



# The Refugee Crisis in Europe: The (Weak) Connection Between Political Media and Twitter Activity?

---

Aleksandar Stojkov, Ss. Cyril and Methodius University (Macedonia)





Figure 1: Syrian refugees on the Macedonian-Greek border attempting to reach Germany

Source: Photo taken by the author.



# Research inspiration

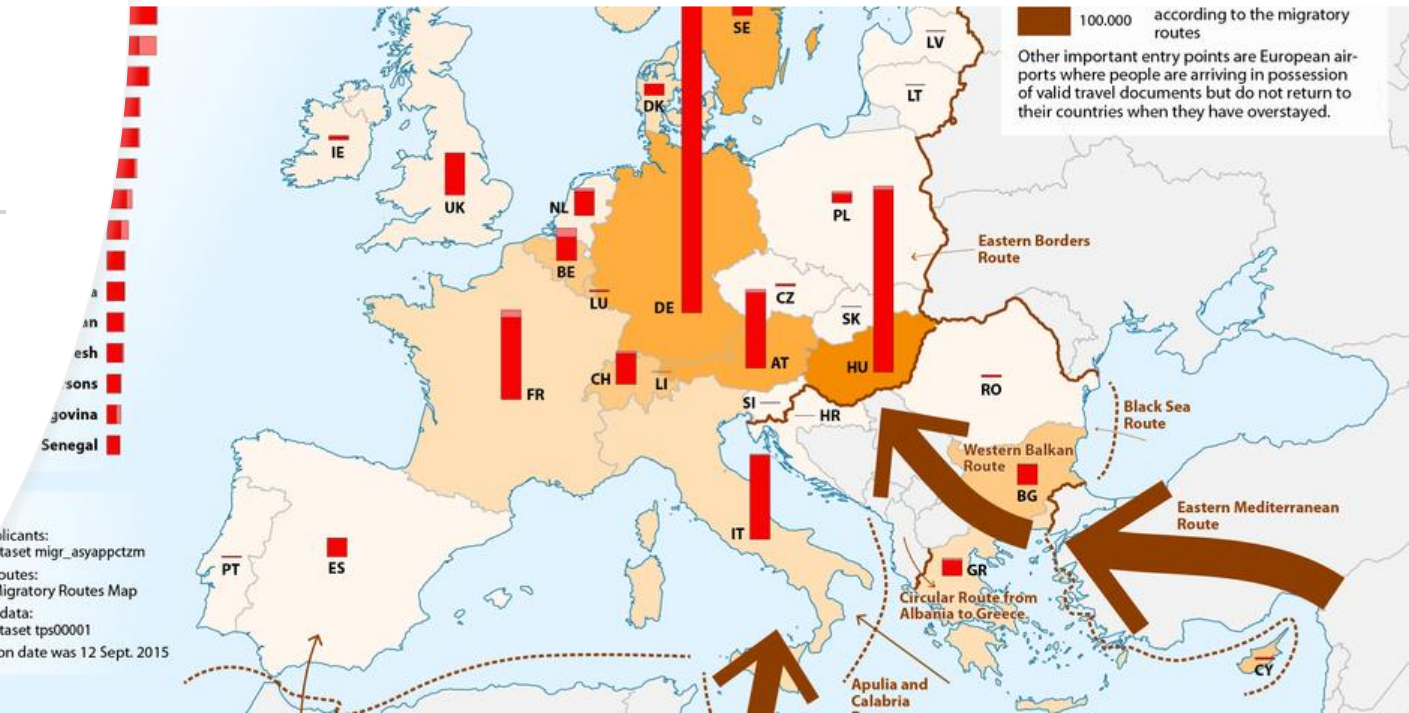


Figure 2: Migratory routes to Europe

Source: Syria's refugee crisis in numbers, *Amnesty International*, September 4<sup>th</sup>, 2015.

# Outline of the Presentation

---



INTRODUCTION



RESEARCH  
CONTEXT



LITERATURE  
REVIEW



METHODOLOGY



RESULTS



# 1. Introduction

---

- The power of news media to set a nation's political agenda and to focus public attention on a few key public issues is well-documented
- News media as primary source of those “pictures in our heads” (Lipmann, 1922) about the larger world of public affairs
- What we know about the world is largely based on what the media decide to tell us.
- More specifically, the result of this mediated view of the world is that the priorities of the media strongly influence the priorities of the public.
- The agenda-setting theory and its hypotheses can inform political communications studies of media influence.

# Our contribution

---

- Our contribution builds on the agenda-setting theory.
- Using a Bayesian way of thinking, we focus on one **issue** – the refugee crisis conversation, one **territory** – Europe – and we look at the different actors' **agendas**.
- Issues are often proxied in the literature by conversations pushed forward in the news media.
- Thus, we will use news media as a proxy, but we will also consider conversations directly generated by other actors, namely political groups and politically-motivated individuals

# Research objectives

---

- **Central research objective:**
  - To examine the opinion formation process and the driving forces of influential thinking on Twitter activity regarding the European refugee crisis 2014/15.
  - This is a methodological contribution (not a search for a potential causal relationship)
- **More specific research questions:**
  - (1) what content starts a conversation?;
  - (2) what are the relative strengths of traditional vs. social media
  - (3) what are the features of an influencer?



## 2. Research Context

- **Increase in asylum applicants.**
- The number of persons seeking asylum from non-EU countries in the EU-28 during the third quarter of 2015 reached 413,800. This was 250,400 more than in the same quarter of 2014. Out of the 430,600 total asylum applicants, 413,800 (96%) were first time applicants.
- **Countries of origin.**
- Citizens of 149 countries sought asylum for the first time in the EU in the third quarter of 2015. Syrians, Afghans, and Iraqis were the top three citizenships of asylum seekers, lodging around 138,000, 56,700 and 44,400 applications, respectively.
- **Destination countries.** The highest number of first-time asylum applicants in the third quarter of 2015 was registered in Germany and Hungary (both with slightly over 108 000 applicants, or 26% each of total applicants in the EU Member States), followed by Sweden (42,500, or 10%), Italy (28,400, or 7%) and Austria (27,600, or 7%).



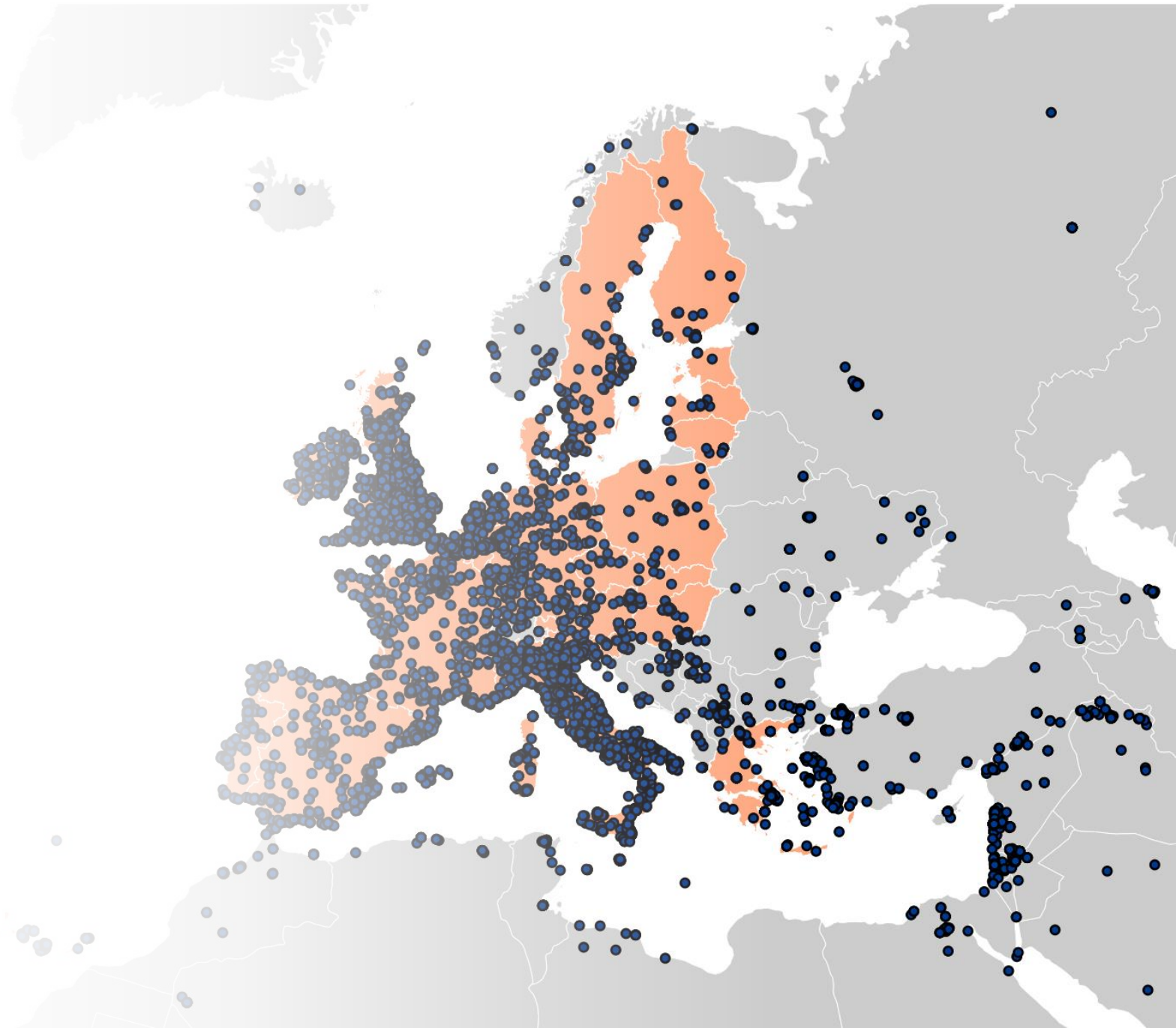




## Refugee Crisis in 2015 on Twitter

---

*The most intensive Twitter activity  
appears in UK (pre-Brexit), Italy,  
Germany, and France*





# 3. Literature Review

---



- Agenda-setting theory
- Media bias
- Agendas and social media
- Twitter



## 3.1. Agenda-setting Theory

- Three types of agenda-setting:
  - Policy agenda-setting or Political agenda setting
  - Media agenda-setting or Agenda building
  - Public/Audience agenda-setting
- The number of agenda-setting studies has increased over time.
- The choice itself of the topics could produce electoral effects, to the extent that most citizens consistently perceive one party to be more competent than the other at handling a given policy issue (the notion of issue ownership - Petrocik, 1996).
- Mass media outlets can influence the agenda of the public (Erbring, Goldenberg, and Miller 1980; Iyengar and Simon 2000).

