

# Towards Spatial Analysis of Opioid Abuse Using Twitter

Sheetal Pandrekar, MS, Manish Valakonda, MS, Xin Chen, MS, Fusheng Wang, PhD  
Stony Brook University, Stony Brook, NY

## Introduction

### Background

- 140+ Americans die daily of Opioid overdose (CDC)
- Trump declared Opioid crisis a public health emergency

### Motivation

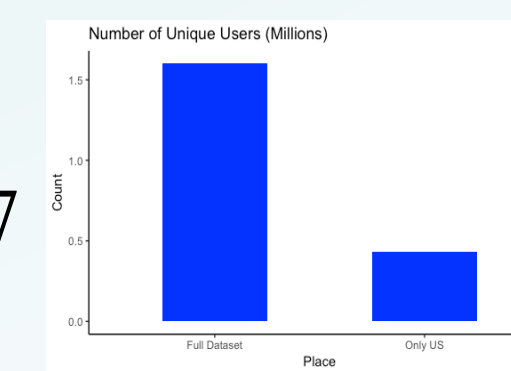
- Social media provides immediate access to public opinions

### Goals

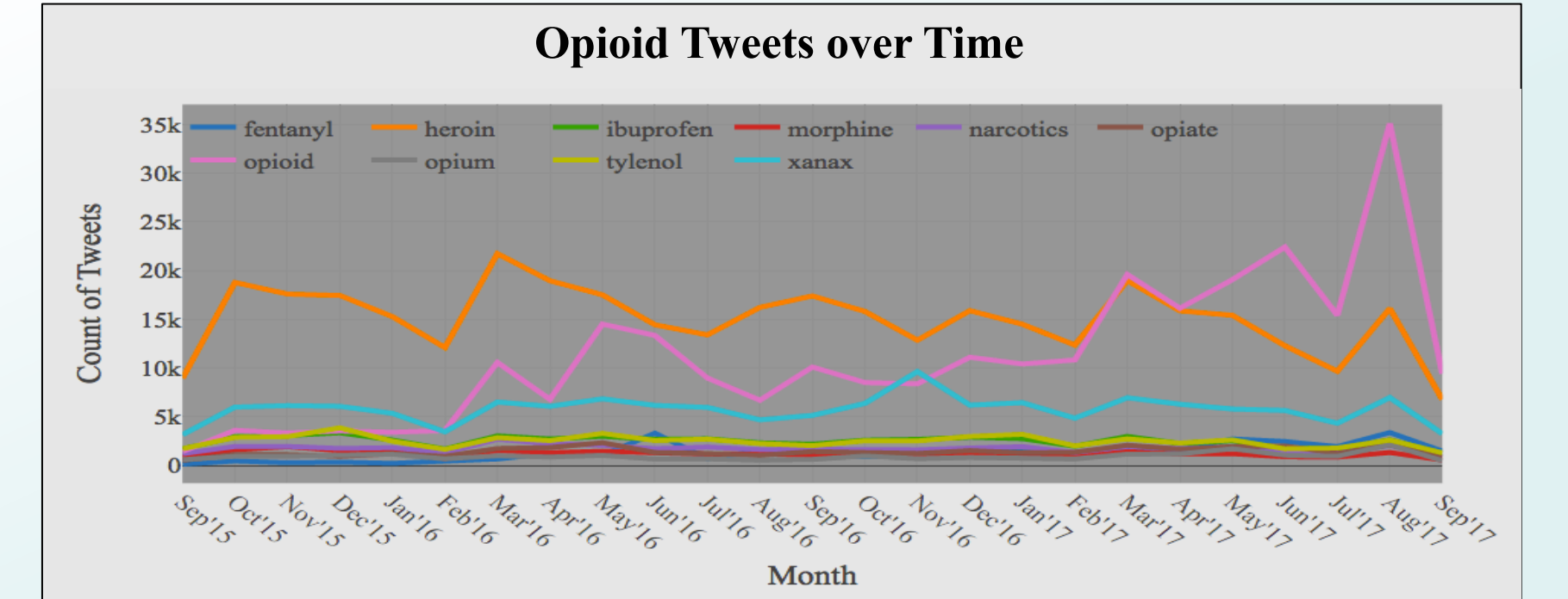
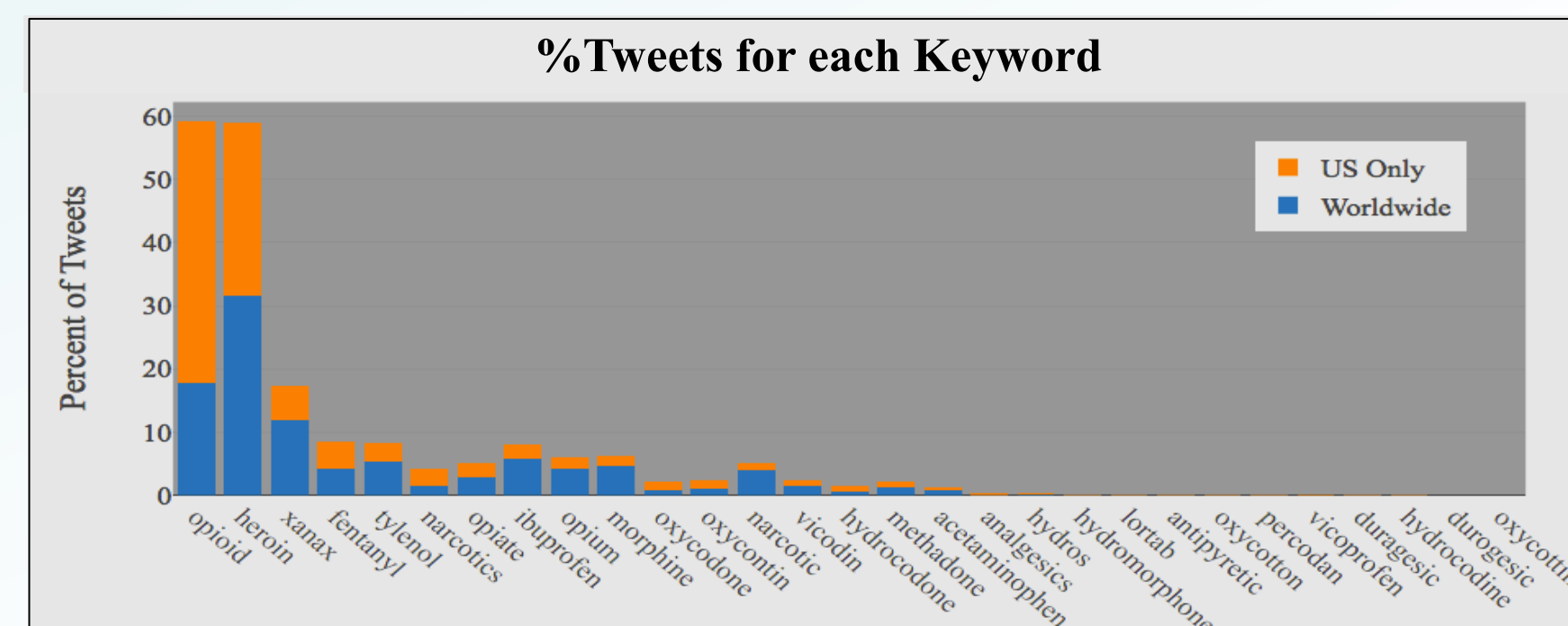
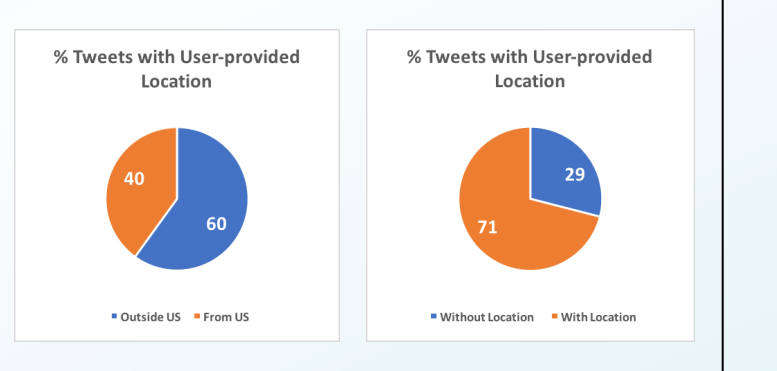
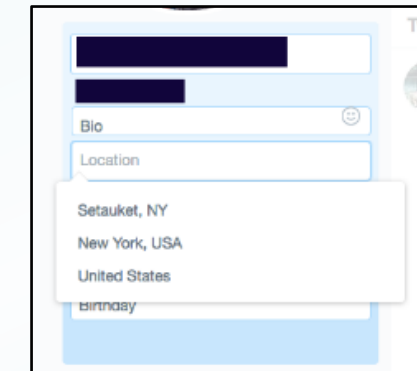
- Leverage social media to discover patterns in drug-abuse
- Develop techniques to enable large scale public health studies

## Exploratory Data Analysis

- ❑ Unique tweets (US): 1.2M
- ❑ Time Frame: 09/2015 to 09/2017
- ❑ Unique Tweets: 4.3M



