Report of Case Study 1: The Study of User Responses to the iOS10 Based on the Twitter

Group 14 members: Shi He, You Zhou,
Xiao-Qun Wang, Zili Ma,Meng Wang

September 22, 2016

1 What data we collected?

Weve collected about 730,000 tweets by using the twitter livestream api. The topic we selected is iOS10. Apple announced iOS10 in this June at WWDC and claimed that it would be the most great update to the iOS devices with multiple changes and new features. Since it was officially released in earlier September. We decided to collected people's tweets to check out their different reactions to this update.

2 Why this topic is interesting or important to us? (Motivations)

Firstly, the release of iOS is one of the biggest events to the whole world. With the new iPhone coming up, this update has attracted everyone who owns iOS devices. As the world's second most favourite mobile operating system, people must have tons of things to say about it after they upgrade their devices with the newest system.

Secondly, compared with Android users, iOS users are most likely to update their system immediately after the officially release. From the viewpoint of business, it will be very important for the company to analyze the very first feedback of customers, to find out the latent problems complained or new features preferred by customers...

Nowadays, the hardware performance gap between the new generation iPhone and other competitors' products is becoming much more narrow than before. The battlefield has been shifted to the software, in other words, competition between iOS and Android gradually. Since Apples iPhone series have dominated the market of mobile phone for years but started to be caught up with by other competitors' products, whether the latest iOS 10 can become popular at the first two weeks will be a key index to predict the sales of its latest product iPhone 7.

3 How did we analyze the data?

In general, the steps of data collecting and processing are as follows. For the description in details, please refer to the notebook file we attached.

1. Collecting: All streaming data collected from twitter API was saved in the JSON file, these are the raw data without filtering and cleaning.
2. Filtering: The collected tweets were divided into two JSON files by identifying whether it was an original tweet or a retweet.
3. Cleaning: Then we used Pandas to read the data from the JSON files and to remove the duplicates. Then JSON data were stored in the format of DataFrame.
4. Storing: We use pandas function to convert DataFrame into cvs file accordingly.
5. Grouping: We divided the tweets into different categories like 'time', 'time zone' and 'Source'.
6. Visualizing: We used matplotlib to generate chart and used nltk to process the text.
7. Analyzing: We used online text analyzing service API to help us analyzing the text.

4 What did we find in the data?

In order to analyze the data, we categorized them into different groups based on the location (time zone), source and sending time of tweets. Again, based on these groups, we conducted emotion analysis among them. It is worth noting that due to the large amount of data we've collected, some samples are taken from each group with the confidence level of 95% and confidence interval of 5.

What's more, we used the sentimental analyzing service provided by Vivek Narayanan and text corpus were based on the IMDB movie reviews. We applied this method to the samples and to gain an insight of what people's attitude.

4.1 The number of tweets v.s. locations (Time zones)

Because not every tweets contained the geological information, we decided to use 'time-zone' to represent the locations for different users. From the data we collected, the top 20 regions to tweet the topic of iOS10 are shown in Table 1. US and Canada made the most of tweets in the time window we observed. It is not surprised that the impact caused by iOS10 release indicated by the number of tweets on the west coast is larger than that on the east coast in north America, because Apple is typical Silicon Valley company with more fans and employees in the west coast. Because we only collected the twitter written in English, it is not surprised that most of the tweets are from the English-speaking regions.

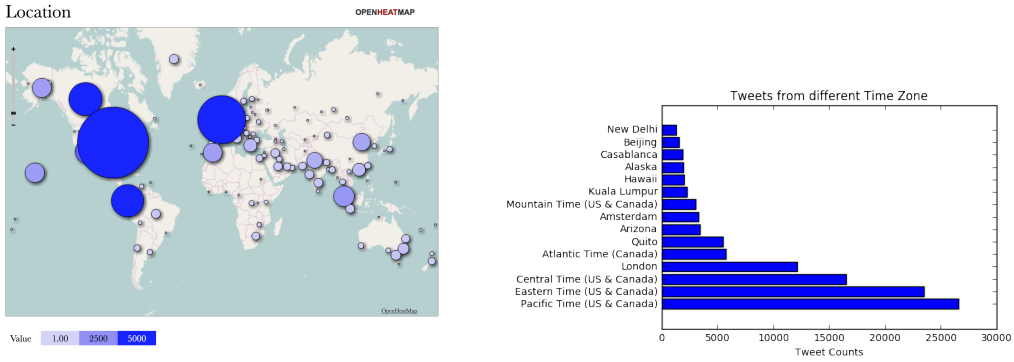
By using the location information extracted from the time zones, we draw the heat map (as shown in Fig. 1) to reflect the activities of international users on twitter to talk about iOS10. It is very interesting that here we find the most popular region to talk about iOS10

