Comparing the Community Areas in Chicago

Looking for Similarities in Further Away Areas and Areas in Transition to Identify Investment Opportunities

Introduction

- Chicago is the 3rd largest city in the United States
- It's the 2nd most visited city in the US, complete with arts, sports, architecture, cuisine, commerce, universities, and more
- Unfortunately, the city is also known for its crime rate, which is generally understood to be localized to specific areas
- For someone looking to invest in Chicago, the size of the city can be daunting, so I've analyzed Community Areas to identify similarities
- A Community Area different from those surrounding it could warrant additional research or consideration

Data

- The data used for this project is from Wikipedia, OSM, ARCGIS, Foursquare, and from the Data Portal from The City of Chicago
- Wikipedia was used to compile a dataframe with all of the 77 Community Areas of Chicago
- OSM and ARCGIS were used to add to the dataframe the latitude and longitude coordinates
- Foursquare was used to load the locations within 1000 meters of the center of the Community Area
- The Data Portal was used to compile a record of every reported crime per Community Area over the last 10 years

Methodology

- *BeautifulSoup* is used to scrape the Wikipedia page for Chicago Community Areas in order to create a dataframe with each of the Community Areas

- Various *pandas* functions were used to compile all 1+ million reported crimes over the last 10 years
    - The data was then separated into Community Area, and further separated into the 5 years from 2009-2013 and 2014-2018
    - Finally, the % increase/decrease in crimes from the first period to the second period per Community Area was calculated

Methodology (cont.)

- *Geocoder* was used to generate the latitude and longitude of each of the 77 Community Areas
  - The *OSM* data was generally correct, but when I noticed some outliers, I did further investigation
  - Some of the locations were for streets of the same name of the Community Area, so I switched the location data to that provided by *arcgis*
  - I then compared the outputs from *OSM* and *arcgis* and a handful showed significant deviations
  - 4 of the *arcgis* locations were determined to be superior to the *OSM* locations, and so the *arcgis* location data for those 4 were used over the *OSM* data
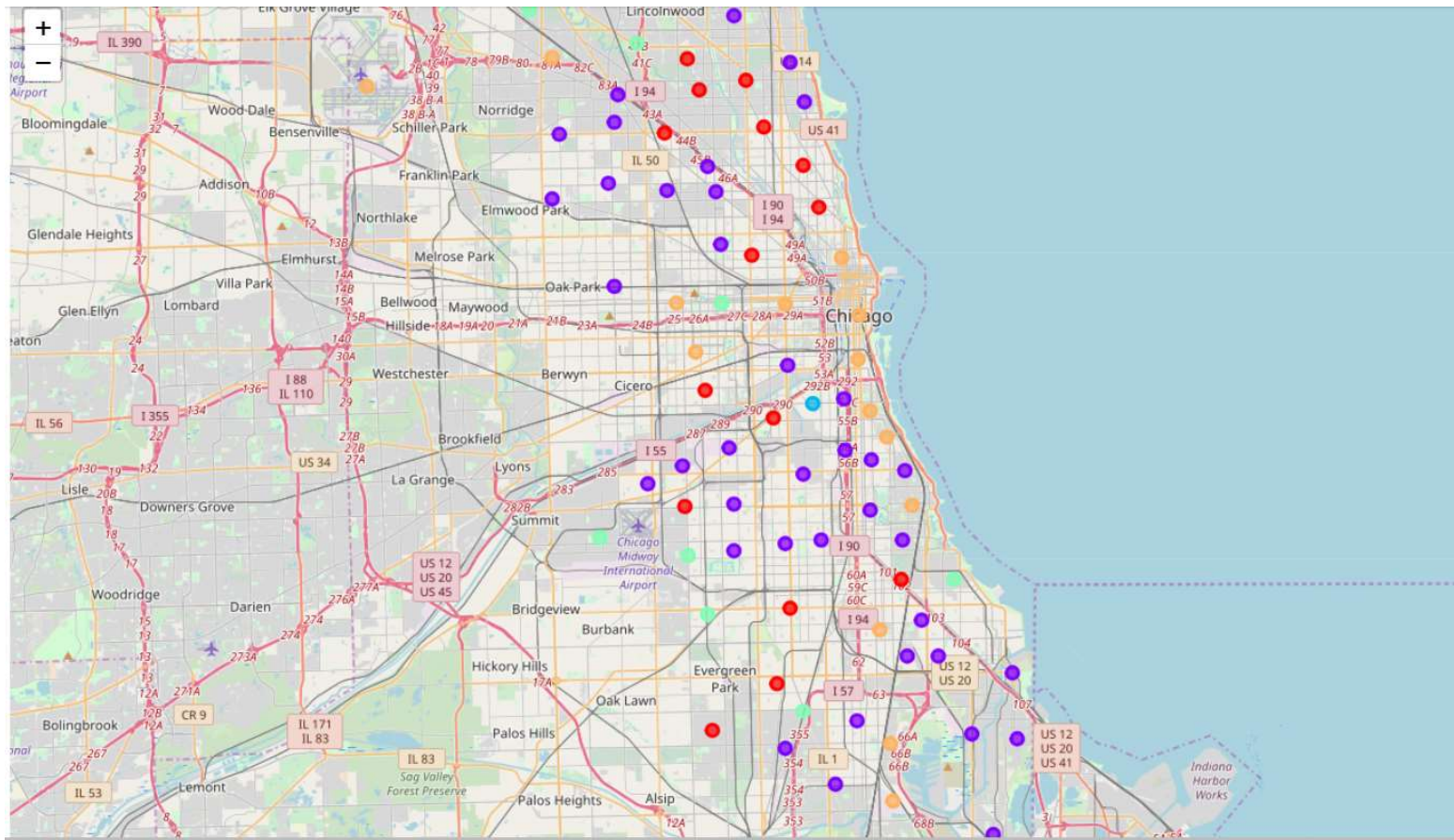
Methodology (cont.)

