

Brexit Data Sentiment Analysis

Question 1

```
library(readr)
library(tm)
library(SnowballC)
library(wordcloud)
library(wordcloud2)
library(readxl)
Brexit <- read_excel("Brexit2017.xlsx")

Brexit2017=Corpus(VectorSource(Brexit$...17))

Brexit2017=tm_map(Brexit2017,content_transformer(tolower))
Brexit2017=tm_map(Brexit2017, removeWords, stopwords('english'))
Brexit2017=tm_map(Brexit2017, removePunctuation)
Brexit2017=tm_map(Brexit2017, removeNumbers)
Brexit2017=tm_map(Brexit2017, stripWhitespace)

StripString=content_transformer(function(x,pattern) gsub(pattern,"",x))
Brexit2017=tm_map(Brexit2017,StripString('[\r\n]'))
Brexit2017=tm_map(Brexit2017,StripString,'[t]')
Brexit2017=tm_map(Brexit2017, StripString, 'http[:]alnum:[]*')

Brexit2017=tm_map(Brexit2017, removewords, c('Brexit', 'brexit'))
Brexit2017=tm_map(Brexit2017,stemDocument)

View(as.matrix(TermDocumentMatrix(Brexit2017)))
DTM= DocumentTermMatrix(Brexit2017)
DTM=removeSparseTerms(DTM, 0.99)
findFreqTerms(DTM, 40)
```

```
> findFreqTerms(m, 40)
[1] "democracy" "gina" "https..." "miller" "back" "brexit"
[7] "court" "need" "news" "supreme" "article" "parliament"
[13] "say" "triggering" "british" "get" "going" "leave"
[19] "people" "voted" "bill" "parlamento" "que" "corte"
[25] "case" "must" "vote" "approval" "government" "ruled"
[31] "will" "ruling" "can" "trigger" "says" "sturgeon"
[37] "may" "rules" "httpstco..." "referendum" "uksupremecourt" "without"
[43] "..." "delay" "process" "act" "judgment" "whether"
[49] "theresa" "via" "business" "parliamentary" "mps" "start"
[55] "para" "america" "britain" "election" "elects" "hacked"
[61] "htt..." "lukeisnotsexy" "norway..." "president" "russia" "tangerine"
[67] "theres" "torn" "amp" "deal" "decision" "scotland"
[73] "trump" "supremecourt" "del" "hard" "good" "one"
[79] "today" "bbcbreaking" "todays" "support" "tories" "uklabour"
[85] "well" "labour" "give" "httpst..." "breaking" "tells"
[91] "want" "skynews" "now" "govt" "live" "just"
[97] "mean" "claimant" "see" "judges" "theresamay" "sovereignty"
[103] "majority" "remain" "lead" "final" "stop" "davis"
[109] "ukip" "boycott" "world"
```

The 100 most common words are the words before "lead" inclusive.

