Network and Sentiment Analysis: iPhone11

Team 40, Section B
Arwen Wang, Kewei Jiang, Kieu Anh Nguyen, Taru Dharra and Vedant Sahay

Introduction and Product Selection

At its fall 2019 event, Apple Inc. announced that the company's latest smartphones iPhone 11, 11 Pro and 11 Pro Max were being released on September 20th. From its release of the first generation of iPhone, the leading technology company has attracted fans from all over the world and received tons of attention.

What comes with attention are customers' reviews and hot debates about the products on the internet. Given the high price of the iPhone compared to other smartphones in the market as well as the cutting edge technology behind the product, the release of the newest generation iPhone always generates discussions on Twitter in both positive and negative ways.

Given the fact that online customer-generated content is becoming increasingly influential, we extracted the most recent 1200 tweets with the hashtag #iPhone11, performed sentiment analysis and social network analysis based on the documents to look into customers' attitude toward the product and generate insights for the company's digital marketing strategies for the product.

Text Mining and Visualization

After extracting 1200 latest tweets about iPhone11, we firstly converted all alphabet to lowercase, then removed punctuation, numbers, stop-words, and URLs from the documents.

After cleaning the data, we transfer the tweets documents into the term-document matrix (Figure 1) which shows how many times each term appears in each document.

	Docs																		
Terms	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
cost	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gold	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
green	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
midnight	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
mysterybox	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
online	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
pro	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
sale	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
unbox	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
crazy	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 1. Term Document Matrix

To take a closer look, we used a bar plot (Figure 2) to show terms that appeared more than 30 times in the documents.

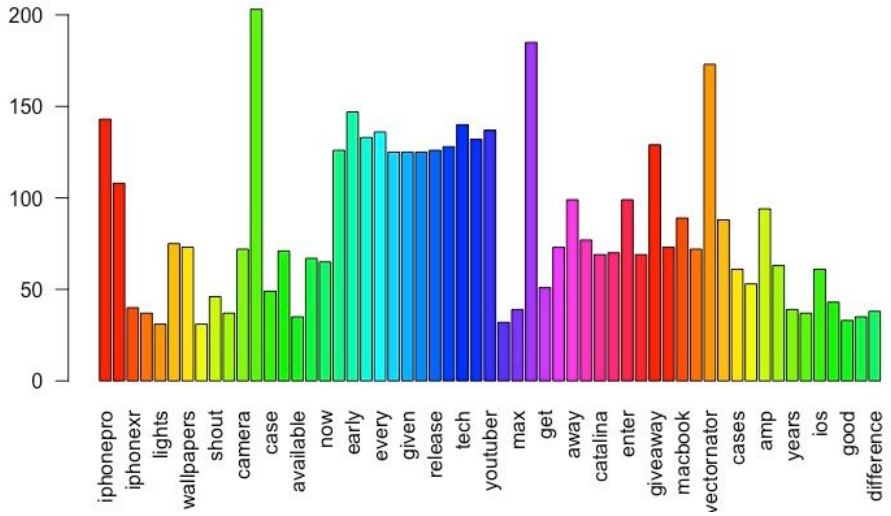

Figure 2. Bar Plot for Terms

We also generated a word cloud (Figure 3) to present the text mining result more intuitively.

Figure 3. Word Cloud

From the bar plot and word cloud shown above, we came to several findings of iPhone11 from discussions on social media:

- Amongst the three types of *iPhone11, 11 Pro and 11 Pro Max* generated more interests among potential customers compared with iPhone11

- The term *Vectornator*, which is the name of a vector graphic design app, appeared 175 times in the documents. This indicates that people paid much attention to apps that are suitable for the new operating system
- Public also has given much attention to the three lense camera and the new iOS operating system of iPhone11, which are two of the most important upgrades in iPhone11 compared with the former generations
- The high frequency of the terms wallpapers and case also indicates that customers are interested in finding accessories for their new phones
- Surprisingly, we found that unboxing and Youtuber are two of the terms that appeared the most in the latest 1200 tweets, which may indicate that social media and customer-generated content have a huge impact on potential customers' behavior
- We noticed that adjective words like *incredible, best, good, great* appeared very frequently, showing people's positive attitude towards iPhone 11

Sentiment Analysis

The text mining and visualization results indicate that users on twitter are mostly showing a positive attitude towards iPhone11. However, to confirm this finding while taking a closer look at the public emotion about this new product, we performed sentiment analysis of the documents.