## 3.1. Agenda-setting Theory (cont'd)

---

- Starting with the seminal contribution by McCombs and Shaw (1972), a host of empirical studies have investigated:
    - Replication of correlations between the media agenda and the public agenda (e.g., Ghanem and Wanta, 2001; Golan and Wanta, 2001; Tedesco, 2001).
    - Identification of contributing factors to different agendas (e.g., Fico and Freedman, 2001; Johnson, Wanta and Boudreau, 2004; Kwansah-Aidoo, 2003)
    - Explanation of the mechanisms of agenda setting (e.g., Brosius and Kepplinger, 1990; McLeod, Becker, and Byrnes, 1974; Shaw and Martin, 1992)
    - Formal development, testing, and refining the theory (e.g., Kosicki, 1993; McCombs & Shaw, 1993; Scheufele, 2000; Scheufele and Tewksbury, 2007).
- (see Kim, Kim and Zhou (2018) for an excellent overview).

## 3.3. Agendas and Social Media

---

- When we talk about media bias and agenda-setting nowadays, what comes in mind is the “fake news” concept. Fake news enter into the issue aspect of the agenda-setting theory.
- The dynamics of **fake news** and **issue framing** on social media are quite interesting (Guess, Nagler, et Tucker 2019), notably to look at the notion of influence (Alizadeh et al. 2020), but our article is more about the agenda-side.
- With the Internet, the first stage was for news media outlets to create a website and push their conversations on the newly appeared social media. Rapidly, social media were taken over by individuals and groups. A stage 2 was to push content generated by individuals to social media. People would decide the story to push and would build up their case by using other sources

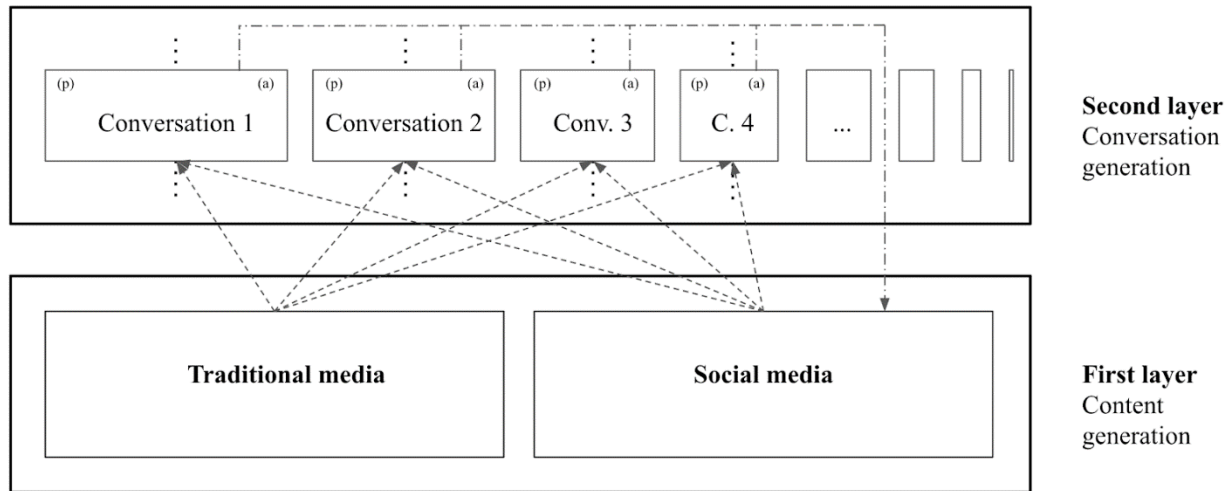


# 3.3. Agendas and Social Media (cont'd)

Dichotomy specific to social media (e.g. Twitter)

(p): passive participation (exogenous content) -----

(a): active participation (endogenous content) - - - - -



- **Traditional media:**

- Almost all newspapers, TV channels, radio stations and magazines publish their relevant content to Twitter.

- **Crowdsourced information:**

- Information generated by people with topics that may not have been covered (yet) by traditional media

## 3.4. Twitter

- **Use of Twitter in Politics through Content Generation**
- Twitter has become an important medium for debating about politics, organizing collective action, and showing support for, or critique of politicians and political issues.
  - Prominent campaigning tool
  - A political communication space
  - Opportunity for other Twitter users to contact them
  - Marketing stunt, to generate positive press
  - Interaction and socialization platform.

# Twitter (Cont'd )

- Increasingly, researchers turn to digital trace data in the analysis of social phenomena. We can group these approaches into two categories.
- First category: Twitter is treated as a sensor in documenting the reactions of users to their direct or mediated experiences through data traces produced by interactions of users with the service (Jungherr, Schoen, et Jürgens 2015; Shamma et al. 2010).
- Second category: Researchers go even as far as to draw inferences on attitudes, affiliations, and opinions of Twitter users based on the data traces of their behavior on the platform (Barberá et Steinert-Threlkeld 2020).



# 4. Methodology

- **Deconstruction of Tweets: The endogenous vs. the exogenous dimension**
- The content generation layer (the first layer) is first composed of exogenous information coming from traditional and social media.
- Then, users generate the conversations in both an active (editorialized) and passive way (retweets, for instance).
- As such, the editorialized content will feed the first layer in our taxonomy, as it corresponds to some new - though related - content generation.
- This pattern illustrates the endogenous dimension on Twitter.



# Methodology

## Content analysis: Natural Language Processing

- **Feature Selection Strategy in Text Classification:**
  - We formulate the feature selection process as a dual objective optimization problem and identify the best number of features for each document
  - We provide a documented framework for conducting text preprocessing in text classification in order to optimize the classifier performances
- **Refining Content Analysis**
  - We estimate four attributes (gender, age, occupation, and interests) of a Twitter user from the contents (a profile document and tweets) generated by the user and the user's social neighbors

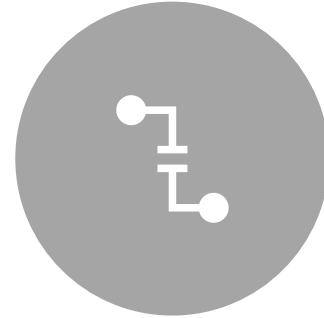
# Top Hashtags

- European-migrantcrisis
- migrants
- refugees
- refugeeswelcome
- refugeecrisis
- réfugiés
- MigrationEU
- migrazione
- rifugiati
- migranti
- fluchtlinge
- asylmissbrauch
- refugiados
- norefugees

# Sentiment and Polarization Analysis



## Computer-assisted linguistic analysis



We describe the most common tools for pre-processing textual data, including stop word removal, stemming, lemmatization, compounding, decompounding, and segmentation.



In each case, the goal is to reduce the scale of the problem by treating words with very similar properties identically and removing words that are unnecessary to our interpretation and our model.



Along with disregarding word order, the so-called “bag-of-words” assumption, these procedures are common preprocessing steps but can differ across languages.

