615 Final report

Ang Li

2017/12/18

```
## Warning: package 'syuzhet' was built under R version 3.4.3
```

The Wedding on Twitter

Model Kate Upton and Houston Astros player Justin Verlander got married on November 4th, 2017. The famous supermodel and pitcher finally got togather at an old church at the Rosewood Castiglion Del Bosco resort in Italy. Although they kept their wedding 'small', celebrities attended it, media reported it, so it still became a hot topic on the social media platform, twitter.In this project,I am going to try to find and study what people are talking about on twitter.

```
###WordCloud for Kate Upton

KateUpton<-readRDS("tweets_kateupton.rds", refhook = NULL)

# str_replace_all(HoustonRockets$text, "@", "")
wordCorpus1 <- Corpus(VectorSource(str_replace_all(KateUpton$text, "@", "")))
wordCorpus1 <- tm_map(wordCorpus1, removePunctuation)
wordCorpus1 <- tm_map(wordCorpus1, content_transformer(tolower))
wordCorpus1 <- tm_map(wordCorpus1, removeWords, stopwords("english"))
wordCorpus1 <- tm_map(wordCorpus1, removeWords, c("kate","upton","kateupton","fuc
k"))
wordCorpus1 <- tm_map(wordCorpus1, stripWhitespace)
#saveRDS(starbucks, file="starbucks.rds")

set.seed(100)
wordcloud(words = wordCorpus1,scale=c(5,0.2),max.words=100, random.order=FALSE,
          rot.per=0.15,use.r.layout=FALSE, colors=brewer.pal(5,"Greens")[4:8])
```


```
###WordCloud for Justin Verlander

JustinVerlander<-readRDS("tweets_JVerlander.rds", refhook = NULL)

# str_replace_all(HoustonRockets$text, "@", "")
wordCorpus2 <- Corpus(VectorSource(str_replace_all(JustinVerlander$text, "@", "")
))
wordCorpus2 <- tm_map(wordCorpus2, removePunctuation)
wordCorpus2 <- tm_map(wordCorpus2, content_transformer(tolower))
wordCorpus2 <- tm_map(wordCorpus2, removeWords, stopwords("english"))
wordCorpus2 <- tm_map(wordCorpus2, removeWords, c("justin", "verlander", "justinv
erlander"))
wordCorpus2 <- tm_map(wordCorpus2, stripWhitespace)
#saveRDS(starbucks, file="starbucks.rds")

set.seed(100)
wordcloud(words = wordCorpus2,scale=c(5,0.2),max.words=200, random.order=FALSE,
          rot.per=0.2, use.r.layout=FALSE, colors=brewer.pal(5,"Greens")[4:8])
```

