Sentiment Analysis on User's Preference on Samsung and Apple Latest Mobile Product (Samsung GalaxyZFold4 and IPhone14) using Twitter Data

Ayomiposi Adebayo

10 October, 2023

OUTLINE

- Introduction
- Data Mining and Cleaning
- Exploratory Analysis
  - *Consumer Engagement*
  - *Time Series Analysis*
  - *Analysis of User's Tweets by Location*
  - *Sentiment Analysis*
- Statistical Analysis
  - *Summary Statistics*
  - *T-test Analysis*
- Discussion of Results
- Conclusion/Recommendations

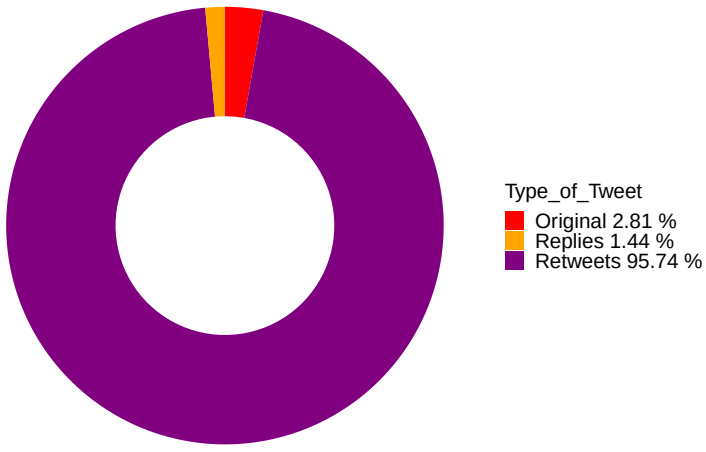
INTRODUCTION

- Social media has become a universal tool employed by several businesses for marketing purposes to boost their brand's image and to remain competitive in the global space (eMarketer, 2014; Shi, Chen, and Chow, 2016).
- Due to the diversity that exist in different businesses, many consumers tend to have different varying opinion on how they relate with their chosen brands on social media (Cathryn and Stander, 2018).
- The explosion of data in the forms of blogs, social media such as Facebook, twitter have provided customers with a platform to express their views or preferences about a product/service which tends to either influence sales of such product/services negatively or positively (Deepali, Kin, and Stephen, 2015).This can then be used by such brands to improve their brands or products.
- The twitter data for both Samsung GalaxyZfold4 and Iphone14 was extracted and mined from Twitter using the **rtweet** package
- The sample size for the raw data generated on tweets related to Iphone14 was 7,355 and 4,442 for the SamsungGalaxyzfold4 tweets

Exploratory Analysis

User Engagement Analysis

Figure 1a: Distribution of GalaxyZFold4 Tweets



Distribution of Iphone14 Tweets

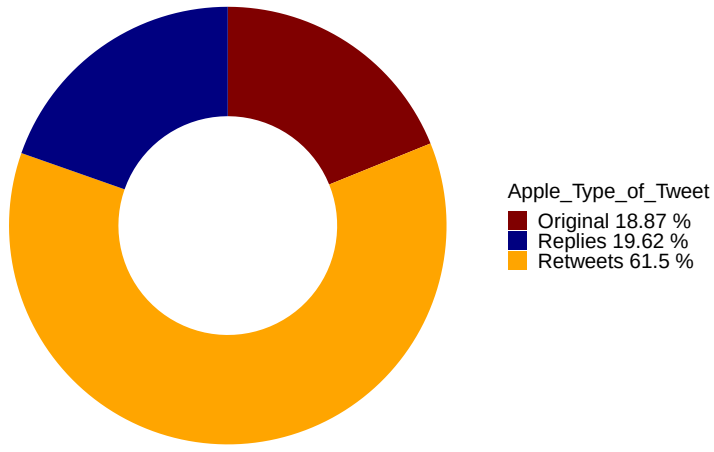


Figure 1: Donut Chat Showing the Percentage of Original tweets,Replies and Retweets of Samsung GalaxyZfold4 and Apple Iphone14

