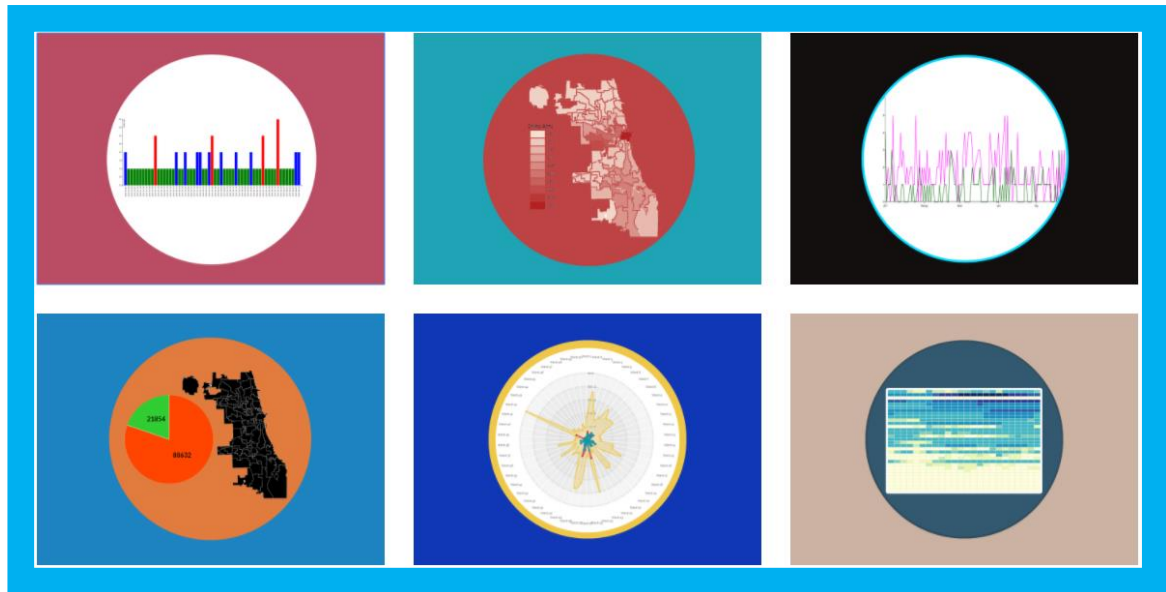


# CHICAGO CRIMES

## INFORMATION VISUALIZATIONS

### PROJECT

---



---

**Rami Bishara: 205721525**

**Orhai Cohen: 038061065**

# WHAT

## - **Background:**

Today's reality crime rate is increasing day by day, we would like to visualize the data in a way that the site's observer/visitor can learn better on how to be more careful around the city and be enriched of today's crimes in Chicago city.

## - **About the data:**

Our dataset contains the information about reported incidents of multiple varied types of crimes (with the exception of murders where data exists for each victim) that occurred in the city of Chicago from the year 2017 to present.

Data is extracted from the Chicago police department's Citizen Law Enforcement Analyzes and Reporting system.

For privacy protection, the addresses are shown at the block level only and specific locations are not identified.

## - **Data Structure:**

The data is a big table of crime incidents of Chicago.

Each incident is described by the crime location, time, date and the crime type.

- Each row contains information about a specific incident.
- Each column contains information of a specific attribute for all the crimes.
- Timeline: 1.1.2017 – 13.6.2017
- Attributes (that we used):
  - Date, primary type, arrest, domestic, ward, latitude, longitude.
  - \* The complete table includes additional attributes that we didn't use.
- The data contains 110486 records of incidents.

ID	Case Number	Date	Block	LCR	Primary Type	Description	Location Description	Arrest	Domestic
1	11053463	08/13/2017 11:50:00 PM	047XX N BROADWAY	0460	BATTERY	SIMPLE	SIDEWALK	✓	✓
2	11053533	08/13/2017 11:49:00 PM	016XX W SHERWIN AVE	0460	BATTERY	SIMPLE	RESIDENCE PORCH/HALLWAY		
3	11053452	08/13/2017 11:48:00 PM	002XX E PEARSON ST	1310	CRIMINAL DAMAGE	TO PROPERTY	RESIDENCE		
4	11053464	08/13/2017 11:48:00 PM	078XX S SOUTH SHORE DR	0498	BATTERY	AGGRAVATED DOMESTIC BATTERY/HAM	RESIDENCE	✓	✓
5	11053426	08/13/2017 11:48:00 PM	020XX W 70TH ST	0420	BATTERY	AGGRAVATED KNIFE/CUTTING INSTR	STREET		✓
6	11053445	08/13/2017 11:45:00 PM	021XX W RANDOLPH ST	0498	BATTERY	DOMESTIC BATTERY SIMPLE	STREET		✓
7	11053472	08/13/2017 11:40:00 PM	003XX W GRAND AVE	0498	BATTERY	DOMESTIC BATTERY SIMPLE	SIDEWALK	✓	✓
8	11053447	08/13/2017 11:40:00 PM	005XX E MARQUETTE RD	0560	ASSAULT	SIMPLE	APARTMENT		
9	11053496	08/13/2017 11:30:00 PM	033XX S COTTAGE GROVE A	0520	THEFT	\$500 AND UNDER	VEHICLE NON-COMMERCIAL		
10	11053826	08/13/2017 11:30:00 PM	033XX N HALSTED ST	0320	ROBBERY	STRONGARM - NO WEAPON	ALLEY		
11	11053436	08/13/2017 11:30:00 PM	030XX W PALMER ST	0810	THEFT	OVER \$500	VEHICLE NON-COMMERCIAL		
12	11053856	08/13/2017 11:30:00 PM	050XX W DICKENS AVE	1320	CRIMINAL DAMAGE	TO VEHICLE	DRIVEWAY - RESIDENTIAL		
13	11053450	08/13/2017 11:30:00 PM	023XX E 71ST ST	0810	THEFT	OVER \$500	STREET		
14	11054577	08/13/2017 11:30:00 PM	055XX S LAFUN ST	0820	THEFT	\$500 AND UNDER	STREET		
15	11054602	08/13/2017 11:30:00 PM	020XX W 70TH ST	1310	CRIMINAL DAMAGE	TO PROPERTY	RESIDENCE		✓
16	11053478	08/13/2017 11:23:00 PM	021XX W RANDOLPH ST	0498	BATTERY	DOMESTIC BATTERY SIMPLE	STREET	✓	
17	11053444	08/13/2017 11:21:00 PM	084XX S ESCANABA AVE	0520	ASSAULT	AGGRAVATED KNIFE/CUTTING INSTR	RESIDENCE		
18	11053431	08/13/2017 11:20:00 PM	000XX W 99TH ST	0498	BATTERY	DOMESTIC BATTERY SIMPLE	RESIDENCE	✓	✓
19	11053548	08/13/2017 11:19:00 PM	071XX S JEFFERY BLVD	0498	BATTERY	DOMESTIC BATTERY SIMPLE	APARTMENT	✓	
20	11053483	08/13/2017 11:17:00 PM	012XX W VICTORIA ST	4367	OTHER OFFENSE	VIOLATE ORDER OF PROTECTION	APARTMENT		✓
21	11053467	08/13/2017 11:15:00 PM	018XX W CERMAK RD	0498	BATTERY	DOMESTIC BATTERY SIMPLE	APARTMENT	✓	✓

