NLP Analysis of Consumer posts on Samsung Galaxy S8 vs. iPhone 8, iPhone X

What were the most important attributes for each product?

Gathering information from the tweets to comprehend the most important attributes will help us know the features that were highlighted the most. This intel will be further used to fetch the sentiments of consumers with regards to some of these important features. We were able to extrapolate a set of important attributes for each product. From this extrapolation we could clearly infer that majority of the consumers prioritise certain features over others when looking at a product. We will be breaking the implementation for this into three parts:

1. **Methodology**
2. **Process**
3. **Key Takeaway**

Methodology: Latent Dirichlet Allocation(LDA) with Term Frequency - Inverse Document Frequency(TF-IDF)

To get the most important attributes for each product we use Latent Dirichlet Allocation(LDA) with Term Frequency - Inverse Document Frequency(TF-IDF). This model gives us a set of topics and the top N (N=20 in our scenario) words for each topic. The distribution of words for each topic can help us infer the context of a set of tweets that are relevant to the topic, and the top words from the topics are extrapolated to get the most important attributes for each product.

TF-IDF: It is a statistical measure that evaluates how relevant a word is to a document in a collection of documents. and is very useful for scoring words in machine learning algorithms for Natural Language Processing (NLP). We can do a similar process by implementing CountVectorizer. But we instead use TF-IDF because it not only focuses on the word count but also gives a higher weightage to important words.

LDA: It is a topic modeling technique in which we first transform the TF-IDF data set into a document-topic distribution set. From this, we try to comprehend the topic to which a document belongs and then we fit the transformed data set to get the important words from each topic. And then when you see that some documents are connected to same set of words. You know they discuss the same topic.

Parameters: num_topics=7, random state=0

Process

We implement LDA with TF-IDF on Twitter and Non-Twitter data separately, this helped us understand the difference in opinion and the context of sentiments on Non-Twitter and Twitter data individually. We were able to gather information on how a consumer's opinion varies with respect to the platform they use.

For each of the following, we segregate the reviews on the basis of the product i.e. Samsung S8 or iPhone 8/X they were referring to, and then implement TF-IDF on the two set of data we get after segregation. The TF-IDF vectors retrieved are then passed to the LDA model to get the 7 topics and their respective distribution of words. The reviews used are tokenized and then lemmatized to make them more abstract, by doing so we also drop the stop words and irrelevant words.

Key Takeaways

```
[Topic 0]
be : 0.007357470942195354
not : 0.0066653025485369005
case : 0.006276912829170179
screen : 0.006253840624785703
new : 0.004788888575690412
work : 0.0045675469932131675
good : 0.004487312963912107
plus : 0.004484081456753619
do : 0.003968161883026633
look : 0.003583006913828301
buy : 0.003422573488274386
like : 0.0033316215131951005
lg : 0.003298040103643031
display : 0.0032124831958477915
app : 0.003183151921097436
use : 0.003111536935116095
camera : 0.003066043187047916
great : 0.003065027679050993
charge : 0.0029548879908169955
it : 0.002906833189128399
```

```
[Topic 1]
bixby : 0.010001064672875561
new : 0.008009392913417533
launch : 0.007099928491603563
assistant : 0.005699932767873606
feature : 0.005561695414692924
note : 0.005231401489446752
flagship : 0.004783911579403594
display : 0.004408824184093643
gb : 0.004246396279996649
button : 0.004232387055168878
company : 0.004012115611832396
plus : 0.003903918656350162
camera : 0.003897599578651476
come : 0.003716908639242747
voice : 0.003663000552236657
pre : 0.003485653872465267
leak : 0.003338820393126851
order : 0.003228966631931305
year : 0.003203620235436359
april : 0.003129957772410165
```

```
[Topic 4]
color : 0.012289749409847402
leak : 0.009779442672239317
androbrix : 0.00956881028722808
tag : 0.006572800385921555
black : 0.006379925533062219
pinter : 0.005978466327718287
orchid : 0.005131963339566801
new : 0.004881544204408177
plus : 0.004774665263447555
pin : 0.004693407756741456
evan : 0.004550744130715374
silver : 0.004499893739529259
blass : 0.004488836347011119
grey : 0.004434681881416921
pinterest : 0.004219362901996522
gold : 0.004105527322216122
arctic : 0.003633587555343694
sky : 0.003591428034641431
photo : 0.003113714584197553
image : 0.003074588890437806
```

```
[Topic 5]
htc : 0.005824055329276447
xiaomi : 0.005453918356928043
vs : 0.005364720759736753
mi : 0.005016148019625597
google : 0.004806577440403316
pixel : 0.004689557995011527
android : 0.004084184031664504
post : 0.004067073306246445
lg : 0.004038402220296013
plus : 0.003997264405183300
huawei : 0.003823627016977041
video : 0.003618201038897726
like : 0.003506192780419474
2017 : 0.003449105324952304
g6 : 0.003057143339741215
new : 0.002961414429850271
follow : 0.002931461096339554
xperia : 0.002631725646949359
xl : 0.002630570164695118
tag : 0.002622776402184198
```

Fig: Some of the relevant topics for Samsung(Non-Twitter forum)

From the above figure we can see that the all the 4 topics refer to something different,

