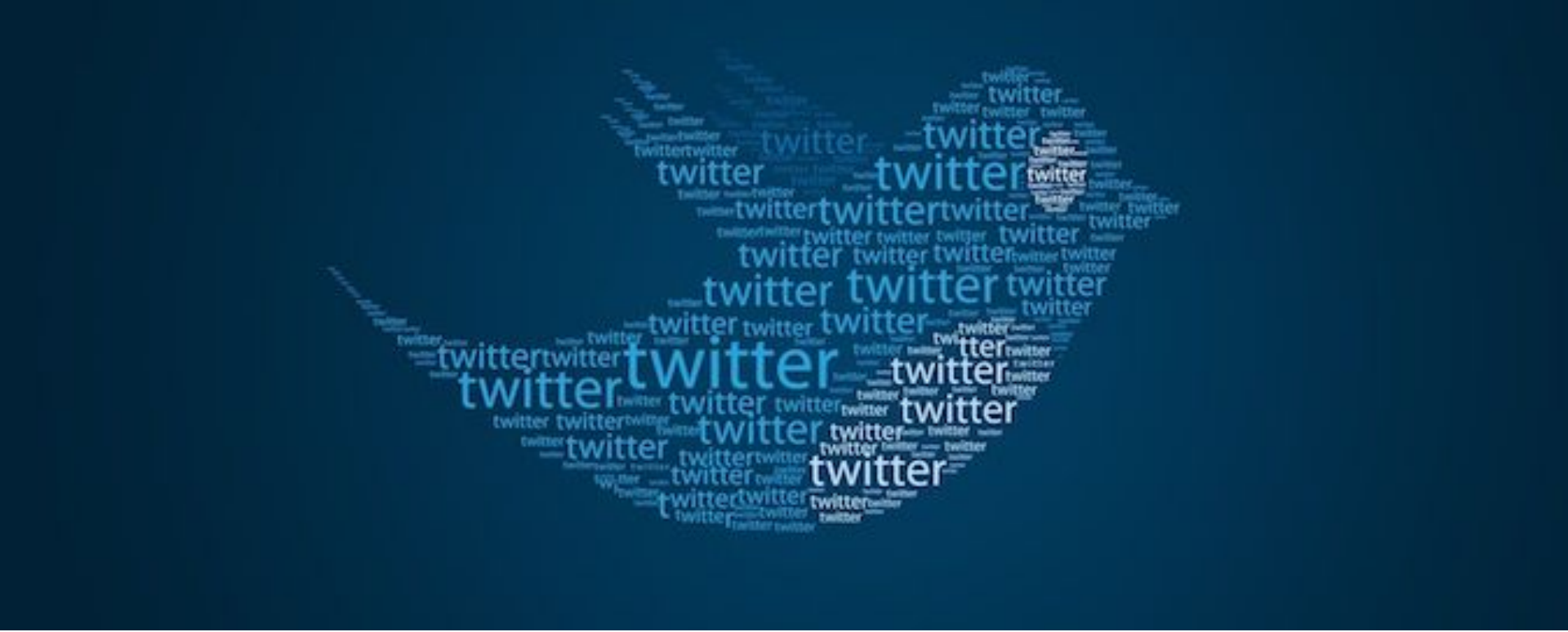


Product Trend Analysis By Mining Twitter Data

By Jaya Sistla

Using R and Tableau

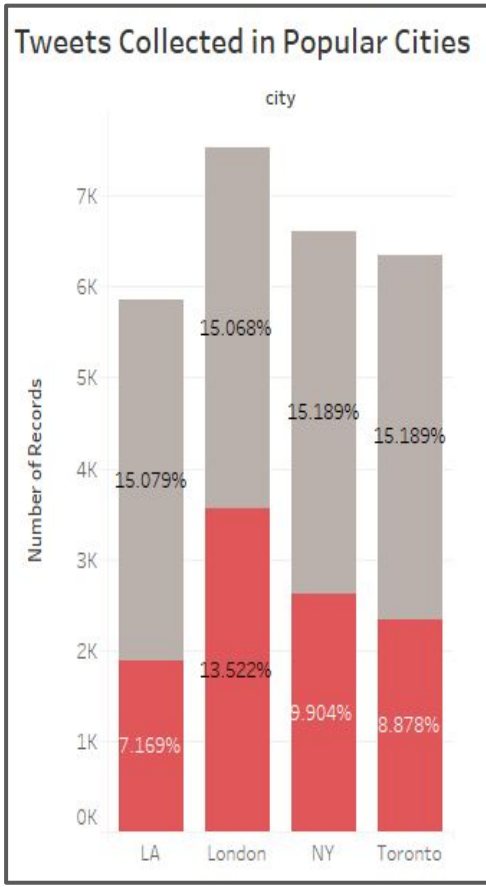
Scrapping Twitter Data

Package Used : **twitterR**

Date Range	Nov 24 - Nov 29
Hashtags	1. Apple + iphone 2. Samsung+ galaxy
Geolocations	1. Los Angeles 2. New York 3. Toronto 4. London
Fields	1. Tweet Text 2. Retweet Count 3. Favorite Count 4. IsRetweet 5. Time

Total Number of Tweets	26,334
Apple	15939
Samsung	10395

Distribution of Tweets
Red : Samsung
Grey : Apple



Preprocessing the Tweets

