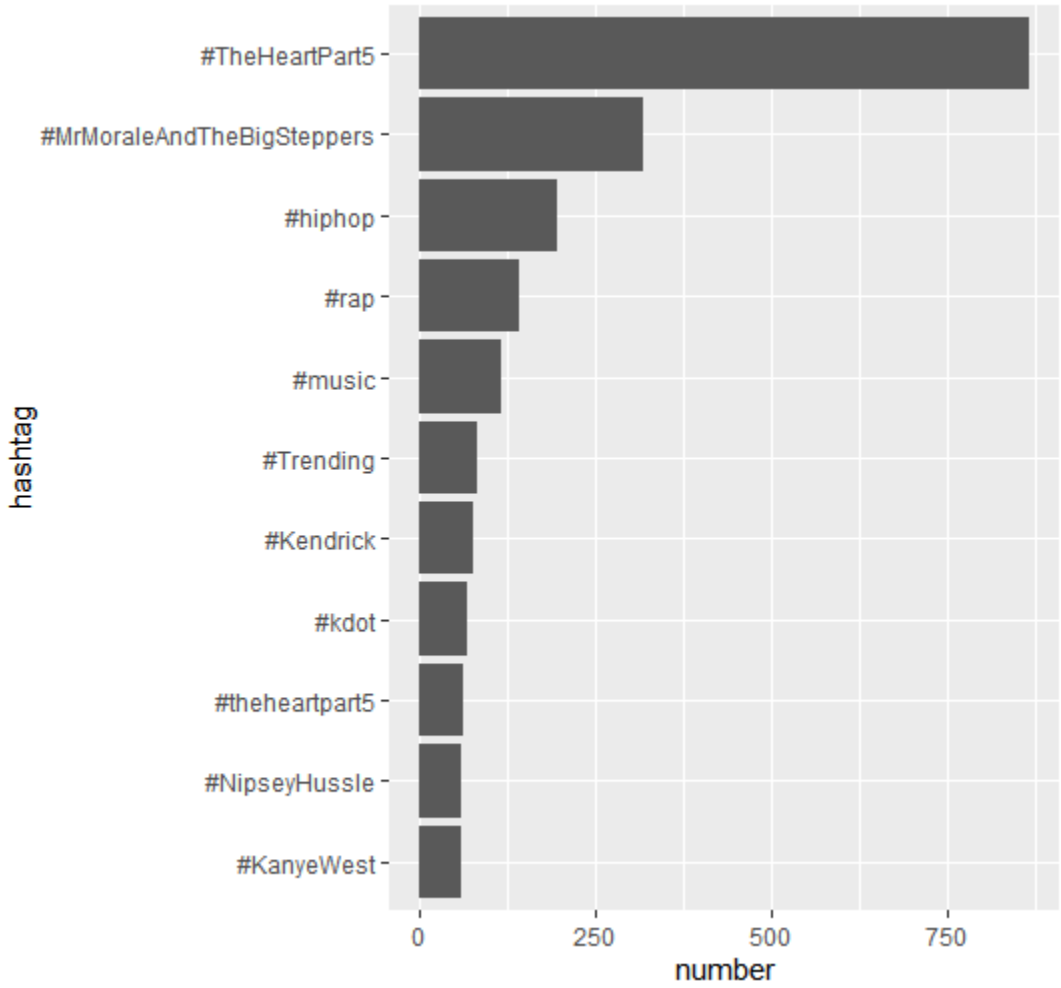
Kendrick Lamar, one of the most revered modern hip - hop artists had been absent from the spotlight for nearly half a decade. That is, until last month, when he came back with an announcement of his new album.



