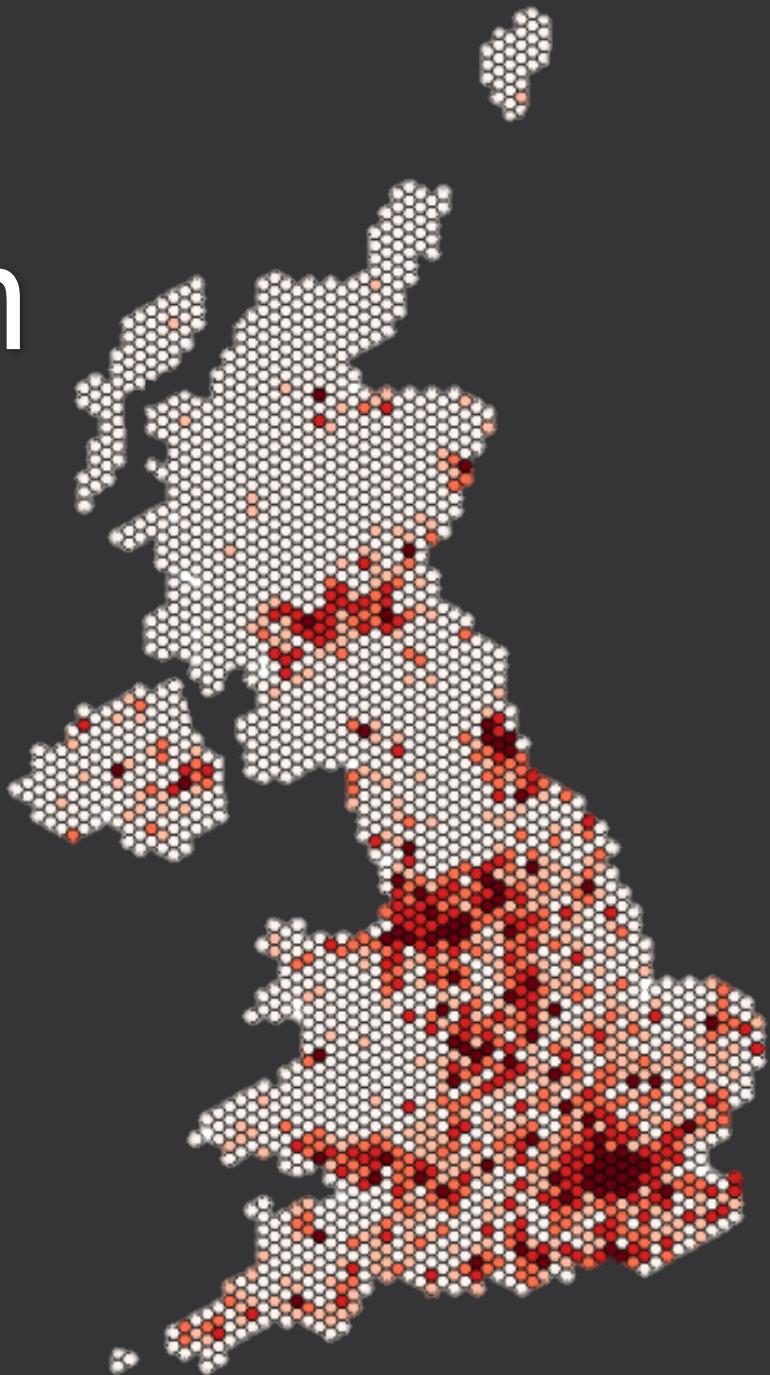


# Twitter Topics Geography of the UK EU Referendum 2016

Rob Webster



# Who is Rob Webster?



BSc Geography (2012 – 2015)



MSc GIS (2016 – 2017)



Previously at Sterling Geo.