- **Topic 0**: Samsung Design - It mainly talks about Camera, Screen, Display.
- **Topic 1**: Bixby - Samsung launched their own personal voice assistant bixby for the first time. Many reviews were based around this topic of conversation.
- **Topic 4**: Color - This topic revolves around different colors of Samsung S8.
- **Topic 5**: Competitors - Many consumers compared Samsung S8 to other products like HTC, Xiomi, Pixel, LG etc.

```
[Topic 1]
plus : 0.01710452872426169
camera : 0.012318021496886822
portrait : 0.011932052924767584
photo : 0.008452216905919869
mode : 0.006680562186324062
new : 0.005908063431804733
zazzle : 0.00589692742109274
8p : 0.005855888202118896
lus : 0.005670406650228611
lighting : 0.005400004244177647
shoot : 0.00521926143792051
video : 0.004855406099155836
tag : 0.004837390693126699
review : 0.0042815039678373535
photography : 0.004204240570229441
light : 0.004161782597569567
take : 0.003924960106278595
shotoniphone : 0.003667408366663592
case : 0.003602611550888841
test : 0.00354004906165988
```

```
[Topic 2]
new : 0.008449374465524438
plus : 0.006696592741169228
wireless : 0.005299915797966896
feature : 0.0043767023171600245
io : 0.004266399051096419
charge : 0.003709295078788494
app : 0.00357662747361624
be : 0.0035687456987492424
11 : 0.0033786785688687715
google : 0.0031938106379559308
camera : 0.0031307529175904933
come : 0.0030350802526205403
year : 0.0029421421999983694
charging : 0.002736890971053126
buy : 0.0026197505470327695
not : 0.002616861649442045
screen : 0.002592961293763455
like : 0.002589445304728717
release : 0.002554854472078246
time : 0.00253228940126656
```

```
[Topic 3]
battery : 0.010412486008741392
plus : 0.010064733519001527
issue : 0.00929873635263414
io : 0.008772980112298786
11 : 0.008475756275457083
charge : 0.008258832994025434
fix : 0.007952285503307593
report : 0.006560283730539801
fast : 0.005996169106191969
follow : 0.005853339030588816
free : 0.005647582804059886
update : 0.004935174566948659
work : 0.004878356478760901
open : 0.004819719750658192
problem : 0.004794276057210107
swell : 0.004698163791444979
crackle : 0.004676619854043172
tag : 0.004639817272137394
freeiphoone8pluss : 0.0044076860343264706
new : 0.0043298833722207
```

```
[Topic 6]
8p : 0.015864134340230464
gb : 0.015820815042790464
lus : 0.015217867735684351
plus : 0.012934004712111025
64 : 0.009452115435024857
tag : 0.0092784322425556
gold : 0.009134672002162978
256 : 0.0077376992749622715
unlocked : 0.00686148973448507
iphonex : 0.006483683814520291
ifttt : 0.005495236962498263
twitter : 0.005470546206345298
new : 0.005092376435899624
000 : 0.004855715469054763
space : 0.004707420126230696
price : 0.004429278760331767
buy : 0.003938418184319336
gray : 0.00392435636239879
silver : 0.0038409513104285787
tech : 0.00342094124843182
```

Fig: Some of the relevant topics Apple(Non-Twitter forum)

In case of Apple, the topics were different:

- **Topic 1**: Camera - Majority of the documents were related to iPhone's camera review for this topic.
- **Topic 2**: Wireless Charging - Apple introduced wireless charging for the first time in iPhone X.
- **Topic 3**: Battery & Noise - Apple iPhone had a lot of issues and consumers did highlight them in their reviews. Two main issues were:
  - Battery Issue: Apple on Friday said it is "aware and looking into" the matter concerning battery problems that were reported on some units of iPhone 8 and iPhone 8 Plus
  - Noise Issue: iPhone 8 or iPhone 8 Plus is making a crackling sound or noise background during phone calls.
- **Topic 6**: Color and Storage - Consumers expressed a lot about the iPhone unique set of colors and storage.

```
[Topic 1]
galaxys : 0.03334532940325821
youtube : 0.01454782901238353
review : 0.013745068669835386
video : 0.0126765154710615
like : 0.011831802212144785
8p : 0.00995965058051581
android : 0.008179761586070646
case : 0.008019971895479917
color : 0.007909151828480331
good : 0.007626715997108348
plus : 0.007208153556029685
lus : 0.0058651555064988875
look : 0.005861801217540614
rumor : 0.00583436563191627
cover : 0.005539414772506087
fingerprint : 0.005488706242730883
leather : 0.005421358741452621
de : 0.004922853635096446
suck : 0.004822732091075164
ouch : 0.00475689414810601
```

```
[Topic 3]
screen : 0.019318986260637938
protector : 0.012848689339570445
button : 0.012446485620540668
be : 0.009136379475845217
bixby : 0.009119948973509305
plus : 0.008741311599517368
glass : 0.008074687671247228
want : 0.007330443111006952
pack : 0.006285595604817682
home : 0.005792923321308479
free : 0.005715082402872187
android : 0.00557076996813499
forget : 0.005367727803952894
what : 0.005274460991527604
upgrade : 0.005206059271907889
not : 0.005113908749083718
pre : 0.0045444203863735785
remap : 0.00452708079451168
app : 0.004516722114631821
case : 0.004487722477230732
```

```
[Topic 6]
iris : 0.01377794233415649
red : 0.01246082762213999
recognition : 0.011944997686444797
photo : 0.011636467946539188
facial : 0.011235143815510246
fix : 0.010815839918708948
scanner : 0.010640884642479706
problem : 0.010059198811096516
display : 0.00855176188896647
new : 0.008453126463318302
screen : 0.00831971124012514
tint : 0.00830802729601053
update : 0.00791367497068016
issue : 0.007545518087624092
report : 0.00725322981101736
unlock : 0.006797375465478193
feature : 0.006104644822466042
trick : 0.006104003232816862
active : 0.006042851016546083
think : 0.006009360656770148
```

Fig: Some of the relevant topics for Samsung(Twitter forum)

Moving on to Twitter reviews, we first look at Samsung:

