A STORM OF TWEETS

ANALYZING TWITTER CONVERSATIONS SURROUNDING HURRICANE IRENE

Erin Irons

Vermont EPSCoR

Summer 2015

Adviser: Dr. Steve Scheinert

Funding provided by NSF Grant EPS-1101317

RACC Research on Adaptation to Climate Change in the Lake Champlain Basin

Vermont EPSCoR CWDD Center for Workforce Development & Diversity AT SAINT MICHAEL'S COLLEGE

Vermont EPSCoR

NSF

Saint Michael's College Founded 1904

The University of Vermont

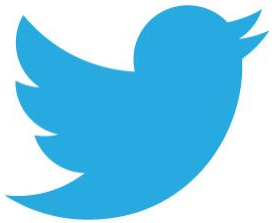
HURRICANE IRENE IN VERMONT

- Hit Vermont August 28th-29th, 2011
- Over 7 inches of rain in some areas
- Worst flooding in the state since 1927
- Approximately $175-250 million in damage
- 89% of Vermont towns had infrastructure damage
- Caused three deaths

HOW IS TWITTER UTILIZED?

- Social Media Platform
- 314 Million Users Worldwide
- 140 Character Messages
- Hashtag Connections
- 500 Million Tweets Posted Per Day

HOW IS THIS RELEVANT?

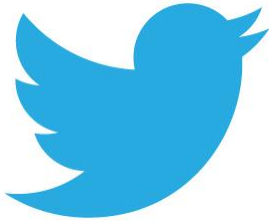


- Natural disasters, especially hurricanes, are worsening due to climate change
  - Irene was the result of this by devastating our local communities
- Twitter and other social media platforms are growing to be become important tools in information sharing during times of emergencies

RESEARCH QUESTIONS

- **How is information being shared during the time of Hurricane Irene in Vermont on Twitter?**
  - What are the sentiments of these tweets?
  - How are features used to share information?
    - What are the most predominant hashtags tweeted?
    - How many average hashtags are there per tweet?
    - Which feature is used the most?

HYPOTHESIS



- Sentiments:
  - Nervous/Concerned, Objective, and Skeptical will be the most characterized sentiment
- Use of Features to share Information
  - Most users will use at least two features per tweet
  - If hashtags are implemented- a combination of hashtags will be used