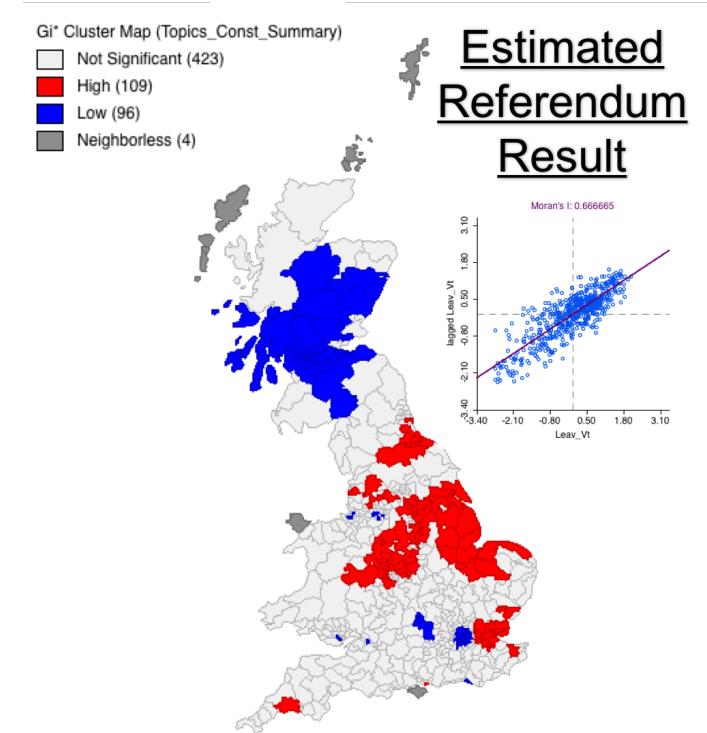
Now Trainee Software Engineer

@IMGeospatial

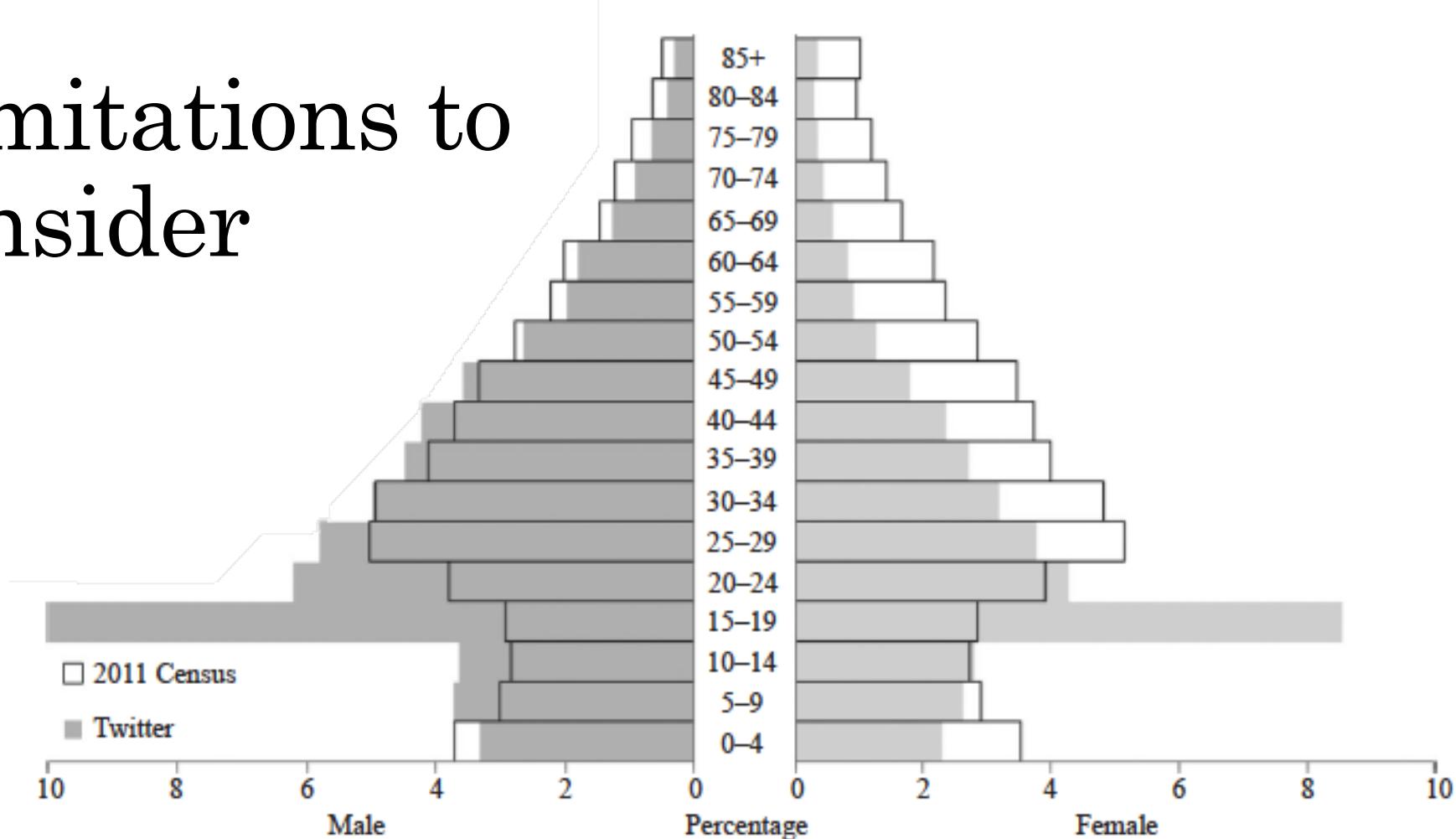
# Why look at EU Referendum using data from Twitter?

