Information Workshop INF 1005: Social Media Analytics
Assignment 1 - Social media network analysis of structured data

Background

On June 23, 2016, the United Kingdom held a referendum on whether the country should leave the European Union, in which 51.9% of voters voted to leave. In March 2017, the UK government formally announced the country's withdrawal. Since then, the UK started the 2-year process of withdrawal but continuously extended the deadline. When Teresa May resigned on June 7, 2019, Boris Johnson preceded as the prime minister and was re-elected on December 12, 2019, general election, promising that the UK will be withdrawing no later than January 31, 2020.

Introduction

The purpose of this paper is to perform an analysis of a current social, political, or technological issue on two social media platforms: Twitter and YouTube. As the UK was approaching the January 31, 2020 deadline set by the European Union for the UK to withdraw from the EU, along with Boris Johnson's recent re-election, Brexit became a topic of trend and interest for social media networks analysis.

We formulated four (4) research questions regards to the trend analysis and media analytics on the keywords "Brexit" or "#Brexit", along with the null and alternative hypotheses of each question for measuring the level of significance of our findings as shown in Table 1.

1. Is there a statistical variation between video duration and view counts on YouTube?
2. Do the dates that the video is viewed influence the view counts?
3. Was there an increase in Twitter users' creation directly after the Brexit referendum in June 2016?
4. Is there a significant association between the location of Tweets and the voter turnout for each United Kingdom constituent country?

On February 03, 2020, we formulated additional two (2) research questions regards to the text and sentiment analysis on the keywords "Brexit" or "#Brexit", along with the null and alternative hypotheses of each question for measuring the level of significance of our findings in Table 1 as well.

5. What are the changes in the top 15 unique words found in tweets before and after Brexit?
6. What are the changes in sentiment on tweets posted before and after Brexit?

Discussion

Network Analysis

The objective of the research is to identify and analyze network associations between information publishers and audiences on Brexit by network analysis.

Text Analysis to evaluate the posters' stance on Brexit, whether this is a polarized group of supporters and opposers on Brexit or, alternatively, a more diverse community. We would also like to look into the quality and quantity of information exchange of Twitter versus Youtube in which content analysis and effectiveness of political propaganda on social media platforms.

Netytic

Using the keyword "'brexit' OR '#brexit'" on Netlytic, we retrieved 5,846 Twitter records between January 13 to January 20, 2020. Netlytic provided that the network density of the dataset is 0.000382, reciprocity is 0.018320, centralization is 0.023860, and the modularity is 0.919600. In other words, it appears that there are limited ties in the network compared to the number of ties that could exist. Moreover, the number of reciprocated ties is low, which implies that the network is less cohesive and people are not responsive to one another, rather they talk unidirectionally. Additionally, the network is not centralized, meaning there are very few people who are getting most of the mentions. Lastly, there is a large number of communities and a high connection within nodes of a community but the connection is sparse between nodes of different communities.

Gephi

Twitter

On Gephi, we calculated In-Degree and Out-Degree values to find potential people being followed/following respectively. The maximum In-Degree was observed for borisjohnson (151), followed by nicholasturgeon (40), bbc (32), june_mummery (28) and lozzafox (27). The maximum Out-Degree was observed for cathyby, jag11814459, shlomoindiana, all being 50. Moreover, Table 3 outlines the top thirty (30) Twitter handles based on indegree, or in other words, the accounts that received many mentions. We have further classified these Twitter handles and Table 3 outlines the respective classifications, which include: politicians, political parties, media outlets, and activist groups.

YouTube

On the YouTube Data Tools, our starting points on the Video Network Module were search query was "brexit" with 6 iterations with an additional craw depth of 0. The tool produced a list of 300 videos networks based on the topic "brexit". The video IDs of the 300 videos were then manually inputted into the Video List Module to generate each video's info and statistics. On Gephi, we analyzed In-Degree and Out-Degree for Youtube video network. View Count was found to be maximum for Entertainment, Education, News & Politics and Film & animation sectors.

R Stats

On R, we used ANOVA for our first research question, using variables of video duration (independent) and view counts (dependent) to see the variation between them. As a result, we could discover that there was no significant variation between the video duration and view counts as the p-value (0.536) was much higher than the alpha (0.05), implying that the video duration does not influence the number of view counts.

For the second research question, we first used the whole data that we retried from the video-list to see a significant difference between the number of view counts (y-axis) based on each published date (x-axis) to create a graph. However, since the graph with all of our data did not really show any significance with the view count, we cleaned the data to see the view counts only in January (the most recent data) out of the other months. As a result, it showed a significantly high view count on the 17th.

For the third research question, we utilized the "user_account_created" data and graphed it yearly to see the historical trends (see Figure 4). Subsequently, we graphed the same data for 2016, the year of the Brexit referendum, to visualize the trends for the respective year. Our hypothesis prior to conducting this analysis was that Brexit would cause an increase in user account creation, as people would refer to

Twitter to express their thoughts and concerns. Accordingly, as evidenced in Figure 4, the months of June and July, the months leading up to and after the Brexit referendum, showed an increase in user accounts created compared to other months. Utilizing regression testing confirmed a statistical significance between Brexit and user account creation. (see Figure 5).

When answering the fourth research question, the variables that we focused on were "user_location" and the results were filtered using the sum function and grepl function. There were 1825 tweets that users did not input their location. Out of 4021 tweets, 84 tweets were "located" from 'wales', 818 tweets were from 'england', 69 tweets 'ireland', 290 from 'scotland', 384 from 'united kingdom', 418 from 'uk' (see Table 2). The proportion of tweets for each United Kingdom constituent country was visualized as a pie chart (see Figure 20). There was a significant number of Tweets from England, with tweets from Scotland is the second highest, followed by Wales, and Ireland has the least amount of tweets. Compared to the turnout reported by BBC News, England has the highest turnout, followed by Wales, and finally Scotland and Ireland. Our data and BBC data both reflect that individuals from England are highly participating in the 'Brexit' discussion. Unfortunately, our data did not capture the same numbers as the election results. Initially, we wanted to perform a test of variance between the turnout from tweet to the actual published referendum turnout, but identifying the exact location of the users were beyond our scope in network analysis. Thus, we were unable to answer our null hypothesis, that "the tweets do reflect the voter turnout for each United Kingdom constituent countries", for our fourth research question.

Text and Sentiment Analysis

The objective of the text and sentiment analysis was to observe the differences in the top 15 unique words and sentiment level on tweets before and after Brexit day on Jan 31, 2020.

