MA615 Final Project

Jiazhou Liu December 17,2017

Introduction

Data analysis about brands is always a hard topic. The most common approach is to start a survey aiming at the general public. It takes a huge amount of time and effort, and the result can be biased as people aren't always serious about answering a survey.

This project is about analyzing a brand's perception based on the data in the well-known social media: Twitter. The brand I choose is Alienware, a popular brand which sells gaming computers and accessories. I would like to know how people think about this brand: Do people feel mostly positive or negative about this brand? What words do they always use to describe this brand?

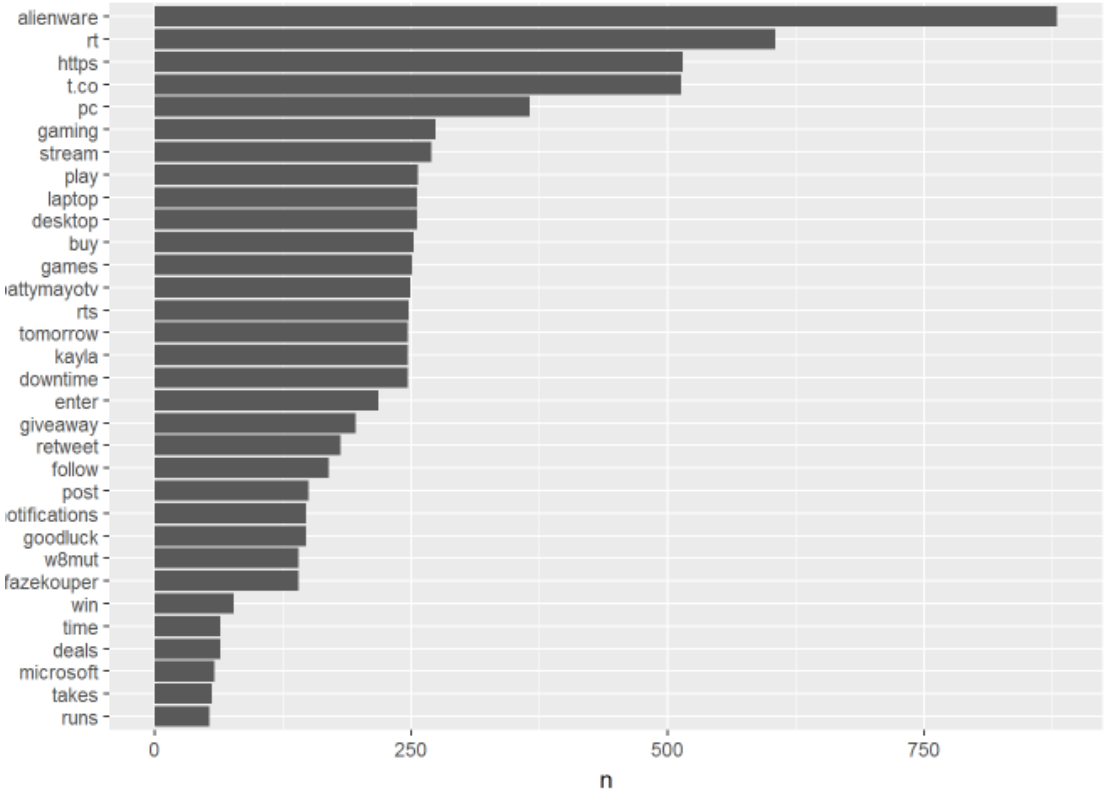
Let's go to the actual analysis!