- Massive source of time series data – Good for looking at events
- Geolocated
- Available to get data through API
- Use R!!!!

```
library(tools)
library(rgdal)
library(GISTools)
#Define Projections
ukgrid = "+init=epsg:27700"
latlong = "+init=epsg:4326"
for(i in 1:19){ #for each week
  Directory <- paste("~/.../WK", i, sep = "") #set to weeks directory
  setwd(Directory)
  filenames <- list.files(, pattern = ".CSV")
  for(j in 1:length(filenames)) { #for each topic
    N = filenames[j]
    file_name = paste(Directory, "/", N, ".CSV", sep = "")
    #read in file
    df = read.csv(file_name, header = TRUE, sep = ",")
```



# Limitations to consider



Age Distribution of Twitter Users in London compared with the 2011 Census. Originally Taken from Longley, Adnan & Lansley, 2015

# Research Questions

- 1 ) How did topics during the EU referendum campaign change on twitter throughout the course campaign and did this change spatially?
- 2) Does twitter highlight any events that appeared to shape the campaign on the platform and where was this most prevalent?
- 3) Is it possible to infer any relationships between twitter topics/intensity and votes at a constituency level?

# Obtaining the data

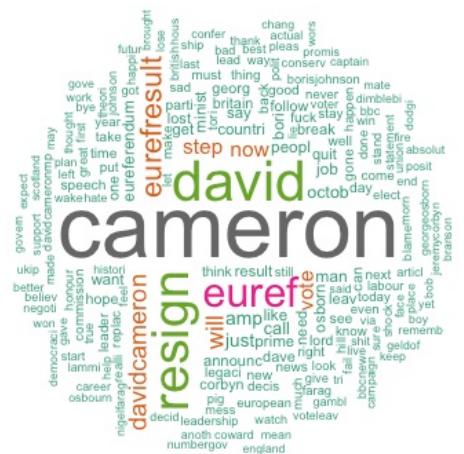
- Given access to Archive of Tweets. Selected tweets from Feb to June 2016.
- Searched and Filtered using Keywords and Hashtags relating to the referendum.
- Also collected a random sample of tweets too.
- Filtered to remove spam accounts, duplicates etc.
- Split tweets into weeks.

## Tweet Search Terms

referendum, eu, brexit , vote, remain, leave, europe, campaign, #strongerin, #remain, #voteremain,  
#votein, #bremain, #labourin, #votestay, #intogther, #labourinforbritain, #greenerin, #yes2eu, #yestoeu,  
#betteroffin, #votein, #ukineu,  
#bremain,#strongerin,#greenerin,#britin,#projectfact,#remaineu,#leavechaos,  
#voteteavelosecontrol,#niineurope,#scotlandineurope,#scotlandineu,#voteteave,  
#leaveeu,#takecontrol,#betteroffout,#voteout,#beleave,#no2eu,#betteroffout,#notoeu,#britainout,#loveeur  
opeleaveeu,#grassrootsout,#projectfear,#scexit

# Topic Modelling

- Used R and the topicmodels package to classify each tweet into topics.
  - LDA topic model applied.
  - Each tweet classified into a Number.
  - Numbers then manually classified to Topic Name using Word Cloud.
  - Unrelated topics removed.



# Analysis Preparation

- Tweets aggregated by area, with an Odds Ratio calculated using the Sample Tweets for each topic.
- Tweets aggregated to two area types
  - Hexagonal Grid – For spatial analysis.
  - Constituency – For comparison with Estimated Referendum Result.
- Spatial Analysis – GeoDa
  - Moran's I
  - Getis-Ord
  - Regression Analysis – Spatial Model

# Summary of Tweets and Topics

- 291,797 Tweets classified as Referendum related.
- 102,941 unique users
- 111 Topics, Peak of 31 topics for a single day.

