R Notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

Setup

```
getwd()
```

```
## [1] "/Users/devinaamangal/Desktop/OneDrive - Southern Methodist University/Computing
in Economics"
```

```
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────────────────────
── tidyverse 1.3.0 ──
```

```
## ✓ ggplot2 3.3.0     ✓ purrr   0.3.3
## ✓ tibble  2.1.3     ✓ dplyr   0.8.5
## ✓ tidyr   1.0.2     ✓ stringr 1.4.0
## ✓ readr   1.3.1     ✓ forcats 0.4.0
```

```
## ── Conflicts ────────────────────────────────────────────────────────────── tid
yverse_conflicts() ──
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(tidytext)
library(wordcloud)
```

```
## Loading required package: RColorBrewer
```

```
library(reshape2)
```

```
##
## Attaching package: 'reshape2'
```

```
## The following object is masked from 'package:tidyr':
##
##     smiths
```

```
library(syuzhet)
```

Data

```
tweets <- read.csv("demonetization-tweets.csv")
```

Cleaning up the Data

```
tweets$stripped_text <- gsub("http.*","",  tweets$text)
tweets$stripped_text <- gsub("https.*","", tweets$stripped_text)
tweets$stripped_text <- gsub("RT*", "", tweets$stripped_text)
tweets$stripped_text <- gsub("&amp", "", tweets$stripped_text)
tweets$stripped_text <- str_to_lower(tweets$stripped_text)
tweets$stripped_text <- gsub("@\\w+", "", tweets$stripped_text)
tweets$stripped_text <- gsub("[ |\t]{2,}", "", tweets$stripped_text)
tweets$stripped_text <- gsub("^ ", "", tweets$stripped_text)
tweets$stripped_text <- gsub(" $", "", tweets$stripped_text)
tweets$stripped_text <- gsub("[[:punct:]]", "", tweets$stripped_text)
```

Removing Stop Words

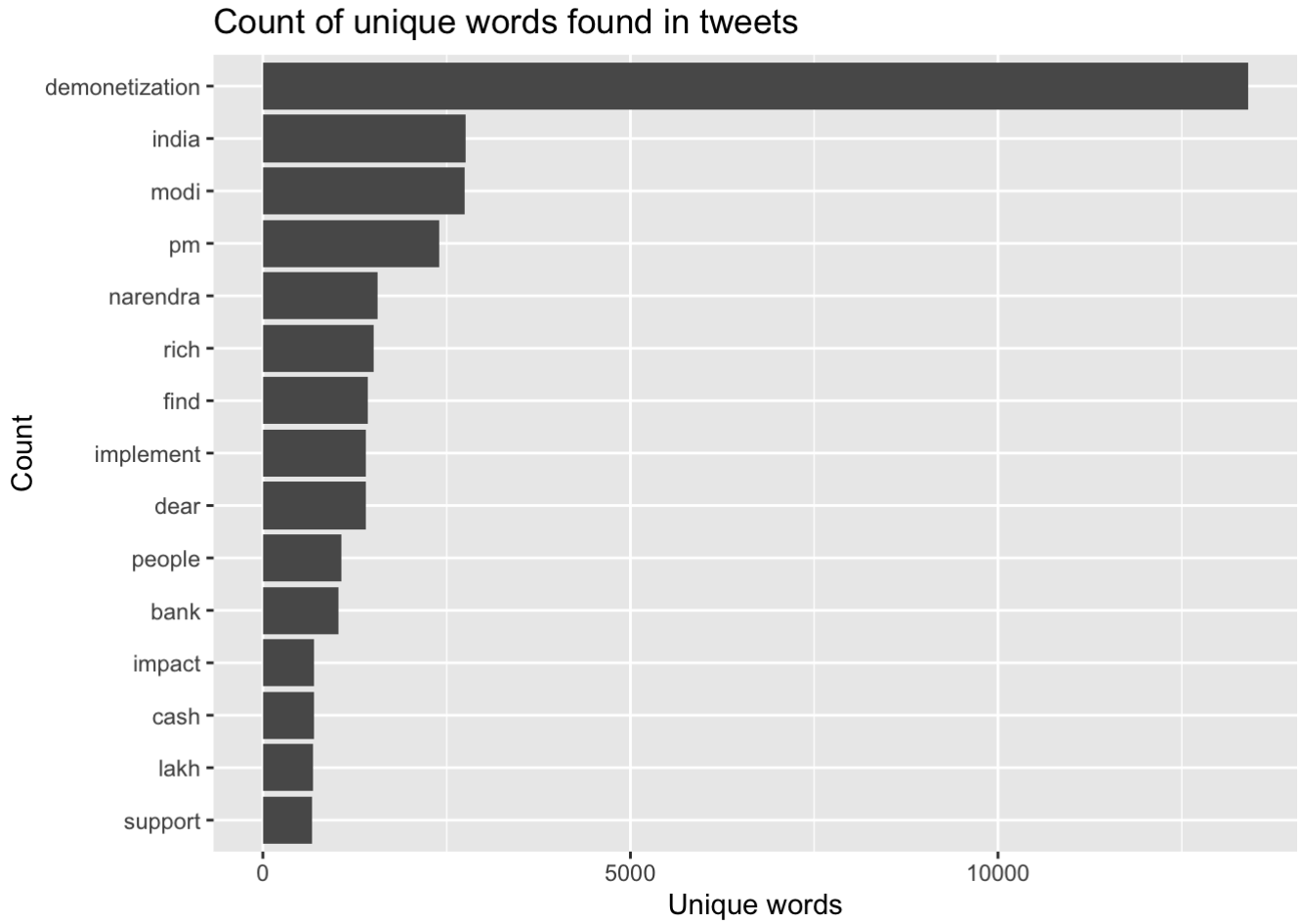
```
tweets_clean <- tweets %>%
  dplyr::select(stripped_text) %>%
  unnest_tokens(word, stripped_text)

SMART_stop_words <- stop_words %>%
  filter(lexicon == "SMART")

tidy_tweets <- tweets_clean %>%
  anti_join(SMART_stop_words)
```

```
## Joining, by = "word"
```

