Project 1

Name: Yina Wei

UNI: yw2922

This project wants to find the differences between two parties and what they have in common. The methods used in the project are topic modelling, emotional clustering and basic data processing method.

Loading pachkages and data from several websites is the first step. After processing data which include deleting the rows that have empty data, adjusting the data into the appropriate format and dividing data into two subsets, the project uses the methods mentioned above to find the relationship between the speeches of presidents from different party.

Step 1

load packages

```
packages.used=c("rvest", "tibble", "qdap",
                "sentimentr", "gplots", "dplyr",
                "tm", "syuzhet", "factoextra",
                "beeswarm", "scales", "RColorBrewer",
                "RANN", "tm", "topicmodels")

# check packages that need to be installed.
packages.needed=setdiff(packages.used,
                        intersect(installed.packages()[,1],
                                  packages.used))
# install additional packages
if(length(packages.needed)>0){
  install.packages(packages.needed, dependencies = TRUE)
}

# load packages
library("rvest")
library("tibble")
library("qdap")
library("sentimentr")
library("gplots")
library("dplyr")
```

```
library("tm")
library("syuzhet")
library("factoextra")
library("beeswarm")
library("scales")
library("RColorBrewer")
library("RANN")
library("tm")
library("topicmodels")

source("../lib/plotstacked.R")
source("../lib/speechFuncs.R")
```

processing data

```
main.page <- read_html(x = "http://www.presidency.ucsb.edu/inaugurals.php")
inaug=f.speechlinks(main.page)
as.Date(inaug[,1], format="%B %e, %Y")
```

```
##  [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## [24] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
## [47] NA NA NA NA NA NA NA NA NA NA NA NA NA
```

```
inaug=inaug[-nrow(inaug),]
main.page=read_html("http://www.presidency.ucsb.edu/nomination.php")
nomin <- f.speechlinks(main.page)
nomin<-nomin[-47,]
main.page=read_html("http://www.presidency.ucsb.edu/farewell_addresses.php")
farewell <- f.speechlinks(main.page)
inaug.list=read.csv("../data/inauglist.csv", stringsAsFactors = FALSE)
nomin.list=read.csv("../data/nominlist.csv", stringsAsFactors = FALSE)
farewell.list=read.csv("../data/farewelllist.csv", stringsAsFactors = FALSE)
```

scrap the texts of speeches from the speech URLs.

```
speech.list=rbind(inaug.list, nomin.list, farewell.list)
speech.list$type=c(rep("inaug", nrow(inaug.list)),
                   rep("nomin", nrow(nomin.list)),
                   rep("farewell", nrow(farewell.list)))
speech.url=rbind(inaug, nomin, farewell)
speech.list=cbind(speech.list, speech.url)
speech.list$fulltext=NA
for(i in seq(nrow(speech.list))) {
  text <- read_html(speech.list$urls[i]) %>% # load the page
    html_nodes(".displaytext") %>% # isloate the text
    html_text() # get the text
  speech.list$fulltext[i]=text
```

```
  # Create the file name
  filename <- paste0("../data/fulltext/",
                     speech.list$type[i],
                     speech.list$File[i], "-",
                     speech.list$Term[i], ".txt")
  sink(file = filename) %>%
  cat(text)
  sink()
}
```

```
sentence.list=NULL
for(i in 1:nrow(speech.list)){
  sentences=sent_detect(speech.list$fulltext[i],
                        endmarks = c("?", ".", "!", "|",";"))
  if(length(sentences)>0){
    emotions=get_nrc_sentiment(sentences)
    word.count=word_count(sentences)
    emotions=diag(1/(word.count+0.01))%*%as.matrix(emotions)
    sentence.list=rbind(sentence.list,
                        cbind(speech.list[i,-ncol(speech.list)],
                              sentences=as.character(sentences),
                              word.count,
                              emotions,
                              sent.id=1:length(sentences)
                              )
    )
  }
}
```

```
sentence.list=
  sentence.list%>%
  filter(!is.na(word.count))
```