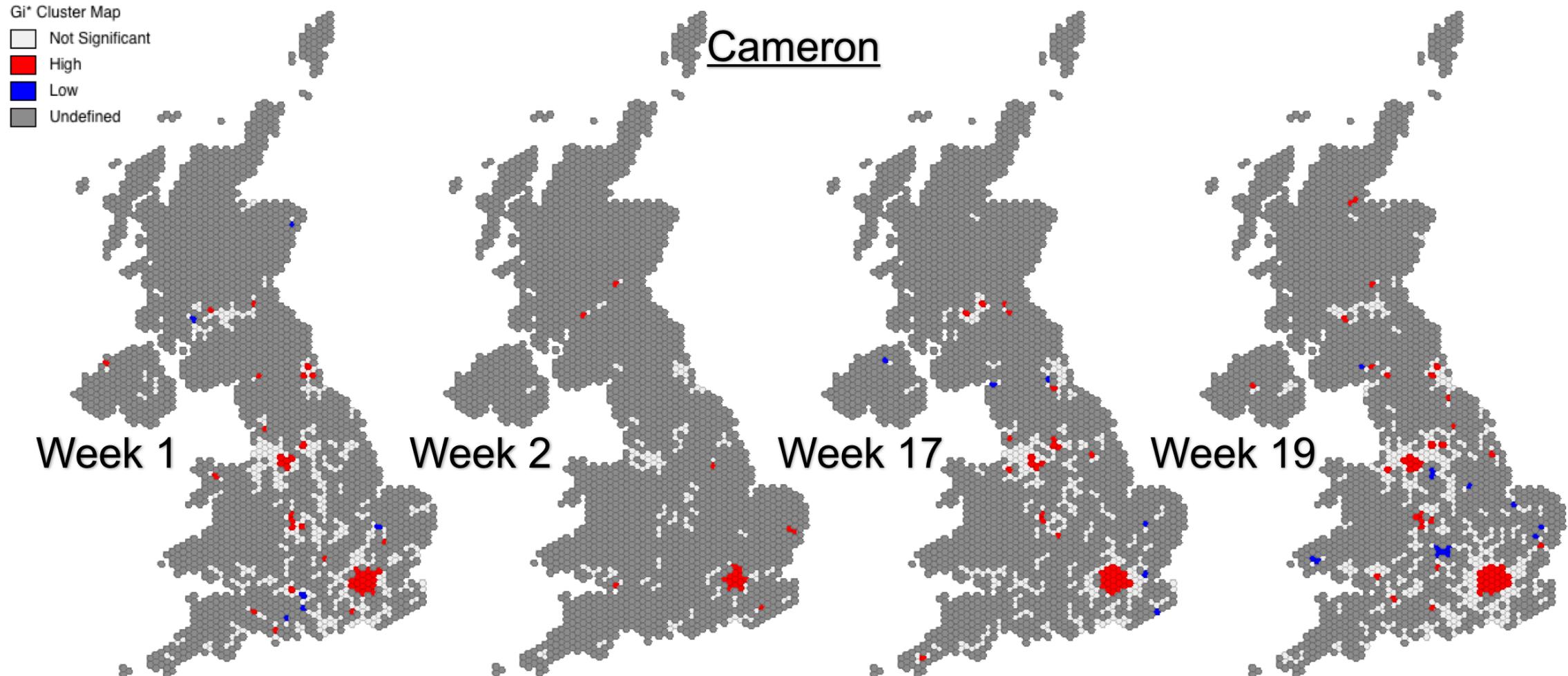
**Table 1** Top 10 Topics During Study Period

Topic	Number of Tweets
Debate	13324
Trump	12946
Trade	10990
Cameron	10696
Rights	9877
Scottish Independence	9151
Elections	8578
Stay	7890
Drop in Pound	6395
NHS	6206

# Summary of Results

- Generally, Twitter activity was restricted to the same areas throughout the campaign

# Evolution of Topic Discussion



# Summary of Results

- Generally, Twitter activity was restricted to the same areas throughout the campaign
- Very few Topics generated large engagement through most of the campaign, restricted to those very engaged politically (Westminster Bubble?).

