Facilitating Community-Informed Opioid Prescription Guidelines

Charet Bolton, Joey Zhou,
Manzoor Mirza, Raja Safiullah,
Yixi Qiu

AGENDA

- Policy Question and Motivation
- Data Description
- Analysis
  - GMM
  - PCA - dimensionality reduction
  - LDA
- Policy Implications

Policy Question & Motivation

Policy Question

What is the **general sentiment** towards the guidelines?

What concerns do people have regarding the drafted *Clinical Practice Guidelines*?

Background & Motivation

- Opioid epidemic
- CDC Clinical Practice Guideline for Prescribing Opioids
- 2022 draft of the guideline
- Public Comments on the draft accepted from 2/3/2022 to 4/11/2022

Source: https://www.hhs.gov/opiods/index.html

Data Description

Data Source

- Extracted from the Regulations.gov API using Python code, and we finally got 4,502 comments from different stakeholders.
- Includes: (1)unique identification of comments, (2) number of duplicates for comments, (3) the text of the comment.

Unstructured Part

- The text of comments is the unstructured part we used to answer the policy questions.
- We conducted (1) lemmanization, (2) lowercase transformation, (3) stopwords filtering on the unstructured comments.
- End up with *27607* lemmas

TF-IDF Representation

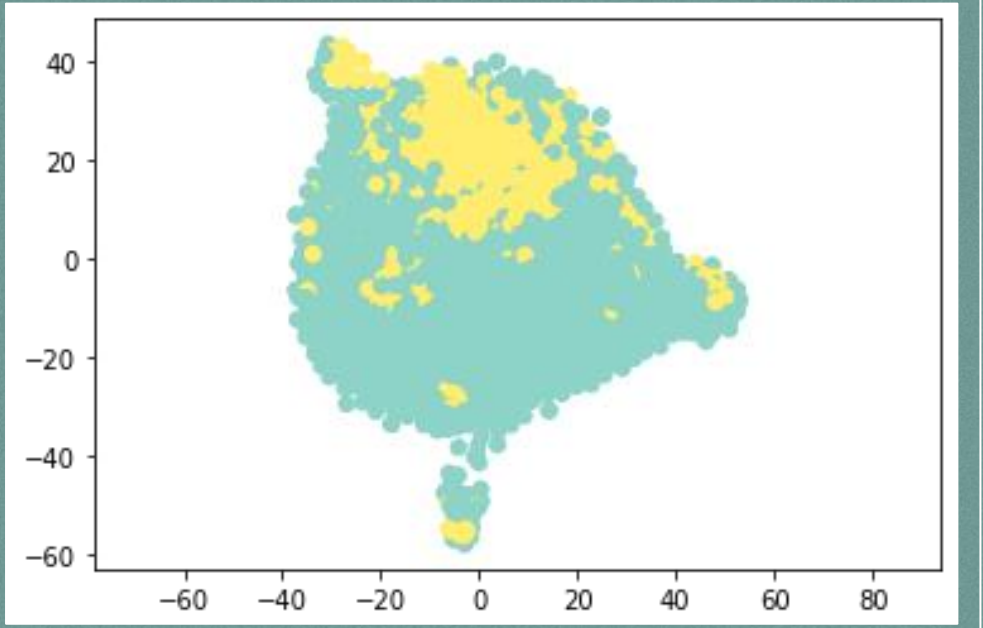
	Words (Features)				
Comments					

4502 Comments

1602 Words (features)

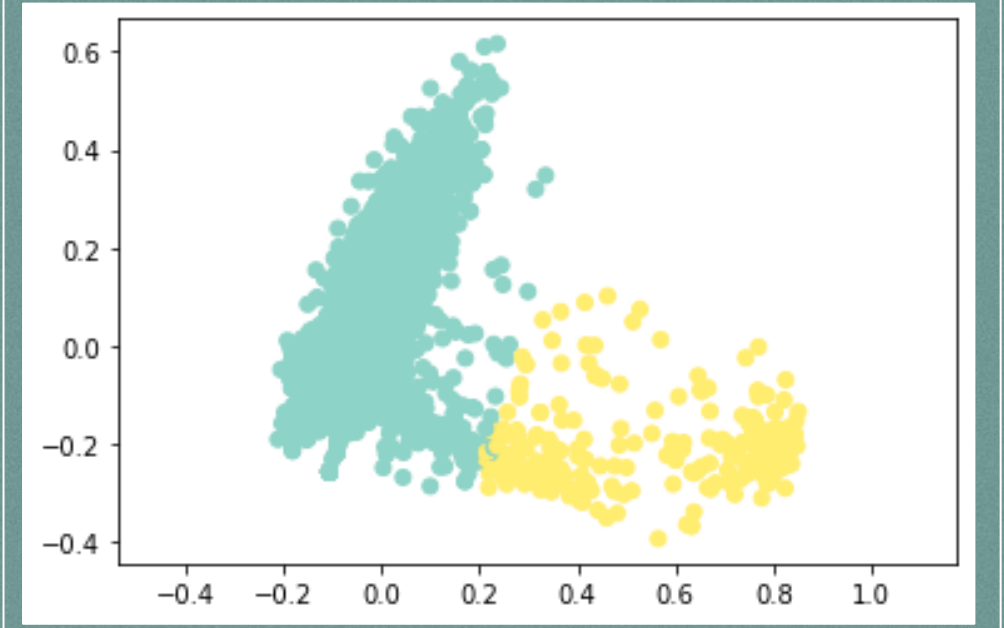
Question 1: Sentiment Analysis

Gaussian Mixture Model Clustering - visualization with TSNE



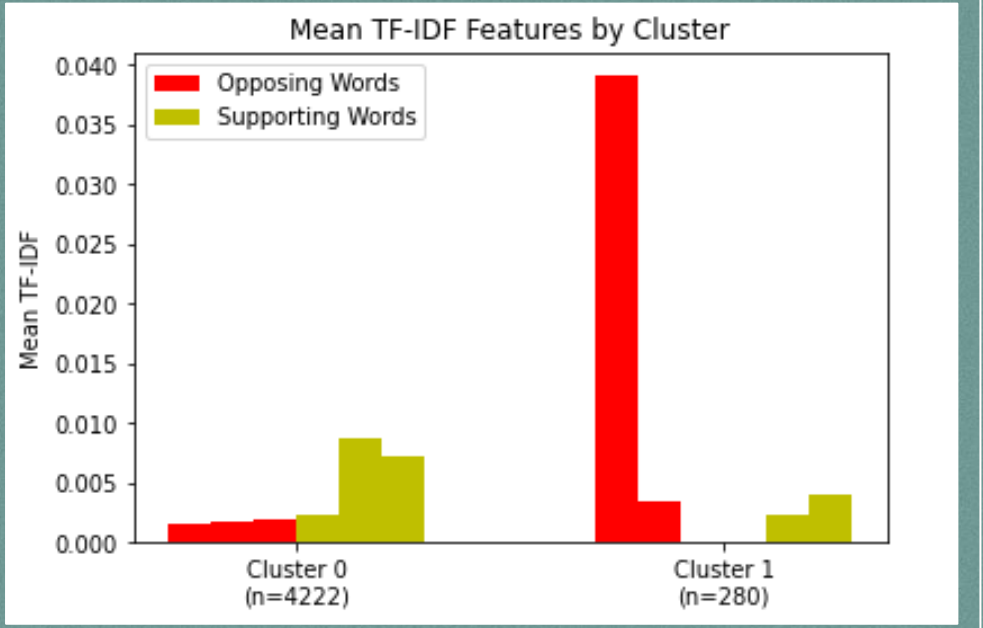
- Assuming Gaussian distribution for the data, we tune the hyperparameter K
- We get K = 2, using the Calinski Harabasz (CH) Index
- Visualize it using TSNE

