Exploratory spatial data analysis of crime in the city of Chicago

Location: City of Chicago

Level of detail: Census tracts

Coordinate Reference System (CRS): NAD 1983 State Plane Illinois East (EPS102671)

Introduction

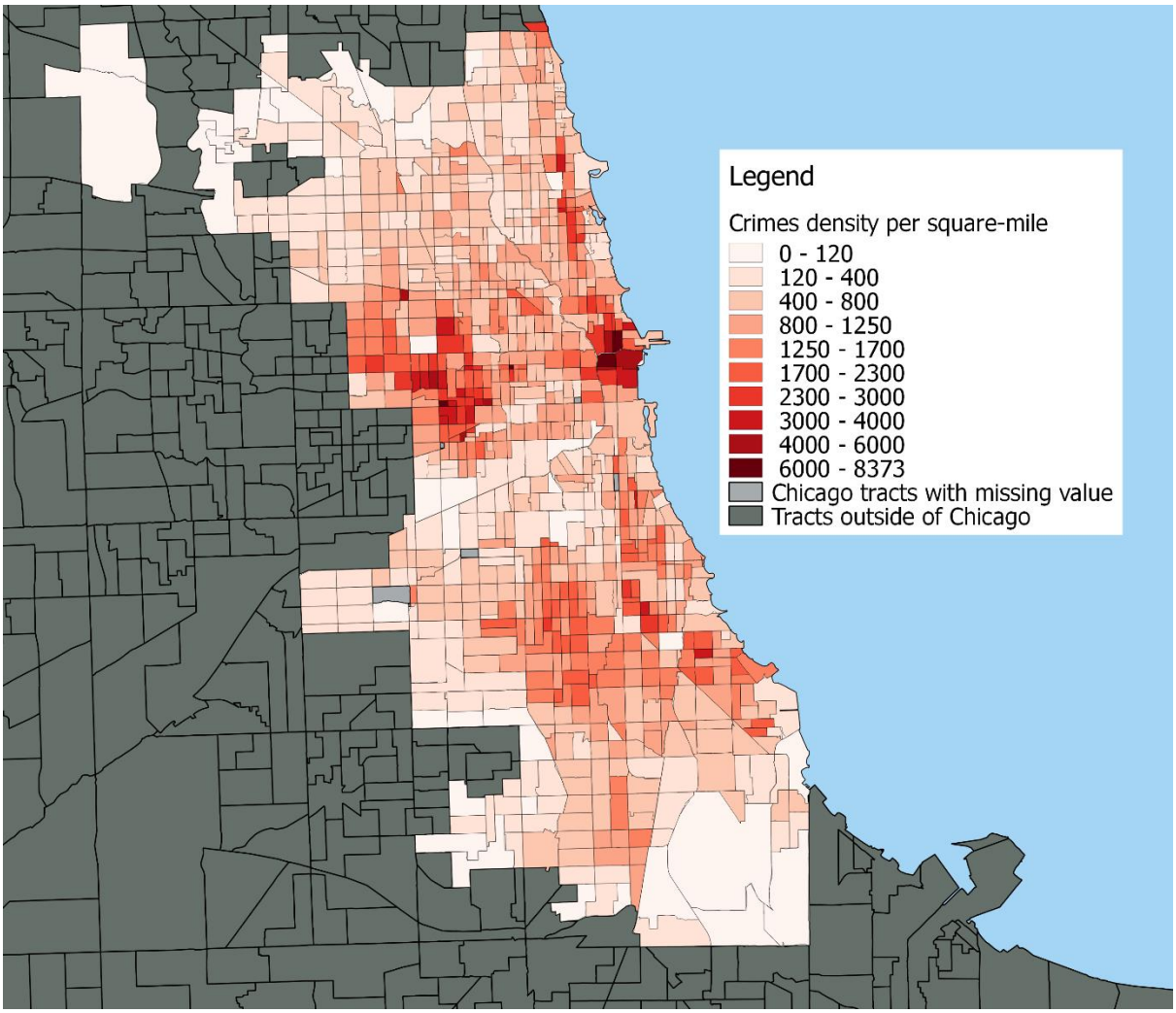
The purpose of this study is to explore crime in Chicago by creating a choropleth map of the crime density to identify potential clusters of criminality, and by using centrographic statistics to better understand the spatial dispersion of selected types of crimes across the city. We will also plot unemployment rate in Chicago to investigate a potential correlation between this variable and crime that may explain the distribution of crime hotspots in the city.

