

# **INCREASING ILLINOIS' APPEAL**

## **TEXT ANALYSIS**

By Kyla Ronellenfitsch

# EXECUTIVE SUMMARY

Illinois has faced a dramatic exodus of citizens for decades. This project aims to understand why this is the case, and what can be done to remedy the problem.

A collection of ~370,000 articles about Illinois were curated for this study. These articles covered a broad range of topics, some relevant to analysis and some not.

After data cleaning and topic modeling, three areas were thoroughly explored due to their presumed relevance to business leaders' and citizens' decisions to move to or away from Illinois. These topics are crime, economics, and leisure. Other topics, such as Covid or sports were deemed either not unique to Chicago or not a significant consideration and were not analyzed at this time.

Recommendations of how to attract both business and citizens to Illinois are included on the final slide of this report.

# METHODOLOGY

## Data Cleaning

- 373069 articles about Illinois were provided for this project
- Both text and title columns were cleaned by selecting English articles and removing special characters, URLs, and HTML
- Duplicate and irrelevant articles were dropped
- This resulted in a total of 209503 articles
- 150,000 articles were used in the final analysis, a decision made for computational efficiency

## Topic Modeling

- NER, LDA, Zero Shot Learning using NLI, and Snorkel were used on a sub-sample of 1000 articles to understand the key themes and entities, and computational efficiency tradeoffs
- One Shot Learning on the 'title' column was used for initial topic modelling
- Key terms used, article distribution and further topic definitions are included on the subsequent slide

## Crime analysis

- LDA was conducted on the crime articles, but found very similar themes
- Article content has been visualized using a word cloud

## Economic analysis

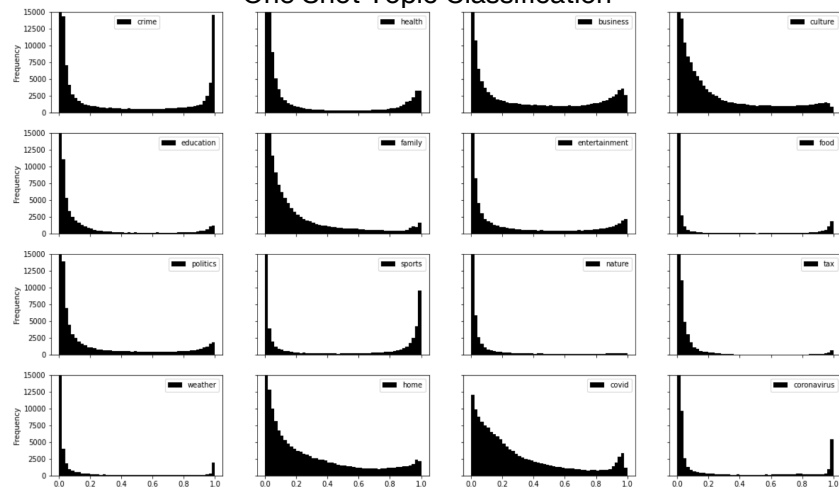
- Different approaches, including LDA, NER, ktrain text summarization, and TFIDF to find similar articles were used to understand major economic themes

## Leisure analysis

- Snorkel and textblob were used to classify positive and negative sentiment
- Text summarization has also been used to provide examples of positive articles related to entertainment, culture, and food

# TOPIC MODELING

One Shot Topic Classification



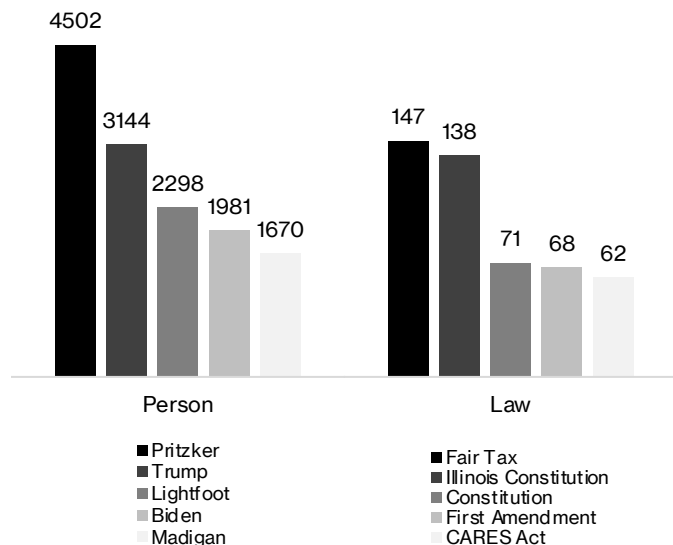
Category	Defined	Articles
Crime	'crime' $\geq 0.8$	28618
Leisure	'culture' $\geq 0.8$ or 'entertainment' $\geq 0.8$ or 'food' $\geq 0.8$	22965
Sports	'sports' $\geq 0.8$	22730
Covid	'covid' $\geq 0.8$ or 'coronavirus' $\geq 0.8$	19605
Economy	'business' $\geq 0.95$ or 'politics' $\geq 0.95$ or 'tax' $\geq 0.8$ and 'sports' $< .9$	11211
Education	'education' $\geq 0.8$	5184
Environment	'weather' $\geq 0.8$ or 'nature' $\geq 0.8$	4683