\*Data as shown in the original database of the Chicago city police department website.

## - **Data manipulation:**

We almost didn't change the data at all, we used it as is, except for deleting a few records that contained NULL elements.

# WHY

- Trying to show correlation between two types of crimes.
- Dangerous wards – trying to map the area of Chicago to find out whether it's risky to go around the city in a specific ward (by the indication of all crimes or by a specific crime).
- Dangerous hours of day shown for each ward.
- Showing the crime volume trends as time goes from 1.1.2017 to present.
- Ranking the wards based on ranking system which calculates how dangerous a ward is (the rank is based on the crime type, whether the police arrested the criminal and if the crime was domestic or not).
- Showing the crime volume by color for each hour of the day. The darkest color indicates highest level of crime occurrences and lightest color indicates the lowest level of crime occurrences.

## **Why is it interesting?**

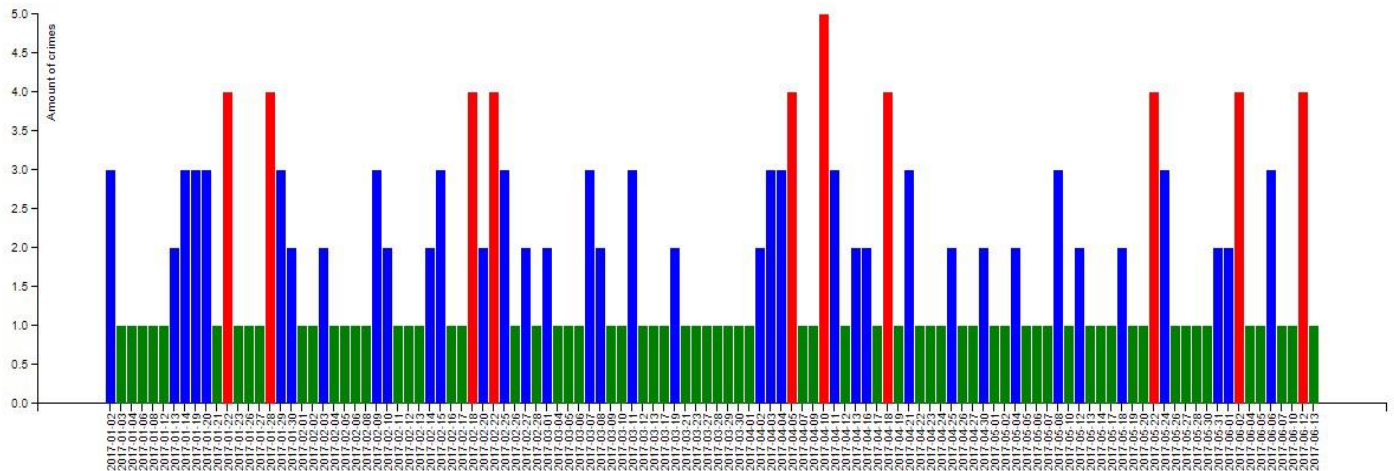
- The visualizations simplify the data in a way that the user can comprehend what 110486 incidents can look like just by looking at one or more pictures and without the need to go through all the data.

## **People who would benefit or could be interested in our visualization:**

- Tourists – upon arriving to Chicago they can tell what are the safer places to visit or the safer hours to be outside.
- Journalists – based on the data, journalists can rely on the data to prove their claims, for example: to prove how efficient the police are and how dangerous a ward is, based on multiple crimes or a specific crime.
- Mayors – The Chicago's mayor may use the information we are visualizing, by emphasizing whether there is increment/decrement in crime rate in the city's life experience. Assist the police with beneficial information about which wards should be taken care of based on the visualizations we are showing.
- Local citizens – can learn about the best hours to be outside and the problematic wards.
- Researchers – can use the visualizations in their articles to prove what they want to write about in their studies.
- Police - can learn where to reinforce the forces in order to decrease the crime rate around the city.  
Can reward a police precinct with the lowest crime rate.  
Can help investigations by showing the relations between two types of crime.

# HOW

- Crime rate increment / decrement through time for each ward and crime.



Type of the graph: Bar Chart.

X-Axis: Date formatted by Year-Month-Day.

Y-Axis: Amount of crimes occurred on a certain date.

We can choose two variables, the crime type and the ward number to filter the results of the visualization.

Crime type:

Ward number:

The dates are ordered from the earliest (left) to the latest (right). And also categorized by three categories.

High  
Medium  
Low

We calculated the Maximum amount of crimes of all days.