in Asia is Kuala Lumpur, the capital city of Malaysia (ranking 10). Beijing, the capital city of China, follows closely behind Kuala Lumpur. while, the 3rd and 4th most active cities on twitter in Asia are Singapore (ranking 17) and Hong Kong (ranking 19), respectively. This behavior of users implies that maybe the local economic level does not directly decide the motive of customers in these region to buy an expensive digital product in Asia.

We chose samples from top 5 groups separately and then we applied the sentimental analysis (as shown in Fig. 2).

Table 1: Top 20 regions making the most tweets about iOS

Time Zone	Tweet Count
Pacific Time (US & Canada)	26568
Eastern Time (US & Canada)	23502
Central Time (US & Canada)	16517
London	12154
Atlantic Time (Canada)	5760
Quito	5503
Arizona	3429
Amsterdam	3310
Mountain Time (US & Canada)	3044
Kuala Lumpur	2232
Hawaii	1989
Alaska	1912
Casablanca	1868
Beijing	1583
New Delhi	1276
Dublin	977
Singapore	956
Athens	901
Hong Kong	859
Sydney	614
Sum	114954



(a) (b)

Figure 1: Heat map of 'iOS' in the world (a) and bar chart of top 15 regions with highest tweet numbers (b)

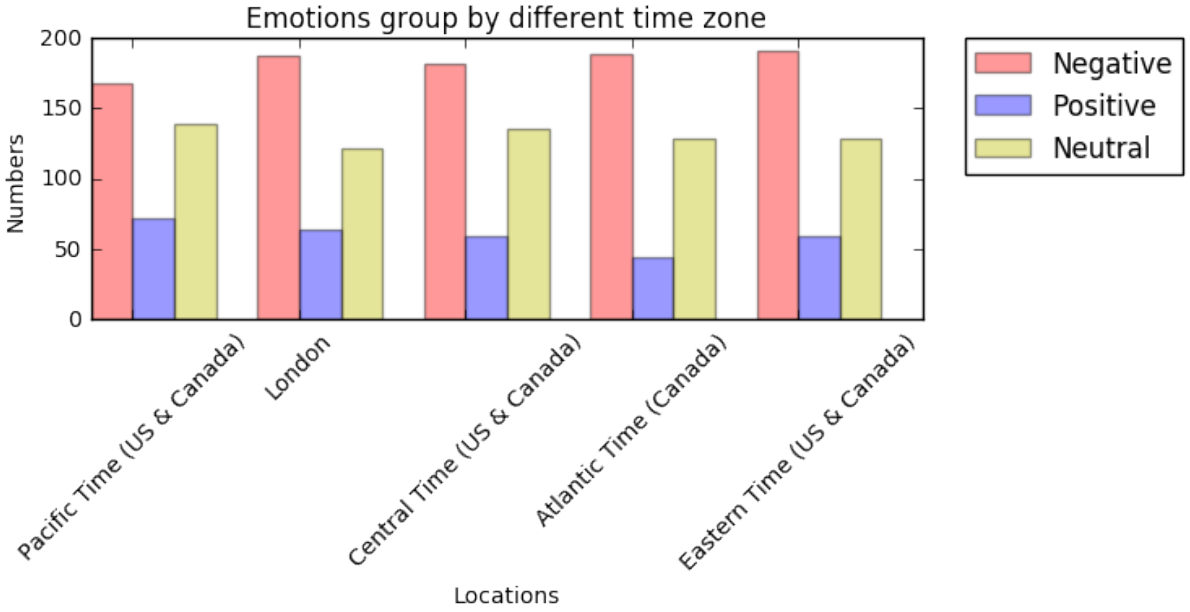