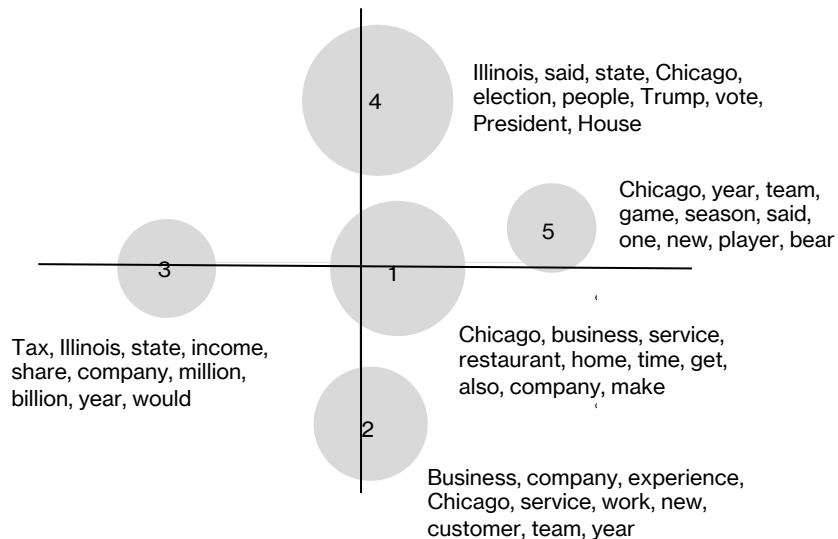
# ECONOMY

*"Illinois has a population problem: Our population has been shrinking faster than any other state except one. Population changes are a measure of governance quality. Over time, individuals migrate toward areas that are governed well and away from regions that are regulated badly. Gov. J.B. Pritzker and House Speaker Michael Madigan are working together to amend the state constitution to increase the tax burden."*

NER



LDA

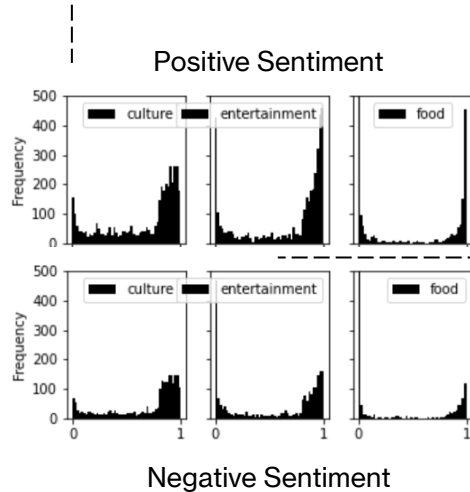


# LEISURE

*"The summer is the best time to visit, because you can fully take advantage of Lake Michigan, exploring the city and the Chicago river on foot. The winter is still full of things to do including Broadway and comedy shows, ice skating at Millennium park or Maggie Daley, or the Kris Kringle Xmas market."*

*"Among the highlighted songs in the show are I Made the Match Myself, Should We Ever Meet, Pride and Sense, Stranger Things Have Happened and Have a Piece of Cake. Chicago Shakespeare Theater has a hit on its hands with this show. Fans of Austen's stories as well as music fans..."*

*"Dining in Chicago is one of our strongest areas, but with so many options to choose from, it can be hard to pick a spot. We made the decision easier for you by narrowing down a list to some of the top restaurants in Chicago. The selection at Little Bad Wolf is to die for. Located in Wicker Park, Ettas Woodfire style cooking is the perfect addition to their innovative menu."*



Positive sentiment related to leisure (including culture, entertainment, and food) far outweighs negative sentiment.

# ATTRACTING RESIDENTS

Woven through these articles are two polarized aspects of life in Illinois.

Crime and safety issues are significant reputational problems for Illinois, and for Chicago in particular. They are obvious potential causes of concern for any prospective resident.

At the same time, quality of life in terms of entertainment, sports, culture, and food is very positive.

Outreach efforts should focus on the livability of the state, while working to alleviate concerns about safety.

The taxation environment as described in these articles is a definite concern and hindrance to attracting entrepreneurs and business to Illinois. However, with the failure of the Fair Tax legislative push and continued use of a flat income tax rate, business decision makers could be convinced that Illinois does provide an hospitable environment for growth.

Beyond taxation, efforts to attract business should focus on alternative elements of competitiveness. such as a highly educated workforce and vibrant urban living for employees.

# ATTRACTING BUSINESSES