dividing data into two subsets

```
dem<-sentence.list[which(sentence.list[,"Party"]=="Democratic"),]
rep<-sentence.list[which(sentence.list[,"Party"]=="Republican"),]

corpus.dem=dem[2:(nrow(dem)-1), ]
sentence.pre1=dem$sentences[1:(nrow(dem)-2)]
sentence.post1=dem$sentences[3:(nrow(dem)-1)]
corpus.dem$snipets=paste(sentence.pre1, corpus.dem$sentences, sentence.post1, sep=" ")
rm.rows=(1:nrow(corpus.dem))[corpus.dem$sent.id==1]
rm.rows
```

```
##  [1]   28   65  198  359  471  561  611  671  749  816  916 1025 1099 1131
## [15] 1251 1306 1403 1464 1566 1677 1800 1905 2257 2551 2799 3136 3519 3922
## [29] 4214 4394 4624 4897 5042 5162 5351 5492 5624 5793 5869 6029 6108 6253
## [43] 6372 6587 6937 7156 7392 7609 7820 7923 7980
```

```
rm.rows=c(rm.rows, rm.rows-1)
corpus.dem=corpus.dem[-rm.rows, ]

corpus.rep=rep[2:(nrow(rep)-1), ]
sentence.pre2=rep$sentences[1:(nrow(rep)-2)]
sentence.post2=rep$sentences[3:(nrow(rep)-1)]
corpus.rep$snipets=paste(sentence.pre2, corpus.rep$sentences, sentence.post2, sep=" ")
rm.rows=(1:nrow(corpus.rep))[corpus.rep$sent.id==1]
rm.rows=c(rm.rows, rm.rows-1)
corpus.rep=corpus.rep[-rm.rows, ]
```

Text mining

```
docs.dem <- Corpus(VectorSource(corpus.dem$snipets))
docs.rep <- Corpus(VectorSource(corpus.rep$snipets))
writeLines(as.character(docs.dem[[sample(1:nrow(corpus.dem), 1)]]))
```

```
## We are provincials no longer. The tragic events of the thirty months of vital turmoil through which
```

```
writeLines(as.character(docs.rep[[sample(1:nrow(corpus.rep), 1)]]))
```

```
## In the President's hometown of Chicago, more than 2,000 people have been the victim of shootings this
```

Text basic processing

```
#remove potentially problematic symbols
docs.dem <-tm_map(docs.dem,content_transformer(tolower))
docs.rep <-tm_map(docs.rep,content_transformer(tolower))

#remove punctuation
docs.dem <- tm_map(docs.dem, removePunctuation)
docs.rep <- tm_map(docs.rep, removePunctuation)

#Strip digits
docs.dem <- tm_map(docs.dem, removeNumbers)
docs.rep <- tm_map(docs.rep, removePunctuation)

#remove stopwords
docs.dem <- tm_map(docs.dem, removeWords, stopwords("english"))
docs.rep <- tm_map(docs.rep, removeWords, stopwords("english"))

#remove whitespace
docs.dem <- tm_map(docs.dem, stripWhitespace)
docs.rep <- tm_map(docs.rep, stripWhitespace)

#Stem document
docs.dem <- tm_map(docs.dem,stemDocument)
docs.rep <- tm_map(docs.rep,stemDocument)
```

Step 2 finding the relationship

(1) Topic modeling

```
dtm.dem <- DocumentTermMatrix(docs.dem)
rownames(dtm.dem) <- paste(corpus.dem$type, corpus.dem$File,
                        corpus.dem$Term, corpus.dem$sent.id, sep="_")

rowTotals <- apply(dtm.dem , 1, sum)

dtm.dem  <- dtm.dem[rowTotals> 0, ]
corpus.dem=corpus.dem[rowTotals>0, ]

dtm.rep <- DocumentTermMatrix(docs.rep)
rownames(dtm.rep) <- paste(corpus.rep$type, corpus.rep$File,
                        corpus.rep$Term, corpus.rep$sent.id, sep="_")

rowTotals <- apply(dtm.rep , 1, sum)

dtm.rep  <- dtm.rep[rowTotals> 0, ]
corpus.rep=corpus.rep[rowTotals>0, ]
```