# Sentiment and Polarization Analysis: Pre-processing Textual Data

---

Common tools	Examples
Stop word removal	<del>“and”</del> and <del>“the”</del>
Stemming	accounts -> account accounting -> account accountants -> account
Lemmatization	The word "better" has "good" as its lemma
(De)compounding	The German word “Kirche,” or church, can be appended to “rat,” forming “Kirchenrat,” one who is a member of the church council, or “pfleger” to form “Kirchenpfleger,” or church warden.
Segmentation	dictionary or statistical methods
Linguistic inquiry and word count (LIWC)	Words have specific psychological meanings (e.g., Weintraub 1989)



# 5. Results

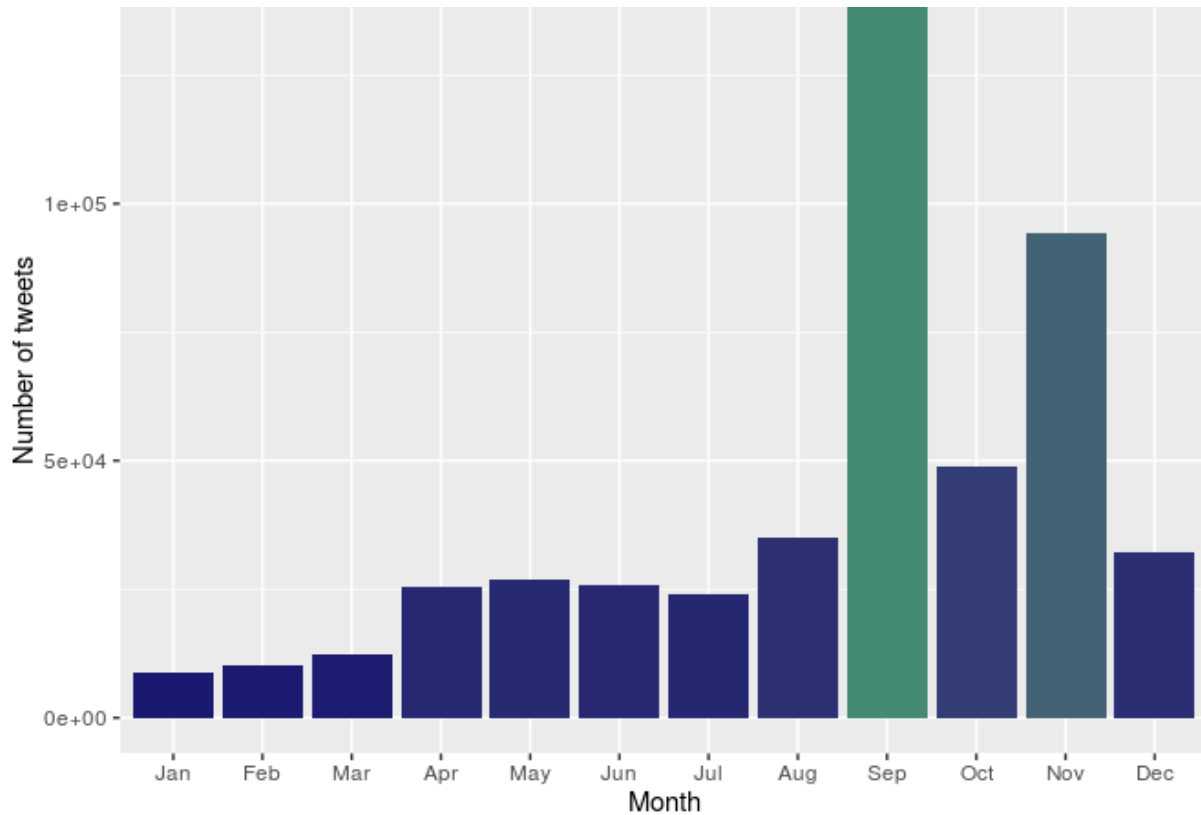
---



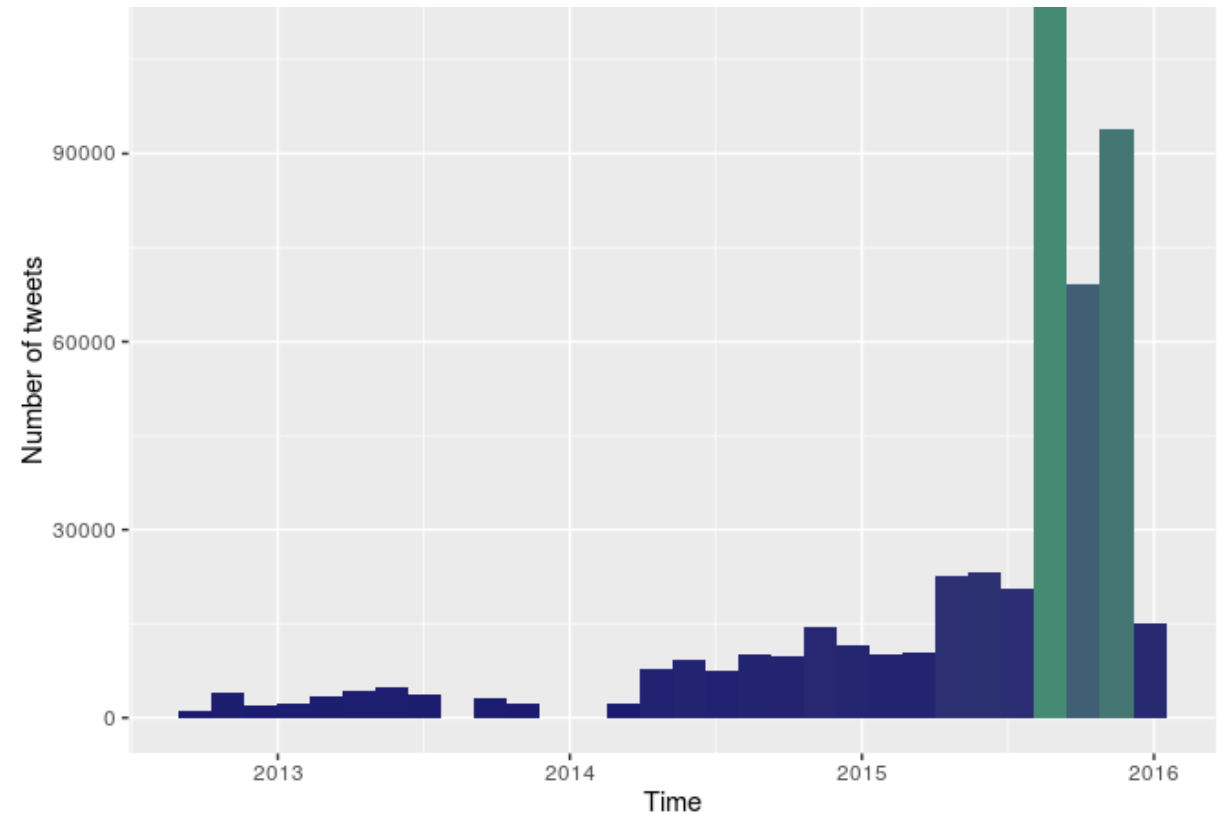
- We collected tweets and their metadata (latitude, longitude, retweets, hashtags, etc.) over a period of 3 years.
- We collected a total of 482,869 messages from September 9<sup>th</sup>, 2012 to December 16<sup>th</sup>, 2015.
- The news reporting seems to have combined empathetic responses in the early stages of the 'crisis', with suspicion or hostility, especially after the terror attacks in Paris on November 13<sup>th</sup>, 2015.