Dataset

I put together the dataset used in this study by joining three files: the first one is a shapefile containing the census tracts of the city of Chicago based on the last census in 2010[1]. The second one is a .csv table which lists all crimes committed in Chicago in 2018 and it is updated live by the Chicago Police Department[2]. The third one is a .csv file from the National Historical Geographic Information System (NHGIS)[3] which includes total workforce and counts of people unemployed in every US census tract as of the latest official survey in 2016. I projected the cloud of crime points on the Chicago tracts shapefile in order to get the number of crimes per census tract in 2018. Then I divided the counts by the surface area of each tract in order to obtain a crime density in “crime / square-mile”. Note that surface areas of census tracts are defined based on

the number of inhabitants who reside in the tract, thus my crime density is equivalent to a count of crime per inhabitants. Afterwards, I joined this set with the NHGIS table containing the number of people unemployed per tract and I divided this feature by the total number of people in the workforce per tract to obtain the unemployment rate per tract. Finally, I obtain a dataset that includes all the Chicago census tracts as of 2010, the crime density per tract using 2018 data to date, and unemployment percentage per tract as of 2016.

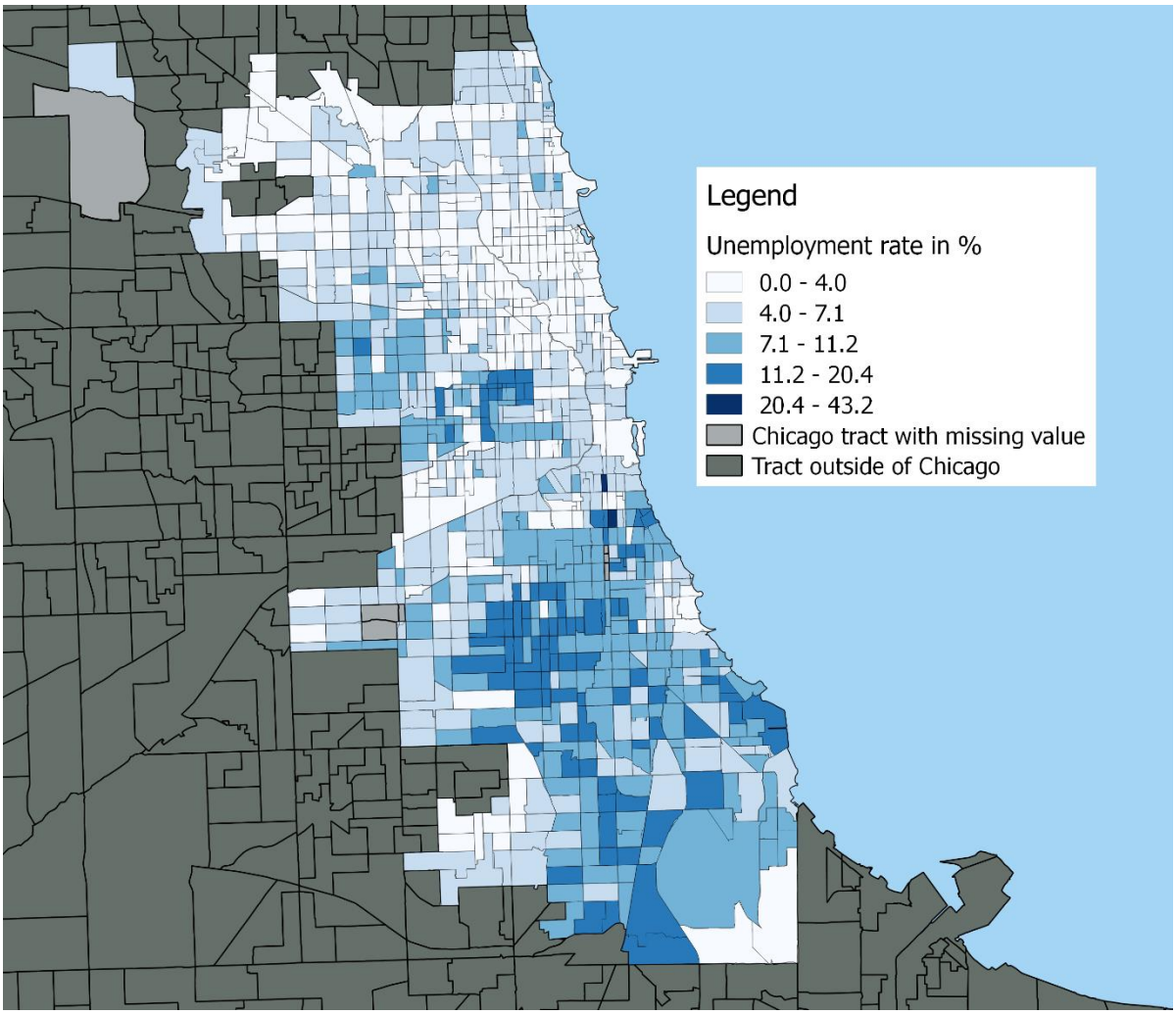
Choropleth map of crime density



The choropleth map above was plotted using the crime density variable. Natural breaks were selected to segment the range of density because they gave the best results in

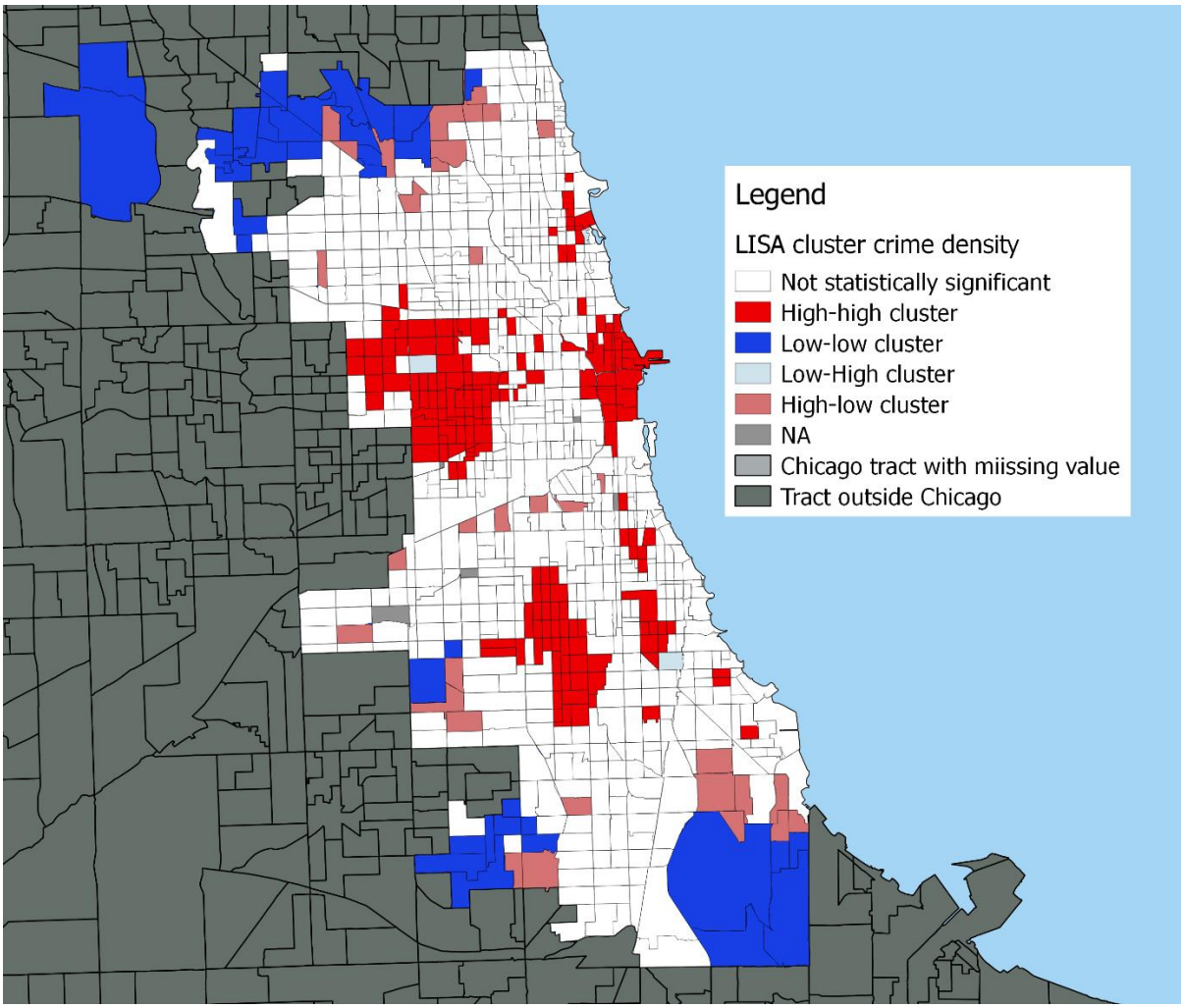
terms of visually identifying areas of high and low criminality. 10 segments were used (more than the standard 5 segments) because the wide range of data spans from 0 to 8373 crimes per square-mile and requires more segments to achieve appropriate level of granularity. The crime map suggests that there are three crime hotspots (clusters of high criminality) in Chicago: one in downtown by the shore, one in the west side of the city and one in the south side. This makes sense because “South Side” is known to be one of the most violent and crime-ridden neighborhood in the whole country and “West Side” is a well-known crime hotspot in Chicago.

