

---

## Data Analytics

# Trump Twitter Archive – Sentiment Analysis

---

# Project Goals

- How many tweets are made from *Android*? And from *iPhone*?
- Is true that the **most negative tweets** are made from *Android*?
- What is the **frequency** of ‘positive’ retweets? And what about favourite?

	source	text	created_at	retweet_count	favorite_count	is_retweet	id_str
0	Twitter for iPhone	Still can't believe that Bloomberg violated a ...	2018-09-01 02:40:23+00:00	5540	21996	False	1035718986871320576
1	Twitter for iPhone	Great day in North Carolina where Republicans ...	2018-09-01 02:36:57+00:00	3885	18946	False	1035718119459893248
2	Twitter for iPhone	The ABC/Washington Post Poll was by far the le...	2018-09-01 02:35:28+00:00	5164	20129	False	1035717747936768000
3	Twitter for iPhone	@Rasmussen_Poll just came out at 48% approval...	2018-09-01 02:25:38+00:00	5717	21855	False	1035715271418413056
4	Twitter for iPhone	<a href="https://t.co/c79zLeREOA">https://t.co/c79zLeREOA</a>	2018-09-01 00:01:21+00:00	8506	31907	False	1035678961349668864



# Project Goals

- How many tweets are made from *Android*? And from *iPhone*?
- Is true that the **most negative tweets** are made from *Android*?
- What is the **frequency** of 'positive' retweets? And what about favourite?

How can we estimate the **sentiment** of an entire corpus of tweets?

	source	text	created_at	retweet_count	favorite_count	is_retweet	id_str
0	Twitter for iPhone	Still can't believe that Bloomberg violated a ...	2018-09-01 02:40:23+00:00	5540	21996	False	1035718986871320576
1	Twitter for iPhone	Great day in North Carolina where Republicans ...	2018-09-01 02:36:57+00:00	3885	18946	False	1035718119459893248
2	Twitter for iPhone	The ABC/Washington Post Poll was by far the le...	2018-09-01 02:35:28+00:00	5164	20129	False	1035717747936768000
3	Twitter for iPhone	@Rasmussen_Poll just came out at 48% approval...	2018-09-01 02:25:38+00:00	5717	21855	False	1035715271418413056
4	Twitter for iPhone	<a href="https://t.co/c79zLeREOA">https://t.co/c79zLeREOA</a>	2018-09-01 00:01:21+00:00	8506	31907	False	1035678961349668864



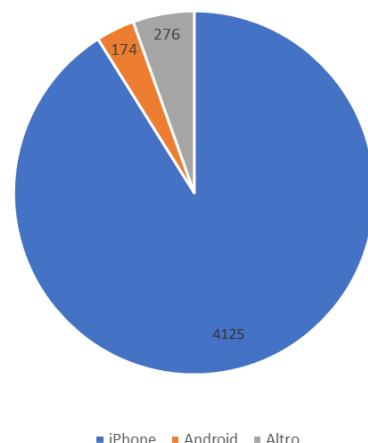
# Dataset

Data from the **Trump Twitter archive** (20/01/2017 - 01/09/2018).

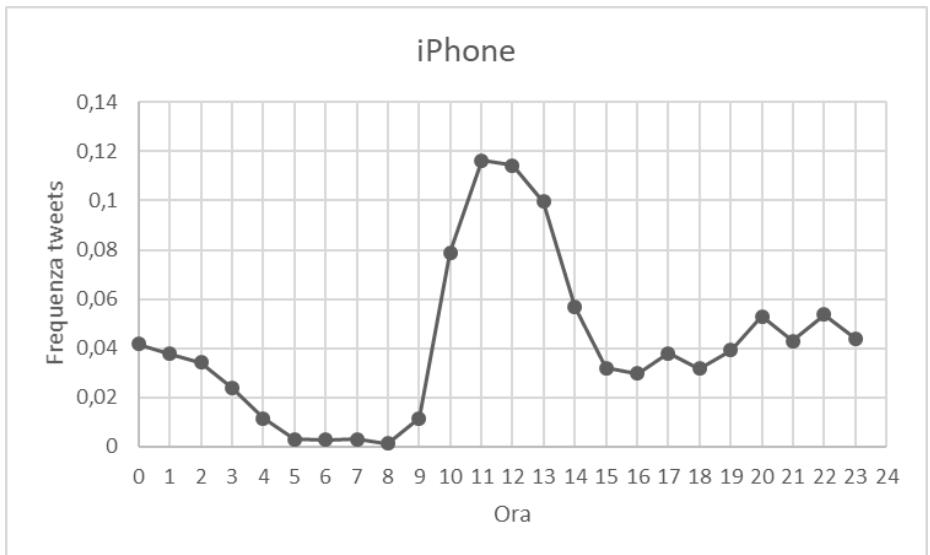
- **4575 tweets**

- 0 Still can't believe that Bloomberg violated a ...  
1 Great day in North Carolina where Republicans ...  
2 The ABC/Washington Post Poll was by far the le...  
3 .@Rasmussen\_Poll just came out at 48% approval...  
4 <https://t.co/c79zLeREOA>

