Case Study 1

Analysis of iPhone 8 and iPhone X using Twitter

Team 9

Tingting Ma
Jiani Gao
Jinyan Lyu
Tianhao Guo
Mo Cheng

1. Motivation

On September 12th, Apple Special Event 2017 was held in the newly established Steve Jobs Theater in Apple Park California and attracted global attention. Millions of people watched the live broadcast of the conference online. In the conference, the expected new products iPhone 8, iPhone X, Apple Watch Series 3 were announced. As one of the most well-designed mobile devices, iPhone has a large number of loyal customers. According to BMO Capital Market's latest reports, the number of iPhone users worldwide has reached 700 million by March 2017. As a result, after the conference, iPhone 8, iPhone X, together with many related topics about the new products soon became trendy on the social network. Twitter, which presents the opinions from users around the world in a real-time data stream, is one of them. So in this case, we took advantage of the large data volume in Twitter draw by the conference, collected data during and after the event, extracted public preference of the two new iPhones, then predicted the market performance based on the result. All raw data was collected via Twitter streaming API, filtered and analyzed using Python and related libraries.

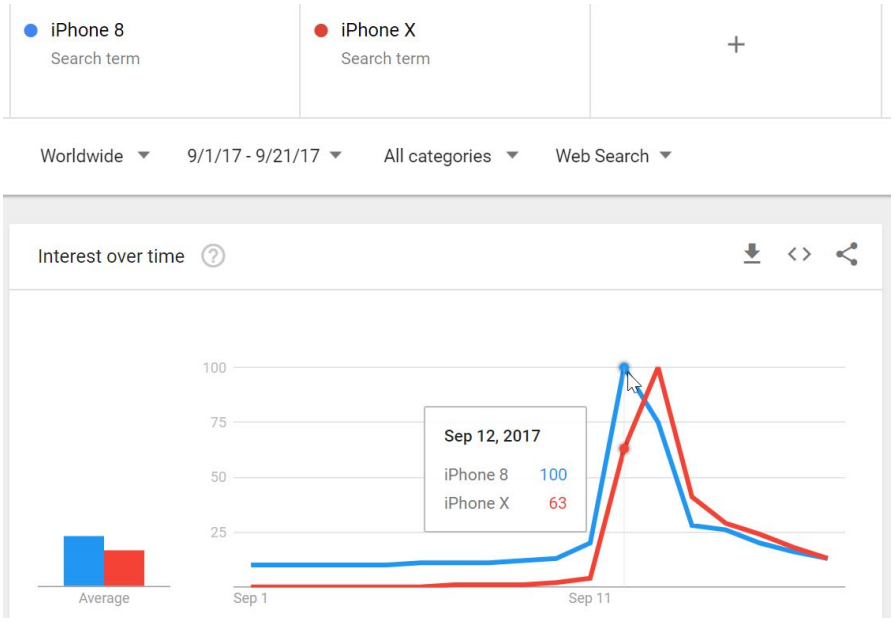

Figure 1: According to Google Trend, the search peak of "iPhone 8" occurred in September 12th, the same day of the conference, while "iPhone X" in the day after.

2. How We Get Data

We chose to collect data on the day of the Apple event on September 12th. Because according to previous experience, the peak of discussion in social networks of Apple new products will occur later in the same day as announced. We collected 10,000 tweets that mention iPhone 8 or iPhone X that afternoon using twitter streaming API.

```
twitter_stream = twitter.TwitterStream(auth=auth)
twitter_stream.statuses.filter(track='iphone 8, iphone x')
```

These two lines of code use Twitter streaming API to get tweets which contain keyword 'iphone 8' or 'iphone x'.

```
file.write(json.dumps(tweet, indent=1))
```

Then we will save these tweets to a file.

3. Analysis

(1) Word Count

First, we calculated the frequencies of the words used in the “text” character of each tweets. During this process, we set some restrictions to make the result more accurate. We removed the special symbols, emoji, punctuation and some meaningless words such as “http”, “n’t”. After that, we sort the frequency of each word and plot a table of the top 30 words with their counts using PrettyTable.

word	frequency
iPhone	6821
X	3405
Y'ALL	955
8	710
unlock	546
Apple	492
x	489
FUCK	478
GO	478
antoniodelotero	477
OK	477
TWEETING	477
PROFILE	477
SAYS	477
TWEETED	477
YET	477
face	474
iphone	439
new	411
people	397
gt	360
skip	335
quick	334
snapchat	334
stories	333
Bluntsandfood	326
7	305
facial	302
5	299
*uses	296

Figure 2: Screenshot of the top 30 words in tweets we collected

(2) Most Popular Tweets

Next, we retrieved the tweets with the largest number of retweet counts. We collected the retweet count of every tweet in a dictionary using “retweet_count” character and sort it. Then we plot a table of the top 10 tweets that are the most popular among our collection.