Question 2

```
library(syuzhet)
```

```
PosNeg=data.frame(text = sample(Brexit2017, as.character),  
stringsAsFactors = FALSE)
```

```
Sentiment=get_sentiment(PosNeg$text)
```

```
> get_sentiment(PosNeg$text)  
[1] 0.00 1.65 0.50 0.50 -0.75 0.50 0.00 0.00 0.00 -0.75 1.10 -0.75 0.00 0.50 0.00 0.00 0.50  
[18] -0.40 1.30 0.00 1.10 0.75 -0.40 -0.40 0.75 -0.40 -0.40 -0.30 -0.30 0.00 -1.50 -0.30 0.50 -0.90  
[35] -0.90 -0.40 0.00 0.25 -0.75 1.25 1.10 1.10 -1.50 0.00 -0.40 1.30 0.00 -1.10 0.00 2.10 0.25  
[52] -3.10 0.00 -0.40 0.00 -1.50 -0.10 0.00 0.50 0.00 -1.00 -0.40 -1.25 -1.25 0.00 -0.50 0.00 0.00  
[69] 0.00 0.00 0.75 0.00 1.10 1.10 -0.50 0.30 -0.75 -0.75 0.50 3.45 0.00 -1.00 0.25 1.55 1.05  
[86] 0.00 1.70 1.70 -0.40 0.15 0.50 0.00 0.65 -1.85 0.50 -1.00 0.50 0.50 0.00 0.00 0.75 0.50  
[103] -1.00 -1.00 -1.00 -1.00 0.00 0.00 0.00 0.00 0.00 0.00 -1.50 -0.50 0.75 -0.10 -0.40 -0.40 -0.40  
[120] 0.00 -1.10 0.30 -0.40 -1.00 0.00 0.50 0.50 0.00 0.95 0.25 0.25 0.25 0.80 -0.90 0.00 -1.10  
[137] -0.25 -0.90 0.00 1.65 -0.50 -0.50 -0.40 0.40 0.50 -0.50 0.10 -1.05 -0.50 0.00 0.00 0.00 0.20  
[154] 1.40 -1.30 0.00 0.00 -0.10 0.15 1.35 1.35 1.35 0.00 -1.70 0.40 -0.75 0.00 0.50 -1.00 0.75  
[171] 1.50 1.30 -1.40 0.80 0.00 -0.10 0.75 1.00 1.15 1.15 1.15 1.15 -0.40 0.25 0.00 0.00 -0.60  
[188] 0.00 0.00 -0.25 0.40 0.40 0.00 0.75 0.80 0.80 0.80 0.70 0.00 -0.50 -0.50 -0.90 -0.40 -0.75  
[205] -0.75 1.00 0.40 0.50 1.00 1.00 1.00 0.75 -0.90 0.25 0.00 -0.75 1.80 0.50 0.05 0.65 -1.50  
[222] 2.20 -0.75 0.00 -0.80 -1.10 0.00 0.40 -0.10 0.00 0.25 -0.30 -0.30 -0.40 1.30 0.00 0.70 1.15  
[239] 0.00 0.50 -1.75 -1.00 0.65 -0.40 -1.55 0.00 0.50 0.00 -1.00 0.00 -1.90 -0.35 -0.10 0.25 0.25  
[256] -1.00 0.00 0.00 0.60 -0.60 -1.00 0.60 3.45 0.00 0.50 -1.00 -0.25 0.50 0.60 1.35 0.50 0.00  
[273] -0.40 -0.55 0.75 0.00 -1.25 -1.25 -0.25 0.00 0.00 0.00 1.25 0.05 1.25 0.00 0.90 0.90 0.75  
  
[273] -0.40 -0.55 0.75 0.00 -1.25 -1.25 -0.25 0.00 0.00 0.00 1.25 0.05 1.25 0.00 0.90 0.90 0.75  
[290] 0.55 0.50 0.00 -1.10 0.10 0.00 1.90 0.75 -0.50 -1.30 0.30 0.30 -0.65 -0.40 1.35 -1.50 0.00  
[307] -0.40 -0.50 0.25 0.25 -0.70 -0.25 0.75 0.75 0.60 0.60 0.60 0.00 1.15 -1.05 0.00 0.00 0.00  
[324] 0.55 0.25 0.10 -0.40 0.40 0.00 -1.75 0.65 -0.25 1.70 1.70 0.50 0.60 1.20 0.00 0.00 1.35  
[341] 1.10 0.00 -1.60 -0.85 0.35 0.50 1.55 0.25 0.60 0.25 0.25 -1.25 -1.25 0.50 -0.50 -1.00 1.00  
