#####################################

# A Twitter-mining tutorial #

# Author: Bas Hofstra #

# Date: 30-10-2015 #

# Task: use twitteR package #

# to download data #

#####################################

# paste this code into your R window

### Step 2: configuring R ###

### C ###

# First run the line below and pick a mirror from which we can download R-packages

# Note: not every mirror will work: France paris will work!

chooseCRANmirror()

# Next we are going to install a few packaces

# Each of these packages is needed to configure R in the correct way

# E.g., twitteR is the package that makes us download Twitter-data

# E.g., xlsx is the package that lets us export R-data tot excel files

# Now run the syntax below, on windows you can select a few lines of

# syntax and press Ctrl+R, on Mac this is Cmd+ENTER

install.packages("twitteR")

install.packages("devtools")

install.packages("foreign")

# Now that we have downloaded the packages, we need to load them into R

# Note the library("xxx") syntax, this indicates that these packages are

# already in our library (see previous step), but that we still need to

# load them into R

# Run the syntax below :

library(twitteR)

library(foreign)

library(devtools)

install\_github("twitteR", username="geoffjentry")

# Now first go back to the tutorial

### Step 4: downloading Twitter-data ###

### F ###

# The Consumer Key (API Key), Consumer Secret (API Secret), Access Token and

# Access Token Secret need to be copied to "X" in the syntax below

# These 'passwords' are your access to the Twitter-API

# When you've set these up correctly, you can for instance make Tweets from

# this R-script

# Run the syntax below:

con <- "x"

cons <-"x"

api <- "x"

apis <- "x"

# Now that you've declared variables with the correct keys, we need to ask

# to Twitter if we've done this correctly

# run the syntax below:

setup\_twitter\_oauth(con, cons, api, apis)

# If all is well, you can choose either a 1 or 2 in the R-console, choose 1 and press enter.

# You are now ready to download twitteR-data!

# Go back to the tutorial

### K ###

# So we want to download information surrounding "#somethinyouwanthere"

# The syntax below will:

# 1. lookup these words on Twitter and download in json code to R

# 2. Will make a nice data frame from this json code

# 3. Will export the dataset to excel format (you must specify the path in the last line of syntax

# please run the three syntax lines below (change "yourpath" to your own directory)

vluchtcrisis <- searchTwitter('#somethinyouwanthere', n = 10000, retryOnRateLimit = 10000)

vluchtcrisis <- do.call("rbind", lapply(vluchtcrisis, as.data.frame))

write.csv(somethinyouwanthere, "yourpath\\somethinyouwanthere.csv")

# Go back to the tutorial

### A ###

# So now we need to obtain an edgelist from the data

# First, only keep the relevant variables (tweeters and who they retweeted to)

# We need to save a list with actors (Tweeters) and a relevant variable (platform)

# We need to save an edgelist (to whom actors tweet)

# The syntax below should work for most of the Twitter data downloaded

# Run the syntax below (change "yourpath" into your directory)

source <- somethinyouwanthere$screenName

Id <- as.data.frame(somethinyouwanthere$screenName)

tweets <- somethinyouwanthere$text

target <- gsub(".\*?@(.\*?):.\*", "\\1", tweets)

target <- gsub("\\??", "", target)

net <- as.data.frame(cbind(source, target))

net <- net[- grep(" ", net$target),]

client <- somethinyouwanthere$statusSource

client <- gsub(".\*?>(.\*?)</a>.\*", "\\1", client)

client <- gsub("Twitter for iPhone", "apple", client)

client <- gsub("Twitter for Android", "nonapple", client)

client <- gsub("Twitter Web Client", "nonapple", client)

client <- gsub("Twitter for iPad", "apple", client)

client <- gsub("TweetDeck", "nonapple", client)

client <- gsub("twitterfeed", "nonapple", client)

client <- gsub("RSScockpit", "nonapple", client)

client <- gsub("iOS", "apple", client)

client <- gsub("Tweetbot for i<U+039F>S", "nonapple", client)

client <- gsub("Facebook", "nonapple", client)

client <- gsub("Hootsuite", "nonapple", client)

client <- gsub("Twitter for Windows Phone", "nonapple", client)

client <- gsub("Engagor", "nonapple", client)

client <- gsub("Echofon", "nonapple", client)

client <- gsub("Twitter for BlackBerry", "nonapple", client)

client <- gsub("Marokko.nl", "nonapple", client)

client <- gsub("vk.nl Mobile on iOS", "apple", client)

client <- gsub("vk.nl Mobile on apple", "apple", client)

client <- gsub("OBI4wan", "nonapple", client)

client <- gsub("Tweetbot for i<U+039F>S", "nonapple", client)

client <- gsub("Twitter for Mac", "apple", client)

client <- gsub("Twitter for Windows", "nonapple", client)

client <- gsub("IFTTT", "nonapple", client)

client <- gsub("Twitter for Windows Phone", "nonapple", client)

client <- gsub("Mobile Web (M5)", "nonapple", client)

client <- gsub("Twitter for Nokia S40", "nonapple", client)

client <- gsub("Plume???for???Android", "nonapple", client)

client <- gsub("Buffer", "nonapple", client)

client <- gsub("dlvr.it", "nonapple", client)

client <- gsub("Twitter for BlackBerry???", "nonapple", client)

client <- gsub("Tweetbot for Mac", "apple", client)

client <- gsub("RoundTeam", "nonapple", client)

client <- gsub("TweetCaster for Android", "nonapple", client)

client <- gsub("Plume???for???Android", "nonapple", client)

client <- gsub("Google", "nonapple", client)

client <- gsub("Plume???for???Android", "nonapple", client)

client <- gsub("WordPress.com", "nonapple", client)

client <- gsub("Mobile Web (M5)", "nonapple", client)

client <- gsub("null", "nonapple", client)

client <- gsub("nonapple???", "nonapple", client)

client <- gsub("???", "", client)

client <- gsub("???", "", client)

client <- gsub("\\(|\\)", "", client)

client <- gsub("\\<|\\>", "", client)

client <- gsub("\\+|", "", client)

client <- gsub(" ", "", client)

client <- gsub("MobileWebM5", "nonapple", client)

client <- gsub("TweetbotforiU039FS", "nonapple", client)

client <- gsub("null", "nonapple", client)

client <- as.data.frame(client)

actors <- cbind(Id, client)

colnames(actors) <- c("Id", "client")

write.table(actors, "yourpath\\somethinyouwanthere\_actors.txt",

row.names = FALSE)

write.table(net, "yourpath\\somethinyouwanthere\_ties.txt",

row.names = FALSE)

# Go back to the tutorial