Retweet Number	Text
137690	RT @brendonSkolat: Beyoncé: *uses iPhone X facial recognition* iPhone X: https://t.co/8ao9moveCd
89774	RT @tobjizzled: iPhone iPhone 3G iPhone 3GS iPhone 4 iPhone 4S iPhone 5 iPhone 5S iPhone 6 iPhone 6S iPhone 7 iPhone 8 iPhone X iPhone XOXO…
79488	RT @memeprovider: iPhone iPhone 3G iPhone 3GS iPhone 4 iPhone 4S iPhone 5 iPhone 5S iPhone 6 iPhone 6S iPhone 7 iPhone 8 iPhone X iPhone XO…
77053	RT @Tolerance: iPhone iPhone 3G iPhone 3GS iPhone 4 iPhone 4S iPhone 5 iPhone 5S iPhone 6 iPhone 6S iPhone 7 iPhone 8 iPhone X iPhone XO TO…
75444	RT @lmScampi: iPhone iPhone 3G iPhone 3GS iPhone 4 iPhone 4S iPhone 5 iPhone 5S iPhone 6 iPhone 6S iPhone 7 iPhone 8 iPhone X iPhone XO TOU…
74485	RT @_YerikaC: me after i sliced my mans face off so i can unlock his iphone x https://t.co/84KUmLBjWB
73201	RT @antoniodelotero: OK BUT HOW THE FUCK Y'ALL TWEETING BUT WHEN I GO TO YOUR PROFILE IT SAYS Y'ALL HAVEN'T TWEETED YET??? https://t.co/Qg2…
70272	RT @lordflaconegro: *sits down* *hears a cracking noise* Me: that better be my tailbone https://t.co/8bjzyXgYAt
62888	RT @TopAchat: 🎁 #Concours 🎁 2 #iPhoneX à gagner ! 😱 RT + Follow @TopAchat Mentionne un ami en réponse (un iPhone X chacun) https://t.co/F…
60045	RT @juanbuis: buying iPhone X in europe? why not add a free weekend trip to NYC https://t.co/K3IbdF4WNL

Figure 3: Screenshot of the top 10 most popular tweets

(3) Most Popular Tweet Entities: Top 10 Hashtags, Top 10 Users Mentioned

Similarly, we collected the contents of hashtags and user mentions respectively using the "hashtags" character and " user_mentions" character of twitter and then plot top 10 hashtags and top 10 user mentions in our collection of tweets.

hashtag	frequency
iPhone8	339
iPhoneX	185
blog	178
amazingarabella	178
AppleEvent	81
Apple	65
StarMoviesSecretScreening	62
TheBigSelfie	62
iPhone9	62
Concours	35

Figure 4: Screenshot of the top 10 hashtags

user mentions	frequency
antoniodelotero	477
Bluntsandfood	326
eimauro	273
emilyferguson_	208
Apple	196
juanbuis	192
holden_a_fork	191
PlNKllD1412	190
Arabelladaho	178
YouTube	159

Figure 5: Screenshot of the top 10 mentioned users

(4) Get the List of All Friends and All Followers of the Twitter User

We picked a user in our data called "TransitoOK" who has thousands of friends and followers. And we counted the number of friends and followers of him and plot 20 followers and 20 friends in two tables. After that, we also computed the mutual friends within the two groups.

```
Fetched 1758 total friends ids for TransitoOK
Fetched 2165 total followers ids for TransitoOK
```