We divided people's emotions into eight categories to evaluate the sentiment of each document (examples given in figure 4).

	anger	anticipation	disgust	fear	joy	sadness	surprise	trust	negative	positive
1	0	0	0	0	1	0	0	1	0	2
2	1	0	0	1	0	1	0	0	1	1
3	0	0	0	0	1	0	0	1	0	2
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	1	0	2
6	0	0	0	0	0	0	0	0	0	0

Figure 4. Sentiment Matrix

To calculate the overall sentiment score of the 1200 documents, we summed the column scores across tweets and used a bar plot to visualize (figure 5):

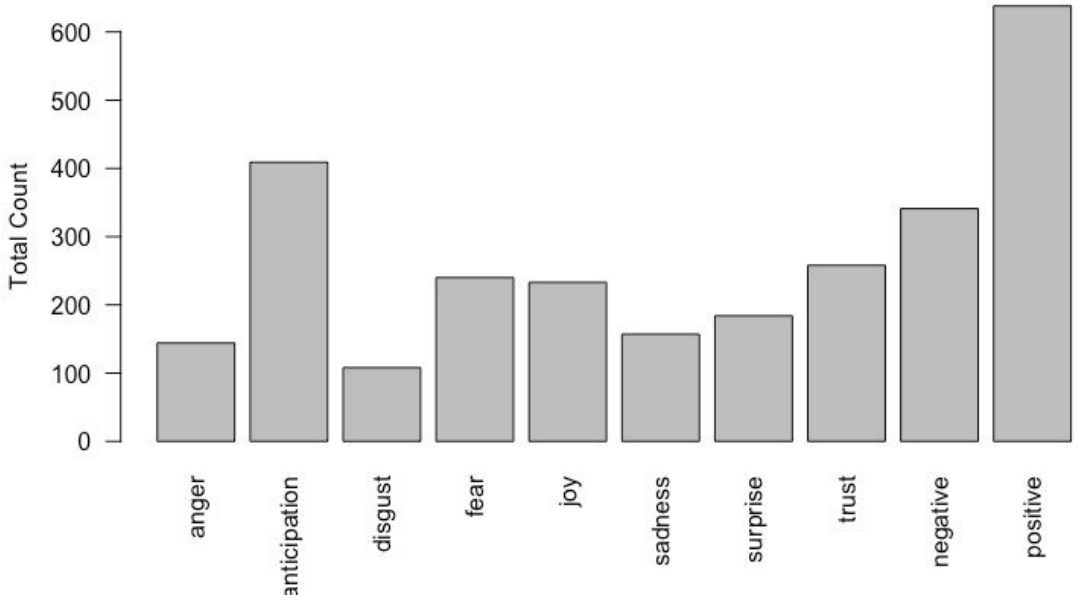

Figure 5. Sentiment Scores for iPhone11 Tweets

As the bar plot shows, positive and anticipation received the highest and the second-highest sentiment scores among the eight kinds of emotions. This finding confirmed our assumption based on the text mining process that twitter users mostly project a positive attitude towards iPhone11. Additionally, we can also tell from the scores that people have trust in the new product (possibly based on their experience of the former generations).

However, it is also shown in the plot that while the majority of the tweets have positive content about iPhone11, some have negative, fear and anger attitude as well. Moreover, it seems like the new product does not give the public much surprise.

Overall, according to the sentiment scores for the eight emotion categories, we can tell that iPhone11 has received a positive review and a good reputation on Twitter.

Network Analysis and Centrality Measures

Besides text mining and sentiment analysis, we also performed network analysis with centrality measures to test the relationships between the tweets.

To do so, we first transformed the term-document matrix into the term-term matrix (figure 6) which shows the occurrences of each pair of terms:

Terms	Terms cost	gold	green	midnight	mysterybox	online	pro	sale	unbox	crazy
cost	4	4	4	4	4	4	4	4	4	0
gold	4	6	5	4	4	4	5	4	4	0
green	4	5	10	7	4	4	5	4	4	0
midnight	4	4	7	8	4	4	4	4	4	0
mysterybox	4	4	4	4	4	4	4	4	4	0
online	4	4	4	4	4	4	4	4	4	0
pro	4	5	5	4	4	4	169	4	4	0
sale	4	4	4	4	4	4	4	7	4	0
unbox	4	4	4	4	4	4	4	4	4	0
crazy	0	0	0	0	0	0	0	0	0	1

Figure 6. Term-Term Matrix

Based on this term-term matrix, we used centrality matrix to plot and evaluate the quality of the network.

