



SPARK ASSIGNMENT

Chicago Crime Analysis



FEBRUARY 18, 2020

GROUP C
Marang Mutloatse

Table of Contents

1	Background and Scenario Description	2
1.1	History of Chicago Crime Dataset	2
1.2	Chicago Crime Dataset Description.....	3
1.3	Purpose of analysis.....	4
2	Methodology: Deep dive.....	5
2.1	Time of Day: Violent and Non-violent Crimes	5
2.2	Area of Chicago: Violent and Non-violent Crimes	5
2.3	Top 10 Police Stations for number of arrests due to domestic violence related	6
2.4	Seasonality with robberies	6
2.5	Socio-Economic Profile of Domestic Violence Communities.....	6
3	Conclusions	7
3.1	Time of Day: Violent and Non-violent Crimes	7
3.2	Area of Chicago: Violent and Non-violent Crimes	7
3.3	Top 10 Police Stations for number of arrests due to domestic violence related	7
3.4	Seasonality with robberies	7
3.5	Socio-Economic Profile of Domestic Violence Communities.....	7
4	Future Work	8
	References	9

1 Background and Scenario Description

1.1 History of Chicago Crime Dataset

This dataset reflects reported incidents of crime (except for murders where data exists for each victim) that occurred in the City of Chicago from 2014 to present 2016, minus the most recent seven days [1]. Data is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system.

CLEAR's goal is to become an integrated, state-of-the-art criminal justice information system supporting the work of justice agencies in Chicago and the State of Illinois [2]. Through CLEAR, the Chicago Police Department aims to impact three major functional areas namely police management, criminal justice integration and community/business partnership [1], [2].

Chicago has the highest violent crimes rate in the United States of America. On average, twenty violent crimes are reported each day and many go unreported. During the 2nd period of Obama's presidency, Chicago violent crimes skyrocketed especially crimes domestic violence. Mayor Rahm Emmmanuel tasked Lori Lightfoot, to do a thorough investigation into violent crimes, robberies and domestic-violence related crimes as

Lori Lightfoot became mayor of Chicago on May 20. Before becoming mayor, Lightfoot worked in private legal practice as a partner at Mayer Brown and held various government positions in the City of Chicago. Most notably, she served as president of the Chicago Police Board and chair of the Chicago Police Accountability Task Force. Moreover, Lightfoot is the first black female and first openly gay leader of the city, which became the largest in United States history to have an openly LGBTQ mayor, and again the largest U.S. city to be headed by a woman. As a native of Chicago, Mayor Lightfoot has always been an advocate against women and child abuse. She task's hire's consultants to utilize the abundant data in CLEAR systems and presents analytical results with some recommendations.

1.2 Chicago Crime Dataset Description

Table 1-1: Chicago Crime between Dataset Field descriptions

	Field	Definitions/Descriptions
1	ID	Unique identifier for the record.
2	Case number	The Chicago Police Department RD Number (Records Division Number), which is unique to the incident.
3	Date (year,month,dayofweek,hour)	Date when the incident occurred. This is sometimes a best estimate.
4	Block	The partially redacted address (---xx) where the incident occurred, placing it on the same block as the actual address.
5	Illinois Uniform Crime Reports (IUCR)	The Illinois Uniform Crime Reporting code. This is directly linked to the Primary Type and Description.
6	Primary Type	The primary description of the IUCR code.
7	Description	The secondary description of the IUCR code, a subcategory of the primary description.
8	Location Description	Description of the location where the incident occurred.
9	Arrest	Indicates whether an arrest was made.
10	Domestic	Indicates whether the incident was domestic-related as defined by the Illinois Domestic Violence Act.
11	Beat	Indicates the beat where the incident occurred. A beat is the smallest police geographic area – each beat has dedicated police beat car. 3 to 5 beats make up a police sector, and three sectors make up a police district. The Chicago Police Department has 22 police districts .
12	District	Indicates the police district where the incident occurred.
13	Ward	The ward (City Council district) where the incident occurred.
14	Community Area	Indicates the community area where the incident occurred. Chicago has 77 community areas .
15	FBI Code	Indicates the crime classification as outlined in the FBI's National Incident-Based Reporting System (NIBRS).
16	X- coordinate	The x-coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection . This location is shifted from the actual location for partial redaction but falls on the same block.
17	Y-coordinate	The y-coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection . This location is shifted from the actual location for partial redaction but falls on the same block
20	Latitude	The latitude of the location where the incident occurred.
21	Longitude	The longitude of the location where the incident occurred.

1.3 Purpose of analysis

The main objectives of this project were to:

1. Explore what periods of the day are of concern for violent crimes and non-violent crimes.
2. Determine which side of Chicago have high violent crime and rates.
3. Investigate which ten police stations have the highest number of arrests for domestic violence related crimes. Moreover, for those crimes where domestic violence is attributed to, what
4. Examine the seasonality of robberies over the three-year between 2014-2016.
5. Outline the profile of the constituents most affected by domestic violence-based on key socio-economic factors including percentage living below the poverty level, percentage living in Crowded Housing, percentage of constituents who do not have a High-School Diploma, income per Capita and Unemployment rate.

2 Methodology: Deep dive

Initial basic profiling of the dataset is done whereby the original dataset is enriched with several external tables. The basic profiling is completed to ascertain whether there is any malformed data and if any cleaning should be necessary. Conversely, enriching the dataset with information pertaining to the community area, police districts and type of crime will assist in gaining actionable insights.