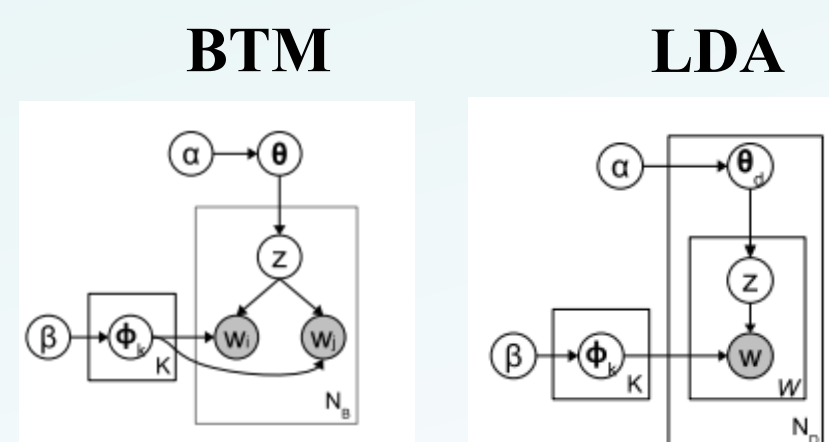
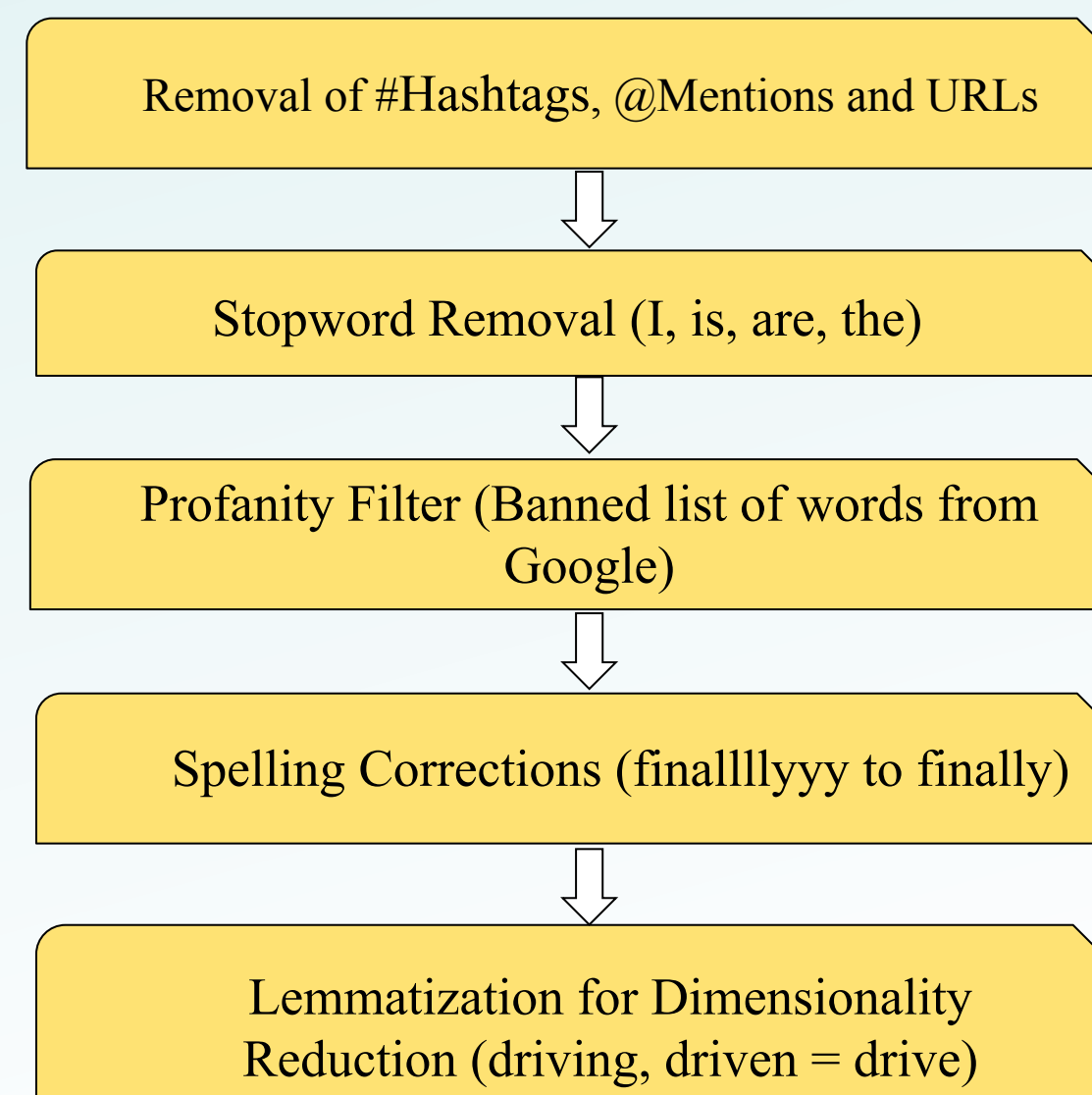
- User-provided location information from tweets



