

# **DS501**

# **Case Study Report**

# **No.1**

**Group 10**

Jiexuan Sun, Chao Xu, Zhaoning Su, Congyang Wang

## **1. What data we collected? - Our Data**

We collected the twitter data by using twitter streaming API. Within a period of time, we collect 1000 tweets with word “iPhone7”. We collect Tim cook’s all twitter friends and followers list.

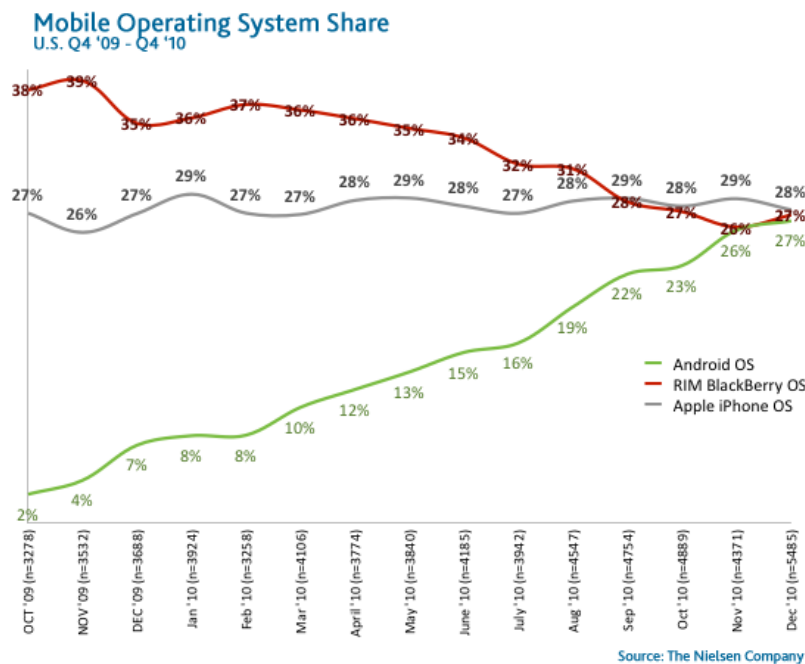
## **2. Why this topic is interesting or important to you? – Our Motivations**

With the first generation iPhone released by Steve Jobs, the Apple Inc. is always in the public view. And in the past years’ September, the new generation iPhone launch event becomes a carnival for the Apple fans. At the same time, it seems that apple’s products start to be fashion icons. But, after Steve Jobs passing away, the following iPhone becomes more and more controversial. For example, some people doubt that the iPhone4 is almost the same product of iPhone4. And starting from that time, people start to saying that “Apple start to give up the innovation and perfect products.” Also, the grow of the Android based mobile Phones gives Apple Inc. great competitive pressures. Samsung’s Galaxy mobile series sale volume jumped to the first tier.

**Table 1: Top Six Global Smartphone Brands by Worldwide Market Share, 1Q16~2Q16**

Company	2Q16		1Q16	
	Ranking	Market Share	Ranking	Market Share
Samsung	1	24.5%	1	28.0%
Apple	2	15.1%	2	14.5%
Huawei	3	9.2%	3	9.3%
OPPO	4	5.6%	4	5.3%
LG	5	5.4%	5	5.2%
BBK/vivo	6	4.7%	6	4.7%
Others	--	35.5%	--	32.8%
<b>Total Production Volume (Unit: k)</b>	<b>314,550</b>		<b>288,960</b>	

Source: TrendForce, Jul., 2016

[http://www.gsmarena.com/trendforce\\_huawei\\_leads\\_chinese\\_smartphone\\_market\\_samsung\\_leads\\_globally-news-19543.php](http://www.gsmarena.com/trendforce_huawei_leads_chinese_smartphone_market_samsung_leads_globally-news-19543.php)

So that's why we have interest on the iPhone. we want to know how people think about these new products, which are released in this September? Do they like it? Which new product is more popular, the iPhone7, iPhone7 Plus? Compared with

Steven Jobs, as the CEO of Apple Inc., is “Tim Cook”, the CEO of Apple Inc., being considered as a great leader?

### **3.How did you analyses the data? – Our Process**

#### **3.1**

First, within a time period, by using the twitter streaming API, we collect 1000 tweets which contains the word – ‘iPhone7’.

#### **3.2**

Second, we compute the frequencies of the words being used in these tweets.

And Plot a table of the top 30 words with their counts. (See Table 1) we found that the top 30 words in those tweets are mostly preposition, article, etc.

It seems that we have no idea what people think about the new iPhone.

Word	Count
#iPhone7	570
the	507
I	398
giveaway!	336
in	334
have	319
RT	301
participated	283
to	172
for	113
is	96
and	95
a	94
7	93
iphone7	92
&amp;P;~~~~~	90
iphone	80
#iphone7	78
of	78
iPhone7	73
just	64
-	62
get	61
new	58
entered	53
@UpsideHQ	53
@nbt88yt	53
@Eztech231	53
your	49
#iPhone7Plus	49