K-Means Clustering - visualization using PCA



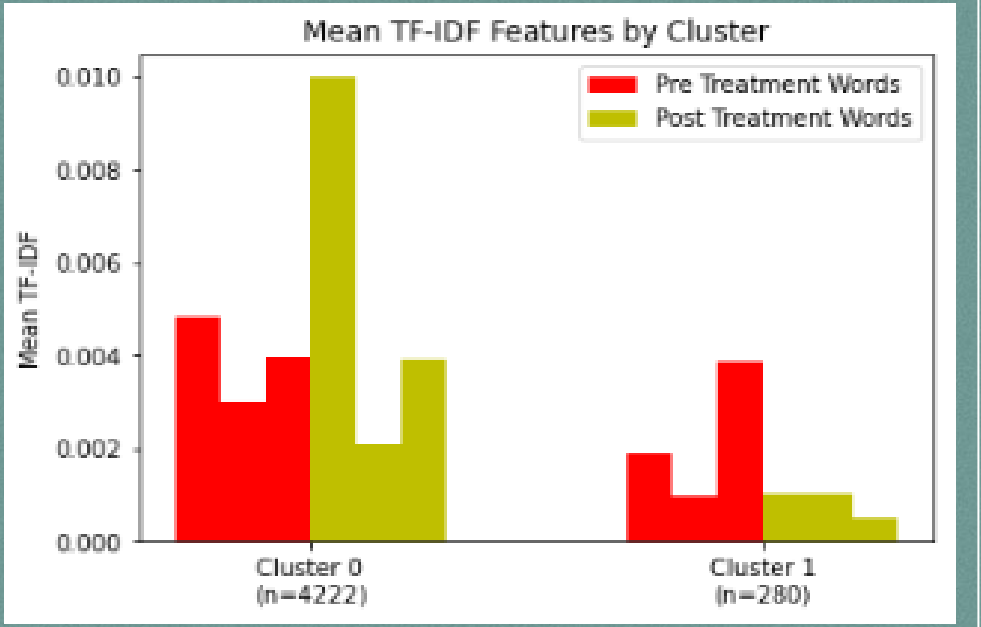
- Perform K means clustering at high dimensionality to minimize loss of information
- Chose K = 2 for clustering, which worked well - forming the clusters as anticipated
- Visualize it using PCA

Characterization of Clusters



- Characterizing clusters based on opposition and support
- Opposing words: 'illegal', 'criminal', 'rescind'
- Supporting words: 'control', 'concerned', 'supporting'

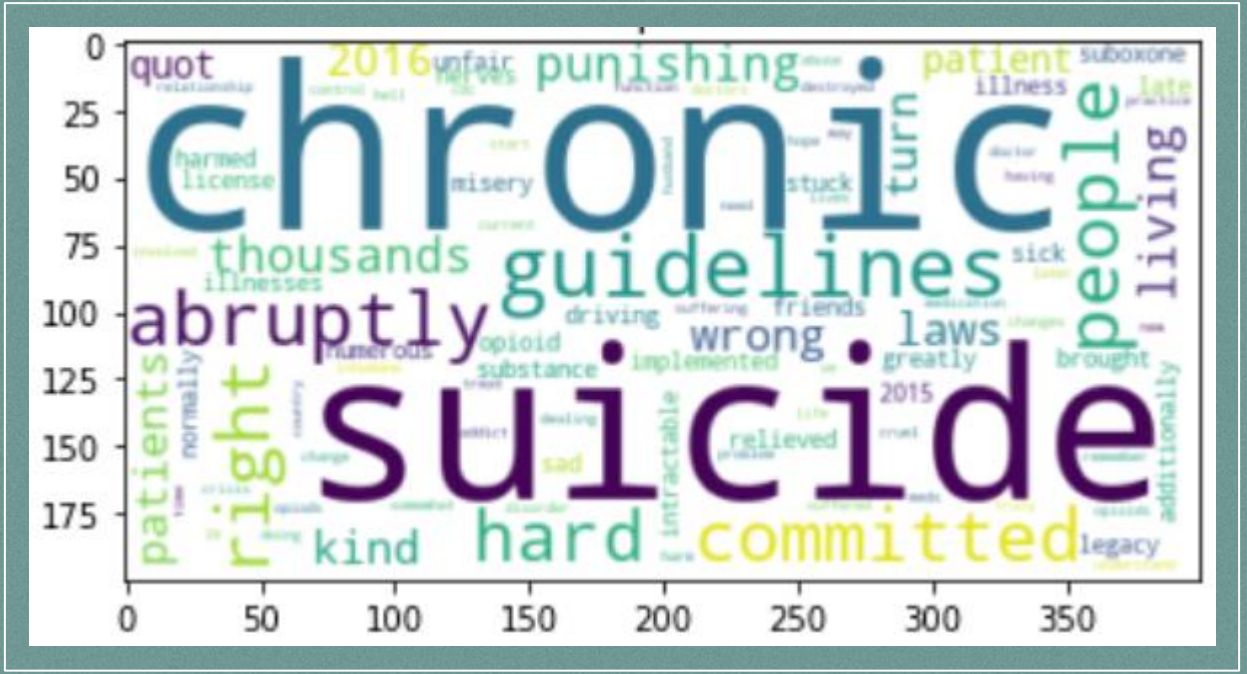
Characterization of Clusters - continued