Figure 2: Sentimental analysis in 5 groups

4.2 The number of tweets v.s. types of devices

As shown in Fig. 3, most of tweets we collected came from iPhone. It is reasonable because the current users of iOS system are the main groups of people concerning upgrading to latest version. Here, we observed that Android users contributed the 3rd most tweets about this iOS event. The interactions between the users of iOS and Android as the two super power phenomenon in the world of smart phones camps are increasingly becoming the most interesting phenomenon, because during the evolution of both types of mobile OS, they always learn from each other and absorb the desirable features from its competitor.

We chose samples from top 5 groups separately and then we applied the sentimental analyse. The result is shown in Fig. 4.

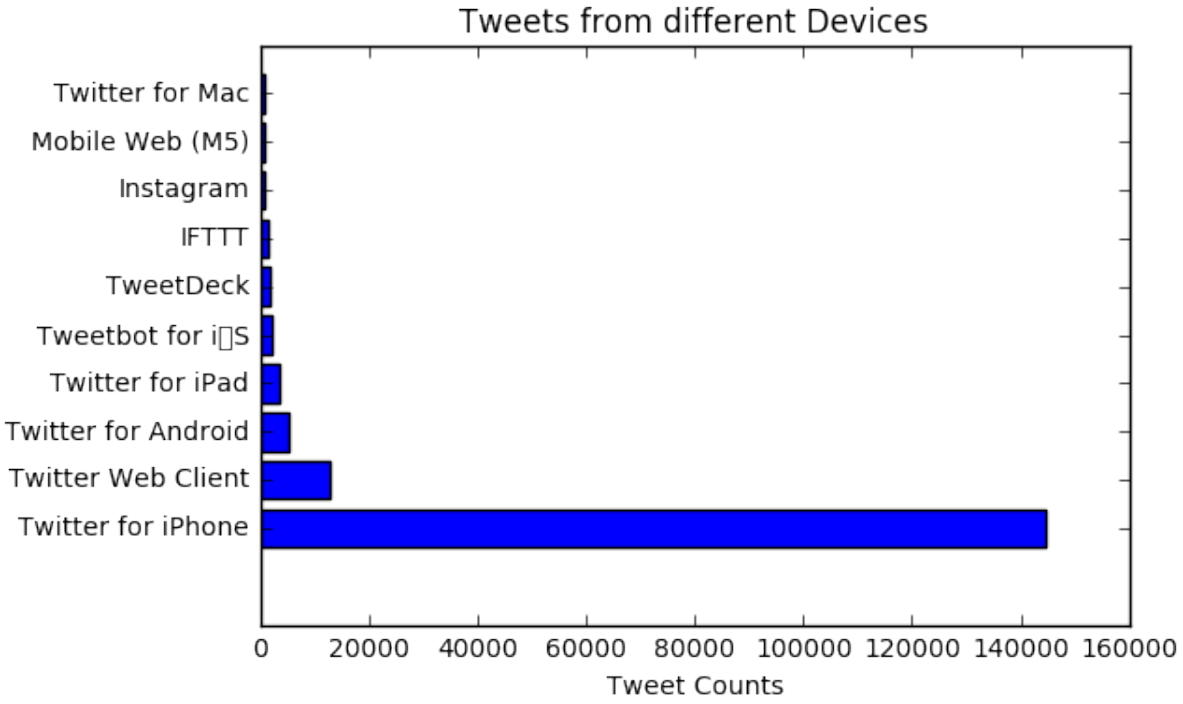


Figure 3: Number of tweets form different devices

There is no surprise that iPhone users are the part of group with the least negative and most positive messages. From the bar chart, we can see that only the tweets from Android shows a higher number of neutral comments than the negative comment number.

