Growing Chicago

A Look into the Opportunities and Existing Risks to a Sustained & Growing Population

Natural Language Processing Final Project

Created by: Matt Norgren

Executive Summary

Host of two World Fairs and World Class Art & History Museums, the third largest city in the United States has its share of amenities. While good things come and go, one major trend that seems to have a foothold on Illinois and more specifically the Chicago-land area is people going (negative population growth). While it was found that people positively reacted to the restaurant life and the numerous sports teams, there are still things left to be desired such as the reductions in the level of crime and concerns about public leadership that might account for this negative trend.

SOURCE DATA OVERVIEW

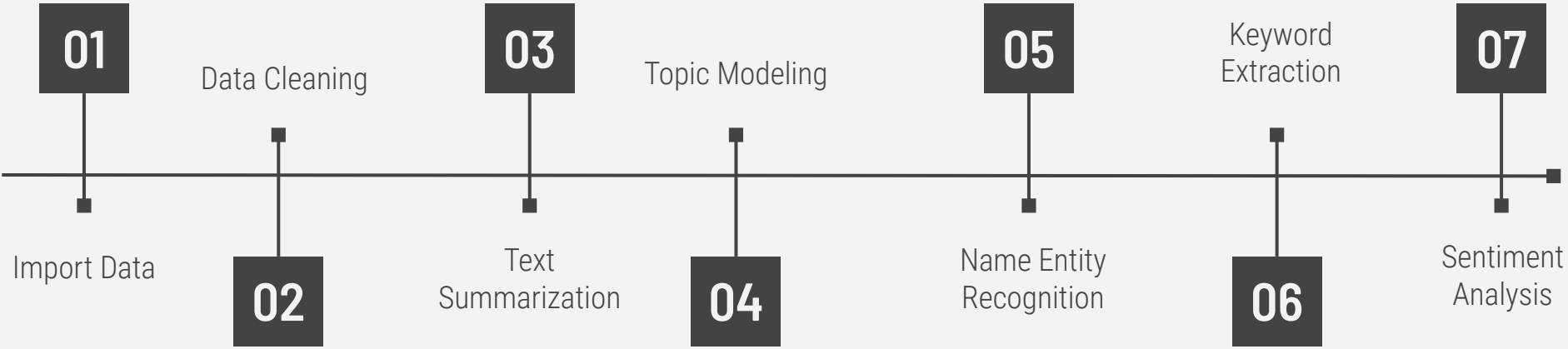
	crawled_date	language	text	title
0	1589155200000	english	\nGov. Jay "Fatso" Pritzker called on all Illi...	All In Illinois
1	1589155200000	english	May 10, 2020 -The Illinois Department of Publi...	The Illinois Department of Public Health Annou...
2	1589155200000	english	Gloria Lawrence said: May 10, 2020 at 1:31 AM\...	Foto Friday: Alton, Illinois
3	1589155200000	english	NBA to follow German soccer league model with ...	Chris Broussard on Michael Jordan returning to...
4	1589155200000	english	Search Minggu, 10 Mei 2020 Pork chops vs. peop...	Pork chops vs. people: Can Americans' appetite...

Comprised of web scraped news articles, the source data is built of:

- 373,068 Articles (Multi-Language)
- 370,586 Articles in English
- 893,157,490 Total Characters
- Robust Titles (Avg. Character Length: 77)
- Descriptive Article Body (Avg. Character Length: 2,316)

And for added measure, citing social media, emojis are embedded throughout 😐

METHODOLOGY



01 Import Raw Data into

02 Data Cleaning via removing punctuation, emojis, and stop words (i.e. this, is, for, at, the)

03 Reduce text by summarizing at 30% of the original size

04 Latent Dirichlet Allocation - Helping identify unobserved relationship between not immediately obvious data points

05 Name Entity Recognition via Spacy provides incredible syntactic parsing providing incredible insight into unstructured data

06 Keyword Extraction via Rake

07 Utilizing Valence Aware Dictionary and sEntiment Reasoner, positive, negative, and strength of emotional sentiment is calculated based on a blend of Naive Bayes, Max Entropy, and Support Vector Machines. It is rule based but can learn as needed.

