Text Analytics Group Assignment



**Submitted to:-**

Prof. Prerna Lal

**Submitted by (Group 4) :-**

Aastha Agarwal(19PGDM002)

Abhinandan Jain(19PGDM004)

Raja Srikrishanan(19PGDM051)

Ritu Mittal(19PGDM052)

Souvik Samanta(19PGDM064)

**Objective:-** To do sentiment analysis of the Mi mobile phones in india.

**Original Data:-**

Since the data set is quite big, we are attaching the screenshot of the original data. Please refer to the CSV file sent along this doc file.

A screenshot of a computer

Description automatically generated

**Cleaned Data:-**

second low budget mi phone first one love brand talk onedesign awesome green killer red also superbattery awesome days backup normal users day heavy userscamera good bad finedisplay good use auto brightness outdoorheating problem fine still problemsound awesome loud clearcharging fine takes time charge mah w charger usual   
## overall screen redmi pretty good considering subrs price phone besides bright screen also bigger compared saw earlier comparable redmi phones redmi comes inch hd ips display screen resolution x pixelsbattery backup good like mh battery backup goodotherwise mobile good fabulous value money comparison handset good  
## spellbound seeing value money love mi also package delivery service packing quite good   
## awesome mobile purchased mom happy surprise guys take option make selection wonderful mobile   
## good phonenice display long battery life good camera gb ram large storage decent price highly recommend  
## smoothly phone works also light weightthin easily use daily lifeimportant thing sapparine blue colour just love whenever see thank much flipkart maximum hours fast delivery think opinion redmi value money phone   
## nice phone good battery life amazing design camera better really satisfied camera quality   
## redmi good phone good processor good gb good g ram  
## stunning look camera good battery powerful performance android pie display awesome value money   
## awesom product point view redmi must buy mobile phone moderate budget people battery display ram rom performs well speaker huge crystal clear voice look hot attractive phone dominates front high budget phones color varients better emerald green varient best awesome appreciated praised whoever see   
## second low budget mi phone first one love brand talk onedesign awesome green killer red also superbattery awesome days backup normal users day heavy userscamera good bad finedisplay good use auto brightness outdoorheating problem fine still problemsound awesome loud clearcharging fine takes time charge mah w charger usual   
## overall screen redmi pretty good considering subrs price phone besides bright screen also bigger compared saw earlier comparable redmi phones redmi comes inch hd ips display screen resolution x pixelsbattery backup good like mh battery backup goodotherwise mobile good fabulous value money comparison handset good  
## spellbound seeing value money love mi also package delivery service packing quite good   
## awesome mobile purchased mom happy surprise guys take option make selection wonderful mobile   
## good phonenice display long battery life good camera gb ram large storage decent price highly recommend  
## super budget mobile like type c port mah big batterygb gb excellent storage device super  
## smoothly phone works also light weightthin easily use daily lifeimportant thing sapparine blue colour just love whenever see thank much flipkart maximum hours fast delivery think opinion redmi value money phone   
## nice phone good battery life amazing design camera better really satisfied camera quality   
## redmi good phone good processor good gb good g ram  
## stunning look camera good battery powerful performance android pie display awesome value money   
## second low budget mi phone first one love brand talk onedesign awesome green killer red also superbattery awesome days backup normal users day heavy userscamera good bad finedisplay good use auto brightness outdoorheating problem fine still problemsound awesome loud clearcharging fine takes time charge mah w charger usual   
## overall screen redmi pretty good considering subrs price phone besides bright screen also bigger compared saw earlier comparable redmi phones redmi comes inch hd ips display screen resolution x pixelsbattery backup good like mh battery backup goodotherwise mobile good fabulous value money comparison handset good  
## spellbound seeing value money love mi also package delivery service packing quite good   
## awesome mobile purchased mom happy surprise guys take option make selection wonderful mobile   
## good phonenice display long battery life good camera gb ram large storage decent price highly recommend  
## smoothly phone works also light weightthin easily use daily lifeimportant thing sapparine blue colour just love whenever see thank much flipkart maximum hours fast delivery think opinion redmi value money phone   
## nice phone good battery life amazing design camera better really satisfied camera quality   
## redmi good phone good processor good gb good g ram  
## stunning look camera good battery powerful performance android pie display awesome value money   
## awesom product point view redmi must buy mobile phone moderate budget people battery display ram rom performs well speaker huge crystal clear voice look hot attractive phone dominates front high budget phones color varients better emerald green varient best awesome appreciated praised whoever see   
## second low budget mi phone first one love brand talk onedesign awesome green killer red also superbattery awesome days backup normal users day heavy userscamera good bad finedisplay good use auto brightness outdoorheating problem fine still problemsound awesome loud clearcharging fine takes time charge mah w charger usual   
## overall screen redmi pretty good considering subrs price phone besides bright screen also bigger compared saw earlier comparable redmi phones redmi comes inch hd ips display screen resolution x pixelsbattery backup good like mh battery backup goodotherwise mobile good fabulous value money comparison handset good  
## spellbound seeing value money love mi also package delivery service packing quite good   
## awesome mobile purchased mom happy surprise guys take option make selection wonderful mobile   
## good phonenice display long battery life good camera gb ram large storage decent price highly recommend  
## super budget mobile like type c port mah big batterygb gb excellent storage device super  
## smoothly