Run LDA

```
#Set parameters for Gibbs sampling
burnin <- 4000
iter <- 2000
thin <- 500
seed <-list(2003,5,63,100001,765)
nstart <- 5
best <- TRUE

# When the number of topics is 10, the result is the best.
k <- 10
```

```{r}
#Run LDA using Gibbs sampling
ldaOut <-LDA(dtm.dem, k, method="Gibbs", control=list(nstart=nstart,
                                                      seed = seed, best=best,
                                                      burnin = burnin, iter = iter,
                                                      thin=thin))
ldaOut.topics <- as.matrix(topics(ldaOut))
table(c(1:k, ldaOut.topics))
write.csv(ldaOut.topics,file=paste("../out/LDAGibbs",k,"DocsToTopics.csv"))
#top 6 terms in each topic
ldaOut.terms <- as.matrix(terms(ldaOut,20))
write.csv(ldaOut.terms,file=paste("../out/LDAGibbs",k,"TopicsToTerms.csv"))
#probabilities associated with each topic assignment
topicProbabilities <- as.data.frame(ldaOut@gamma)
write.csv(topicProbabilities,file=paste("../out/LDAGibbs",k,"TopicProbabilities.csv"))
terms.beta=ldaOut@beta
terms.beta=scale(terms.beta)
topics.terms=NULL
for(i in 1:k){
  topics.terms=rbind(topics.terms, ldaOut@terms[order(terms.beta[i,], decreasing = TRUE)[1:7]])}
#topic terms for Democratics
topics.terms
ldaOut.terms
```