Text Analysis

For the text analysis, our team utilized TwitterR to gather an additional 5000 tweets from Twitter with the hashtag brexit. but not including retweets due to retweets would limit the R package from gathering unique twitter posts. Subsequently for both Twitter data before and after Brexit, we cleaned the text to remove URLs, punctuation and stop words, change all tweets to lower case, and clear the word "brexit". Once the data was cleaned, we created a bar graph to outline the top fifteen (15) unique words found in tweets after Brexit. Our analysis shows that the most common words found in tweets after Brexit are "EU" and "UK". (See Table 4, Figure 18 & Figure 19) . We can observe that "boris" and "johnson" are some of the most tweeted words before Brexit. Words like "vote", "citizens", "people" and "leave" reflect public action and behaviour towards this issue. After Brexit has words like "politics" and "news" which can be due to news channels talking about the announcement and next steps.

The results of text mining reflected the frequency of words used along the topic of Brexit during the text analysis. Understanding and analyzing why all 30 words appeared frequently in the 10000 tweets from two different sequences of event is, however, beyond our scope of this paper. Unless those words have a weighting factor, we can only explain that the words in Table 4 reflect how frequently the words showed up in tweets that are related to "Brexit" but not the importance of those 30 words on the topic "Brexit" in general.

Sentiment Analysis

Once the text analysis was complete, we referred to sentiment analysis to identify and categorize opinions expressed in our two datasets. As evidenced by Table 5, our sentiment analysis confirmed that "positive" sentiment had the highest percentage of expression than "negative".

Given the fact that the United Kingdom ("U.K.") officially left the European Union ("EU") on January 31, 2020, our team was interested in comparing the data we gathered before and after the official departure. Accordingly, we utilized the data gathered prior to the U.K.'s official departure to conduct text and sentiment analysis and compared the results against the data gathered subsequent to the U.K.'s official departure. The top 3 sentiments on tweets before and after Brexit were "trust", "positive", and "negative". While the list of unique words does not provide much context, the sentiment analysis conversely identified a decrease of 0.52% and an increase of 1.27% in positive and neutral sentiment respectively, as well as an increase of 0.66% in negative sentiment. Table 5 further expands on the differences between sentiments before and after the U.K.'s official departure. The level of "anticipation' was also noticeable higher before Brexit but also includes negative tweets like *'There is nothing good coming out of Brexit. It is just harming Britain socially and economically, causing suffering, hugely damaging our reputation abroad, and taking away rights from our children. I know it's hard to abandon an ideology but it's time to face reality.'* that scored high in "anticipation". On the contrary, the sentiment "fear", which also had a noticeable difference after Brexit, did reflect the feeling in tweets like *'@BrutumF Sadly I fear this all kicking off again now with Brexit and the restored border on N.Ireland. So sad. Like truly. I grew up in the Basque country surrounded by terrorism and political turmoil. Horrific. \nAfter following Yang I now realize that Brexit is a symptom, not the disease.'*.

The sentiment analysis program had limitations. Firstly, sentiment analysis cannot detect irony and sarcasm which may be problematic against British irony and sarcasm, such that may be more negative opinions veiled by positive words. Secondly, sentiment analysis cannot detect implicit sentiment such as misspellings and exclamation marks and even shorthanded text due to Twitter having a character limit. Third limitation is negation which some twitter users would use to emphasize an emotion, positive or negative.

Conclusion

The purpose of this paper was to perform a network analysis of Brexit on Twitter and YouTube, followed by a text and a sentiment analysis. Among the four research questions on network analysis, only research question 3 was able to see the significance between the dependent variable user_account created and independent variable number of account mentioning brexit, thus rejecting the null hypothesis. We failed to reject the null hypothesis for research question 1, 2, and 4. For the text analysis, we observed common words like "eu" and "uk" which cater to the location of users. We also observed "boris", "johnson" and "vote" before Brexit showing public action. Post Brexit, some of the unique words are "news", "politics" and "love". As for the sentiment analysis, "negative" and "neutral" sentiment increases and "positive" sentiment dropped after Brexit day. The limitation of the sentiment program, however, did produce errors on in the scoring such as the "anticipation"..

References

1. BBC News (2016). EU Referendum Result. Accessed on Jan 27. https://www.bbc.com/news/politics/eu_referendum/results
2. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
3. Jeff Gentry (2015). twitteR: R Based Twitter Client. R package version 1.1.9. https://CRAN.R-project.org/package=twitteR
4. Jockers ML (2015). _Syuzhet: Extract Sentiment and Plot Arcs from Text_. <URL: https://github.com/mjockers/syuzhet>
5. https://yougov.co.uk/topics/lifestyle/articles-reports/2019/01/11/half-americans-wouldnt-be-able-tell-british-person.

Appendix

Table 1. Research questions, hypotheses, tests, and variables

	Hypotheses	Variance Test	Variables
Research Question 1	**Ho**: There is a statistical variation between the video duration and view counts **Ha**: There is no statistical variation between length of video (Duration) and view counts	ANOVA	duration vs. view counts
Research Question 2	**Ho**: Dates significantly affect view counts **Ha**: Dates do not significantly affect view counts	NA	View counts vs published date
Research Question 3	**Ho**: There is a significant spike on account creation in June 2016. **Ha**: There is no significant change in accounts created in June 2016.	Regression	User account creation date vs. number of accounts mentioned 'brexit'
Research Question 4	**Ho**: The tweets do reflect the voter turnout for each United Kingdom constituent countries **Ha**: The twits do not reflect the voter turnout for each United Kingdom constituent countries	N/A	Twitter user location vs. each UK country turnout
Research Question 5	**Ho**: top 15 unique words in tweets after Brexit is no different than before Brexit **Ha**: top 15 unique words in tweets after Brexit is different than before Brexit	N/A	frequency of top 15 unique words before Brexit vs. frequency of top 15 unique words afterBrexit
Research Question 6	**Ho**: There is no change in sentiment scores from tweets before Brexit **Ha**: There is a change in sentiment scores from tweets before Brexit	N/A	sentiment score before Brexit vs. sentiment score after Brexit

```
> videolist <- read.csv("Youtube_videolist.csv")
> colnames(videolist) #to see all the column names
 [1] "position"           "channelId"          "channelTitle"       "videoId"
 [5] "publishedAt"        "publishedAtSQL"     "videoTitle"         "videoDescription"
 [9] "videoCategoryId"    "videoCategoryLabel" "duration"           "durationSec"
[13] "dimension"          "definition"         "caption"            "thumbnail_maxres"
[17] "licensedContent"    "viewCount"          "likeCount"          "dislikeCount"
[21] "favoriteCount"      "commentCount"
> #Does view counts(dependent) matter based on video duration(independent) on YouTube?
> str(videolist$viewCount)
 int [1:301] 4989 58441 9801 73436 51936 111563 204774 142366 3343875 1729 ...
> anovamodel <- aov(viewCount ~ durationSec, data = videolist)
> summary(anovamodel)
             Df    Sum Sq   Mean Sq F value Pr(>F)
durationSec   1 7.839e+11 7.839e+11   0.383  0.536
Residuals   297 6.077e+14 2.046e+12
2 observations deleted due to missingness
```