Filter the tweets which contains words related to deals and coupons

Coupons	40
Giveaway	273
Deals	1288
Giveawayends	787
Both products	365

Custom Stop Words :
BestBuy, Save,

Processes the tweets to remove following from the tweet text

Un recognised Variables :
Convert from Latin1 to ASCII , if cannot be converted delete

Remove

1. URLS
2. # Symbols
3. Punctuation symbols
4. Control Characters
5. Digits
6. White spaces
7. Stop words

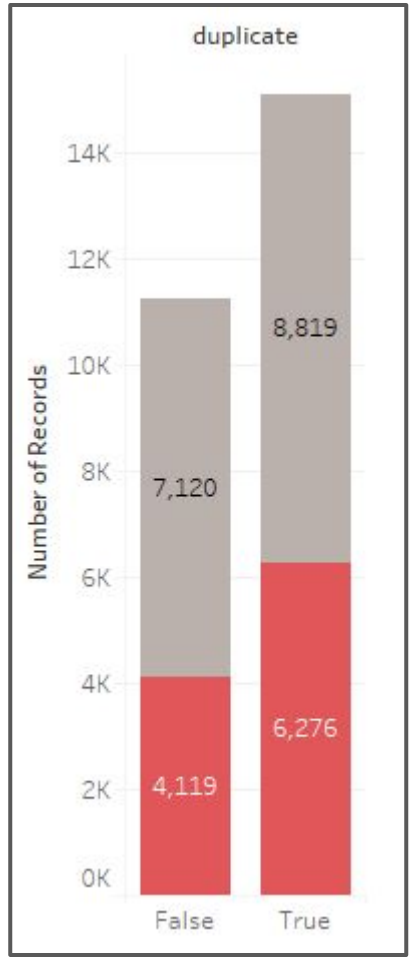
Retweet and Duplicate Data

Retweet : Share the tweet with original tweet tagged.