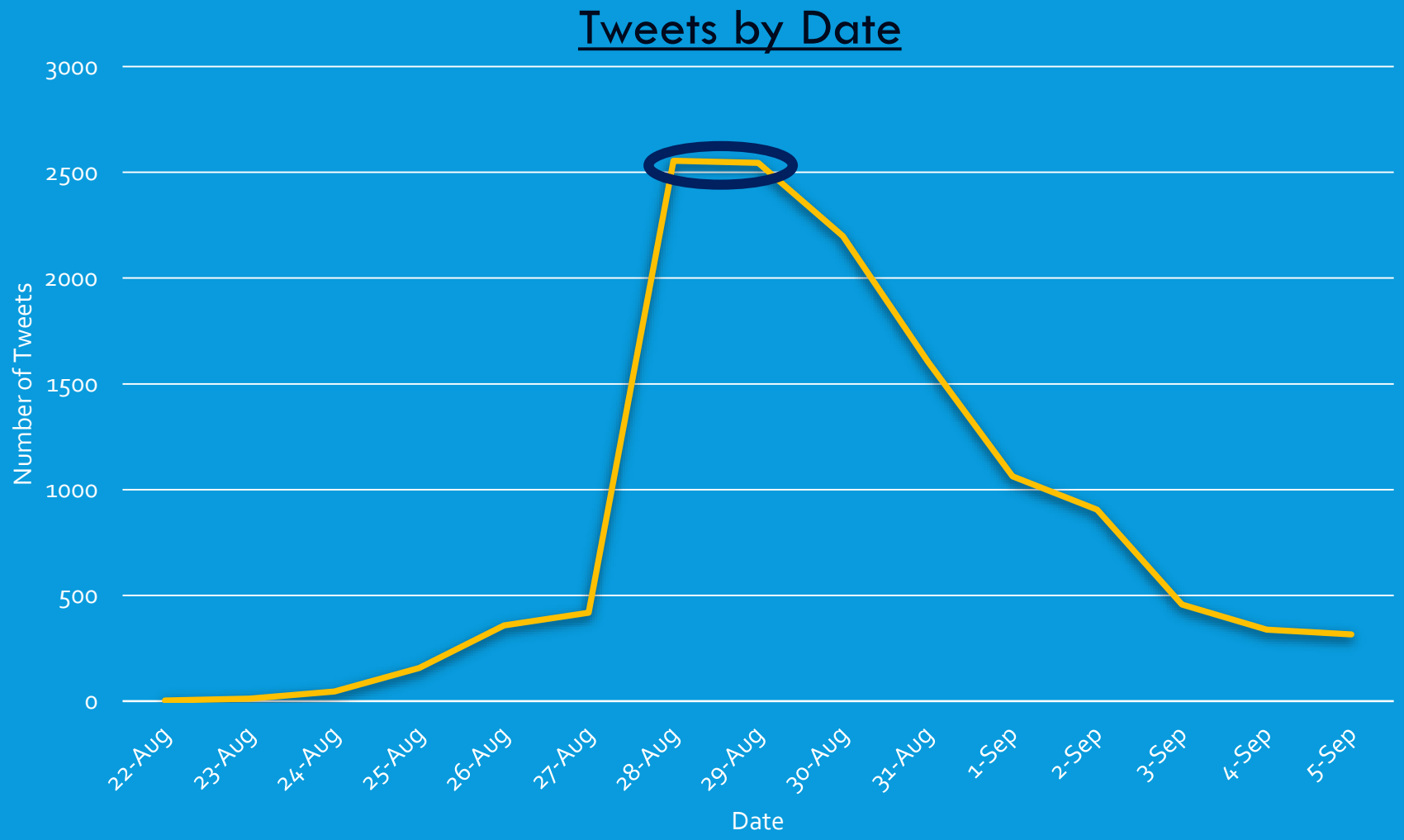
METHODS

- Literature Reviews
- Manually Creating a Tweet Database
  - Search by date
  - Search terms
    - (“Vermont” AND “Hurricane OR Irene”)
    - (“VT” AND “Hurricane OR Irene” NOT “Vermont”)
- Manually labeling sentiments to tweets, counting hashtags, links, and user handles
- Used Python code to create a database of hashtags
  - Used Google Drive’s WordCloud Generator to find out which hashtags appeared most