Fig.1. Video duration vs video counts analysis

```
library(ggplot2)
library(tidyr)
library(tidyverse)
library(dplyr)
channelTitle_viewCount <- read.csv("Youtube_videolist.csv")
#the correlation between the published date and the view counts
channelTitle_viewCount
```

```
df.cleaned <- channelTitle_viewCount[-c(254, 301),]
df.cleaned <- separate(df.cleaned, publishedAt, into = c("date", "time"), sep="T")


df.cleaned <- separate(df.cleaned, date, into = c("year", "month", "day"), sep = "-")
df.cleaned <- subset(df.cleaned, year > 2019)
df.cleaned
```

```
ggplot(data = df.cleaned) + geom_bar(mapping = aes(x = day, y = viewCount), stat =
"identity") + ggtitle("Number of View Counts in January")
```

Fig.2. R codebase

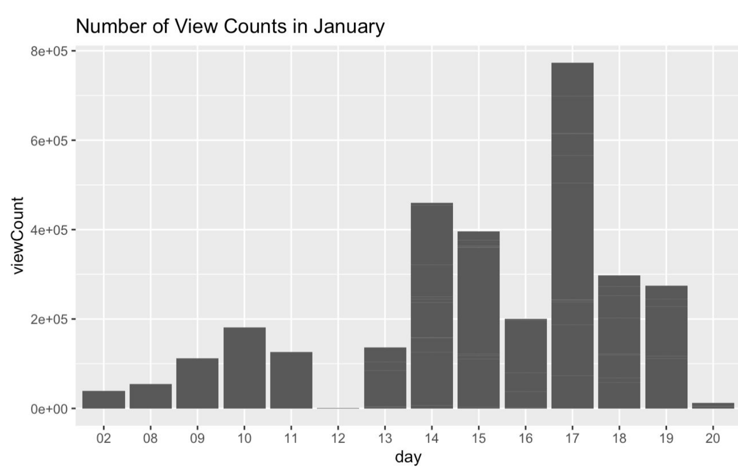

Fig.3. Graph between number of viewcounts vs days in January

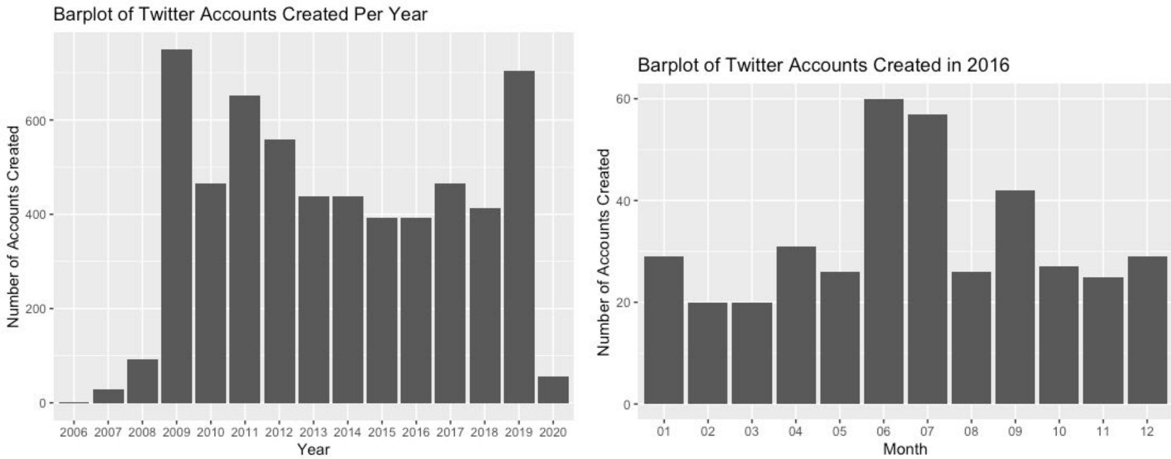

Fig.4. Visualization of Twitter data over the years.

```
brexit <- read.csv(file.choose(), header=T)
brexit <- brexit%>%
  separate(user_created_at, into = c("Year", "Month", "Date", "Time"), sep = "-
brexit <- brexit%>%
  separate(brexit_referendum_month, into = c("ref_Year", "ref_Month", "ref_Date
brexit <- brexit%>%
  separate(brexit_triggered_date, into = c("trig_Year", "trig_Month", "trig_Dat

year_2016 <- subset(brexit, Year == 2016)
month_vector <- c(year_2016$Month)
month_vector <- as.numeric(month_vector)
ref_month_vector <- c(year_2016$ref_Month)
ref_month_vector <- as.numeric(ref_month_vector)

# Does brexit relate to account creation in 2016
model <- lm(month_vector ~ ref_month_vector, data=year_2016)
summary(model)          as.numeric(x, ...)
```

```
Call:
lm(formula = month_vector ~ ref_month_vector, data = year_2016)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.6862 -2.6862  0.3138  2.3138  5.3138 

Coefficients: (1 not defined because of singularities)
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)        6.6862     0.1581    42.3   <2e-16 ***
ref_month_vector       NA         NA      NA       NA    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.13 on 391 degrees of freedom
```