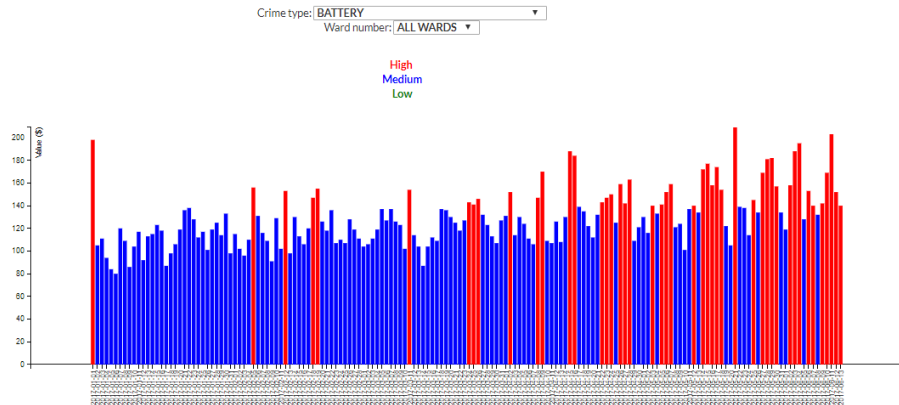
High  $\geq \frac{2}{3} * \text{Maximum}$ .

$\frac{2}{3} * \text{Maximum} > \text{Medium} \geq \frac{1}{3} * \text{Maximum}$ .

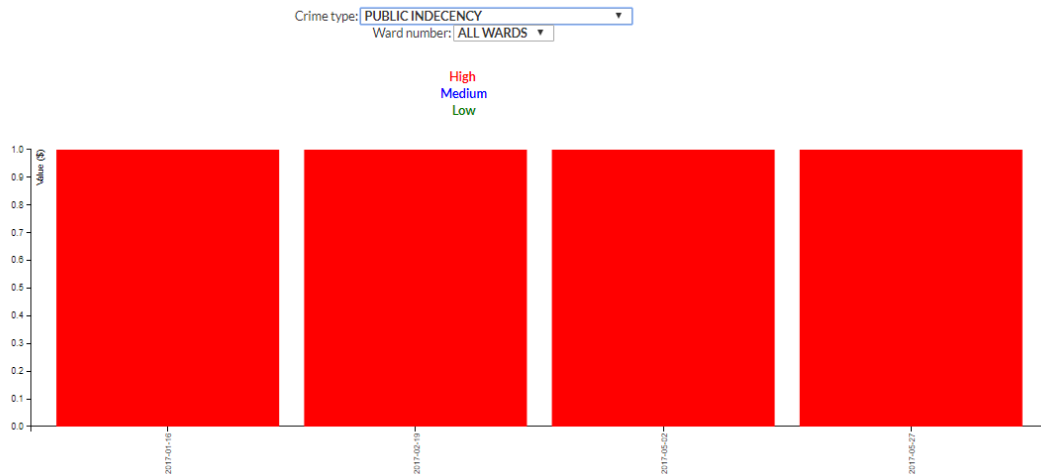
Low  $< \frac{1}{3} * \text{Maximum}$

## What can we learn?

The following example shows the increment of the amount of the battery crime type in all wards. We can tell that we started around 100 crimes at the beginning of the year and ended up with around 170 crimes at the middle of June.

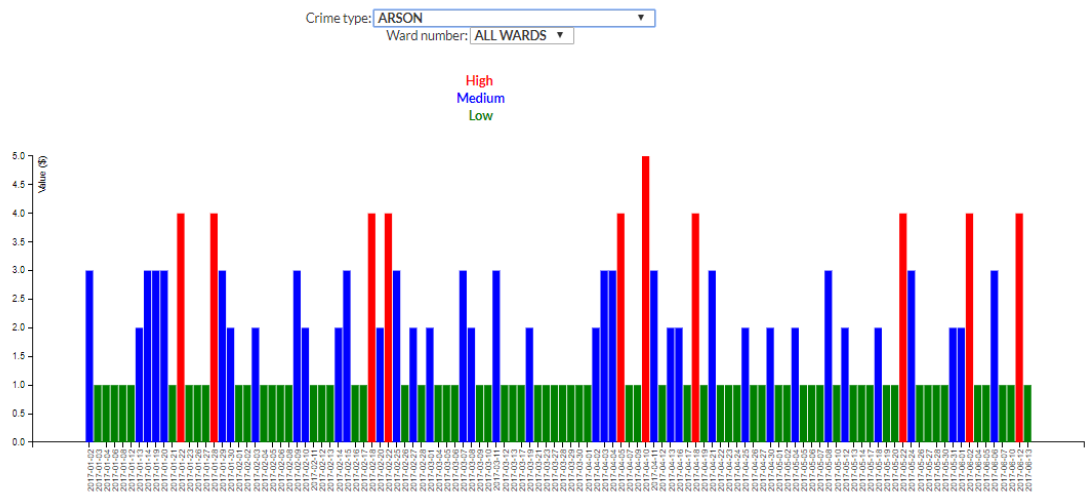


**An exceptional example would be**, in the case of “Public Indecency” for all wards, we can’t really learn anything significant because they were only four cases since the beginning of 2017.



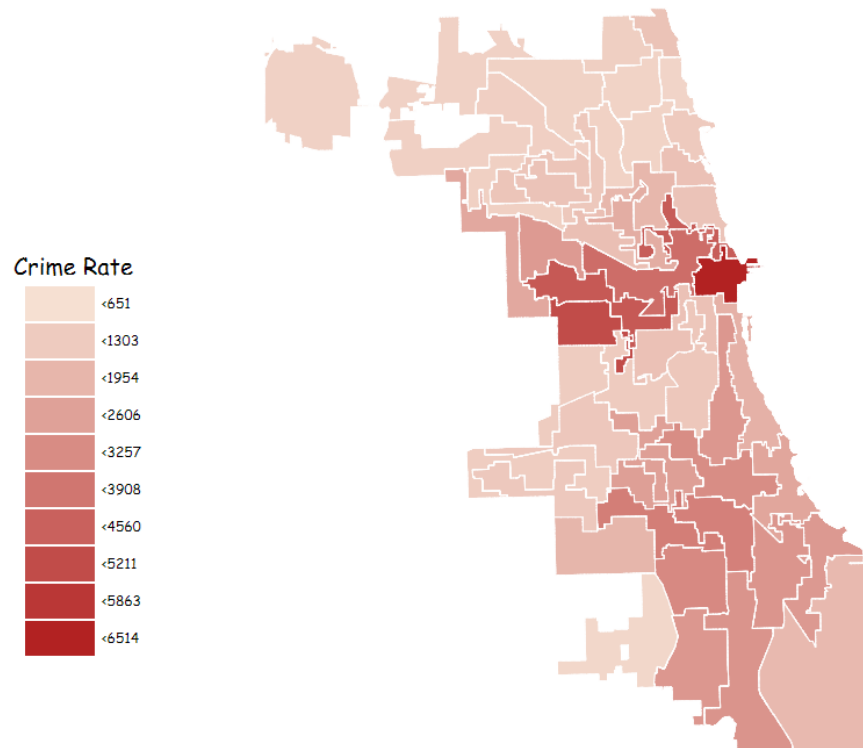
## Disadvantages

- If there are too many crimes, the dates would be very crowded that it is hard to read them.
- In this case, we cannot really tell whether there was an increment or decrement in crimes, so it is inconclusive to determine.