[358] 1.00 -0.30 -0.30 -0.90 0.00 1.05 0.00 -1.15 0.00 -0.80 -0.80 -1.15 0.00 0.25 1.25 -0.55 0.00  
[375] 0.25 0.00 1.35 0.00 -0.45 0.00 0.70 -1.00 0.00 0.00 -0.20 1.35 1.35 1.35 0.25 -0.40 0.00  
[392] 0.00 0.40 -0.55 1.25 0.00 -0.90 -0.05 -1.90 1.00 -1.00 0.00 -0.60 0.25 -0.50 -0.50 -0.25 -0.50  
[409] -1.25 -1.10 0.00 -0.15 0.60 -0.10 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 -2.00 0.00 0.70  
[426] -0.30 0.00 0.00 -0.40 -0.40 0.00 0.75 0.50 0.75 0.00 0.00 0.00 0.50 1.15 -0.75 1.25 1.25  
[443] 1.25 1.25 1.25 2.50 1.35 -0.55 1.25 0.15 1.30 0.50 -0.40 -0.25 0.00 -0.75 1.75 -0.85 -1.00  
[460] -0.25 -0.10 0.25 -1.50 0.00 0.60 0.60 1.20 1.20 -1.50 -1.00 0.00 0.50 0.35 -1.50 0.00 0.25  
[477] -0.50 -1.00 0.00 1.50 -2.35 1.10 -0.40 0.25 0.50 0.00 -0.65 -1.00 1.35 1.50 -0.75 -0.25 -0.25  
[494] -1.10 0.00 1.85 -0.75 -0.75 0.50 -0.75 -1.60 -1.50 0.20 -1.15 -1.15 -1.15 0.00 0.25 0.25 -0.25  
[511] -0.50 0.00 0.25 0.25 0.25 0.50 -0.50 0.75 0.00 -0.50 -0.75 0.50 0.50 0.15 0.15 0.25 0.25  
[528] 0.25 0.25 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.25 -1.55 0.50 0.40 0.00 0.00 -0.40 -0.10  
[545] 1.00 1.00 0.00 0.00 0.00 0.00 0.50 0.00 0.00 0.00 0.25 -0.40 0.00 0.40 -0.40 0.25 0.25  
[562] 0.25 0.00 0.00 1.70 1.70 0.00 0.00 -1.00 -0.40 -0.65 0.75 -0.75 0.00 0.00 0.00 0.00 0.45  
  
[562] 0.25 0.00 0.00 1.70 1.70 0.00 0.00 -1.00 -0.40 -0.65 0.75 -0.75 0.00 0.00 0.00 0.00 0.45  
[579] 0.45 0.00 -1.00 0.00 -0.65 0.45 0.50 -0.10 1.00 1.00 0.00 0.00 0.00 0.00 -1.50 0.65 0.40  
[596] 0.00 0.00 -1.10 -0.55 0.10 -0.50 -0.65 1.30 -1.10 0.60 0.60 0.75 -0.75 -0.75 -1.10 -0.75 -0.75  
[613] 0.00 -1.00 0.60 0.40 0.00 0.00 0.00 0.80 1.60 1.60 1.60 0.00 0.00 -0.50 -0.90 -1.75 -1.50  
[630] -0.50 -2.00 0.00 -1.00 0.00 -1.50 0.00 0.00 -1.35 0.35 1.45 0.75 1.60 1.60 1.60 0.00 0.00  
[647] -0.25 0.50 0.00 -0.50 0.80 0.40 -2.50 1.25 -0.25 0.00 -0.90 0.00 -0.25 0.75 -0.55 0.25 0.20  
[664] -1.90 0.00 0.00 0.25 1.00 1.00 1.00 1.00 1.90 -0.90 -0.50 0.10 0.00 0.00 -0.75 -0.75 -0.75  
[681] -0.75 -0.60 -1.00 0.50 0.75 0.00 0.00 0.00 -0.75 0.25 0.25 0.40 0.40 0.10 1.35 -1.10 -0.50  
[698] 0.00 -1.30 -1.50 -0.50 -1.30 -0.10 -0.10 -0.10 -0.40 0.00 -0.40 -0.40 -0.75 -0.25 0.00 -0.65 1.10  
[715] 1.60 0.00 -0.40 0.90 0.00 0.00 0.75 -1.00 0.50 0.50 0.00 0.50 -0.90 0.10 0.00 -1.10 -0.75  
