

ANEXO

```

#-----
#=====Scrip R=====
#-----

#=====Librerias=====

library(twitter)
library(tidyverse)
library(tidyr)
library(tidytext)
library(NLP)
library(tm)
library(SnowballC)
library(RColorBrewer)
library(wordcloud)
library(haven)
library(sqldf)
library(janeaustenr)
library(stringr)
library(sjmisc)
library(googleLanguageR)
library(readxl)
library(ggplot2)
library(stringr)
library(igraph)
library(treemapify)
#=====Api Twitter=====

api_key<-'
api_secret<-'
acces_token<-'
acces_token_secret<-'

setup_twitter_oauth (api_key,api_secret,acces_token,acces_token_secret)

#=====Filtros o reglas de busqueda=====

twettsBukele<- searchTwitter ('Gobierno de Bukele', n=5000, lang = "es",locale = "El Salvador",
maxID=NULL)

twettsEncuestaDF_Buk<-twListToDF(twettsBukele)

baseBuk<-twettsEncuestaDF_Buk

twettsNayib<- searchTwitter ('Gobierno de Nayib', n=5000, lang = "es",locale = "El Salvador",
maxID=NULL)

twettsEncuestaDF_Nay<-twListToDF(twettsNayib)

baseNay<-twettsEncuestaDF_Nay

#---Uninedo ambas bases

base<-bind_rows(baseBuk,baseNay, .id = NULL)

#=====Definiendo el corpus de la base de twitter

corpus=Corpus(VectorSource(base$text))

length(corpus)

content(corpus[[1]])

#=====Depurando la base del corpus

corpus=tm_map(corpus,removePunctuation)
corpus=tm_map(corpus,tolower)
corpus=tm_map(corpus,removeWords, stopwords (kind="es"))
corpus=tm_map(corpus,stemDocument)

#=====Construyendo la matriz del corpus para preparar la wordcloud

frecuency=DocumentTermMatrix(corpus)
frecuency
inspect(frecuency[10:15,1:15])
findFreqTerms(frecuency,lowfreq = 25)
sparse=removeSparseTerms(frecuency,0.995)
sparse
inspect(sparse[10:15,1:15])
sparse2=data.frame(as.matrix(sparse))
sparse2

#=====word cloud

```

```

positiva=subset(sparse2,sparse2$nyib>0)
positiva2=as.data.frame(colSums(positiva))
positiva2$words=row.names(positiva2)
colnames(positiva2)=c("freq","words")

wordcloud(positiva2$words,positiva2$freq,random.order = FALSE,
           colors=brewer.pal(7,"Dark2"),min.freq=1,scale=c(5,0.48),max.words = 4000)

wordcloud(positiva2$words,positiva2$freq,random.order = FALSE,
           colors=brewer.pal(9,"Set1"),min.freq=1,scale=c(4,1),max.words = 3000)

#=====Polaridad

#--libreria de polaridad: janeaustenr,dplyr,stringr
sentiments
get_sentiments("nrc")
get_sentiments("afinn")
get_sentiments("bing")
get_sentiments("loughran")

#----extrayendo lexico español
lexico<-read_sav("lexico.sav")

#----se pega con el lexico y positiva 2
basePolar<-inner_join(positiva2,lexico,by=c("words"="palabra"))

basePolar<-basePolar%>% distinct(words, .keep_all = TRUE)

basePolar<-basePolar%>%na.omit(words)

sum(basePolar$freq)

#-----Graficando

ggplot(data = basePolar)+
  geom_bar(mapping = aes(x=sentimiento,fill=sentimiento))+
  labs(title ="Frecuencia de polaridad del léxico empleado en los twitters",
        subtitle ="Muestra 8,123 twitter del 16 de Septiembre al 12 de octubre, 84 palabras claves
(Frq: 5,365)",
        x = "Sentimientos", y = "Frecuencia")

#-----Gráfico Vertical
#---tabla

tabla<-frq(basePolar,sentimiento)
tabla<-data.frame(tabla)

ggplot() +
  geom_bar(data=tabla,aes(x=reorder(val,valid.prc),y=valid.prc,fill=frq),stat='identity',position =
"dodge") +
  scale_fill_distiller(palette = "Reds")+
  labs(title ="Frecuencia de polaridad del léxico empleado en los twitters",
        subtitle ="Muestra 8,123 twitter del 16 de Septiembre al 10 de octubre, 84 palabras claves
(Frq: 5,365)",
        x = "Sentimientos", y = "Porcentajes")+
  geom_text(aes(x=tabla$val, y=tabla$valid.prc, label=tabla$valid.prc,
                hjust=ifelse(sign(tabla$valid.prc)>0,1,0)),
            position = position_dodge(width=1)) +
  scale_y_continuous() +
  coord_flip()

#ymax=a$raw.prc
#---Horizontal (no lleva coord_flip)
ggplot() +
  geom_bar(data=tabla,aes(x=reorder(val,valid.prc),y=valid.prc,fill=frq),stat='identity',position =
"dodge") +
  scale_fill_distiller(palette = "YlOrBr")+
  labs(title ="Frecuencia de polaridad del léxico empleado en los twitters",
        subtitle ="Muestra 8,123 twitter del 16 de Septiembre al 10 de octubre, 84 palabras claves
(Frq: 5,365)",
        x = "Sentimientos", y = "Porcentajes")+
  geom_text(aes(x=tabla$val, y=tabla$valid.prc, label=tabla$valid.prc,
                hjust=ifelse(sign(tabla$valid.prc)>0,0.5,0)),
            position = position_dodge(width=1)) +
  scale_y_continuous()

#=====Análisis exploratorio=====

base<-read_sav("base.sav")

dplyr::glimpse(base)