Analysis of User's Tweets by Location

Figure 4a: Location of users tweeting about #Iphone14

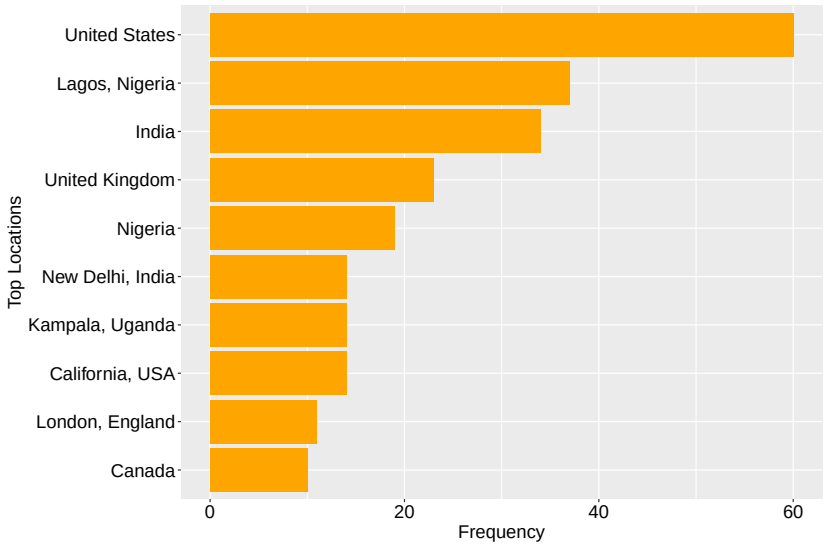


Figure 4b: Location of users tweeting about #Sams

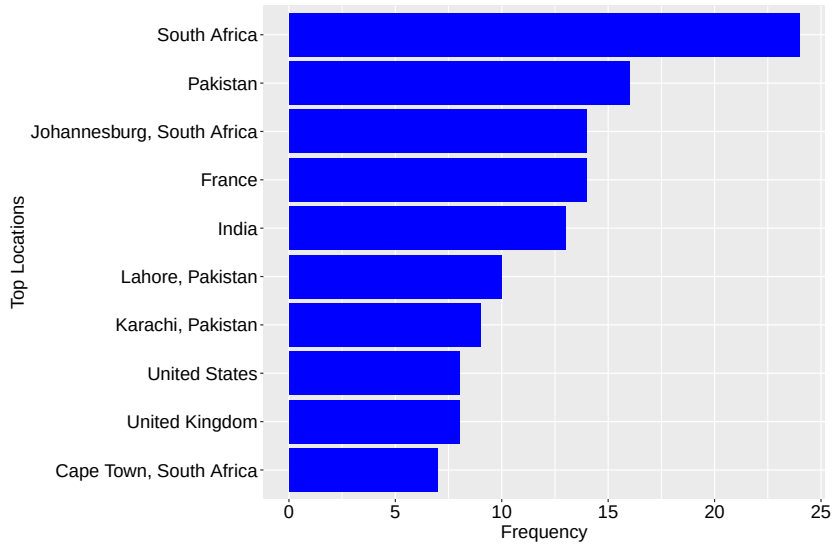


Figure 2: Plot Showing Users Tweets for Galaxyfold4 and Apple iphone14 by Location

User's Engagement and Analysis of Tweets by Location (Interpretation)

- The User Engagement analysis was explored to understand how users interact and engage on different topics, brands and products on Social media such as twitter
- Figure 1a shows a great level of retweet engagement for the tweets posts generated for SamsungGalaxyZFold4 tweets but less replies compared to Figure 1b where a uniform level of engagement was observed for the tweets, re-tweets and replies generated for Iphone14
- Results from further analysis on User engagement showed that most of the users retweeted posts on SamsungGalaxyZfold4 because a raffle draw was conducted at the period
- In addition, Figure 2a shows that the number of users tweeting about the new "Iphone14" was higher in Nigeria and United States while Figure 3b indicates that conversations on the new product "Samsung GalaxyZfold4" was observed to be more prominent in France and United States
- This findings suggest that a higher proportion of Iphone users or Iphone enthusiast might be from these locations. Also, United States was observed to have tweet conversations activity for both products.
- Another possible reason for the uneven data set is that people in a particular geolocation tend to tweet more about a particular product compared to the other