Fig.5. R codebase to analyze year-based data

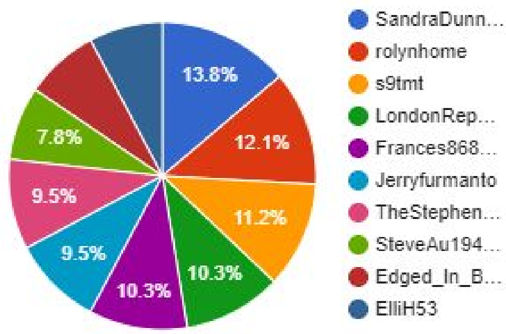

Fig. 6. Source Data: Top Ten Posters

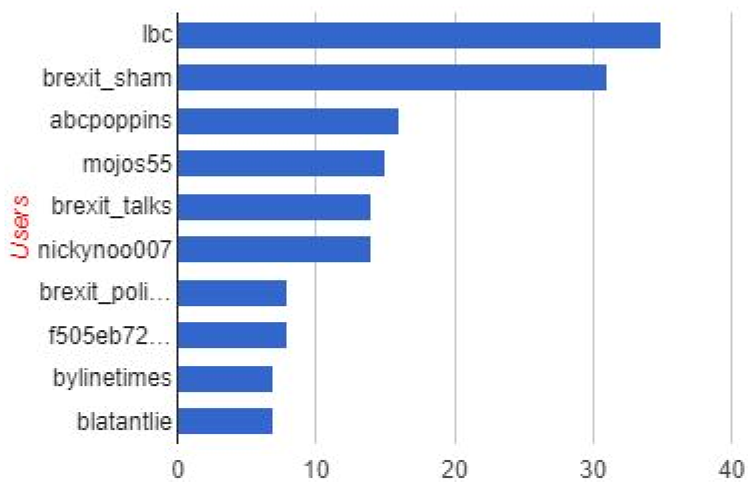

Fig. 7. Network: Top 10 Posters Mentioned in Messages

Id	Label	Interval	followers	following	In-Degree	Out-Degree	Degree	Dynamic In-Degree	Dynamic Out-Degree	Dynamic Degree
n1701	borisjohnson				151	0	151	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2543	nicolasturgeon				40	0	40	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1650	bbc				32	0	32	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2283	june_mummery				28	0	28	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2390	lozzafox				27	0	27	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1060	lbc		379206	1888	25	1	26	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1547	adamposen				25	0	25	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2500	mrjamesob				23	0	23	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n331	brexit_sham		49547	8081	22	2	24	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2222	jeremycorbyn				21	0	21	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2546	nigel_farage				21	0	21	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1684	bestforbritain				19	0	19	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1724	brexitparty_uk				19	0	19	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2235	joannaccherry				19	0	19	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n3109	youtube				19	0	19	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n2795	scotsecofstate				18	0	18	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n3107	yougov				18	0	18	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1668	beindorein				17	0	17	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..
n1758	carolecadwalla				17	0	17	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2..

Fig. 8. In-Degree calculation for Twitter network

Id	Label	Interval	followers	following	In-Degree	Out-Degree	Degree	Dynamic In-Degree	Dynamic Out-Degree	Dynamic Degree
n102	cathyby		4143	1833	6	50	56	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n104	jag11814459		349	559	6	50	56	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n310	shlomoindiana		177	28	2	50	52	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n309	charlie_45uk		1852	1838	1	50	51	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n16	meallyhen		1414	1256	0	50	50	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n100	andycapp321		16	23	0	50	50	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n105	krys51		746	1181	0	50	50	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n101	bigbaybear		3359	2052	5	49	54	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n103	gazza41		575	1061	5	49	54	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n20	apritions		2844	4086	0	49	49	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n7	deightonsue		322	355	1	48	49	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n8	paulthatistall		1795	2006	1	48	49	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n35	bremaininspain		19069	15014	3	47	50	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n168	f505eb72b9cd...		764	1442	8	46	54	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n159	lunaperla		10613	9776	0	27	27	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n672	inabster		32806	35783	1	26	27	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n4	jimmyc558240...		6	55	0	26	26	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n56	littleegret1		317	819	0	21	21	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.
n55	kimsussex3		2742	3052	5	20	25	<[2020-01-13, 202...	<[2020-01-13, 2020...	<[2020-01-13, 2.