Figure 6: Screenshot of the number of friends and followers of user "TransitoOK"

friend_id	friend_screen_name
32953007	gendarmeria
4224187773	gustavo_osman
220410227	Estacion905
256260076	acostamaxi
105932149	virginiaarce
90185866	marianoemartin
130866630	josedel_rio
780858179018035200	lanacionmas
34016889	Maikeme
180564455	Taxivista
59655672	bertellof
59655753	cgmira
190607660	EleonoraCole
75178495	pablitolescano
96187825	AdalRamones
149288563	segregustavo
116828061	pibesdesistemas
50126679	telefenoticias
4883842013	JuevesIntratabl
333573183	Julyweich

Figure 7: Screenshot of the 20 friends of user "TransitoOK"

follower_id	follower_screen_name
137814421	JARandazzo
310089312	kevinlytlemusic
858357868084461570	BsAs_recuerdo
35710956	leonxxxiii
896452093061541888	tenib_col
38149258	EstherArboix
265992839	FerCieguno
815003085114404864	tonmasci
105932149	virginiaarce
2788411175	trotzkyvengaran
378935962	CamDellafontana
256260076	acostamaxi
160236027	nicoalberio
3374269216	DigitalTrendsEs
180564455	Taxivista
149595417	gaburierus
59655753	cgmira
785837696	Danisam74Daniel
580366572	pablo18carp
864890840	GastonVisceglio

Figure 8: Screenshot of the 20 followers of user "TransitoOK"

mutual_friend_id	mutual_friend_screen_name
220410227	Estacion905
256260076	acostamaxi
105932149	virginiaarce
180564455	Taxivista
59655753	cgmira
3330077703	am1300radio
221170428	soygabito
170159116	Julianvaccarini
514876685	marcelo_aiello
855017036208013314	CarlosSubosky
50765618	pablogaston
2306952265	nahdrian
208754920	pablocigliutti
1731098514	ciharispe
151197934	HernanRolon
1472641098	morenojuan95
896575289244954624	cristian2016141
137177411	FlorLippo
3400652687	CeliStella
1568121829	ddeurieta
886924919967289346	carlosmfrugoni
4017174549	AySA_Oficial
858013333684977669	NPolicialesBA
175108094	cesarmarchione
119239623	dantrila

Figure 9: Screenshot of the mutual friends within the two groups

4. Business Questions

(1) Top Five Languages in Tweets

One of the information useful for app developers regarding new iPhone product is that "what languages should be first internationalized when an iPhone app is developed?" To answer this question, we may look into the data that what language do twitter users use to tweet right after iPhone new products were released. The more tweets in one language may suggest that the larger number of iPhone users in that language.

We did frequency analysis to the language used by these 10,000 sample tweets and used Python Matplotlib library to plot bar charts. We picked top five languages to present in the chart.

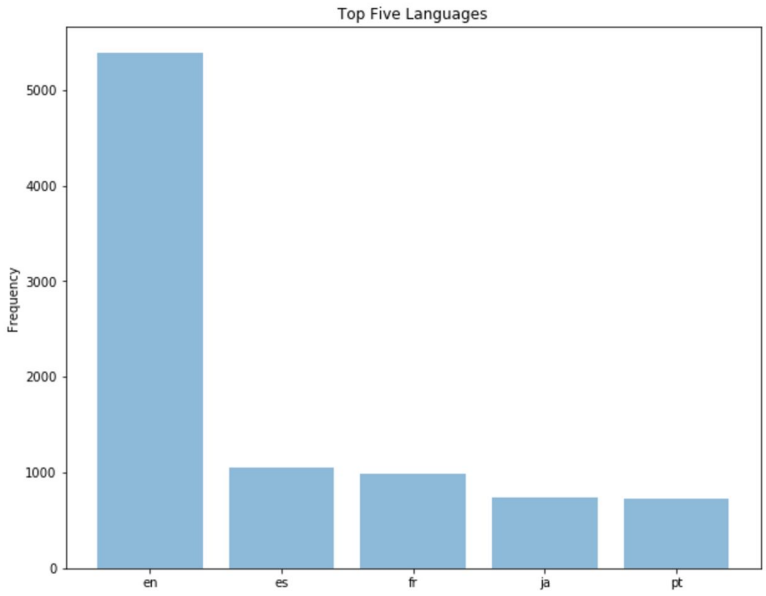

Figure 10: The top five languages used to tweet were English, Spanish, French, Japanese, and Portuguese.

