

# Individual Project 2

Mamerto F. Gonzales Jr.

2022-12-06

#Individual Project 2

Instructions:

Extract 10000 tweets from Twitter using twitterR package including retweets. Subset the retweets and the original tweets into a separate file Plot the retweets and the original tweets using bar graph in vertical manner. Include legends

##Setup credentials and connect to twitter app

```
CONSUMER_SECRET <- "hFxH49SOh25H4krmBJix4FWIZ50YikL1v3t322gAowOGSsKeIa"
CONSUMER_KEY <- "r5TgXYf8eGB48dKKH18Y98Y6j"
ACCESS_SECRET <- "MkWbNGyID8ExRmhmZY8p7hStIESPZs99FVMm5KX4byaDV"
ACCESS_TOKEN <- "1595021548640882692-ETdY47BvvH5sbYGp8jHqJamfYSNYLG"
setup_twitter_oauth(consumer_key = CONSUMER_KEY,
                    consumer_secret = CONSUMER_SECRET,
                    access_token = ACCESS_TOKEN,
                    access_secret = ACCESS_SECRET)
```

##Get 10,000 observations including retweets.

```
GamingTweets <- searchTwitter("#Gaming",
                              n = 10000,
                              since = "2022-12-6", until = "2022-12-13",
                              lang = "en")

GamingTweetsDF <- twListToDF(GamingTweets)
GamingSubsets <- GamingTweetsDF %>%
  select(screenName, text, created, statusSource, isRetweet)
class(GamingSubsets)
names(GamingSubsets)
View(GamingSubsets)
head(GamingSubsets)[1:5]
head(GamingSubsets$text)[1:5]
```

##Subset the retweets and the original tweets into a separate file

```
GamingRetweet <- GamingSubsets %>%
  select(screenName, text, created, statusSource, isRetweet) %>%
  filter(isRetweet == "TRUE")
GamingNonRetweet <- GamingSubsets %>%
  select(screenName, text, created, statusSource, isRetweet) %>%
  filter(isRetweet == "FALSE")
```

##Plot the retweets and the original tweets using bar graph in vertical manner. Include legends

```
SourceEncode <- function(x) {
  if(grepl(">Twitter for iPhone</a>", x)){
    "iphone"
  }else if(grepl(">Twitter for iPad</a>", x)){
    "ipad"
  }else if(grepl(">Twitter for Android</a>", x)){
    "android"
  } else if(grepl(">Twitter Web Client</a>", x)){
    "Web"
  } else if(grepl(">Twitter for Windows Phone</a>", x)){
    "windows phone"
  }else if(grepl(">dlvr.it</a>", x)){
    "dlvr.it"
  }else if(grepl(">IFTTT</a>", x)){
    "ifttt"
  }else if(grepl(">Facebook</a>", x)){
    "facebook"
  }else {
    "others"
  }
}

GamingSubsets$tweetSource = sapply(GamingSubsets$statusSource, SourceEncode)

GamingSubsetSource <- GamingSubsets %>%
  select(tweetSource) %>%
  group_by(tweetSource) %>%
  summarize(count=n()) %>%
  arrange(desc(count))

ggplot(GamingSubsets[GamingSubsets$tweetSource != 'others',], aes(tweetSource, fill = tweetSource)) +
  geom_bar() +
  theme(legend.position="right",
        axis.title.x = element_blank(),
        axis.text.x = element_text(angle = 45, hjust = 1)) +
  ylab("Number of Tweets and Retweets") +
  ggtitle("Tweets by Source")
```