phone works also light weightthin easily use daily lifeimportant thing sapparine blue colour just love whenever see thank much flipkart maximum hours fast delivery think opinion redmi value money phone   
## nice phone good battery life amazing design camera better really satisfied camera quality   
## redmi good phone good processor good gb good g ram  
## stunning look camera good battery powerful performance android pie display awesome value money   
## ordered mobile husband liked lot nice moddle look vise also good battery back good camera quality good thank u flipkart delivery time   
## awesome need buy   
## good  
## like   
## super battery backup value money fastest delivery thank flipkart  
## lovely phone  
## product good camera good overal ok  
## super phone quality good performancesuper camera batteryvalue money   
## nice products  
## good one price slightly high  
## awesome display new type phone camera awesome performance better delivered time amazing must buy  
## excellent camera excellent performance redmi given tough competition brands pricevery good display lag free pubg finally worth money  
## giving review month usage hope helps usual budget phone redmi will definitely feel soon hold phone hand upgrade phone budget k phone will fill difference instantly use using oneplus might feel redmi tried everything give premium quality phone still lacks behind oneplus phone aspectsdisplay amoled displa  
## nice phone battery good camera good front camera excellent back just average night mode cool product perfect  
## detailed review days use k pro killer opossitivessd provides dazzling performance will work like charm even one second lagcamera best market price range pop also good enoughamoled display just amazing screen vibrant p videos will look amazing full scale display also good curve bottom filling display fullymeans clearance bottom even though   
## mobile best battery draning fast camera quality good mp zoom photo performance good  
## good photography mobile awesome features low light photography cool description must important give review phone cool phone flagship killer belive hype true master pubg users must buy phone lag free mobile fps hdr support ping mobile loved   
## says k pro overpriced nincompoopprosexcellent camera portrait mode awesomelow light camera good one plus bettersd beastbest processordesign awesomeevery colour veriant someting specialdisplay awesome can watch netflix hdl support boxapp drawer also available box running poco launcher charging time hr min cons w charger box ir  
## phone looks good unique degine fluid performancecameras expectations love triple camerasdisplay gorgeousfast charging box charger pretty good satisfiedspeakers loud enough least performance top notch snapdragon chipset also network aggression dual gpsearpeas also loud enough first impressions pretty good  
## bought boyfriend loving review using entry level smartphone since last years perception like others device overpriced using couple weeks totally changed opinion worth every single penny spent flipkart delivery got delivered th dayperformance fast score untutudisplay play ap k video   
## well awesome phone battery life just amazing got awesome display well performance camera quality good enough great product price  
## genuine review camera phone camera ok dont expect much camera buget can still give u good picturesperformance think performance handle daily task easily bht handle heavy games like pubg bht can run pubg lite smoothly also functiong goodbattery battery backup good can easily give u days battery backup easily displaydisplay good much resposive sometime lags littlr bit g  
## bought family member quite happy product supports dual sim camera quality appropriate honestly battery back good though isnt obvious mi device  
## phone value money low budget phone dont expect much good phone process application smoothly battery life becoming imo days phone design battery camera things tht can improved battery lasts long full brightness camera decent iphone camera still job done good light abd exposure will help  
## functions good phone heavy slip hand easily drawbackmi said aura grip design grip forms phone hand true   
## good  
## nice  
## bad product faced charging issue  
## phone cover accessories provided quality charger also good  
## nice  
## gifted father nd smart phone simple expectations trust beyond expectations cool buget oriented device good camera battery backup normal daily usagecamera display battery value money specs   
## best value money mobile camera awesome battery back good work smooth  
## overall good product rate thenfront camera rear camera display design performance overall good product range lack finger print sensor flipkart site page written full hd display box nthg specified know last least fast delivery flipkart thanq flipkart   
## camera good everything else just fine according price nice built quality looks display battery os processor good lite usage   
## super mobile must buy good free fire also smoothly running  
## nice product  
## vary good quality  
## nice  
## good phone budget  
## great product redmi thankyou fastest delivery expected will get product soon thankyou   
## ordered mobile husband liked lot nice moddle look vise also good battery back good camera quality good thank u flipkart delivery time   
## awesome need buy   
## good  
## like   
## super battery backup value money fastest delivery thank flipkart  
## lovely phone  
## product good camera good overal ok  
## super phone quality good performancesuper camera batteryvalue money   
## nice products  
## good one price slightly high  
## super  
## awesome product liked charging slow comes w charger use slow charging  
## really nice mobile camera glorious style   
## gifted father nd smart phone simple expectations trust beyond expectations cool buget oriented device good camera battery backup normal daily usagecamera display battery value money specs   
## best value money mobile camera awesome battery back good work smooth  
## overall good product rate thenfront camera rear camera display design performance overall good product range lack finger print sensor flipkart site page written full hd display box nthg specified know last least fast delivery flipkart thanq flipkart   
## camera good everything else just fine according price nice built quality looks display battery os processor good lite usage   
## super mobile must buy good free fire also smoothly running  
## nice product  
## vary good quality