# Distribution of Tweets

Overall distribution of tweets

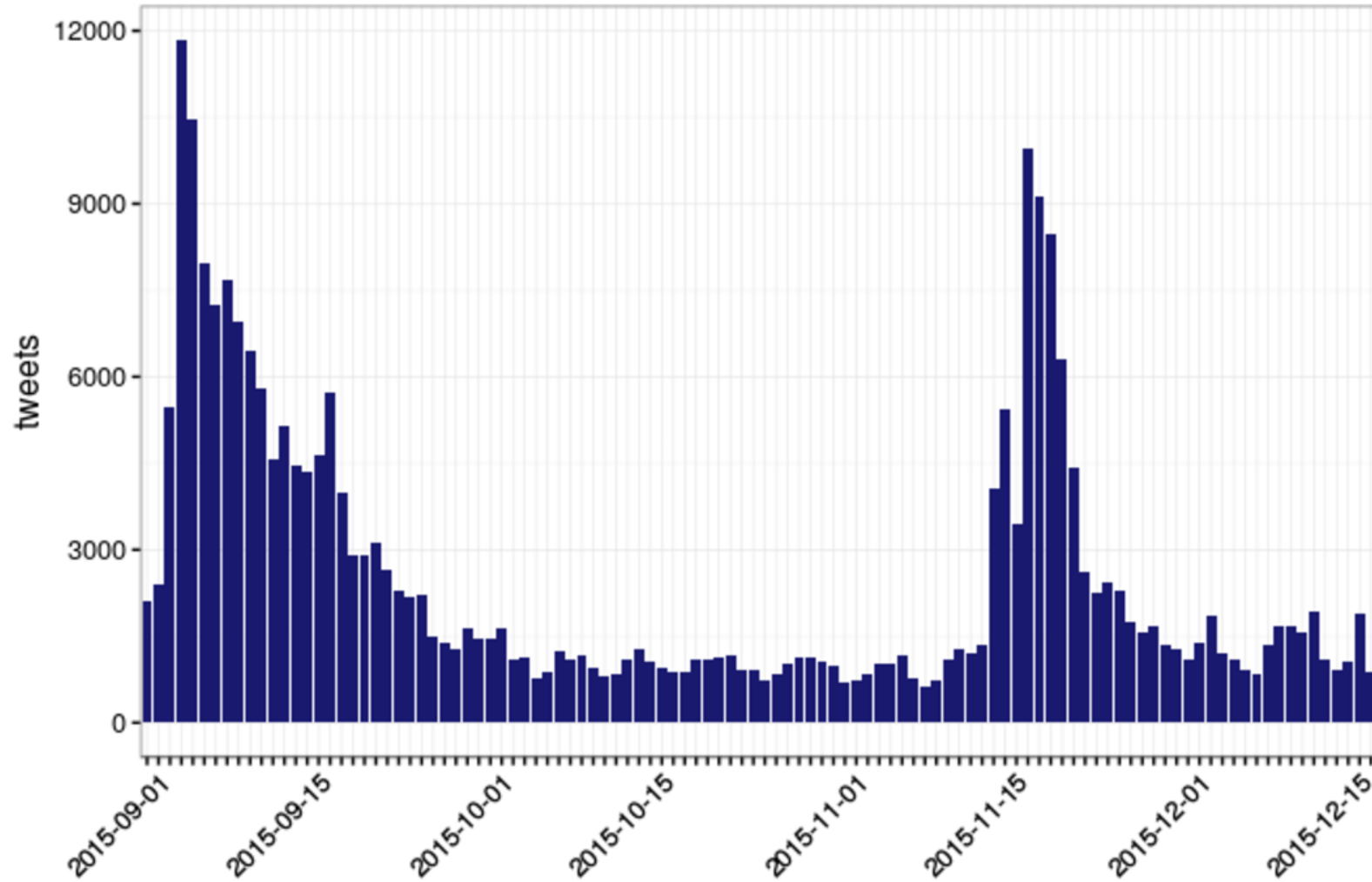


Distribution of tweets by month



# Results

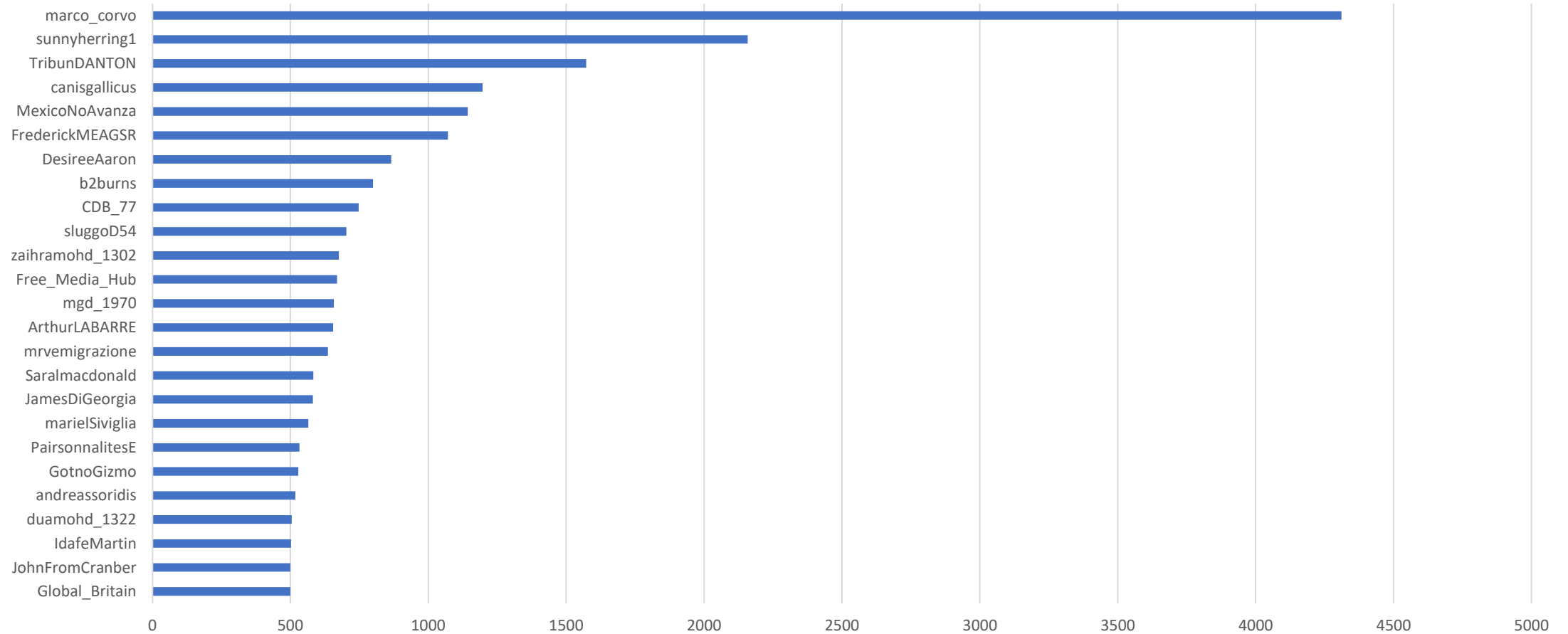
Number of Tweets by Date



Day with the highest amount of tweets:  
September 3rd,  
2015 with 11,842

# Results:

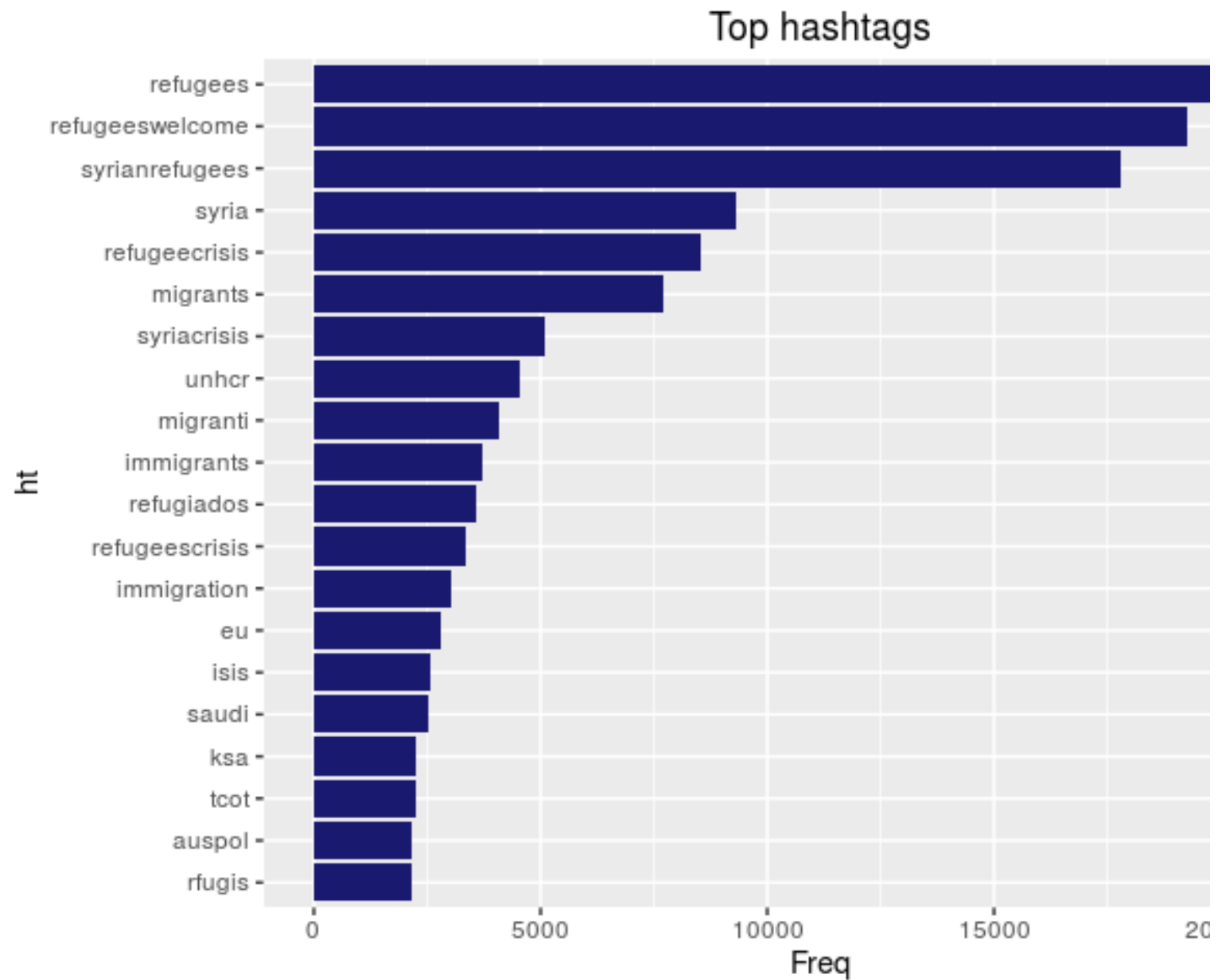
## Top users



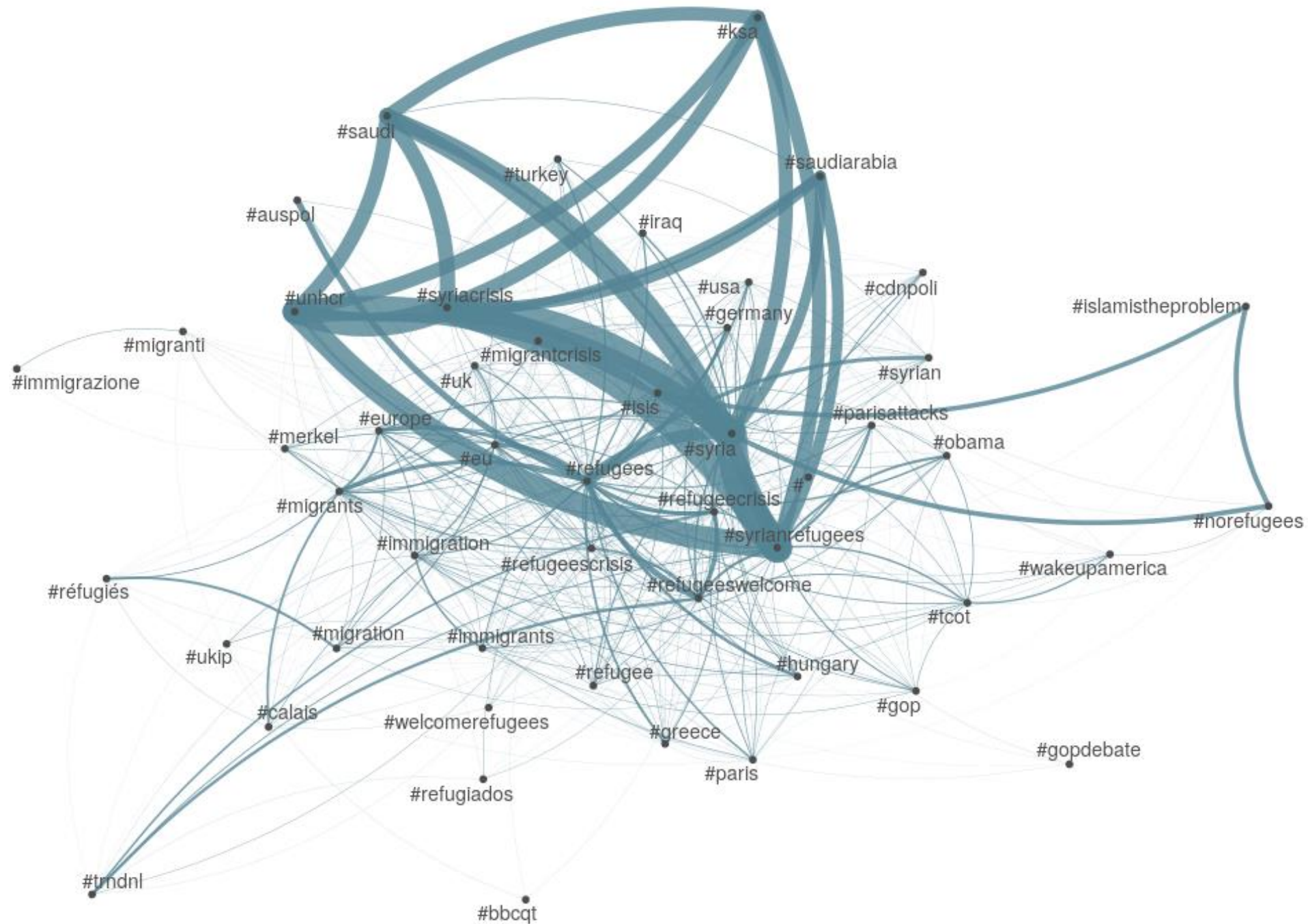
# Traditional media: Newspapers

France	Germany	Italy	Spain	United Kingdom	United States	Other
@lemondefr	@SPIEGELONLINE	@SkyTG24	@muyinteresante	@BBCBreaking	@washingtonpost	@AJELive