**A Suggestion:** If we have used an average line on this graph, we could have told the trend behind.

- Showing crimes occurrences for all the map by visualizing all crimes types differed by color.



Type of the graph: Choropleth Map.

The map describes crime occurrences by color intensity, each ward may be colored with a range of colors [lightest, darkest]. The lightest color indicates for low rate of crime and the darkest color indicates for high rate of crime.

The legend shows the quantities that each color indicates, in total there are 10 colors.

The map is zoomable and movable so the user can tell the difference between the boundaries of each ward in a better way.

When you point the mouse on a specific ward, the ward's number will be shown near the ward area so you can tell which ward you are looking at.

### **Filtering:**

- ☒ **ALL CRIME TYPES**   ☐ THEFT   ☐ BATTERY   ☐ DECEPTIVE PRACTICE   ☐ ASSAULT  
☐ OTHER OFFENSE   ☐ CRIMINAL DAMAGE   ☐ ROBBERY   ☐ MOTOR VEHICLE THEFT  
☐ OFFENSE INVOLVING CHILDREN   ☐ BURGLARY   ☐ PROSTITUTION  
☐ CRIMINAL TRESPASS   ☐ CRIM SEXUAL ASSAULT   ☐ HOMICIDE   ☐ ARSON  
☐ WEAPONS VIOLATION   ☐ PUBLIC PEACE VIOLATION   ☐ NARCOTICSE  
☐ LIQUOR LAW VIOLATION   ☐ SEX OFFENSE   ☐ GAMBLING   ☐ STALKING  
☐ INTERFERENCE WITH PUBLIC OFFICER   ☐ KIDNAPPING   ☐ INTIMIDATION  
☐ CONCEALED CARRY LICENSE VIOLATION   ☐ OBSCENITY   ☐ HUMAN TRAFFICKING  
☐ NON-CRIMINAL   ☐ PUBLIC INDECENCY   ☐ OTHER NARCOTIC VIOLATION  
☐ NON-CRIMINAL (SUBJECT SPECIFIED)

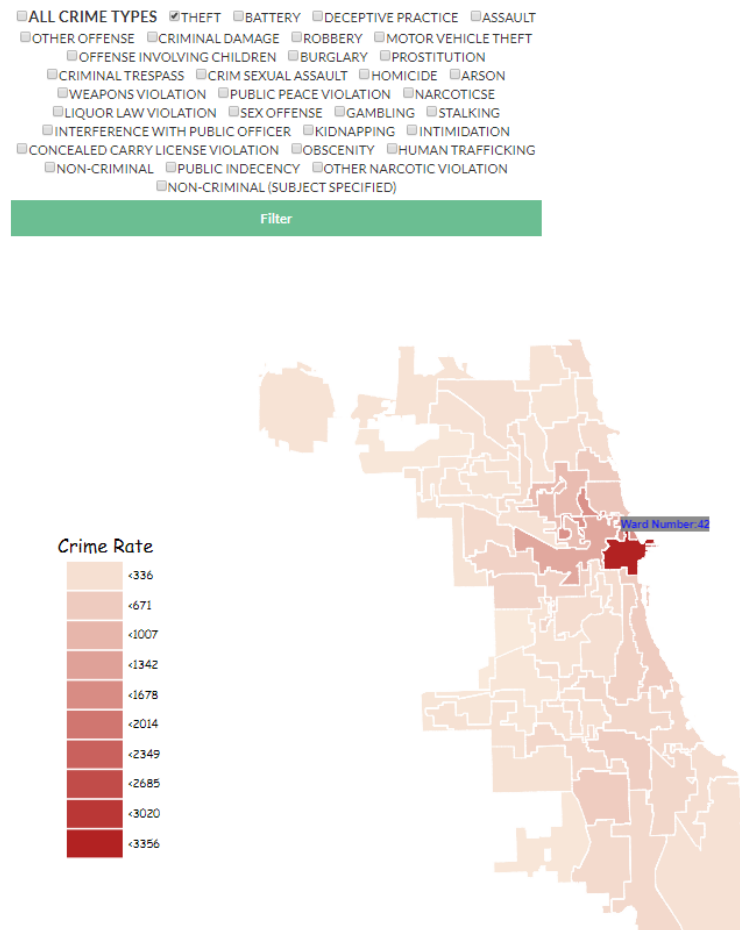
Filter

You can choose multiple crime types to see the correlation between them on the map, or you can choose all crime types (When “ALL CRIME TYPES” option is checked, all the other boxes will be unchecked for obvious reasons).



## What can we learn?

The following example, shows the “popularity” of theft on ward 42 related to all the other wards of Chicago which relatively have a lower theft rate. In conclusion, the center of theft is in ward 42, on the same principle, we can learn centers of other types of crimes, for example: The center of prostitution is in ward 24 and the center of gambling is in ward 28.