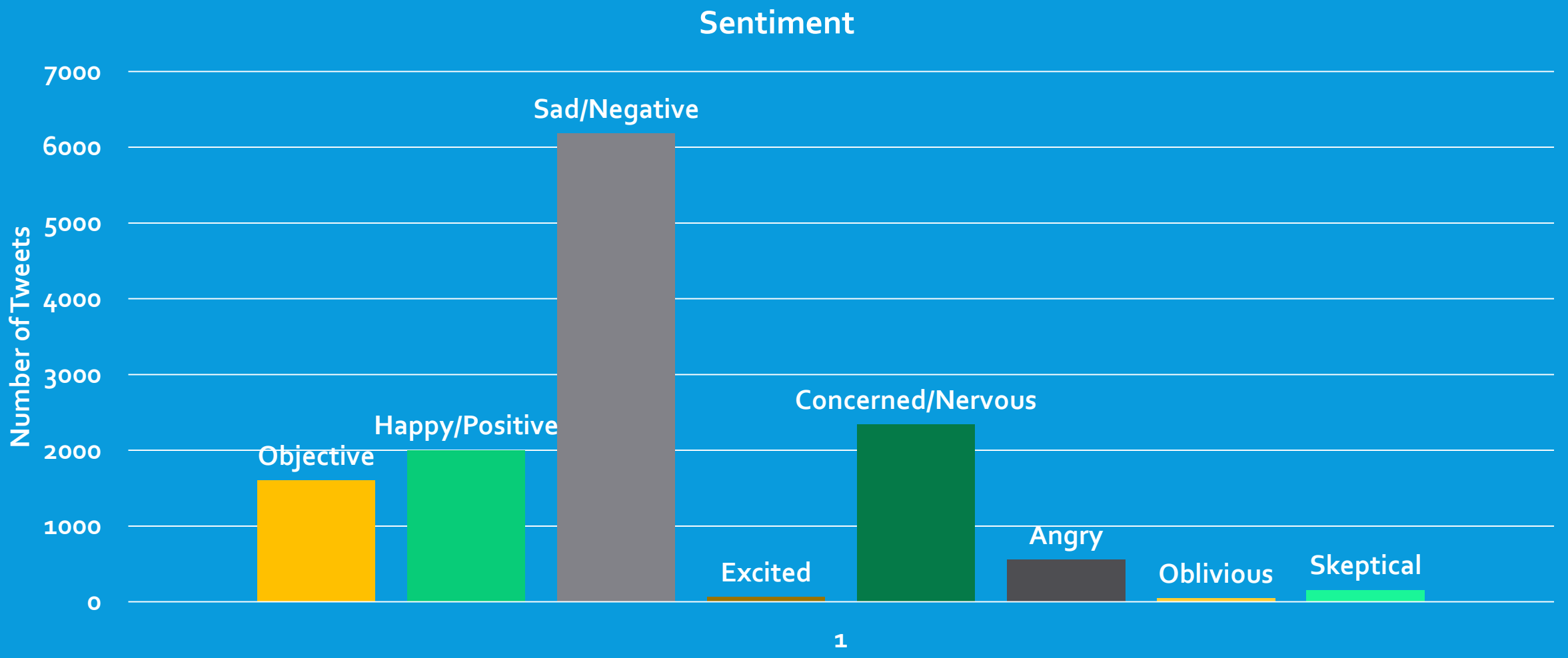
RESULTS

- 12,958 tweets from August 22nd, 2011- September 5th, 2011 collected

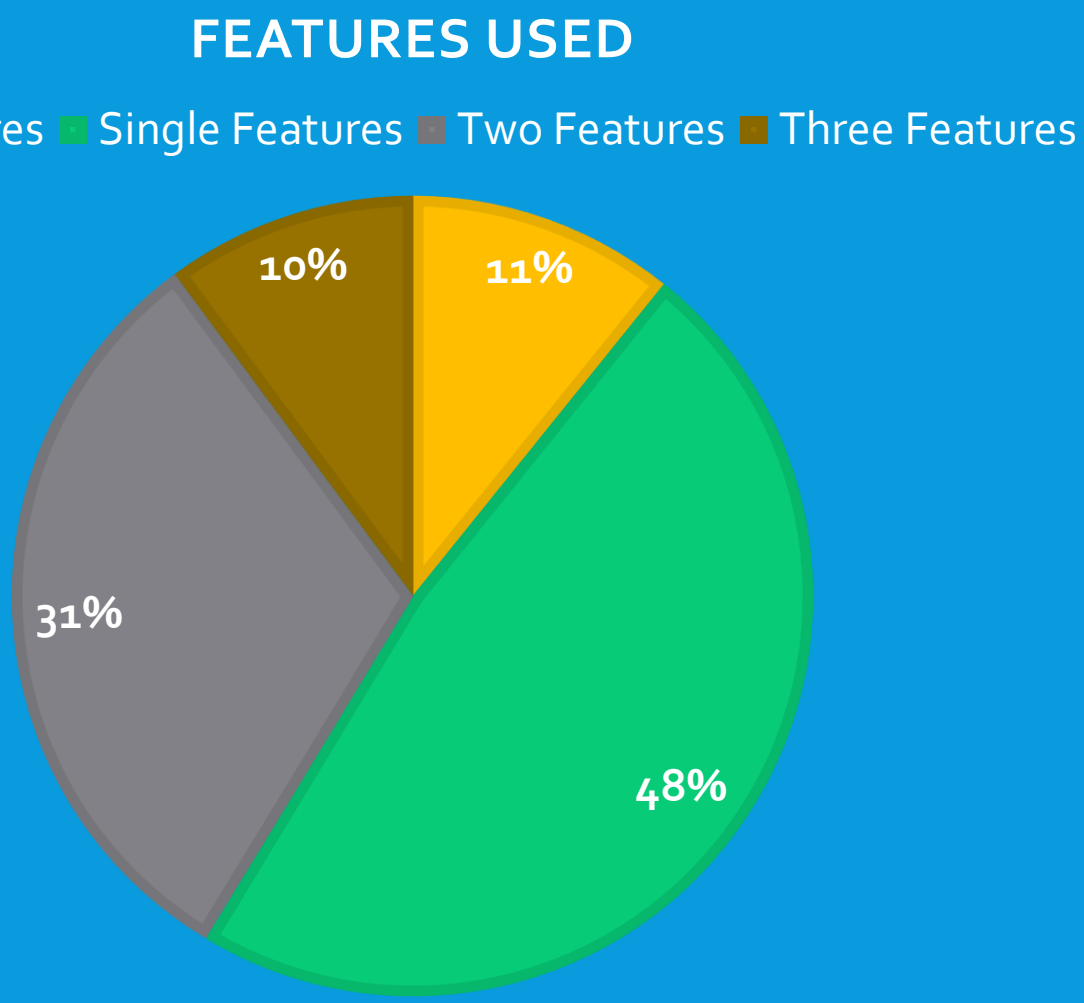
RESULTS


Tweets by Date
3000
2500
2000
1500
1000
500
0
Number of Tweets
22-Aug
23-Aug
24-Aug
25-Aug
26-Aug
27-Aug