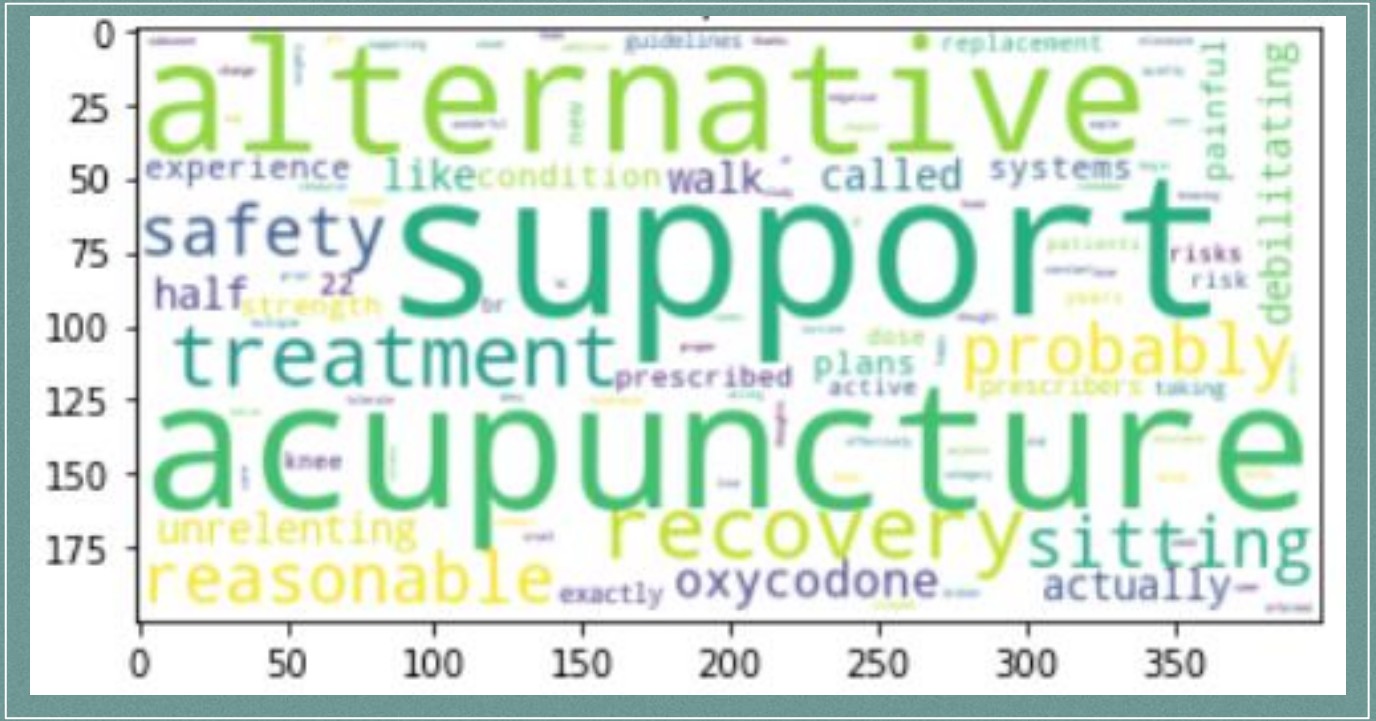
- Characterizing clusters based on aspects of CDC used for designing guidelines
- Pre-treatment words: 'start', 'visit', 'short'
- Post-treatment words: 'follow', 'potential', 'harm'