Fig. 9. Out-Degree calculation for Twitter network

Fig. 10. Out-Degree for Twitter network

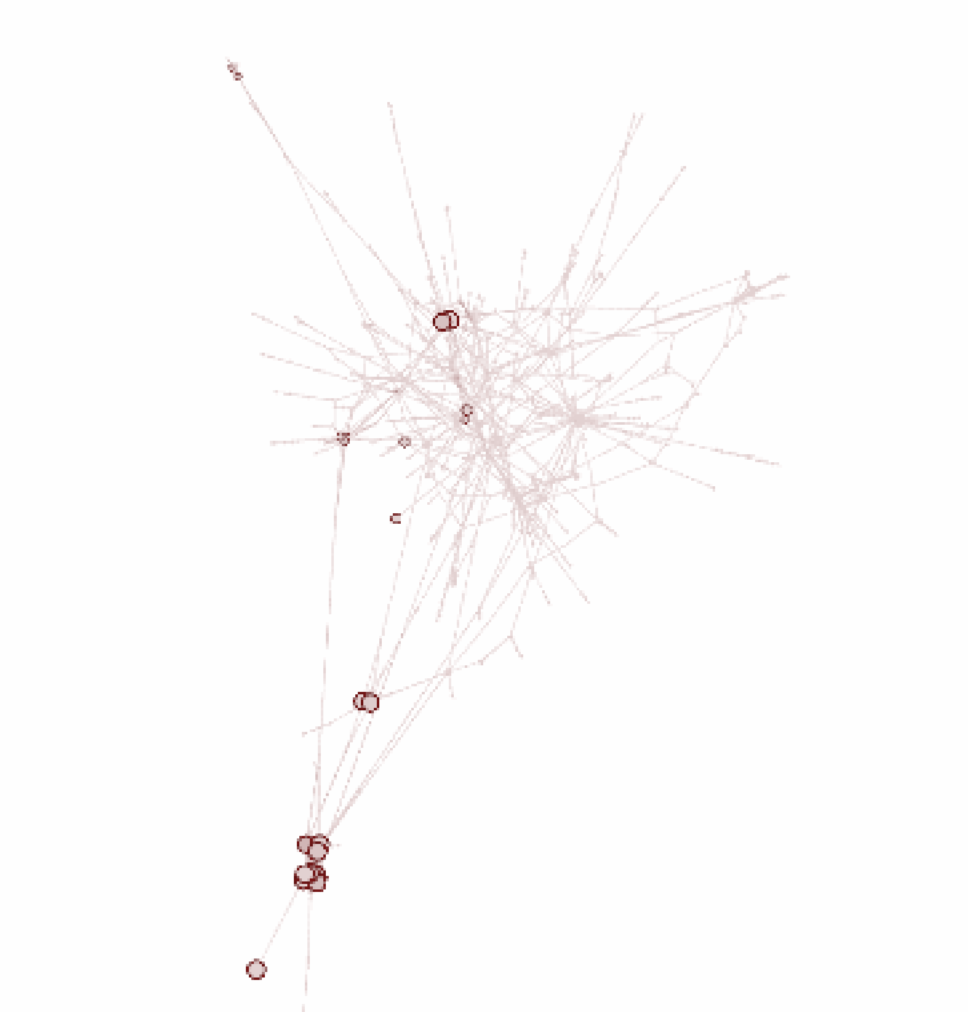

Fig. 11. Out-Degree visualization for Twitter network

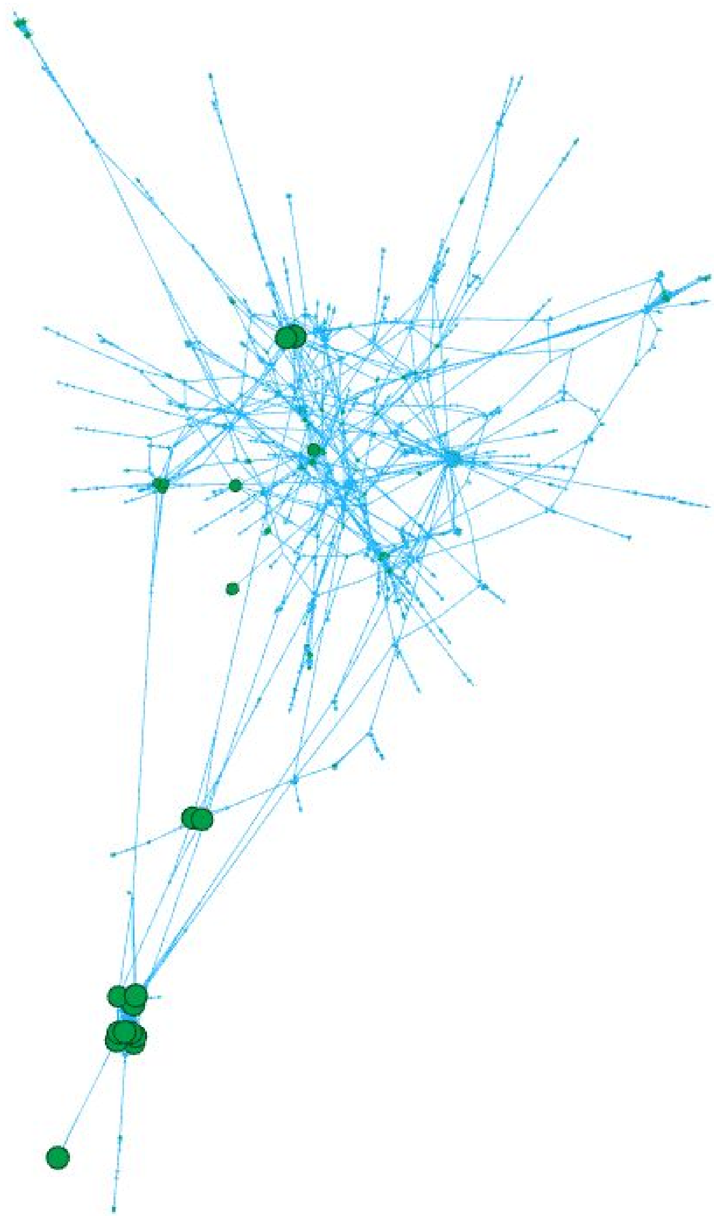

Fig. 12. In-Degree visualization for Twitter network

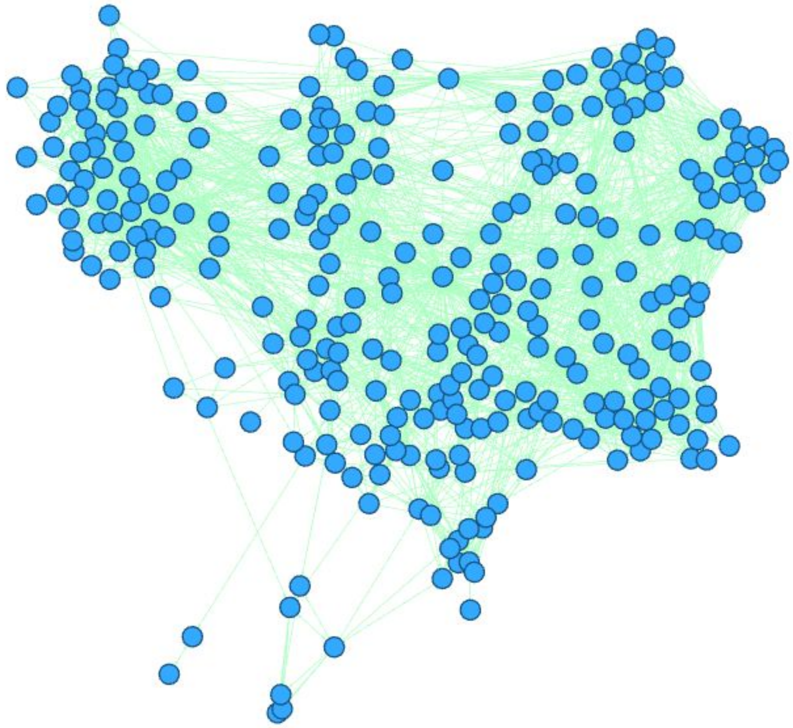

Fig. 13. In-Degree visualization for Youtube network

Id	Label	Interval	isseed	seedrank	publishedat	channeltitle	channelid	videocategorylabel	viewcou...	likecount	dislikecount	dislikelikeratio	favoritecount	commentcount
iAgKHSNqx...	Brexit: Last...		yes	14	14664042...	LastWeekT...	UC3XTzVz...	Entertainment	16663712	171632	30650	0.17858	0	35798
HaBQfSAVt0s	Brexit III: L...		yes	59	15504750...	LastWeekT...	UC3XTzVz...	Entertainment	13207041	157034	7753	0.049371	0	19903
fyVz5vgqBhE	Brexit II: La...		yes	203	14972490...	LastWeekT...	UC3XTzVz...	Entertainment	9421139	137010	5287	0.038588	0	8466
m3_l2rfApYk	Brexit Briefly		yes	53	14685773...	CGP Grey	UC2C_jShtL...	Education	4657483	100384	5563	0.055417	0	15903
GLp7LtXm...	Brexit: Why...		yes	270	15208447...	PragerU	UCZWlSUN...	Education	4145971	68744	11309	0.164509	0	10984
BZF9eg35lXI	John Oliver...		yes	228	15499605...	The Late S...	UCMtFAi84...	Entertainment	4032070	40808	1602	0.039257	0	2908
7eoDwvl0...	Brexit expl...		yes	9	15370010...	Channel 4 ...	UCTrQ7HX...	News & Politics	3339011	39902	3177	0.07962	0	10093
J1Yv24cM2...	□□□ Brexit ...		yes	65	15522310...	CGP Grey	UC2C_jShtL...	Education	2697888	88531	1506	0.017011	0	11430
E5S1EMmC...	Brexit (201...		yes	189	15448103...	HBO	UCVTQuK2...	Film & Animation	1992091	25899	3351	0.129387	0	4017
UYonSZ8s3_o	Brexit: Fact...		yes	93	15435789...	Pindex	UCDg_EVB...	Film & Animation	1913999	40549	8134	0.200597	0	13604
_HDFegpX...	Brexit: End...		yes	220	15581657...	Pindex	UCDg_EVB...	Film & Animation	1787672	51091	9817	0.192147	0	15300
dcwuBo4Pv...	Why Brexit ...		yes	19	14710217...	TED	UCAuUUnT...	News & Politics	1716209	23936	15128	0.632019	0	9539
oqT1lneQ2...	Inside Brex...		yes	18	14823335...	Financial Ti...	UCoUxsWa...	News & Politics	1639151	5290	1581	0.298866	0	4489
mVy7faNK...	Fascinating...		yes	232	14758509...	Dillie Keane	UCEZYraifZ...	People & Blogs	1162120	9342	1239	0.132627	0	
EpYrclfgX2I	What could...		yes	16	15475828...	Channel 4 ...	UCTrQ7HX...	News & Politics	1129875	8626	1960	0.22722	0	6261
tLj6GHb1tt0	Forget polit...		yes	118	15485112...	Channel 4 ...	UCTrQ7HX...	News & Politics	1073417	5278	1665	0.31546	0	8633
68h9pSKn...	Chaos In P...		yes	125	15676820...	Comedy C...	UCsv4YOF...	Comedy	1048335	18385	935	0.050857	0	1301
OQSMr-3G...	Facebook s...		yes	223	15601838...	TED	UCAuUUnT...	News & Politics	857421	22421	3286	0.146559	0	4117
O8ZBKA9_...	¿Qué es el ...		yes	260	14668124...	Platzi	UC55-mxU...	Science & Techn...	835557	24468	2376	0.097106	0	3473

