Election Tweet analysis

Maryam Bappah

2023-07-19

## About Dataset

## The data was extracted from twitter (by someonelse), hence why i load it in as a csv file. The dataset has 19 rows (variables) and 6672 columns (observations). To gain insights into the language used in the tweets, we perform a text analysis. Our approach involves breaking down the tweets into individual words and examining their frequency. By counting the occurrences of each word, we can identify the top 15 most frequently used words.Additionally, i conducted a sentiment analysis to assess the sentiment expressed in the tweets. This analysis helps understand the overall emotional tone conveyed by the Twitter users. We utilize a pre-trained sentiment model to assign sentiment scores to the tweets, indicating whether they are positive, negative, or neutral.

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this: ## Load File as CSV

text <- read.csv("C:\\Users\\KYELA GADI\\Desktop\\#NigeriaElections2023 - Twitter Search ~ Twitter (1) new.csv", stringsAsFactors = FALSE)  
head(text)

## Row.ID Tweet.ID Username  
## 1 1 1.636262e+18 ElTigrei  
## 2 2 1.636261e+18 ElTigrei  
## 3 3 1.636261e+18 NewsCentralTV  
## 4 4 1.636260e+18 NewsCentralTV  
## 5 5 1.636258e+18 PeterObiUSA  
## 6 6 1.636258e+18 TheICIR  
## Tweet  
## 1 "I paid all of them (that came forward)"\n\nZero IQ. Negative EQ.\n\nTransactional monster!!! \n\n#NigeriaElections2023 #NigeriaDecides2023 #GRV4Lagos  
## 2 "Snap out of it"\n\nZero IQ. Negative EQ.\n\nMonster!!! \n\n#NigeriaElections2023 #NigeriaDecides2023 #GRV4Lagos  
## 3 Join us by 8am WAT as we evaluate the Presidential polls and INEC?s approach to the Gubernatorial elections.\n\nWatch LIVE:\n : DSTV CH 422, StarTimes CH 274, AVO CH 23\n : https://t.co/rJcgTPHuWB\n\n#Breakfastcentral #NewsCentralTV #AfricaFirst #NigeriaElections2023 #nigeriadecides https://t.co/zAhTqdTIew  
## 4 Join us by 8am WAT as we evaluate the Presidential polls and INEC?s approach to the Gubernatorial elections.\n\nWatch LIVE:\n : DSTV CH 422, StarTimes CH 274, AVO CH 23\n : https://t.co/rJcgTPHuWB\n\n#Breakfastcentral #NewsCentralTV #AfricaFirst #NigeriaElections2023 #nigeriadecides https://t.co/zAhTqdTIew  
## 5 The People?s President @PeterObi at Zion Ministry Lagos Yesterday. It is visible to the visually impaired, audible to the deaf ??? the Nigerians Nationwide Voted ? Massively for #PeterObi #NigeriaElections2023 #NigeriaDecides2023 #INECElectionResult https://t.co/W0vReCsequ? https://t.co/Ll9Seevu3f  
## 6 CLAIM: The governorship candidate of the Young Progressives Party, Sunny Ofehe, claims Delta state has the second-highest debt profile in Nigeria.