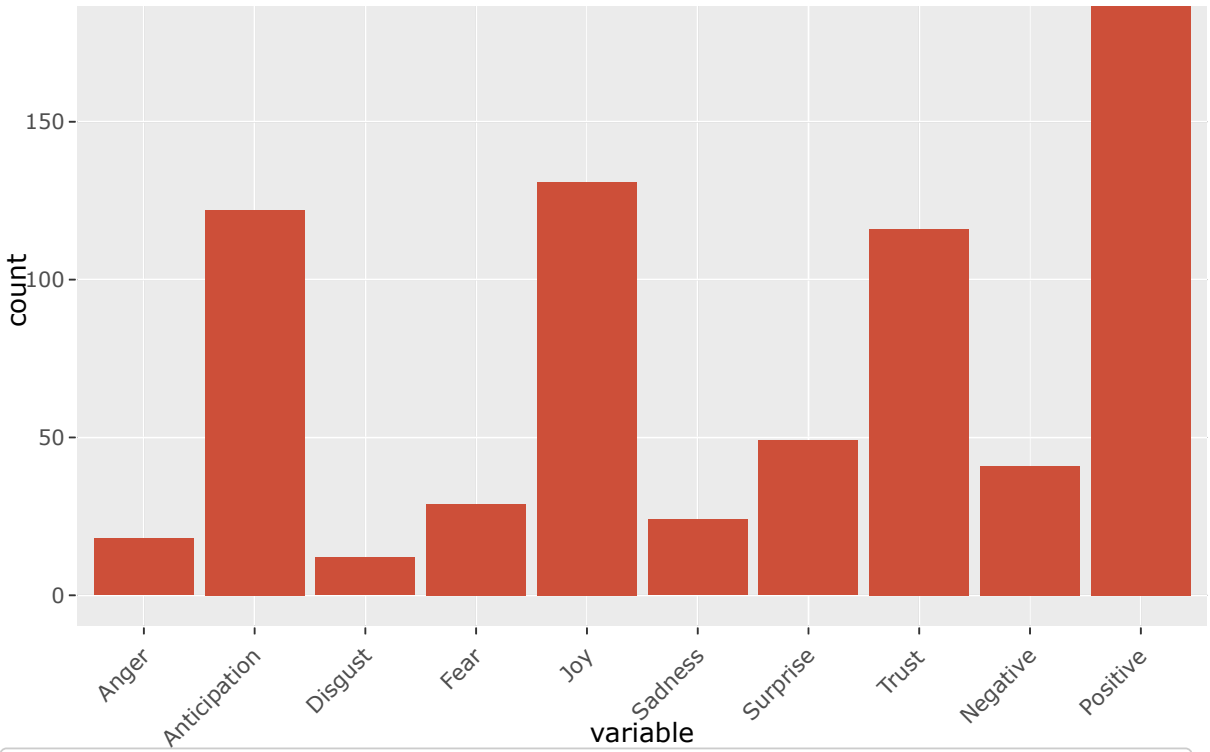
For Kate's wordcloud, for the plot we can see that, except those words that describe or compliment Kate herself, such as 'amazing', 'gorgeous', the top 3 most frequently mentioned words are 'Justin', 'Verlander' and 'wedding'; For Justin, except for those words like 'Astros', which is the name of his team, the top 3 frequently mentioned words are 'Kate', 'Upton' and 'wedding'. It shows that on twitter, people are definitely talking about their relationship, and their wedding.

```
### Sentiment Plot
eee<-readRDS('player.sentiment.rds')
k<-as.data.frame(eee[1])
j<-as.data.frame(eee[2])
colnames(k)<-c("Anger","Anticipation","Disgust","Fear","Joy","Sadness","Surprise"
,"Trust","Negative","Positive")
colnames(j)<-c("Anger","Anticipation","Disgust","Fear","Joy","Sadness","Surprise"
,"Trust","Negative","Positive")

k1<-reshape2::melt(data = k, measure.vars = c("Anger","Anticipation","Disgust","F
ear","Joy","Sadness","Surprise","Trust","Negative","Positive"))

k2 = k1 %>%
  select(variable, value)%>%
  group_by(variable)%>%
  summarise(count=sum(value))

j1<-melt(data = j, measure.vars = c("Anger","Anticipation","Disgust","Fear","Joy"
,"Sadness","Surprise","Trust","Negative","Positive"))

j2 = j1 %>%
  select(variable, value)%>%
  group_by(variable)%>%
  summarise(count=sum(value))


p1<-ggplot(k2, aes(variable, count))+
  geom_col(fill="tomato3")+
  ggtitle("Sentiment Plot for Kate") +
     theme(plot.title = element_text(lineheight=.8, face="bold"),axis.text.x = el
ement_text(angle = 45,hjust = 1,size=9))
```

```
p2<-ggplot(j2, aes(variable, count))+
  geom_col(fill="tomato3")+
  ggtitle("Sentiment Plot for Justin") +
     theme(plot.title = element_text(lineheight=.8, face="bold"),axis.text.x = el
ement_text(angle = 45,hjust = 1,size=9))
```

```
ggplotly(p1)
```