Choropleth map of unemployment rate



The choropleth map above was plotted using the unemployment rate variable. We chose unemployment as an explanatory variable because its correlation with criminality is well-known and extensively covered in academic literature. For consistency, we use natural breaks to plot this data but this time only with 5 segments given that the unemployment range is much smaller, from 0% to 43.2%. Unemployment clusters can be seen matching the crime hotspots of West Side and South Side, but the Downtown crime hotspot is invisible on the map. This was expected because Downtown is typically where the highest earners reside, and its cost of living is very high. We may thus draw the conclusion that the types of crime in West Side and South Side must be similar to each other, but both different from the crimes observed downtown. We may further argue that crimes in South Side and West Side are related to the poor economic situation of these neighborhoods while crimes in Downtown are on the contrary triggered by the wealth and dynamics of this area.

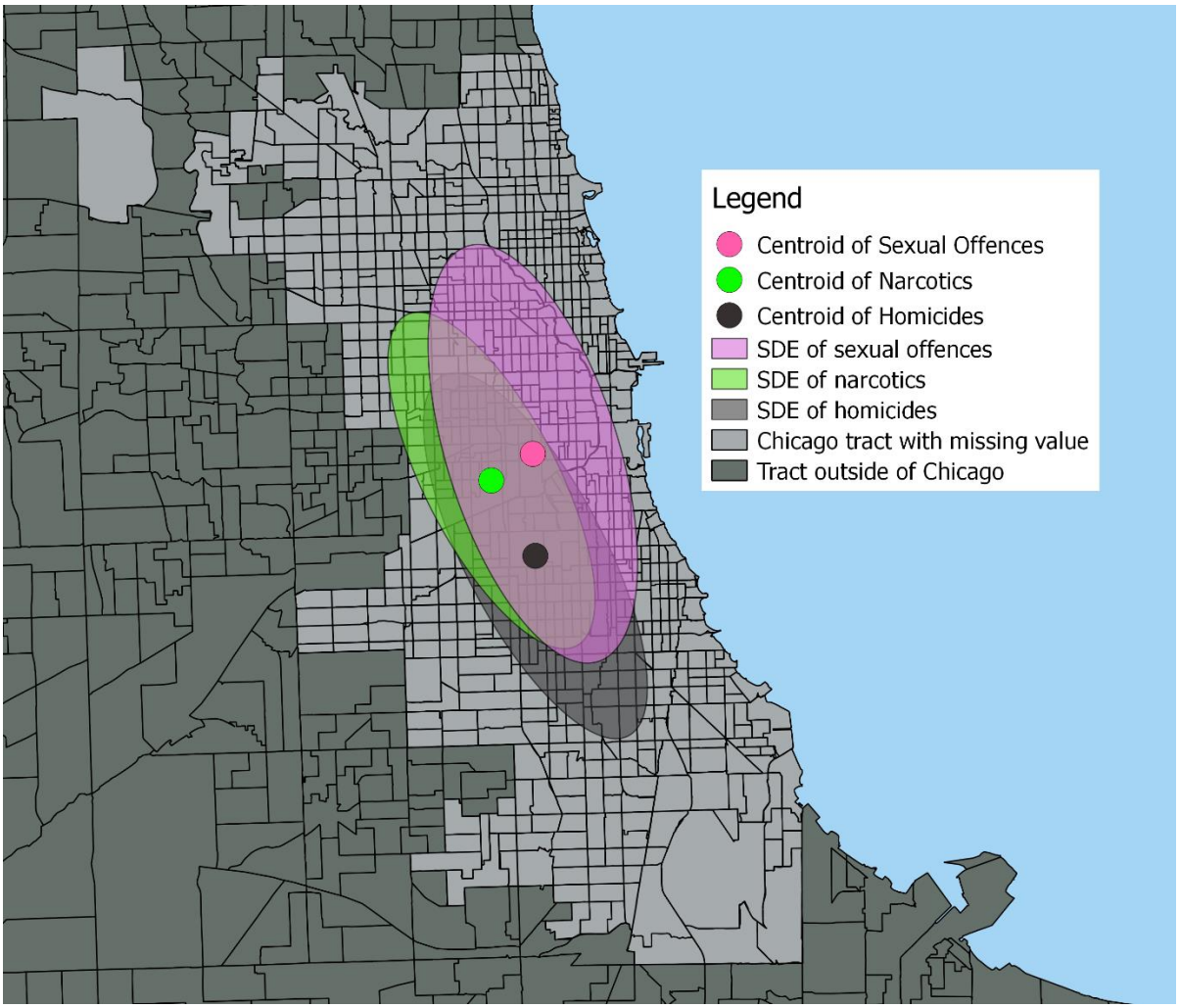
LISA cluster map of crime density



Based on the previous choropleth map of crime we suspect that there are three clusters of crime in Chicago, in West Side, South Side, and Downtown. The LISA map above was plotted using the crime density variable to verify this hypothesis using a queen weight matrix. The matrix was of order two as we believe that the clusters of high criminality have at least two layers of contiguous tracts. The results shown on the map validate our hypothesis as we identify three statistically significant clusters of high criminality located around the west side, south side and downtown area of Chicago. The results also show two clusters of low criminality in Chicago O'Hare's international airport and in Wolf Lake memorial park. This is typically expected because airports are