Degree Centrality

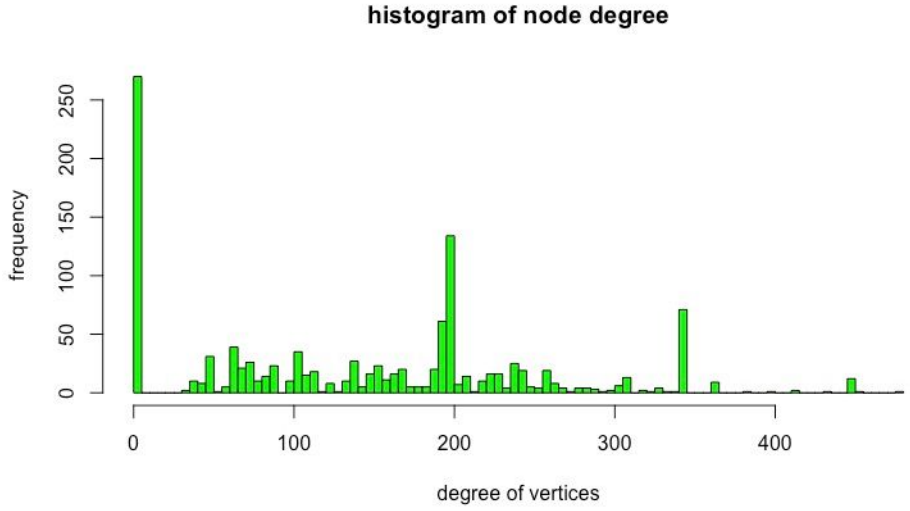

Figure 7. Histogram of Nodes Degree

As per the bar plot above, we see that the majority of the nodes have zero degrees of centrality indicating that most of the nodes in the network do not have a neighbor. However, there are spikes such as at degree ~195 and ~335 indicating a dense network of nodes.

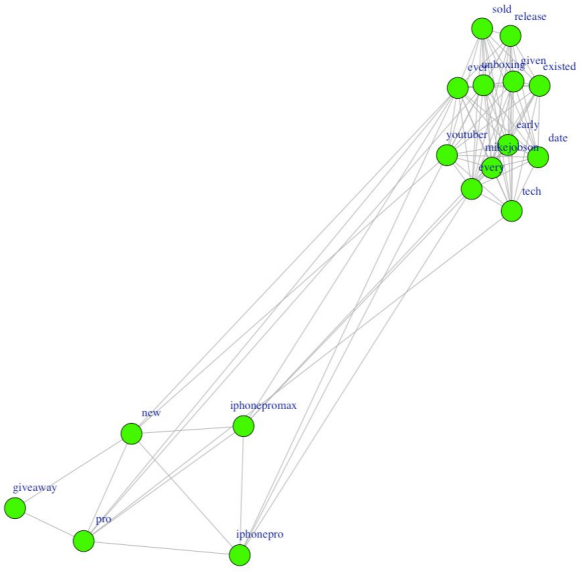

Figure 8. Network based on node degree

Based on the network graph above, we see a dense cluster with '*Unboxing*' as of the nodes with high degree of centrality. This can be expected as with every release of a new iPhone, unboxing videos are anticipated and viewed on platforms such as Youtube and hence it is not surprising that '*Youtube*' and '*Unboxing*' are neighbors.