@TF1	@ProSieben	@repubblicait	@elmundoes	@BBCWorld	@cnnbrk	@AP
@M6	@tagesschau	@Corriere	@voguespain	@Reuters	@WSJbreakingnews	@breakingnews
@FRANCE24	@zeitonline	@fattoquotidiano	@diarioas	@BBCNews	@CBSTopNews	
@canalplus	@BILD	@sole24ore	@actualidadrt	@guardian	@ABCNewsLive	
@lefigaro	@SPIEGEL_EIL	@Internationale	@el_pais	@FinancialTimes	@nytimes	
	@SZ	@ilmessaggero		@SkyNews	@WSJ	
	@sternde			@SkyNewsBreak	@FoxNews	
	@welt			@ReutersLive	@ABC	

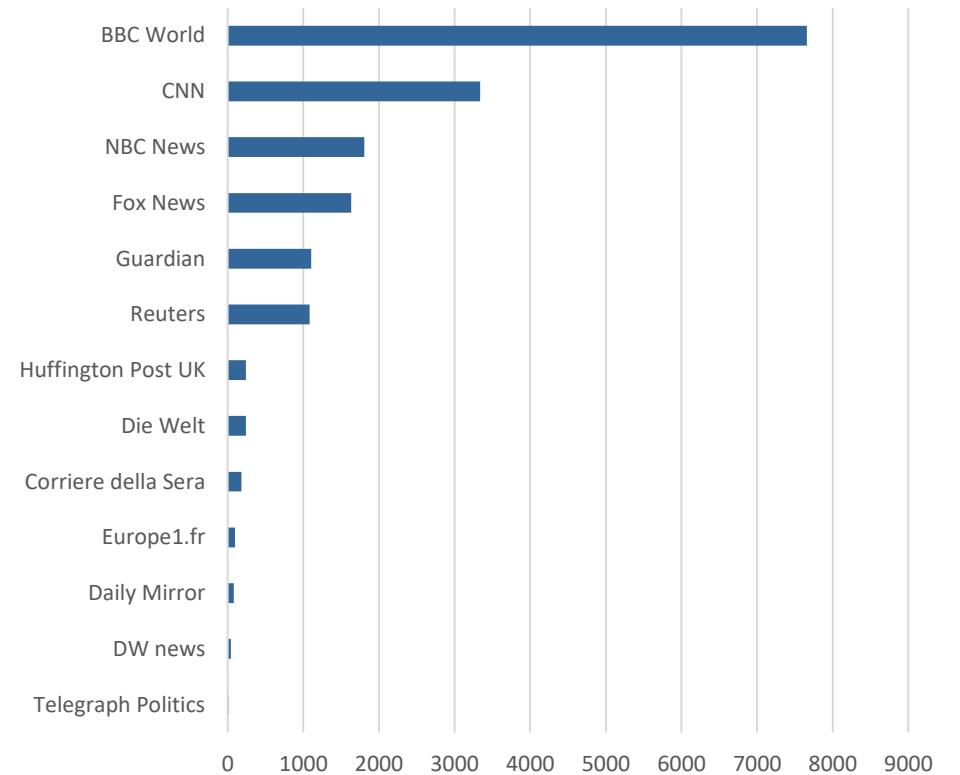
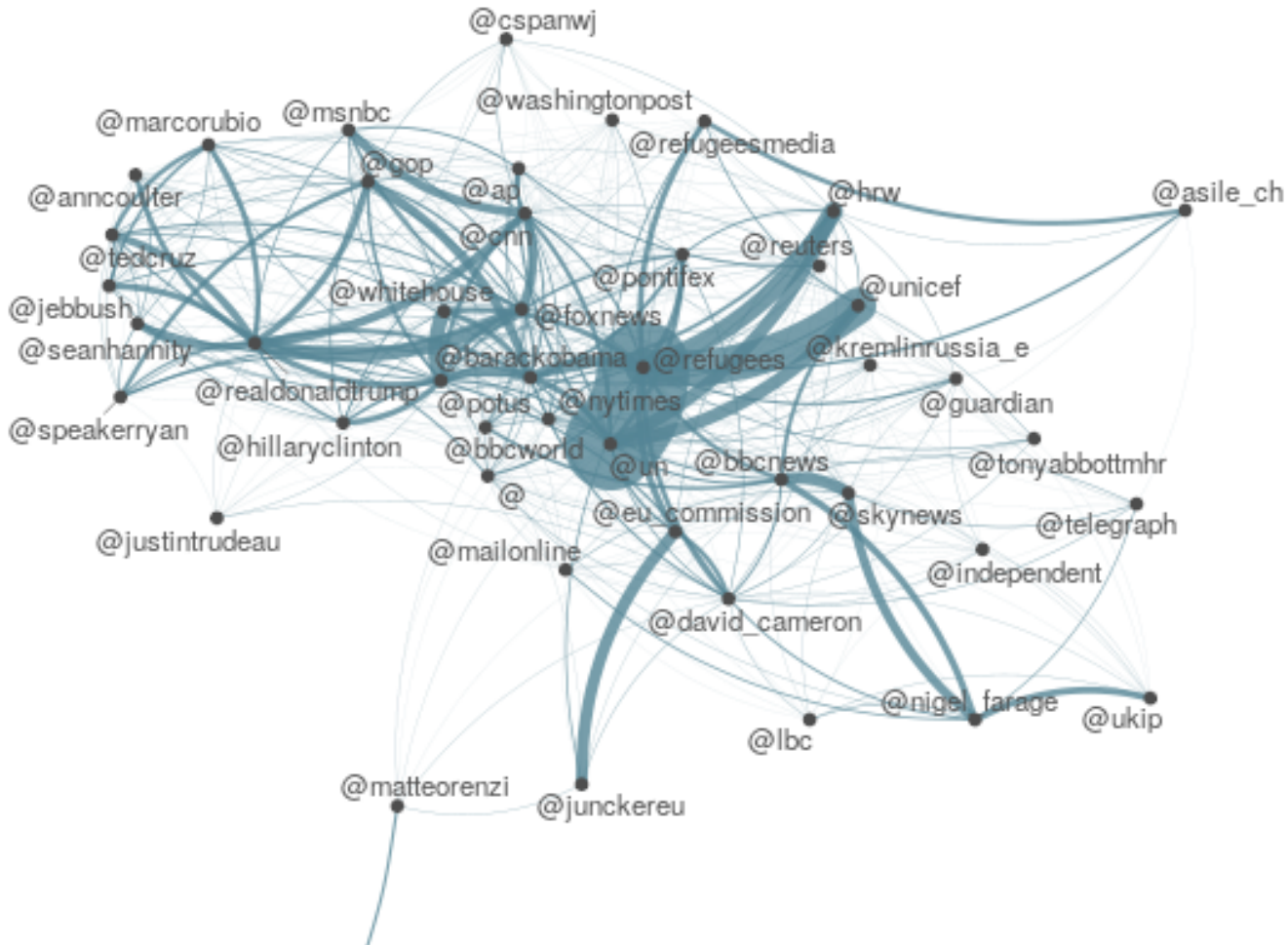
# Results: Top hashtags