## Number of Topic Classified Tweets Per Day

NO. TWEETS

60000

50000

40000

30000

20000

10000

0



DATE

Gi\* Cluster Map

- Not Significant
- High
- Low
- Undefined

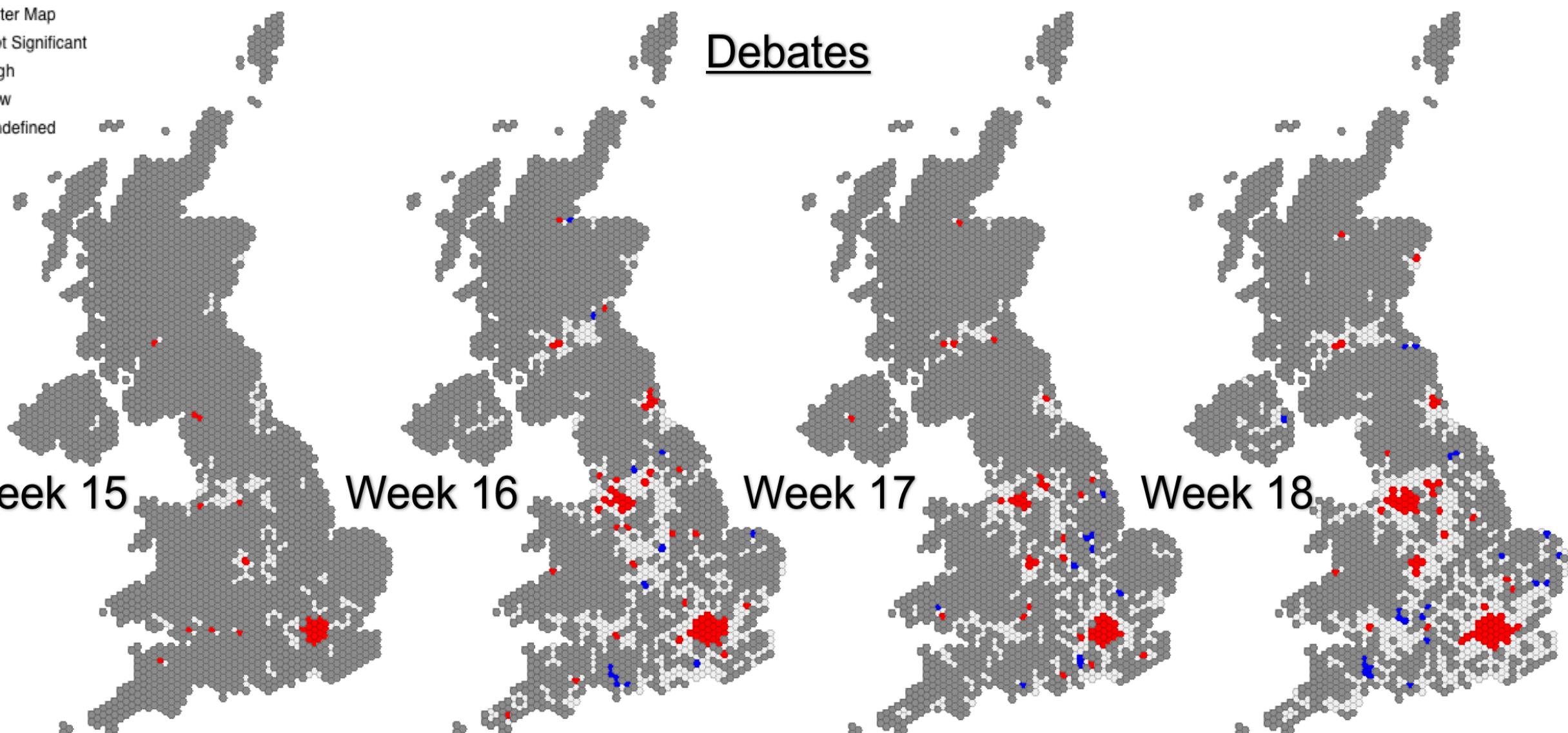
Week 15

Week 16

## Debates

Week 17

Week 18



## Proportion of Tweets by Hour for Jun 21<sup>st</sup> 2016

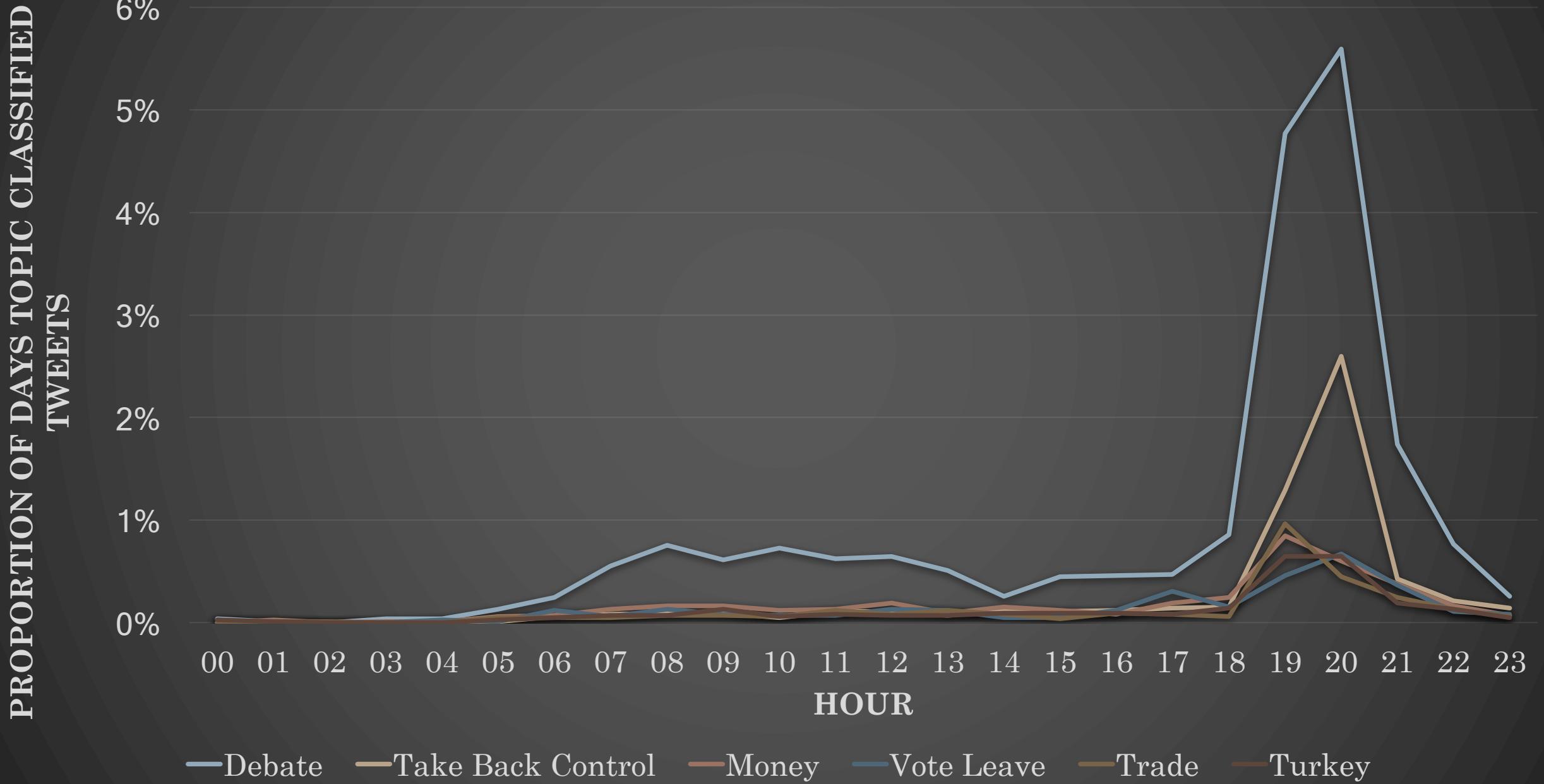
PROPORTION OF DAYS TOPIC CLASSIFIED  
TWEETS