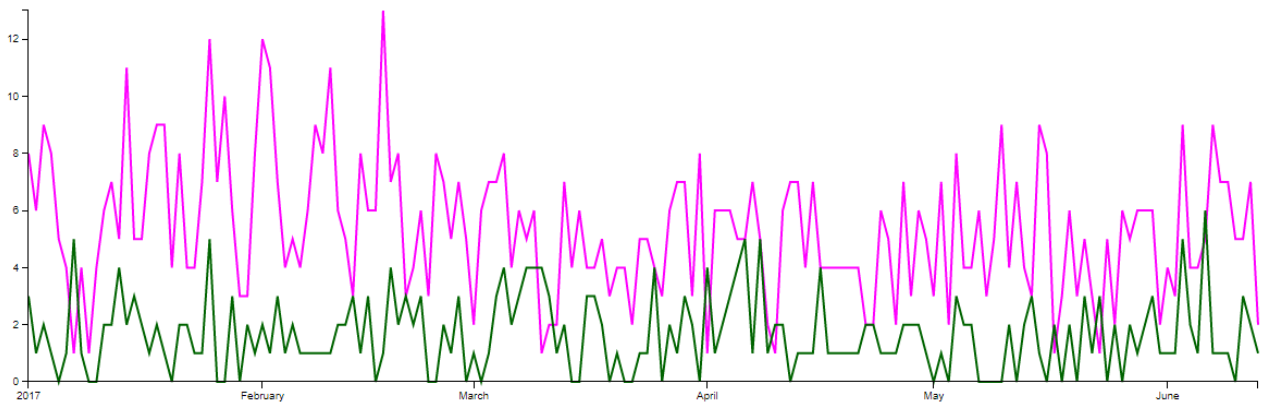


## Disadvantages

The visualization doesn't consider the importance of wards with low population to area ratio. For example, ward 10 has a low level of crime but it is not necessarily safer there, since the ward's population is very low related to its area.

In addition, wards are zoomable but you can't really see which crime happened on a certain street, nor you can see the streets on the ward itself.

- Showing the relation between two types of crime by showing all the crimes that happened in the same street on the map.



Type of the graph: Line Chart.

X-Axis: Month of crime occurrence.

Y-Axis: Amount of crimes occurred on a specific date.

We can choose three variables: first crime type, second crime type and the ward number, to filter the results of the visualization.

First Crime type:

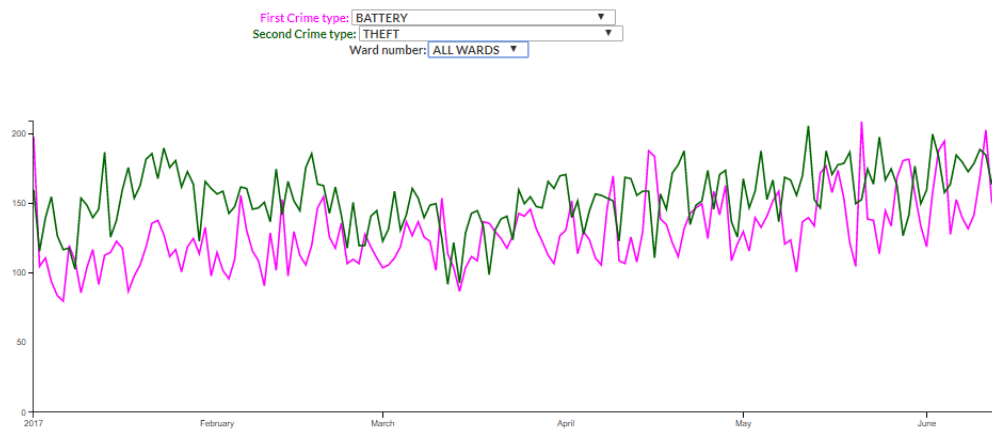
Second Crime type:

Ward number:

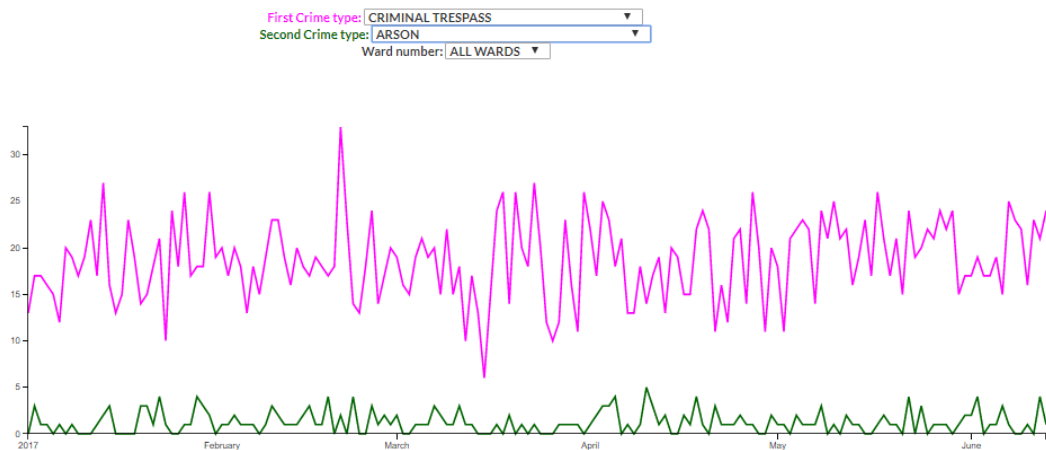
First crime type is represented by the pink color and the second crime type is represented by the green color.

## What can we learn?

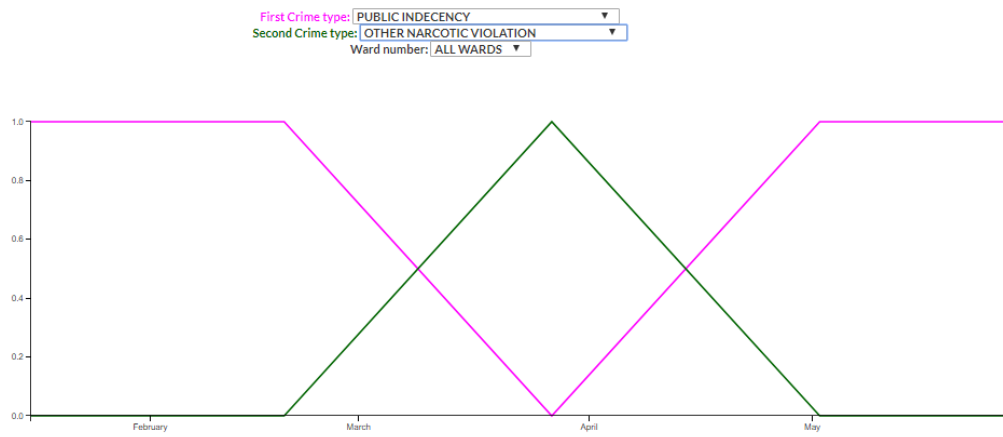
We can see that all over Chicago, choosing “ALL WARDS”, there is a high correlation between battery and theft. It’s also true per specific ward.



