



# Topic Modeling Politicians' Tweets

Luke Von Glan  
5/24/19



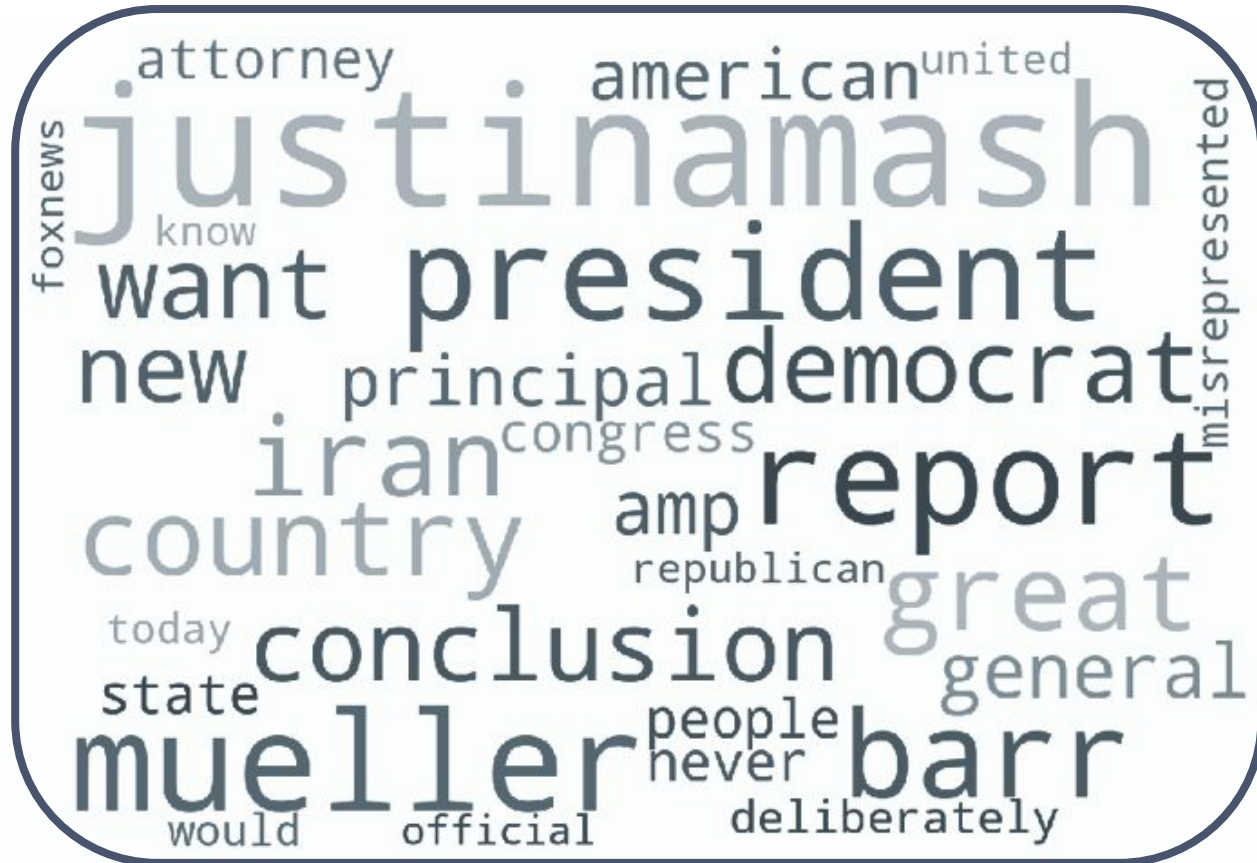
# Results

All about the president

Little separation of topics

Not much high-level discourse

# Our Political Discourse



# Process

## FIND

Used C-SPAN's lists of politician accounts to get Twitter ID's for 655 of the highest-ranking politicians in the US

# Process

## FIND

Used C-SPAN's lists of politician accounts to get Twitter ID's for 655 of the highest-ranking politicians in the US

## GATHER

Collected > 1.5M tweets that were from, @, or retweets of these accounts to store in MongoDB

# Process

## FIND

Used C-SPAN's lists of politician accounts to get Twitter ID's for 655 of the highest-ranking politicians in the US

## GATHER

Collected > 1.5M tweets that were from, @, or retweets of these accounts to store in MongoDB

## SCRUB

Cleaned tweets to remove punctuation, empty words, and get consistent word use

# Process

## FIND

Used C-SPAN's lists of politician accounts to get Twitter ID's for 655 of the highest-ranking politicians in the US