- **Topic 1:** Case and Color: The topic in general was about samsung cases i.e. consumers did care about the protective cases too. From this we can infer that accessories are also an important topic of discussion for a product.
- **Topic 3:** Screen Protector
- **Topic 6:** Samsung introduced iris scan for face recognition, which was a unique feature.

```
[Topic 1]
plus : 0.021097307192294412
wireless : 0.012447722405777097
buy : 0.011816587077099174
new : 0.01048236937395689
charging : 0.009885217721010276
not : 0.008772185178145743
do : 0.00785517741681556
get : 0.00686632590979197
tomorrow : 0.006720031498577046
gb : 0.005795454910411997
look : 0.005770411436783007
fuck : 0.005469659841767412
charge : 0.00519336927908247
unveil : 0.005149203594571049
like : 0.004798975771851015
come : 0.004605703106692385
be : 0.004431987071075232
review : 0.0043869998410744826
good : 0.004377550147578269
event : 0.0043348005843658385
```

```
[Topic 4]
io : 0.014608523340477336
11 : 0.012587524477161069
leak : 0.011810871823219797
new : 0.011005262993653432
review : 0.010667045585601125
screen : 0.009955374237369616
camera : 0.00961085174680499
touch : 0.009408489245825718
rumor : 0.007341789861663493
plus : 0.007268912553130715
display : 0.006799688569670996
ar : 0.006214564370644017
design : 0.005963137241127566
curve : 0.005779737188345295
big : 0.005552601106065854
cost : 0.005258571547057031
com : 0.005217314195712317
dual : 0.00494866800562747
change : 0.004941720711198238
feature : 0.004925013628184181
```

Fig: Some of the relevant topics for Apple(Twitter forum)

Incase of Apple reviews on Twitter:

- **Topic 1:** Wireless Charging
- **Topic 4:** Rumours/ Leak - Many reviews on Twitter for Apple were based around the rumours of price and discussed features, design prior to the release. This is was mainly in context of the iPhone pictures leaked before launch.

From all the analysis done we can gather that the most important attributes for Samsung were **Camera, Screen Design, Bixby, Color** while those for Apple were **Wireless-Charging, Color, Storage and Issues (Battery and Noise.)**. We also notice that each product had a unique set of features that made them different and the consumers sentiments did highlight these points. Like for Samsung it was bezel display or bixby and for Apple it was wireless charging, color.

Comparing Twitter and Non-Twitter data we can infer that Twitter reviews were more neutral and were expressing on the basis of the hype of the launch event of the two products. Where as the Non-Twitter sentiments were focused on people who actually bought the product were expressing sentiments on the basis of hand-on experience of the product.

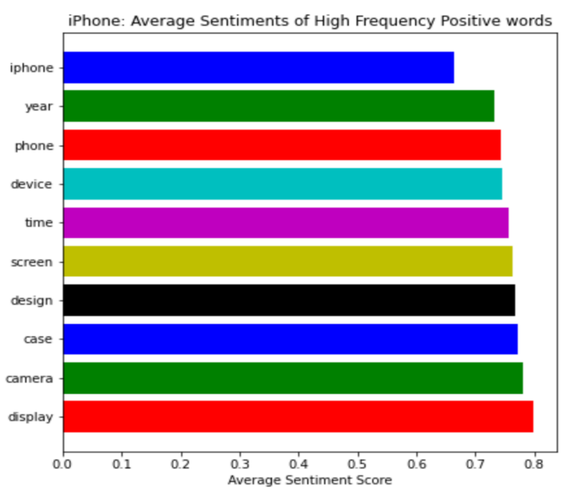
What attributes were liked the most and what attributes weren't?

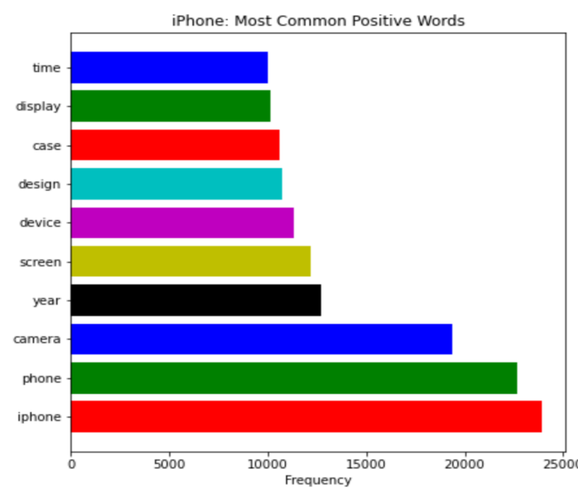
Approach

We started out by separately extracting the nouns and adjectives from tweets referring to each phone, further processing them using a cleaner function and then also storing the sentiment scores of the tweets containing the words. We then took the top 10 most occurring nouns and analyzed the average sentiment scores for the tweets they were mentioned in. We stored words associated with positive and negative sentiments in separate lists for both apple and samsung devices. We also separately extracted the most common positive and negative adjectives used for each phone to get an idea about customer satisfaction.
Finally, we displayed the most common positive/negative nouns(i.e. Attributes used in positive/negative contexts),their frequencies,average sentiment scores, as well as the most common positive/negative adjectives for each device.