Fig. 14. Data Laboratory for Youtube network

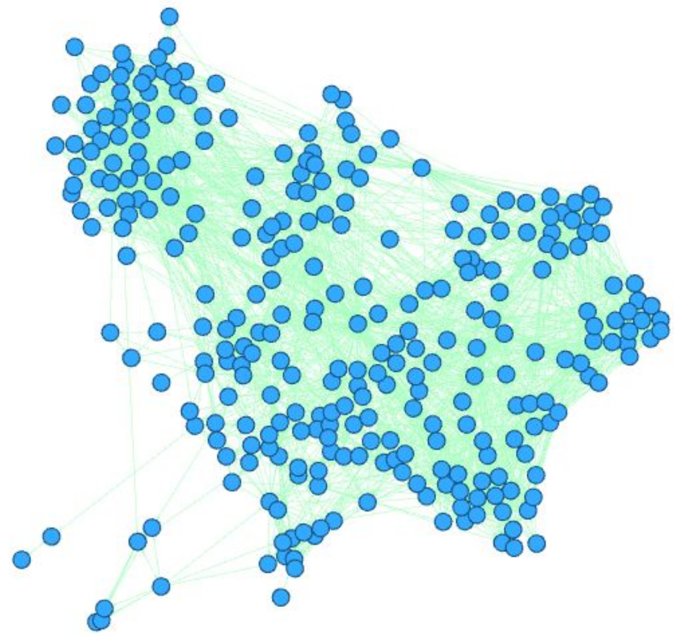

Fig. 15. Out-Degree visualization for Youtube network

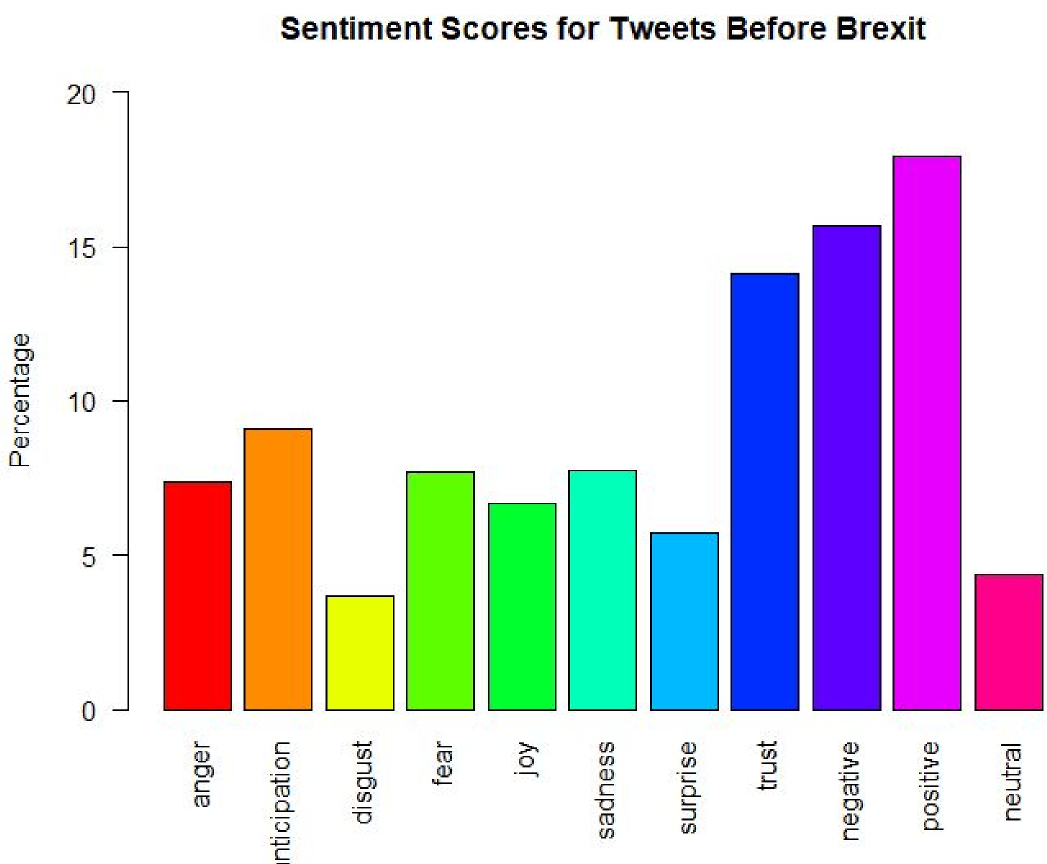

Fig. 16. Sentiment Analysis Before Brexit

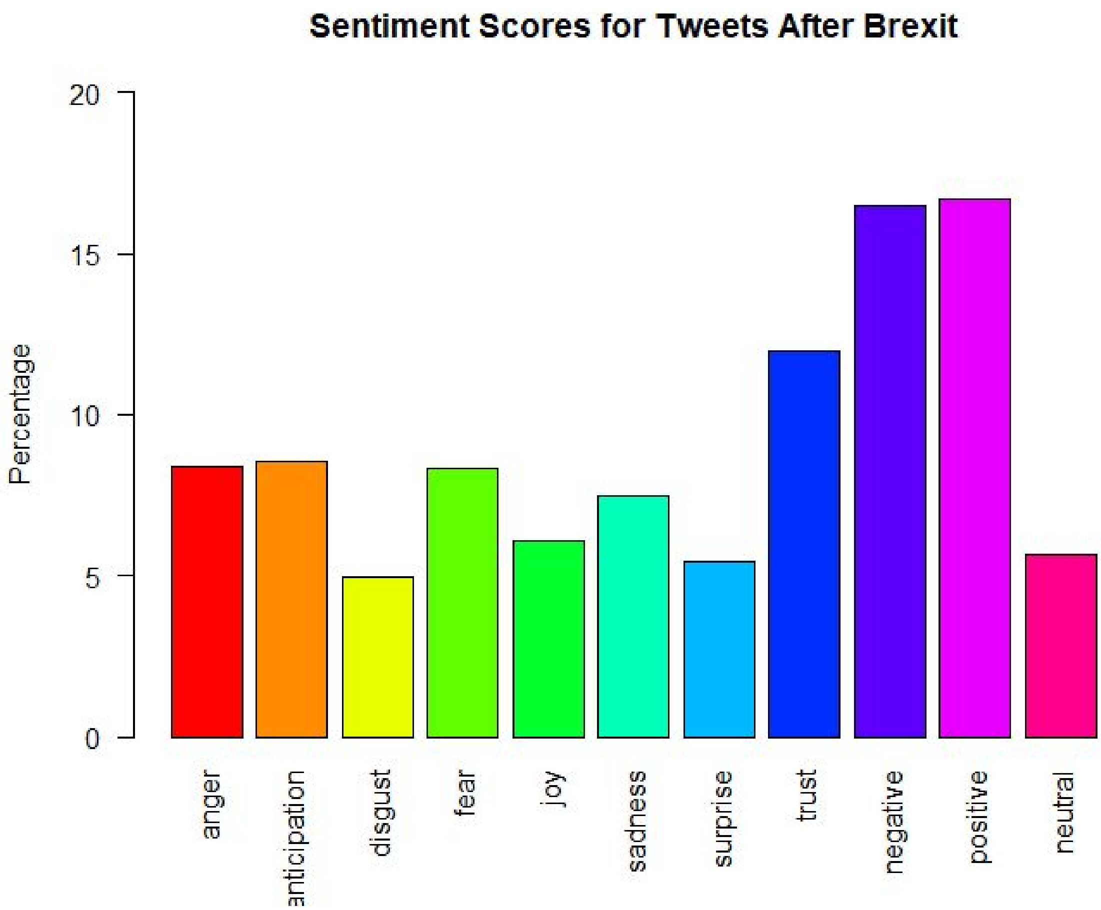

Fig. 17. Sentiment Analysis After Brexit

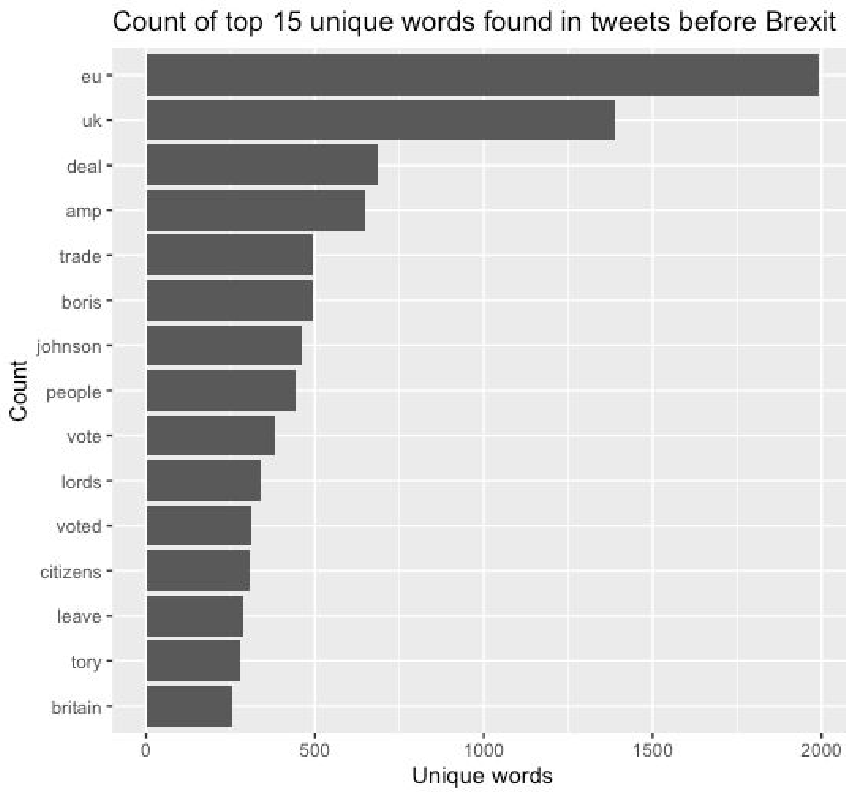

Fig 18. Unique Words Before Brexit

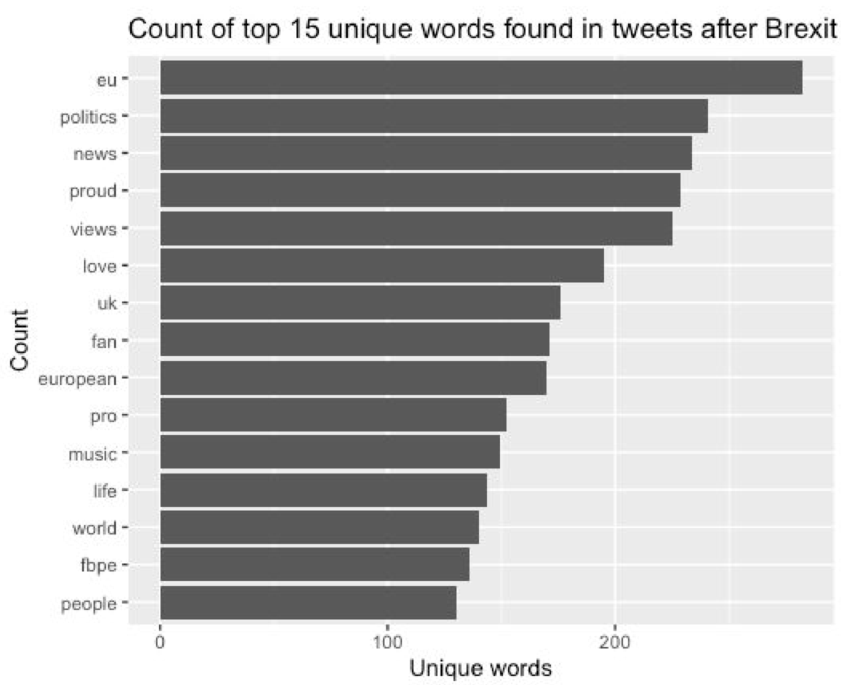

Fig. 19. Unique Words After Brexit

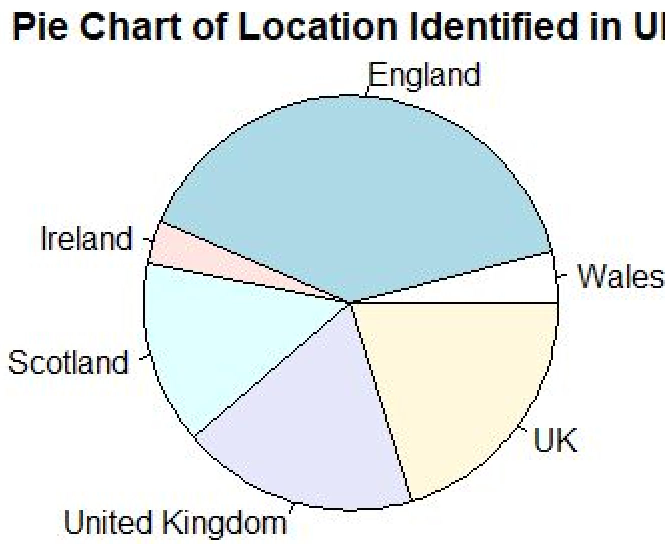

Fig. 20 Visual Representation of tweets from the United Kingdom constituent country identified from user_location

Table 2. Twitter location data

Location Identified on Twitter "user_location"	**Number of Tweets**	% of turnout by territories	% of nation turnout from BBC News
No Location Input	1825	NA	NA
Has location Input	4021	NA	NA
Locations part of UK territories	2071	NA	NA
united kingdom	384	18.54%	NA
uk	418	20.18%	NA
england	818	39.50%	73.3%*
ireland	69	3.33%	62.7%*
scotland	290	14.39%	67.2%*
wales	84	4.05%	71.7%*

Table 3. Twitter Network Indegree Analysis

Twitter Handle	In-degree	Bio	Classification
borisjohnson	151	Prime Minister of the United Kingdom and Conservative leader. Member of Parliament for Uxbridge and South Ruislip	Politician
nicolasturgeon	40	"First Minister of Scotland, @theSNP Leader and MSP for Glasgow Southside. Loves 📚"	Politician
bbc	32	Our mission is to enrich your life and to inform, educate and entertain you, wherever you are.	Media Outlet
june_mummery	28	Brexit Party MEP for East of England. Business owner and Founder of REAF - Renaissance of the East Anglian Fisheries⚓	Politician
lozzafox	27	Wokey McWokeface 👉🏼“A Grief Observed” is out now. 💿 here 👉🏼 https://listnin.co/AGriefObserved 🎵 🎥 https://youtu.be/E0NrIEt8rrw	Musician