Another example to show for “ALL WARDS” option, that there is no correlation between criminal trespass and arson, it is also true per specific ward and it also means that each ward that has high criminal trespass rate is likely to have low arson rate.



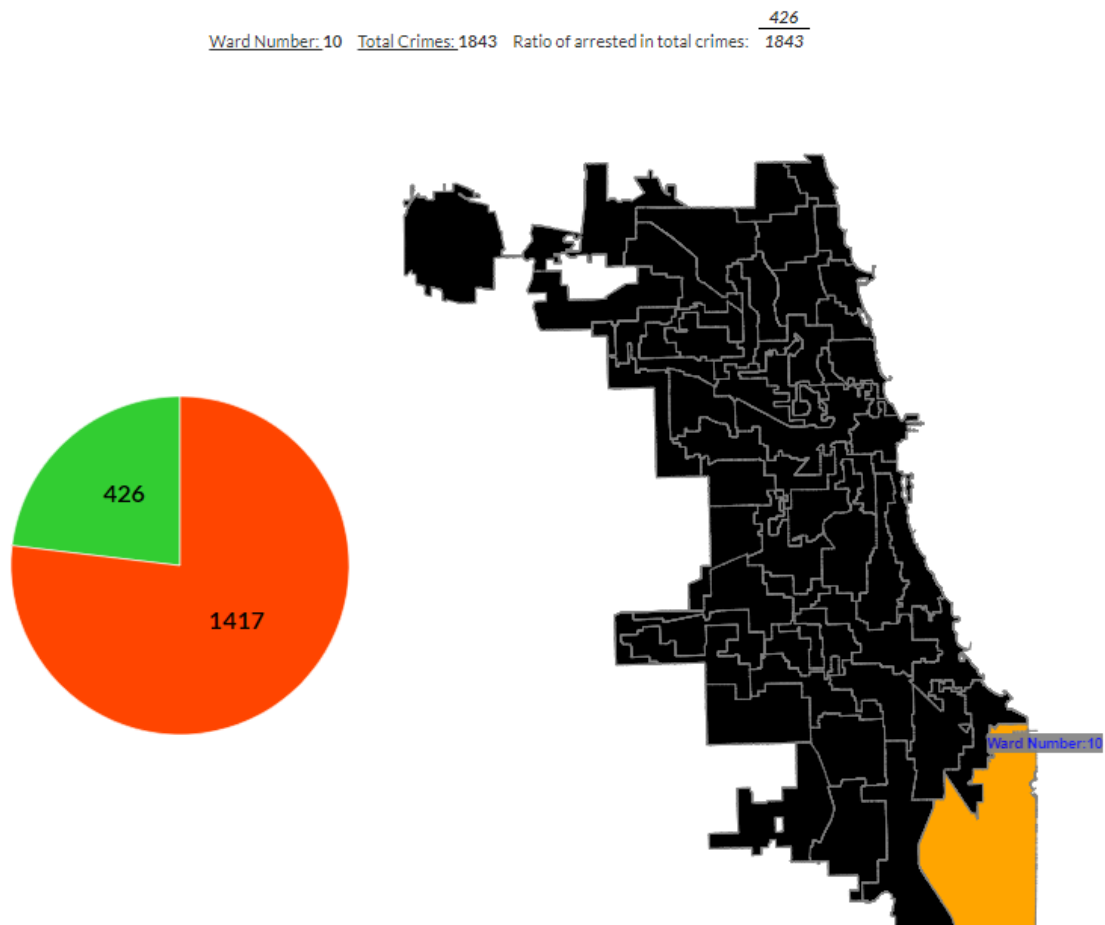
**An exceptional example**, in the case of public indecency and other narcotic violation, we can't really conclude about the relation between those two crimes since there weren't a lot of crimes there.



**Disadvantages:**

The visualization is limited to two crime types, so you can't see the correlation between 3 or more crime types.

- How efficient the police are in each ward where arrest rate is the indication.



Type of the graph: Pie Chart and Map.

The visualization shows the ratio of crimes that police made an arrest and the crimes that the police did not make an arrest.

The pie chart shows the number of arrests on the green area and the number of non-arrests on the orange area.

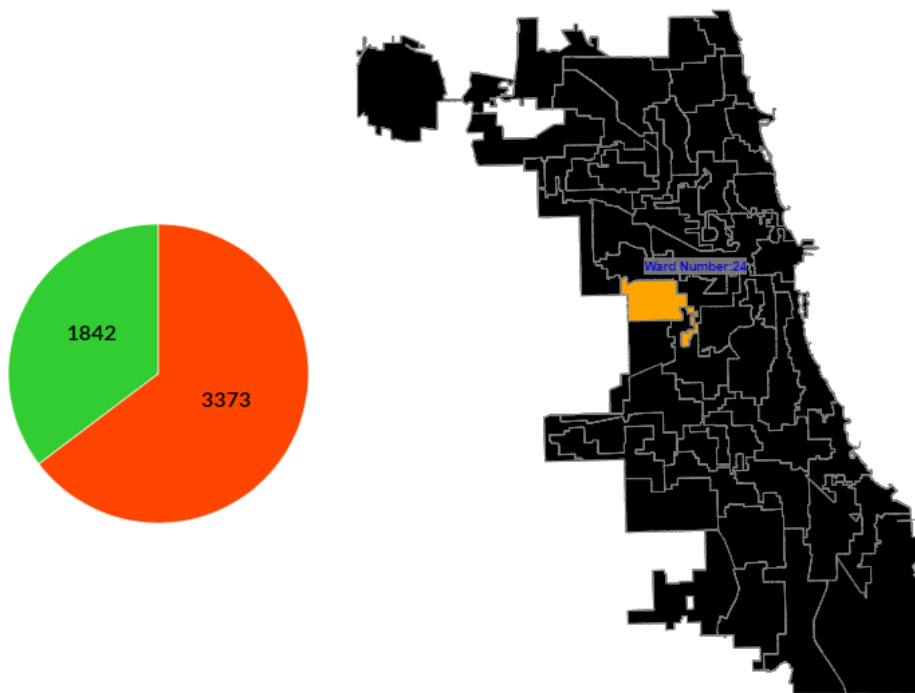
When hovering on the map, the ward number will be shown together with the corresponding arrests to non-arrests ratio of that ward. If you are not hovering on the map, the pie chart will show the general arrests to non-arrests ratio for all wards.