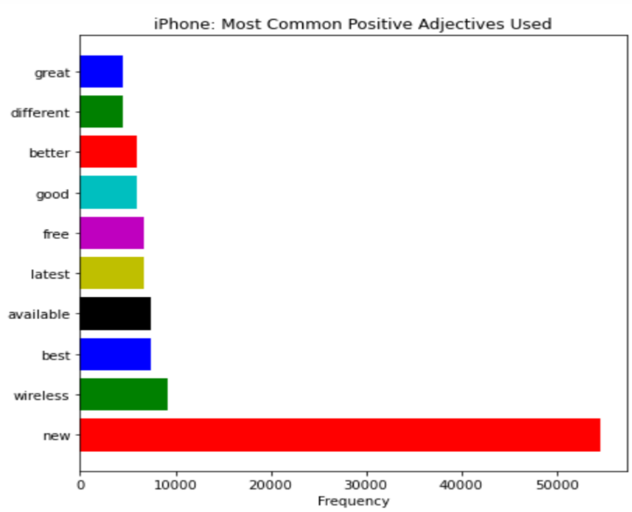
IPHONE

Most Common Positives Nouns/Attributes

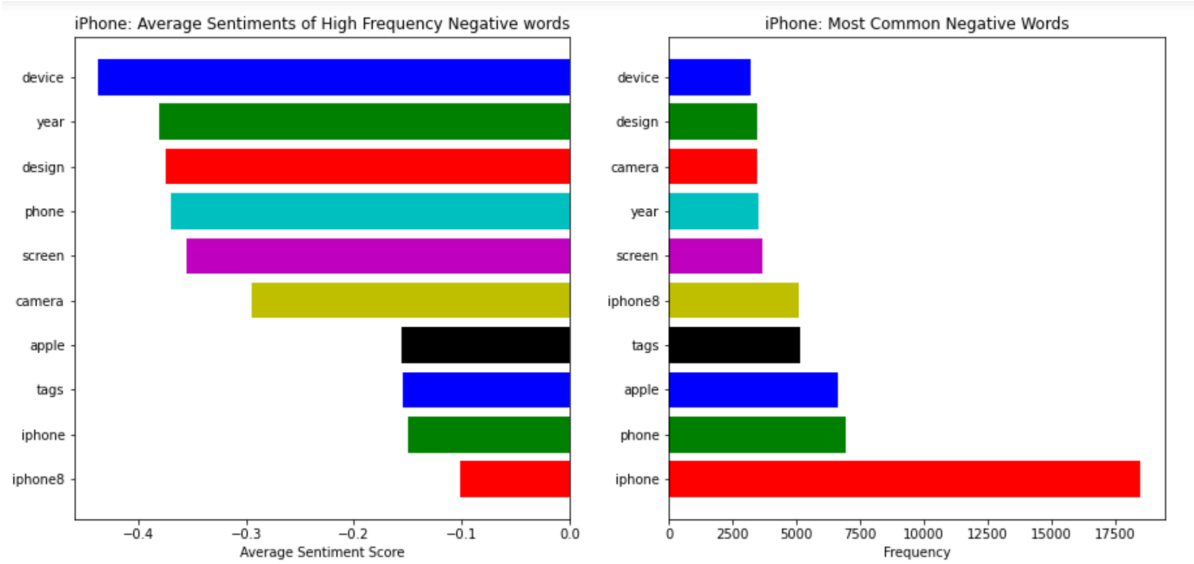




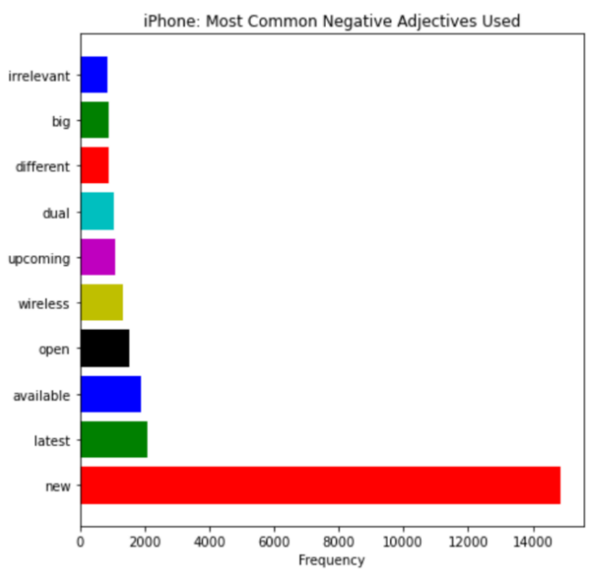
Most common Positive Adjectives



Most Common Negative Nouns/Attributes



Most common Negative Adjectives

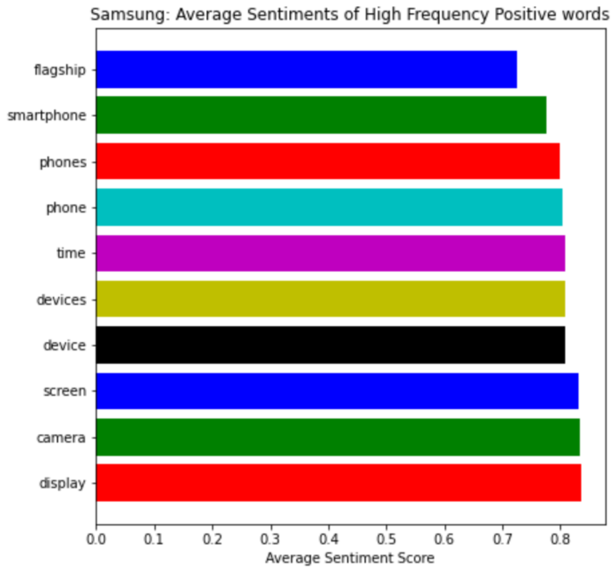


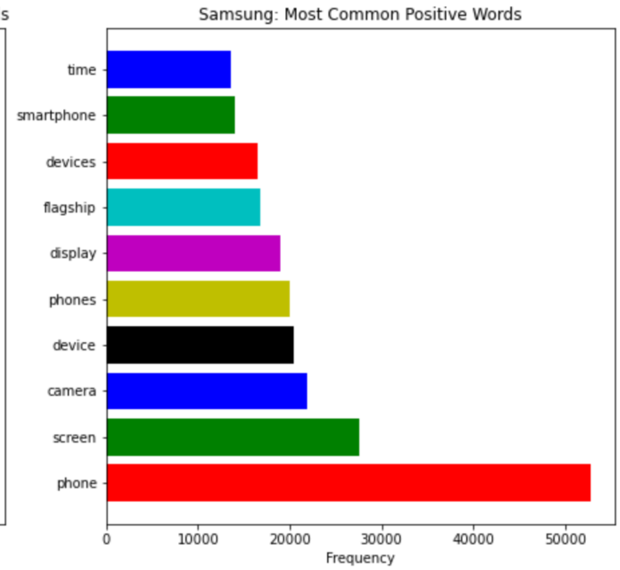
For Samsung, the list of both positive and negative words contain “camera”, “screen” and “design”, indicating that they have mixed reviews but still generate a lot of buzz, which could be a huge benefit to both the brands. Additionally, case and display were used in a positive context and feature in a lot of tweets, so it indicates that those attributes were huge successes for the iPhone.

