

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 together at an old church at the Rosewood Castiglione 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", "fuck"))
```

```
wordCorpus1 <- tm_map(wordCorpus1, stripWhitespace)
```

```
#saveRDS(starbucks, file="starbucks.rds")
```

```
set.seed(100)
```

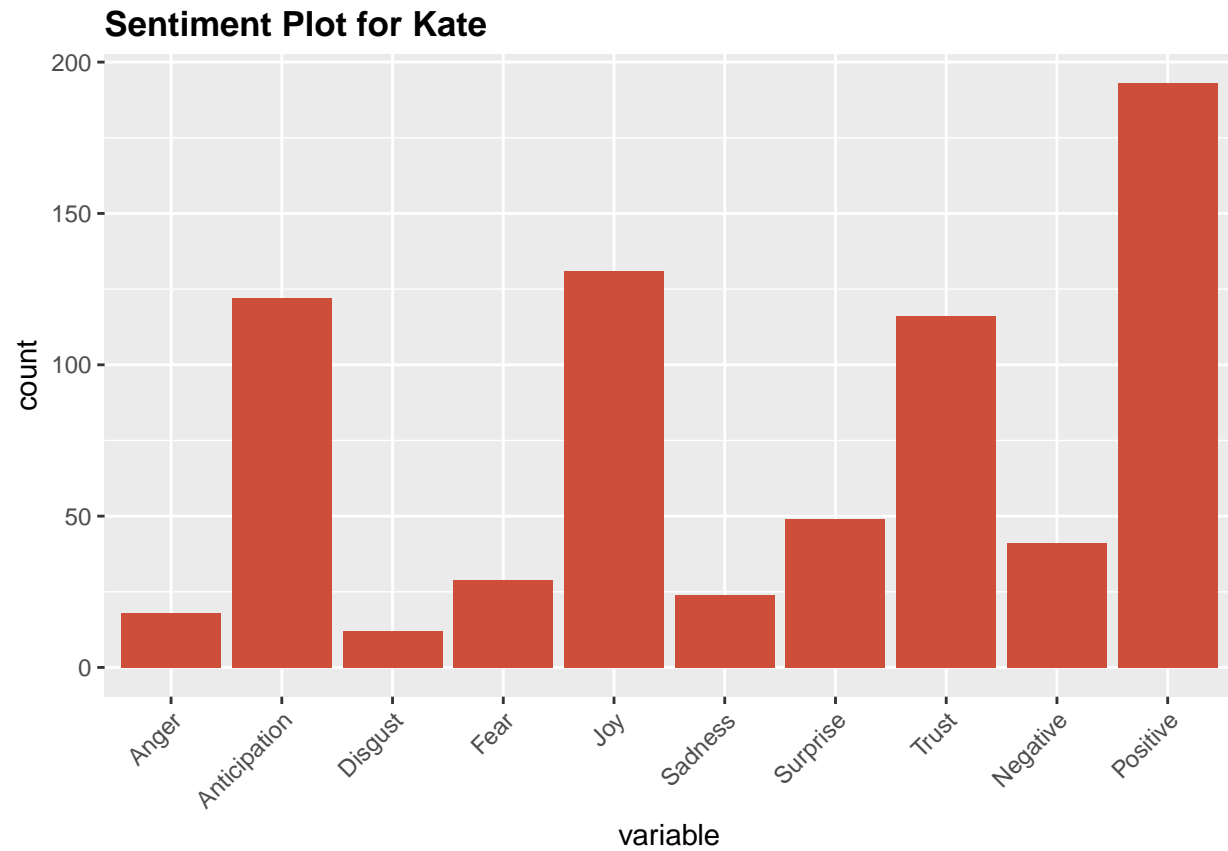
```
wordcloud(words = wordCorpus1, scale=c(5,0.2), max.words=100, random.order=FALSE,  
          rot.per=0.15, use.r.layout=FALSE, colors=brewer.pal(5, "Greens")[4:8])
```



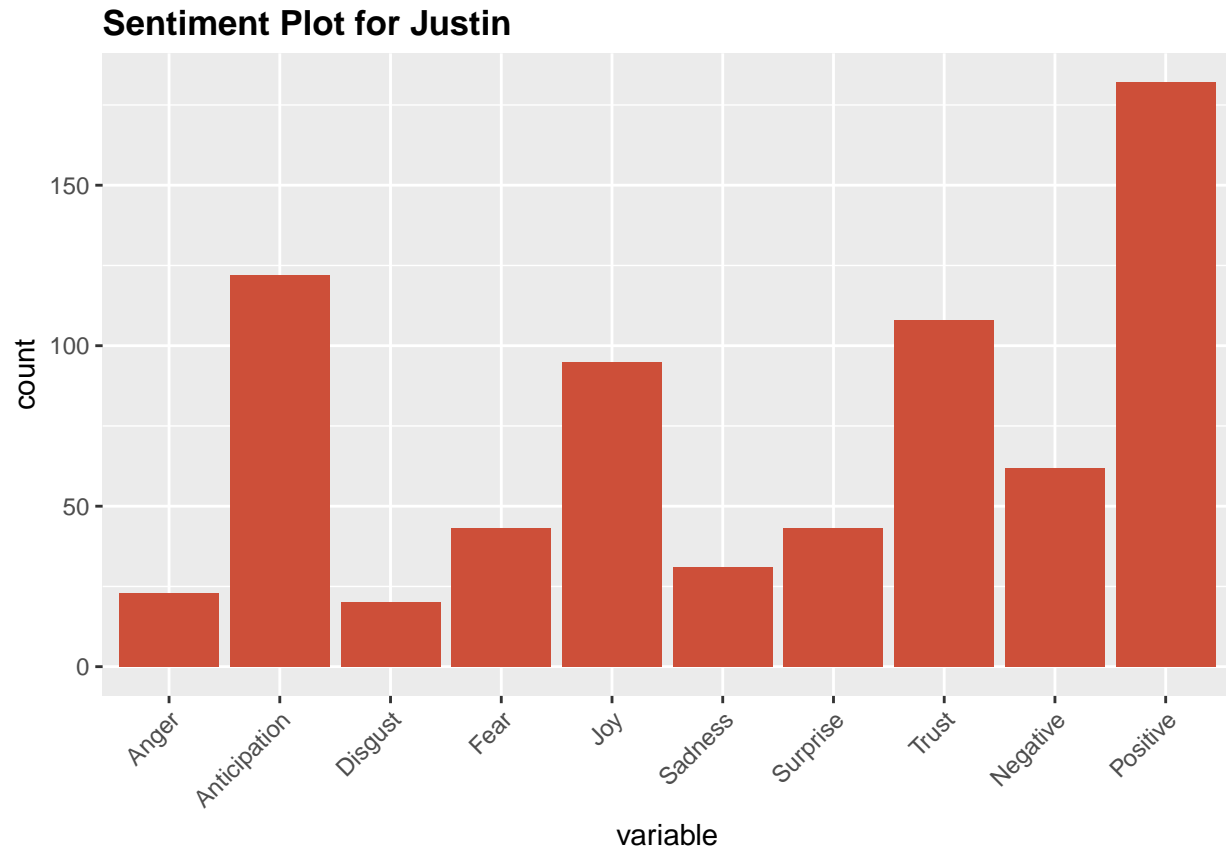
```
geom_col(fill="tomato3")+
ggtitle("Sentiment Plot for Kate") +
  theme(plot.title = element_text(lineheight=.8, face="bold"),axis.text.x = element_text(angle = 45,
```

```
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 = element_text(angle = 45,
```

p1



p2



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)
qplot(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
```



```
#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 twitter 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 worldwide, which means this topic is popular and well-known for recent days.