**Data Scraping:-** Flipkart reviews are scraped out to do the analysis.

**Code:-**

## Loading libraries

library(dplyr)

library(rvest)

library(stringr)  
library(tidyverse)

library(dplyr)  
library(SnowballC)  
library(textstem)

library(tm)

library(NLP)  
library(reshape2)

library(ggplot2)  
library(tidyr)  
library(igraph)

library(ggraph)  
library(tidytext)  
library(wordcloud)

library(textdata)

## Scraping Reviews from FLIPKART

# Defining scraper function  
scrape <- function(url){  
flk <- read\_html(url)  
  
flk %>% html\_nodes("[class='\_2xg6Ul']") %>% html\_text() %>% gsub("[^\x01-\x7F]", "", .) -> ttl  
  
flk %>% html\_nodes("[class='qwjRop']") %>% html\_text() %>% gsub("READ MORE", "", .) %>% gsub("[^\x01-\x7F]", "", .)-> rv  
  
rvw <- data.frame("Title" = ttl, "Review" = rv)  
return(rvw)  
}  
  
# Generating links on which links for individual phones are found  
url <- c()  
for (i in seq(1,5)){  
 ur <- paste("https://www.flipkart.com/mobiles/mi~brand/pr?sid=tyy%2C4io&otracker=nmenu\_sub\_Electronics\_0\_Mi&page=",i) %>% gsub(" ", "", .)  
 url <- c(url,ur)  
}  
  