## Methods

### Topic Modelling

#### Text Preprocessing Steps

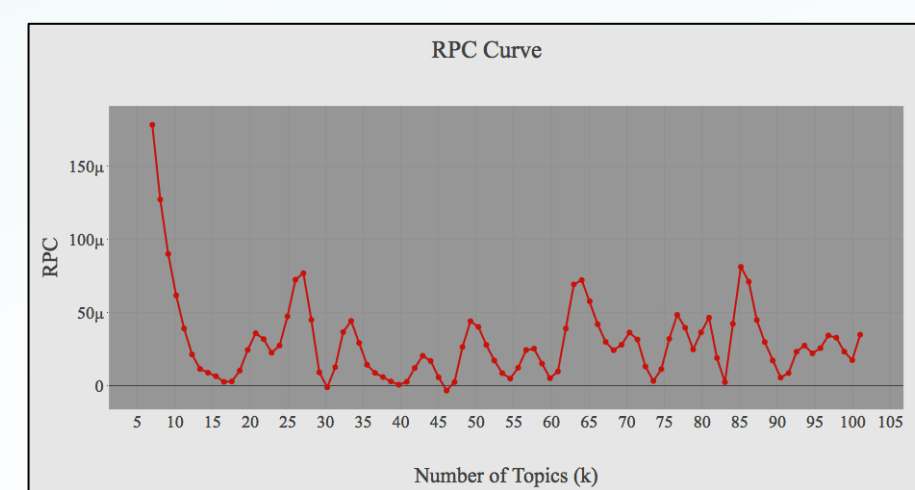


- Size of Vocabulary  $N = 42246$
- Number of Topics  $K = 20$
- Document-Topic Density  $\alpha = 50/K = 2.5$ 
  - High- $\alpha$  = more topics per document
- Topic-Word Density  $\beta = 200/N = 0.0005$ 
  - High- $\beta$  = more words per topic

#### Parameter Tuning

RPC=Rate of Change of Perplexity

$$RPC(i) = \left| \frac{P_i - P_{i-1}}{P_{i-1}} \right|$$

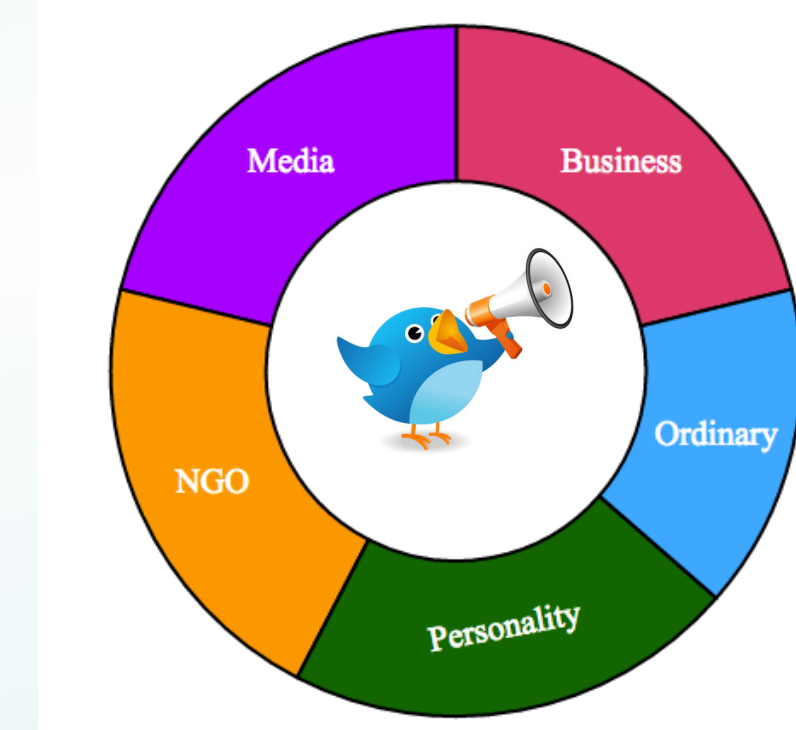


Perplexity of the model

$$p(\tilde{W}|M) = \exp \left( - \frac{\sum_{m=1}^M \log p(\tilde{w}_m|M)}{\sum_{m=1}^M N_m} \right)$$

$$\log p(\tilde{w}_m|M) = \sum_{i=1}^V n_{m,i}^{(0)} \log \left( \sum_{k=1}^K \phi_{k,i} \cdot \theta_{m,k} \right)$$

