CA- 1

INTM517 – Business Analytics

Name - Nowneesh T

Section - Q3E45

Roll No – RQ3E45A08

Question:

Q1. Automotive industry is one of the largest and highly competitive economic sectors in the world. Due to the high competition, automotive companies are moving toward using social media sites to reach further customers and advertise their products in considerably short time.

Twitter is one of the highest growing social media websites in the world. Analyse the sentiment of customers of these two companies with the help of 500 tweets downloaded from each company's official Twitter handle and analysed in R. In this analysis, user will able understand customer sentiments from the social media engagement of a brand (In this particular case tweeter). Collect the data from twitter using word "Mercedes" and "Audi".

1. Which emotion is having highest and lowest number of terms
2. Create a word cloud represents which term is having highest frequency
3. Find whether the feedback is positive or negative.
4. Create a visualization for above scenerios.
5. Analyse the data and mention 3-4 points to conclude the complete analysis.

Answer:

Program code:

```
library(twitteR)

library(ggplot2)

library(syuzhet)

library(tm)

appname <- "R_Demo_Data"

key <- "u7BLabVpFhhv6s7WbHA8cLvHe"

secret <- "qGVlE47azdmGRh6TuZI1Io0rEIgufp0Xm0HguCUjXA4LtMNCJM"
```

```r
access<-"1291627776009592832-5gOVKJF0AfBQEfyMm39EFQZc9J8X47"

access_secret<-"iZYa1zJ44jI7go1WZykHmU0q0vjgEfHcINNhYcrtWUKe6"


setup_twitter_oauth(key, secret, access, access_secret)


tweets_tech <- searchTwitter("Mercedes", n=500,lang = "en")


a <- twListToDF(tweets_tech)
# library(tm)


corpus = iconv(a$text, "latin1", "UTF-8")
corpus<- Corpus(VectorSource(corpus))


# corpus==>Documents/Docs
# VectorSource==>vector
# a$text==> row/records


toSpace <- content_transformer(function (x , pattern ) gsub(pattern, " ", x))
docs=corpus
docs <- tm_map(docs, toSpace, "/")
docs <- tm_map(docs, toSpace, "@")
docs <- tm_map(docs, toSpace, "\\|")
corpus=docs
corpus<- tm_map(corpus,tolower)
corpus<-tm_map(corpus,removePunctuation)# remove puntuations like , .
corpus<- tm_map(corpus,removeNumbers)
cleanset<-tm_map(corpus,removeWords,stopwords('english'))# remove common words
removeURL<- function(x)gsub('http[[:alnum:]]=','',x)
cleanset<-tm_map(cleanset,content_transformer(removeURL))
```

```r
x=cleanset

tdm1<-TermDocumentMatrix(cleanset)

tdm1 # display information

tdm1<-as.matrix(tdm1)

#============================================================

v=sort(rowSums(tdm1))

library(wordcloud)

w<-data.frame(names(v),v)

colnames(w)<-c('word','freq')

set.seed(1234)

wordcloud(words=w$word,freq=w$freq)

library(wordcloud2)

letterCloud(w,
       word="R",
       size=5,
       color="rainbow")

letterCloud(w, word = "WORDCLOUD2", wordSize = 1)

wordcloud2(w, size=10,color = "random-light", backgroundColor = "grey")

wordcloud2(w,size=5,shape = 'pentagon')

g=w

#============================================================

write.csv(tdm1,"tdm1.csv")

library(syuzhet)

data=read.csv("tdm1.csv")

mysentiment_tech<-get_nrc_sentiment((data$X))


#calculationg total score for each sentiment


Sentimentscores_tech<-data.frame(colSums(mysentiment_tech[,]))
```

```r
#======================================

tweets_tech <- searchTwitter("Audi", n=500,lang = "en")


a <- twListToDF(tweets_tech)
# library(tm)
corpus = iconv(a$text, "latin1", "UTF-8")
corpus<- Corpus(VectorSource(corpus))
# corpus==>Documents/Docs
# VectorSource==>vector
# a$text==> row/records
toSpace <- content_transformer(function (x , pattern ) gsub(pattern, " ", x))
docs=corpus
docs <- tm_map(docs, toSpace, "/")
docs <- tm_map(docs, toSpace, "@")
docs <- tm_map(docs, toSpace, "\\|")
corpus=docs
corpus<- tm_map(corpus,tolower)
corpus<-tm_map(corpus,removePunctuation)# remove puntuations like , .
corpus<- tm_map(corpus,removeNumbers)
cleanset<-tm_map(corpus,removeWords,stopwords('english'))# remove common words
removeURL<- function(x)gsub('http[[:alnum:]]=','',x)
cleanset<-tm_map(cleanset,content_transformer(removeURL))
x=cleanset


tdm1<-TermDocumentMatrix(cleanset)
tdm1 # display information
```

```r
tdm1<-as.matrix(tdm1)


#=============================================

v=sort(rowSums(tdm1))

library(wordcloud)
w<-data.frame(names(v),v)


colnames(w)<-c('word','freq')
set.seed(1234)
wordcloud(words=w$word,freq=w$freq)


library(wordcloud2)



letterCloud(w,
        word="R",
        size=5,
        color="rainbow")

letterCloud(w, word = "WORDCLOUD2", wordSize = 1)
wordcloud2(w, size=10,color = "random-light", backgroundColor = "grey")
wordcloud2(w,size=5,shape = 'pentagon')


#===========================



write.csv(tdm1,"tdm1.csv")
```

```
library(syuzhet)

data=read.csv("tdm1.csv")

mysentiment_tech1<-get_nrc_sentiment((data$X))

#calculationg total score for each sentiment

Sentimentscores_tech1<-data.frame(colSums(mysentiment_tech1[,]))

#====================================


names(Sentimentscores_tech)<-"Mercedes"

names(Sentimentscores_tech1)<-"Audi"

Sentimentscores_tech<-
cbind("sentiment"=rownames(Sentimentscores_tech),Sentimentscores_tech,Sentimentscor
es_tech1)

rownames(Sentimentscores_tech)<-NULL

#********************************************************************
***********

ggplot(data=Sentimentscores_tech,aes(x=sentiment,y=Mercedes))+

  geom_bar(aes(fill=sentiment),stat = "identity")+

  theme(legend.position="none")+

  xlab("Sentiments")+ylab("Mercedes")+ggtitle("Mercedes Analysis")

ggplot(data=Sentimentscores_tech,aes(x=sentiment,y=Audi))+

  geom_bar(aes(fill=sentiment),stat = "identity")+

  theme(legend.position="none")+

  xlab("Sentiments")+ylab("Audi")+ggtitle("Audi Analysis")
> View(Sentimentscores_tech)
> View(g)
> View(w)
```

