

# Final Project

*Yuehann Xiao*

*4/24/2018*

## Introduction

The reason that I wish to do this topic is that recently more and more school shootings begin to exist. It happened without any preparation and many students and faculties have been killed by the shooter. From my perspective, it is an important social issue that everyone needs to pay attention to in order to have a stable society in the future.

## What I did in this Final Project

Word Cloud for real-time data Bar plot for the Top 10 words in the word cloud Sentiment Analysis and Pie Chart for real-time data Comparison Word cloud for data from 2012 to now Comparison of top 10 words in the word cloud from 2012 to now Time Series on video related to school shootings

```
library(tuber)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(tidytext)
library(stringr)
library(lubridate)

##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##   date

library(dplyr)
library(tm)

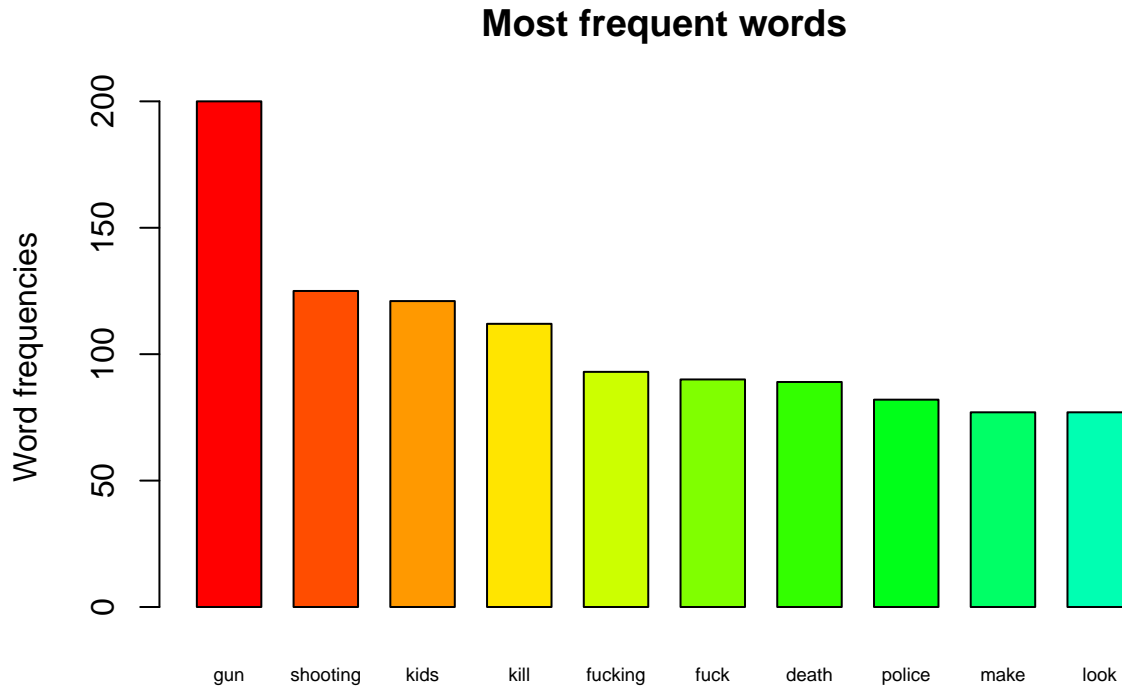
## Loading required package: NLP

library(SnowballC)
library(wordcloud)

## Loading required package: RColorBrewer
```



show the top 10 words in the word cloud and their frequency

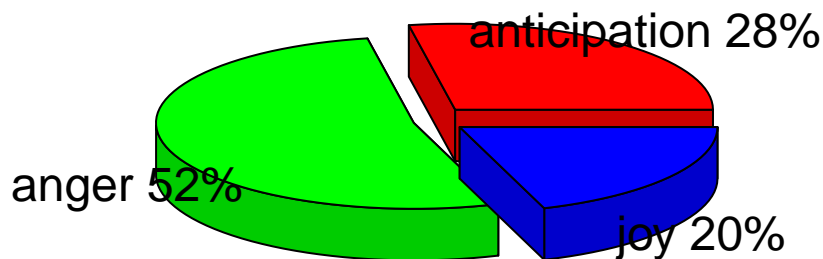


In order for more clear undersing about the word cloud, I did a barplot to directly show the frequency of the top 10 words. From this bar plot we can notice that besides any emotional words, the gun has been mentioned the most which indicates that people are begin to pay attention on the gun using.