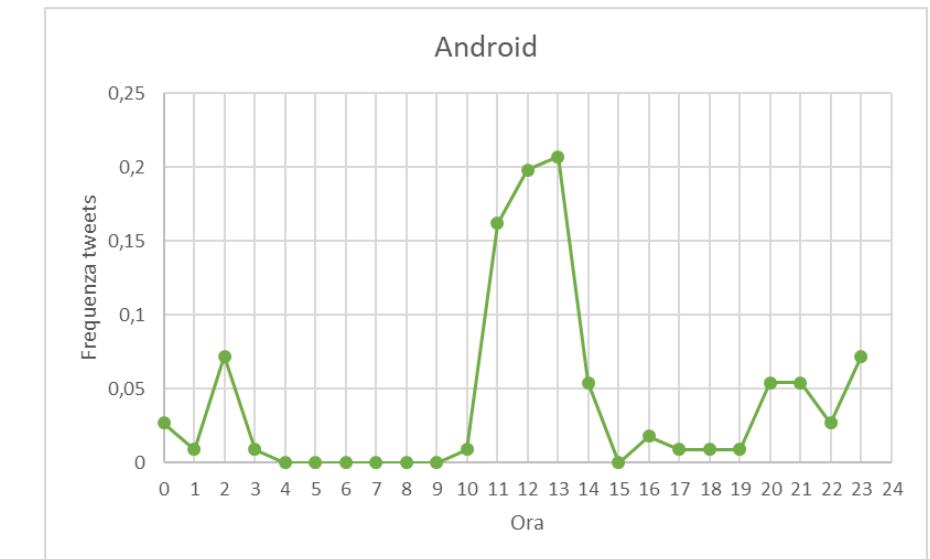
Distribuzione dei tweet sulla base della piattaforma utilizzata



- **4125 tweets** from *iPhone* (no iPad)
- **174 tweets** from *Android*



Distribuzione dei tweet sulla base dell'ora Apple



Distribuzione dei tweet sulla base dell'ora Android

# iPhone Tweets



Frequency	Token	Frequency	Token	Frequency	Token	Frequency	Token	Frequency	Token
5639	.	5639	.	887	I	559	great	949	gre
4830	,	4830	,	725	'	376	people	547	new
4550	the	2811	!	654	...	330	Trump	426	peopl
2855	to	887	I	628	The	307	President	366	country
2826	and	817	:	559	great	267	Fake	365	presid
2811	!	725	'	530	RT	258	Democrats	344	thank
2181	of	708	&	463	"	250	News	337	stat
1656	a	654	...	453	"	232	many	335	big
1545	in	628	The	376	people	224	Great	327	job
1539	is	559	great	357	...	208	Thank	327	democr

TweetTokenizer

Stopwords

Punctuation

Few more  
Stopwords

Stemming



# Android Tweets



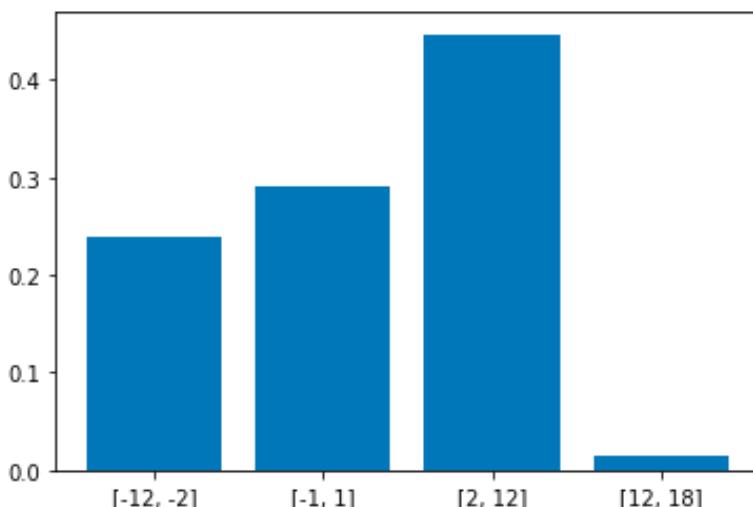
Frequency	Token
32	<b>new</b>
25	<b>fak</b>
22	<b>gre</b>
22	<b>peopl</b>
20	<b>country</b>
17	<b>bad</b>
16	<b>med</b>
15	<b>obam</b>
14	<b>big</b>
14	<b>elect</b>