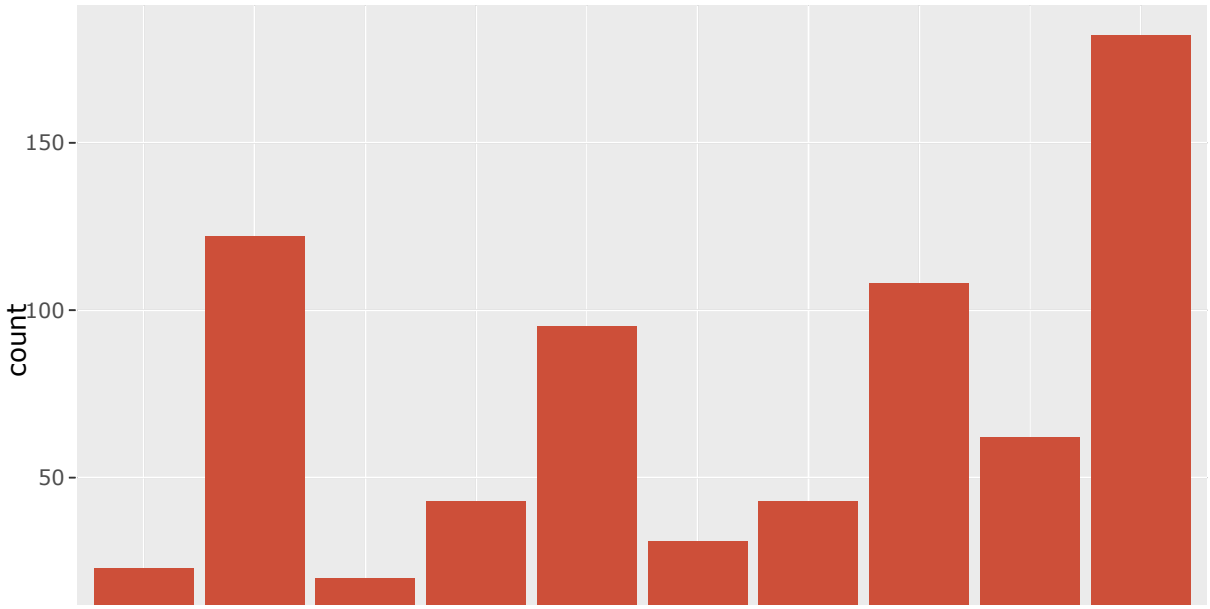
```
## We recommend that you use the dev version of ggplot2 with `ggplotly()`
## Install it with: `devtools::install_github('tidyverse/ggplot2')`
```

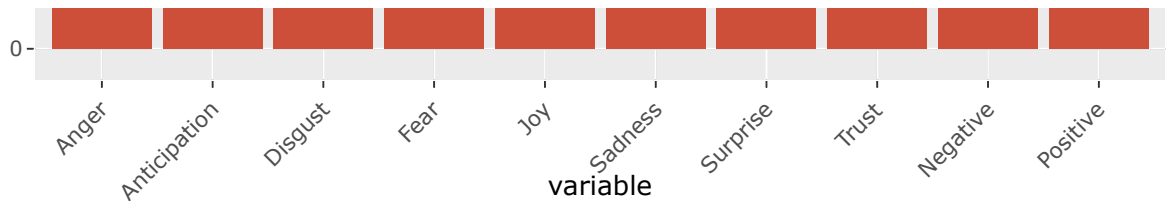



```
ggplotly(p2)
```

```
## We recommend that you use the dev version of ggplot2 with `ggplotly()`
## Install it with: `devtools::install_github('tidyverse/ggplot2')`
```

Sentiment Plot for Justin





From the graph of sentiment for Kate, we can see that the word that related with positive emotions are greatly exceed negative emotions, especially for emotions like 'Joy', 'Trust', 'Anticipation'.

Similar with the graph of Kate, from the graph of sentiment for Justin, we see that the word that related with positive emotions are greatly exceed negative emotions (I notice negative emotions uprise a little bit), and emotions like 'Joy', 'Trust', 'Anticipation' remains high in our sample.

```
## Map
wed<-readRDS("tweets_wedding.rds")
wed$longitude<-as.numeric(wed$longitude)
wed$latitude<-as.numeric(wed$latitude)
q<-qmplot(longitude, latitude, data = wed,
       colour = I('red'), size = I(3), darken = .3)
```

```
## Using zoom = 2...
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/0/1.png
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/1/1.png
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/2/1.png
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/3/1.png
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/0/2.png
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/1/2.png
```