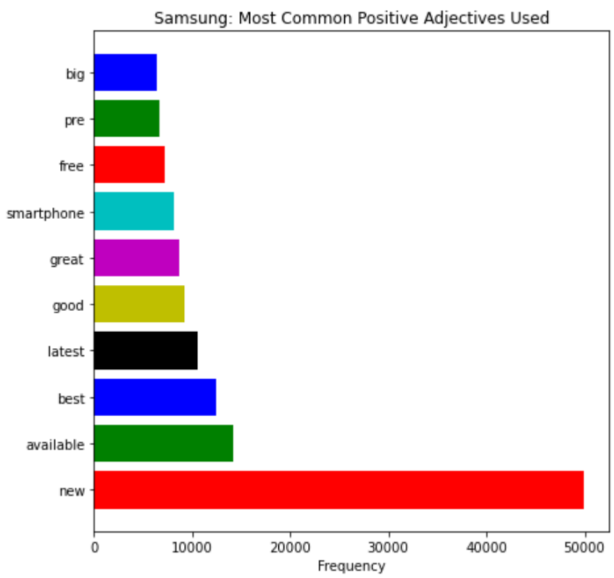
You also similarly see words like wireless(referring to the iPhone’s wireless charging capabilities) appear in positive and negative contexts(mixed reviews), and the words “irrelevant”, “big” and “iphone8” feature in the negative lists. This could possibly be referring to internet chatter in 2017 when people were rather disappointed that the “iphone8” was not a huge improvement over the iphone7(and was hence “irrelevant”). Also note that the size(“big”) has been used in a negative context for the iPhone.

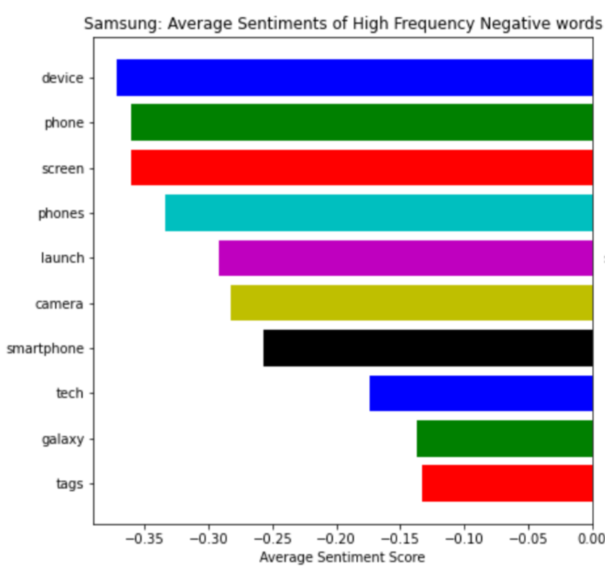
SAMSUNG:

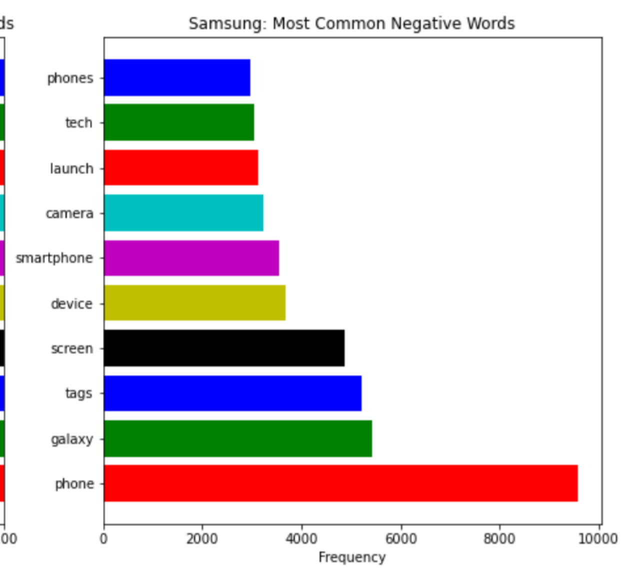
Most Common Nouns/Attributes

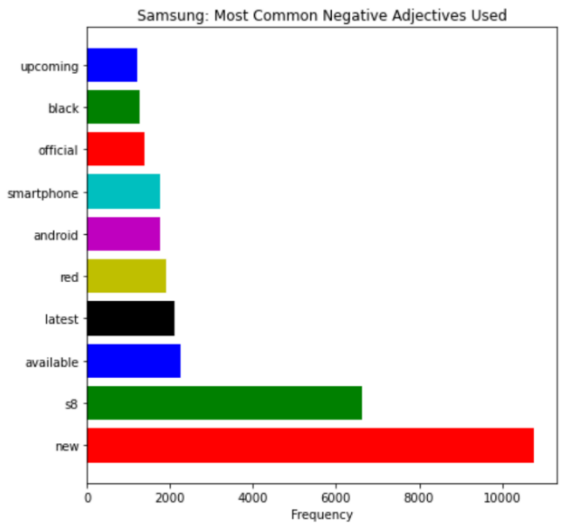

Samsung: Average Sentiments of High Frequency Positive words
flagship
smartphone
phones
phone
time
devices
device
screen
camera
display
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Average Sentiment Score


Samsung: Most Common Positive Words
time
smartphone
devices
flagship
display
phones
device
camera
screen
phone
0
10000
20000
30000
40000
50000
Frequency


Samsung: Most Common Positive Adjectives Used
big
pre
free
smartphone
great
good
latest
best
available
new
0
10000
20000
30000
40000
50000
Frequency


Samsung: Average Sentiments of High Frequency Negative words
device
phone
screen
phones
launch
camera
smartphone
tech
galaxy
tags
−0.35
−0.30
−0.25
−0.20
−0.15
−0.10
−0.05
0.00
Average Sentiment Score


Samsung: Most Common Negative Words
phones
tech
launch
camera
smartphone
device
screen
tags
galaxy
phone
0
2000
4000
6000
8000
10000
Frequency



Note that samsung also has camera and screen are very frequent words used in positive and negative contexts. The display of the phone is also a positive word used very frequently, indicating that it was a hit with customers. Note that "big"(possibly referring to the size of the phone) is used in a positive context here(unlike the iPhone), and the words "Red" and "Black" feature in the list of negative words, indicating that consumers were displeased with those color choices.