Time Series Analysis

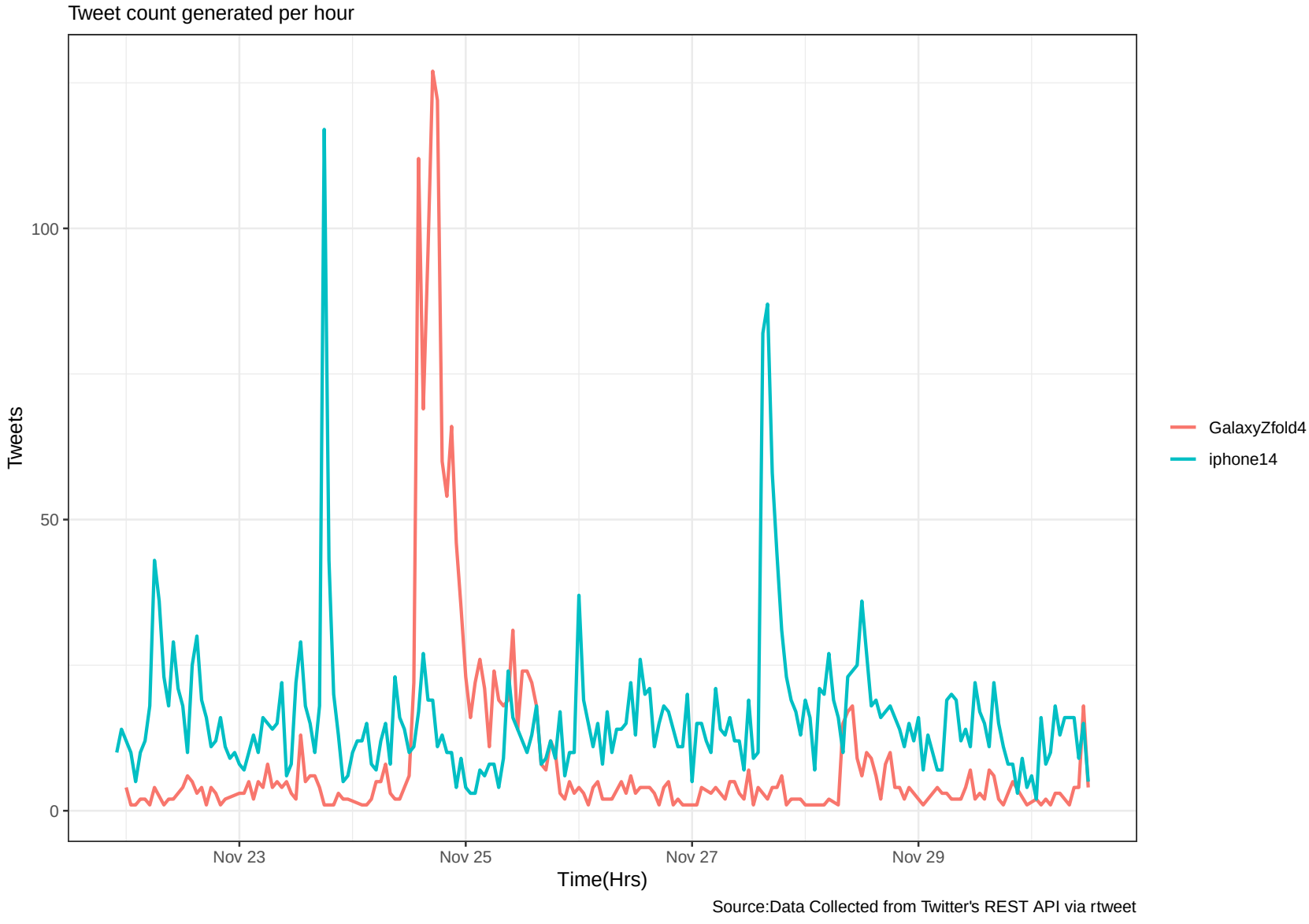


Figure 3: Plot Showing Time Series Analysis for Galaxyfold4 and iphone14

Time Series Analysis(Interpretation)