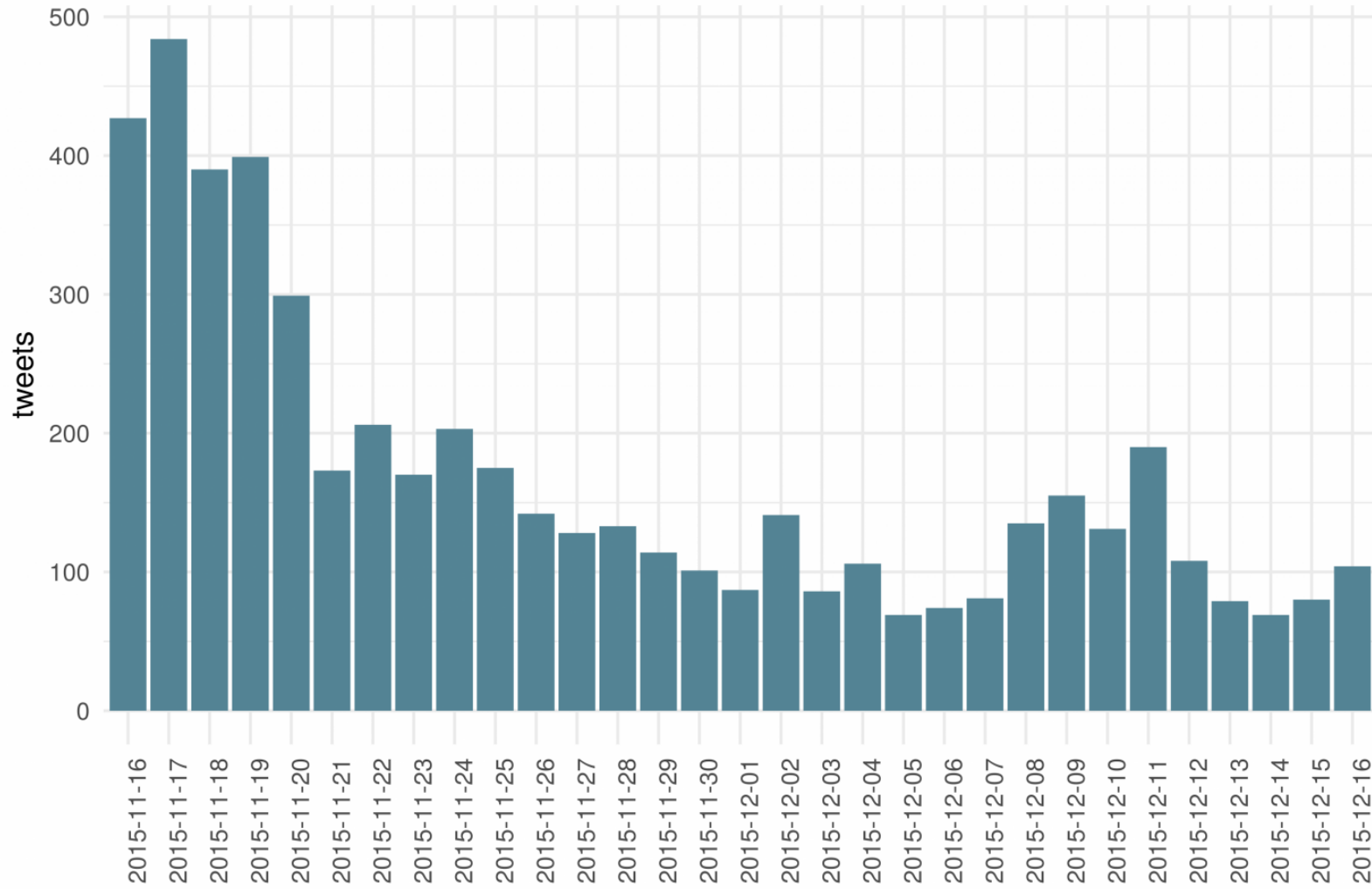
# The Hashtags Network



# Top User Network







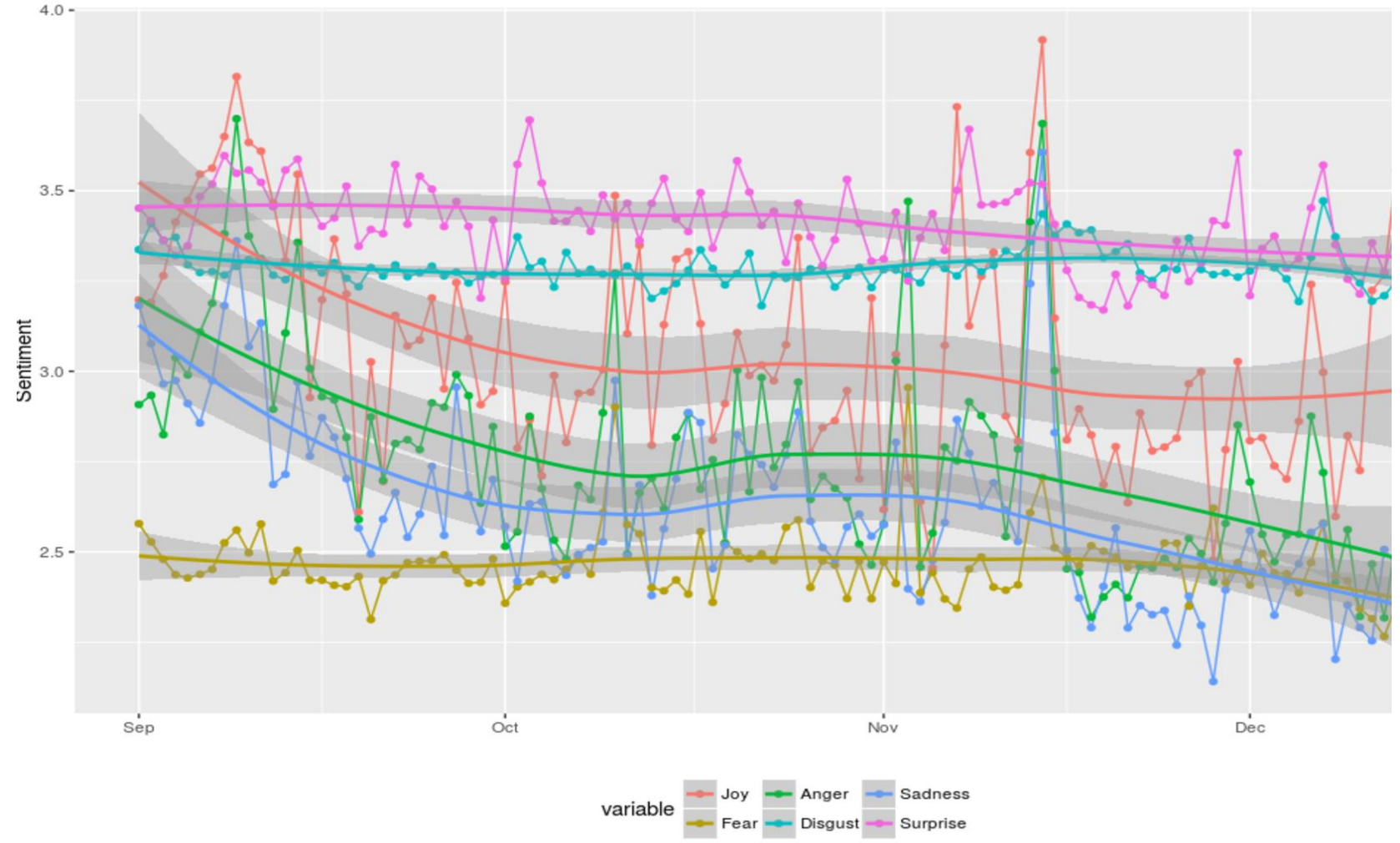
A spike after the terror attacks in Paris on November 13th, 2015