4.3 The number of tweets v.s. time

The time analysis of tweet number shown in Fig. 5 shows that this topic stayed popular from the day when iOS10 release (Sep. 13th), and became even hotter to spread on the second day of WWDC (Sep. 14th). Then, the tweet number decreased in the third day.

Again, We chose samples from groups separately and then we applied the sentimental analysis (as shown in Fig. 6).
According to the sentimental analysis, It is not hard to find out that people are not satisfied with iOS10, because we can find in every categories almost 40% or more people's

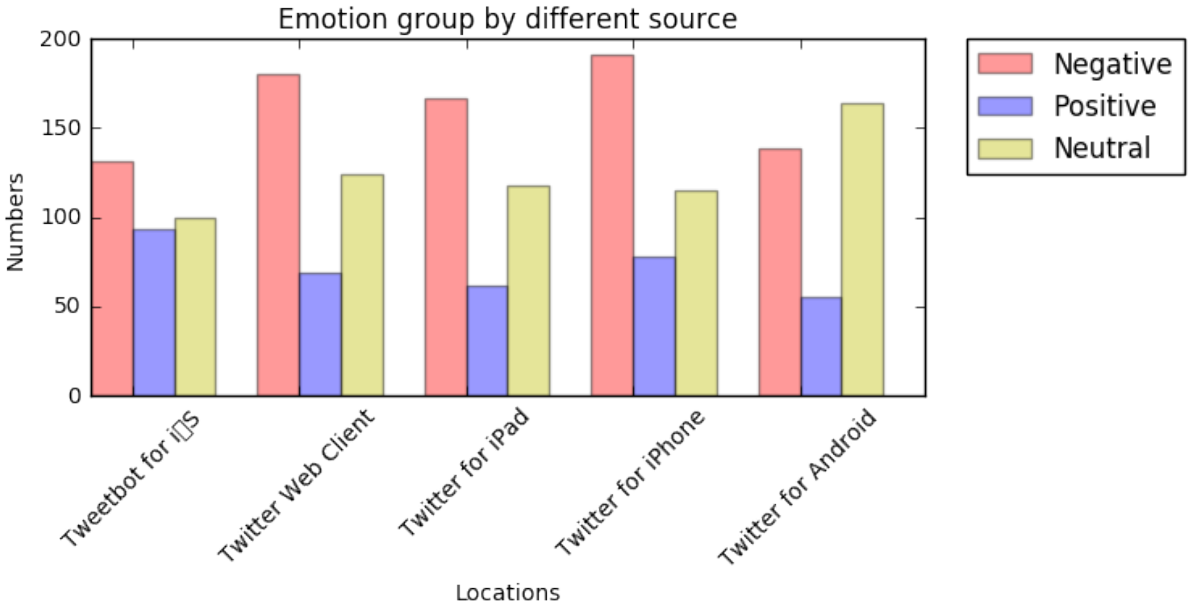


Figure 4: Sentimental analysis among 5 device groups