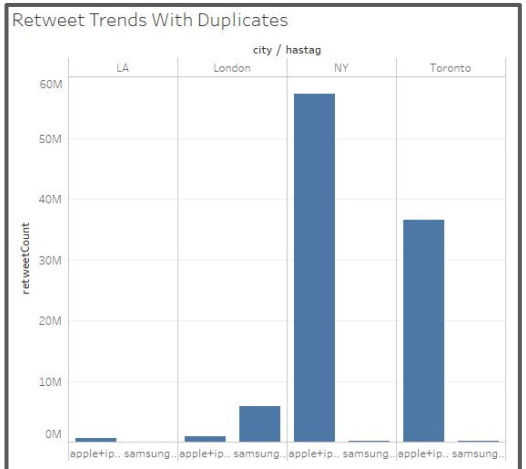
Share Tweet : Copy and paste the tweet

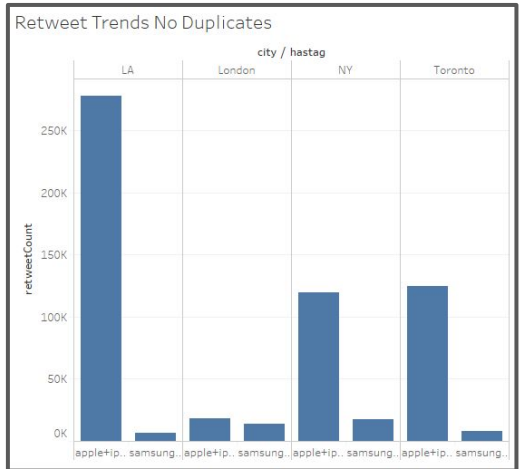
Retweet + ShareTweet = Duplicated Data

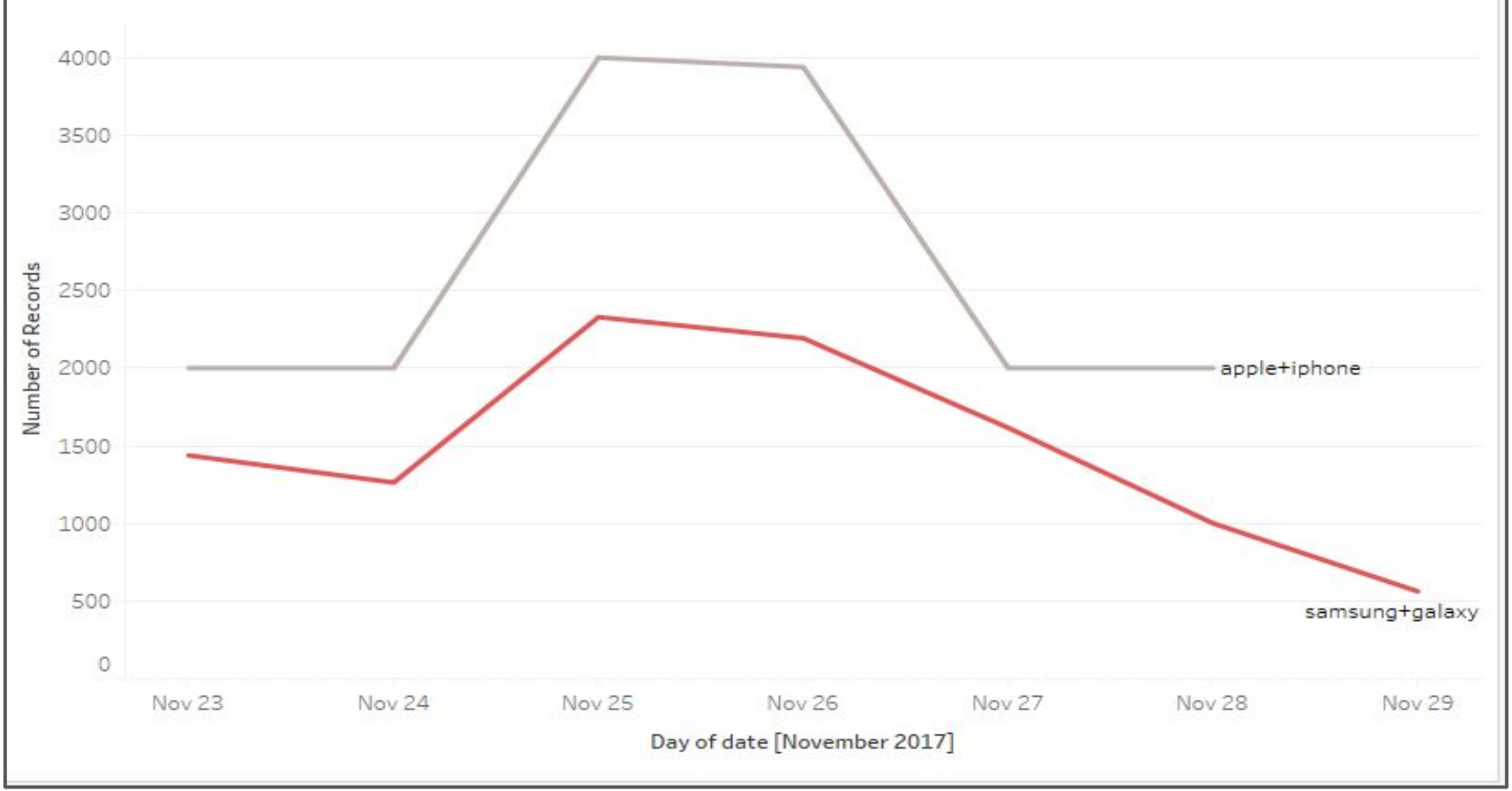
Retweets Depict influence of any Events



Trends in Retweet







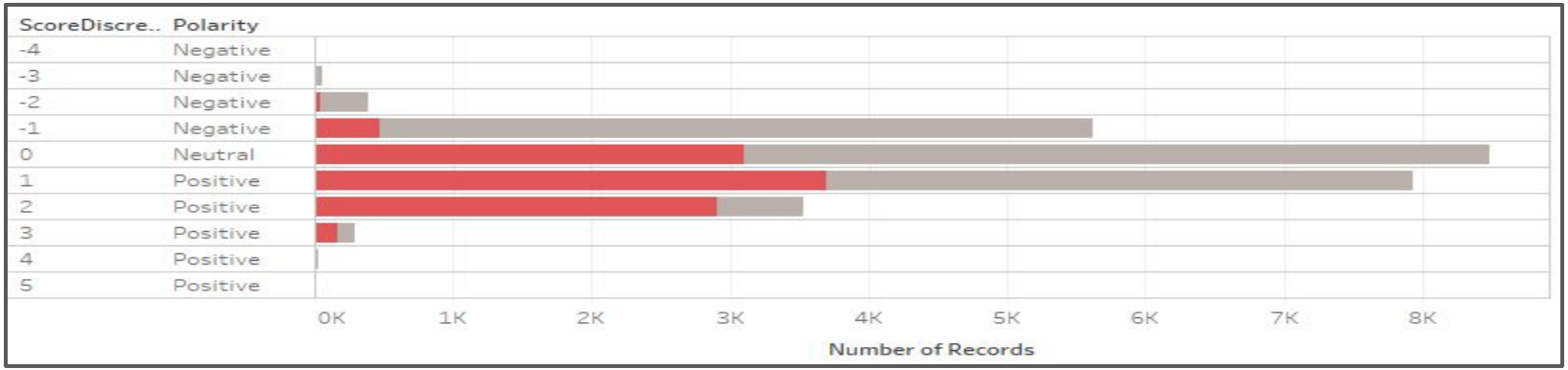
Variation of Tweets in Time Series : Black Friday Effect

Sentiment Analysis

Lexicon Based Analysis : Based on the Lexicon Data provided by **University of Illinois** at chicago. They have provided lists of Positive and negative words which is used for Classification of Tweets into positive , negative and neutral tweets.