Takeaway

The camera, screen and display are the most talked-about words, so we should really focus on them while designing our product. The design and case of the iPhone were hits, and the size of the Samsung phone was reviewed positively(unlike the iPhone). Similarly, the color choices of Samsung phones were viewed negatively compared to the iPhone. We should take these factors into consideration as well.

What was the sentiment for these products?

To calculate the sentiment for the two products that we are analyzing – Samsung Galaxy S8 and the Apple iphones X & 8, we will be using Vader.

VADER (Valence Aware Dictionary for Sentiment Reasoning) is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion. It is available in the NLTK package and can be applied directly to unlabeled text data. After we have lemmatized the tweets, we directly use VADER's **SentimentIntensityAnalyzer()** which takes in a string and returns a dictionary of scores in each of four categories:

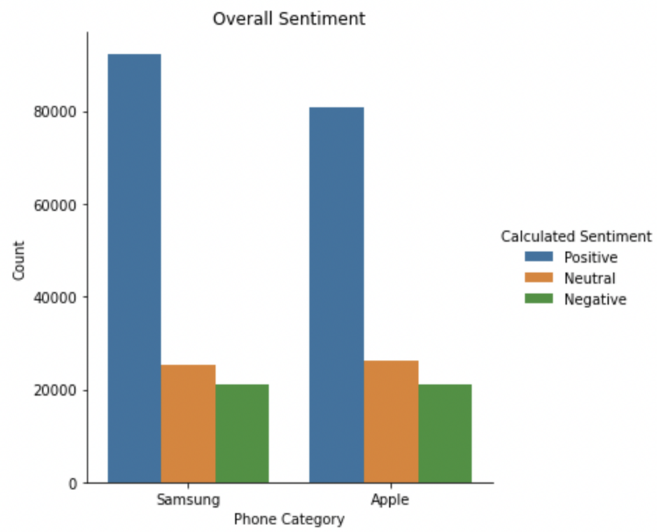
- negative
- neutral
- positive
- compound *(computed by normalizing the scores above)*

In our case, we will use the compounded score of the tweets and categorize the tweets into three categories – Positive, Neutral and Negative.

VADER's calculates compound scores on a scale of {-1, 1}, with -1 being associated with extremely negative sentiment, whereas 1 being associated with extremely positive sentiment.

We will then categorize these tweets as Calculated Sentiment based on their compound scores as such –
Negative – Compound Score is lesser than 0
Neutral – Compound Score = 0
Positive – Compound Score is greater than 0

The overall sentiment of the two products after calculating the scores and categorizing them is as follows -

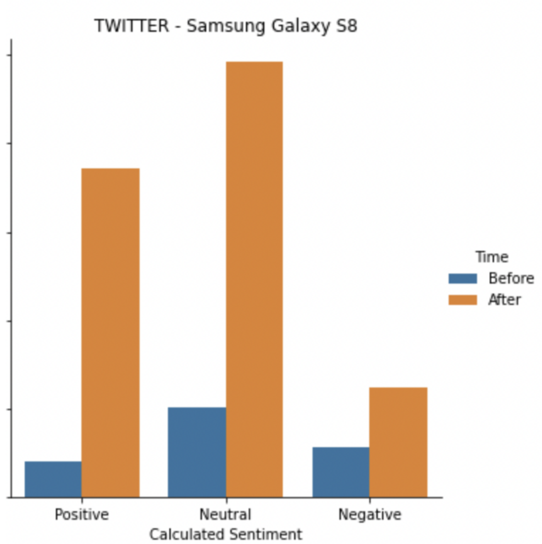


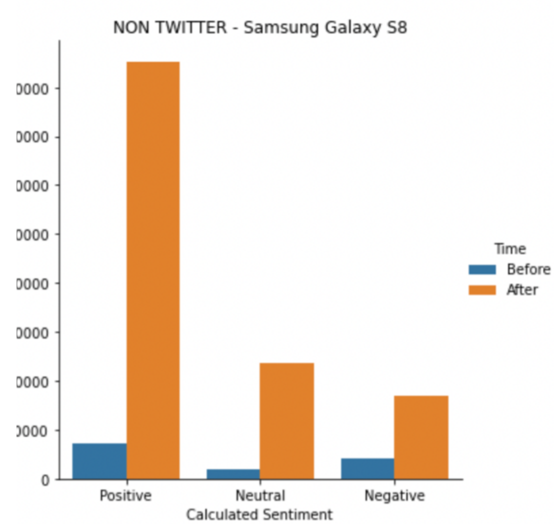
The plot indicates that the overall sentiment for both the products was majorly positive. The response for all three sentiments for both the product was comparable.

Did the sentiment change BEFORE and AFTER the launch of the products? How?

To understand the sentiment change of the two products before and after launch, we find the individual sentiment change on the different types of data we have.