## Sentiment Analysis

### Pie Chart for sentimental analysis

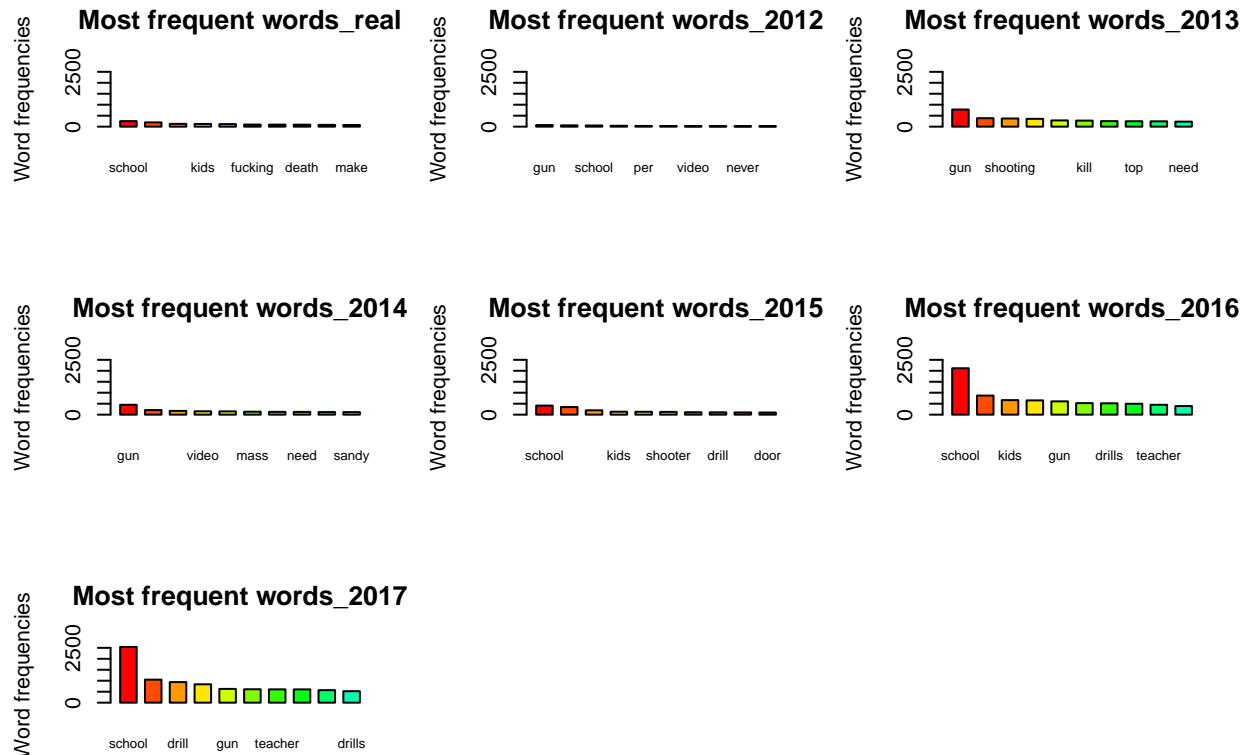
#### Real Time: Pie Chart of sentiment analysis



According to the sentiment analysis, although the postive word “like” has been mentioned 200 times, but “like” in this context auxiliary verb that do not contain real meaning. I could conclude that, as we expected, most comments are negative and intense. We could also conclude from the pie chart that anger has occupy a