\n\nVERDICT: INCORRECT\nSOURCE: Debt Management Office\n\n#NGFactCheckers #NigeriaElections2023 #DeltaDecides2023 #ICIRElectionCoverage https://t.co/9SF3XiNRNm  
## Time Tweet.Type Retweeted.by Number.of.Retweets  
## 1 3/16/2023 8:02:32 AM Tweet 0  
## 2 3/16/2023 8:00:52 AM Tweet 0  
## 3 3/16/2023 7:58:18 AM Retweet tolulopeab 1  
## 4 3/16/2023 7:56:47 AM Tweet 1  
## 5 3/16/2023 7:48:28 AM Retweet NwachukwuColl18 46  
## 6 3/16/2023 7:47:22 AM Retweet B\_otoibhi 1  
## Hashtags  
## 1 NigeriaElections2023 NigeriaDecides2023 GRV4Lagos  
## 2 NigeriaElections2023 NigeriaDecides2023 GRV4Lagos  
## 3   
## 4 Breakfastcentral NewsCentralTV AfricaFirst NigeriaElections2023 nigeriadecides  
## 5   
## 6   
## Mentions Name Location  
## 1 bRinE St. Ranger Over My Head  
## 2 bRinE St. Ranger Over My Head  
## 3 News Central TV Africa  
## 4 News Central TV Africa  
## 5 PeterObi PeterObi\_Save Nigeria Group USA New York, USA  
## 6 The ICIR Abuja, Nigeria  
## Web  
## 1 https://t.co/yHiwdRBF6r  
## 2 https://t.co/yHiwdRBF6r  
## 3 https://t.co/kYBDaJlw8D  
## 4 https://t.co/kYBDaJlw8D  
## 5   
## 6 https://t.co/2LZmiZCpL6  
## Bio  
## 1 S.O. Teric | Building a new ledge for legend  
## 2 S.O. Teric | Building a new ledge for legend  
## 3 . .\nDSTV CH 422, StarTimes CH 274, AVO CH 23\nWe are telling complete and balanced African stories...  
## 4 . .\nDSTV CH 422, StarTimes CH 274, AVO CH 23\nWe are telling complete and balanced African stories...  
## 5 Nigerians in America Diaspora support for Peter Obi Presidential Bid in 2023. #POTownsquare Space Host. https://t.co/k6jLB3CNQH  
## 6 The International Centre for Investigative Reporting (#THEICIR) is an independent, nonprofit news agency that seeks to promote transparency and accountability.  
## Number.of.Tweets Number.of.Followers Number.Following Location.Coordinates  
## 1 70332 135 135   
## 2 70332 135 135   
## 3 31096 11089 232 +07.1881+021.09375/  
## 4 31096 11089 232 +07.1881+021.09375/  
## 5 6341 82305 964 +40.71427-074.00597/  
## 6 64175 40965 217 +09.05785+007.49508/