### Classification of Twitter Users

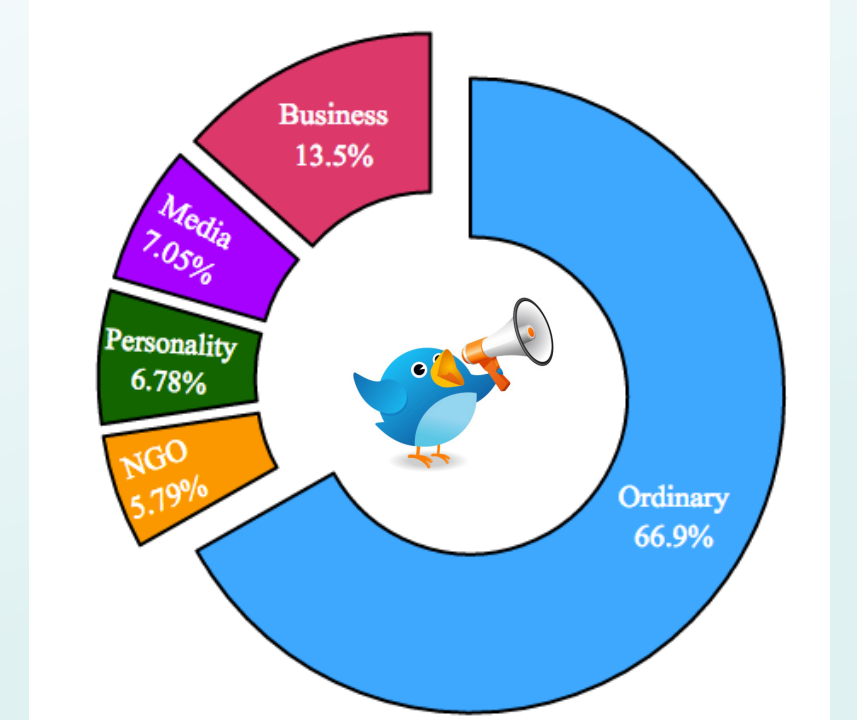


Network features	Activity features
Indegree (#followers)	Average number of tweets per day
Outdegree (#friends)	Number of favorited tweets
Interaction features	User's Influence
Fraction of Retweets	Number of Favorited tweets
Fraction of @mentions	Number of retweets for the user
Fraction of posts with URL in them	Is the user verified?

Training Data Size = 2645  
Labelled with AMT

Class/Classifier	Logistic Regression	Decision Tree	kNN
Business	0.67	0.73	0.71
Media	0.86	0.91	0.95
NGO	0.50	0.64	0.56
Personality	0.92	0.95	1.0
Ordinary	0.78	0.77	0.63
Average F1	0.75	0.80	0.77