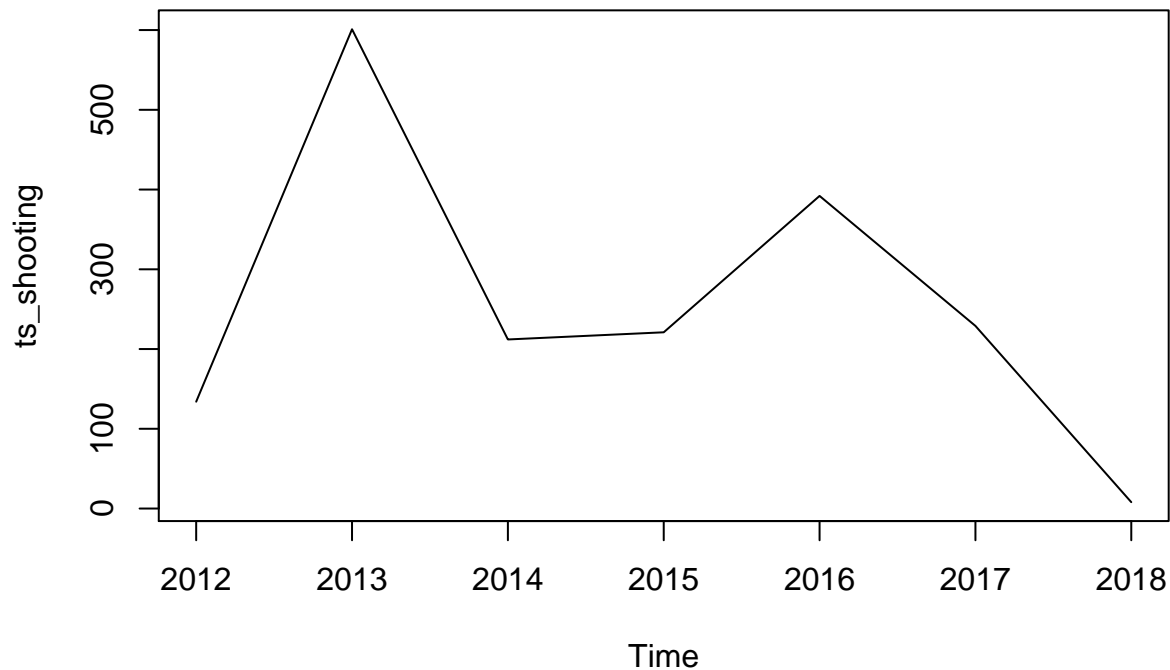
show the top 10 words in the word cloud and their frequency



According to these graphs, although different years are having different most frequent words, we could clearly see that the word “gun” has been highly mentioned in the 2012, 2013, 2014, 2015, and 2018, which is the real time data. Although in 2016 and 2017 the word “gun” has not been mentioned as the most frequent word, it is still in the rank of the top ten mentioned words. This indicates that as the school shooting happening, the most comments, besides the expression of sympathy to the children, would be the eagerness of the gun control or things related to gun. It is a continuous topic from 2012 to 2018 that people keep discussing and debating about. the real time data is only from 04/01 till 04/22, so it has smaller amount of words compare to the rest. 2012 also has small amount of words because it is long time ago and people might not be that active online to comment during 2012

## Time Series on video related with school shootings

### Time Series for Related School Shootings' Videos



According to the plots, we could see that 2013 there is a sharp increase in the 2013 about the videos that related with school shootings. According to the data in the past, I guess that this is because the number of school shootings is doubled from 2012 and 2013, so there were large amount of attention on this issue so the number of videos increase in the youtube. 2014 and 2015 have similar number of videos about school shootings and this might also because. number of school shootings in 2014 and 2014 are similar. The number of videos about school shootings is increasing from 2016 and 2017 because the number of school shootings also decreased. However, the number of videos about school shootings in 2018 are decreasing because this is only part of the data in 2018. Since the number of school shootings has already reached 17 cases in 2018, I predicted that the number of videos about school shootings would also increase compare to previous years.

## Conclusion

From these comments on Youtube, people have paid a lot of attentions on school shootings which are one of the major issues among the society. According to our various word clouds and bar plots, there are lots of discussions on gun. Although it is hard to detect that whether the gun represents “gun control” or any other term, we admitted that “gun” is a major topic that have been emphasized in those comments. Moreover, as people express their sympathies and anger through their comments, I hope more and more people could engage in this societal issue and to provide more protections for students in attempt to create a safe environment for them to study. Students, who are studying in the school and have not step into the society yet, should have a great memory for their school life instead of having emotions like anxious and fear based on the random shootings that would happen around them.