6%  
5%  
4%  
3%  
2%  
1%  
0%

00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23

HOUR

Debate Take Back Control Money Vote Leave Trade Turkey

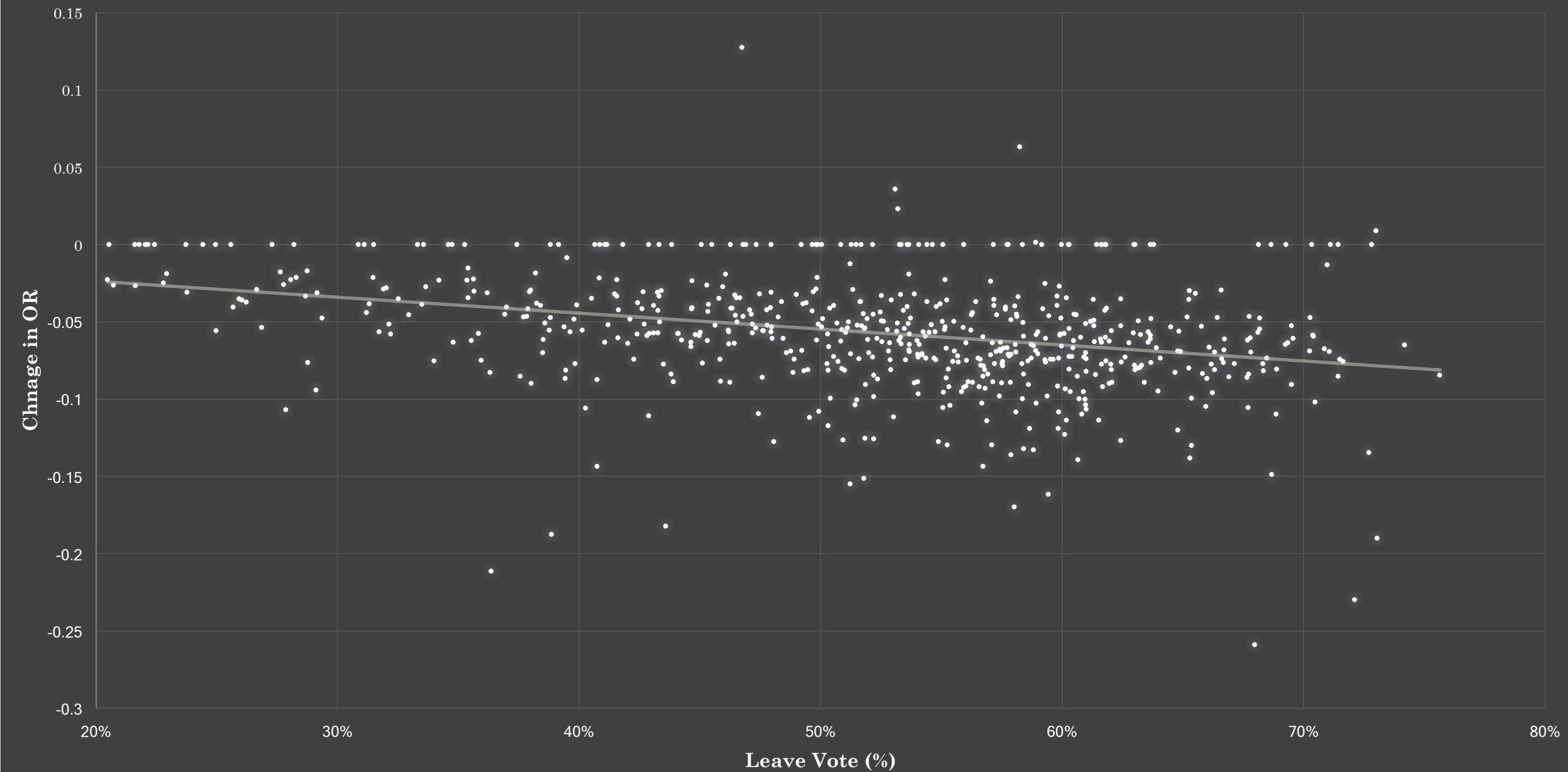


# Summary of Results

- Generally, Twitter activity was restricted to the same areas throughout the campaign
- Very few Topics generated large engagement through most of the campaign, restricted to those very engaged politically (Westminster Bubble?).
- Clear switch in engagement after the Referendum result with topics. Unsurprising, it was unexpected!!