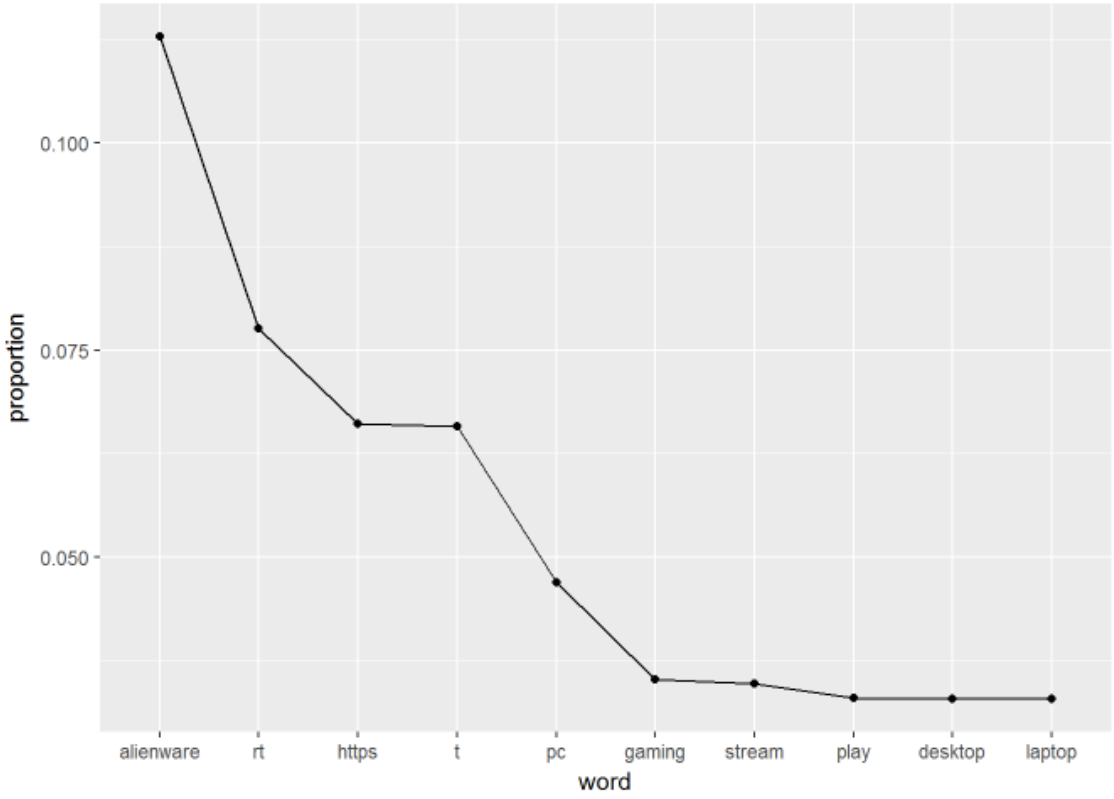
Analysis

Firstly, I need to retrieve data from twitter. I use the twitterR package for the purpose. Then I transform tweets to separate words for further analysis by using tidytext package.

```r
# Get 1000 tweets about Alienware since 2010-01-01 in English.
tweets<-searchTwitter('Alienware',since='2010-01-01',n =1000,lang="en")

# Transform tweets list into a data frame
tweets.df <- twListToDF(tweets)

#Transform tweets into lines of seperate words.
tidy_tweets <- tweets.df %>%
  unnest_tokens(word,text)

#Disregard not meaningful words,such as"the" before we do any analysis.
data(stop_words)

tidy_tweets <-  tidy_tweets %>%
  anti_join(stop_words)
```

Then, I create two graphs which show the most frequent words in tweets about Alienware.





From these graphs, I can see that “gaming, play, games” appear frequently, and this means Alienware’s brand marketing is quite successful as it’s closely associated with gaming computers in people’s minds.

To further explore people’s attitude towards Alienware, I try to find the ratio between positive and negative words in tweets about Alienware.