Types of Users in Training Data

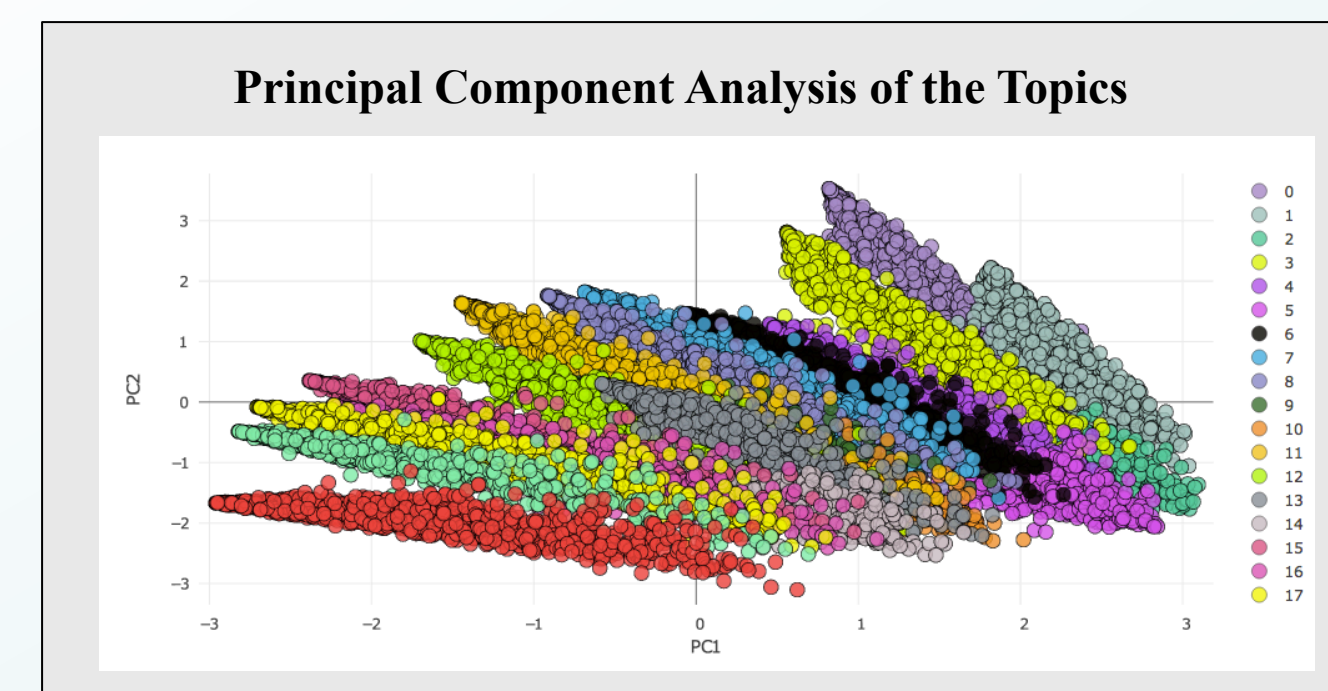
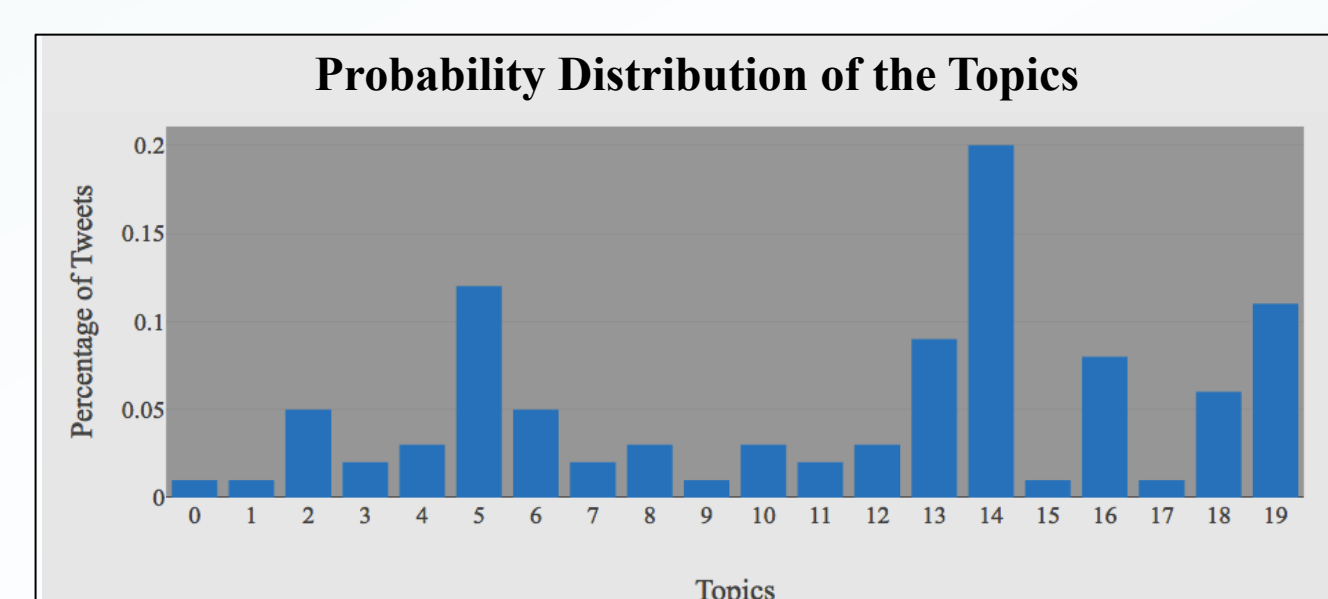