As shown in the chart above, the most users used English to tweet about coming iPhone 8 and iPhone X, more than half of tweets (over 5,000 twitters out of 10,000) are in English. The second most used language is Spanish, followed by French, Japanese, and Portuguese. We also may infer from this plot that popularity and demand of iPhone 8/X in those English, Spanish, Japanese and Portuguese-speaking countries are high. The result may not accurate due to other influence factors like the local time when we sample data, availability of twitter in each area and population difference in different language. But this information can in some degree give us a sense of popularity in different language speaking area.

(2) Top 5 Active Time Zones

Every year, the Apple new product releasing conference is a hot topic for apple fans and iPhone users. People talk about new releasing topics a lot after the releasing conference online. In marketing aspect, if we want to let more audience be able to know the new product at the first time, we may consider what time of the releasing conference is the best. In sales aspect, we want to know which countries have the most enthusiasm and demand for new product, so that apple can fill the stock in advance.

To answer these questions, we did frequency analysis on time zone to see numbers of users in different time zones and used Python Matplotlib library to plot bar charts. We picked top five time zones to present in the chart.

As we can see in the chart, the top five numbers of users are in these time zones: Pacific Time (US and Canada), Eastern Time (US and Canada), Central Time (US and Canada), London and Amsterdam. This chart infers that most twitter users that follow iPhone new products are from US, Canada and Europe. To make more audience able to watch the live new product releasing conference, Apple may choose a daytime that fits US and Europe time (London and Amsterdam). And we also know from this plot that US, Canada, England and Netherland, have the most enthusiasm for the new product, Apple should prepare to have more salesman and product in these areas. The results may not be accurate due to the influence factors like time we sample data, the availability of Twitter in each area, and user habit.

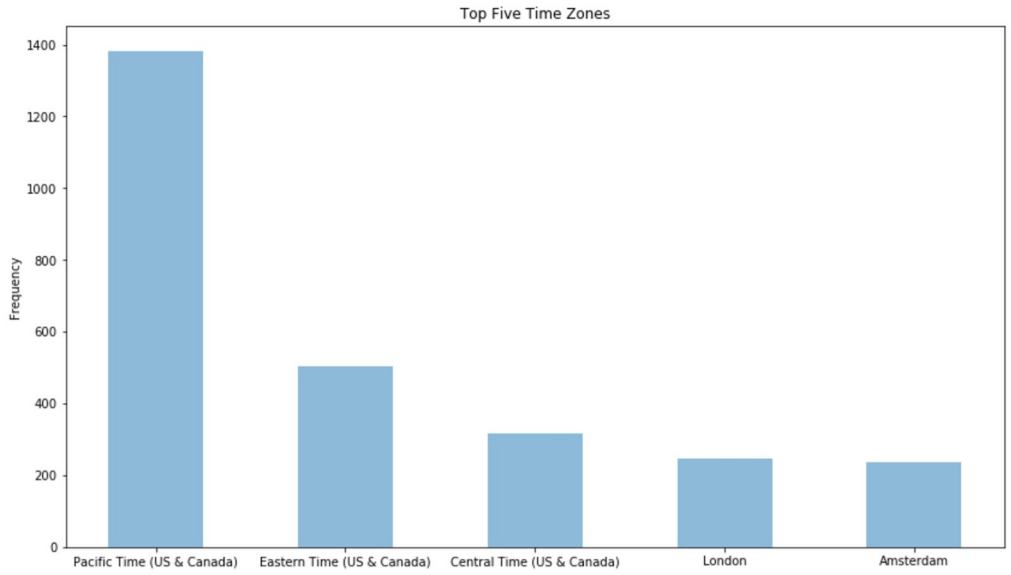

Figure 11: The top five active time zones are: Pacific Time , Eastern Time, Central Time, London and Amsterdam.

(3) Who's Paying Attention

We also want to know more about the people who were paying attention to this event, especially what device are they using. Apple has a lot of fans around the world and the customer loyalty has been high in the past decades. We can get the data of which twitter client are people using in the tweets we get. We classify the sources to Apple devices, Android and other. Apple devices including iPhone, iPad and Mac. Many people are using third-party twitter clients like Tweetbot and Twitterrific. Since these apps are only available on iOS and Mac, we need to count them to Apple devices. In our collected data, tweets sent from Apple devices are more than twice of those sent from Android. Most current Apple customers will continue buying new iPhone to stay in the Apple ecosystem. People who are using Android phones but send tweets about the event may buy an iPhone in the future.