# Gathering urls on which scraping will be done  
url\_list <- c()  
for (i in url)  
{  
 flk <- read\_html(i)  
 ff <- flk %>% html\_nodes("[class='\_31qSD5']") %>% html\_attr("href")  
 j <- 1  
 nn <- c()  
 for (i in ff)  
 {  
 nn[j] <- paste("https://www.flipkart.com",i) %>% gsub(" ", "", .)  
 j <- j + 1  
 }  
 url\_list <- c(url\_list, nn)  
}  
  
# Defining blank data frame for capturing data  
textdata <- data.frame(Title=character(), Review=character())  
  
# Running scraper  
for (i in url\_list){  
textdata <- rbind(textdata,scrape(i))}

## Preprocessing the data frame for text data

colnames(textdata)

str(textdata)

# Load the data as a corpus  
docs <- VCorpus(VectorSource(textdata$Review))  
  
#Text transformation  
toSpace <- content\_transformer(function (x , pattern ) gsub(pattern, " ", x))  
docs <- tm\_map(docs, toSpace, "/")  
docs <- tm\_map(docs, toSpace, "@")  
docs <- tm\_map(docs, toSpace, "\\|")  
  
#Cleaning the text  
# Convert the text to lower case  
docs <- tm\_map(docs, content\_transformer(tolower))  
# Remove numbers  
docs <- tm\_map(docs, removeNumbers)  
# Remove english common stopwords  
docs <- tm\_map(docs, removeWords, stopwords("english"))  
  
# Remove punctuation  
docs <- tm\_map(docs, removePunctuation)  
# Eliminate extra white spaces  
docs <- tm\_map(docs, stripWhitespace)  
  
for(i in 1:100){  
 writeLines(as.character(docs[[i]]))  
}

# Build a term-document matrix

dtm <- TermDocumentMatrix(docs)  
m <- as.matrix(dtm)  
v <- sort(rowSums(m),decreasing=TRUE)  
d <- data.frame(word = names(v),freq=v)  
head(d, 10)

# Generate the Word cloud  
set.seed(1234)  
wordcloud(words = d$word, freq = d$freq, min.freq = 1,  
 max.words=200, random.order=FALSE, rot.per=0.35,   
 colors=brewer.pal(8, "Dark2"))

names(textdata$Review) = "text"  
textdata<-mutate(textdata,text= as.character(Review))  
data\_bigrams <- textdata %>%  
 unnest\_tokens(bigram, text, token = "ngrams", n = 2)%>%  
 count(bigram,sort=TRUE)  
head(data\_bigrams,50)

# Deleting stop words  
bigrams\_separated <- data\_bigrams %>%  
 separate(bigram, c("word1", "word2"), sep = " ")  
  
bigrams\_filtered <- bigrams\_separated %>%  
 filter(!word1 %in% stop\_words$word) %>%  
 filter(!word2 %in% stop\_words$word)  
  
# New bigram counts:  
bigram\_counts <- bigrams\_filtered %>%   
 count(word1, word2, sort = TRUE)

head(bigram\_counts,50)

## combine words  
bigrams\_united <- bigrams\_filtered %>%  
 unite(bigram, word1, word2, sep = " ")  
head(bigrams\_united, 50)

# SENTIMENT ANALYSIS

# Tokenise  
textdata$Review<-as.character(textdata$Review)  
review\_tidy <- textdata %>% unnest\_tokens(word, Review)  
# View(textdata)  
# View(review\_tidy)  
  
# Count  
review\_tidy %>% count(word) %>% arrange(desc(n)) %>% head(., 50)

# Remove stop words -  
# http://snowball.tartarus.org/algorithms/english/stop.txt  
# http://www.lextek.com/manuals/onix/stopwords1.html  
  
review\_tidy2 <- review\_tidy %>% anti\_join(stop\_words)