### What can we learn?

In most of the wards, the number of arrests is less than a quarter of the number of total crimes, which means that the police efficiency is low when it comes to arresting the criminal.

**An exceptional example** would be, in ward number 42, the number of arrests is more than a quarter of the number of total crimes, so we can conclude that the police are efficient in ward 42.

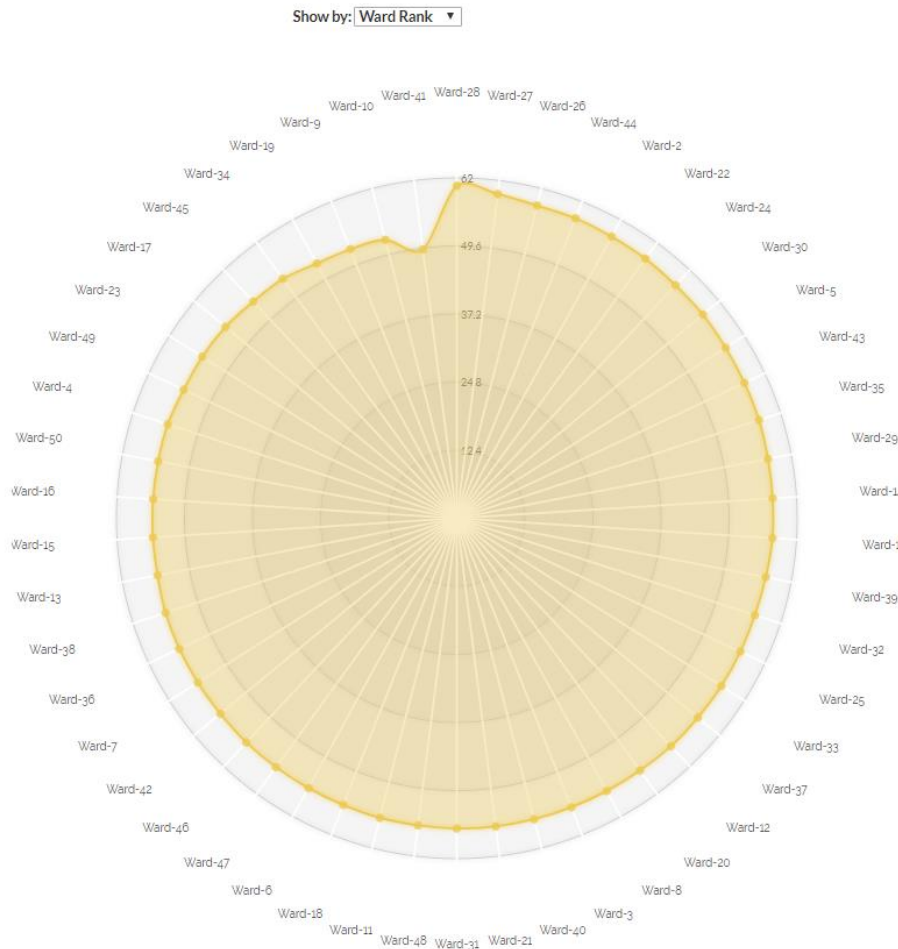
Ward Number: 24    Total Crimes: 5215    Ratio of arrested in total crimes:  $\frac{1842}{5215}$



### Disadvantages

You can't see the statistics of a smaller scale areas like street or a block.

- Rank ward areas by using weights for each type of crime or where the crime has occurred (domestic or not) and whether the criminal was arrested or not.



Type of the graph: Radar Chart.

The visualization shows two types of radar charts:

1. The first option is, ranks of each ward (Yellow) ordered from the highest rank to the lowest rank clockwise.
2. The second option is, crime rate for each ward ordered from the highest to the lowest clockwise (Yellow). The chart also includes the number of domestic crimes in each ward (Blue) and the number of captured criminals (Red) on the same scale of the number of crimes.

The maximum number on the scale is the maximum of all ranks or maximum of crime rate (depends on the option you choose).

If you hover the mouse on each dot on the chart a number with the rank/number of crimes will appear.

### How did we calculate the rank?

First, we searched online the maximum sentence for each crime type in the state of Illinois and we made a list of each crime and its maximum sentence.

(The list will be presented in the next page).

Later, we used the following formula:

$$x \in Crimes, y \in Wards$$
$$R_{xy} = \left( \frac{\text{number of } x \text{ crimes in ward } y}{(\text{Total crimes in ward } y)/2} \right) * (\text{The maximum sentence of } x)$$

$$T_y = (\text{Total crimes in ward } y)/2$$

$$A_y = \text{number of arrests in } y$$

$$D_y = \text{number of domestic crimes in } y$$

$$Rank_y = \frac{50 * (T_y - \frac{A_y}{2} + T_y - \frac{D_y}{2})}{T_y} + \sum_{\forall x \in Crimes} R_{xy}$$

\* We used the maximum sentence because that way we can quantify how serious the crime is in the eyes of the state of Illinois. Basically, the weight for each crime is the maximum sentence.

\* **What do we measure?** The rank measures how unsafe the ward is. The more crimes with a high sentence related to the total amount of crimes will produce a higher rank. Another thing that might increase the rank is a low rate of arrests or a low rate of domestic crimes.

### What can we learn?

We can tell that ward 42 has the highest crime rate.

We can also tell that the police are not that efficient because related to the amount of crimes, the number of arrests is relatively low in each ward.

We can understand that it is safer to stay at home because most of the crimes aren't domestic.

The winning ward in the best rank contest goes to ward 28. It is the most unsafe ward by our grading system.

The safest ward by our grading system is ward 41.