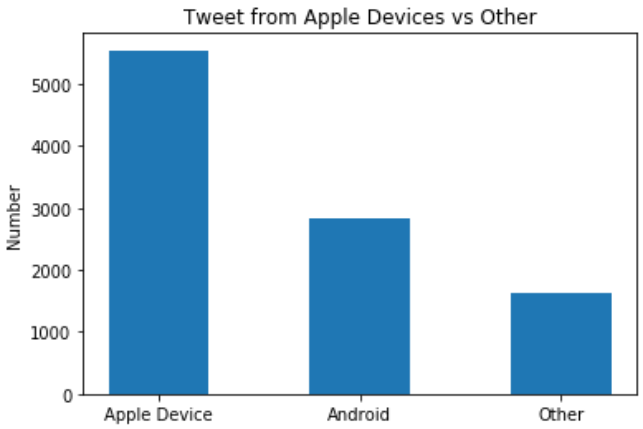

Figure 12: For all collected tweets, those sent from Apple devices are more than twice of those sent from Android.

(4) iPhone 8 or iPhone X

Unlike previous releases, to celebrate the 10th birthday of the iPhone product line, Apple skipped the version number 9 and introduced the completely newly designed flagship model iPhone X together with iPhone 8. According to figure 1 of Google Trends, it was not until 24 hours before the conference that the iPhone X plan start to reveal to public (almost zero search record before September 11th). However, during and after the release, the popularity of iPhone X kept climbing to soon passed iPhone 8 series, and reached the peak in the next 24 hours. The result of our analysis of tweets (figure 13) also shows people's interests in iPhone 8 were overwhelmed by iPhone X. These facts together suggests that although money is the major concern, the market shows strong interests in iPhone X.

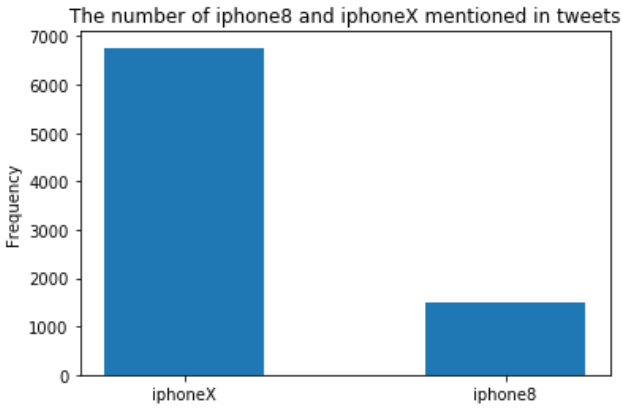

Figure 13: For all collected tweets, "iPhone X" was mentioned more than 4 times more frequently than "iPhone 8".

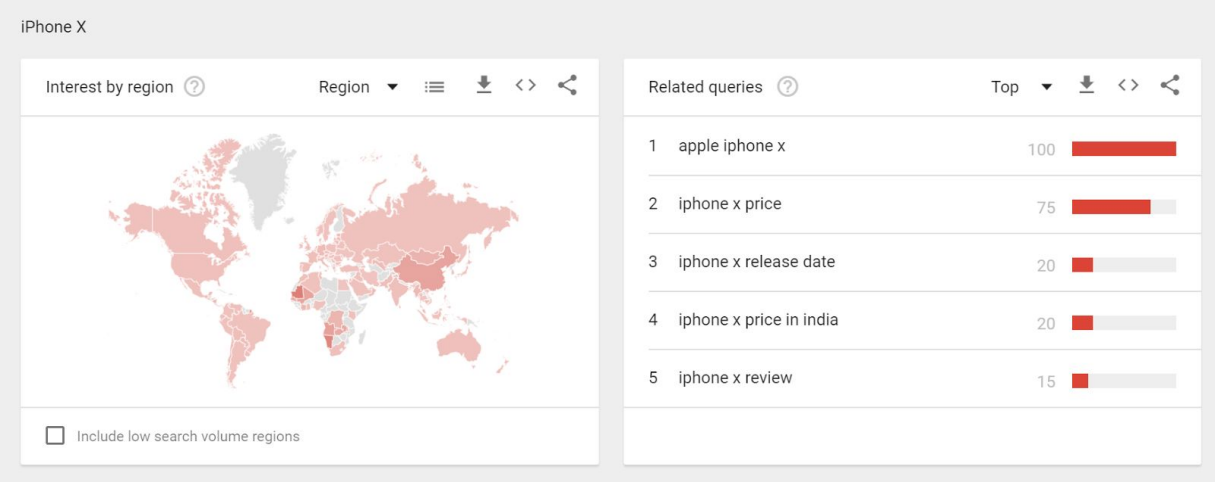

Figure 14: As Google Trend shows, the most related queries with iPhone X were focused on price.