According to the results I have run in .rmd file:

```
> topics.terms
      [,1]       [,2]         [,3]       [,4]       [,5]       [,6]     [,7]
 [1,] "bless"    "abil"       "often"    "devot"    "deep"     "enjoy"  "countrymen"
 [2,] "world"    "war"        "forc"     "nuclear"  "weapon"   "east"   "attack"
 [3,] "human"    "poor"       "belief"   "repres"   "tyranni"  "minor"  "brave"
 [4,] "famili"   "care"       "get"      "pay"      "school"   "health" "bill"
 [5,] "govern"   "state"      "institut" "legisl"   "subject"  "object" "domest"
 [6,] "mani"     "two"        "past"     "littl"    "seen"     "almost" "near"
 [7,] "democrat" "republican" "truth"    "campaign" "issu"     "senat"  "run"
 [8,] "never"    "watch"      "behind"   "messag"   "legaci"   "book"   "entrust"
 [9,] "busi"     "increas"    "resourc"  "standard" "reduc"    "market" "dollar"
[10,] "new"      "togeth"     "promis"   "build"    "generat"  "goal"   "endur"
```

```
> ldaOut.terms
      Topic 1   Topic 2    Topic 3     Topic 4    Topic 5     Topic 6   Topic 7      Topic 8
 [1,] "shall"   "nation"   "peopl"     "work"     "govern"    "year"    "presid"     "will"
 [2,] "may"     "must"     "right"     "famili"   "state"     "time"    "parti"      "can"
 [3,] "citizen" "world"    "live"      "job"      "power"     "day"     "know"       "everi"
 [4,] "upon"    "peac"     "life"      "children" "interest"  "mani"    "say"        "one"
 [5,] "polit"   "war"      "men"       "help"     "constitut" "first"   "democrat"   "countri"
 [6,] "spirit"  "forc"     "believ"    "million"  "unit"      "thing"   "tonight"    "peopl"
 [7,] "faith"   "unit"     "freedom"   "make"     "union"     "last"    "want"       "now"
 [8,] "prosper" "strength" "great"     "care"     "upon"      "made"    "republican" "never"
 [9,] "hope"    "end"      "human"     "tax"      "principl"  "now"     "word"       "make"
[10,] "purpos"  "also"     "free"      "that"     "foreign"   "long"    "much"       "need"
[11,] "public"  "fear"     "nation"    "get"      "preserv"   "still"   "thank"      "way"
[12,] "great"   "arm"      "see"       "pay"      "protect"   "histori" "elect"      "take"
[13,] "respect" "danger"   "secur"     "home"     "general"   "even"    "said"       "stand"
[14,] "seek"    "home"     "equal"     "keep"     "duti"      "ever"    "truth"      "alway"
[15,] "servic"  "well"     "justic"    "economi"  "institut"  "ago"     "tell"       "person"
[16,] "part"    "polici"   "democraci" "school"   "public"    "back"    "leader"     "respons"
[17,] "confid"  "becom"    "man"       "health"   "limit"     "two"     "love"       "find"
[18,] "peopl"   "defens"   "just"      "educ"     "feder"     "hope"    "proud"      "whether"
[19,] "high"    "effort"   "differ"    "valu"     "may"       "weve"    "campaign"   "give"
[20,] "countri" "strong"   "common"    "put"      "form"      "ive"     "offic"      "serv"
      Topic 9     Topic 10
 [1,] "law"       "america"
 [2,] "busi"      "american"
 [3,] "congress"  "new"
 [4,] "mean"      "let"
 [5,] "econom"    "futur"
 [6,] "use"       "chang"
 [7,] "administr" "can"
 [8,] "control"   "togeth"
 [9,] "without"   "promis"
[10,] "act"       "come"
[11,] "increas"   "build"
[12,] "labor"     "ask"
[13,] "countri"   "better"
[14,] "industri"  "generat"
[15,] "part"      "land"
[16,] "resourc"   "good"
[17,] "system"    "today"
[18,] "nation"    "old"
[19,] "made"      "opportun"
```

```{r}

ldaOut1 <-LDA(dtm.rep, k, method="Gibbs", control=list(nstart=nstart,
                                                       seed = seed, best=best,
                                                       burnin = burnin, iter = iter,
                                                       thin=thin))
ldaOut.topics <- as.matrix(topics(ldaOut1))
table(c(1:k, ldaOut.topics))
write.csv(ldaOut.topics,file=paste("../out/LDAGibbs",k,"DocsToTopics.csv"))
#top 6 terms in each topic
ldaOut.terms <- as.matrix(terms(ldaOut1,20))