28-Aug
29-Aug
30-Aug
31-Aug
1-Sep
2-Sep
3-Sep
4-Sep
5-Sep
Date

RESULTS

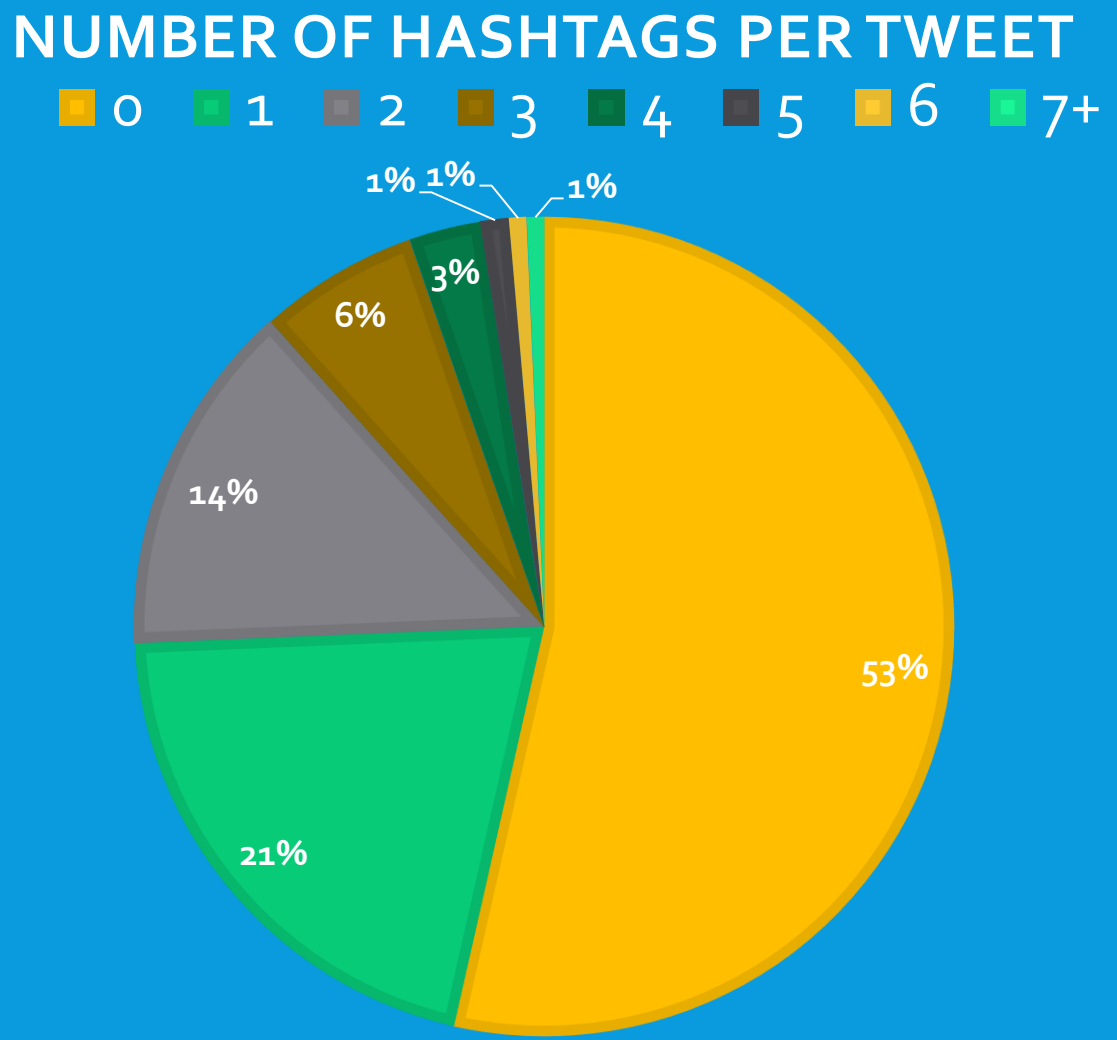

Sentiment
Number of Tweets
7000
6000
5000
4000
3000
2000
1000
0
Objective
Happy/Positive
Sad/Negative
Excited
Concerned/Nervous
Angry
Oblivious
Skeptical
1