lbc	25	Leading Britain's Conversation.	Media Outlet
adamposen	25	President, Peterson Institute for International Economics; proclaims on monetary, fiscal and trade policy, advises investors. Tweets/RT only on own behalf.	Academic
mrjamesob	23	Podcast: https://l-bc.co/2SGnJ7B Book: http://po.st/HowtoBeRight Radio: http://LBC.co.uk	Reporter
brexit_sham	22	Using information, education and agitation to expose the shills and charlatans responsible for stealing our EU rights and freedoms. #brexitshambles #brexitscam	Activist Group
jeremycorbyn	21	Leader of the Labour Party.	Politician
nigel_farage	21	"Leader of @BrexitParty_UK."	Politician

bestforbritain	19	"We are the UK's leading #StopBrexit campaign▪Our CEO is @pimlicat ▪Tweets by @pranman (~pm)▪Guests: @sturdyAlex (~AA), @JonLis1 (~JL), @localnotail (~JT)"	Activist Group
brexitparty_uk	19	We Are Ready to Change Politics for Good. Register your support today: http://thebrexitparty.org	Political Party
joannaccherry	19	SNP MP for Edinburgh South West, Queen's Counsel and Feminist.	Politician
youtube	19	Like and subscribe.	Media
scotsecofstate	18	Official account of the Office of the Secretary of State for Scotland.	Politician
yougov	18	Exploring what the world thinks, every day	Research Group
beindorein	17	🏴#FBPE #FBSI🇪🇺	Layman
carolecadwalla	17	Late adopter. Late giver-upper. Guardian & Observer writer.	Reporter
guyverhofstadt	17	"MEP @RenewEurope and Brexit Coordinator for @Europarl_EN #IAmEuropean🇪🇺"	Politician
keithbrownsnp	17	SNP MSP for the Clackmannanshire & Dunblane constituency, Depute Leader of the SNP. Retweets are not endorsements	Politician
telegraph	17	Think ahead with the latest news, comment, analysis and video	Media

