Influence of Police Violence Against Unarmed Black Victims on Twitter Conversation

Data Science Institute
COLUMBIA UNIVERSITY

Shadi Fadaee, Shimeng Feng, Jose Lopez, Andrea Navarrete, Omar Pardo
Colin Wayne Leach & Courtney Cogburn

Data Science Capstone Project With Barnard College & School of Social Work - Columbia University

1. Introduction

The goal of this project is to understand Twitter activity before and after the events related to police use of force against unarmed black victims. We approach the problem by applying topic modeling and sentiment analysis techniques to understand the development of themes and their corresponding emotions. We have focused on over 8.5 million tweets from August 2014 to cover the event of the police officer (Darren Wilson) shooting a black victim (Michael Brown) that happened in Ferguson, MO.

Figure 1. Timeline of events for August, 2014

2. Topic Modeling

We used the Mallet - LDA model to cluster tweets into different themes to understand the topics. The model with the maximum coherence grouped the data into 12 topics. The Word Cloud on Fig. 2 contains the most relevant terms for each of the 4 selected topics. Fig. 3 shows the evolution of topics throughout August, how they drove the conversation, and how they were affected by different events.
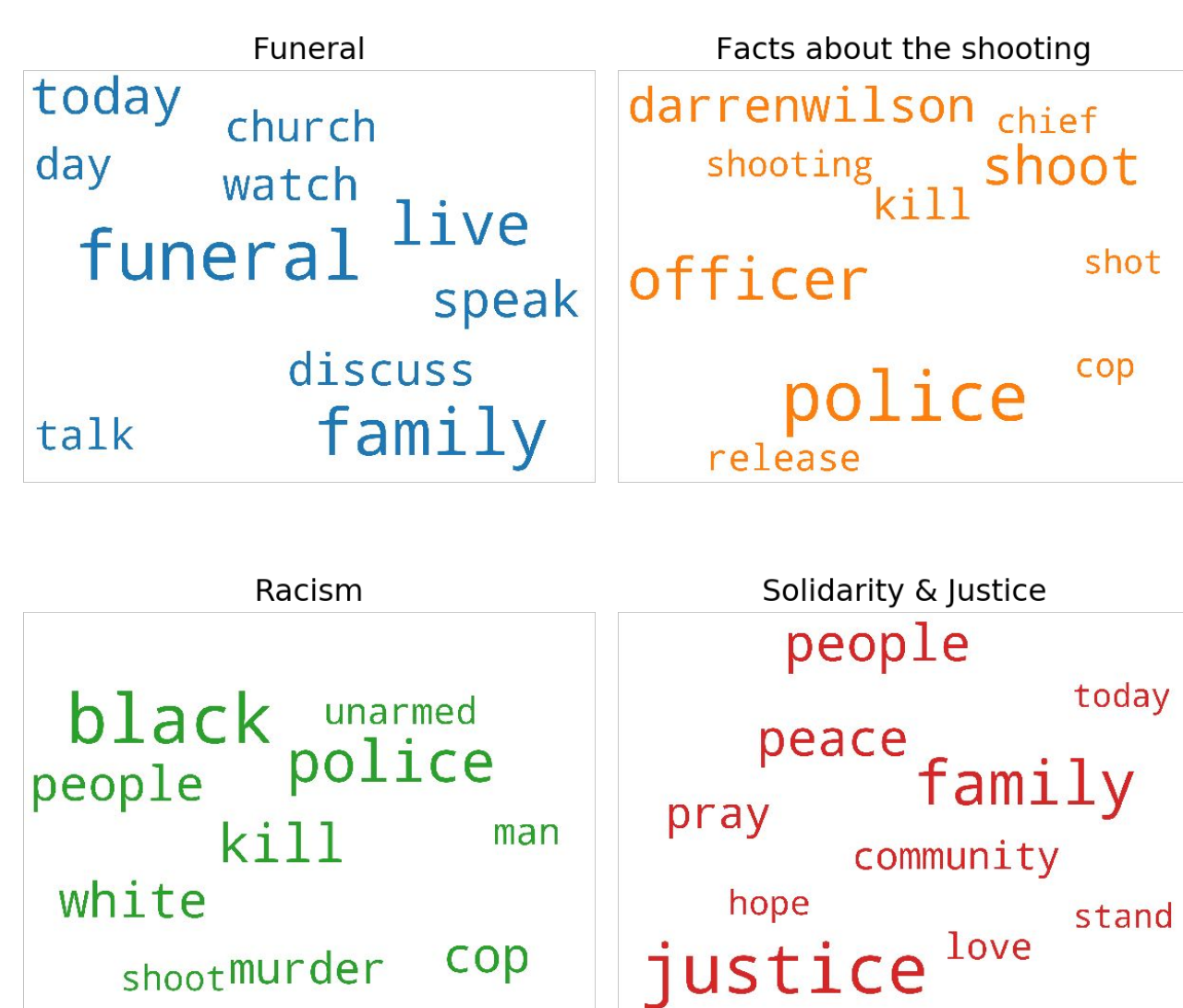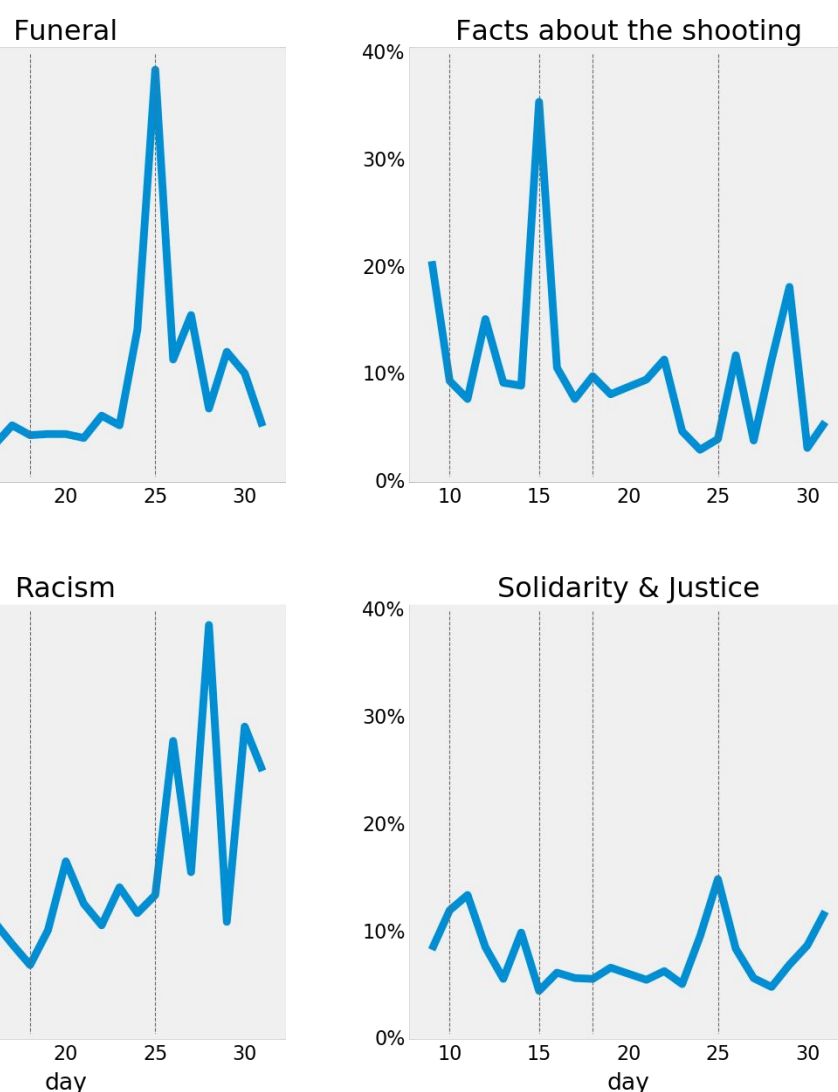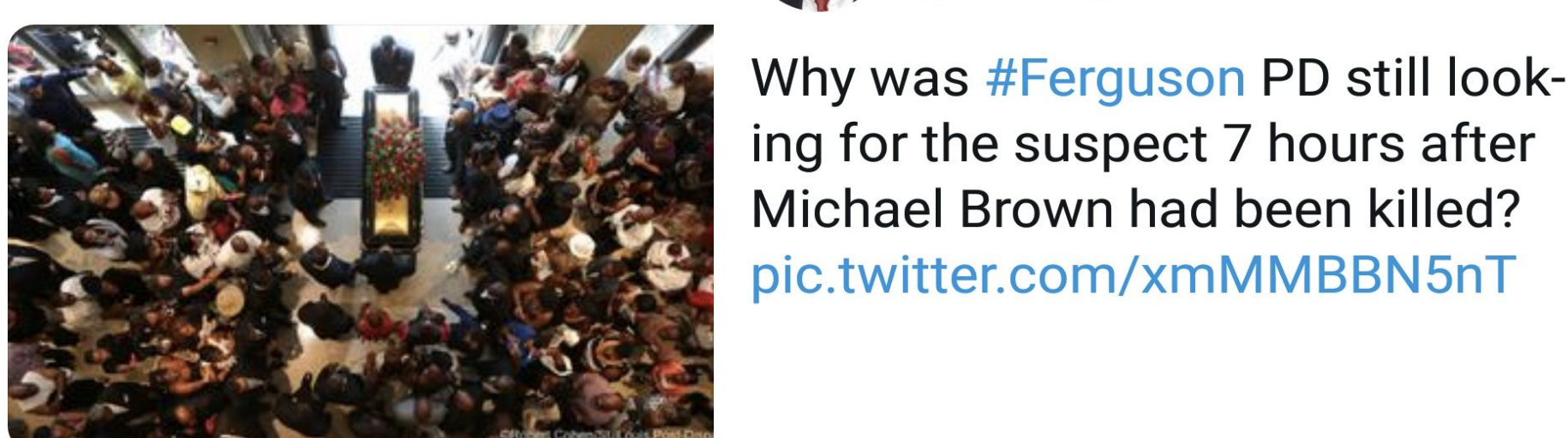
Figure 2. Word Cloud of tweet topics