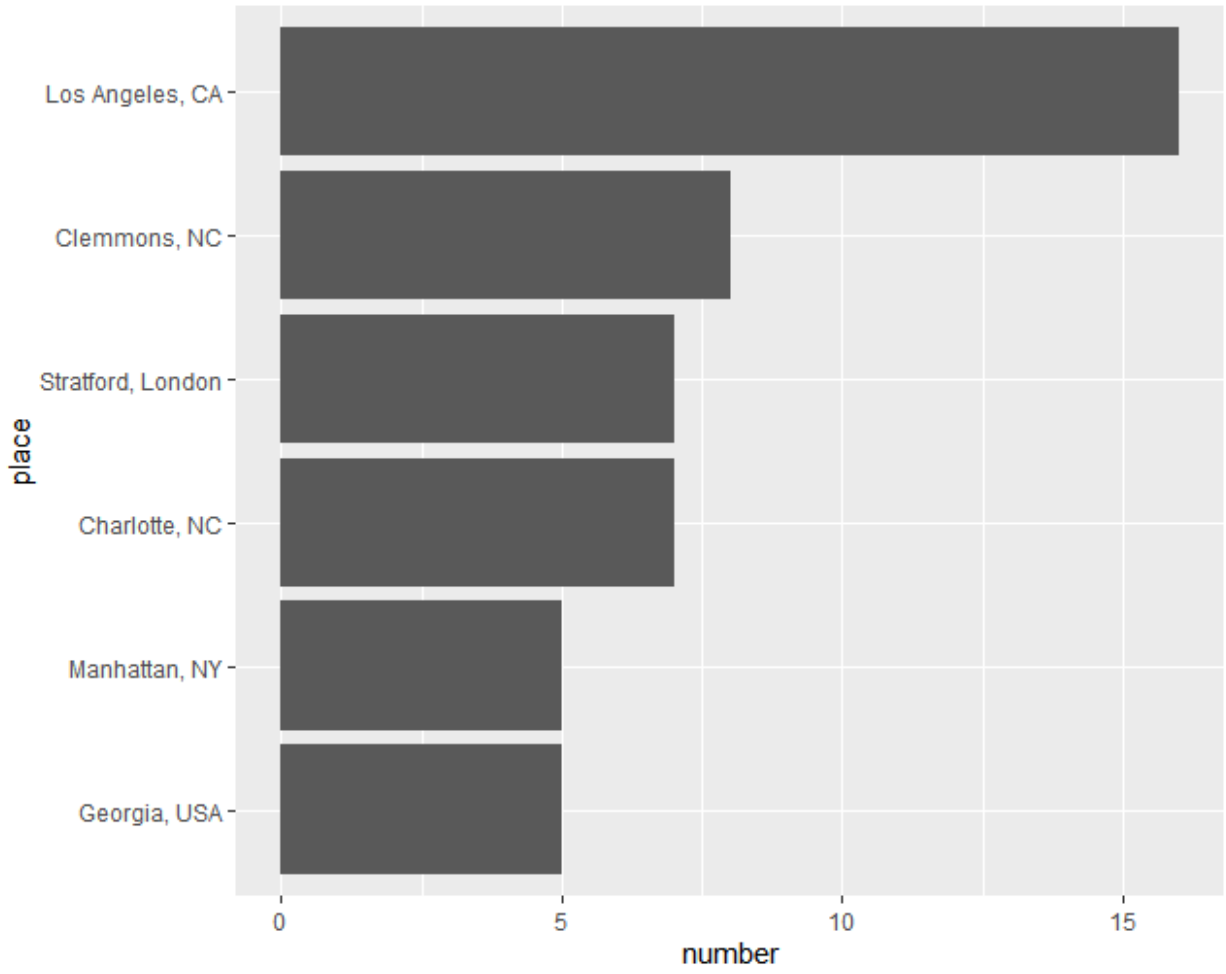
This is a graph showing the frequency of tweets from the artist over the past decade.
A significant rise in the graph can be seen in 2011 - 2012, the first year marking his first studio album, and the second one denoting the release of his critically acclaimed modern classic - "good kid, m.A.A.d. city".
Next, there are increases in the years 2015-2016 and 2017-2018, these years also marking releases of his works, or works that he was heavily involved in.
The frequency drops drastically after 2018, nearly flatlining, with a small increase being spotted at the right end, which corresponds to now.