# Some Tweets

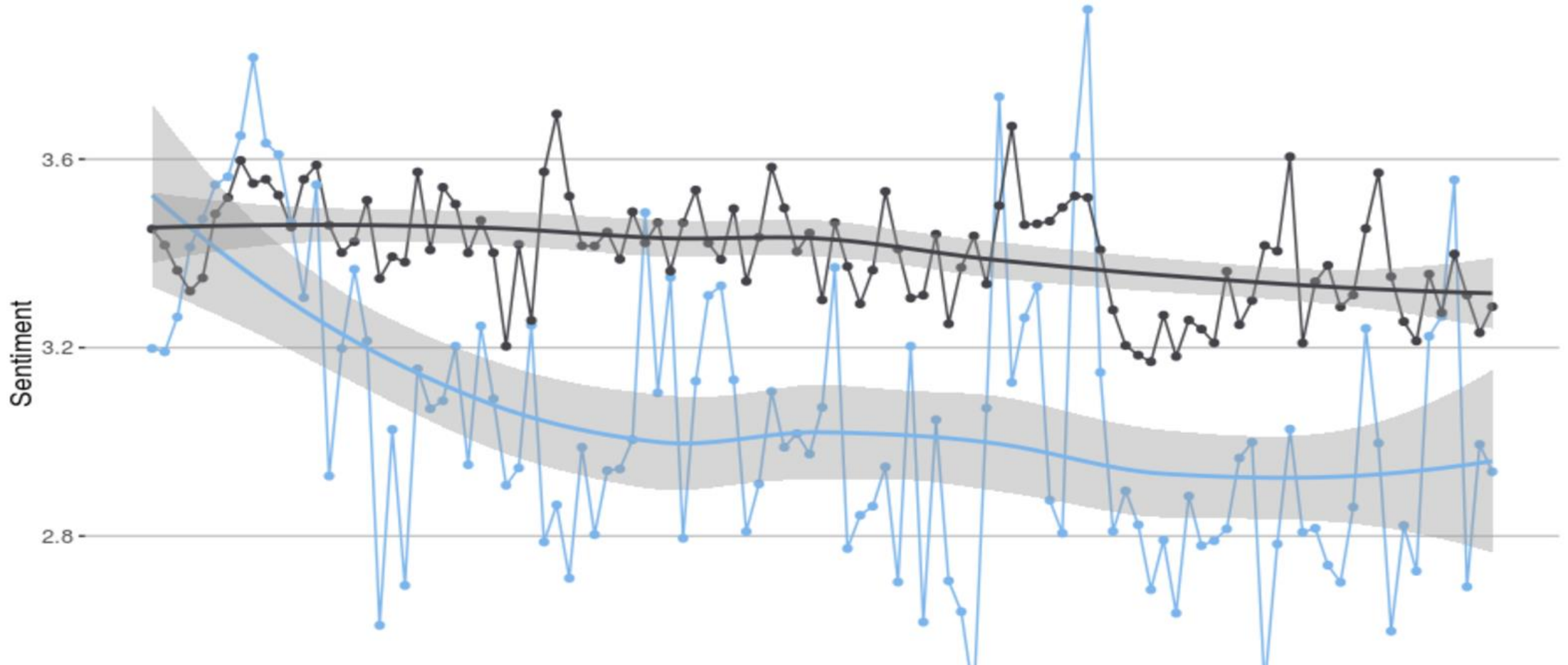
- “I know those on the left will regard me as a racist; however, immigrants are posing a real problem here in Britain, and we must curtail it”
- “I also wonder whether we can correlate the rise in crime to the number of immigrants here in Britain ??”
- “Christian fairytales aren't perceived truthfully. refugees want human freedom, not forced imprisonment! #qanda #god”

# Sentiment Analysis

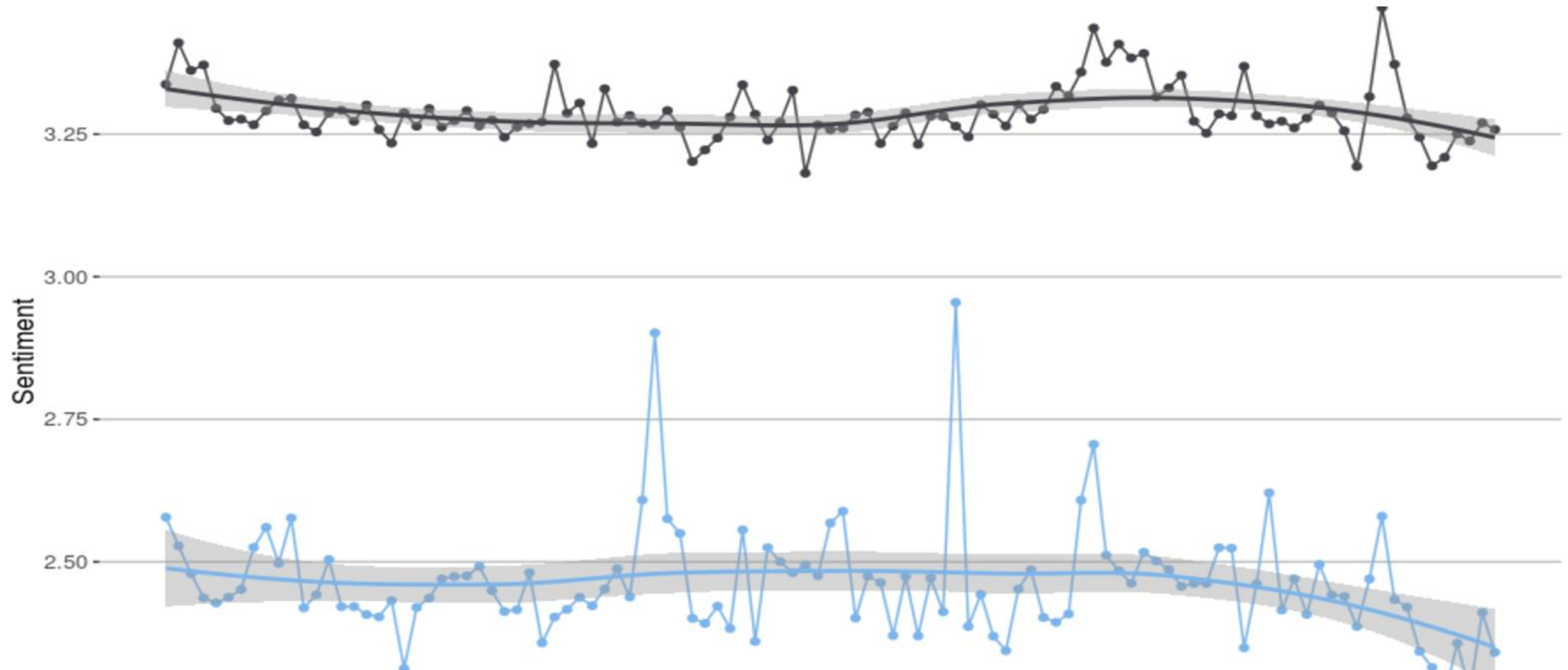
- Joy
- Fear
- Anger
- Disgust
- Sadness
- Surprise