Figure 3. Time series of tweet topics

Figure 4. Sample tweets per topic

3. Sentiment Analysis

We applied the Vader scoring approach to discover the emotions and the emotional intensity score for each tweet. Compound scores represent a unidimensional measure of sentiment for a tweet. It is a normalized, weighted composite score.

The trend of the compound score in Fig. 5 reveals that sentiment starts very negative and evolves to more positive scores. The compound score in Fig. 6 shows users relate the 'Solidarity & Justice' topic with a Positive sentiment, while 'Racism' and 'Facts about shooting' relate to a negative one. 'Funeral' has a neutral score as it describes a fact.
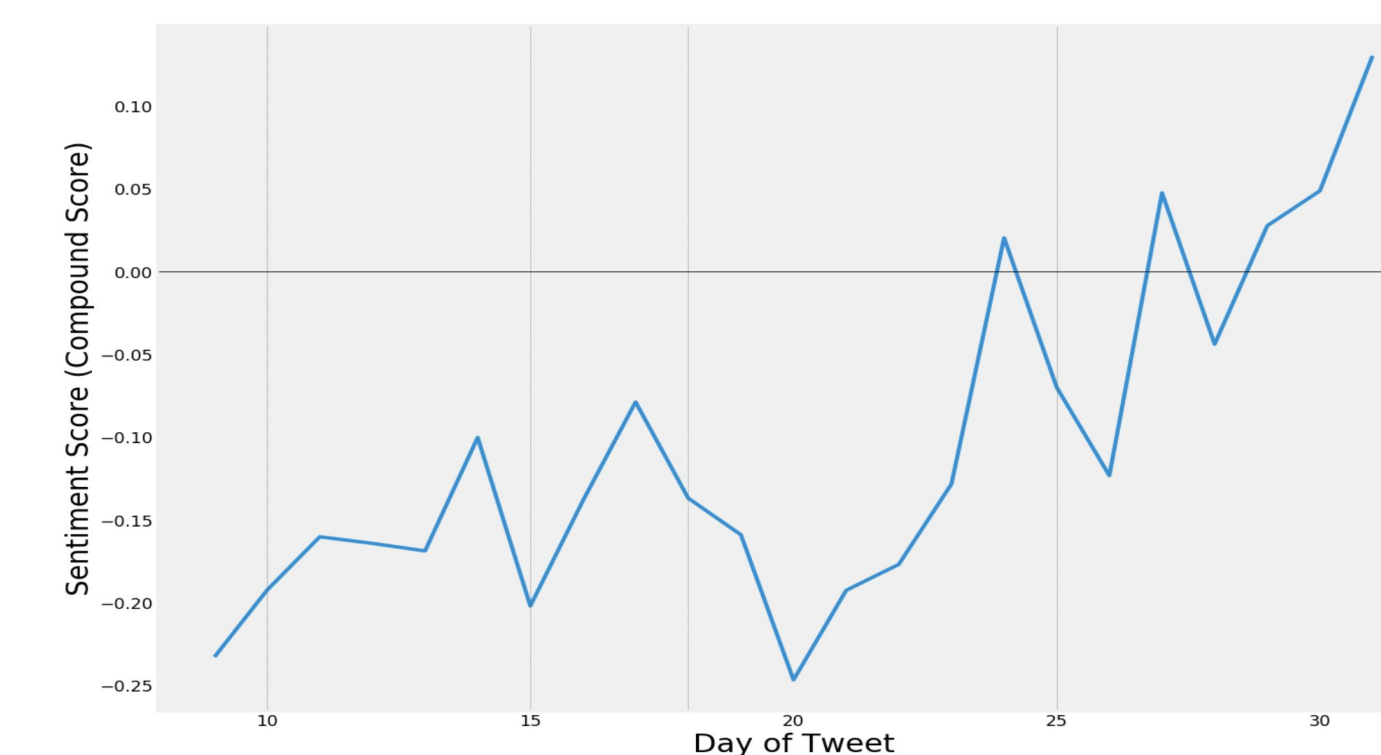
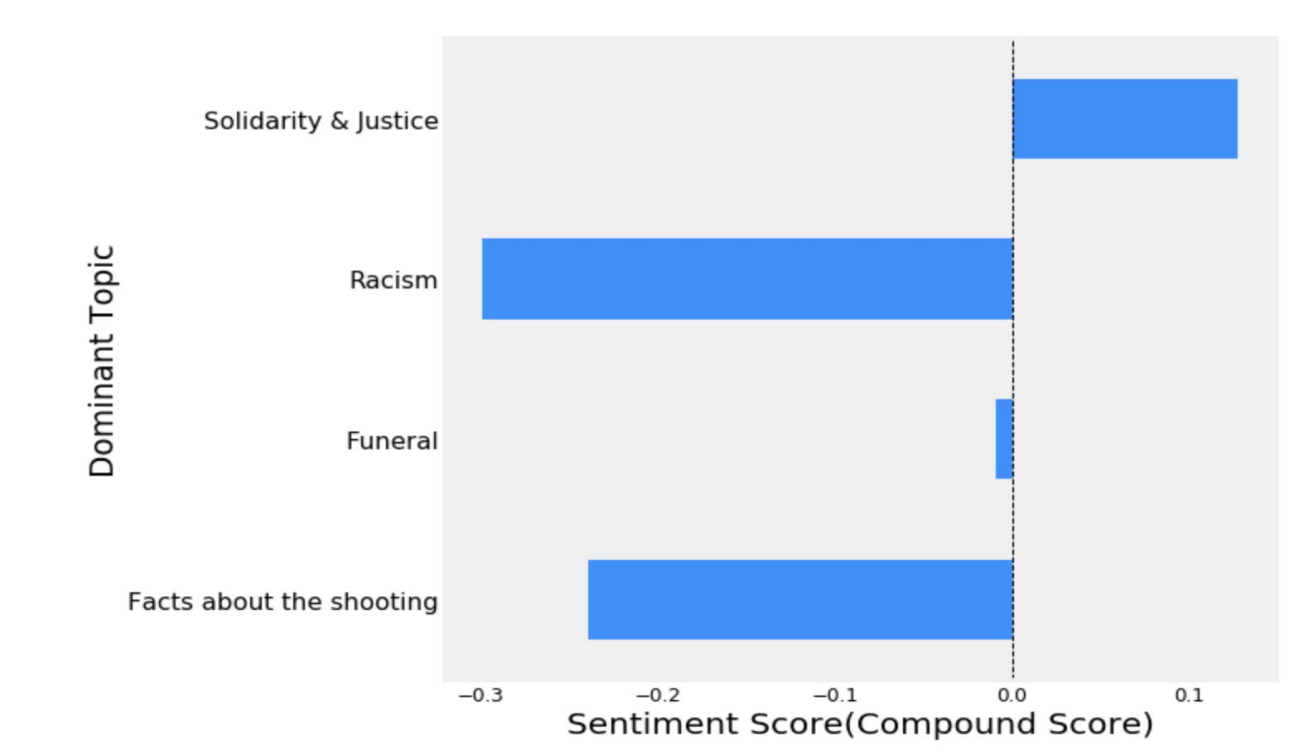**Figure 5. Sentiment Score of Tweets Over Time (Compound Score)**

Figure 6. Positive & Negative emotions in topics (Compound Score)

