ASSIGNMENT 5
Voyant analysis of political discourse
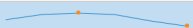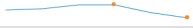
The purpose of this assignment is to analyse word trends over time in Barack Obama's and Donald Trump's tweets using Voyant, an open-source text analysis tool highly appreciated by digital humanists because of its effective interface and incorporated analytics tools.

Both former presidents have left their mark on American politics, so it is no surprise that their online presence on digital platforms such as Twitter has been a source of information up for debate. In this analysis, the focus will be on the election years for both politicians, exploring various key terms, patterns and contradictory practices.

The corpus for each former president contains several text files organised by years: Obama (6 files) 2011-2016 and Trump (9 files) 2011-2019. For both corpora, stopwords from the English library were eliminated along with unnecessary occurrences such as 'http', 'https', '&amp'.

OBAMA

For context, Barack Obama's term of office ran from 2009 to 2017. The first-ever African American president of the United States of America was reelected in 2012 with one of the most successful modern campaigns since the begging of the 2000s.

The Obama corpus contains 6 txt. documents with 503,433 total words and 36,123 unique word forms.

One of the most useful visualisation tools provided by Voyant is the creation of wordclouds, which always offer interesting results in discovering potential keywords and trends. The first wordcloud was plotted over the whole corpus to identify the main key terms over the years: Obama, president, watch, today, new, whlive, people, Americans etc.

	Term	Count	Trend
1	president	6566	
2	obama	7982	
3	watch	2074	
4	today	1689	
5	new	1209	
6	whlive	1185	
7	people	1156	
8	americans	1154	
9	live	1082	
10	vp	1068	
11	american	1053	
12	time	956	
13	america	956	
14	just	952	
15	health	938	
16	day	893	

From the perspective of elections, it can be observed in the above line graph showing trends of occurrences of the words *president* (red) and *Obama* (green). There is an upwards trend for both words during the years 2012-2013 and then a descending trend after the year 2014. This can be associated with the fact that Obama was reelected in 2012, so his popularity hit a peak value.

Taking a look at Obama's online discourse during 2011 and 2012 we can assume that the main concern at the time for the former president was his reelection campaign since *president* and *Obama* are the most frequent words. Internal issues such as *taxation* seem to be topics of discussion.

2011

2012

TRUMP

In the case of the Trump corpus, we can immediately identify certain words in the Cirrus visualization characteristic of his political discourse, also by observing the table view of the most frequently used words in the corpus.

	Term	Count	Trend
1	great	5393	
2	trump	4834	
3	president	2492	
4	thank	2435	
5	just	2271	
6	thanks	2156	
7	people	2139	
8	false	1774	
9	donald	1616	
10	america	1560	
11	country	1558	
12	rt	1555	
13	new	1552	
14	time	1465	
15	like	1430	

An interesting placement in the wordcloud can be observed in the year 2011, when Trump announced he would not run for president in the 2012 election, putting an end to what he described as "unofficially campaigning". In February 2012, Trump endorsed Mitt Romney for president. We see a change of rhetoric when Trump was allegedly running for president alongside Obama and then when he was not actively campaigning. Trump's online discourse seemed to tackle different subjects such as foreign affairs (mentions of *China*, *Lybia*, and *Iraq*) and internal problems such as the employment issue in America (mention of *jobs*).

2011   2012

barackobama   obama

DT_tweets_2011		127
		16
DT_tweets_2012		430
		360
DT_tweets_2013		12
		302
DT_tweets_2014		8
		313
DT_tweets_2015		3
		134
DT_tweets_2016		2
		84
DT_tweets_2017		1
		57
DT_tweets_2018		1
		121
DT_tweets_2019		34
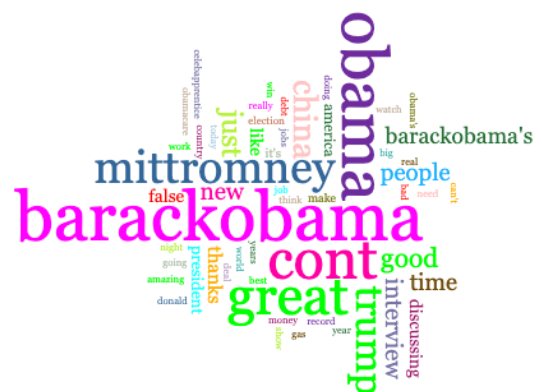
Compared to Barack Obama's discourse, which did not mention Trump that often, the latter was quite invested in what we could presume was finding flaws in Obama's first term at the White House. Results displayed by the bubble line visualization show that indeed the year 2012 has the highest number of occurrences (430) of Barack Obama in the corpus. Moreover, looking at what terms collocate with *barackobama* in the Trump corpus, we sense that Trump's distrusts of the renewed Obama office is further strengthened by how keywords have been used in different phrases in the corpus. In the below WordTrees, by default, the most common term is used as the first word tree root, this way assigning attributes such as *"competent"*, *pathetic*, *outrageous* and *reckless* to Obama.

Term	Collocate	Count (context)
barackobama	cont	35
barackobama	mittromney	28
barackobama	barackobama	21
barackobama	china	17
barackobama	wants	16
barackobama	campaign	15
barackobama	economy	14
barackobama	america	14
barackobama	country	13
barackobama	budget	13
barackobama	release	12
barackobama	keystone	12
barackobama	great	12
barackobama	debt	11
barackobama	interview	10
barackobama	timetogettough	9
barackobama	thanks	9