1. Which emotion is having highest and lowest number of terms

     Use `View(Sentimentscores_tech)`

	sentiment	Mercedes	Audi
1	anger	39	25
2	anticipation	68	44
3	disgust	19	18
4	fear	41	34
5	joy	52	28
6	sadness	32	33
7	surprise	28	21
8	trust	72	49
9	negative	73	60
10	positive	133	87

The highest emotion for Mercedes is Trust
The lowest emotion for Mercedes is Disgust
The highest emotion for Audi is Trust
The lowest emotion for Audi is Disgust

2. Create a word cloud represents which term is having highest frequency

The word cloud here it is able to print ascending which it will start from low frequency to high frequency.
So, you are able to see the size difference in words.

The below table will represent highest frequency.

Mercedes Frequency

	word	freq
mercedes	mercedes	297
https	https	287
tco	tco	281
mercedesbenz	mercedesbenz	115
benz	benz	93
guydealership	guydealership	76
audi	audi	74
top	top	74
ford	ford	73
power	power	72

Audi Frequency

	word	freq
audi	audi	489
https	https	158
tco	tco	158
guydealership	guydealership	138
ford	ford	129
top	top	129
source	source	127
dependable	dependable	126
land	land	125
least	least	125

Word cloud for Mercedes

Word cloud for Audi

3. Find whether the feedback is positive or negative.

	sentiment	Mercedes	Audi
8	trust	72	49
7	surprise	28	21
6	sadness	32	33
10	positive	133	87
9	negative	73	60

For both of Mercedes and Audi have positive feedback.