[732] -0.90 -0.85 0.00 0.50 -1.00 -0.25 0.00 -0.40 0.00 1.75 0.75 -0.70 -1.10 -0.50 0.00 0.00 0.40  
[749] -0.60 0.00 1.00 1.00 0.45 -0.25 -0.25 -1.50 0.00 -0.25 0.60 0.65 -0.75 -0.75 -0.75 -0.75 -1.00  
[766] 0.50 0.50 0.00 0.00 -0.10 -1.00 -0.90 -1.35 -0.25 -0.50 0.00 0.00 0.00 0.00 0.50 -1.00 1.00  
[783] 0.00 0.25 -0.10 -0.10 -0.10 -0.95 1.10 -0.95 1.00 0.50 0.50 -0.30 1.90 -1.00 0.00 0.00 1.30  
[800] 1.30 0.65 0.00 0.00 0.00 -0.30 -0.75 0.35 1.25 -0.75 0.50 -2.00 0.00 0.60 0.60 -0.25 0.00  
[817] 0.00 1.00 1.00 -1.50 -0.75 -0.10 0.00 -0.40 -0.40 -0.85 -1.00 0.00 0.25 -0.50 -1.00 0.80 -0.25  
[834] 0.00 0.00 0.00 0.00 -1.50 -1.75 0.00 0.25 0.00 0.50 0.60 -1.05 -1.05 -1.05 -1.00 -0.10 -0.65  
[851] 0.00 0.00 0.00 -0.10 -0.10 -0.10 -1.00 -1.60 -1.60 0.80 1.65 3.45 1.25 0.00 -1.10 0.00 -1.75
```

```
[766]  0.50  0.50  0.00  0.00 -0.10 -1.00 -0.90 -1.35 -0.25 -0.50  0.00  0.00  0.00  0.00  0.50 -1.00  1.00
[783]  0.00  0.25 -0.10 -0.10 -0.10 -0.95  1.10 -0.95  1.00  0.50  0.50 -0.30  1.90 -1.00  0.00  0.00  1.30
[800]  1.30  0.65  0.00  0.00  0.00 -0.30 -0.75  0.35  1.25 -0.75  0.50 -2.00  0.00  0.60  0.60 -0.25  0.00
[817]  0.00  1.00  1.00 -1.50 -0.75 -0.10  0.00 -0.40 -0.40 -0.85 -1.00  0.00  0.25 -0.50 -1.00  0.80 -0.25
[834]  0.00  0.00  0.00  0.00 -1.50 -1.75  0.00  0.25  0.00  0.50  0.60 -1.05 -1.05 -1.05 -1.00 -0.10 -0.65
[851]  0.00  0.00  0.00 -0.10 -0.10 -0.10 -1.00 -1.60 -1.60  0.80  1.65  3.45  1.25  0.00 -1.10  0.00 -1.75
[868]  0.75  0.50  0.50 -1.20  0.00  0.00  1.10  1.00  1.00  0.60  1.70  1.70 -1.00 -0.75  1.00  1.35  1.10
[885] -1.50  2.10 -1.75  0.15  0.80  0.50  0.50 -0.25 -0.75  1.70  1.70 -0.25 -1.00 -1.10  0.50 -0.75  1.00
[902] -1.10  0.25  0.30  0.30  0.00 -0.55  0.25  0.25  0.25  0.25  0.25  0.25 -0.40  0.00  1.30  1.30  1.35
[919]  1.35  1.35  1.35  0.25  0.50  0.50  0.25  0.25  0.25 -0.40  0.10  1.30  0.00  0.00  0.75  0.25  0.60
[936] -0.75 -1.50 -0.50  0.00  0.60  0.00  0.30  0.30  1.55  0.50  0.50  0.00  0.00  0.00  0.00  0.00  0.00
[953]  0.00  0.00  0.00  0.00  0.90  0.00 -0.25  0.25  0.25  0.00  1.00  0.50  0.75 -1.00 -0.75  0.00 -0.40
[970] -1.00  0.50  0.75  0.50 -0.80 -1.85 -1.85  0.50  1.35  0.50  0.50  0.65  0.50  0.25  1.25  1.55 -0.25
[987]  1.30 -1.10  0.40 -1.10 -1.25  0.50  0.50 -1.10 -1.00 -1.10  0.00  1.35  0.25  0.25
```