- This chart describes the frequency of tweets gotten for these products over time. this analysis was done to determine the interest level of the products by Users
- Fig.3 shows that the period for the tweets generated for Samsung GalaxyZfold4 within the 60 minutes interval was observed to have started from November 19th and engagement continued which gradually increased on 22nd of November and it peak b November 23rd, engagement on the tweet continued and gradually reduced from 25th of November down to 27th of November
- Fig.3 shows that the period for the tweets generated for Iphone14 within the 60 minutes interval was observed to have started from November 19th with more engagement per time throughout.
- The peak period was observed on 27th of November into 28th of November however, more people tweeted about the product on the 21st of November because there's an even spread on this day
- From the time series plot for both products, the Apple product Iphone14 has more engagements than Samsung product GalaxyZfold4.Twitter users are engaging more on the Apple products compared to the samsung products
- This could imply that there are more users of the Apple product Iphone14 than that of the Samsung users SamsungGalaxyZfold4

Sentiment analysis

Word Cloud

Figure 4: Word Cloud for Samsung Galaxyfold4 and Apple iphone14

Sentiment Analysis of Tweets for both Samsung Galaxyfold4 and Iphone14

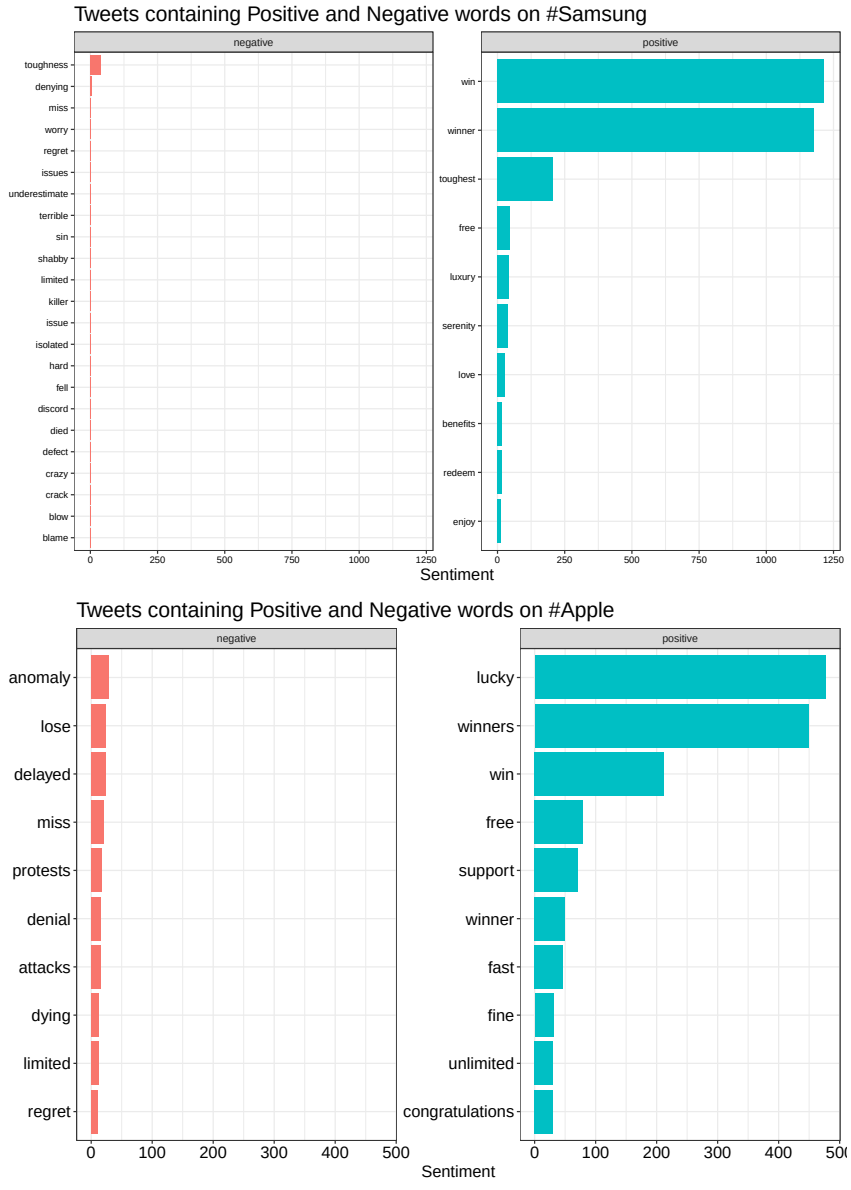


Figure 5: Tweets containing Positive and Negative words for both Samsung Galaxyfold4 and Iphone14

Sentiment Analysis(Interpretation)

- Sentiment Analysis is the process of examining the opinions, views and sentiments of users or consumers across social media platforms on different business topics, products, or services (Hu and Liu, 2004).
- This presents the opinion or views of people on a particular product or topic based on the frequency of positive and negative words gotten from the sentiment analysis
- Figure 5 shows that people used more of positive words while engaging these two latest products on Twitter
- Tweets generated on SamsungGalaxyZfold4 had the highest level of positive sentiment as only a few negative words were observed from the tweets. In the same vein, Iphone14 tweets also had positive tweets on "winning or being lucky" but also had negative words centered on "losing/delay the competition"
- The sentiment analysis of users was highly dependent on the activities trending in twitter at the period of this analysis