```

```

colnames(base)

as.Date.POSIXct(base$created)

#-----Por día

twitts<-base%>%
  mutate(mes_año = as.Date(created))%>%
  mutate(mes_año = format(mes_año,"%Y-%m-%d"))%>%
  group_by(screenName,mes_año)%>%
  summarise(n = n())%>%
  arrange(desc(n))

ggplot(data = twitts)+
  geom_bar(mapping = aes(x=mes_año,fill=n))+
  labs(title = "Tendencia de los twitters por fecha de análisis",
        subtitle = "Muestra 8,123 twitter del 16 Sep al 12 de Oct",
        x = "Fecha", y = "Frecuencia de Twitters")+
  theme(axis.text.x = element_text(angle = 90, hjust = 1))

twitts2<-base%>%
  mutate(mes_año = as.Date(created))%>%
  mutate(mes_año = format(mes_año,"%Y-%m-%d"))%>%
  group_by(mes_año)%>%
  summarise(n = n())%>%
  arrange(desc(n))

ggplot() +
  geom_bar(data=twitts2,aes(x=reorder(mes_año,n),y=n,fill=n),stat='identity',position = "dodge") +
  scale_fill_distiller(palette = "YlOrBr")+
  labs(title = "Tendencia de los twitters por fecha de análisis",
        subtitle = "Muestra 8,123 twitter del 16 Sep al 12 de Oct",
        x = "Sentimientos", y = "Porcentajes")+
  geom_text(aes(x=twitts2$mes_año, y=twitts2$n,label=twitts2$n,
                hjust=ifelse(sign(twitts2$n)>0,0.75,0)),
            position = position_dodge(width=1)) +
  scale_y_continuous()+
  theme(axis.text.x = element_text(angle = 90, hjust = 1))

#-----Por cuenta de usuario
twitts3<-base%>%
  mutate(mes_año = as.Date(created))%>%
  mutate(mes_año = format(mes_año,"%Y-%m-%d"))%>%
  group_by(screenName)%>%
  summarise(n = n())%>%
  arrange(desc(n))%>%
  filter(n>18)

ggplot() +
  geom_bar(data=twitts3,aes(x=reorder(screenName,n),y=n,fill=n),stat='identity',position = "dodge") +
  scale_fill_distiller(palette = "YlOrBr")+
  labs(title = "Frecuencia por arriba de 18 twitters por nombre de cuenta",
        subtitle = "Muestra 8,123 twitter del 16 Sep al 12 de Oct",
        x = "Nombre Cuenta", y = "Número de twitters")+
  geom_text(aes(x=twitts3$screenName, y=twitts3$n, label=twitts3$n,
                hjust=ifelse(sign(twitts3$n)>0,1,0)),
            position = position_dodge(width=1)) +
  scale_y_continuous() +
  coord_flip()

#---Por favoritos y retwit

twitts4<-base%>%
  group_by(screenName)%>%
  summarise(Favoritos=sum(favoriteCount),Retwiteados=sum(retweetCount))%>%
  arrange(desc(Favoritos))%>%
  filter(Favoritos>100)

ggplot() +

geom_bar(data=twitts4,aes(x=reorder(screenName,Favoritos),y=Favoritos,fill=Retwiteados),stat='identit
y',position = "dodge") +
  scale_fill_distiller(palette = "Reds")+
  labs(title = "Frecuencia de Favoritos por arriba de 100 nombre de cuenta",
        subtitle = "Muestra 8,123 twitter del 16 Sep al 12 de Oct",
        x = "Nombre Cuenta", y = "Número de twitters")+
  geom_text(aes(x=twitts4$screenName, y=twitts4$Favoritos, label=twitts4$Favoritos,
                hjust=ifelse(sign(twitts4$Favoritos)>0,1,0)),
            position = position_dodge(width=1)) +
  scale_y_continuous() +
  coord_flip()

#---Nube de palabras

```

```

nube<-positiva2%>%
  group_by(words)%>%
  arrange(desc(freq))%>%
  filter(freq>100)

ggplot() +
  geom_bar(data=nube,aes(x=reorder(words,freq),y=freq,fill=freq),stat='identity',position = "dodge")
+
  scale_fill_distiller(palette = "Reds")+
  labs(title ="Palabras por arriba de una frecuencia de 100 en los twitters analizados",
        subtitle ="Muestra 8,224 twitter del 15 al 26 de Septiembre",
        x = "Frecuencias", y = "Palabras")+
  scale_y_continuous() +
  coord_flip()

#-----Grafico de tree
ggplot(nube,
aes(fill = words,
    area = freq,
    label = words)) +
  geom_treemap() +
  geom_treemap_text(colour = "white",
                    place = "centre") +
  labs(title = "Palabras más frecuentes (8,123 twitts)") +
  theme(legend.position = "none")

#-----Gráfico de pastel

polar_pie <- ggplot(as.data.frame(basePolar), aes(x = factor(1), fill = factor(sentimiento))) +
  geom_bar(width = 1)

polar_pie + coord_polar(theta = "y") + ggtitle("Sentimiento Gobierno Nayib (100 días)", subtitle =
"Twitters del 16 de Septiembre al 12 de Octubre (27 días)") +
  ylab("Y") + xlab("X") + scale_fill_brewer(palette = "RdYlGn") +
  theme(plot.title = element_text(size=12, face='bold'))

```