MATH513 Practical Presentation
Strategic Twitter Analysis: Samsung and Apple

10570155, 10696253, 10701983

10/12/2020

Introduction

- Samsung and Apple
- Flagship phones chosen
  - S20FE
  - iPhone12
  - S20

Tools Utilised

- Rstudio
- RTweet
- Twitter Developer API
- GitHub

Research

Choosing Twitter for Analysis

- Open API Access compared to others
- Almost all data is public
- Advanced filtering and queries
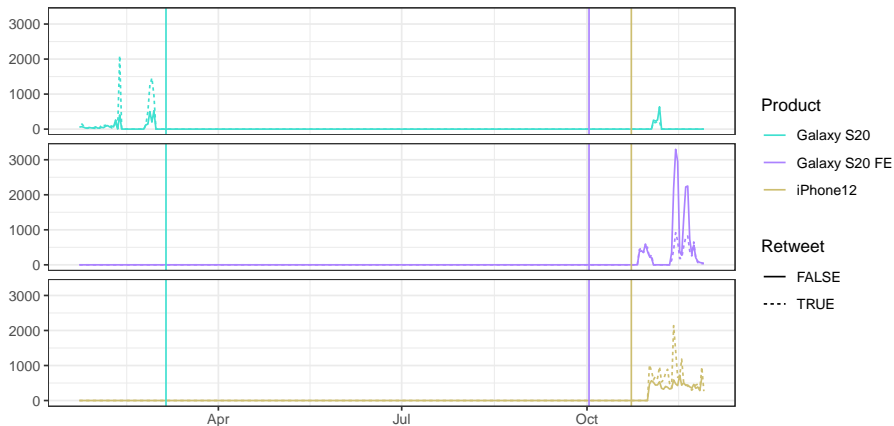- Generous Rate limiting

Determining Hashtags

- @SamsungMobile
  - #GalaxyS20FE and #GalaxyS20
- @Apple -
- @tim_cook -
  - #iPhone12

Time Periods for Data Collection

Frequency of Twitter Statuses

Twitter status counts aggregated using 1−day intervals

Source: Data collected from Twitter's REST API via rtweet

Data Cleaning and Feature Engineering

Data Cleaning

- Duplicate tweet and user observations were removed
- Tweet text and user bios were cleaned
  - Removed links, hash-tags, emojis, and user mentions

Feature Engineering

- Users were marked as potential bots
- User country was extracted from the location of their profile
- Tweets were marked as potential spam
- Hash-tags were extracted from the tweet text
- Product features were extracted from the tweet text
  - Display, Battery, Camera, Price, and 5G Capability

Summary of Data

Total Tweets: 73690 after data cleaning

Total Features: 5 (Display, Battery, Camera, Price, and 5G)

Table 1: Summary of Tweet Data

Product	Number of Tweets	% Spam Tweets	% Feature Tweets
Galaxy S20	13147	3%	20%
Galaxy S20 FE	28923	19%	19%
iPhone12	31620	13%	7%