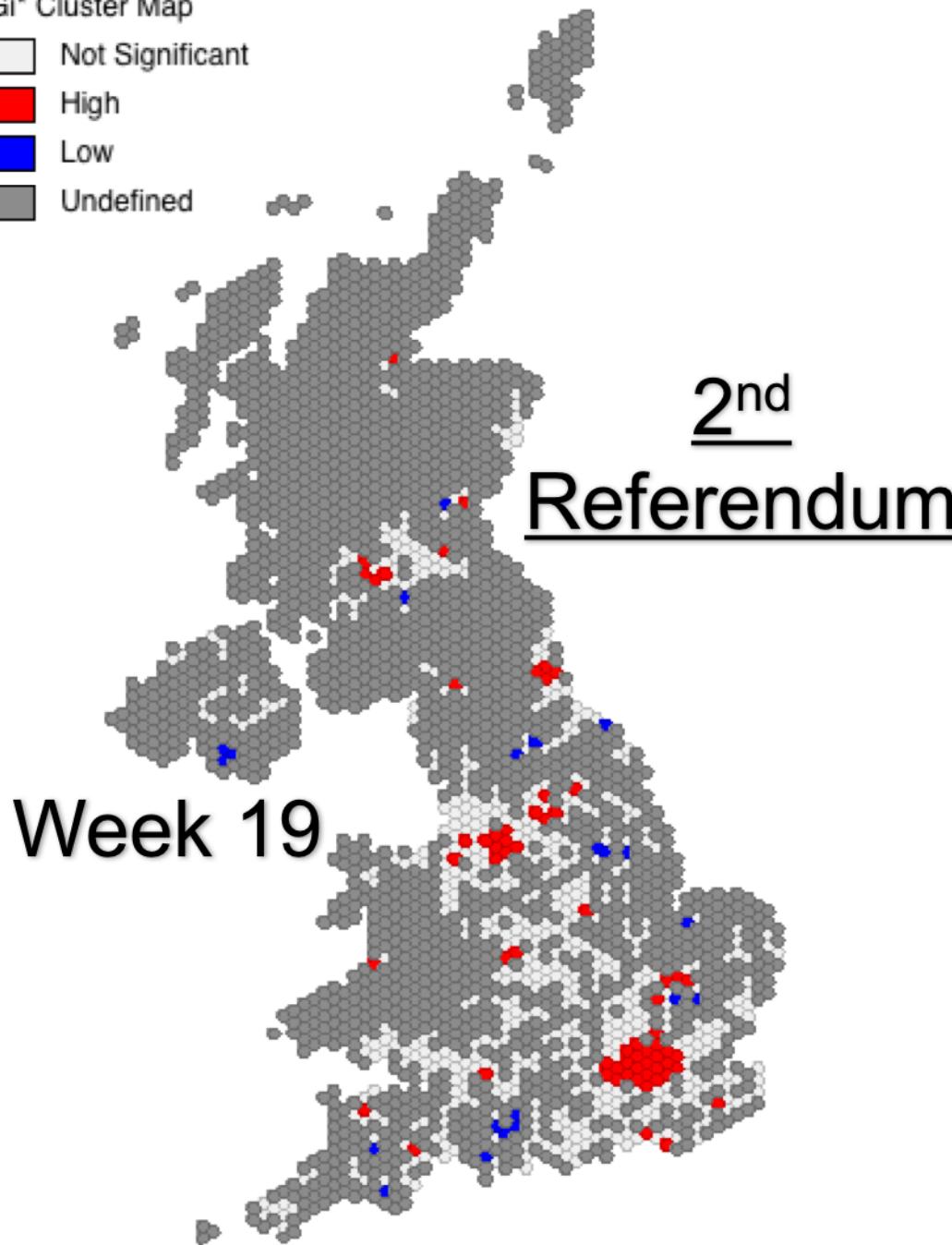
Topics Extracted		
Week 18 (Before the Result)	Week 19 (After the result)	
Debate	Another Referendum	Lies
Geldof Farage	Boris Johnson	Move
Jo Cox	Cameron	NHS
Law	Change	Passport
NHS	Control	People Racist
People	Democratic	Pound Drops
Political Parties	Election	Racist
Referendum	Fear Hate	Sad
Stronger In	Feelings	Scottish Independence
Take Back Control	Free Movement	Shock
Trade	Future	Thought
Turkey	Generations	Trade
Vote Leave	Good Bad Thing	Trump
Workers Rights	Happen	Wake Up to Result
	Important Vote	Working Class
	Labour Corbyn	