## Results

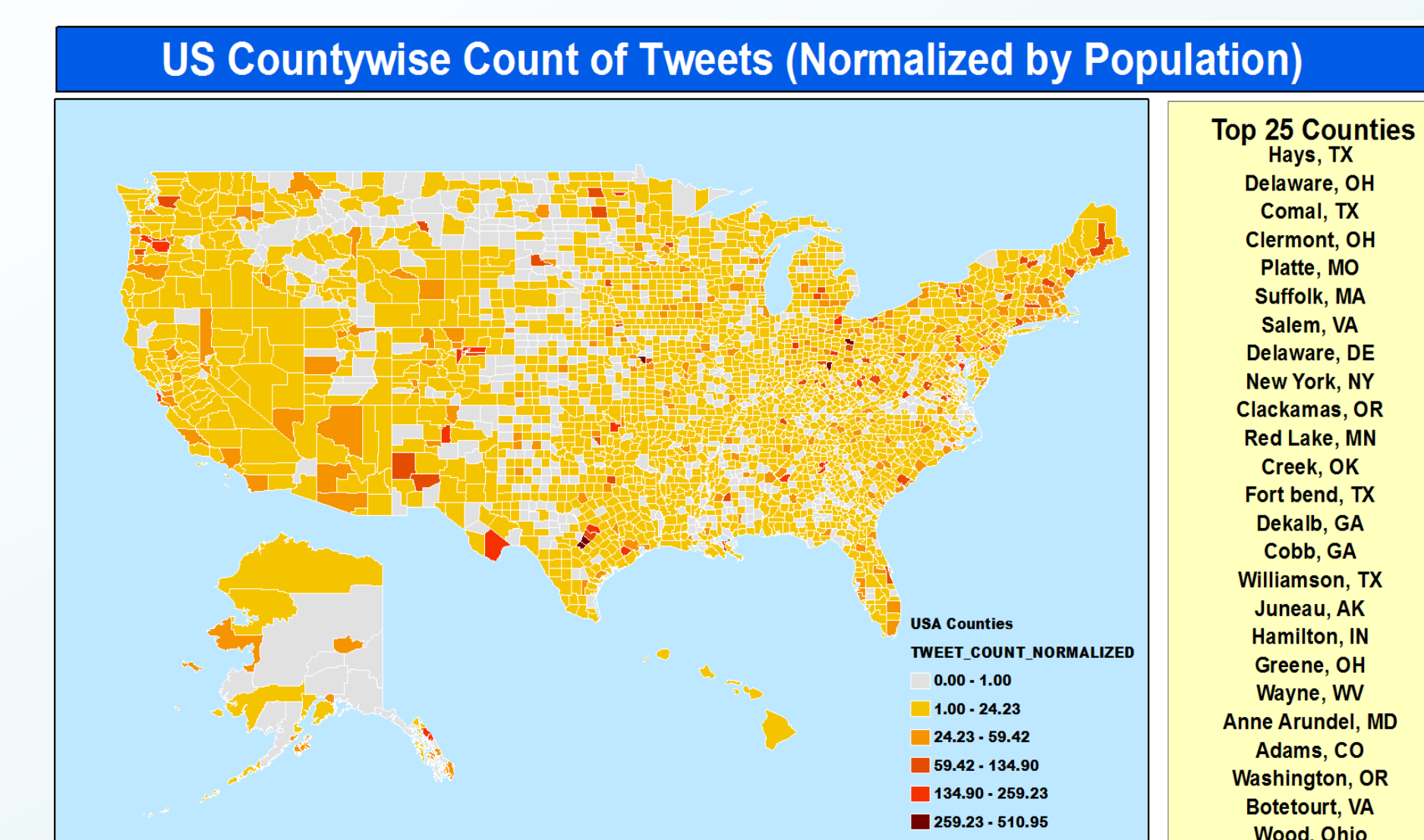
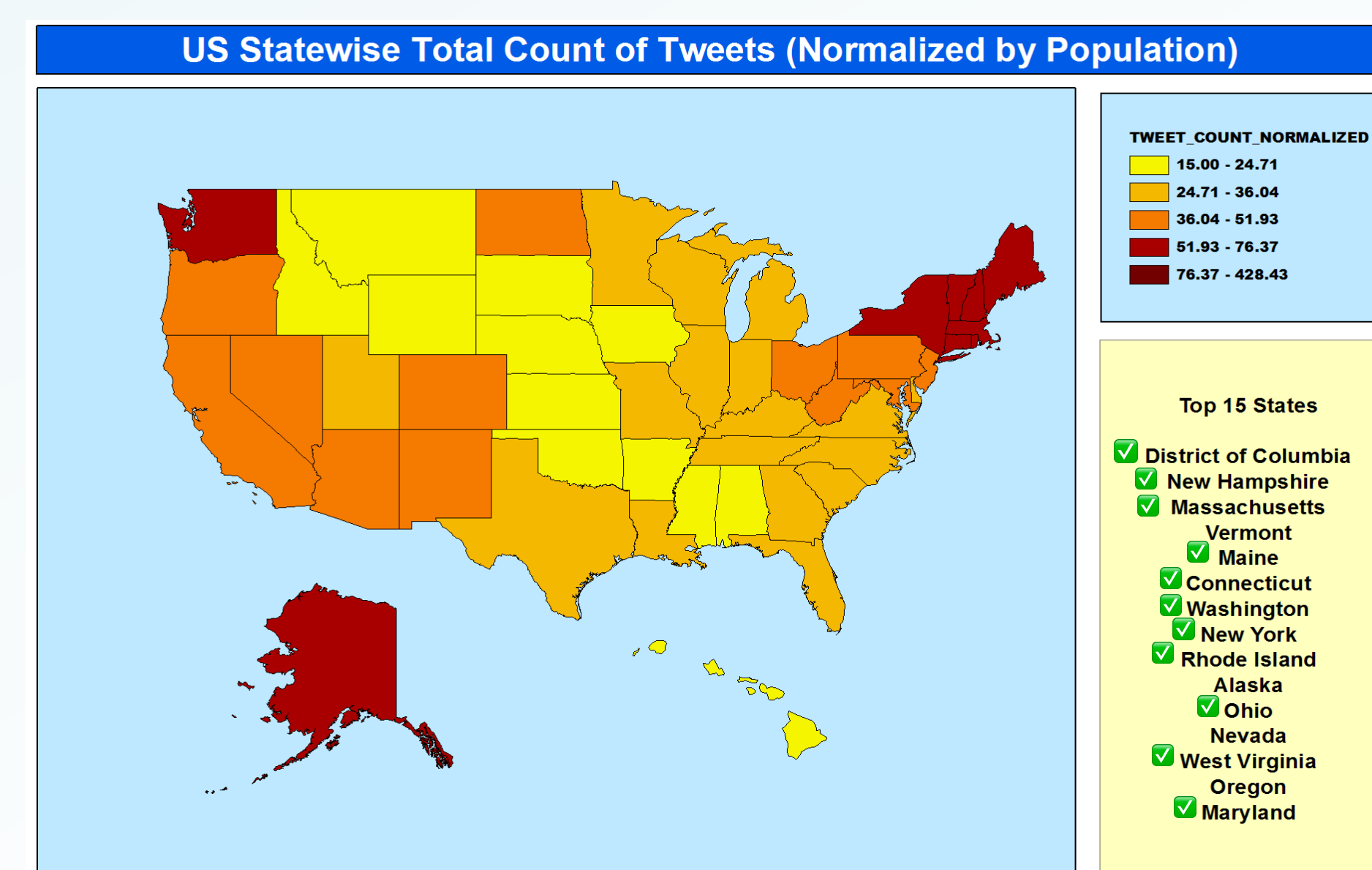
### Content Analysis