Question 2: Concern Identification

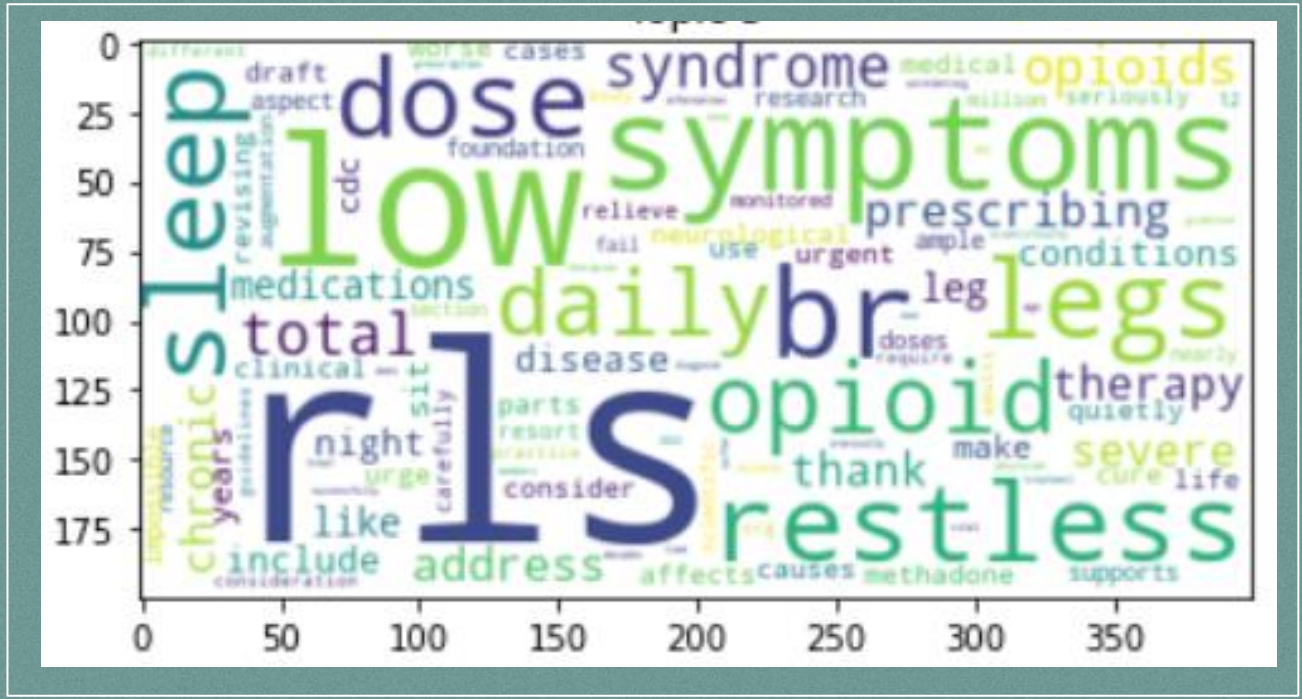
Latent Dirichlet Allocation



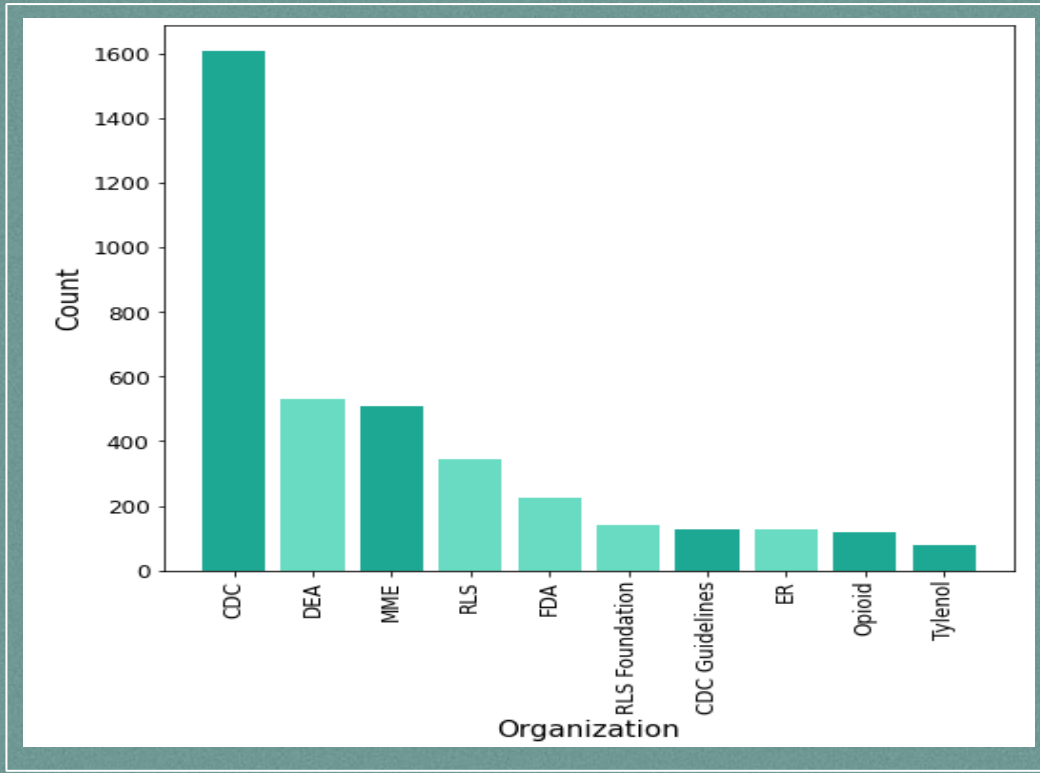
Latent Dirichlet Allocation



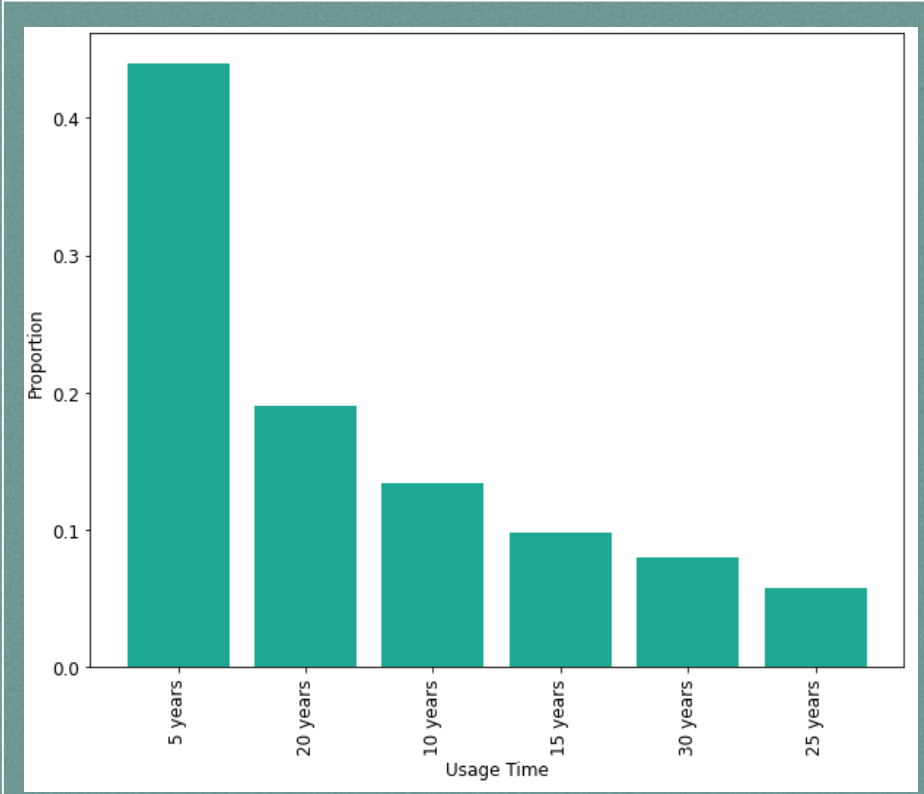
Latent Dirichlet Allocation



Top Entities with ORG label



Opioid Medication Usage Distribution



Conclusion

POLICY IMPLICATIONS

Two distinct clusters and key agency identification

STAKEHOLDERS

Majority supports the new guidelines

SENTIMENTS

Restless Leg Syndrome, alternatives, chronic use

CONCERNS

THANK YOU!

Appendix, if questions