RESULTS


FEATURES USED
No Features
Single Features
Two Features
Three Features
11%
48%
31%
10%

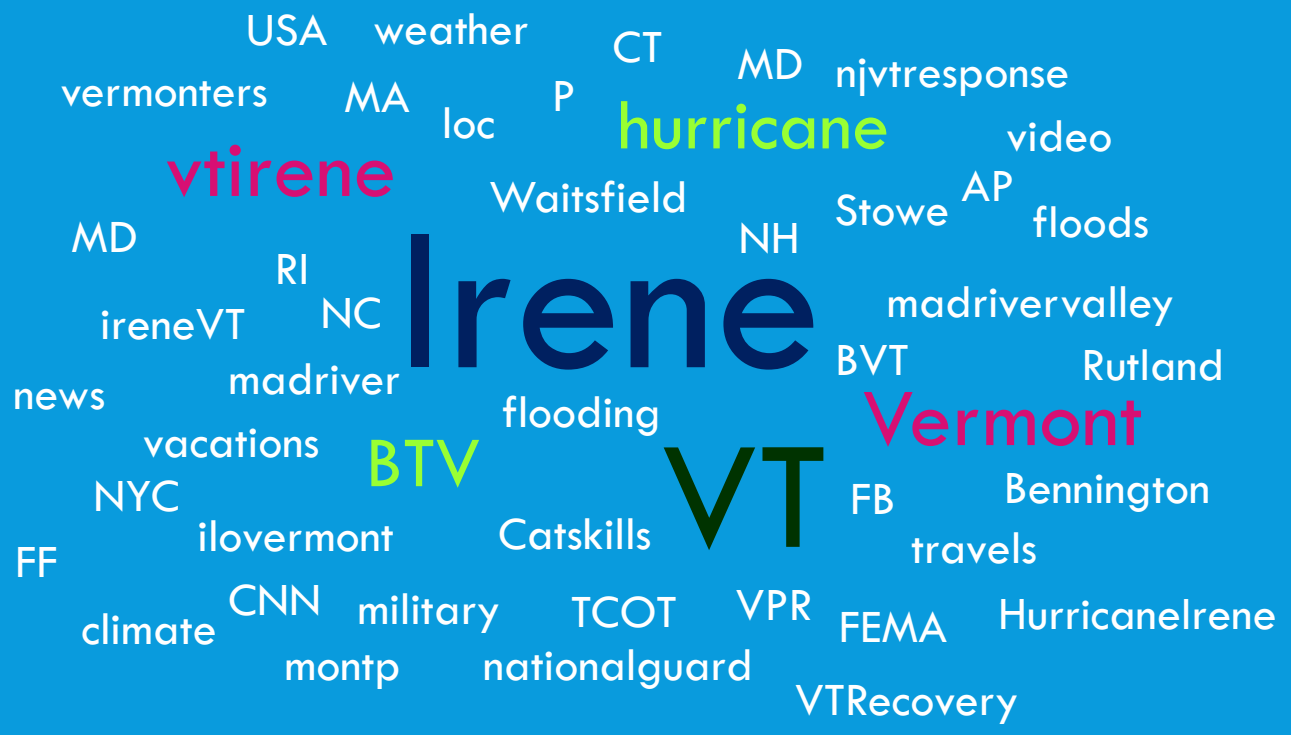
RESULTS

RESULTS



RESULTS

- 12,123 total hashtags used

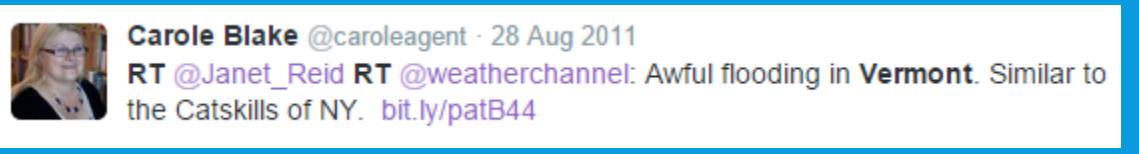


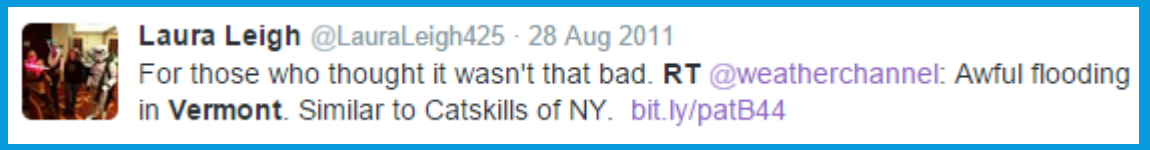
CONCLUSION

- Over half of tweets had a negative sentiment
  - Most tweets use at least one feature
    - Links are used the most to share information
  - If users are using hashtags,
  - The number of tweets is continuously growing as the date gets closer to when the hurricane hits Vermont, but then continuously decreases in the

LIMITATIONS

- What Could've Been Done Better?
  - Human Error
  - Search Terms Appear in Usernames
  - Repeated Tweets – denoted by RT
  - Subjectivity
  - Hashtag Reoccurrences
    - #Irene vs. #Irene's





ACKNOWLEDGEMENTS

- Thank you to CWDD and RACC at Vermont EPSCoR for the opportunity to research climate change this summer, as well to NSF for the grant funding for our projects and program

- Thank you to my mentor Dr. Steve Scheinert for always being a great resource and pillar of support

- Thank you to the other Q3 mentors for constant support and always giving invaluable and helpful suggestions

- Thank you to my Q3 colleagues for always offering insights, advice, support, and comic relief

REFERENCES

Agarwal, Apoorv, and Jasneet Singh Sabharwal. "End-to-End Sentiment Analysis of Twitter Data." In 24th International Conference on Computational Linguistics, p. 39. 2012.

Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. "Sentiment analysis of twitter data." In Proceedings of the workshop on languages in social media, pp. 30-38. Association for Computational Linguistics, 2011.

Dodds, Peter Sheridan, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth. "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter." PloS one 6, no. 12 (2011): e26752.