4. Create a visualization for above scenerios.

Mercedes Analysis
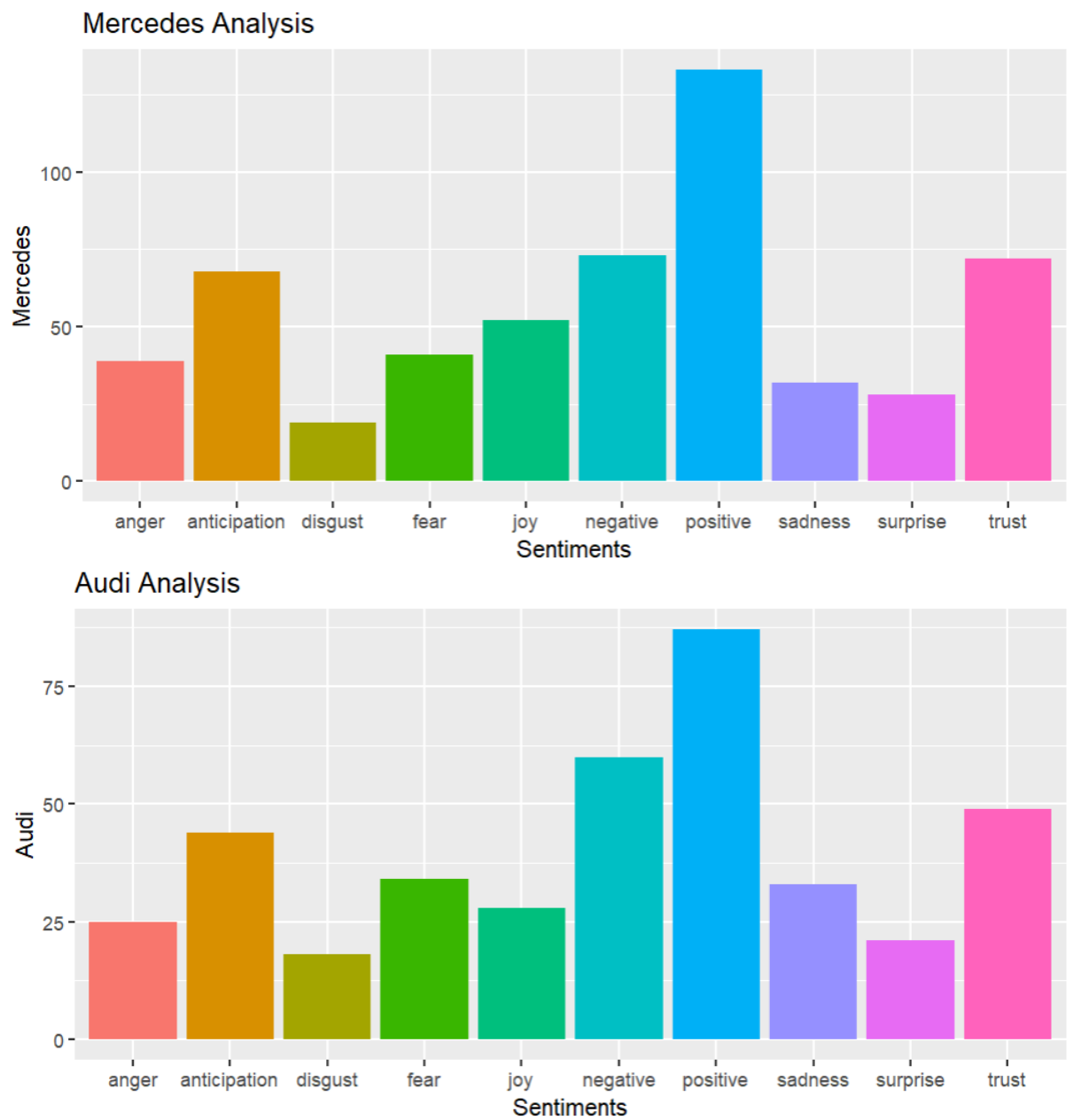
Audi Analysis

5. Analyse the data and mention 3-4 points to conclude the complete analysis.

1] The people are mostly given the positive feedback for both Mercedes and Audi.

2] As compare to other emotion, trust was high. It may arrive because of service given by both the companies.

3]The low emotion was disgust.

4]the companies are competing each other wisely. Sometimes they are differing themselves with price and their marketing strategies.

5] Most of discussion was between competing companies like ford. The people are comparing automobiles with other companies.