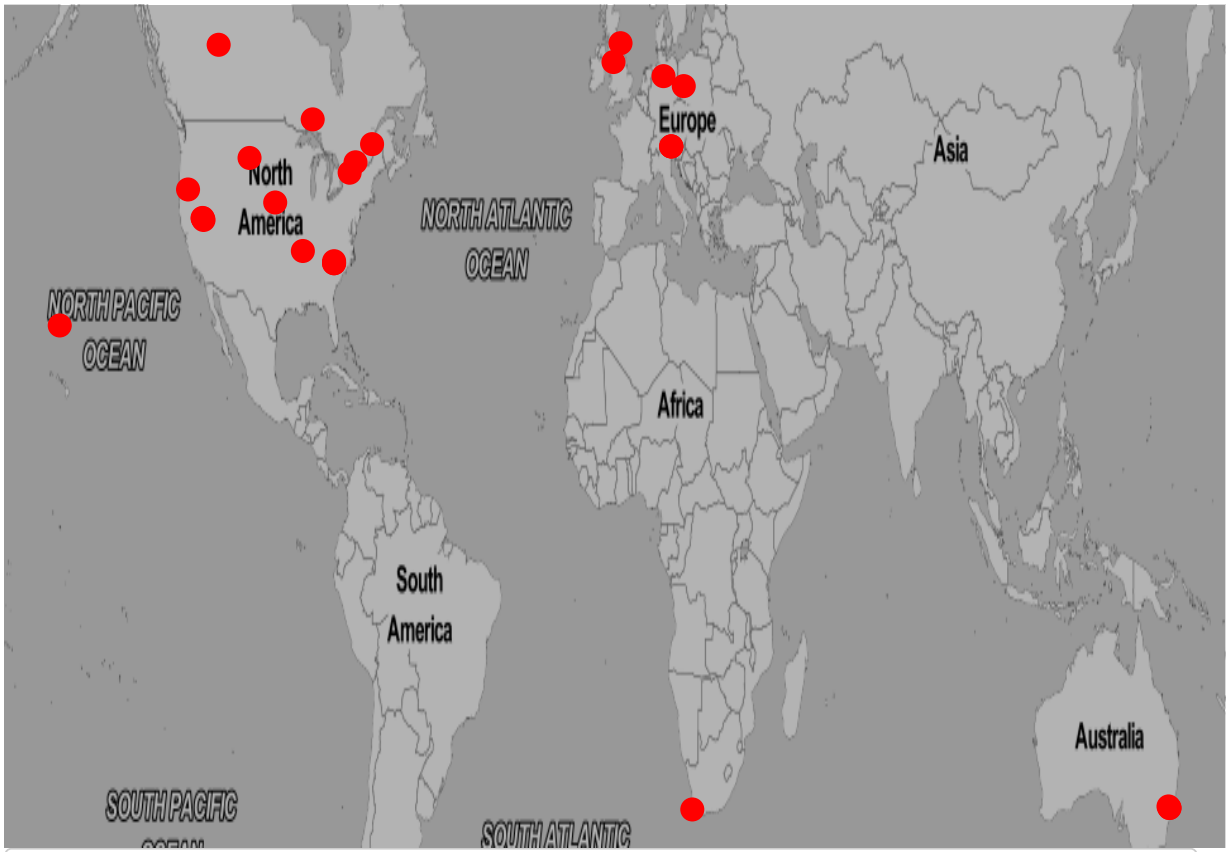
```
## Map from URL : http://tile.stamen.com/toner-lite/2/2/2.png
```

```
## Map from URL : http://tile.stamen.com/toner-lite/2/3/2.png
```

```
## Warning: `panel.margin` is deprecated. Please use `panel.spacing` property
## instead
```

```
ggplotly(q)
```

```
## We recommend that you use the dev version of ggplot2 with `ggplotly()`
## Install it with: `devtools::install_github('tidyverse/ggplot2')`
```




```
#ggmap(qmap()) + geom_point(aes(x=longitude, y=latitude, colour='tomato3'), data=
wed)
```

This is the map that shows the location of people who have interaction with kate's and Justin's twiiter account. Due to the privacy policy of twitter, I cannot obtain the geographic locations of all the people. However, we can still find something from even these a few points. Those points are spotted on several continents, such as North America,Europe or Oceania. It means that Kate and Justin's wedding are discussed and concerned worldwidely, which means this topic is popular and well-known for recent days.