## GATHER

Collected > 1.5M tweets that were from, @, or retweets of these accounts to store in MongoDB

## SCRUB

Cleaned tweets to remove punctuation, empty words, and get consistent word usage

## MODEL

Entered tweets into 6 different topic models

# Most Distinct Model Topics

## LSA Count Vectorizer Model

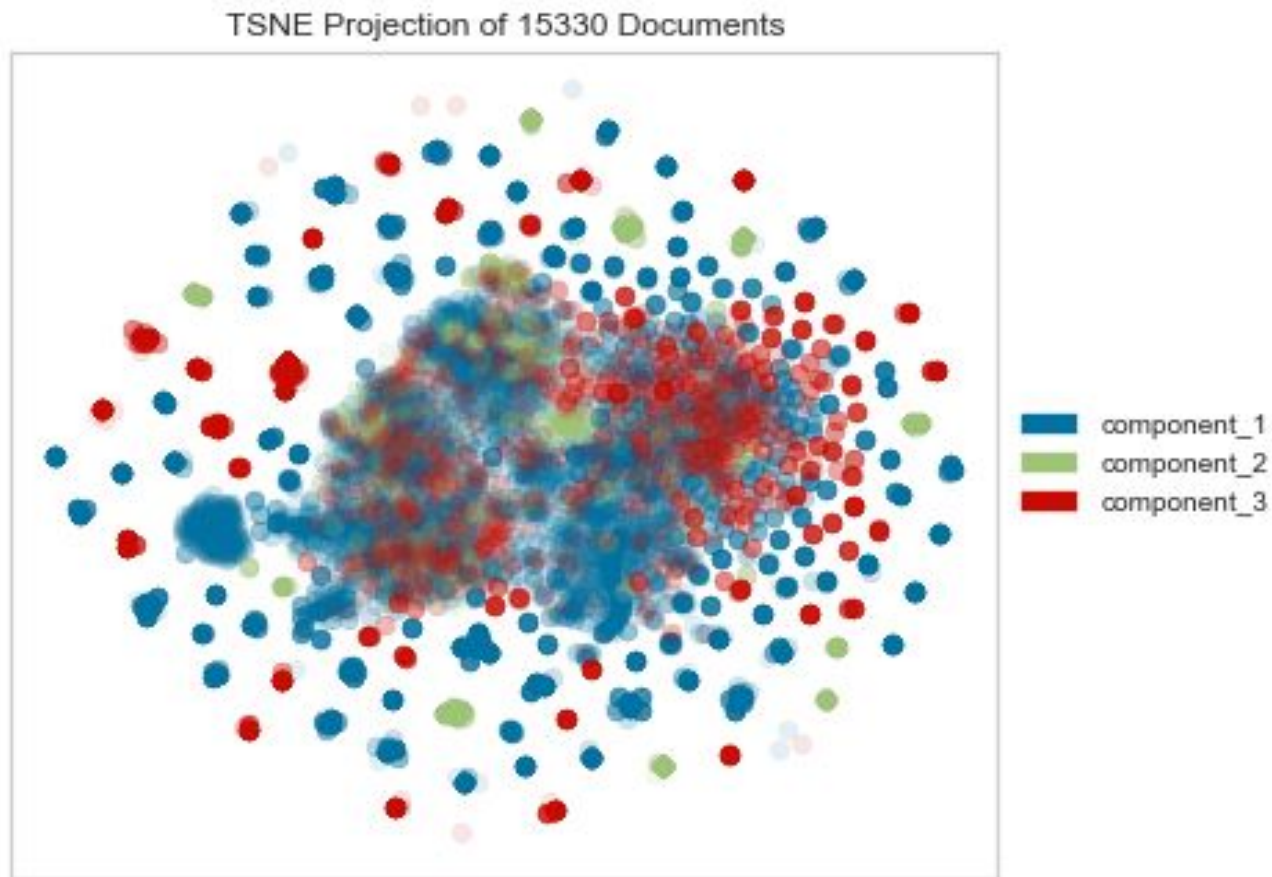
**Topic 1:** justinamash, mueller, report, barr, president

**Topic 2:** iran, want, state, united, never

**Topic 3:** country, congress, amp, justinamash, democrat



# Visualizing Topic Differences



# Takeaways

- President's account drives the political discourse on Twitter
- Much less diversity of topics and influential accounts than expected
- Hot take about the president is a great way to go viral