Table 1

Count	Screen Name	Text
2033	YOONAYAcom	RT @YOONAYAcom: [YOONAYA gift] 160920 128G iPhone7 Plus for YoonA <a href="https://t.co/sRSft2jWAw">https://t.co/sRSft2jWAw</a>
1434	ChiefKeef	RT @ChiefKeef: Flexing on Hoe ass niggas With this new iphone7 shit 🤔🤔🤔 <a href="https://t.co/6tefmN8Qoq">https://t.co/6tefmN8Qoq</a>
1433	ChiefKeef	RT @ChiefKeef: Flexing on Hoe ass niggas With this new iphone7 shit 🤔🤔🤔 <a href="https://t.co/6tefmN8Qoq">https://t.co/6tefmN8Qoq</a>
1412	luki_mankin	RT @luki_mankin: Get your FREE #iPhone7 #iPhone7Plus ONLY TODAY. LimitedEdition and limited quantity. <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a> <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a>
1411	luki_mankin	RT @luki_mankin: Get your FREE #iPhone7 #iPhone7Plus ONLY TODAY. LimitedEdition and limited quantity. <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a> <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a>
1410	luki_mankin	RT @luki_mankin: Get your FREE #iPhone7 #iPhone7Plus ONLY TODAY. LimitedEdition and limited quantity. <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a> <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a>
1409	luki_mankin	RT @luki_mankin: Get your FREE #iPhone7 #iPhone7Plus ONLY TODAY. LimitedEdition and limited quantity. <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a> <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a>
1408	luki_mankin	RT @luki_mankin: Get your FREE #iPhone7 #iPhone7Plus ONLY TODAY. LimitedEdition and limited quantity. <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a> <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a>
1407	luki_mankin	RT @luki_mankin: Get your FREE #iPhone7 #iPhone7Plus ONLY TODAY. LimitedEdition and limited quantity. <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a> <a href="https://t.co/w415l1lME0">https://t.co/w415l1lME0</a>
1110	KeeganAllen	RT @KeeganAllen: iPhone7 had room for Jack. Smh <a href="https://t.co/5Vmzh6WrCl">https://t.co/5Vmzh6WrCl</a>

Table 2

### 3.3

Third, we plot a table of the top 10 tweets that with the largest number of retweet counts among our collection.(See Table 2) Interestingly, we found that someone is keep retweet his/ her own tweet. Obviously, this tweet account are ‘Zombie’ fans, which are those artificial friends. Those tweets become to ‘noise’ and they will effect out analysis.

### 3.4

Fourth, we get the list of all friends and all followers of Tim cook. And we Plot 20 out of his followers (See Table 3) and 20 out of his friends (See Table 4). And, we try to Compute the mutual friends within the two groups. Surprisingly, there is no mutual friend for Tim cook. (See Table 5)How could that happen! After carefully checking, we find that the python package (tweet) we use limit us to retrieve up to 5000 friends/followers' information.

user_choose_follows		user_choose_friends	
user_id	screen_name	user_id	screen_name
779043546728759297	Vj_saxena	90420314	ActuallyNPH
1084351826	beckyskr	29450962	repjohnlewis
503206587	orchidfrost	12	jack
779043100996100096	cupello_lenore	14130366	sundarpichai
779043048160366593	kenyattakechi3	3257608936	AppleNews
21202107	Jake_Mathers	3309375033	AppleSupport
35488827	AishaJ7	1205322793	mspear96
779042100348203008	BGrant2016	21872269	laurenepowell
779041418106974208	MichellemSong	16303106	StephenAtHome
779041528178245634	BrandManaging16	21288052	zanelowe
779040393551720449	MinulRobim	1605	sama
735684333021724672	24HourBrew	1582853809	HistoryInPics
910839768	pritammshr	2204034272	Beats1
779039735893397505	nexodoc_global	338084918	Pharrell
779041622424072192	ipn1433	1182675871	lisapjackson
756769427610185728	SobhyAbraham	19608297	HRC
726636744645435392	SivakumarSubra7	69231187	UNHumanRights
4909826322	BiserRangelov	20571756	satyanadella
766144766132322304	SamRamo9424	21133007	OneRepublic
772033629723635712	bnuz11	914061	levie

Table 3

Table 4

```

mutual_friends
+-----+-----+
| user_id | screen_name |
+-----+-----+
+-----+-----+

```

Table 5

### 3.5

Fifth, we try to find that which country of citizen pay more attention on the new iPhone. We plot a table to show the tweet's distribution based on different countries. (See table 6)

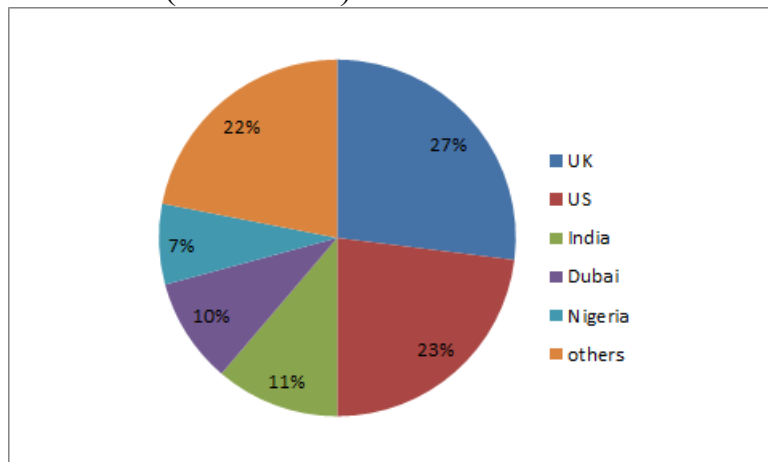


Table 6

## 4. What did you find in the data? – Our Final Result

Even though, because of those 'noise' - including those Common vocabulary(e.g. 'the', 'a', 'an'...) and the 'zombie' fans effect our data purity. We cannot get our final result. But We know that we should remove those noise next time while collecting data. That's what we learn from this project.

## Reference

1. The Nelson Company
2. <http://www.gasmarena.com>