Playing with vizNetwork

Michael Puerto

11/2/2020

```
library(pacman)
```

```
## Warning: package 'pacman' was built under R version 3.6.3
```

```
p_load(tidyverse,tidytext,tm,lubridate,stringr, text2vec,jsonlite,widyr,quanteda,visNetwork,igraph,ggrap


file_names = dir("C:/Users/macia/Documents/twitch_clips/October Clips")

G_path = "C:/Users/macia/Documents/twitch_clips/October Clips/"
data = NULL


for(i in file_names){
  L_path = paste(G_path,as.character(i),sep = "")
  data_temp <- fromJSON(as.character(L_path))

  chat_temp <- data_temp$comments$message$body
  user_temp <- data_temp$comments$commenter$display_name
  date_temp <- data_temp$comments$commenter$created_at
  streamer_temp <- data_temp$streamer$name

  #------------- This was me tyring to pull out the user bages ( Moderator/subcriber etc..)
  # for(i in 1:length(data_temp$comments$message$user_badges)){
  #  id_temp = list(data_temp$comments$message$user_badges[[i]]$`_id`)
  #  id = c(test,id_temp)


   # }

  data_temp = data.frame("body"= chat_temp,"user"=user_temp,"date"=date_temp,"streamer" = streamer_temp
  #assign(paste(i,"data",sep = "_"),data_temp) # this will create new dataframes in the environment
  data = bind_rows(data,data_temp)

}


data %>% glimpse()
```

```
## Rows: 46,319
## Columns: 4
## $ body     <chr> "+2", "OMEGALUL", "OUT OF MOJO", "DOOR OMEGALUL", "OMEGALU...
## $ user     <chr> "blakeleigh321", "Humorous_Chimp", "mayodongs", "mprQQ", "...
```

```
## $ date     <chr> "2017-01-04T23:24:33.150485Z", "2013-03-29T16:35:36.802546...
## $ streamer <chr> "Jerma985", "Jerma985", "Jerma985", "Jerma985", "Jerma985"...
```

```
emote_data_bttv <- read_csv('C:/Users/macia/Documents/MSIA-19/Git/Reddit-and-Twitch/Data Collection/bttv
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   emote_name = col_character(),
##   emote_image = col_character()
## )
```

```
emote_data_ffz <- read_csv('C:/Users/macia/Documents/MSIA-19/Git/Reddit-and-Twitch/Data Collection/ffz_
```

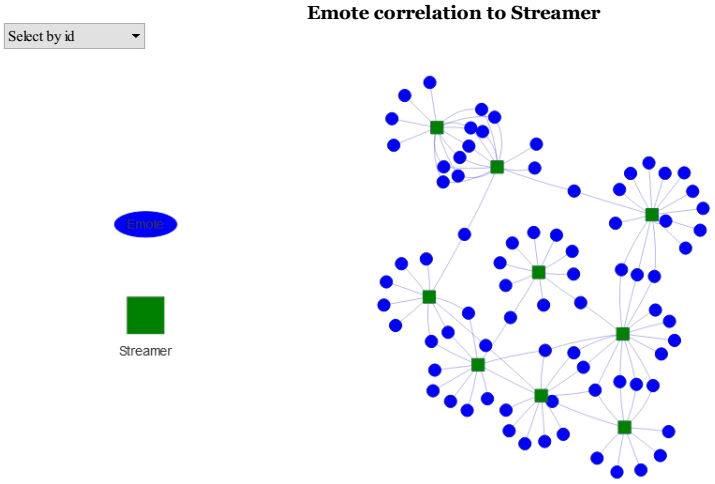
```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   emote_name = col_character(),
##   emote_link = col_character()
## )
```

```
emote_data <- emote_data_ffz %>% full_join(emote_data_bttv, by = "emote_name") %>% select(-c(X1.x,X1.y)
  mutate(emote_name = str_to_lower(emote_name))
```

```
tokens <- data %>%
  unnest_tokens(word,body)%>%
  filter(str_detect(word,"^[:alpha:]"))


word_cors <- tokens %>%
  group_by(word) %>%
  filter(n() >= 10 ) %>%
  pairwise_cor(word, streamer, sort = T)#, sort = TRUE)

top_10 <-word_cors %>% mutate("streamer" = case_when(
  item2 == "trainwreckstv" ~ "trainwreckstv", # ahh,
  item2 == "esfandtv" ~ "esfandtv",
  item2 == "forsen" ~ "forsen",
  item2 == "mizkif" ~ "mizkif",
  item2 == "ludwig" ~ "ludwig",
  item2 == "moonmoon" ~ "moonmoon",
  item2 == "xqcow" ~ "xqcow",
  item2 == "sykkuno" ~ "sykkuno",
  item2 == "vadikus007" ~ "vadikus007",
  item2 == "loltyler1" ~ "loltyler1",
  TRUE ~ "WHO?"
)) # there are 83 unique streamers in the dataset, we should filter this some how. Either top 20, or ma
# Build a scraper that grabes the names of emotes for eache of the streamers?


#top_10 %>% group_by(streamer) %>% count() # strange numbers here, each streamer has same number?, beca
 # from 2 mil rows
 # to about 2k rows
```

```
test<-top_10 %>%
  mutate(contains_emote = case_when(item1 %in% emote_data$emote_name ~ 1, TRUE ~ 0)) %>%
  filter(contains_emote == 1) %>% # filtering for only emotes!
  filter(streamer != "WHO?")%>%
  group_by(streamer) %>% top_n(10,wt = correlation) %>%
  graph_from_data_frame()# %>%
 # ggraph(layout = "fr") +
 # geom_edge_link(aes(edge_alpha = correlation), show.legend = FALSE) +
 #  geom_node_point(aes(color = streamer), size = 5) +
 #  geom_node_text(aes(label = name), repel = TRUE) +
 #  theme_void()



streamers = c("trainwreckstv","esfandtv","forsen","mizkif","ludwig","moonmoon","xqcow","sykkuno","vadiku
test_viz <- toVisNetworkData(test)
test_viz$nodes <- test_viz$nodes %>% mutate("group" = case_when(
  label %in% streamers ~ "Streamer",
  TRUE ~ "Emote"))
#test_viz checking the dataframe
visNetwork(nodes = test_viz$nodes, edges = test_viz$edges, main = "Emote correlation to Streamer")%>%
  visGroups(groupname = "Streamer", color = "green", shape = "square") %>%
  visGroups(groupname = "Emote", color = "blue")%>%
  visOptions(highlightNearest = list(enabled = T, hover = T),nodesIdSelection = T)%>%
  visLegend()
```




```
#emote_data

# Now need to figure out which words are related to the emotes?

#path_to_images <- "https://raw.githubusercontent.com/datastorm-open/datastorm-open.github.io/master/vi
#data.frame(id = 1:4,
#                    shape = c("image", "circularImage"),
#                    image = paste0(path_to_images, 1:4, ".png"),
#                    label = "I'm an image")

# Now I need to download the images? build a scraper to download images
```

What I need to do, I'd like to show, all the popular emotes that are shared among the streamers. - I need to filter the text to only include emotes, - Or I can have the emotes correlated with streamers, and the text associated with the emote!?

- Will i need to do this manually?
- Frankerfacez seems like it can be scraped.
- BTTV written in js.