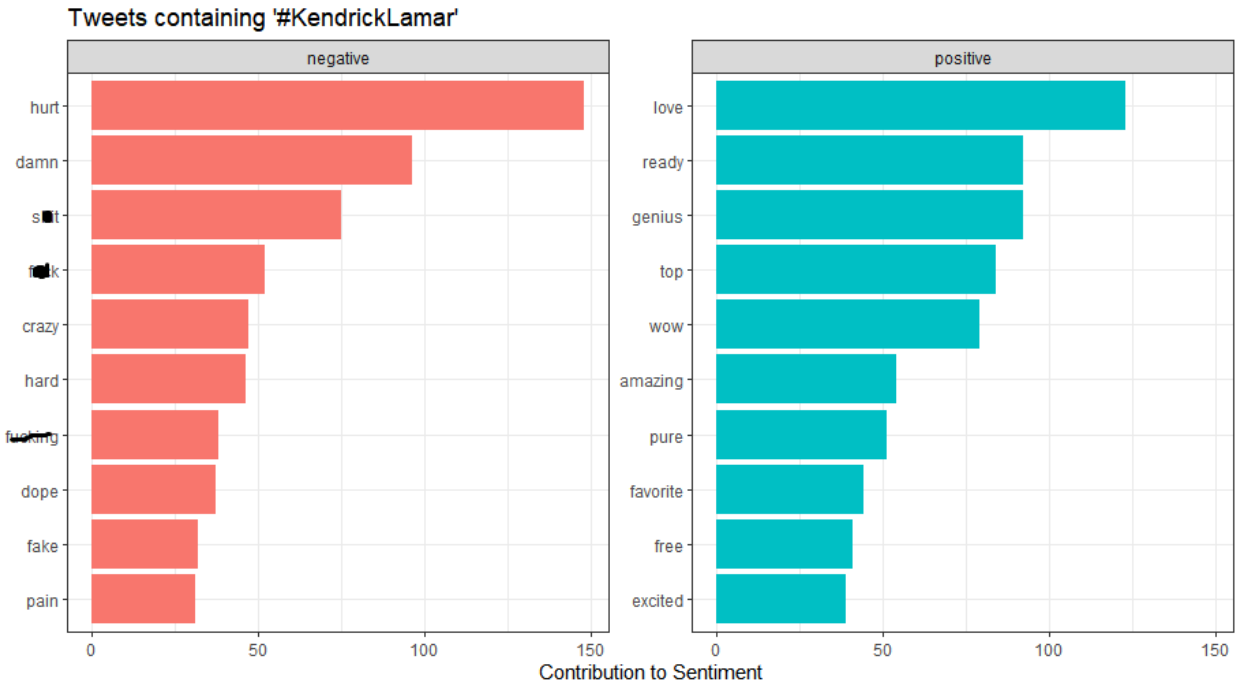
Here is another frequency plot. This time, this graph showcases the use of the hashtag #KendrickLamar over the past few days. The plot once again achieves taller heights at the beginning of the x-axis. This is, because, in the latter half of May 8th the artist released his new single, which became the center of attention for that day. Additionally, we can again see how the graph starts to increase at the end, as the album is projected to be released on May 13th. (The graph dropping is not relevant in this case, as there are no tweets during May 13th while this document is being made.)



This barplot shows us other hashtags that were used along with #KendrickLamar. The number one hashtag is the name of the single he released May 8th, and the number two is the name of the upcoming album. It is to be noted that the name of the single is seen two times on this plot, and we can also see names like Nipsey Hussle and Kanye West, which are tightly intertwined with the subject matter of said single. This can lead us to believe that most of the conversation right now is going around the single and not the album, even though the album will be released in hours. Observations could be made about this graph after tomorrow.



This graph showcases the top 5 places tweeted from. The sample size is really small, as a lot of locations were not specified in tweets. Furthermore, running bigger amounts of tweets on my laptop took more than a substantial amount of time. Still, though, the graph makes sense. Kendrick's audience is mostly American, and he is especially recognized in Los Angeles, California, as it is his home.



This is a sentiment analysis of the tweets containing the hashtag #KendrickLamar.