write.csv(ldaOut.terms,file=paste("../out/LDAGibbs",k,"TopicsToTerms.csv"))
#probabilities associated with each topic assignment
topicProbabilities <- as.data.frame(ldaOut1@gamma)
write.csv(topicProbabilities,file=paste("../out/LDAGibbs",k,"TopicProbabilities.csv"))
terms.beta=ldaOut1@beta
terms.beta=scale(terms.beta)
topics.terms=NULL
for(i in 1:k){
  topics.terms=rbind(topics.terms, ldaOut1@terms[order(terms.beta[i,], decreasing = TRUE)[1:7]])}
#topic terms for republicans
topics.terms
ldaOut.terms
```

```
> topics.terms
      [,1]        [,2]     [,3]       [,4]       [,5]       [,6]        [,7]
 [1,] "tell"      "dont"   "hard"     "start"    "parent"   "wont"      "weve"
 [2,] "free"      "women"  "ideal"    "achiev"   "societi"  "revolut"   "rise"
 [3,] "constitut" "execut" "enforc"   "appoint"  "honest"   "judgment"  "sentiment"
 [4,] "feder"     "cut"    "dollar"   "incom"    "billion"  "deficit"   "currenc"
 [5,] "foreign"   "labor"  "tariff"   "wage"     "commerc"  "benefit"   "island"
 [6,] "respons"   "common" "balanc"   "object"   "posit"    "process"   "broad"
 [7,] "ask"       "god"    "across"   "light"    "destini"  "road"      "uniti"
 [8,] "consequ"   "perhap" "creed"    "england"  "11th"     "transcend" "champion"
 [9,] "war"       "arm"    "threaten" "middl"    "defeat"   "hate"      "tyranni"
[10,] "presid"    "said"   "next"     "four"     "sinc"     "win"       "came"
```

```
> ldaOut.terms
      Topic 1    Topic 2    Topic 3     Topic 4    Topic 5     Topic 6      Topic 7    Topic 8
 [1,] "work"     "peopl"    "law"       "govern"   "busi"      "must"       "american" "will"
 [2,] "know"     "new"      "shall"     "tax"      "upon"      "can"        "america"  "countri"
 [3,] "just"     "free"     "right"     "system"   "polici"    "nation"     "let"      "one"
 [4,] "want"     "men"      "upon"      "need"     "product"   "great"      "live"     "make"
 [5,] "say"      "home"     "state"     "feder"    "industri"  "peopl"      "today"    "everi"
 [6,] "children" "well"     "may"       "provid"   "import"    "power"      "futur"    "now"
 [7,] "like"     "better"   "constitut" "econom"   "foreign"   "respons"    "see"      "time"
 [8,] "way"      "opportun" "order"     "economi"  "trade"     "continu"    "come"     "good"
 [9,] "famili"   "progress" "citizen"   "program"  "effect"    "administr"  "stand"    "believ"
[10,] "back"     "great"    "duti"      "propos"   "protect"   "purpos"     "ask"      "chang"
[11,] "job"      "life"     "congress"  "reduc"    "labor"     "act"        "god"      "made"
[12,] "love"     "prosper"  "question"  "public"   "issu"      "leadership" "heart"    "differ"
[13,] "thing"    "problem"  "govern"    "congress" "interest"  "use"        "promis"   "alway"
[14,] "get"      "equal"    "offic"     "secur"    "countri"   "among"      "look"     "polit"
[15,] "school"   "build"    "servic"    "million"  "necessari" "support"    "man"      "much"
[16,] "keep"     "still"    "public"    "creat"    "tariff"    "time"       "help"     "can"
[17,] "put"      "fail"     "respect"   "money"    "employ"    "effort"     "togeth"   "done"
[18,] "help"     "turn"     "execut"    "increas"  "may"       "without"    "tonight"  "mani"
[19,] "oppon"    "yet"      "elect"     "save"     "increas"   "clear"      "pledg"    "anoth"
[20,] "even"     "women"    "enforc"    "pay"      "direct"    "principl"   "spirit"   "take"
      Topic 9    Topic 10
 [1,] "world"    "year"
 [2,] "peac"     "presid"
 [3,] "nation"   "parti"
 [4,] "war"      "unit"
 [5,] "freedom"  "state"
 [6,] "never"    "republican"
 [7,] "forc"     "first"
 [8,] "caus"     "last"
 [9,] "seek"     "friend"
[10,] "liberti"  "ago"
[11,] "strength" "one"
[12,] "defens"   "past"
[13,] "human"    "accept"
[14,] "strong"   "mani"
[15,] "militari" "said"
[16,] "danger"   "trust"
[17,] "republ"   "two"
[18,] "face"     "histori"
[19,] "arm"      "say"
[20,] "preserv"  "now"
```