### Change in OR (Before to After Result)



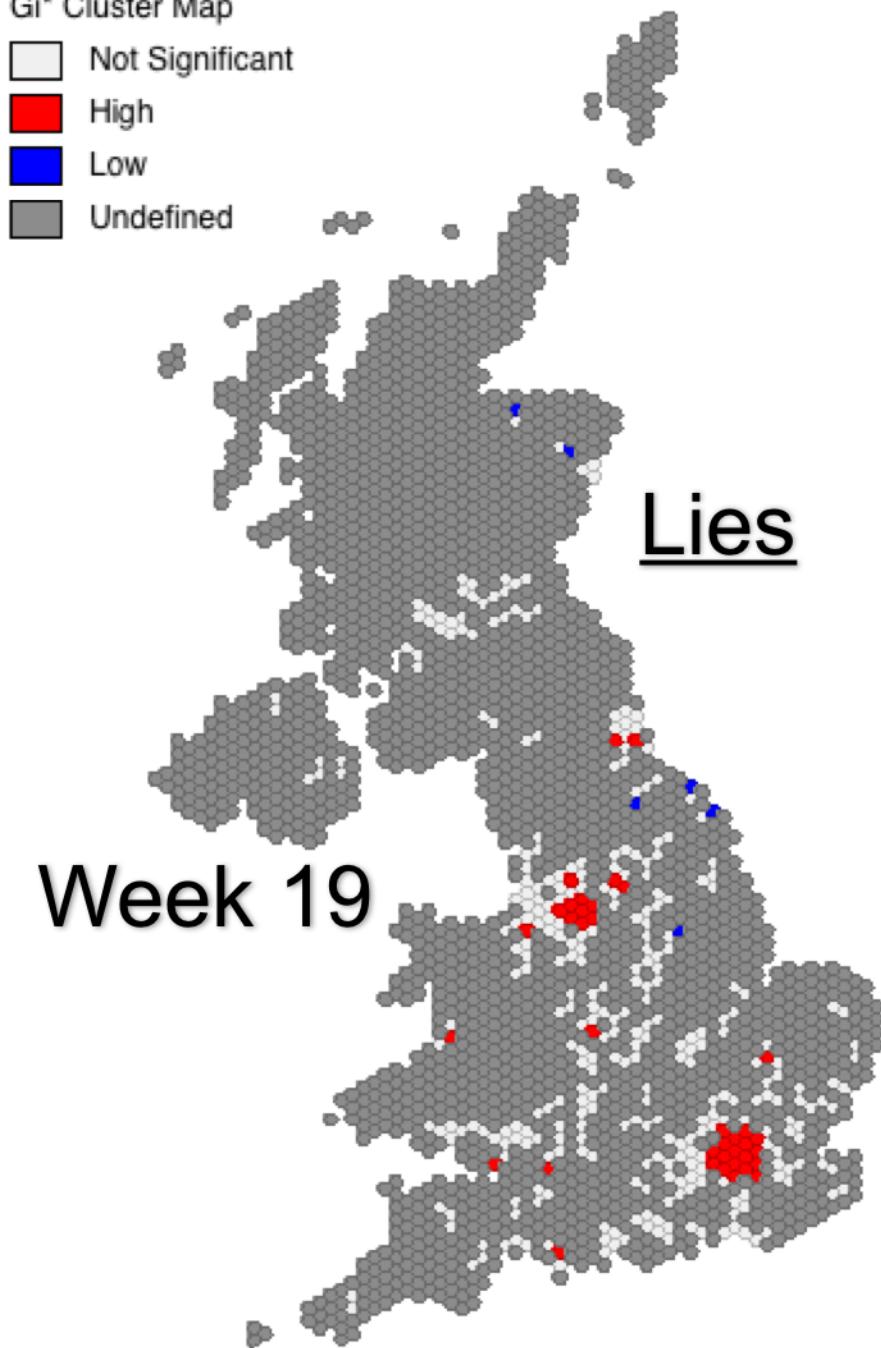
Gi\* Cluster Map

- Not Significant
- High
- Low
- Undefined



Gi\* Cluster Map

- Not Significant
- High
- Low
- Undefined



# Reflections on results

- GWR should have been used, for more rigorous testing between constituency voting and twitter topics. Spatial Error model an base indicator.
- More work needs to be done on the topic modelling, a good process but needs finesse and understanding of representativeness. Subjective.
- Perhaps another approach to examining events through social media could have been explored. Of all the events through the campaign that generated the widest engagement, it was the Result, highlights difficulty of measuring topics/opinions on the platform and making it useful.
- Always needs to be framed within twitter, not trying to project onto beyond the platform onto general population (e.g. opinion polling using social media).

Thank You and Enjoy the rest of  
GISRUK!!! ☺

Any questions feel free to chat or contact me elsewhere.



@RobWebsterUK



[www.linkedin.com/in/rob-webster-geo](https://www.linkedin.com/in/rob-webster-geo)

[rob.a.webster1@gmail.com](mailto:rob.a.webster1@gmail.com)