safe areas where many police officers and private companies ensure safety of travelers, and Wolf Lake area has a park and a golf course which are gated areas with low pedestrian traffic.

Centroids and standard deviational ellipses for three common types of crimes.



The map above shows the location of the centroids of three of the most frequent types of crimes in Chicago in 2018: homicides, sexual offences and narcotics use. These groups are respectively plotted in black, pink and green. Moreover, we added standard deviational ellipses around each group centroid to gain further insights into their dispersion across the city. We observe that the black centroid of homicides is located near South Side which makes sense as gang violence is highly prevalent in this area.

When looking at the pink centroid for sexual offences, we notice a slight shift towards the North East which suggests that these offences tend to occur more frequently near Downtown. We may argue that Downtown nightlife causes more sexual offences, which matches our previous speculation that, in Downtown, wealth and economic dynamic is the root cause of criminality. Finally, we look at the green centroids of narcotics use which is shifted towards the North West near West Side. The shapes and orientations of the three ellipses do not seem to differ from each other and merely seem to follow the outline of the city along the lake.

Resources

[1] : Chicago census tracts 2010

https://www.cityofchicago.org/city/en/depts/doit/dataset/boundaries_-_censustracts.html

[2] : Chicago crime data 2018

https://data.cityofchicago.org/Public-Safety/Crimes-2018/3i3m-jwuy

[3] : 2016 unemployment data per census tracts

https://www.nhgis.org/