

Twitter Analytics

Ayo Lewis, Brandon Kwon, and Clara Seo

5/10/2021

Rise of Social Media during the Pandemic

A pandemic, the death of George Floyd, the rise of hate crimes, and a presidential election called Americans to action. These events aggregated into a surging movement of activism that called for social change. The pandemic added a level of complexity (i.e., lockdowns and movement restrictions) that ultimately resulted in many of these calls to action to take place online.

The past 15 months proved that social media can be an agent of change beyond advertisements and awareness. We all witnessed the riveting rise of digital activism nationwide during a pandemic. Users feel that they have a safe space to interact, feel connected, be comforted, distract themselves, stay engaged, and find inspiration - without any risk of contagion. People are staying at home without much to do except to watch Netflix and go on Twitter.

Activism and protest thrived in digital spaces

Harness social media to demand justice.

Why Twitter

Twitter allows for greater visibility to social movements. The entry costs are low (free to make an account).

Twitter is among the top 3 social networking apps in the US with 353.1 million monthly and 192 million daily active users currently (February 2021). Twitter generates vast amounts of user-generated language data, making it the perfect platform to conduct textual/sentiment analysis. Twitter has considerable advantages over other social media platforms (e.g., Instagram, Facebook, TikTok) for analysis:

- 1) Every tweet is limited to a maximum of 280 characters, which provides us with a relatively homogeneous corpora.
- 2) There are millions of active users that post tweets and retweet/favorite/comment on organic tweets, which generates an accessible and large data sample.
- 3) The tweets are publicly available, accessible, and retrievable via APIs.

```
covid_exp_tweets <- read_csv("data/twitter_data/tweets_covid_exp.csv")
```

```
##
## -- Column specification -----
## cols(
##   word = col_character()
## )
```

```
asian_exp_tweets <- read_csv("data/twitter_data/tweets_asian_exp.csv")
```

```
##
## -- Column specification -----
## cols(
##   word = col_character()
```

```
## )
black_exp_tweets <- read_csv("data/twitter_data/tweets_black_exp.csv")
```

```
##
## -- Column specification -----
## cols(
##   word = col_character()
## )
```

```
# gives you a bar chart of the most frequent words found in the tweets
covid_bar <- covid_exp_tweets %>%
  count(word, sort = TRUE) %>%
  top_n(30) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col(fill = "blue") +
  xlab(NULL) +
  coord_flip() +
  labs(y = "\n Count", x = "Unique words", title = "Pandemic") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60, vjust = 0.5, hjust=1))
```

```
## Warning: `count_()` was deprecated in dplyr 0.7.0.
## Please use `count()` instead.
## See vignette('programming') for more help
## Selecting by n
```