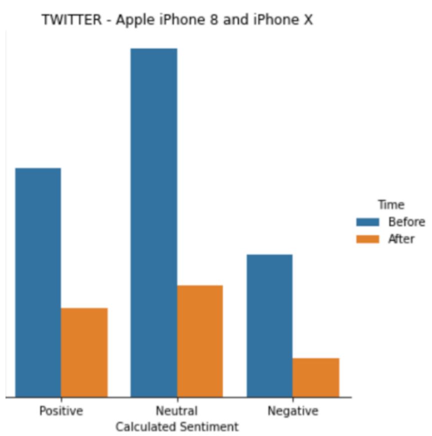
For Samsung Galaxy S8

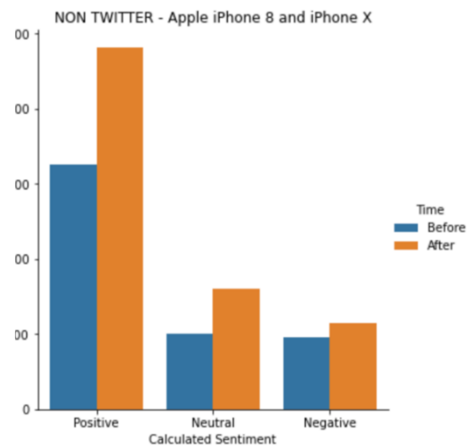




By plotting the graph for the Twitter data, we can see that the sentiment of the Galaxy S8 changed drastically post launch. A major reason for this that the data collected post launch was high in quantity as opposed to that of pre-launch. Nevertheless, the sentiment for the product overall Neutral or Positive.
In the case of non-Twitter data, we can see that the primary sentiment amongst those mentioning the Galaxy S8 was positive. As the non-twitter data mostly comprises of reviews and blogposts, we may extrapolate that the people using the product were overall positive about it in their responses.

For Apple iPhone X & 8

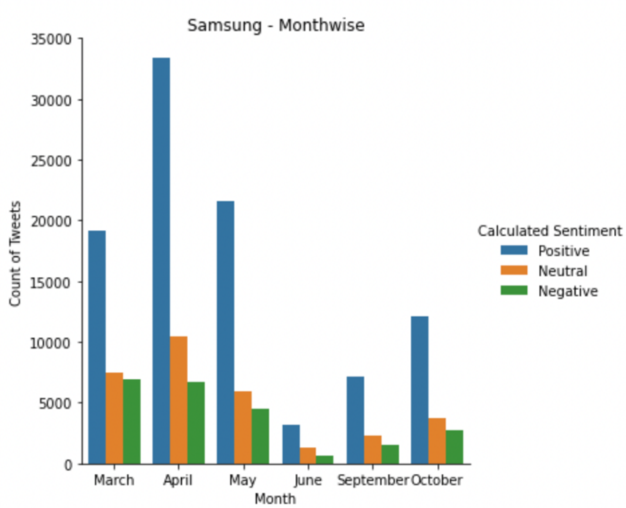




Twitter data about the Apple iPhones was primary tending towards being Neutral or Positive before the launch of the product. Since there isn't a lot of data available after the iPhones were launched, the recorded sentiment isn't a valid approximation of the overall post-launch sentiment. However, from whatever data we do have, we can see that the sentiment is more Neutral or Positive than Negative.
For non-twitter data, we can see that a lot of the response for Apple has been positive before the launch. The sentiment did not change much after the launch as we can see that the primary sentiment is still positive. However, it is curious to notice that even though we didn't have much data for Apple iPhones after they were launched, they were still talked about a lot.

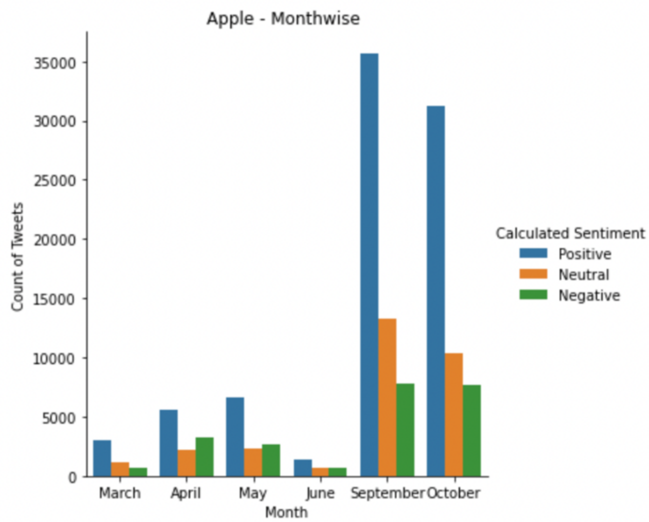
<u>MONTHLY SENTIMENT ANALYSIS</u>

For Samsung Galaxy S8



As the Samsung Galaxy S8 was released on March 29th, we have a lot of data to analyze the sentiment of the product after it's launch. However, as we have data just one month before the launch of the product, we don't have a lot of data to analyze the sentiment before the launch of the product.

For Apple iPhone X & 8



The Apple iPhones were released on September 22nd, which made it easier for use to understand the sentiment of the product before its launch. As the data we have is only until 31st October, which is just a month after the launch of the iPhones, we don't have substantial data to analyze the sentiment post-launch.

EXTRA-CREDIT - ADOPTION UPDATE PREDICTION:

Approach

In order to determine whether we can predict the adoption uptake of the products using the pre-release data, we ended up using a supervised learning approach.

Labeling

Though we didn't have any pre-generated labels, we decided to label those tweets that had a high positive sentiment or the ones that contained synonyms and tense variations of "buy" as "1" or adopted. The other tweets were labeled as "0"(not adopted).