2.1 Time of Day: Violent and Non-violent Crimes

Identifying the time of day for a high volume of violent and non-violent crimes is key for police to understand the daily trends and subsequent deployment of necessary resources. Chicago Police dept works on twelve-hour shifts, either 06h00 – 18h00 or 18h00 – 06h00.

As a result for our analysis, by splitting the day into four six-hour bands (00h00 – 06h00; 06h00 – 12h00; 12h00 – 18h00; 18h00 – 00h00) for our analysis, and by enriching the data from the IUCR dataset which splits the type of crime into Indexed (**I**) (Violent) and Non-indexed (**N**) (Non-Violent) crimes. Index/Violent crimes are typically associated with crime such as Assault and Battery. These indexed crimes also have FBI codes: the FBI also has jurisdiction over these crimes and may deploy resources to minimize these crimes.

2.2 Area of Chicago: Violent and Non-violent Crimes

Identifying the areas of Chicago for a high volume of violent and non-violent crimes is key for Mayor Rahm Emmanuel and the Chicago legislature to understand the dividing of the police budget. City council member approve budgets, but the finances are split into “North Side”, “South Side” and “Central”.

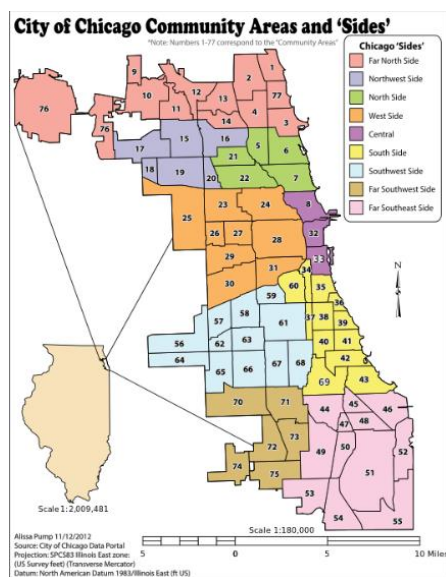


Figure 2-1: Chicago Community Area

By grouping each community area, then subsequently making filtering only the domestic-violence related crimes where there had been arrests, we calculated the conviction rate as rate of all crimes.

2.3 Top 10 Police Stations for number of arrests due to domestic violence related

Obtaining the top 10 Police stations to calculate the reports of domestic-violence related crimes and the conviction rate is vital in assessing the accountability of the police station. By enriching the data with police district and department name, then subsequently making filtering only the domestic-violence related crimes where there had been arrests, we calculated the conviction rate as rate of all crimes for each department and sorted from the highest to lowest.

2.4 Seasonality with robberies

We generated the season's from the data according to months whereby (1 December – 28 Feb : Winter; 1 March – 31 May: Spring; 1 June - 31 August: Autumn; 1 September – 30 November: Spring). Then by filtering out the primary-type of crime, we obtained the seasonality for the three year period.

2.5 Socio-Economic Profile of Domestic Violence Communities

Socio-Economic census data for the period between 2014 – 2016 was enriched into the original data as census was according to community area. The general statistics including minimum, maximum, average appended to crimes where domestic violence crime had occurred, and these were grouped according to the community area sides as previously developed.

3 Conclusions

Based on the results obtained, the following conclusions can be made:

3.1 Time of Day: Violent and Non-violent Crimes

- Over the period between 2014 -2016 at least 65 % of violent crimes occurred in the evening and early morning.

3.2 Area of Chicago: Violent and Non-violent Crimes

- At least 55 % of violent crimes occurred in the South Side of Chicago.

3.3 Top 10 Police Stations for number of arrests due to domestic violence related

Below are results.

districtName	Domestic Violence Arrest Rate	Domestic Violence Cases/Reported Crimes
South Chicago	384	1.28
Grand Central	345	1.15
Gresham	335	1.12
Harrison	317	1.06
Calumet	309	1.03
Englewood	304	1.01
Chicago Lawn	286	0.95
Austin	286	0.95
Grand Crossing	276	0.92
Ogden	263	0.88

- South Chicago, Grand Central, Gresham, Harrison and Calumet have the highest arrests due to domestic violence related crimes.

3.4 Seasonality with robberies

- Summer has the highest number of robberies, with Autumn in second, Winter in third and Spring in last place.

3.5 Socio-Economic Profile of Domestic Violence Communities

Moreover, there are clear ties between economic indicators and crime, a fact that should be considered when developing perceptions about the amount of crime within Chicago.

4 Future Work

Based on these conclusions, the following recommendations are made:

In addition, more work should be completed to truly understand the relationship between crime and per capita income of a community area. The variability, or lack of variability, of these relationships suggests that these associations are more complicated than what we have uncovered. However, there are clear ties between economic indicators and crime, a fact that should be considered when developing perceptions about the amount of crime within Chicago

References

- [1] "Chicago Data Portal," 2010. [Online]. Available: <https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data>. [Accessed: 08-Nov-2019].
- [2] W. G. Skogan *et al.*, "Policing Smarter Through IT: Learning from Chicago' s Citizen and Law Enforcement Analysis and Reporting (CLEAR) System," *Inst. Policy Res. Northwest Univ.*, 2003.