## Load Libraries

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(stringr)

## Data Cleaning (Removing Hashtags, https, @ etc.)

remove\_reg <- "&amp;|&lt;|&gt;"  
text <- text %>%  
 mutate(Tweet = str\_remove\_all(Tweet, remove\_reg)) %>%  
 mutate(Tweet = tolower(Tweet)) %>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("@\\w+"),"" )) %>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("http\\w+"),"" )) %>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("://t.co/\\w+"),"" ))%>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("<\\w+"),"" )) %>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("/+[a-zA-Z0-9<>]+"),"")) %>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("fa[a-zA-Z0-9+<>]+"),"" )) %>%  
 mutate(Tweet = str\_replace\_all(Tweet, regex("#[a-zA-Z0-9<>]+"),"" ))  
head(text)

## Row.ID Tweet.ID Username  
## 1 1 1.636262e+18 ElTigrei  
## 2 2 1.636261e+18 ElTigrei  
## 3 3 1.636261e+18 NewsCentralTV  
## 4 4 1.636260e+18 NewsCentralTV  
## 5 5 1.636258e+18 PeterObiUSA  
## 6 6 1.636258e+18 TheICIR  
## Tweet  
## 1 "i paid all of them (that came forward)"\n\nzero iq. negative eq.\n\ntransactional monster!!! \n\n   
## 2 "snap out of it"\n\nzero iq. negative eq.\n\nmonster!!! \n\n   
## 3 join us by 8am wat as we evaluate the presidential polls and inec?s approach to the gubernatorial elections.\n\nwatch live:\n : dstv ch 422, startimes ch 274, avo ch 23\n : \n\n   
## 4 join us by 8am wat as we evaluate the presidential polls and inec?s approach to the gubernatorial elections.\n\nwatch live:\n : dstv ch 422, startimes ch 274, avo ch 23\n : \n\n   
## 5 the people?s president at zion ministry lagos yesterday. it is visible to the visually impaired, audible to the deaf ??? the nigerians nationwide voted ? massively for ?   
## 6 claim: the governorship candidate of the young progressives party, sunny ofehe, claims delta state has the second-highest debt profile in nigeria.\n\nverdict: incorrect\nsource: debt management office\n\n   
## Time Tweet.Type Retweeted.by Number.of.Retweets  
## 1 3/16/2023 8:02:32 AM Tweet 0  
## 2 3/16/2023 8:00:52 AM Tweet 0  
## 3 3/16/2023 7:58:18 AM Retweet tolulopeab 1  
## 4 3/16/2023 7:56:47 AM Tweet 1  
## 5 3/16/2023 7:48:28 AM Retweet NwachukwuColl18 46  
## 6 3/16/2023 7:47:22 AM Retweet B\_otoibhi 1  
## Hashtags  
## 1 NigeriaElections2023 NigeriaDecides2023 GRV4Lagos  
## 2 NigeriaElections2023 NigeriaDecides2023 GRV4Lagos  
## 3   
## 4 Breakfastcentral NewsCentralTV AfricaFirst NigeriaElections2023 nigeriadecides  
## 5   
## 6   
## Mentions Name Location  
## 1 bRinE St. Ranger Over My Head  
## 2 bRinE St. Ranger Over My Head  
## 3 News Central TV Africa  
## 4 News Central TV Africa  
## 5 PeterObi PeterObi\_Save Nigeria Group USA New York, USA  
## 6 The ICIR Abuja, Nigeria  
## Web  
## 1 https://t.co/yHiwdRBF6r  
## 2 https://t.co/yHiwdRBF6r  
## 3 https://t.co/kYBDaJlw8D  
## 4 https://t.co/kYBDaJlw8D  
## 5   
## 6 https://t.co/2LZmiZCpL6  
## Bio  
## 1 S.O. Teric | Building a new ledge for legend  
## 2 S.O. Teric | Building a new ledge for legend  
## 3 . .\nDSTV CH 422, StarTimes CH 274, AVO CH 23\nWe are telling complete and balanced African stories...  
## 4 . .\nDSTV CH 422, StarTimes CH 274, AVO CH 23\nWe are telling complete and balanced African stories...  
## 5 Nigerians in America Diaspora support for Peter Obi Presidential Bid in 2023. #POTownsquare Space Host. https://t.co/k6jLB3CNQH  
## 6 The International Centre for Investigative Reporting (#THEICIR) is an independent, nonprofit news agency that seeks to promote transparency and accountability.  
## Number.of.Tweets Number.of.Followers Number.Following Location.Coordinates  
## 1 70332 135 135   
## 2 70332 135 135   
## 3 31096 11089 232 +07.1881+021.09375/  
## 4 31096 11089 232 +07.1881+021.09375/  
## 5 6341 82305 964 +40.71427-074.00597/  
## 6 64175 40965 217 +09.05785+007.49508/

## tokenizing, annotating, and parsing text using pre-trained models.

library(udpipe)

## Warning: package 'udpipe' was built under R version 4.3.1

udmodel <- udpipe\_download\_model(language = "english")

## Downloading udpipe model from https://raw.githubusercontent.com/jwijffels/udpipe.models.ud.2.5/master/inst/udpipe-ud-2.5-191206/english-ewt-ud-2.5-191206.udpipe to C:/Users/KYELA GADI/Desktop/english-ewt-ud-2.5-191206.udpipe

## - This model has been trained on version 2.5 of data from https://universaldependencies.org

## - The model is distributed under the CC-BY-SA-NC license: https://creativecommons.org/licenses/by-nc-sa/4.0

## - Visit https://github.com/jwijffels/udpipe.models.ud.2.5 for model license details.

## - For a list of all models and their licenses (most models you can download with this package have either a CC-BY-SA or a CC-BY-SA-NC license) read the documentation at ?udpipe\_download\_model. For building your own models: visit the documentation by typing vignette('udpipe-train', package = 'udpipe')