Feature Engineering

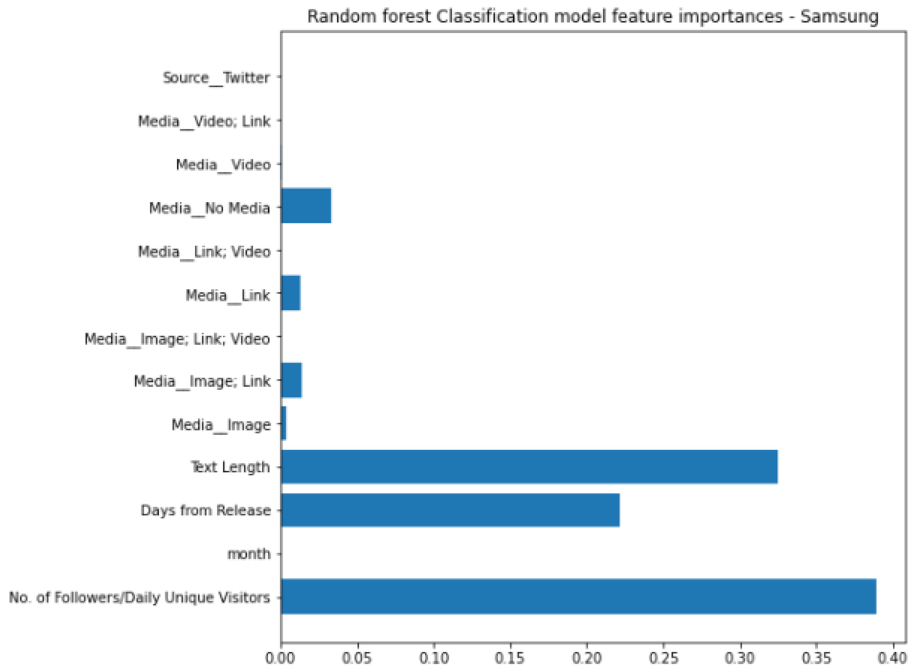
We first dropped irrelevant features that wouldn't contribute to our prediction, and ended up keeping the source type, media type, and the text of the tweet as our main features to engineer.

We used TF-IDF(described previously) to vectorize our lemmatized text data column into a number of numerical feature columns.

We also generated some of our own features, like the length of the text of the tweet and the difference in days between the tweet date and the product launch date.

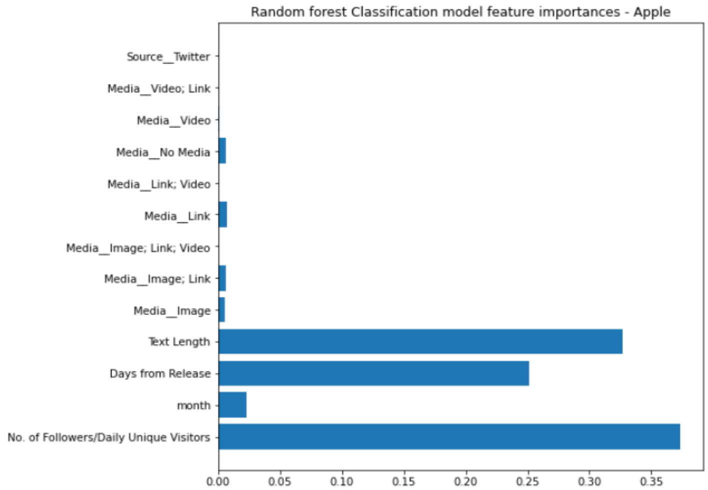
We trained a random forest classifier on our data because it is known to deal well with high dimensional data and large datasets. We trained our model on pre-release data and validated it on our post-release data.

Samsung:




```
{'0': {'precision': 0.8271356783919598,
  'recall': 0.9939613526570048,
  'f1-score': 0.9029072956664836,
  'support': 828},
 '1': {'precision': 0.0, 'recall': 0.0, 'f1-score': 0.0, 'support': 172},
 'accuracy': 0.823,
 'macro avg': {'precision': 0.4135678391959799,
  'recall': 0.4969806763285024,
  'f1-score': 0.4514536478332418,
  'support': 1000},
 'weighted avg': {'precision': 0.6848683417085427,
  'recall': 0.823,
  'f1-score': 0.7476072408118485,
  'support': 1000}}
```


Apple:




```
{'0': {'precision': 0.7971086739780658,
  'recall': 0.9690909090909091,
  'f1-score': 0.87472647702407,
  'support': 1650},
 '1': {'precision': 0.23880597014925373,
  'recall': 0.037825059101654845,
  'f1-score': 0.06530612244897958,
  'support': 423},
 'accuracy': 0.779064158224795,
 'macro avg': {'precision': 0.5179573220636597,
  'recall': 0.503457984096282,
  'f1-score': 0.4700162997365248,
  'support': 2073},
 'weighted avg': {'precision': 0.6831858357148785,
  'recall': 0.779064158224795,
  'f1-score': 0.7095625551787911,
  'support': 2073}}
```

As can be seen from the feature importances and classification reports for both products above, we are able to obtain an average accuracy of about, which is not great for binary classification. Furthermore, though the recall rate for class 0(no adoptions) is pretty high, the recall rate for class 1(adoption) is pretty low, indicating that the model does not perform very well when it comes to predicting adoptions(True Positives).
The most important features for this classification are the Text length,Days from release, no. of unique visitors/followers and certain sources like blogs.

Insights and Improvements:

The model's performance can be drastically improved by using more robust ways of dealing with class imbalances(There were more labels of "0" than there were of "1"). We can use oversampling/undersampling techniques or SMOTE, but this might be more convoluted than usual since we're dealing with text data.
The labeling processes can also be improved. We would ideally have liked to have pre-generated labels, since high positive sentiments or even the presence of the word "buy" may not be truly indicative of an adoption. We could improve it through an n-gram analysis(look for n-grams indicative of adoption).
It would also be great to have more non-text features to analyze for our supervised learning predictions. Since we're labeling the data based on certain text features, and using TFIDF to vectorize the tweet texts, there could be some data leakage involved(i.e. The labeling is conducted on the basis of some text features, and some of those features are vectorized and used for prediction, which can make the model biased.). Having pre-made adoption labels or more non-text features to predict adoptions would help resolve this issue.

In the future, it would also be interesting to explore unsupervised learning methods or time-based algorithms like ARIMA to predict adoption uptake.