Betweenness Centrality

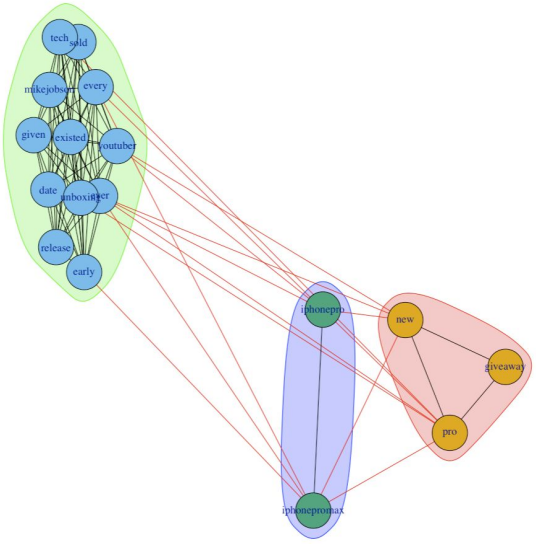

Figure 9. Network based on edge betweenness

As per the network above, we see that the green cluster with blue nodes is a dense arrangement of nodes with high betweenness centrality. Nodes such as '*Youtuber','early','sold'* (in the leftmost cluster) are connected to other networks that have a high betweenness amongst their nodes(iPhone Pro and iPhone Pro Max). Since these nodes

act as bridges between clusters, removal from the network will cause the most disruptions as they lie on the largest number of path taken.

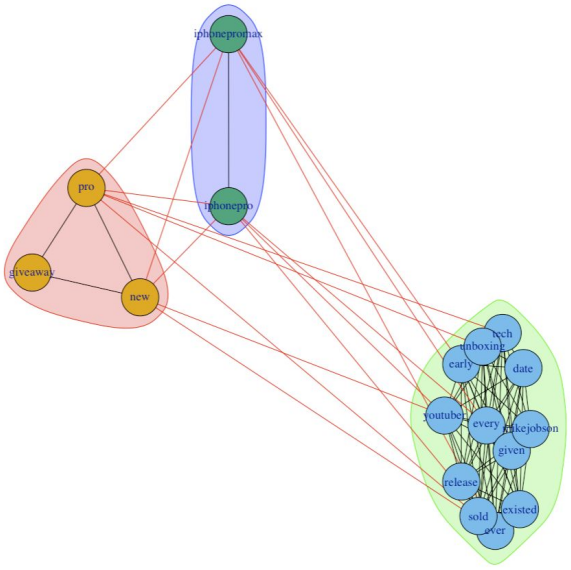

Figure 10. Network based on Label Propagation

Label propagation is an approach in which every node is assigned one of k labels. The method then proceeds iteratively and re-assigns labels to nodes in a way that each node takes the most frequent label of its neighbors in a synchronous manner. The method stops when the label of each node is one of the most frequent labels in its neighborhood. The network generated by the label propagation method yielded the same result as using betweenness centrality.

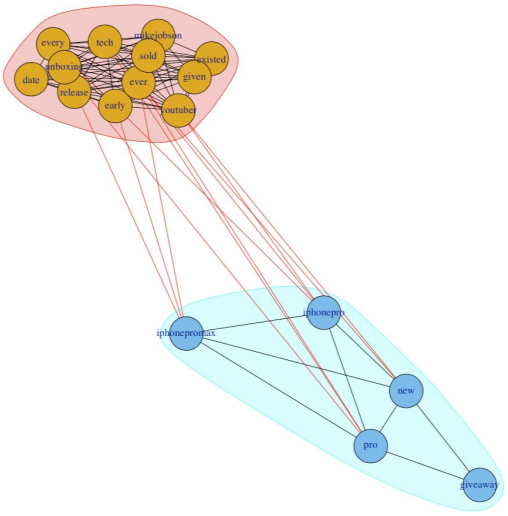

Figure 11. Network based on Greedy Algorithm

The Greedy method starts with nodes in separate clusters, and then merges clusters together in a greedy fashion. This method has less clusters compared to the last two methods since it

joined the two smaller clusters that appears for the betweenness centrality method and the label propagation method to one bigger cluster. In the new bigger cluster, we can see that this group is simply talking about the phone and a new giveaway.

Conclusion

- Our sentiment analysis showed that the general sentiment about Iphone 11 is very positive. People are generally excited about the launch of this product and expressed trust towards the brand.
- A big part of what people are talking about is the online buzz around unboxing videos on Youtube.
- A lot of talk about the phone is about it's new camera features and it's sophistication.
- People also seem to be excited about the photo application that works well with the new camera feature.
- Given the fact that social media and consumer-generated content are becoming increasingly influential, companies should invest heavily into Market Mix models to understand the customer attributes and seasonality.