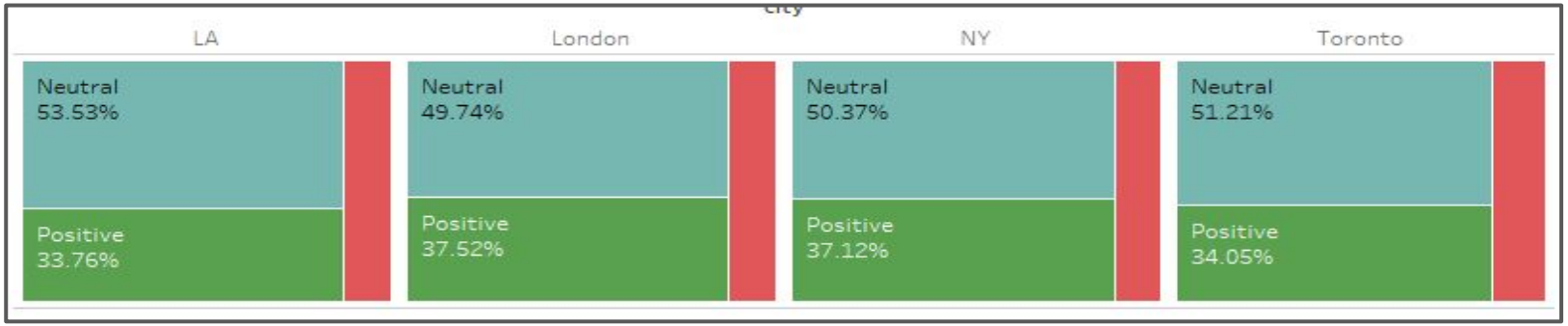
Processing Sentiment : Each tweet is split into words and matched against positive and negative words. **Sentiment score is the difference of {Positive Match - Negative Match }**

If the Score is 0 then “Neutral” , >0 “Positive” , <0 “Negative ”