Table 2: Summary of User Data

Number of Users	% Bot Users	Unique Countries
35051	>1%	163

Results - Sentiment Analysis - All Tweets

Distribution of Sentiment Score Across Tweets
Tweets faceted by product

Mean: 0.140508

Mean: 0.222527

Mean: 0.13162

Density

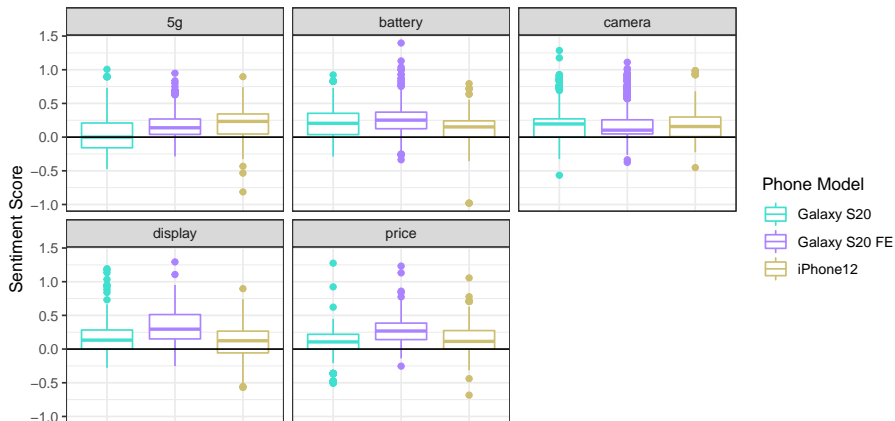Sentiment Score

Product: Galaxy S20   Galaxy S20 FE   iPhone12
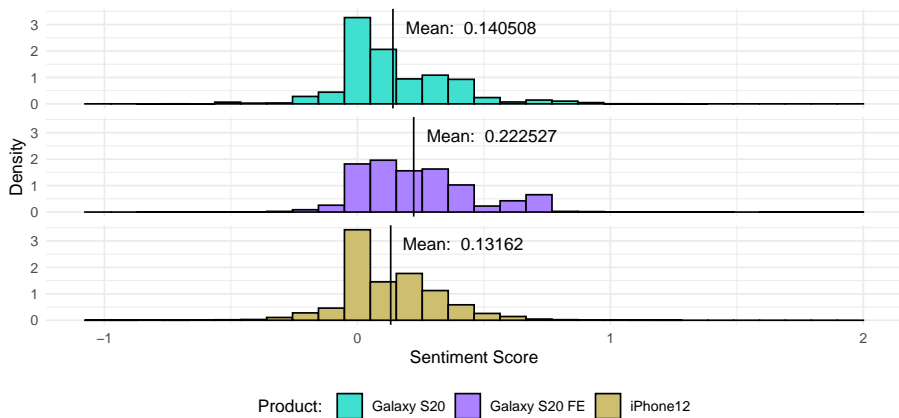
Source: Data collected from Twitter's REST API via rtweet

Results - Sentiment Analysis - Features

Distribution of Sentiment Score Across Tweets
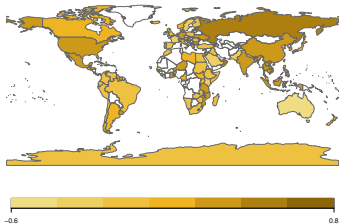Tweets faceted by mentioned product features

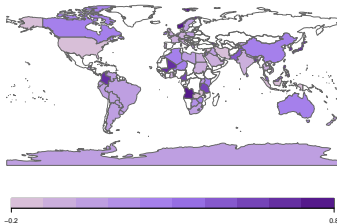Source: Data collected from Twitter's REST API via rtweet

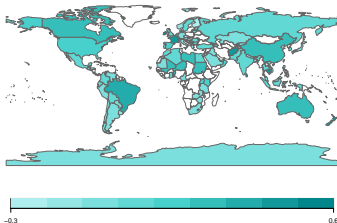# Results - Global Sentiment By Product

iPhone12 Sentiment by Country

Galaxy S20 FE Sentiment by Country

Galaxy S20 Sentiment by Country

Statistical Test - T-Test

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  vals$s20
## D = 0.40721, p-value < 2.2e-16
## alternative hypothesis: two-sided

##
##  One-sample Kolmogorov-Smirnov test
##
## data:  vals$s20fe
## D = 0.45695, p-value < 2.2e-16
## alternative hypothesis: two-sided

##
##  One-sample Kolmogorov-Smirnov test
##
```

Conclusions

Twitter data useful Region

- Twitter data provides up to date information for companies to analyse customer feedback
- Data can provide useful information to guide product teams when analysed correctly
- With the world map data, which can be broken down into more specific regions, we can see key areas that need to have more targeted marketing or improve marketing to increase the sentiment

Apple

Samsung BTS

Improvements & Further Study

Improvements

- Google Maps API
- Look at mentions of apple in samsung and vice versa
- Examination of average income and sentiment by region
- Increased number of tweets with more targeted dates before and after the release dates

References

- Ahmed, Wasim (2019). *Using Twitter as a data source: an overview of social media research tools* Available at: https: //blogs.lse.ac.uk/impactofsocialsciences/2019/06/18/using-twitter- as-a-data-source-an-overview-of-social-media-research-tools-2019/ (Accessed: 07 December 2020)

- Dalla Valle, Luciana (2020). *MATH513 Lecture and Tutorial Code* Available at: https://dle.plymouth.ac.uk/course/view.php?id=49628 (Accessed: 01 October 2020)

- Fuchs, Matti (2018) *Doing your first sentiment analysis in R with Sentimentr* Available at: https://towardsdatascience.com/doing-your- first-sentiment-analysis-in-r-with-sentimentr-167855445132 (Accessed: 06 December 2020)

- Rinker, Tyler (2020). *R Documentation - sentiment_by* Available at: https://www.rdocumentation.org/packages/sentimentr/versions/2.7. 1/topics/sentiment_by (Accessed: 06 December 2020)

References

- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/
- RStudio (2020). *R Markdown Cheat Sheet* Available at: https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf (Accessed: 10 October 2020)
- RStudio (2014). *R Markdown Reference Guide* Available at: https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf (Accessed: 10 October 2020)
- Swider, Matt (2020). *Twitter hack exploits Apple, Elon Musk and other prominent accounts* Available at: https://www.techradar.com/news/twitter-hack-2020 (Accessed: 10 October 2020
- Twitter (2020). *API Documentation* Available at: https://developer.twitter.com/en/docs/twitter-api (Accessed: 10 October 2020)
- Young, Michelle (2017). *Twitter Data Mining: A Guide to Big Data*