## Downloading finished, model stored at 'C:/Users/KYELA GADI/Desktop/english-ewt-ud-2.5-191206.udpipe'

udmodel <- udpipe\_load\_model(file = udmodel$file\_model)  
tidy\_text <- udpipe\_annotate(udmodel, x = text$Tweet)  
tidy\_text <- as.data.frame(tidy\_text)

head(tidy\_text)

## doc\_id paragraph\_id sentence\_id sentence  
## 1 doc1 1 1 "i paid all of them (that came forward)"  
## 2 doc1 1 1 "i paid all of them (that came forward)"  
## 3 doc1 1 1 "i paid all of them (that came forward)"  
## 4 doc1 1 1 "i paid all of them (that came forward)"  
## 5 doc1 1 1 "i paid all of them (that came forward)"  
## 6 doc1 1 1 "i paid all of them (that came forward)"  
## token\_id token lemma upos xpos feats  
## 1 1 " " PUNCT `` <NA>  
## 2 2 i I PRON PRP Case=Nom|Number=Sing|Person=1|PronType=Prs  
## 3 3 paid pay VERB VBD Mood=Ind|Tense=Past|VerbForm=Fin  
## 4 4 all all DET DT <NA>  
## 5 5 of of ADP IN <NA>  
## 6 6 them they PRON PRP Case=Acc|Number=Plur|Person=3|PronType=Prs  
## head\_token\_id dep\_rel deps misc  
## 1 3 punct <NA> SpaceAfter=No  
## 2 3 nsubj <NA> <NA>  
## 3 0 root <NA> <NA>  
## 4 3 obj <NA> <NA>  
## 5 6 case <NA> <NA>  
## 6 4 nmod <NA> <NA>

## More libraries

library(tidytext)  
library(tm)

## Warning: package 'tm' was built under R version 4.3.1

## Loading required package: NLP

## Converting Tweet column to a corpus

text\_corpus <- Corpus(VectorSource(text$Tweet))  
head(text\_corpus)

## <<SimpleCorpus>>  
## Metadata: corpus specific: 1, document level (indexed): 0  
## Content: documents: 6

## Convert Corpus to lowercase

text\_corpus <- tm\_map(text\_corpus, tolower)

## Warning in tm\_map.SimpleCorpus(text\_corpus, tolower): transformation drops  
## documents

## More data cleaning (removing top words)

text\_corpus <- tm\_map(text\_corpus, removeWords,   
 c("http", "rt", "re", "amp"))

## Warning in tm\_map.SimpleCorpus(text\_corpus, removeWords, c("http", "rt", :  
## transformation drops documents

text\_corpus <- tm\_map(text\_corpus, removeWords,   
 stopwords("english"))

## Warning in tm\_map.SimpleCorpus(text\_corpus, removeWords, stopwords("english")):  
## transformation drops documents

text\_corpus <- tm\_map(text\_corpus, removePunctuation)

## Warning in tm\_map.SimpleCorpus(text\_corpus, removePunctuation): transformation  
## drops documents

text\_df <- data.frame(text\_clean = get("content", text\_corpus),   
 stringsAsFactors = FALSE)  
newtext\_df <- cbind.data.frame(text, text\_df)  
head(newtext\_df)

