

Twitter Data HashTag Extraction using R: Prepared for Internet Policy Africa Meet March 1 2018

Richard Ngamita

March 01 2018

```
# Twitter provides the REST search api for searching tweets from Twitter's search index. This is different than using the streaming filter API, in that the later is real-time and starts giving you results from the point of query, while the former is retrospective and will give you results from past, up to as far back as the search index goes (usually last 7 days). While the streaming API seems like the thing to use when you want to track a certain query in real time, there are situations where you may want to use the regular REST search API. You may also want to combine the two approaches, i.e. start 2 searches, one using the streaming filter API to go forward in time and one using the REST search API to go backwards # in time, in orde
```

```
# List of packages required for this Twitter Search API data analysis
```

```
pkg <- c("twitter", "ROAuth", "base64enc")
```

```
# Check if packages are not installed and assign the names of the packages not installed to the variable new.pkg
```

```
new.pkg <- pkg[!(pkg %in% installed.packages())]
```

```
# If there are any packages in the list that aren't installed, install them
```

```
if (length(new.pkg)) {  
  install.packages(new.pkg, repos = "http://cran.rstudio.com")  
}
```

```
# Load libraries for accessing the Twitter search API.
```

```
library(twitter)
```

```
library(ROAuth)
```

```
library(base64enc)
```

```
# connect to twitter api and authenticate. Select one for direct authentication go to: https://apps.twitter.com to get these credentials.
```

```
api_key <- "QKwrEqeCi8TPXNY8qGrIcZgXE"
```

```
api_secret <- "RxpELmc2qetpfNG20dkK8vYEd4GgIFXiH8zHnpGiFKLGZcVcs1"
```

```
access_token <- "3910492041-mQ8HHYqgwHmayYZR16vMzPRDMSnCwaP54eNUg1v"
```

```
access_token_secret <- "B55w0tLONnib4nI8UFi5EK6R1YaXvMLqzk0Yf7LP0WjYW"
```

```
setup_twitter_oauth(api_key, api_secret, access_token, access_token_secret)
```

```
## [1] "Using direct authentication"
```

```
# extract hashtag data. The maximum for Twitter is 3200, but we use 50 for the case of this tutorial
```

```
tweets <- searchTwitter("#mugabe", n=50)
```

```
# convert variable to data frame
```

```
tweets.df <- do.call(rbind, lapply(tweets, as.data.frame))
```

```
# save extracted data to csv file
```

```
# Windows compute d:/internet_policy.csv or right path to writable drive.
```

```
write.csv(tweets.df, "~/internet_policy.csv")
```

```
head(tweets.df, n = 20)
```

```

##
## 1 Remember this picture guys? #tbt \nLook at #Mugabe and #zuma at the l
## 2 .@ZANUPF_Official #Zimbabwe #ZanuPF Gov too arrogant to respond to deliver > via #ElephantD
## 3 .@ZANUPF_Official #Zimbabwe #ZanuPF Gov too arrogant to respond to deliver > via #ElephantD
## 4 RT @Kawasaki__900: @gavdavis @NoTolls1 @NickolausBauer @ZeldalaGrangeSA #Malema made a n
## 5 Dr Chipo Dendere speaking at the USA Foreign Affairs Committee on Zimbabwe After Mugabein
## 6 Dr Chipo Dendere speaking at the USA Foreign Affairs Committee on Zimbabwe After Mugabein
## 7 Dr Chipo Dendere speaking at the USA Foreign Affairs Committee on Zimbabwe After Mugabein
## 8 Dr Chipo Dendere speaking at the USA Foreign Affairs Committee on Zimbabwe After Mugabein
## 9 Dr Chipo Dendere speaking at the USA Foreign Affairs Committee on Zimbabwe After Mugabein
## 10 My adorable toothless puffball #Labadiah is one month old today! May the almighty God sho
## 11 RT @Kawasaki__900: @gavdavis @NoTolls1 @NickolausBauer @ZeldalaGrangeSA #Malema made a n
## 12 Dr Chipo Dendere speaking at the USA Foreign Affairs Committee on Zimbabwe After Mugabein
## 13 @ProfJNMoyo Same can be said about #Mugabe & ALL old or new #ZanuPF. Unrepentant & unap
## 14 [ #Actualité ] L'ancien président du #Zimbabwe, Robert #Mugabe, affirme que sa femme pleu
## 15 [ #Cameroun ] L'ancien président du #Zimbabwe, Robert #Mugabe, affirme que sa femme pleu
## 16 [ #Cameroun ] L'ancien président du #Zimbabwe, Robert #Mugabe, affirme que sa femme pleu
## 17 [ #Actualité ] L'ancien président du #Zimbabwe, Robert #Mugabe, affirme que sa femme pleu
## 18 RT @Kawasaki__900: @gavdavis @NoTolls1 @NickolausBauer @ZeldalaGrangeSA #Malema made a n
## 19 RT @Kawasaki__900: @gavdavis @NoTolls1 @NickolausBauer @ZeldalaGrangeSA #Malema made a n
## 20 #Zimbabwe #ZanuPF Gov too arrogant to respond to delivered > via #ElephantDawn Petition > 2
## favorited favoriteCount replyToSN created truncated
## 1 FALSE 0 <NA> 2018-03-01 08:48:16 FALSE
## 2 FALSE 3 <NA> 2018-03-01 07:49:41 TRUE
## 3 FALSE 2 <NA> 2018-03-01 07:40:48 TRUE
## 4 FALSE 0 <NA> 2018-03-01 07:17:20 FALSE
## 5 FALSE 3 povozim 2018-03-01 07:16:04 TRUE
## 6 FALSE 2 povozim 2018-03-01 07:16:03 TRUE
## 7 FALSE 2 povozim 2018-03-01 07:12:50 TRUE
## 8 FALSE 5 povozim 2018-03-01 06:50:50 TRUE
## 9 FALSE 10 <NA> 2018-03-01 06:48:11 TRUE
## 10 FALSE 3 <NA> 2018-03-01 06:46:33 TRUE
## 11 FALSE 0 <NA> 2018-03-01 06:35:45 FALSE
## 12 FALSE 20 <NA> 2018-03-01 06:33:30 TRUE
## 13 FALSE 0 ProfJNMoyo 2018-03-01 06:02:23 TRUE
## 14 FALSE 0 <NA> 2018-03-01 05:29:39 TRUE
## 15 FALSE 0 <NA> 2018-03-01 05:29:30 TRUE
## 16 FALSE 0 <NA> 2018-03-01 05:29:21 TRUE
## 17 FALSE 0 <NA> 2018-03-01 05:29:21 TRUE
## 18 FALSE 0 <NA> 2018-03-01 05:12:45 FALSE
## 19 FALSE 0 <NA> 2018-03-01 05:12:18 FALSE
## 20 FALSE 0 <NA> 2018-03-01 05:11:16 TRUE
## replyToSID id replyToUID
## 1 <NA> 969132200225574912 <NA>
## 2 <NA> 969117457830105089 <NA>
## 3 <NA> 969115222190338048 <NA>
## 4 <NA> 969109315138310144 <NA>
## 5 969108990138421248 969108994198515713 829685494514647040
## 6 969108182151950337 969108990138421248 829685494514647040
## 7 969102645964300288 969108182151950337 829685494514647040
## 8 969101979736854529 969102645964300288 829685494514647040
## 9 <NA> 969101979736854529 <NA>
## 10 <NA> 969101568426610689 <NA>
## 11 <NA> 969098848911216640 <NA>