Summary:

For the presidents from Democratic Party, the topics include public citizens, world peace, people's interest, family, government power, laws and regulations and how to make America better.

For the presidents from Republican Party, the topics include women, liberty, opportunities, economy, world peace, foreign policy, job, tax.

Compared to Democratic Party, the presidents of Republican were more focused on the reality problems such as jobs, tax system, opportunities of economy. Also, they emphasized ethnicity, sexism, liberty.

What they have in common is that they both mentioned the topics like world peace, family, promises of making America great again and government administrations.

#(2) frequency of short or long sentences of two parties

```{r}
memberdem<-unique(dem$President)
par(mar=c(4, 11, 2, 2))
sentence.list.dem=filter(dem, File%in%memberdem)
sentence.list.dem$File=factor(sentence.list.dem$File)
sentence.list.dem$FileOrdered=reorder(sentence.list.dem$File,
                                      sentence.list.dem$word.count,
                                      mean,
                                      order=T)
beeswarm(word.count~FileOrdered,
         data=sentence.list.dem,
         horizontal = TRUE,
         pch=16, col=alpha(brewer.pal(9, "Set1"), 0.6),
         cex=0.55, cex.axis=0.8, cex.lab=0.8,
         spacing=5/nlevels(sentence.list.dem$FileOrdered),
         las=2, xlab="Number of words in a sentence.")
```

As the result showed in rmd:

```
memberrep<-unique(rep$President)
par(mar=c(4, 11, 2, 2))
sentence.list.rep=filter(rep, File%in%memberrep)
sentence.list.rep$File=factor(sentence.list.rep$File)
sentence.list.rep$FileOrdered=reorder(sentence.list.rep$File,
                                      sentence.list.rep$word.count,
                                      mean,
                                      order=T)
beeswarm(word.count~FileOrdered,
         data=sentence.list.rep,
         horizontal = TRUE,
         pch=16, col=alpha(brewer.pal(9, "Set1"), 0.6),
         cex=0.55, cex.axis=0.8, cex.lab=0.8,
         spacing=5/nlevels(sentence.list.rep$FileOrdered),
         las=2, xlab="Number of words in a sentence.")
```

as the result in rmd:

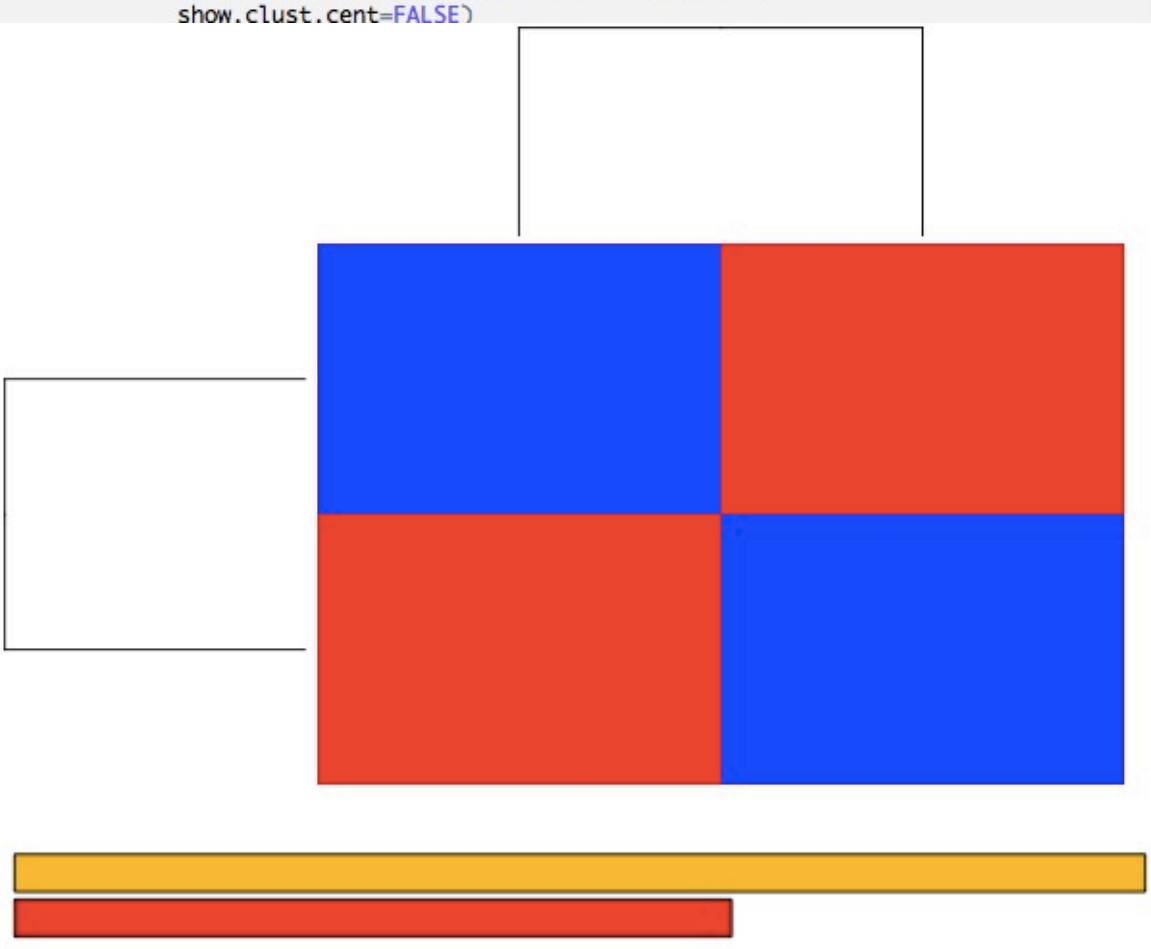
Summary:

Although they use short and long sentences randomly. In general, the presidents of Democratic tend to be more likely to use short sentences than that of Republican.

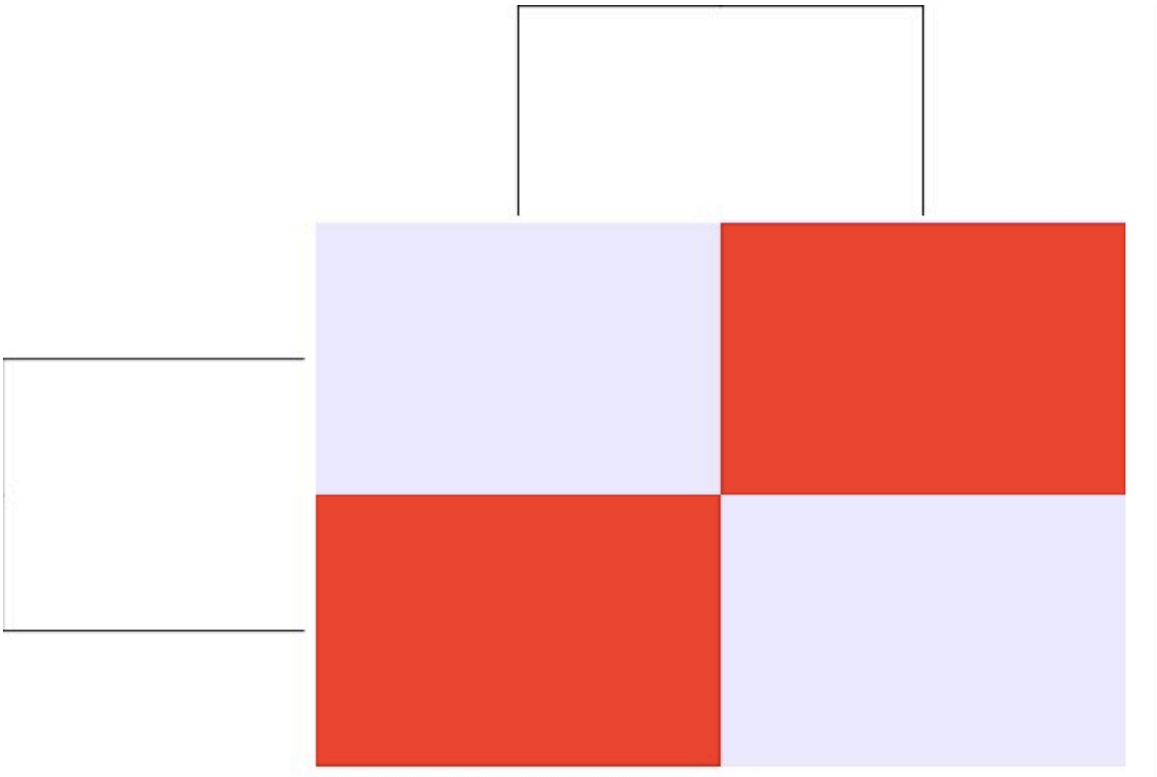
```
#(3) cluster of positive/negative
```{r, fig.width=2, fig.height=2}
heatmap.2(cor(dem%>%select(negative:positive)),
          scale = "none",
          col = bluered(100), , margin=c(2, 2), key=F,
          trace = "none", density.info = "none")
par(mar=c(4, 6, 2, 1))
emo.means=colMeans(select(dem, negative:positive)>0.01)
col.use=c("red2", "darkgoldenrod1")
barplot(emo.means[order(emo.means)], las=2, col=col.use[order(emo.means)], horiz=T)
presid.summary=tbl_df(dem)%>%
  filter(File%in%memberdem)%>%
  summarise(
    negative=mean(negative),
    positive=mean(positive))
presid.summary=as.data.frame(presid.summary)
rownames(presid.summary)=as.character((presid.summary[,1]))
km.dem=kmeans(presid.summary[,-1], iter.max=200,2)
fviz_cluster(km.dem, stand=F, repel= TRUE,
             data = presid.summary[,-1], xlab="", xaxt="n",
             show.clust.cent=FALSE)
heatmap.2(cor(rep%>%select(negative:positive)),
          scale = "none",
          col = bluered(100), , margin=c(2, 2), key=F,
          trace = "none", density.info = "none")
par(mar=c(4, 6, 2, 1))
emo.means=colMeans(select(rep, negative:positive)>0.01)
col.use=c("red2", "darkgoldenrod1")
barplot(emo.means[order(emo.means)], las=2, col=col.use[order(emo.means)], horiz=T)
presid.summary=tbl_df(rep)%>%
  filter(File%in%memberrep)%>%
  summarise(
    negative=mean(negative),
    positive=mean(positive))
presid.summary=as.data.frame(presid.summary)
rownames(presid.summary)=as.character((presid.summary[,1]))
km.rep=kmeans(presid.summary[,-1], iter.max=200,2)
fviz_cluster(km.rep, stand=F, repel= TRUE,
             data = presid.summary[,-1], xlab="", xaxt="n",
             show.clust.cent=FALSE)
```
```