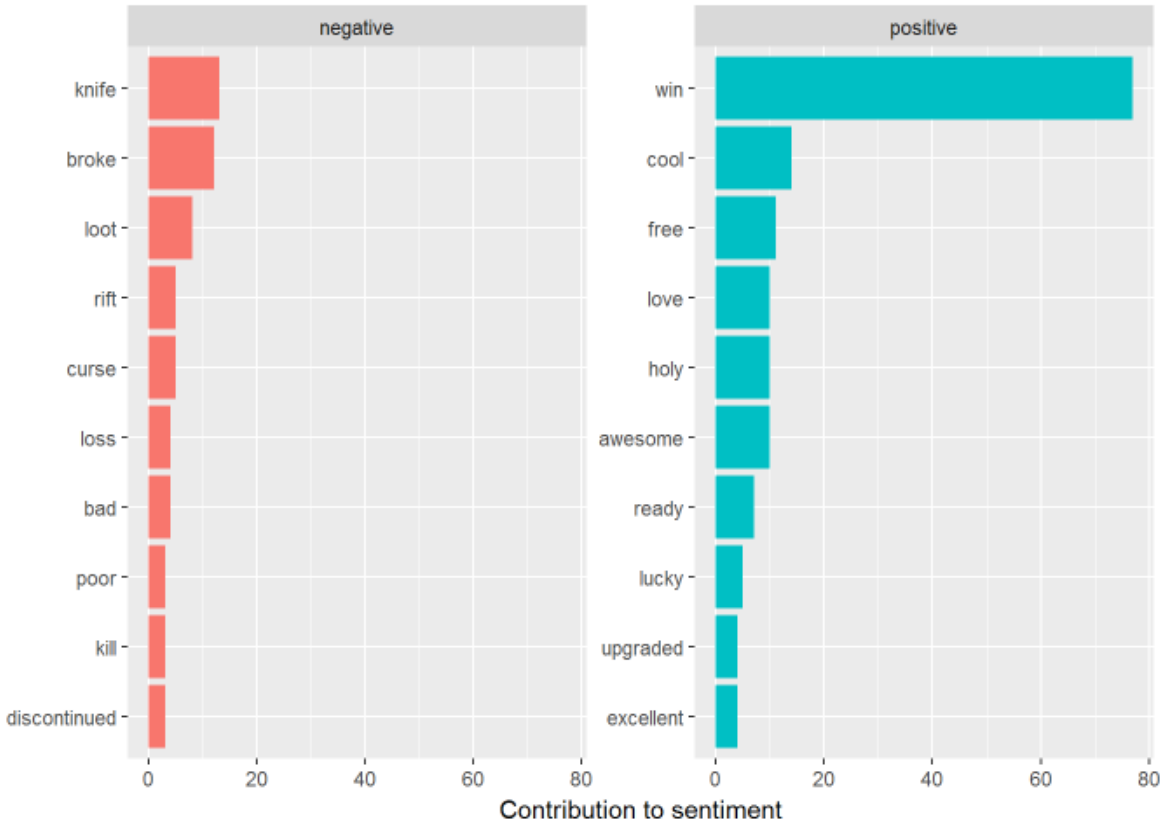
```
#Count how many positive words and negative words in tweets,
#and show top 10 frequent words of both types in seperate plots.
tidy_tweets%>%
  inner_join(get_sentiments("bing")) %>%
  count(sentiment, sort = TRUE)
```

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

```
## # A tibble: 2 x 2
##   sentiment     n
##       <chr> <int>
## 1  positive   218
## 2  negative   136
```

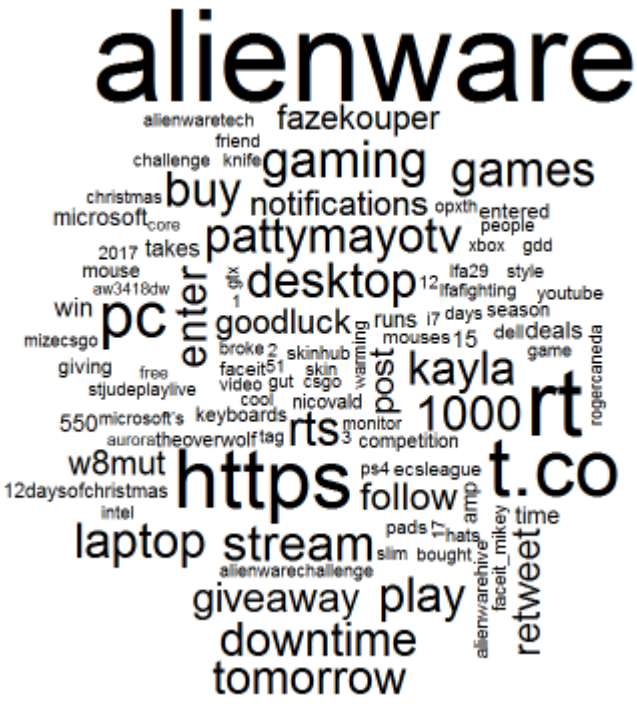
The ratio between positive and negative words equals 218/136=1.6, which means the general perception about Alienware is positive.