```

```

## 12          <NA> 969098282990489600          <NA>
## 13 968962293768900614 969090451125997568      3026140439
## 14          <NA> 969082213403938816          <NA>
## 15          <NA> 969082179098677248          <NA>
## 16          <NA> 969082141496799232          <NA>
## 17          <NA> 969082139072454656          <NA>
## 18          <NA> 969077961046593536          <NA>
## 19          <NA> 969077850874810369          <NA>
## 20          <NA> 969077589544521728          <NA>
##                                     statusSource
## 1 <a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>
## 2          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 3          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 4 <a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>
## 5          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 6          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 7          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 8          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 9          <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 10         <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 11 <a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>
## 12         <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
## 13 <a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>
## 14         <a href="http://www.aucunlait.net/" rel="nofollow">aucunlait Auto post</a>
## 15         <a href="http://www.camer24.de" rel="nofollow">Camer24</a>
## 16         <a href="http://www.aucunlait.net/" rel="nofollow">aucunlait Auto post</a>
## 17         <a href="http://www.camer24.de" rel="nofollow">Camer24</a>
## 18 <a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>
## 19 <a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>
## 20         <a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>
##      screenName retweetCount isRetweet retweeted longitude latitude
## 1 Beefybeef911          0      FALSE      FALSE          NA          NA
## 2   Ingrid1Ser          2      FALSE      FALSE          NA          NA
## 3   asmithruby          1      FALSE      FALSE          NA          NA
## 4   rest_again          6       TRUE      FALSE          NA          NA
## 5     povozim           0      FALSE      FALSE          NA          NA
## 6     povozim           0      FALSE      FALSE          NA          NA
## 7     povozim           0      FALSE      FALSE          NA          NA
## 8     povozim           1      FALSE      FALSE          NA          NA
## 9     povozim           6      FALSE      FALSE          NA          NA
## 10 voziedwards          0      FALSE      FALSE          NA          NA
## 11     MWT_CT           6       TRUE      FALSE          NA          NA
## 12     povozim           7      FALSE      FALSE          NA          NA
## 13 MableKayiya          0      FALSE      FALSE          NA          NA
## 14   aucunlait          0      FALSE      FALSE          NA          NA
## 15     Camer24          0      FALSE      FALSE          NA          NA
## 16   aucunlait          0      FALSE      FALSE          NA          NA
## 17     Camer24          0      FALSE      FALSE          NA          NA
## 18   crimemind3          6       TRUE      FALSE          NA          NA
## 19 AndreEllewee          6       TRUE      FALSE          NA          NA
## 20 SharonHoole          0      FALSE      FALSE          NA          NA

```