## Joining, by = "word"

# Count  
word\_count<-review\_tidy2 %>% count(word) %>% arrange(desc(n))  
  
# Visualization  
ggplot(word\_count, aes(x = word, y = n)) + geom\_col()

# Count  
word\_count2 <- review\_tidy2 %>% count(word) %>% filter(n>200) %>% arrange(desc(n))  
ggplot(word\_count2, aes(x = word, y = n)) + geom\_col() + ggtitle("Word Counts")

# SENTIMENT Analysis  
get\_sentiments("bing")

get\_sentiments("afinn")

get\_sentiments("loughran")

get\_sentiments("nrc")

sentiment\_review1 <- review\_tidy2 %>% inner\_join(get\_sentiments("bing"))

## Joining, by = "word"

sentiment\_review1<-sentiment\_review1 %>% count(sentiment)%>% filter(n>200)%>% arrange(desc(n))  
  
sentiment\_review2 <- review\_tidy2 %>% inner\_join(get\_sentiments("loughran"))

## Joining, by = "word"

sentiment\_review2<-sentiment\_review2 %>% count(sentiment)%>% filter(n>200)%>% arrange(desc(n))  
  
sentiment\_review3 <- review\_tidy2 %>% inner\_join(get\_sentiments("nrc"))

## Joining, by = "word"

sentiment\_review3<-sentiment\_review3 %>% count(sentiment)%>% filter(n>200)%>% arrange(desc(n))  
  
sentiment\_review3 %>% count(sentiment) %>% arrange(desc(n))

pos\_neg <- sentiment\_review3 %>% count(sentiment) %>% filter(sentiment %in% c("positive", "negative"))

pos\_neg

ggplot(sentiment\_review1, aes(x = n, y = sentiment)) + geom\_col() + ggtitle("BING library results") + xlab("Counts")

ggplot(sentiment\_review2, aes(x = n, y = sentiment)) + geom\_col() + ggtitle("Loughran library results") + xlab("Counts")

ggplot(sentiment\_review3, aes(x = n, y = sentiment)) + geom\_col() + ggtitle("NRC library results") + xlab("Counts")

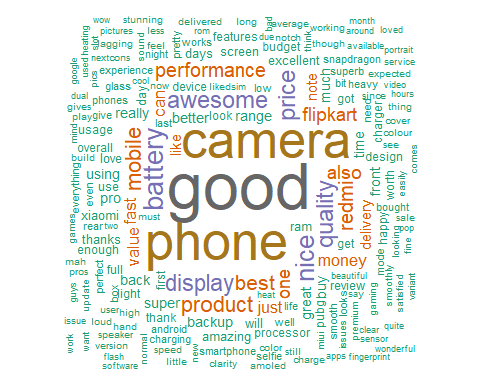
**Data analysis and Insights:-**

Document term matrix:-

(Displaying top 10 words)

## word freq  
## good good 940  
## phone phone 755  
## camera camera 722  
## battery battery 320  
## nice nice 309  
## quality quality 276  
## price price 264  
## awesome awesome 261  
## display display 250  
## mobile mobile 233

**Word Cloud:**



In the word cloud the most prominent words are good, camera, phone, display, product, awesome etc. This conveys that the these are the most talked about things/features that the in reviews.