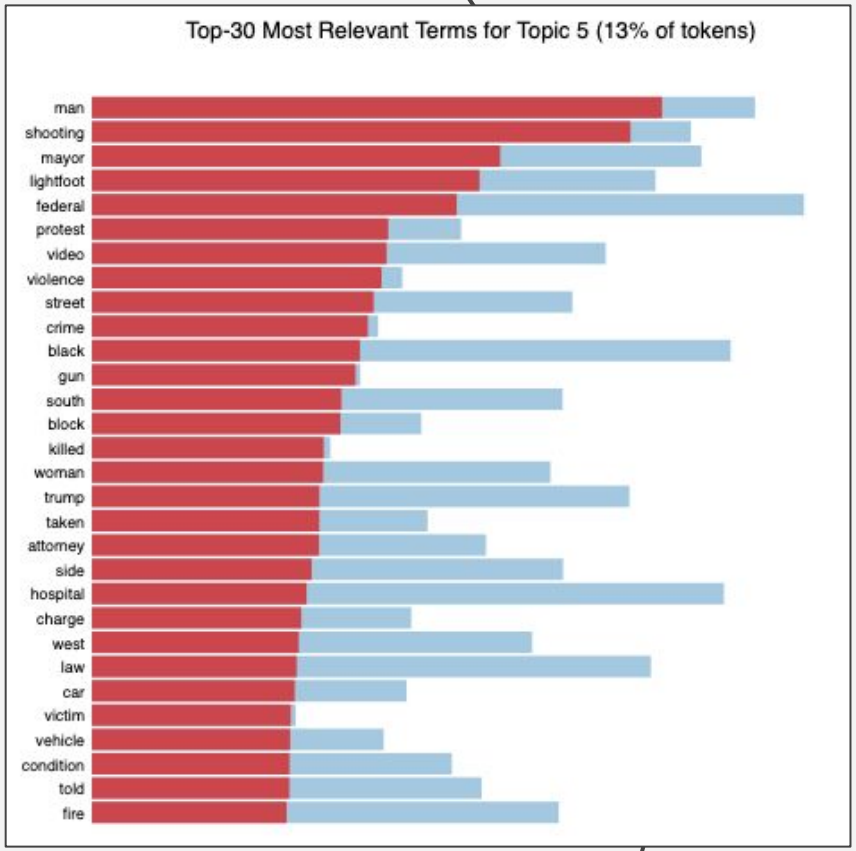
EXODUS

Recurring Topics and Negative Sentiments that are associated with Illinois and Chicago-land.

Given that there is a recurring negative sentiment related to crime, the judicial system, and political figures

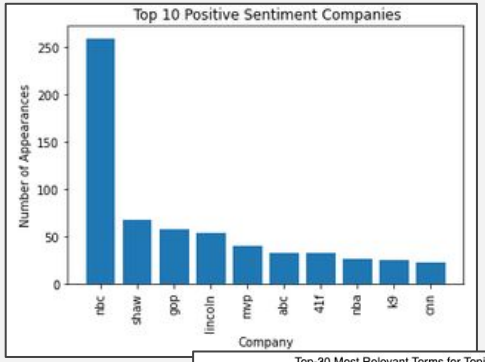
Positive actions that could start turning around this trend:

- Delve deeper into how noted political figures factor into negative sentiments (The why) (i.e. corruption, lack of action)
- Create a taskforce to review and deploy Police & Community resources to areas in need
- Identify areas where positively identified sentiment items can be further invested in and developed.