- The Foursquare API was used to compile 100 locations within 1000 meters of the center of each Community Area
- A dataframe with the mean value of locations per Community Area was created
- *K-means* analysis was used to cluster the 77 Community Areas into 5 distinct clusters, and plotted onto a map of Chicago
- When no discernable pattern was observed, the % increase/decrease in crime was added to the dataframe
- *K-means* was then re-run and re-mapped

Results

- The results showed Community Areas in proximity to downtown were generally in similar clusters
- Three Areas that warrant further analysis are Chatham, Pullman, and Riverdale
  - Many of the Community Areas in Cluster 4 are downtown – in and around The Loop
  - Chatham, Pullman, and Riverdale, however, are all in Cluster 4 but are far into the South Side of the city
- Another Area that stands out is Bridgeport, the only Area that ended up in Cluster 2

Results

Discussion

- I decided to look further into Chatham, Pullman, and Riverdale
- As of 2015, the estimated population of Riverdale was just 7,090, so I decided to exclude that Area from further analysis
- Upon further investigation into Chatham, while the population of the Area has an above-city-average level of education, the Area has a below-city-average level of income per resident
  - Chatham is on the Red Line, the public transit line that goes straight to downtown
  - Reported crime fell ~17% in 2014-2018 versus 2009-2013
  - 60% of the residents are 20-64, with an above-average Unemployment Rate

Conclusion

- To make a real estate and business expansion investment decision, a lot of analysis must be done

- Given this preliminary analysis that highlighted some potential of Chatham, I believe more research is warranted into the type of business that would succeed

- With some of the problems the area has faced with Unemployment, there would be the ability to pick up real estate cheaper here than in other already-developed areas

- While there would likely be higher risk investing in this Area, given an educated underutilized population, there could also be high reward

References

- https://en.wikipedia.org/wiki/Community_areas_in_Chicago
- https://foursquare.com/
- https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2
- https://www.cmap.illinois.gov/documents/10180/126764/Chatham.pdf