4. LIWC categories across topics

We used 41 categories around psychological constructs (e.g., affect, cognition, drives) and iterated for each tweet to find the list of words (word stems) of every LIWC category by topic and time. Some of the categories by topic over time are shown in Fig. 7 and 8.
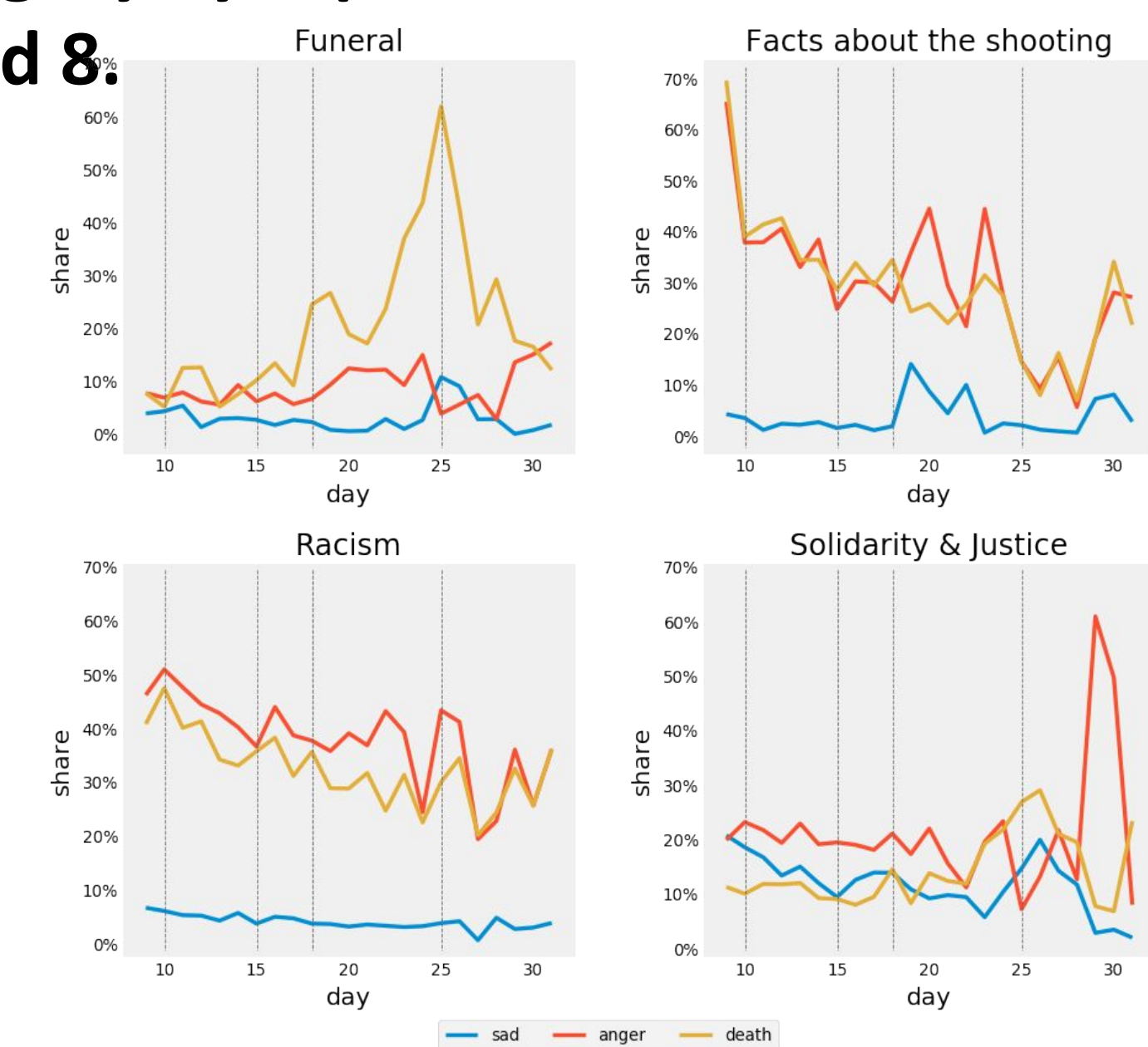
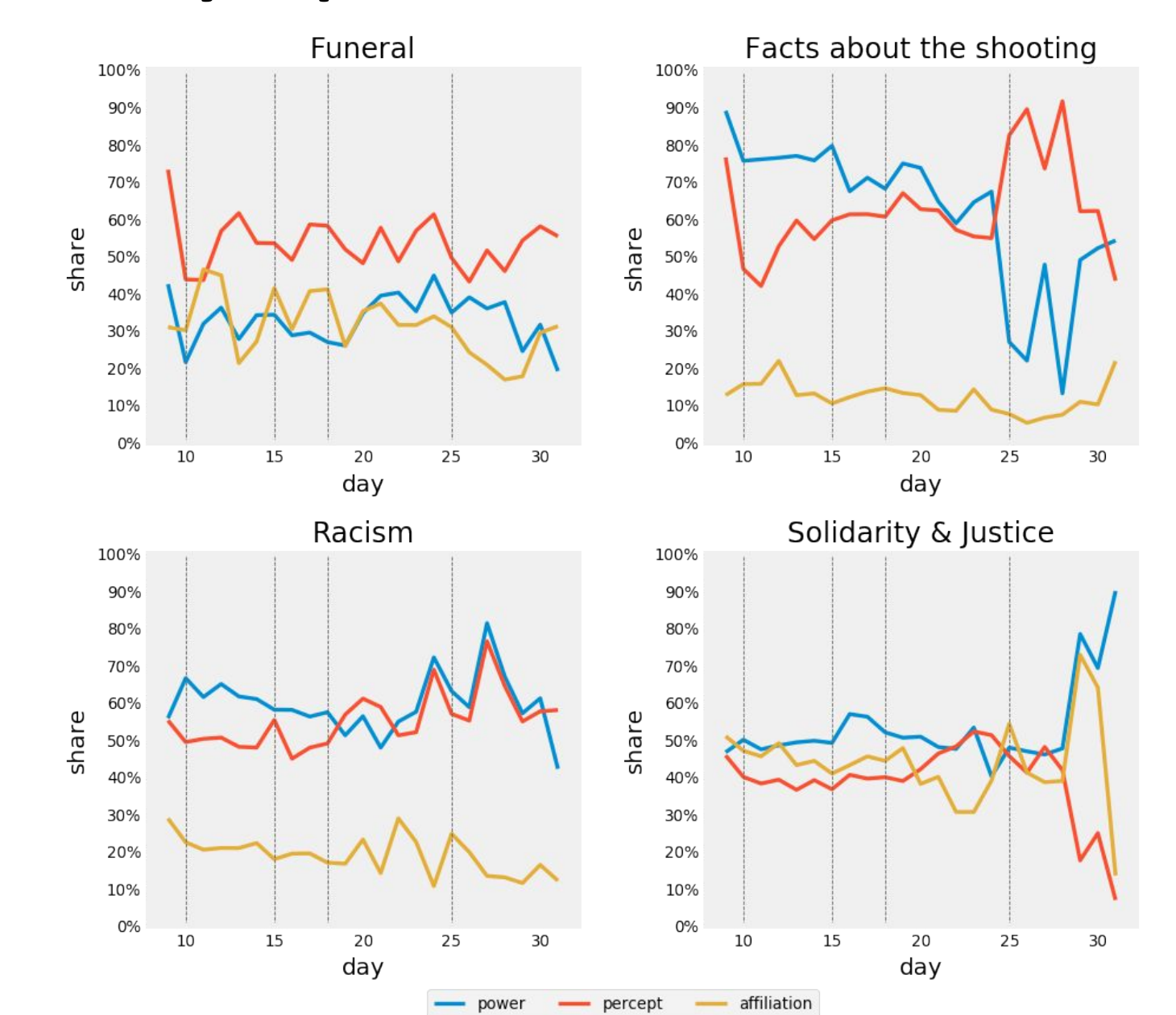**Figure 7. Progression of emotions by topic**

Figure 8. Progression of drives by topic

Figure 7 shows that the 'Funeral' topic has an increase of Death words around Michael Brown's funeral, while Anger words increased for the 'Solidarity' Topic after this event. In a similar way, Fig. 8 shows how words related to Power, Percept and Affiliation change after the funeral.

5. Conclusion

This study demonstrates that there are lags between events and their emotional response in Twitter. This might be due to the fact that users take time to absorb, process and respond to situations through social media. One of the most important episodes in this study is the Funeral of Michael Brown. No significant emotion was observed during the event, but the spike after the funeral shows how the community was shaken.

References

1. Freelon, D.,et al. (2016). Beyond the Hashtags: #Ferguson, #Blacklivesmatter, and the Online Struggle for Offline Justice. SSRN Electronic Journal. doi: 10.2139/ssrn.2747066
2. Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.
3. Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015). Linguistic Inquiry and Word Count: LIWC2015.
4. Topic Modeling in Python with Gensim. (2018, December 4). Retrieved from https://www.machinelearningplus.com/nlp/