(5) Internet's Attitude towards the Release and Face ID

The event got a lot of attention around the world but people's attitudes towards new iPhone varied. We used TextBlob to analyze the polarity of the tweets we collected. The polarity score is a float within the range [-1.0, 1.0]. Value 0.0 means neutral. Negative value means negative attitude and postive value means postive. Regarding the new iPhone, we can see most people like it. The new feature FaceID caused lots of controversy. In the chart, we noticed that the percentage of negative tweets increased.

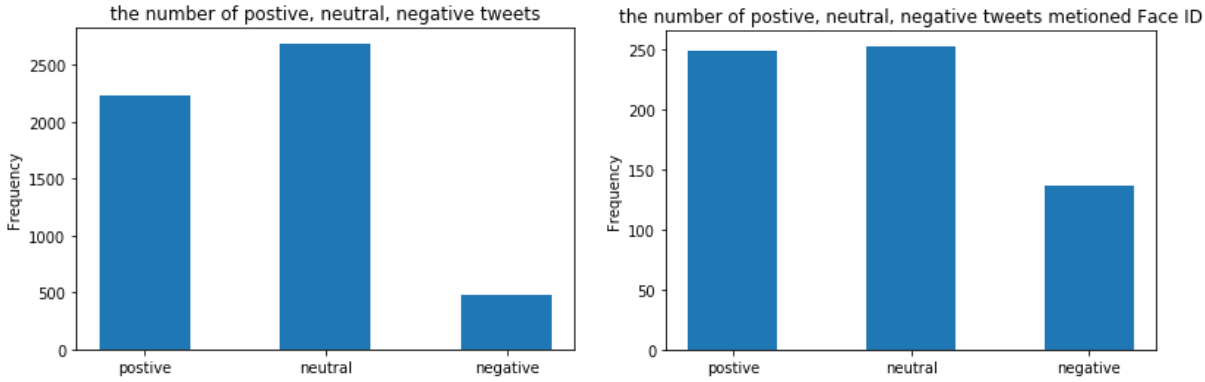

Figure 15: Attitudes towards new iPhone and FaceID

(6) Subjectivity of tweets

As we seen in the previous step, we can get the polarity of tweets. We can also get the subjectivity which is a float within the range [0.0, 1.0]. Value 0.0 means very objective and value 1.0 means very subjective. As we see in the scatter plot, most points drop in the area represents positive and subjective.

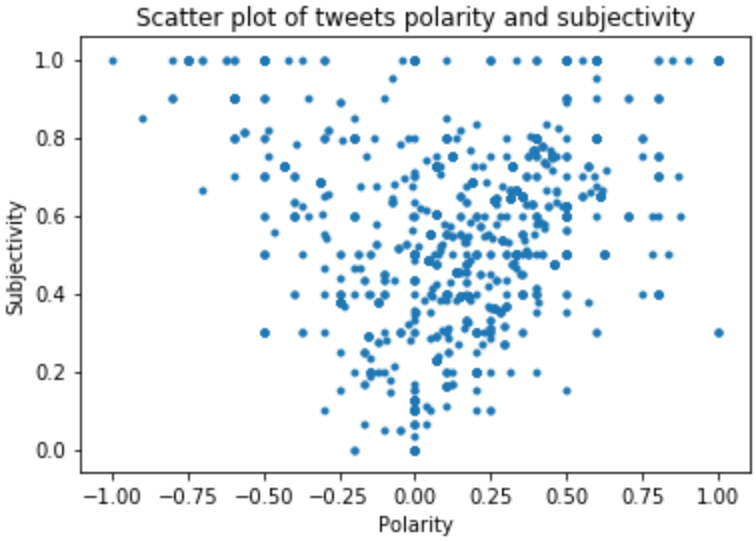

Figure 16: tweets polarity and subjectivity