### Disadvantages

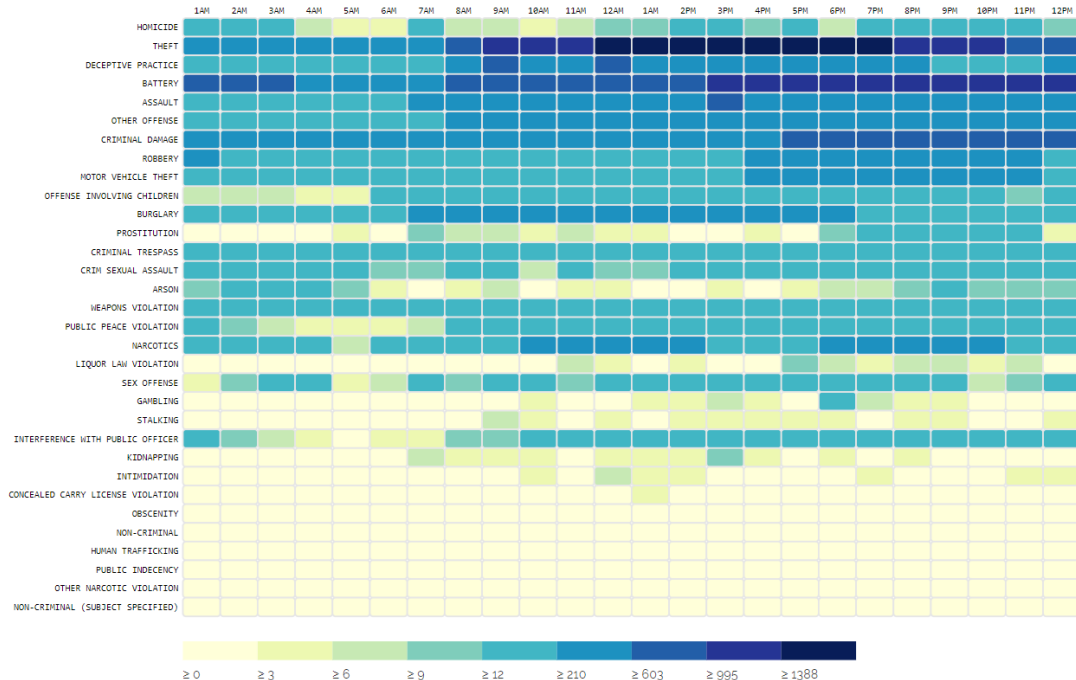
When choosing the second option (crime rate for each ward), we cannot see the differences between each ward clearly when it comes to the amount of arrests and the amount of domestic crimes.



### **List of crime type and its sentence**

<b>HOMICIDE</b>	<b>60 years</b>
<b>THEFT –</b>	<b>15 years</b>
<b>DECEPTIVE PRACTICE</b>	<b>3 years</b>
<b>BATTERY</b>	<b>30 years</b>
<b>ASSAULT</b>	<b>30 days</b>
<b>OTHER OFFENSE</b>	<b>1 years</b>
<b>CRIMINAL DAMAGE</b>	<b>15 years</b>
<b>ROBBERY</b>	<b>55 years</b>
<b>MOTOR VEHICLE THEFT</b>	<b>30 years</b>
<b>OFFENSE INVOLVING CHILDREN</b>	<b>15 years</b>
<b>BURGLARY</b>	<b>15 years</b>
<b>PROSTITUTION</b>	<b>1 year</b>
<b>CRIMINAL TRESPASS</b>	<b>2 years</b>
<b>CRIM SEXUAL ASSAULT</b>	<b>60 years</b>
<b>ARSON</b>	<b>15 years</b>
<b>WEAPONS VIOLATION</b>	<b>5 years</b>
<b>PUBLIC PEACE VIOLATION</b>	<b>Jail</b>
<b>NARCOTICS</b>	<b>30 years</b>
<b>LIQUOR LAW VIOLATION</b>	<b>10 years</b>
<b>SEX OFFENSE</b>	<b>30 years</b>
<b>GAMBLING</b>	<b>3 years</b>
<b>STALKING</b>	<b>5 years</b>
<b>INTERFERENCE WITH PUBLIC OFFICER</b>	<b>5 years</b>
<b>KIDNAPPING</b>	<b>30 years</b>
<b>INTIMIDATION</b>	<b>5 years</b>
<b>CONCEALED CARRY LICENSE VIOLATION</b>	<b>5 years</b>
<b>OBSCENITY</b>	<b>5 years</b>
<b>NON-CRIMINAL</b>	<b>Jail</b>
<b>HUMAN TRAFFICKING</b>	<b>11 years</b>
<b>PUBLIC INDECENCY</b>	<b>1 year</b>
<b>OTHER NARCOTIC VIOLATION</b>	<b>Jail</b>
<b>NON-CRIMINAL (SUBJECT SPECIFIED)</b>	<b>Jail</b>

- Showing the crime volume by color for each hour of the day. The darkest color indicates highest level of crime occurrences and lightest color indicates the lowest level of crime occurrences.



Type of the graph: Crime hour heatmap.

The visualization shows the amount of crimes in each hour of the day for every crime type.

Each row is a different crime type.

Each column is a different hour of the day.

The legend underneath the heatmap shows the quantities of crimes for each color, where lighter color indicates a lower crime rate and a dark color represents a higher crime rate.

### What can we learn?

- All crimes are divided between 2 main types:
  - o Crimes that are common and occur a lot around Chicago at all time.
  - o Crimes that are so rare that they don't occur at all or barely appear in the police database.
- For each crime type with high crime rate we can learn the crime rate distribution around the day.
- For a specific crime, one can learn when it is safer to be outside, because each crime type has hours with low rate related to the rest of the day.  
For example, if one would like to avoid arson, the hour 10AM is the best time to visit Chicago because statistically less crimes of that type happen in 10AM.

**An exceptional example**, usually one would think that most of the thievery will happen in the dark hours of the day, but as the map shows, thievery mostly happen in daylight (12AM – 7PM).

### **Disadvantages**

The crimes that rarely occur do not affect the map at all and makes the bottom part look boring and not interesting and we cannot recognize any pattern of crime occurrences around the day.

In conclusion, we cannot see the crime distribution of crime types with low occurrences, because the legend is limited to 9 levels and there is a very high rate for other crime types, so all the low rate crime types have the same light color.