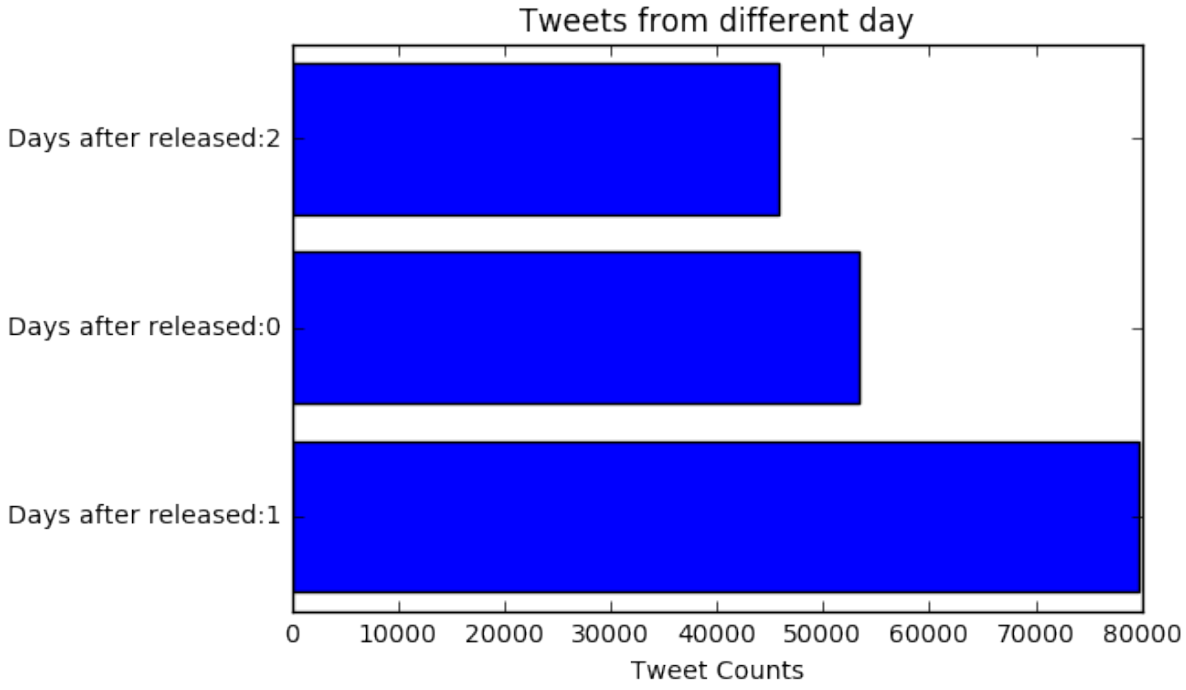


Figure 5: Number of tweets form different posted time

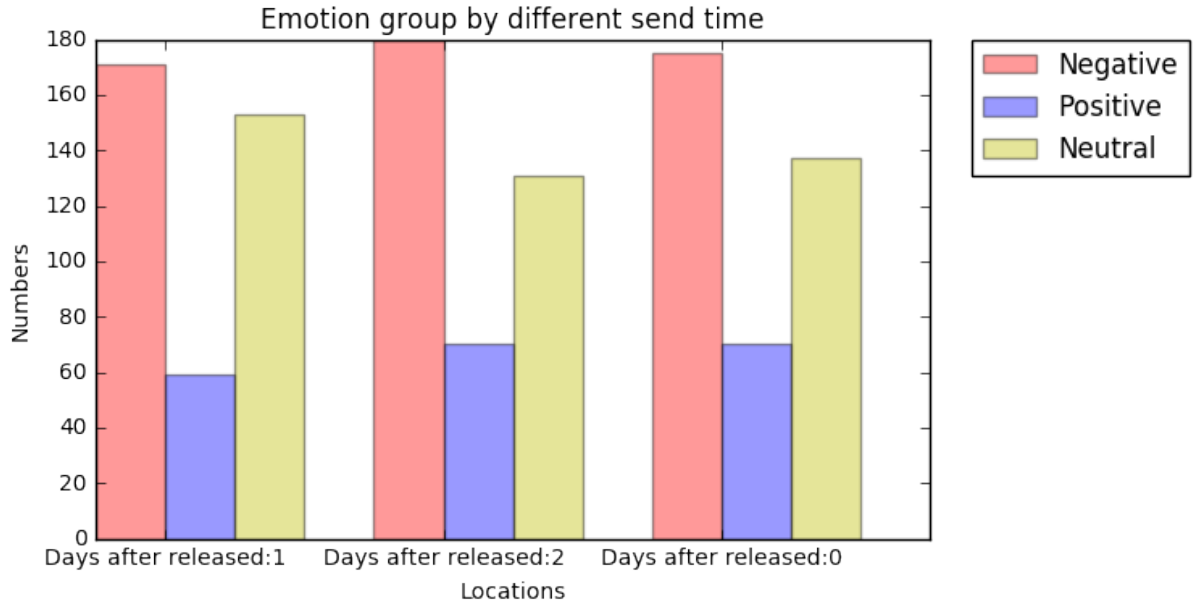


Figure 6: Sentimental analysis among 5 posted time

tweet contain negative information. Compare to that, the positive tweets can't make it to even 30%. This is kind of a failure. But, people in the neutral zone are just like the voters in swing states. If Apple can win those peoples heart, then the iOS10 would not be a total fiasco. But how should Apple make next decision? We came up with the idea to check up what people's reaction to the new features of iOS, whether it is a good change or not.