In order to dive deeper, I create plots to find out which words contribute to positive and negative perception.



From the plots, I see that “broke” stands out as a negative word, and I think Alienware should improve the reliability of their computers. I also see that “win” show up as a positive word, and it means the performance of Alienware computers are outstanding.

To have more vivid representations, I create two word clouds, which clearly show the words appearing most frequently, and the words contributing to positive or negative sides. A table is also offered to give quantitative information.


alienware
alienwaretech
fazekouper
friend
challenge
knife
gaming
games
christmas
buy
notifications
opxth
entered
microsoft
core
people
pattymayotv
2017
takes
xbox
gdd
mouse
aw3418dw
desktop
12
lfa29
style
lfafighting
youtube
win
pc
enter
goodluck
runs
i7
days
season
mizecsgo
broke
2
skinhub
mouses
15
dell
deals
game
giving
free
faceit
51
skin
warming
post
kayla
rt
stjudeplaylive
video
gut
csgo
cool
nicovald
1000
rogercaneda
550
microsoft's
keyboards
rts
monitor
3
competition
auroratheoverwolf
tag
w8mut
https
ps4
ecsleague
t.co
12daysofchristmas
follow
amp
intel
time
pads
17
hats
laptop
stream
slim
bought
alienwarehive
faceit_mikey
retweet
alienwarechallenge
giveaway
play
downtime
tomorrow