libdems	15	Data protection: http://libdems.org.uk/privacy	Political Party
uklabour	15	We are the party for the many, not the few. 🌹	Political Party
bbcpolitics	14	The best of the BBC's political coverage. Find us on Facebook too: https://facebook.com/BBCPolitics	Media Outlet
conservatives	14	👍 Get Brexit Done. 🇬🇧 Unleash Britain's Potential. 🌳 Join the Conservatives today!	Political Party
jessphillips	14	"Labour MP for Birmingham Yardley http://jessphillips.net"	Politician
the3million	14	"We are the largest grassroots organisation of EU27 citizens in the UK, campaigning to protect #citizensrights together with @BritishinEurope RTs not endorsement"	Activist Group
bbclaurak	13	I know it's fashionable, but even in 2019 there is nothing big or clever about shooting the messenger - tweets or retweets here aren't necessarily my view	Reporter
edwardjdavey	13	#LibDems Acting Leader of Commons, MP for #Kingston & #Surbiton, former Sec of State #Energy & #ClimateChange, husband, father, campaigner on #ClimateEmergency	Politician

Table 4. Text Analysis

Top 15 unique Words before Brexit	Top 15 unique words after Brexit
eu	eu
uk	politics
deal	news
amp	proud

trade	views
boris	love
johnson	uk
people	fan
vote	european
lords	pro
voted	music
citizens	life
leave	world
tory	fbpe
britain	people

Table 5: Differences in Sentiment Before and After Brexit using "syuzhet" R package

Sentiment	**Sentiment before Brexit (%)**	**Sentiment After Brexit (%)**	**Difference in sentiment (%)**
Neutral	4.39	5.66	-1.27
Trust	8.84	7.71	+1.14
Disgust	3.21	4.06	-0.84
Negative	8.70	9.36	-0.66
Joy	5.48	4.83	+0.65
Positive	9.90	9.38	+0.52
Anticipation	6.62	6.20	+0.41
Fear	5.54	5.91	-0.37
Surprise	4.87	4.52	+0.35
Anger	5.74	6.07	-0.32
Sadness	5.78	5.68	+0.11