Topic	Sample Tweet
14	Mexican drug cartels have taken over much of the heroin market in the US
5	A new vaccine could make the brain immune to heroin and opioids
19	The opioid crisis could get worse if trump repeals obamacare via @vicenews
13	President's commission to combat the opioid crisis delays report again #drugs
16	Police dogs being harmed by dangerous drug fentanyl <a href="http://ref.gl/osnfmvgt">http://ref.gl/osnfmvgt</a>
18	Michigan governor creates state panel to address opioid crisis
2	Freemansburg police to host opioid awareness forum Monday #drugs
6	Elephant sedative carfentanyl is the latest and deadliest heroin ...
4	Seattle officials endorse first safe heroin injection site in us <a href="http://ift.tt/2bzfmty">http://ift.tt/2bzfmty</a>
8	No wonder the CIA hates mexican cartels. Cutting into their afghan opium profits.
10	Police arrest two in Harrison county in undercover fentanyl operation
12	Dailycaller rate of babies born with opioid withdrawal rising at alarming pace <a href="http://trib.al/9uolj">http://trib.al/9uolj</a>



### Location-based Analysis



Statistically significant States for Opioid overdose deaths (2015)

- ✓ Connecticut
- ✓ Florida
- ✓ Illinois
- ✓ Kentucky
- ✓ Louisiana
- ✓ Maine
- ✓ Maryland
- ✓ Massachusetts
- ✓ Michigan
- ✓ New Hampshire
- ✓ New Jersey
- ✓ New York
- ✓ North Carolina
- ✓ Ohio
- ✓ Pennsylvania
- ✓ Rhode Island
- ✓ Tennessee
- ✓ Washington
- ✓ West Virginia

Src: Center of Disease Control and Prevention