**Bi-grams**

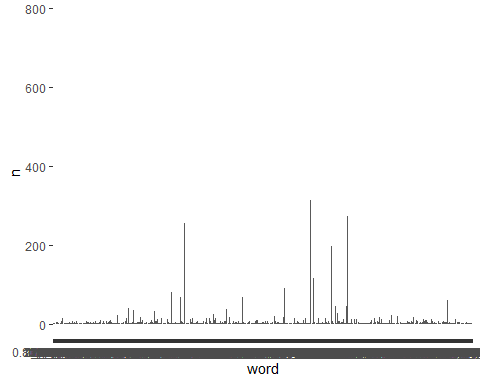
## word1 word2 n  
## 1 camera quality 94  
## 2 price range 81  
## 3 battery backup 74  
## 4 redmi note 70  
## 5 battery life 52  
## 6 front camera 46  
## 7 nice phone 45  
## 8 awesome phone 42  
## 9 rear camera 38  
## 10 fast delivery 36  
## 11 nice mobile 29  
## 12 note 7 29  
## 13 6gb ram 25  
## 14 build quality 25  
## 15 fast charger 25  
## 16 nice product 25  
## 17 note 5 25  
## 18 7 pro 23  
## 19 2 days 22  
## 20 mobile phone 22  
## 21 night mode 22  
## 22 display quality 20  
## 23 low light 20  
## 24 phone camera 20  
## 25 5 pro 19  
## 26 k20 pro 19  
## 27 amoled display 18  
## 28 google camera 18  
## 29 48 mp 17  
## 30 awesome camera 17  
## 31 selfie camera 17  
## 32 1 month 16  
## 33 mind blowing 16  
## 34 battery performance 15  
## 35 pros 1 15  
## 36 super fast 15  
## 37 excellent camera 14  
## 38 gorilla glass 14  
## 39 nice camera 14  
## 40 redmi 5a 14  
## 41 redmi 8 14  
## 42 note 4 13  
## 43 portrait mode 13  
## 44 4000 mah 12  
## 45 4gb ram 12  
## 46 charging time 12  
## 47 heavy usage 12  
## 48 mi phone 12  
## 49 superb phone 12  
## 50 white color 12

**Combined Bigrams**

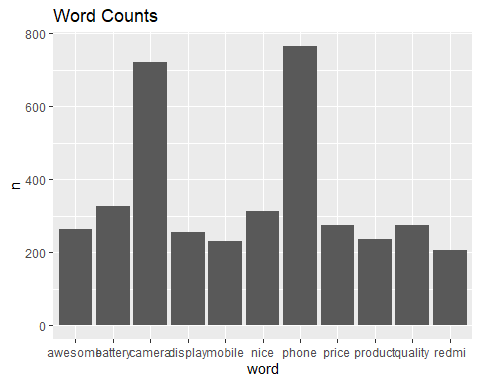
## bigram n  
## 1 camera quality 94  
## 2 price range 81  
## 3 battery backup 74  
## 4 redmi note 70  
## 5 battery life 52  
## 6 front camera 46  
## 7 nice phone 45  
## 8 awesome phone 42  
## 9 rear camera 38  
## 10 fast delivery 36  
## 11 nice mobile 29  
## 12 note 7 29  
## 13 6gb ram 25  
## 14 build quality 25  
## 15 fast charger 25  
## 16 nice product 25  
## 17 note 5 25  
## 18 7 pro 23  
## 19 2 days 22  
## 20 mobile phone 22  
## 21 night mode 22  
## 22 display quality 20  
## 23 low light 20  
## 24 phone camera 20  
## 25 5 pro 19  
## 26 k20 pro 19  
## 27 amoled display 18  
## 28 google camera 18  
## 29 48 mp 17  
## 30 awesome camera 17  
## 31 selfie camera 17  
## 32 1 month 16  
## 33 mind blowing 16  
## 34 battery performance 15  
## 35 pros 1 15  
## 36 super fast 15  
## 37 excellent camera 14  
## 38 gorilla glass 14  
## 39 nice camera 14  
## 40 redmi 5a 14  
## 41 redmi 8 14  
## 42 note 4 13  
## 43 portrait mode 13  
## 44 4000 mah 12  
## 45 4gb ram 12  
## 46 charging time 12  
## 47 heavy usage 12  
## 48 mi phone 12  
## 49 superb phone 12  
## 50 white color 12

These bigrams are giving better understanding of the most talked about features like, front camera, camera quality, battery performance, price value.