NORMALITY

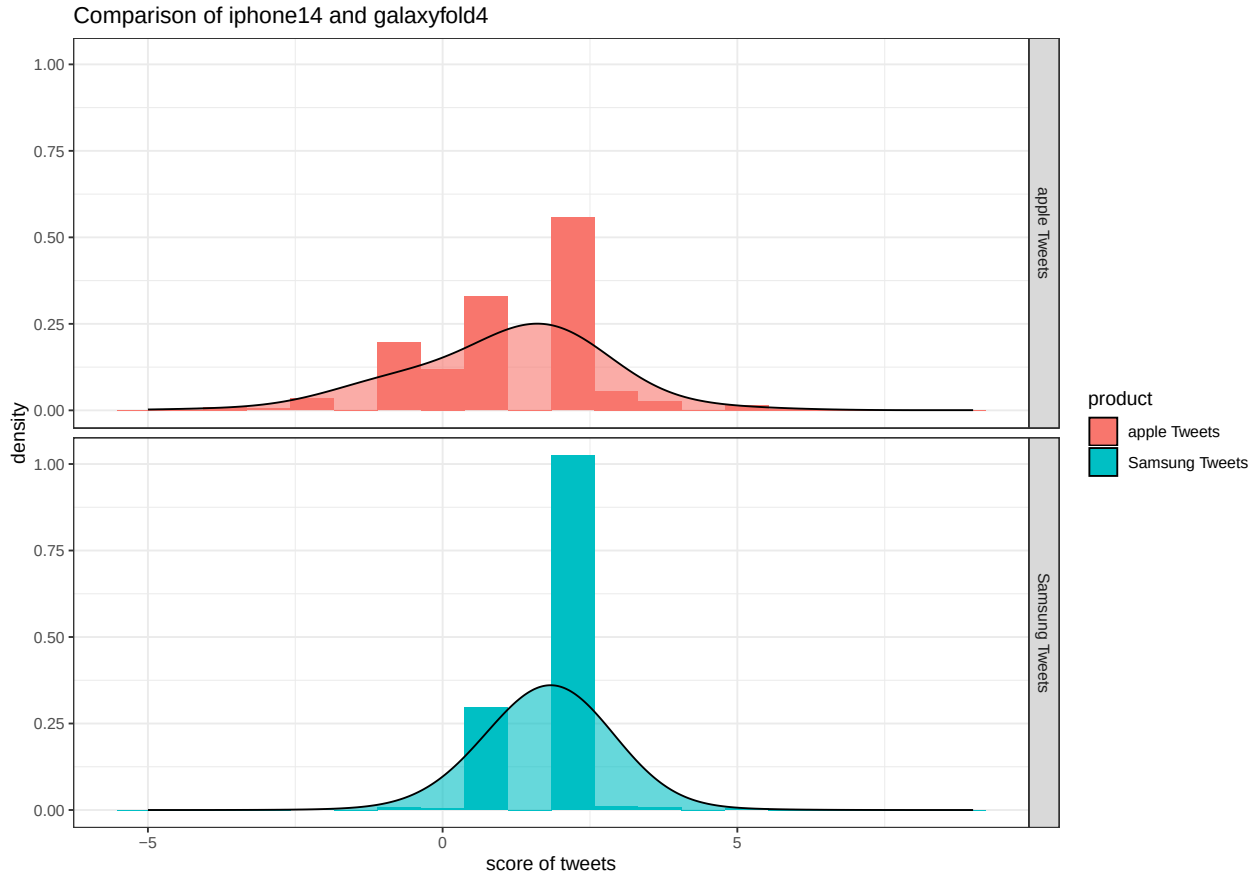


Figure 6: Comparison of Samsung GalaxyZfold4 and Apple iphone14

Statistical Analysis

T-TEST ANALYSIS

- T test analysis is done to check if there exist an underlying difference of the mean sentiment scores between iphone14 and galaxyzfold4
- An unpaired statistical t-test analysis will be the most appropriate statistical test because we have 2 independent groups
- The choice of the statistical t-test is dependent on if the sentiment scores are normally distributed and has equal variances
- Figure 6 revealed that both the sentiment scores for the products are not normally distributed as the Iphone14 sentiment scores are negatively skewed suggesting more negative opinions compared to the Samsung GalaxyZfold4 while sentiment scores for Samsung Galaxyzfold4 were concentrated around the mean (2.0) indicating more positive tweets
- This was confirmed by Shapiro-wilk test used to check for normality of the distribution , the p-value was determined as <2.2e-16. The null hypothesis that distribution is normal is rejected
- Therefore, the non-parametric t-test (Mann-Whitney Wilcoxon test) will be used for this statistical analysis
- The p-value determined by the Mann-Whitney Wilcoxon test c(W = 829821.5), NULL, 9.567400633576 35e-56, c(`location shift` = 0), two.sided, Wilcoxon rank sum test with continuity correction, score by product is less than 0.05, hence we reject the null hypothesis. The findings reveals that there is a difference in the preference of Users for both SamsungGalaxyZfold4 and Iphone14

Key Insights and Strategic Suggestions

- **User Interactions** vary widely on different topics/products on social media within a particular period of time, hence the sentiment analysis would be determined by the nature of engagement or activity on twitter at the point at which the data was extracted.
- Based on the statistical analysis,there exists a difference in User's preference for both products. Users had more positive opinion for galaxyzfold4, Hence the observed number of retweets indicating **user engagement**
- This study revealed that the twitter engagement was greatly determined by the **promotional messages** trending on twitter at the period at which the data was extracted.
- The companies can leverage more on the **advertisement potentials** of the social media space to promote their new offerings by engaging more through flash messages, branding, visibility etc

Project Limitation

- It must be noted that the following limited the robust analysis conducted on the two business products:

a. Inadequacy in the time in which the data was gathered and the difference in size of data
b. Social media data is not representative of the consumer population.
c. Unavailability of information in the tweets about specific functionality of the products
d. Representation of opinion not guaranteed because of promotional activity which swayed opinion in favour of one brand.