Ghiassi, M., J. Skinner, and D. Zimbra. "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network."Expert Systems with applications 40, no. 16 (2013): 6266-6282.

Go, Alec, Lei Huang, and Richa Bhayani. "Twitter sentiment analysis."Entropy 17 (2009).

Gerken, James. "Hurricane Irene 2 Years Later: Vermont Makes Second Anniversary Of Flooding And Damage." The Huffington Post. August 28, 2013. Accessed March 11, 2016. http://www.huffingtonpost.com/2013/08/28/hurricane-irene-2-years-later_n_3827088.html.

Khan, Aamera ZH, Mohammad Atique, and V. M. Thakare. "Combining lexicon-based and learning-based methods for Twitter sentiment analysis."International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE) (2015): 89.

Kolbert, Elizabeth. "Hurricane Irene and Global Warming: A Glimpse of the Future? - The New Yorker." The New Yorker. August 28, 2011. Accessed March 11, 2016. http://www.newyorker.com/news/news-desk/hurricane-irene-and-global-warming-a-glimpse-of-the-future.

Lemonick, Michael. "Irene's Potential for Destruction Made Worse by Global Warming, Sea Level Rise | OnEarth Magazine." On Earth. August 26, 2011. Accessed March 11, 2016. http://archive.onearth.org/blog/irene-destruction-global-warming-sea-level-rise.

O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series." ICWSM 11, no. 122-129 (2010): 1-2.

Saif, Hassan, Yulan He, and Harith Alani. "Alleviating data sparsity for twitter sentiment analysis." CEUR Workshop Proceedings (CEUR-WS. org), 2012.

Saif, Hassan, Yulan He, Miriam Fernandez, and Harith Alani. "Contextual semantics for sentiment analysis of Twitter." Information Processing & Management 52, no. 1 (2016): 5-19.

Sobkowicz, Pawel, Michael Kaschesky, and Guillaume Bouchard. "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web." Government Information Quarterly 29, no. 4 (2012): 470-479.

QUESTIONS?