Top 10 words

```
tidy_tweets %>%
  count(word, sort = TRUE) %>%
  top_n(15) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(x = word, y = n)) +
  geom_col() +
  xlab(NULL) +
  coord_flip() +
  labs(x = "Count",
       y = "Unique words",
       title = "Count of unique words found in tweets")
```

```
## Selecting by n
```



WordCloud

```
tidy_tweets %>%
  anti_join(stop_words) %>%
  count(word) %>%
  with(wordcloud(word, n, max.words = 100))
```

```
## Joining, by = "word"
```

```
## Warning in wordcloud(word, n, max.words = 100):
## edu00a0u00bdedu00b8u00a9edu00a0u00bdedu00b8u00a5edu00a0u00bdedu00b8u00a2edu00a0u00bde
du00b8u00addemonetization
## could not be fit on page. It will not be plotted.
```


```
tidy_tweets %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  acast(word ~ sentiment, value.var = "n", fill = 0) %>%
  comparison.cloud(colors = c("darkred", "darkgreen"),
                   max.words = 100)
```

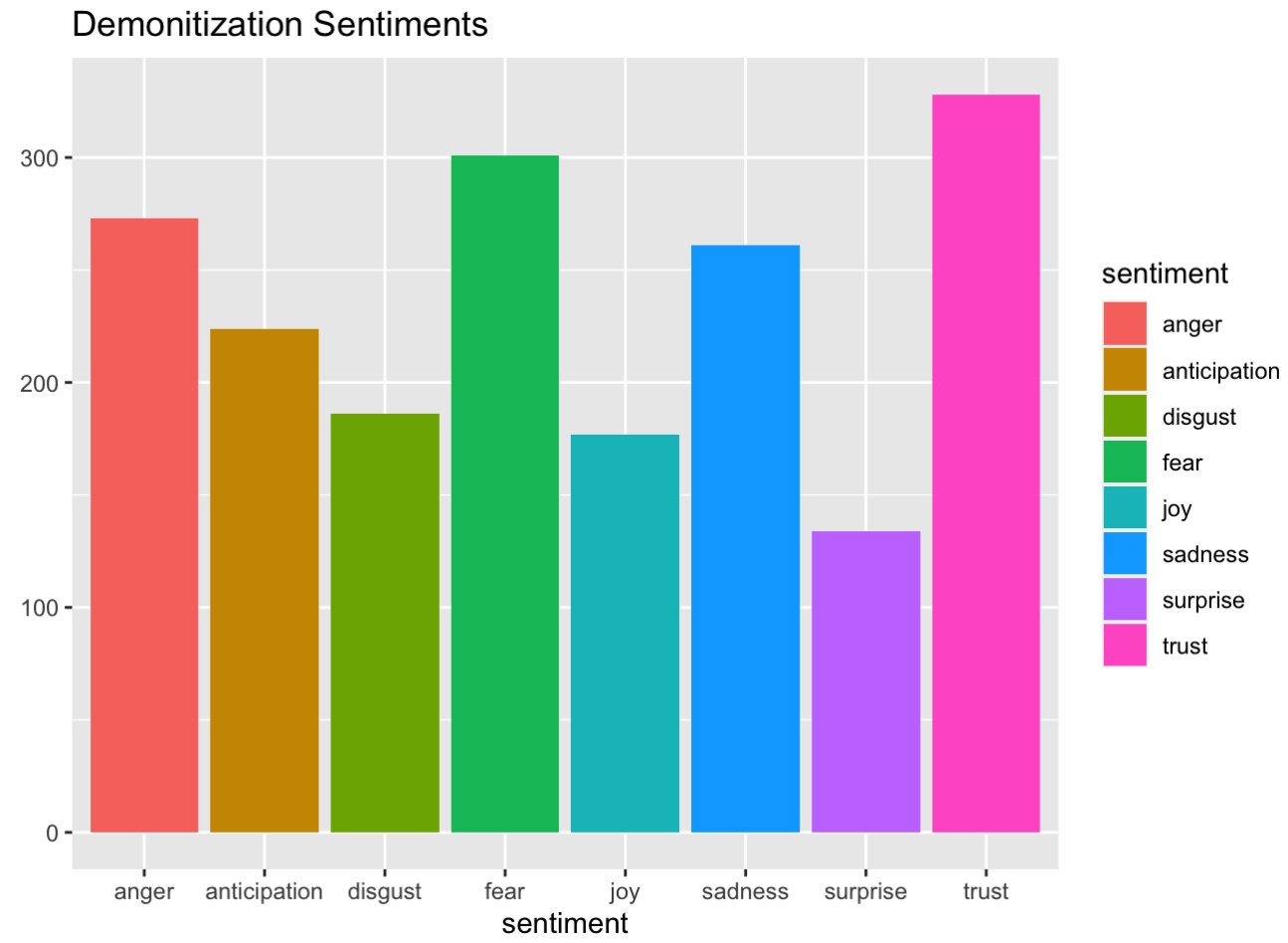
```
## Joining, by = "word"
```

Sentiment Analysis

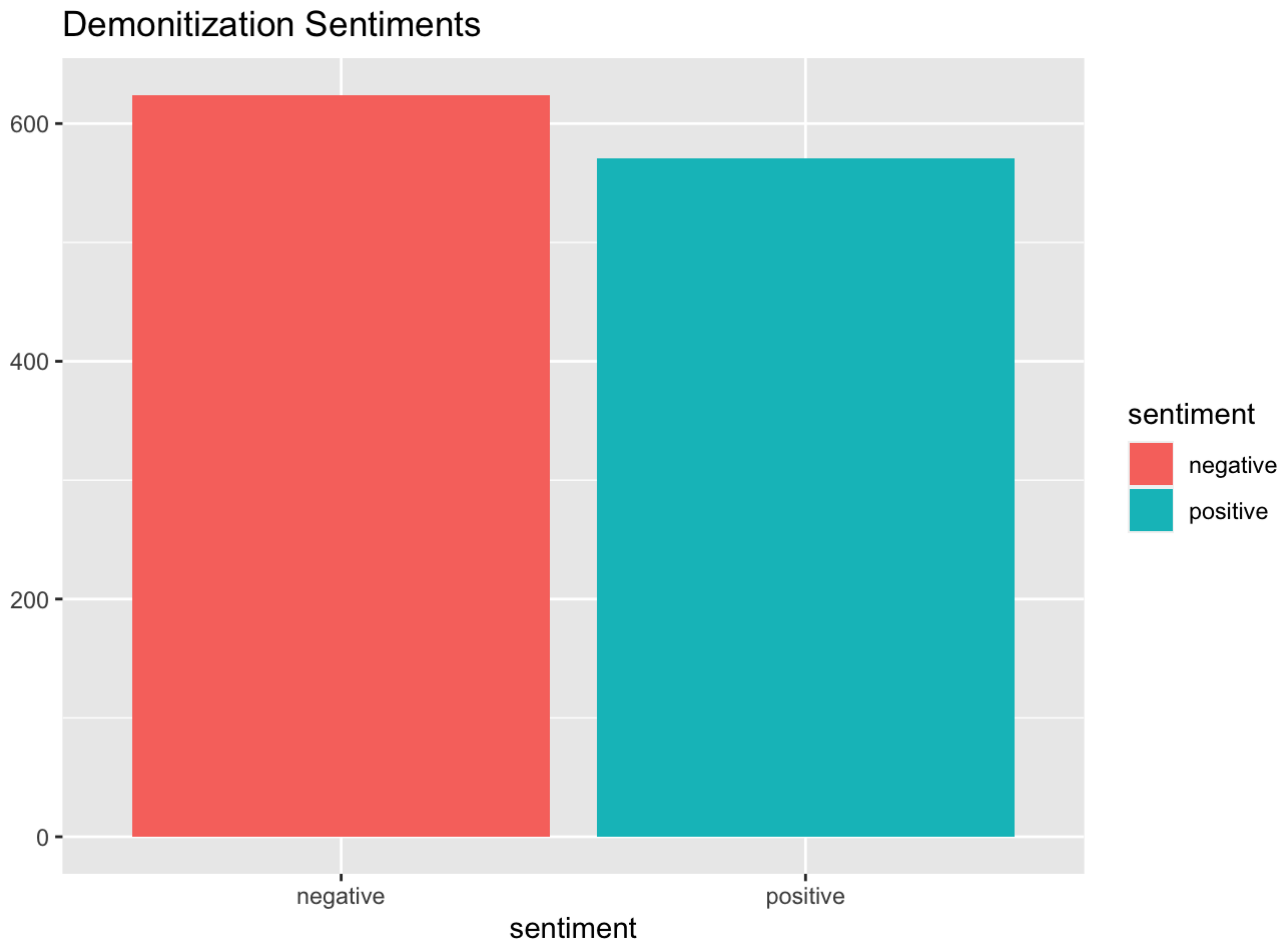
```
tweets_sentiments <- get_nrc_sentiment(as.character(tidy_tweets))
tweets_sentiments1 <- data.frame(t(tweets_sentiments))
new_sentiment <- data.frame(rowSums(tweets_sentiments1))
names(new_sentiment)[1] <- "count"
new_sentiment <- cbind("sentiment" = rownames(new_sentiment), new_sentiment)
rownames(new_sentiment) <- NULL
```

Plots for the Sentiment Analysis

```
qplot(sentiment,
      data = new_sentiment[1:8,],
      weight = count,
      geom = "bar",
      fill = sentiment) +
  ggtitle("Demonitization Sentiments")
```




```
qplot(sentiment,
      data = new_sentiment[9:10,],
      weight = count,
      geom = "bar",
      fill = sentiment) +
  ggtitle("Demonitization Sentiments")
```


Demonitization Sentiments
600
400
200
0
negative
positive
sentiment
sentiment
negative
positive