The analysis uses a method that was utilized in phase 03 of this project.
As seen here, the positive words contribute more to the sentiment, and even if there are a lot of negatives on the plot, one can deduce that many of the words mentioned can be used in a positive context. (An example of this being dope, crazy, etc.)

The Code:

```
install.packages("rtweet")
install.packages("dplyr")
install.packages("tidyr")
install.packages("tidyverse")
install.packages("tidytext")
install.packages("textdata")
install.packages("forestmangr")

library(rtweet)
library(dplyr)
library(tidyr)
library(tidyverse)
library(tidytext)
library(textdata)
library(forestmangr)
library(ggplot2)

rm(list=ls())

#retrieving the tweets that use the hashtag

tweets <- search_tweets(q = "#kendricklamar",
                        n = 5000,
                        include_rts = FALSE,
                        `-filter` = "replies",
                        lang = "en")



#Getting tweets from the timeline of kendrick lamar and graphing them over the years

k_tweets <- get_timeline("@kendricklamar", n= 2000)

colnames(k_tweets)[colnames(k_tweets)=="screen_name"] <- "Twitter_Account"

#Graphing the frequency of tweets over years

ts_plot(group_by(k_tweets, Twitter_Account), "month") +
```

```
  theme_minimal() +
  theme(plot.title = element_text(face = "bold")) +
  labs(
    x = NULL, y = NULL,
    title = "Frequency of Tweets from Kendrick Lamar Over the Past Decade"
    )


#Graphing the frequency of tweets including #KendrickLamar over the past few days

ts_plot(tweets, by = "hours") +
  theme_minimal() +
  theme(plot.title = element_text(face = "bold")) +
  labs(x = NULL, y = NULL,
       title = "Frequency of tweets with a #KendrickLamar hashtag",
       subtitle = paste0(format(min(tweets$created_at), "%d %B %Y"), " to ",
format(max(tweets$created_at),"%d %B %Y")),) +
  theme_minimal()


#Plotting the most used hashtags along with #KendrickLamar

hash <- tweets %>%
  unnest_tokens(hashtag, text, "tweets", to_lower = FALSE) %>%
  filter(str_detect(hashtag, "^#"),
         hashtag != "#KendrickLamar",
         hashtag != "#kendricklamar") %>%
  count(hashtag, sort = TRUE) %>%
  top_n(10)

hash <- hash %>%
  rename(
    number = n
  )

dfh <- data.frame(hash)

dfh$hashtag <- factor(dfh$hashtag, levels = dfh$hashtag[order(dfh$number)])

ggplot(dfh, aes(x = number, y = hashtag)) +
```

```
  geom_bar(stat = "identity")



#Plotting the cities most tweeted from


locations <- tweets %>%
  filter(!is.na(place_full_name)) %>%
  count(place_full_name, sort = TRUE) %>%
  top_n(5)

locations <- locations %>%
  rename(
    number = n,
    place = place_full_name
  )

dfl <- data.frame(locations)

ggplot(dfl)

dfl$place <- factor(dfl$place, levels = dfl$place[order(dfl$number)])

ggplot(dfl, aes(x = number, y = place)) +
  geom_bar(stat = "identity")


#Sentiment analysis of tweets containing the hashtag


tweets.Celeb <- tweets %>% select(screen_name, text)

tweets.Celeb$stripped_text <- gsub("http\\S+","",tweets.Celeb$text)
tweets.Celeb_stem <- tweets.Celeb %>%
  select(stripped_text) %>%
  unnest_tokens(word, stripped_text)
cleaned_tweets.Celeb <- tweets.Celeb_stem %>%
  anti_join(stop_words)
```

```
bing_celeb = cleaned_tweets.Celeb %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  ungroup()
bing_celeb %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(title = "Tweets containing '#KendrickLamar' ",
       y = "Contribution to Sentiment",
       x = NULL) +
  coord_flip() + theme_bw()
```