4.4 Sentimental analysis for the new iOS10's new features

In order to know the users' attitude to some of the iOS10's new features, again, we used the sentimental analysis method.

Firstly, we filtered the tweets with four keywords and the result list is as follows:

1. "music" for the study of customers' attitude to new music app;
2. "emoji" for the study of customers' attitude to iOS10's new design of emoji;
3. "unlock" for the study of customers' attitude to iOS10's new ways to unlock;
4. "ui" for the study of customers' attitude to iOS10s new UI style.

Correspondingly, the number of each keyword's search result are shown in Table 2. The most comments about iOS10's new features focused on the change of emoji, it is probably because it is easiest features to make comments, everyone can express their ideas about these simple cartoons, while for other features in our study, such as UI, users need to have some experience and knowledge about the old version of iOS, because what UI stands for may even not make sense to some people.
Then, we conducted sentimental analysis on above results respectively. The return results are displayed as follows:

Table 2: The number of different tweets of different keywords

Keyword	Number
music	2207
emoji	16550
unlock	3840
ui	253

```
music : {'Neutral': 448, 'Positive': 475, 'Negative': 1284},
emoji : {'Neutral': 4168, 'Positive': 2044, 'Negative': 10338},
unlock: {'Neutral': 968, 'Positive': 482, 'Negative': 2390},
ui    : {'Neutral': 802, 'Positive': 544, 'Negative': 1186}.
```

The bar chart of each keyword is shown in Fig. 7. The percentage table of each attitude is shown in Table 3. Generally, the majority of the comments about the new iOS are negative. From the histograms, we can see that for all the keywords, users who have negative attitude are much more than people who have positive attitude.

The highest negative percentages occurred in the topic of emoji and unlock. UI has the lowest negative percentage, probably there was not much changes of it. Does the result with high fraction of complaint to the new iOS represent the true attitude of users who are using iOS10? Probably not, people tend to spend more time to complain about something than praise it, especially when the thing they complain about has been a classical pop icon like iPhone and iOS.

However, we can still have some conclusions in the complaint rate: Emoji and unlock interface are considered to have the most changes in iOS10, so they will understandably gain most of the comments, no matter in positive side or negative side. Compared with other positive rate, the least value falls on Emoji and Unlock, so we could say that most of the twitter users' are not fond of iOS10's Emoji and unlock features. In contrast, Apple Music has gained the most positive response, suggesting that Apple Music has been improved better and more user-friendly.

At some level, this negative shows the failure of Apple making people comfortable with these changes. People don't like too much aggressive changes to their daily using devices. That makes people uncomfortable. What's more, people mainly using iPhone to listen to music and send text with emojis. When people finally figure out how to unlock the phone, they open the brand new music app and see total redesign of the emojis, they have much probability to give negative response to the new iOS.

Table 3: The number of different tweets of different keywords

Percentage	Neutral	Positive	Negative
Music	20.3%	21.5%	58.2%
Emoji	25.2%	12.4%	62.5%
Unlock	25.2%	12.6%	62.2%
UI	31.7%	21.5%	46.8%