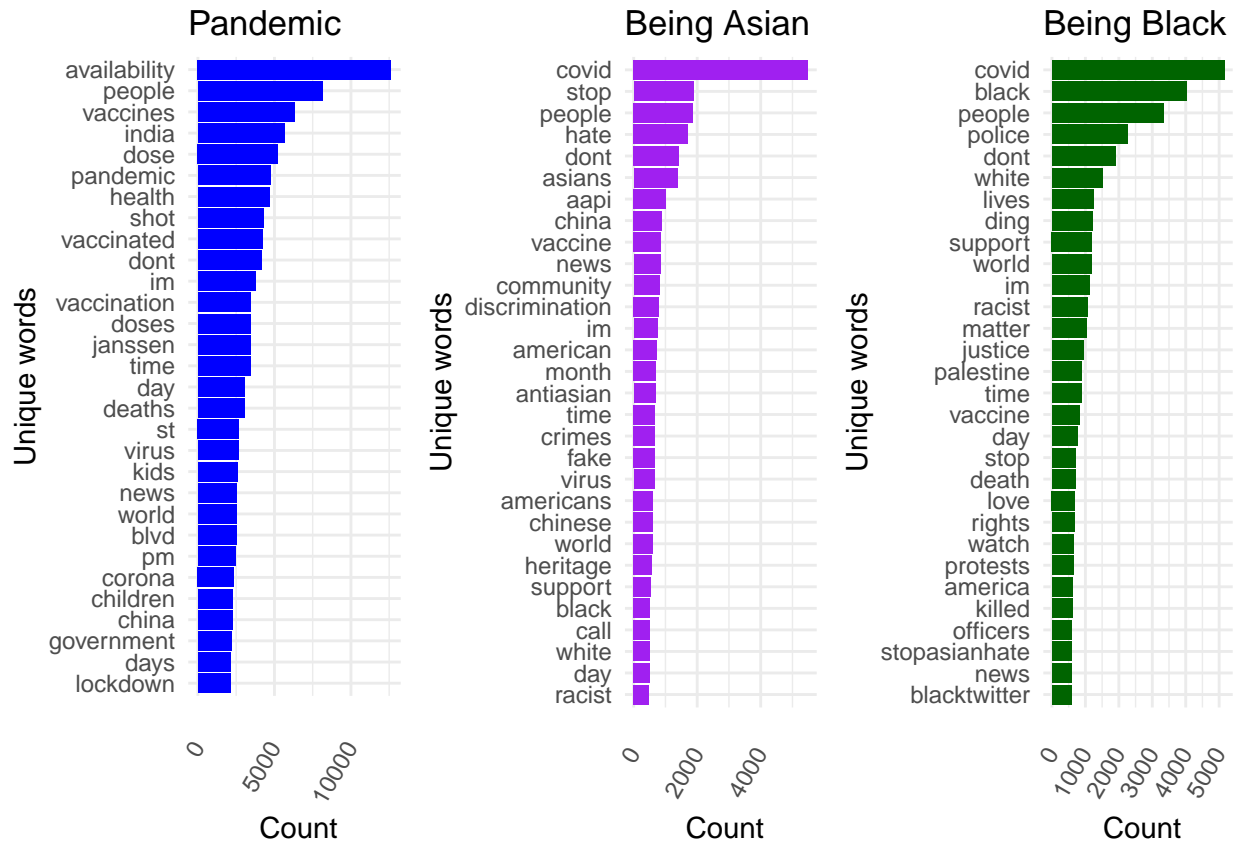
```
asian_bar <- asian_exp_tweets %>%
  count(word, sort = TRUE) %>%
  top_n(30) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col(fill = "purple") +
  xlab(NULL) +
  coord_flip() +
  labs(y = "\n Count", x = "Unique words", title = "Being Asian") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60, vjust = 0.5, hjust=1))
```

```
## Selecting by n
```

```
black_bar <- black_exp_tweets %>%
  count(word, sort = TRUE) %>%
  top_n(30) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col(fill = "darkgreen") +
  xlab(NULL) +
  coord_flip() +
  labs(y = "\n Count", x = "Unique words", title = "Being Black") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60, vjust = 0.5, hjust=1))
```

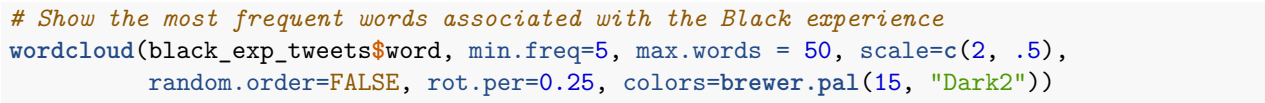
```
## Selecting by n
```

```
gridExtra::grid.arrange(covid_bar, asian_bar, black_bar, ncol=3)
```



A few interesting observations:

- The word “availability” is the most frequently used word in the pandemic category. Given the Twitter data was pulled from first and second week of May 2021, it makes sense since the country is going through mass vaccine distribution efforts.
- Words about the criminal justice system, “police”, “justice”, “officers” appeared in top 30 of the “Being Black” category. This tells us how the black experience during the pandemic has not only been about combating health issues but also fighting against institutionalized racism, police brutality, medical bias, and so much more. The current civil unrest is deeply connected to the racial disparities that have been brutally exposed by the coronavirus crisis. This textual data finding supports American Psychological Association (APA) President Sandra L. Shullman’s comment that “we are living in a racism pandemic.” For Blacks, they are going through a pandemic on a pandemic.
- In “Being Asian” category, “black” appeared in the top 30. Conversely, in the “Being Black” category, “stopasianhate” appeared in the top 30. This shows that the racism is long-lasting, pervasively prominent, and institutionalized in the US. It shows how communities of POC’s (people of color) come together to form the Black-Asian solidarity and aims to stand against and ultimately dismantle white supremacy. This finding supports the rise in support towards and participation in the fight for the liberties of all marginalized groups. A Chinese American New York-based fashion designer Dao-Yi Chow, who was one of the organizers of Running to Protest’s “Black & Asian Solidarity” rally, remarks on this shift occurring in AAPI community towards Black-Asian unity. “We have new generations of young people and activities... who are down and will show up for the Black and brown struggle, who will continue to fight for the greater social justice movement not only when we’re under attack but when everyone is under attack.
- The word “im” is part of the top 20 in all three categories. It tells us that many users are tweeting

[illegible]

- capture the tone of your tweets and understand how they balance out

5