Positive numbers are positive tweets, negative numbers are negative tweets and 0.00 are neutral tweets.

Question 3

```
BrexitPolarity<- BrexitPolarity[order(BrexitPolarity),]
view(BrexitPolarity)
```

```
Top10Neg=head(BrexitPolarity, 10)
```

```
Top10Neg
```

```
Top10Pos=tail(BrexitPolarity, 10)
```

```
Top10Pos
```

```
anger : num 0 0 1 1 0 0 0 0 0 1 ...
anticipation: num 0 0 1 2 0 1 0 0 0 0 ...
disgust : num 0 0 0 0 0 0 0 0 0 1 ...
fear : num 0 0 1 2 0 0 0 0 0 1 ...
joy : num 0 0 0 0 0 0 0 0 0 0 ...
sadness : num 0 0 0 0 1 0 0 0 0 1 ...
surprise : num 0 0 0 0 1 0 0 0 0 0 ...
trust : num 0 1 0 1 0 0 0 0 0 0 ...
negative : num 0 0 0 0 1 0 0 0 0 1 ...
positive : num 0 3 1 1 0 1 0 0 0 0 ...
```

inflamm inflating https://2drelupac logical relationship suffolkroyal https://toub... finance european view sense sticky loose agree stringtrottya nymies trying brexitbathurst highest theguardian commons european presentat... taxation. parliamento... backreturn 3 https://inwpa3w3w businesswoman 2017 progressive theguardian diegoastevard markets blamed bigger chinaewest1 kingskronke thomasbrakenoschabablands saying income has bargain britain's theatre asia charliebeckit... northern chukumauna https://tscotliff truth sh... ekashanun farmingasmallomun sueddetcheads adamsnap piers chukumauna https://tscotliff mipenycok reminder panel views trevdick airbus gordon threats https://tscotliff citizens katinatrona crackin brexites fightinghis speech... stevenagnew harkaway theresamay john others colour supporter criticize alexalex1991 parliamentorch predictions higher national final takes down hubo get moment fesaalun Course https://c4dxc9... scottpruska britannique coming live evidence standing zaccaron giving turn follow channetrecuiseisen stronger anyone jaundit drop actually listening uncinffured catsdanielhananays discuss unless gamillier may brit... labourwills thrive talk thank monkey issue spread and junes holding biggesttogether irelands lessusa blog wed virality media greenparty rule jacksondyer expansion breaking goodies die six x colleagues seriousparliament's death... clarity aprova... the... https://c0gmiv6... victory second ainclinks to workisundo inflicts...tom boxen david expansion ithe... landmark submission decimann risks to workisundo inflicts...tom boxen david brink op... point longformand expect lead mths argued bank risks to workisundo inflicts...tom boxen david blatantly remain constituents benefits brexiting merited industrial group bump muss... investors partner hearing vn sp intend substantive 50... bent see calls mean diorexio 100 international nartow lftidesimensions... pledge german challenge fuertza... anecody... https://c06b6107 confirm... breast... yeshrexit...mental... threatening country... groovymand economy test https://c03fzyfuyv trade fantastic neuberger ruled sans allifkha jacobfurling internet mps cant hannahb4l1vmp paulnuttallscup https://c03eathoh... british paper...received... pointverdict quintin falls started in... jonaqueral... https://c03e32qwt... decided undermining... interview... second... r... path campaign... https://c03f7tseuz... then first congoedprace... night heart brainfirst stand... asten... aronbanks judici... ag... lawshttps://c03f7tseuz... then first congoedprace... night heart brainfirst stand... asten... aronbanks judici... ag... betw crowning catalonia are the... anything clearly Liverpool due... drawing... absence rights angry claimant davidg525951945 mannewesdec 2sweetabouit... democracy ppar counsils disclosure priorities non regierung camlodia microphone benbradshaw businessinsu support snap goodwinmollymp former enacted nicola stewartshiseng wereundivided activation place https://c03... blueyedspulman secretary scottlands consulted will westmonstary estimated socialmedia place https://c03... blueyedspulman secretary scottlands power commonst... will holyroodday thatcher botsgutgers lords lawres bloomberg west lawyers que independence jospavard f beardyoff loss something angiemadeupillets lefts campaigners gina lying video hopes happens acappeddemocrac mittag duncanthead source britains perfect before gina bating

```
(Brexit$...17=Corpus(DirSource('Tweets')))
```



```
wordcloud(Brexit$...17, min.freq=50, max.words=100, vfont=c('gothic italian','plain'),
random.order=FALSE, scale=c(5,0.5), rot.per=.3, color=brewer.pal(8,'Dark2')[c(2,4,8)])
```

This word cloud is looking at words that have a minimum frequency of 50 words so it is much cleaner than the previous word cloud.

A word cloud visualization with the word 'brexit' in a large, dark grey, gothic-style font at the top. Below it, the words 'supremecourt' and 'article50' are written in a smaller, orange, sans-serif font. 'supremecourt' is positioned to the right of 'brexit', and 'article50' is positioned below 'brexit'.

Shows most popular hashtags used-

```
wordcloud(Brexit$...20, min.freq=50, max.words=100, vfont=c('gothic italian','plain'),
random.order=FALSE, scale=c(5,0.5), rot.per=.3, color=brewer.pal(8,'Dark2')[c(2,4,8)])
```