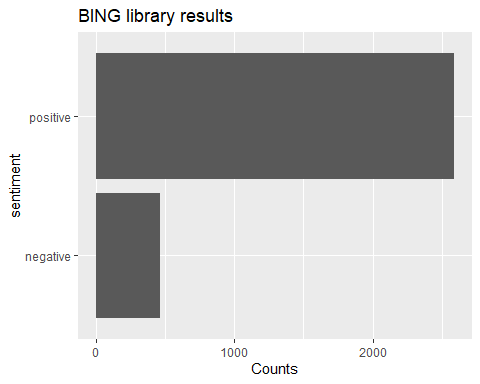
**Visualization**

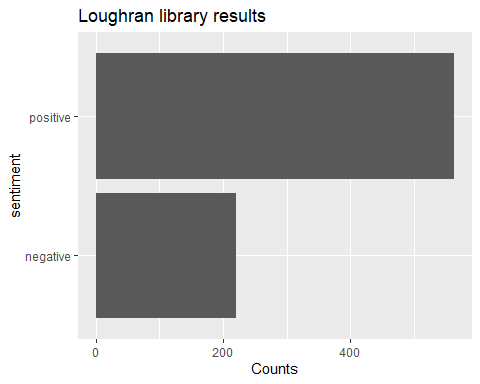


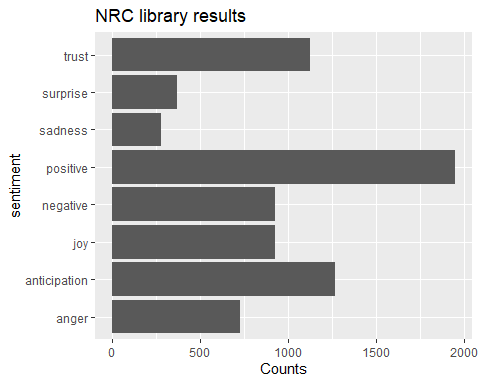
**Word count**



**Sentiment analysis**







**Inferences and Closing notes**

* The inferences regarding the analysis conducted shows that there is a reasonably large degree of positivity surrounding the reviews that discuss the MI phones. There are negative reviews as well, however they are fewer and far in between the rest of the positive reviews. There are approximately 5 positive reviews for every 2 negative reviews within the sample size (5 : 2 :: positive : negative)
* The techniques employed such as the bi-grams, word cloud and the lexicon analyses also gives us ample insights into the overall perception of the products.
  + From the word cloud inference, we are able to draw the key topics that are coming out of the reviews about features:
    - Budget – Worth, Awesome, Value for money, etc.
    - Camera quality – Rear camera as well as front-facing camera
    - Battery life – Good quality and high standby time
    - Aesthetics and Build – Pretty, premium look, lightweight, etc.
    - Performance – Snapdragon processor, fast for gaming, loud audio, etc.
  + The above sentiments look positive and are almost an exhaustive list of the most common features that are sought after by people purchasing mobile phones.
* Analysis of bigrams also reassures the interpretations drawn from the above statements. Higher frequencies are observed from bigrams such as camera quality, price range, nice/awesome phone and battery life (features like camera quality here likely indicates positiveness due to the lower price point and greater number positive reviews compared to negative showing the efficacy of the product and highlighting the value for money aspect).
* Lastly, the inferences drawn from the NRC graph indicate the emotional stance of the customers. A range of words are being used to indicate emotions that are both positive and negative. Positive emotions are encapsulated by words such as anticipation (customers eagerly awaiting product launch), trust (customers purchase the product knowing or having heard of its reliability), joy (customers display satisfaction of purchasing the products), etc. Negative emotions are captured by words like sadness (customers feel disappointed for some reason), fear (customers are skeptical or second-guessing the product and expecting issues to arise), anger (customers are experiencing difficulties or are unsatisfied with the product), etc.