3) cluster of emotions

```{r, fig.width=2, fig.height=2}
heatmap.2(cor(dem%>%select(negative:positive)),
          scale = "none",
          col = bluered(100), , margin=c(2, 2), key=F,
          trace = "none", density.info = "none")
par(mar=c(4, 6, 2, 1))
emo.means=colMeans(select(dem, negative:positive)>0.01)
col.use=c("red2", "darkgoldenrod1")
barplot(emo.means[order(emo.means)], las=2, col=col.use[order(emo.means)], horiz=T)
presid.summary=tbl_df(dem)%>%
  filter(File%in%memberdem)%>%
  summarise(
    negative=mean(negative),
    positive=mean(positive))
presid.summary=as.data.frame(presid.summary)
rownames(presid.summary)=as.character((presid.summary[,1]))
km.dem=kmeans(presid.summary[,-1], iter.max=200,2)
fviz_cluster(km.dem, stand=F, repel= TRUE,
             data = presid.summary[,-1], xlab="", xaxt="n",
             show.clust.cent=FALSE)
```



```
heatmap.2(cor(rep%>%select(negative:positive)),
          scale = "none",
          col = bluered(100), , margin=c(2, 2), key=F,
          trace = "none", density.info = "none")
par(mar=c(4, 6, 2, 1))
emo.means=colMeans(select(rep, negative:positive)>0.01)
col.use=c("red2", "darkgoldenrod1")
barplot(emo.means[order(emo.means)], las=2, col=col.use[order(emo.means)], horiz=T)
presid.summary=tbl_df(rep)%>%
  filter(File%in%memberrep)%>%
  summarise(
    negative=mean(negative),
    positive=mean(positive))
presid.summary=as.data.frame(presid.summary)
rownames(presid.summary)=as.character((presid.summary[,1]))
km.rep=kmeans(presid.summary[,-1], iter.max=200,2)
fviz_cluster(km.rep, stand=F, repel= TRUE,
             data = presid.summary[,-1], xlab="", xaxt="n",
             show.clust.cent=FALSE)
```



Summary:

The frequency of positive words versus nagative words are the same between two parties.
The presidents tend to use positive words and the frequency of the positive words is far more than that of negative words.