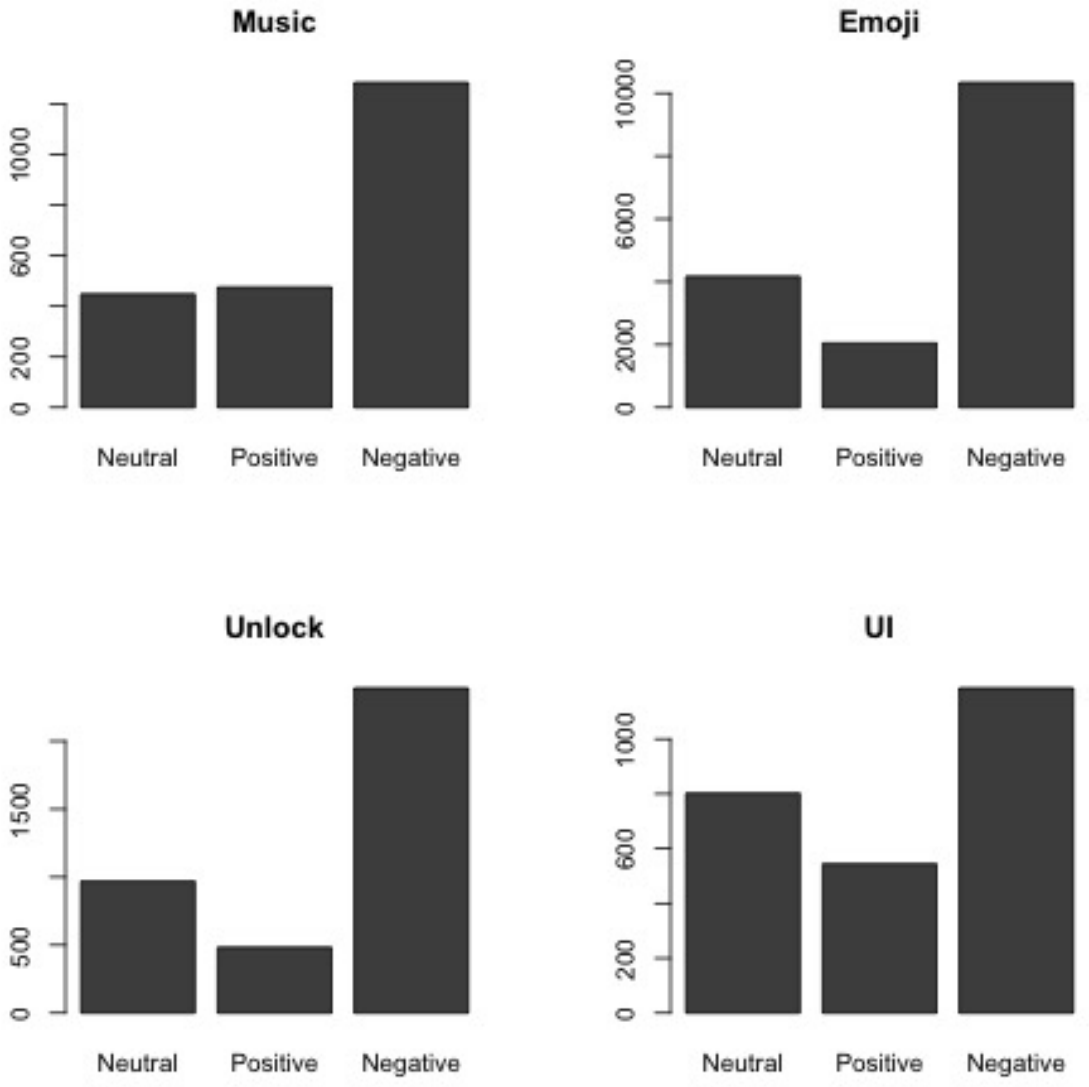


Figure 7: Sentimental analysis of iOS10's new features

4.5 Conclusions

In this report, we described in details the methods we used and the results in the Case 1 study of our group.

In general, we used the live-stream API tools provided by twitter.com to colleted the popular topic we are interested. Here, we choose the topic of iOS10. Because the time window of its release matched exactly our pace of course. By crawling from the API for 3 days, we extracted about 730,000 tweets to conduct our study. The original responses given in JSON format was recognized and filtered by python and the relative libs to get the really target of interest, namely the original tweets sent by users in this study.

We used word counts to visualize the trend hidden behind the data samples. Here, we analyse the word counts from several different dimensions, such as, time zones, source type of the devices, time and so on.

The sentimental analysis is demonstrated in our report to measure the feedback coming from the twitter users. However, the result of the users' attitude to the new iOS seems not valid. Because the sample space we chose here is strongly biased because only twitter users are analyzed in our study.