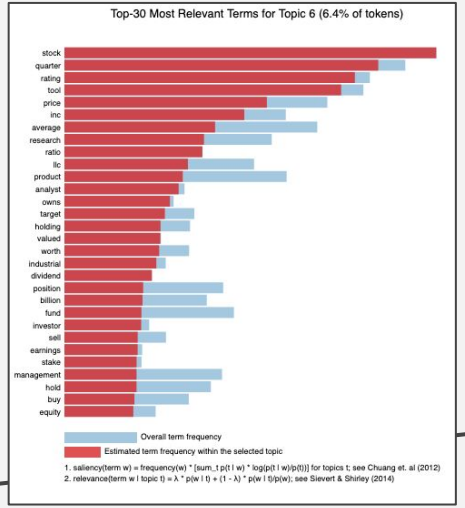


OPEN FOR BUSINESS

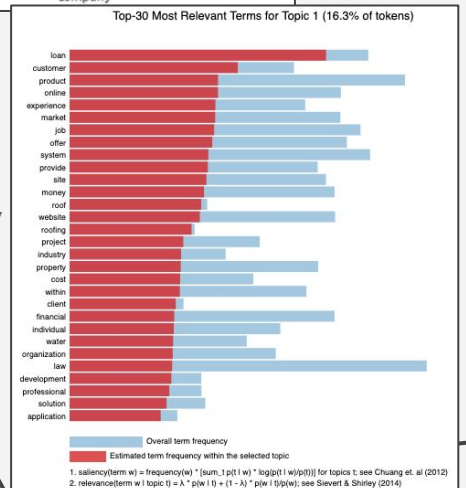
Why Businesses should relocate & continue to stay



INFORMED INVESTORS



OPPORTUNITY



TAXES

INFORMED INVESTORS Chicago-Land has a significant amount of business acumen and the companies to back it up. Informed customers are going to challenge businesses to improve and it has the customer base to reap the benefits.

OPPORTUNITY An incredible test market ripe for the picking. Driven by pure economics there is a considerable demand and room for healthy competition. Companies have ample land and clients to expand to.

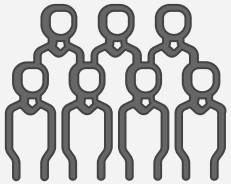
TAXES While most don't see taxes as a benefit, corporations making the investment in the community consistently reap tax benefits from the communities and state legislators.

A PLACE CALLED HOME

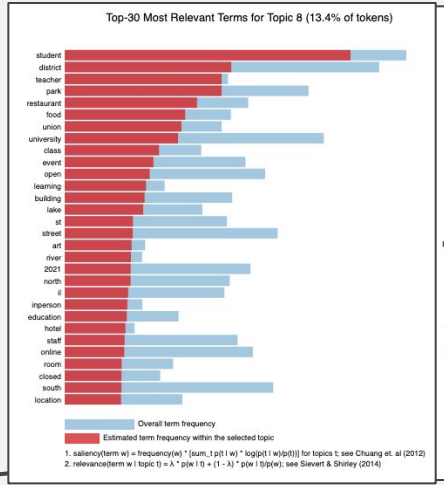
Why people should relocate & continue to stay

COMMUNITY

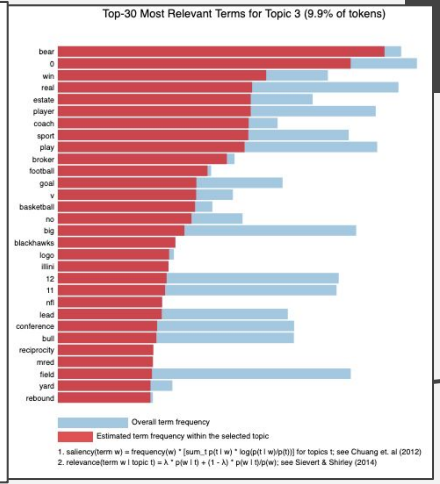
Entities	
american	727
democrat	647
republican	426
christian	54
mexican	35
african american	29
japanese	28
indian	20
european	20
chinese	19
armenian	16
gop	15
british	13
muslim	11
german	10
jamaican	9
hawaiian	9
russian	8
canadian	8



PARKS & RESTAURANTS



SPORTS CULTURE



COMMUNITY continues to grace the top of the positive sentiment list. Investing in the city and suburbs communities with parks and sports areas will further incentivize families to take a chance. This investment should pay dividends to the feeling of safety and reduce crime.

SPORTS culture and having a franchise to invest in is one of the way to settle in. Increase marketing campaigns and events to promote and give back to the community. A community within the community.

RESTAURANTS are top notch, they help residents relax and attract people from outside the city centers. Refinement and diversity are key to this space and programs to help underprivileged groups get started could pay dividends.