Extract reviews of any product from ecommerce website like snapdeal and amazon

Perform sentimental analysis

library(rvest)

library(XML)

library(magrittr)

# Amazon Reviews ################

##Extracting the reviews of the Samsung galaxy s10 mobile phone

url <- "https://www.amazon.in/Samsung-Galaxy-Black-128GB-Storage/product-reviews/B07KXC7WQZ/ref=cm\_cr\_arp\_d\_paging\_btm\_next\_2?showViewpoints=1&pageNumber"

amazon\_reviews <- NULL

## Running a for loop to get the data from the ten payes of the amazon Reviews

for (i in 1:10){

murl <- read\_html(as.character(paste(url,i,sep="=")))

rev <- murl %>%

html\_nodes(".review-text") %>%

html\_text()

amazon\_reviews <- c(amazon\_reviews,rev)

}

write.table(amazon\_reviews,"galaxy.txt",row.names = F)

## in the above line we have made the csv file with the review for sentimental analysis

getwd()

#### emotion mining

install.packages("syuzhet")

library("syuzhet")

library(lubridate)

library(ggplot2)

library(scales)

library(dbplyr)

library(reshape2)

txt = readLines(file.choose())

x <- iconv(txt, "UTF-8")

s <- get\_nrc\_sentiment(x)

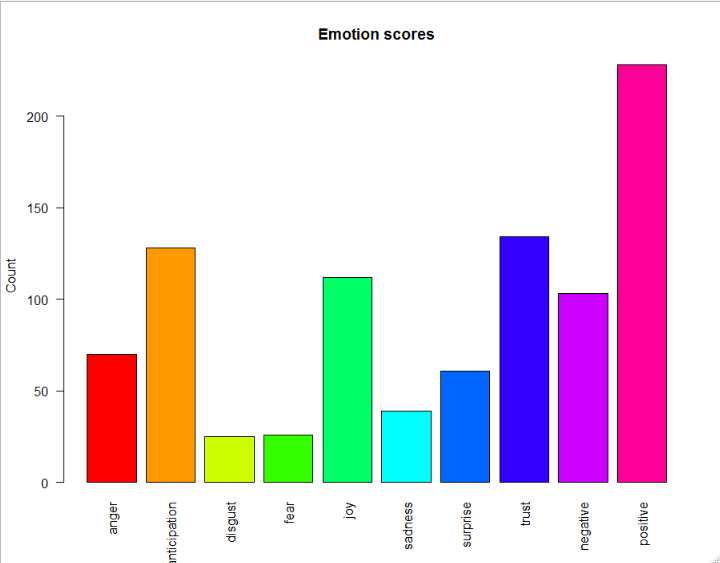
head(s)

x[4]

get\_nrc\_sentiment('excellent')

# Bar plot for emotion mining

barplot(colSums(s), las = 2, col = rainbow(10), ylab = 'Count', main = 'Emotion scores')



Inferences from the graphs

* The people are very anticipated about the product as it new release with new add features in it from there Reviews
* As the company is well known brand in the world the trust is expected from it
* And the most popular sentiment from it is the positive sentiment
* Overall we can the product galaxy s10 have very positive word of mouth

## Extracting the data from the snap deal

##Her e the product is digital analog watch

surl\_1 <- "https://www.snapdeal.com/product/skmei-black-analogdigital-watch-for/630545550665/reviews?page="

surl\_2 <- "&sortBy=HELPFUL&ratings=4,5#defRevPDP"

snapdeal\_reviews <- NULL

for (i in 1:30){

surl <- read\_html(as.character(paste(surl\_1,surl\_2,sep=as.character(i))))

srev <- surl %>%

html\_nodes("#defaultReviewsCard p") %>%

html\_text()

snapdeal\_reviews <- c(snapdeal\_reviews,srev)

}

write.table(snapdeal\_reviews,"digital.txt",row.names = FALSE)

getwd()

##sentimental analysis ##

txt = readLines(file.choose())

x <- iconv(txt, "UTF-8")

s <- get\_nrc\_sentiment(x)

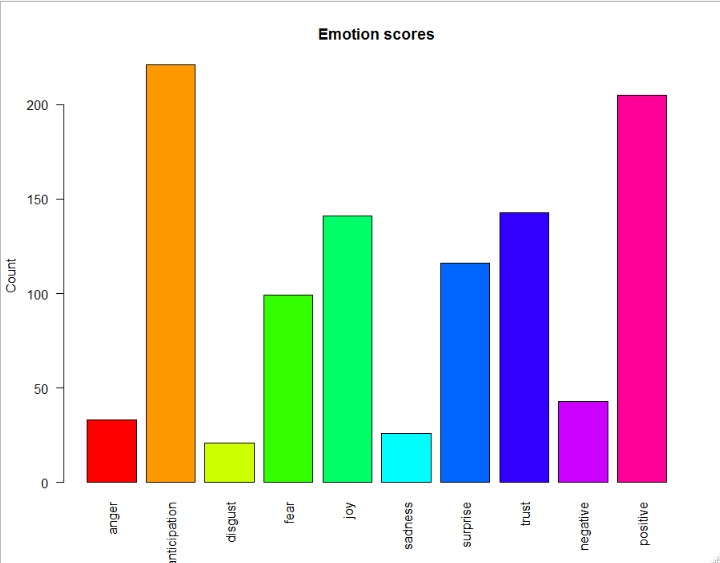
head(s)

x[4]

get\_nrc\_sentiment('excellent')

# Bar plot for emotion mining

barplot(colSums(s), las = 2, col = rainbow(10), ylab = 'Count', main = 'Emotion scores')



Inferences from the graphs

* As the product is digital analog watch the people expectation from it is more and by seeing other sentiments the product as reached the expectation of the people.
* The people included anticipation and positive words in there Reviews
* Overall sentiment for the product is positive