# Sentiment analysis: Joy + Surprise

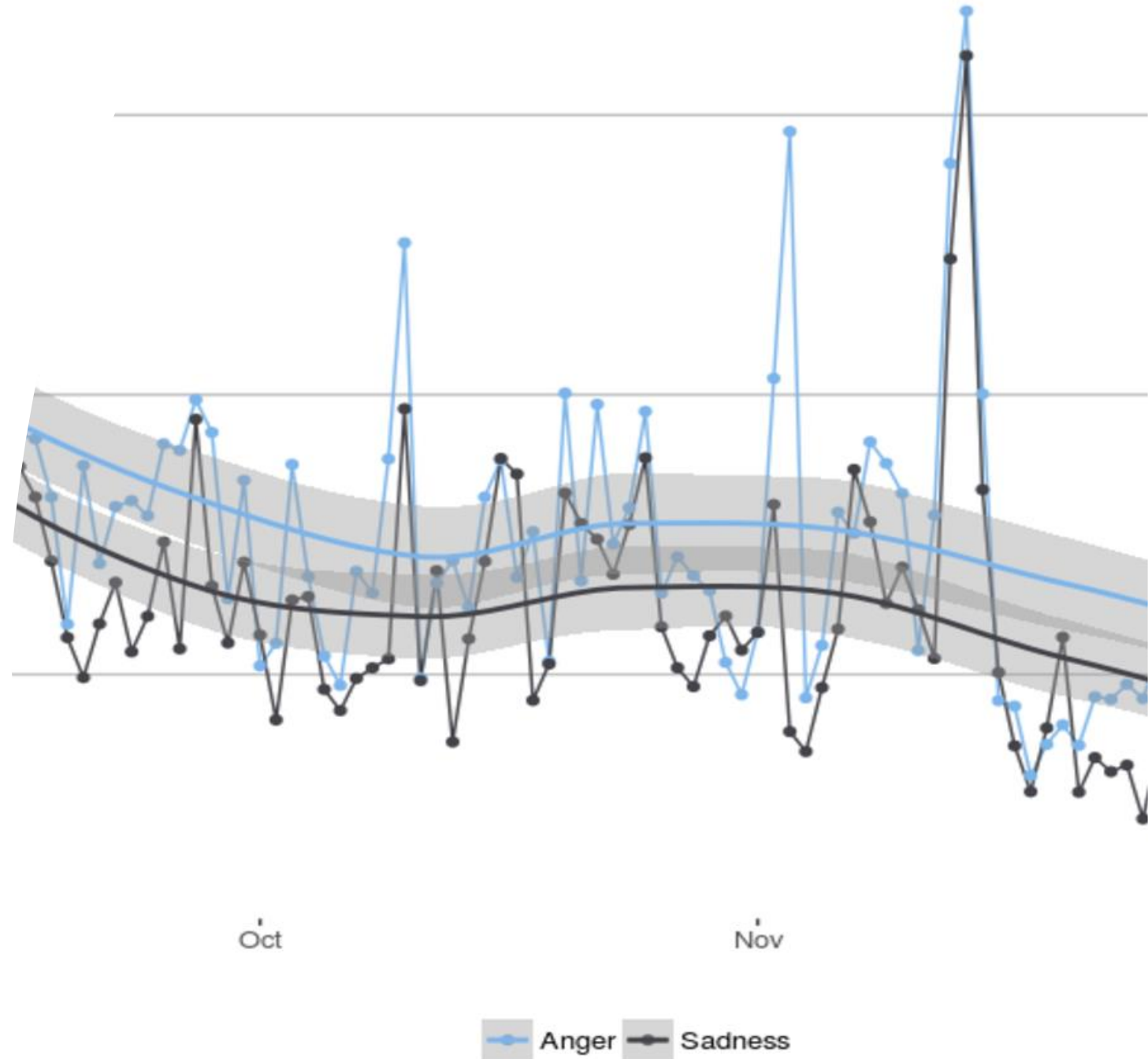


# Sentiment analysis: Disgust plus fear



Sentiment analysis:  
Sadness plus  
fear

---



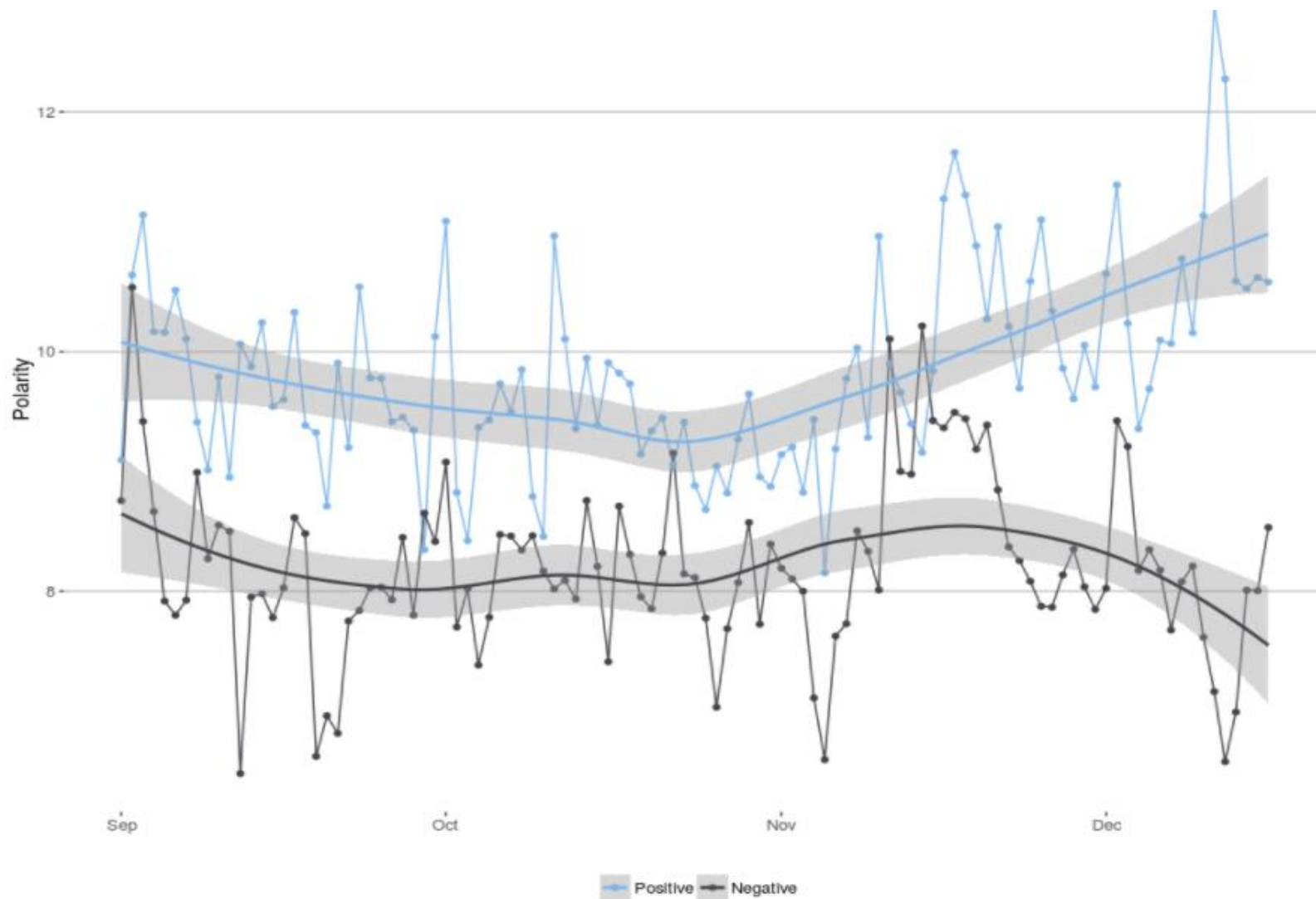
# Results:

## Polarity Analysis

---

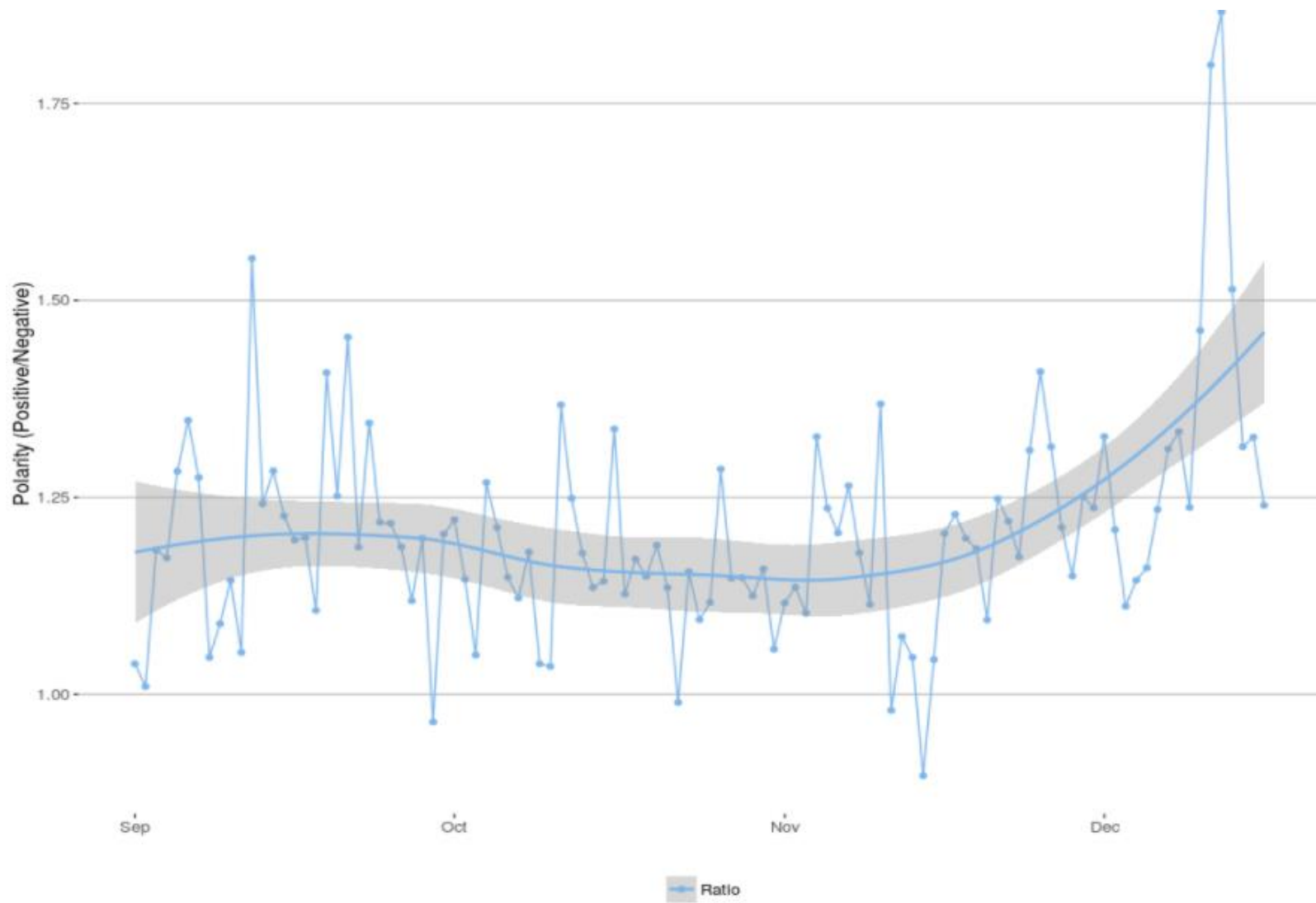
- The second step of our sentiment detection approach is **polarity classification**, i.e., predicting positive or negative sentiment on subjective tweets.
- In this section, first we analyze the quality of the polarity labels provided by the three sources, and whether their combination has the potential to bring improvement.
- Two questions we investigate regarding these sources are: (1) how useful are these polarity labels? and (2) does combining them bring improvement in accuracy?
- We take the following aspects into consideration:
  - Labeler quality
  - Number of labels provided by the labelers
  - Labeler bias
  - Different labeler bias





Polarity  
Analysis:  
**Positive +  
Negative**





Polarity  
Analysis:  
**Ratio**

# Conclusion

---

- The conversations are mostly from individuals in volume, but **at the core of the network, we find traditional news media.**
- Agenda-setting theory has had an impressive theoretical and methodological expansion as well as diverse applicability in various fields such as politics, business, and culture.
- The next step in the development of this paper is about creating a dataset of sentiment and polarity indices and test whether they could have some exploratory power for certain general election results.