# Capstone project

Crime patterns in Chicago

# Introduction

The aim of analysis was to provide answers to below questions and explain the difference between overall crimes and homicides in Chicago including the following aspects:

- When crimes are committed?
- Which community areas are the most dangerous? (number of crimes per 1000 inhabitants)
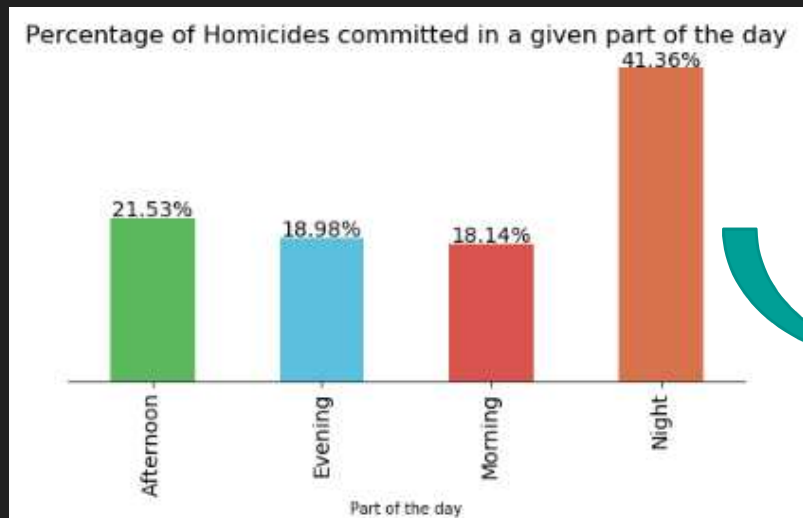- Are police departments properly located? (Foursquare location)

# Data

The sources of data are official websites:

- https://data.cityofchicago.org – juxtaposition of crimes in 2018
- https://www.chicago.gov – Chicago census data

Data were loaded from csv files using pandas package. Both dataframes were combined and cleansed from missing data.

# When crimes are committed?



Percentage of Homicides committed in a given part of the day

- Slight difference between seasons(summer 28% vs winter 22 %)
- Significant relation for time of the crime (homicides only)

# Which community areas are the most dangerous?

| | Community Area | Community Area Description | Community Area Population | Number of Homicides | Crime Rate | Crime Rate Normalized | Postion in Total Crimes | Community Area Homicide Cluster |
|---|---|---|---|---|---|---|---|---|
| 1 | 25 | Austin | 98514 | 56 | 0.57 | 0.64 | 2 | 3 |
| 2 | 54 | Riverdale | 6482 | 13 | 2.01 | 0.61 | 20 | 3 |
| 3 | 26 | West Garfield Park | 18001 | 26 | 1.44 | 0.58 | 5 | 3 |
| 4 | 68 | Englewood | 30654 | 33 | 1.08 | 0.56 | 12 | 3 |
| 5 | 29 | North Lawndale | 35912 | 32 | 0.89 | 0.50 | 3 | 3 |
| 6 | 27 | East Garfield Park | 20567 | 22 | 1.07 | 0.45 | 10 | 3 |
| 7 | 53 | West Pullman | 29651 | 26 | 0.88 | 0.44 | 22 | 3 |
| 8 | 67 | West Englewood | 35505 | 23 | 0.65 | 0.36 | 8 | 2 |
| 9 | 43 | South Shore | 49767 | 25 | 0.50 | 0.34 | 7 | 2 |
| 10 | 61 | New City | 44377 | 23 | 0.52 | 0.33 | 27 | 2 |
| 11 | 49 | Roseland | 44619 | 22 | 0.49 | 0.31 | 15 | 2 |
| 12 | 71 | Auburn Gresham | 48743 | 23 | 0.47 | 0.31 | 14 | 2 |
| 13 | 40 | Washington Park | 11717 | 10 | 0.85 | 0.29 | 17 | 2 |
| 14 | 23 | Humboldt Park | 56323 | 21 | 0.37 | 0.27 | 13 | 2 |
| 15 | 66 | Chicago Lawn | 55628 | 20 | 0.36 | 0.26 | 21 | 2 |