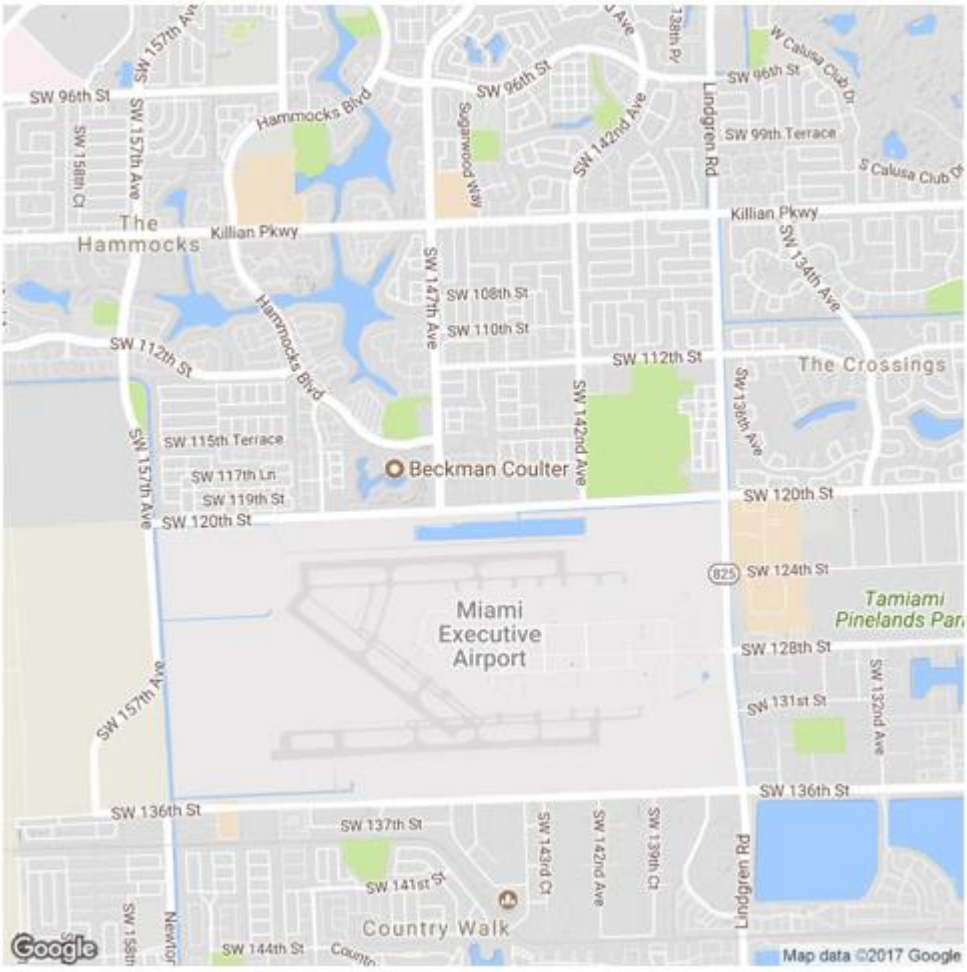
negative

positive

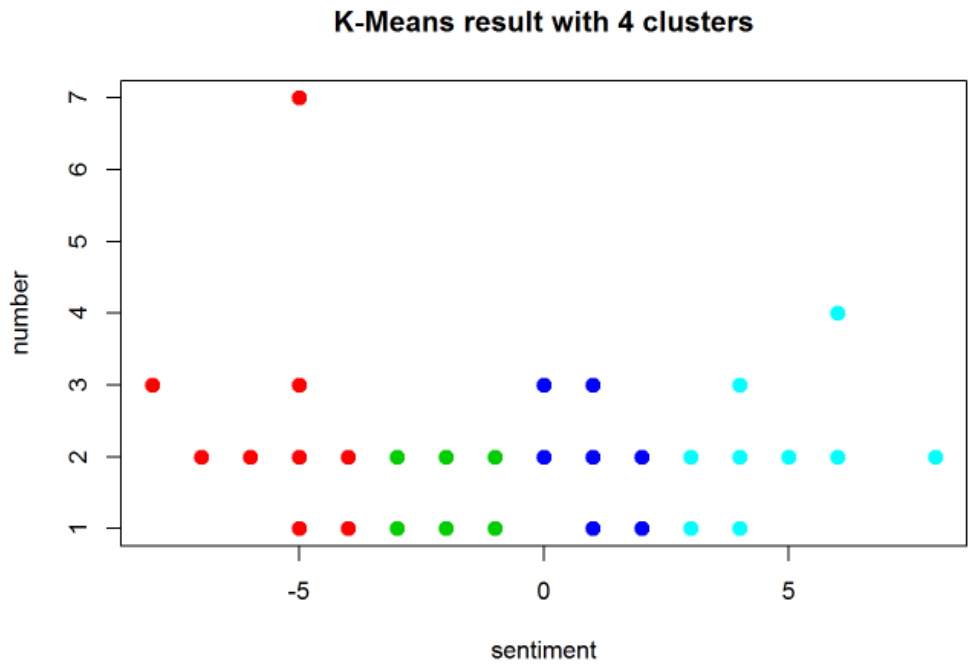
words of sentiment

word	sentiment	n
win	positive	77
cool	positive	14
knife	negative	13
broke	negative	12
free	positive	11
awesome	positive	10
holy	positive	10
love	positive	10
loot	negative	8
ready	positive	7
curse	negative	5
lucky	positive	5
rift	negative	5

Because tweets normally don't contain geographical information, so a map of data distribution isn't applicable in this case. Therefore, I use ggmap to draw a map of Alienware company's location, which is at Hammocks, Florida.



Finally, I create cluster plot that defines some clusters of sentimental tweets about Alienware. X-axis represents
the sentiment score: higher score means a more positive attitude. Y-axis represents the number of sentimental words in that tweet.



Conclusion

According to the data from twitter, I find that public's perception about Alienware is generally positive. People always use "cool" and "win" to describe this brand, and it implies the outlook of Alienware products is stylish, and the performance of them is excellent. However, I also notice people use "broke" to describe it, which means the quality of Alienware products still needs to be improved.

Most importantly, this brand is closely associated with "games" and "gaming" in people's tweets, and I consider the marketing of Alienware is therefore successful, as the position of Alienware is clear in people's mind: a stylish brand of top-level-performance gaming computers.