## Row.ID Tweet.ID Username  
## 1 1 1.636262e+18 ElTigrei  
## 2 2 1.636261e+18 ElTigrei  
## 3 3 1.636261e+18 NewsCentralTV  
## 4 4 1.636260e+18 NewsCentralTV  
## 5 5 1.636258e+18 PeterObiUSA  
## 6 6 1.636258e+18 TheICIR  
## Tweet  
## 1 "i paid all of them (that came forward)"\n\nzero iq. negative eq.\n\ntransactional monster!!! \n\n   
## 2 "snap out of it"\n\nzero iq. negative eq.\n\nmonster!!! \n\n   
## 3 join us by 8am wat as we evaluate the presidential polls and inec?s approach to the gubernatorial elections.\n\nwatch live:\n : dstv ch 422, startimes ch 274, avo ch 23\n : \n\n   
## 4 join us by 8am wat as we evaluate the presidential polls and inec?s approach to the gubernatorial elections.\n\nwatch live:\n : dstv ch 422, startimes ch 274, avo ch 23\n : \n\n   
## 5 the people?s president at zion ministry lagos yesterday. it is visible to the visually impaired, audible to the deaf ??? the nigerians nationwide voted ? massively for ?   
## 6 claim: the governorship candidate of the young progressives party, sunny ofehe, claims delta state has the second-highest debt profile in nigeria.\n\nverdict: incorrect\nsource: debt management office\n\n   
## Time Tweet.Type Retweeted.by Number.of.Retweets  
## 1 3/16/2023 8:02:32 AM Tweet 0  
## 2 3/16/2023 8:00:52 AM Tweet 0  
## 3 3/16/2023 7:58:18 AM Retweet tolulopeab 1  
## 4 3/16/2023 7:56:47 AM Tweet 1  
## 5 3/16/2023 7:48:28 AM Retweet NwachukwuColl18 46  
## 6 3/16/2023 7:47:22 AM Retweet B\_otoibhi 1  
## Hashtags  
## 1 NigeriaElections2023 NigeriaDecides2023 GRV4Lagos  
## 2 NigeriaElections2023 NigeriaDecides2023 GRV4Lagos  
## 3   
## 4 Breakfastcentral NewsCentralTV AfricaFirst NigeriaElections2023 nigeriadecides  
## 5   
## 6   
## Mentions Name Location  
## 1 bRinE St. Ranger Over My Head  
## 2 bRinE St. Ranger Over My Head  
## 3 News Central TV Africa  
## 4 News Central TV Africa  
## 5 PeterObi PeterObi\_Save Nigeria Group USA New York, USA  
## 6 The ICIR Abuja, Nigeria  
## Web  
## 1 https://t.co/yHiwdRBF6r  
## 2 https://t.co/yHiwdRBF6r  
## 3 https://t.co/kYBDaJlw8D  
## 4 https://t.co/kYBDaJlw8D  
## 5   
## 6 https://t.co/2LZmiZCpL6  
## Bio  
## 1 S.O. Teric | Building a new ledge for legend  
## 2 S.O. Teric | Building a new ledge for legend  
## 3 . .\nDSTV CH 422, StarTimes CH 274, AVO CH 23\nWe are telling complete and balanced African stories...  
## 4 . .\nDSTV CH 422, StarTimes CH 274, AVO CH 23\nWe are telling complete and balanced African stories...  
## 5 Nigerians in America Diaspora support for Peter Obi Presidential Bid in 2023. #POTownsquare Space Host. https://t.co/k6jLB3CNQH  
## 6 The International Centre for Investigative Reporting (#THEICIR) is an independent, nonprofit news agency that seeks to promote transparency and accountability.  
## Number.of.Tweets Number.of.Followers Number.Following Location.Coordinates  
## 1 70332 135 135   
## 2 70332 135 135   
## 3 31096 11089 232 +07.1881+021.09375/  
## 4 31096 11089 232 +07.1881+021.09375/  
## 5 6341 82305 964 +40.71427-074.00597/  
## 6 64175 40965 217 +09.05785+007.49508/  
## text\_clean  
## 1 paid came forward\n\nzero iq negative eq\n\ntransactional monster \n\n   
## 2 snap \n\nzero iq negative eq\n\nmonster \n\n   
## 3 join us 8am wat evaluate presidential polls inecs approach gubernatorial elections\n\nwatch live\n dstv ch 422 startimes ch 274 avo ch 23\n \n\n   
## 4 join us 8am wat evaluate presidential polls inecs approach gubernatorial elections\n\nwatch live\n dstv ch 422 startimes ch 274 avo ch 23\n \n\n   
## 5 peoples president zion ministry lagos yesterday visible visually impaired audible deaf nigerians nationwide voted massively   
## 6 claim governorship candidate young progressives party sunny ofehe claims delta state secondhighest debt profile nigeria\n\nverdict incorrect\nsource debt management office\n\n