The level of danger for community areas is different for homicides and total crimes . "Loop" which is at first place for all crimes is out of the top 15 in case of homicides. The highest homicide rate is in Austin which is 2nd in total crimes. The second place at homicide list is for "Riverdale" although it has the middle rate of total crimes. The community areas were clustered into 4 groups (low risk, medium risk, high risk and very high risk).
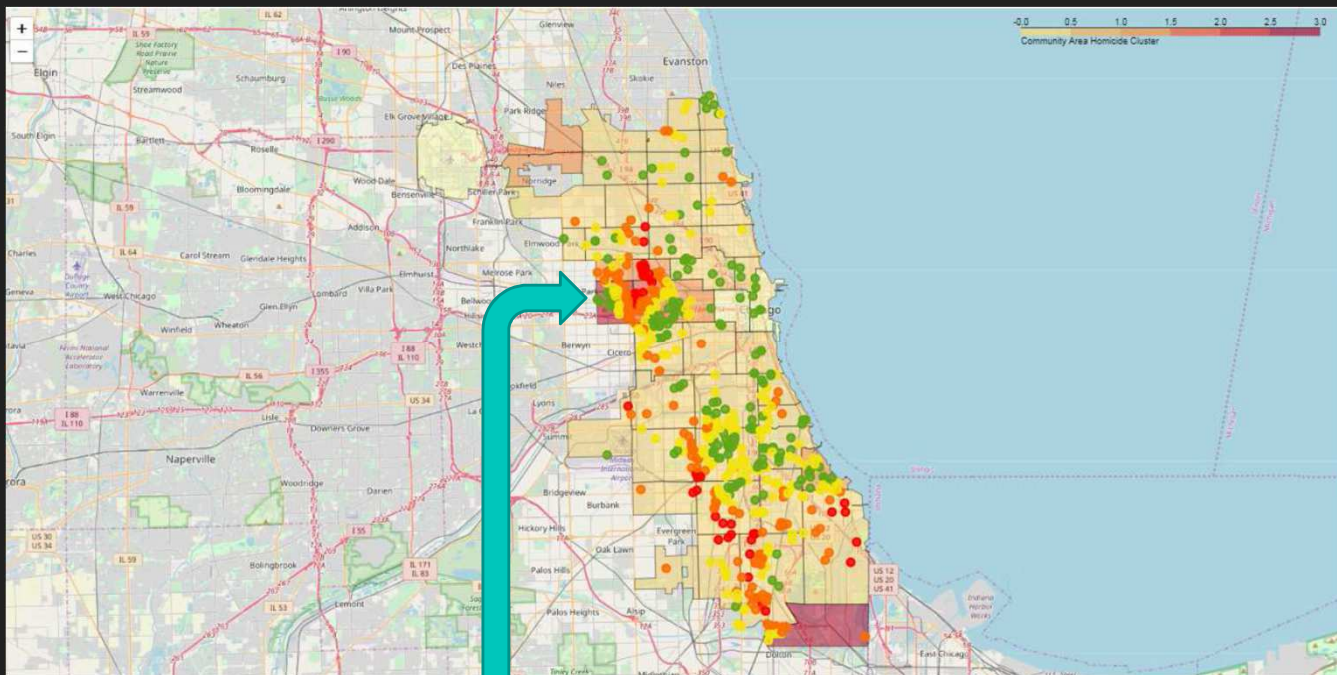
# Are police departments properly located?

| | Case Number | Season | Part of the day | Primary Type | Description | Block | Location Description | Arrest | Beat | District | Community Area | Community Area Description | Community Area Population | Latitude | Longitude | Community Area Homicide Cluster | Distance [m] | Police Department Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | JC476281 | Fall | Evening | HOMICIDE | FIRST DEGREE MURDER | 037XX N ASHLAND AVE | AUTO | False | 1922 | 19 | 6 | Lake View | 94368 | 41.949456 | -87.668921 | 0 | 1470.0 | 1 |
| 1 | JB156223 | Winter | Morning | HOMICIDE | FIRST DEGREE MURDER | 077XX S KILBOURN AVE | STREET | True | 834 | 8 | 70 | Ashburn | 41081 | 41.751973 | -87.734194 | 0 | 2388.0 | 2 |
| 2 | JB164505 | Winter | Night | HOMICIDE | FIRST DEGREE MURDER | 000XX N MAYFIELD AVE | STREET | True | 1513 | 15 | 25 | Austin | 98514 | 41.880996 | -87.771819 | 3 | 265.0 | 0 |
| 3 | JB534354 | Fall | Afternoon | HOMICIDE | FIRST DEGREE MURDER | 009XX N MONTICELLO AVE | STREET | True | 1112 | 11 | 23 | Humboldt Park | 56323 | 41.898270 | -87.717625 | 2 | 2068.0 | 2 |
| 4 | JB407292 | Summer | Night | HOMICIDE | FIRST DEGREE MURDER | 005XX E 47TH ST | STREET | True | 221 | 2 | 38 | Grand Boulevard | 21929 | 41.809437 | -87.613093 | 2 | 1674.0 | 2 |

| | min_Distance | max_Distance | Police Department Cluster | Police Department Cluster new |
|---|---|---|---|---|
| 0 | 28.0 | 941.0 | 0 | 0 |
| 1 | 955.0 | 1646.0 | 2 | 1 |
| 2 | 1665.0 | 2412.0 | 1 | 2 |
| 3 | 2460.0 | 4123.0 | 3 | 3 |

Each homicide site was assigned to one of four clusters based on distance between site and police station (short, medium, distant, very distant).

# Are police departments properly located?



It was found that there was significant number of distant sites at the high risk area at the middle western part of the city.

# Recommendation

The recommendation is to increase the number of police patrols at night in 27 specific blocks defined by below addresses. The blocks were selected according to the following rules:

- ○ - 19 out of 27: community areas cluster-2 and police department cluster-3

  042XX W HIRSCH ST , 073XX S BELL AVE , 010XX N PULASKI RD , 042XX W HADDON AVE , 013XX N KILDARE AVE , 072XX S WESTERN AVE , 002XX W 95TH ST , 088XX S PAULINA ST , 009XX N KEDVALE AVE , 041XX W POTOMAC AVE , 012XX N KEELER AVE , 097XX S HARVARD AVE , 023XX W 72ND ST , 073XX S CLAREMONT AVE , 042XX W AUGUSTA BLVD , 007XX N SPRINGFIELD AVE , 111XX S NORMAL AVE ,   008XX N HARDING AVE , 012XX N KILDARE AVE

- ○ - 8 out of 27: community areas cluster-3 and police department cluster-3

  045XX W ADAMS ST , 045XX W WEST END AVE , 042XX W CARROLL AVE , 122XX S STATE ST , 001XX E 124TH ST , 045XX W LAKE ST , 044XX W WEST END AVE , 119XX S LA SALLE ST