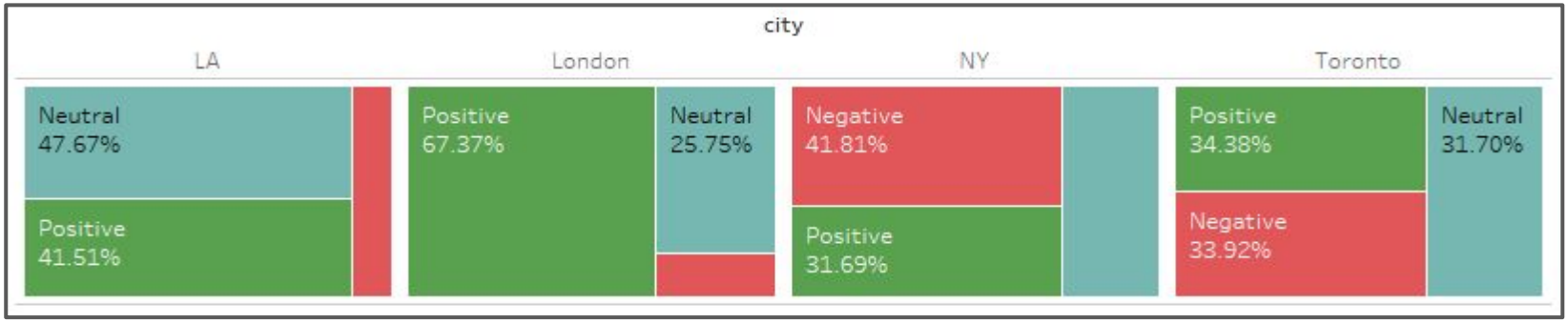


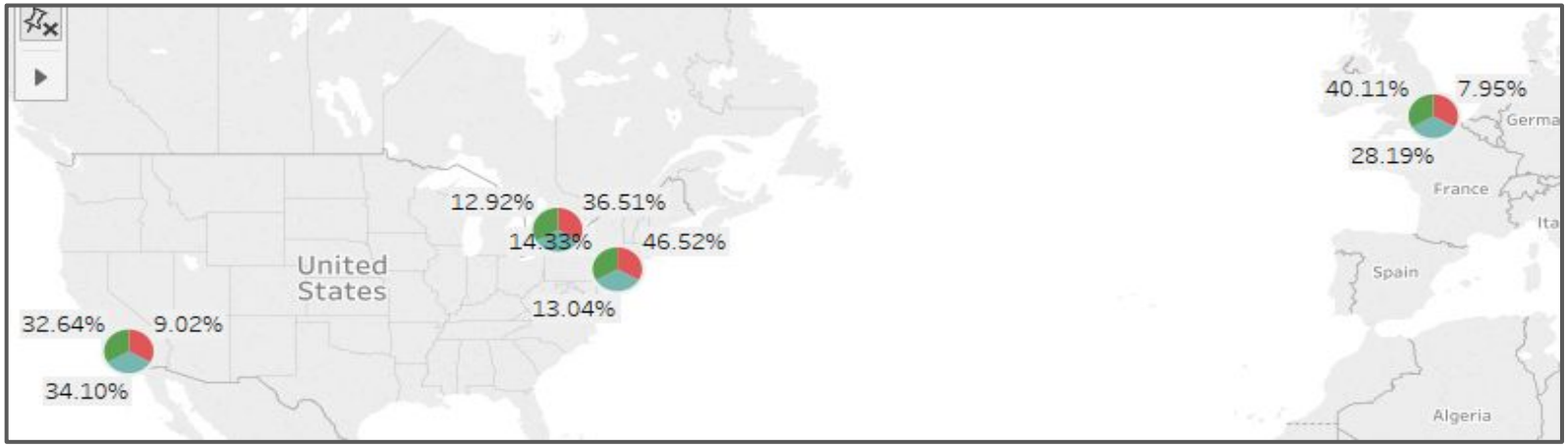
Sentiment Analysis in Different Cities

With Duplicates

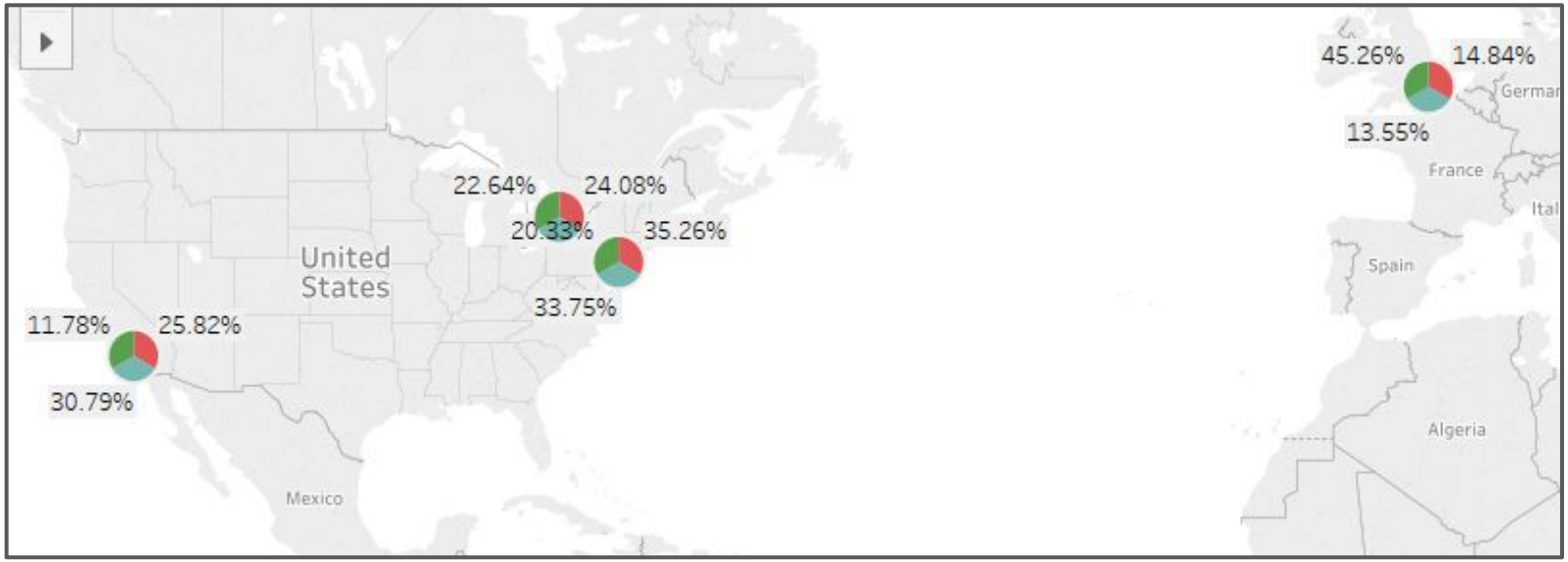


No Duplicates





Apple
Iphone



Samsung
Galaxy

Negative

This is so stupid I have an iPhone plus which is bigger and can use it with one hand It isnt that hard I have s

This is Apples awful phone experience iphone iphonex

Why is my Apple iPhone unable to access WiFi Internet and how to fix it Troubleshooting Guide

Neutral

AmazonPrime Shipped Samsung Galaxy Note US Version Factory Unlocked Phone Screen GB Mid

AmazonPrime Shipped Samsung Galaxy Note US Version Factory Unlocked Phone Screen GB US W

AmazonPrime Shipped Samsung GalaxyS Unlocked GB US Version Midnight Black US Warranty aff

Positive

WIGHIfunny I love the Note Had the Notes since the Note and Im pleased And the Galaxy S is a really good p

VZWSupport How much for one line iphone x GB Unlimited data and apple care and protection

VZWSupport Ilinrac How much for one line iphone x GB Unlimited data and apple care and protection

Predictive Analysis

Tried few algorithms like SVM and random forest.

Steps Involved

1. Convert the sentiment score into 3 classes, =0 “Neutral” > 0 “Positive” and <0 “Negative”
2. Divide the Data into Train and test
3. Converted the text part of data into Document term matrix.
4. Accuracy was around 60 % , precision and recall values are not good as data is highly imbalanced.

Packages Used :