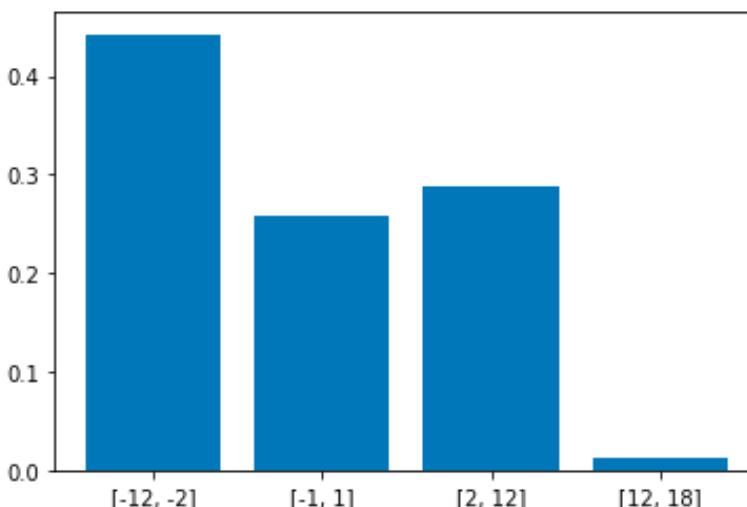
## Lexicon-based (AFINN) Approach



Sentiment	Range	Frequency
Negative	[-12, -2]	0.24
Neutral	[-1, 1]	0.29
<b>Positive</b>	<b>[2, 12]</b>	<b>0.45</b>
Extreme Positive	[12, 18]	0.01



Sentiment	Range	Frequency
Negative	[-12, -2]	0.44
Neutral	[-1, 1]	0.26
Positive	[2, 12]	0.29
Extreme Positive	[12, 18]	0.01



## Lexicon-based (AFINN) Analysis



Sentiment	% Retweet
Positive	0.34
Neutral	<b>0.42</b>
Negative	0.24

Sentiment	% Favourite
Positive	0.37
Neutral	<b>0.41</b>
Negative	0.22



Sentiment	% Retweet
Positive	0.27
Neutral	0.26
Negative	<b>0.47</b>

Sentiment	% Favourite
Positive	0.30
Neutral	0.25
Negative	<b>0.45</b>

# Summary

- The majority of the tweets from *iPhone* are positive (**45%**);
- The majority of the tweets from *Android* are indeed negative (**44%**).
- The majority of the reetweets (**42%**) and favourite (**41%**) from *iPhone* are neutral.
- The majority of the reetweets (**47%**) and favourite (**45%**) from *Android* are negative.

The screenshot shows a news article from The Guardian. At the top, there are buttons for 'Search jobs', 'Sign in', 'Search', and 'International edition'. Below that is the 'The Guardian' logo. The main headline reads: 'Donald Trump 'writes angrier and more negative Twitter posts himself''. A sub-headline states: 'Republican candidate uses words such as 'crazy', 'weak' and 'dumb' on Android device, while staff post calmer tweets on iPhone, data suggests'. The article features a large photo of Donald Trump with his mouth open, giving a thumbs-up. On the left, there's a sidebar with 'Social media' links and a bio for Jasper Jackson. The main text area includes sections for 'most viewed' articles and a 'Read The Guardian without interruption on all your devices' sidebar.

## Text analysis of Trump's tweets confirms he writes only the (angrier) Android half

I don't normally post about politics (I'm not particularly savvy about polling, which is where data science has had the largest impact on politics). But this weekend I saw a hypothesis about Donald Trump's twitter account that simply begged to be investigated with data:

The image displays two tweets from Donald J. Trump's Twitter account. The first tweet, posted on Aug 5, 2016 at 8:59 PM, contains hyperbolic language like 'Good luck #' and '#OpeningCeremony.pic.twitter.com'. It has 27,391 likes. The second tweet, posted on Aug 6, 2016 at 11:11 AM, is more neutral or positive, wishing the Olympic team good luck. It has 4,451 likes. Both tweets are marked with a blue checkmark and a small profile picture of Donald Trump.

When Trump wishes the Olympic team good luck, he's tweeting from his iPhone. When he's insulting a rival, he's usually tweeting from an Android. Is this an artifact showing which tweets are Trump's own and which are by some handler?