## Load more libraries

library(ggplot2)

##   
## Attaching package: 'ggplot2'

## The following object is masked from 'package:NLP':  
##   
## annotate

library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ forcats 1.0.0 ✔ readr 2.1.4  
## ✔ lubridate 1.9.2 ✔ tibble 3.2.1  
## ✔ purrr 1.0.1 ✔ tidyr 1.3.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ ggplot2::annotate() masks NLP::annotate()  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(tidytext)

## Seperate Individual Words

tidy\_tweet <- text\_df %>%  
 unnest\_tokens(word, text\_clean)  
tidy\_tweet <- tidy\_tweet %>%  
 anti\_join(stop\_words)

## Joining with `by = join\_by(word)`

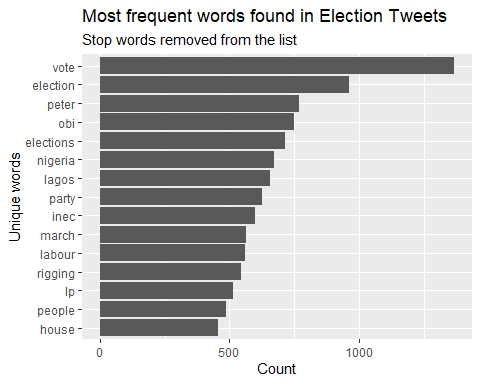
head(tidy\_tweet)

## word  
## 1 paid  
## 2 forward  
## 3 iq  
## 4 negative  
## 5 eq  
## 6 transactional

## Creating a bar chart of most frequent words

tidy\_tweet %>%   
 count(word, sort = TRUE) %>%  
 top\_n(15) %>%  
 mutate(word = reorder(word, n)) %>%  
 ggplot(aes(x = word, y = n)) +  
 geom\_col() +  
 xlab(NULL) +  
 coord\_flip() +  
 labs(y = "Count",  
 x = "Unique words",  
 title = "Most frequent words found in Election Tweets",  
 subtitle = "Stop words removed from the list")

## Selecting by n



## #Creating a sentiment bar chart

library(syuzhet)

## Warning: package 'syuzhet' was built under R version 4.3.1

# Converting tweets to ASCII to trackle strange characters  
tidy\_tweet <- iconv(tidy\_tweet, from="UTF-8", to="ASCII", sub="")  
# removing retweets, in case needed   
tidy\_tweet <-gsub("(RT|via)((?:\\b\\w\*@\\w+)+)","",tidy\_tweet)  
# removing mentions, in case needed  
tidy\_tweet <-gsub("@\\w+","",tidy\_tweet)  
ew\_sentiment<-get\_nrc\_sentiment((tidy\_tweet))

## Warning: `spread\_()` was deprecated in tidyr 1.2.0.  
## ℹ Please use `spread()` instead.  
## ℹ The deprecated feature was likely used in the syuzhet package.  
## Please report the issue to the authors.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.

sentimentscores<-data.frame(colSums(ew\_sentiment[,]))  
names(sentimentscores) <- "Score"  
sentimentscores <- cbind("sentiment"=rownames(sentimentscores),sentimentscores)  
rownames(sentimentscores) <- NULL  
ggplot(data=sentimentscores,aes(x=sentiment,y=Score))+  
 geom\_bar(aes(fill=sentiment),stat = "identity")+  
 theme(legend.position="none")+  
 xlab("Sentiments")+ylab("Scores")+  
 ggtitle("Total sentiment based on scores")+  
 theme\_minimal()