1. **tm**
2. **“RtextTools”**

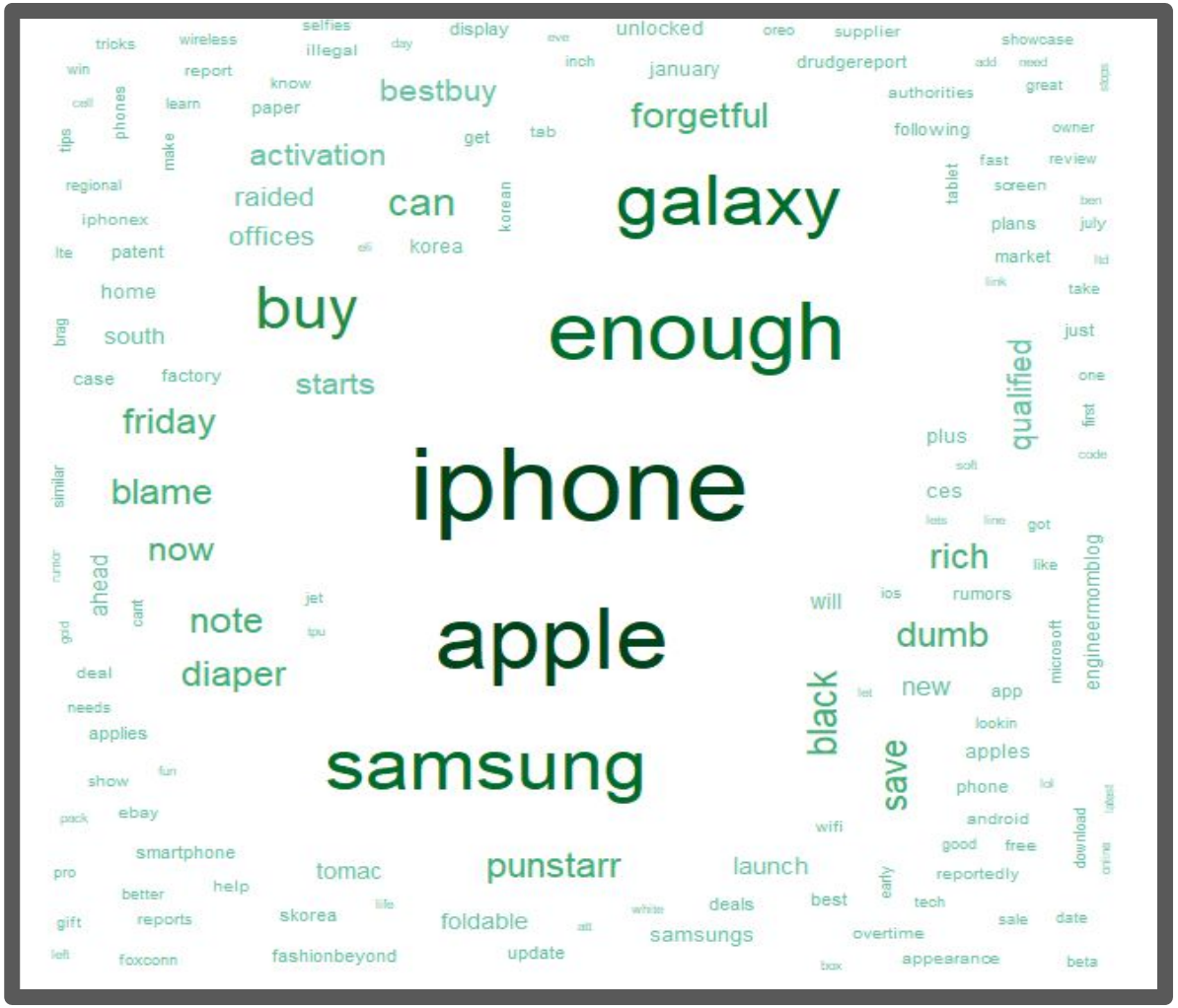
References

1.https://rpubs.com/mbhargav68/TweetClassificationModel

2.https://boraberan.wordpress.com/2013/12/24/sentiment-analysis-in-tableau-with-r/

3.https://stackoverflow.com/questions/1958267/how-to-do-multi-class-classification-using-support-vector-machines-svm

4. http://tidytextmining.com/sentiment.html


selfies
display
unlocked
oreo
supplier
showcase
tricks
wireless
illegal
day
ever
inch
january
drudgereport
add
need
win
report
know
bestbuy
authorities
great
stops
cell
phones
learn
paper
forgetful
following
owner
tips
make
get
tab
activation
fast
review
tablet
regional
raided
can
korean
galaxy
screen
ben
iphonex
offices
korea
plans
july
lte
patent
market
bid
link
home
take
buy
enough
brag
south
just
case
factory
starts
one
qualified
friday
first
plus
code
soft
similar
blame
iphone
ces
lets
line
got
rumor
now
rich
like
ahead
cant
jet
will
ios
rumors
engineermomblog
grid
note
tpu
apple
dumb
microsoft
deal
diaper
new
app
black
needs
let
lookin
applies
apples
save
show
fun
phone
lol
pack
ebay
android
wifi
smartphone
good
free
download
pro
tomac
punstarr
launch
reportedly
online
help
early
better
deals
best
tech
life
white
gift
reports
skorea
foldable
att
samsungs
sale
date
overtime
left
foxconn
fashionbeyond
update
appearance
beta
box

Questions ?