CSE 442

# Assignment #2

## Dataset: Crimes in Chicago

This dataset contains records from 335,706 cases that were reported in the city of Chicago in the year 2012. The raw data contained information on date of crime, location, primary type of crime, description, whether an arrest was made and whether it was a domestic crime. Other information was included to help uniquely identify the records. At times, there are two records for a single incident, perhaps when there are multiple parties involved in the crime.

## Initial Analysis Questions

1. **Overarching Question:** What factors influence crimes?
2. How does the time of year, week and day influence crime rates?
3. Are there specific crimes that correlate with one another? Is there potential for over policing biases to exist in these areas?
4. How do the crime patterns relate to the geography of Chicago? Are there areas that are known for specific kinds of crimes?

## Discoveries and Insights

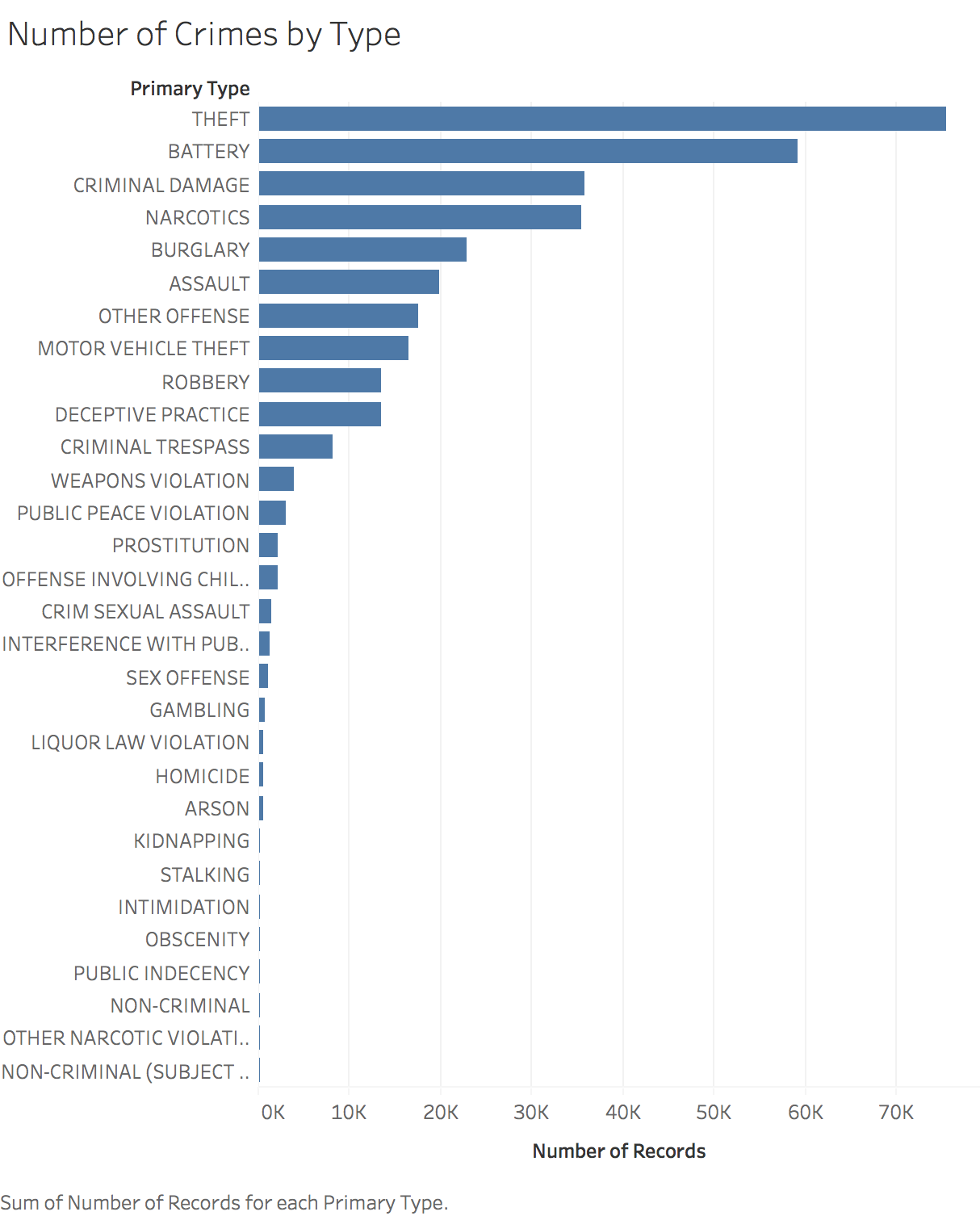


Figure 1: resize to lower length, trim down to 15

Figure 2: re-export to show label on narcotics

Type of CRIME

Figure 2a: add graph of categories

Figure 3: text labels as well

Figure 4: rearrange label to 3

Figure 5: reference line on average, flip

Figure 6 and 7 fit to width

There is an uneven distribution of crimes across the different types of crimes (crime type is based off Primary Types as indicated in each record). Chicago has a reputation for rates of gun violence and homicide, however this chart show that the number of homicide cases are negligible compared to that of theft and battery. One hypothesis for this difference is the value that our society and media places on addresses cases such as homicide and sexual assault; smaller crimes are often overlooked. For sake of scientific reliability, the five most common types of crimes are used for analysis as they are more reliable indicators of crime trends over time (with the exception to one analysis below).

When each of the crimes are sorted into larger categories as defined by the LegalMatch Law Library, personal crimes, property crimes and drug crimes vary greatly in the rates of arrest. Most notable, in 2012, there was a 99% arrest rate among drug crimes (35,225 out of 35,487). Whereas the rate of arrest for personal crimes and property crimes are 22% and 13% respectively.

The number of crimes per category are graphed throughout the months of 2012. Line graphs were chosen over area charts to lessen errors with aggregated effects. One trend shows that the number of crimes in theft, battery, burglary and criminal damage all increase as the months warm, peaking in July possibly due to overall more outdoor activities and inter-personal interactions. Narcotics’ activity peaks in February where all other crime activities dip (fewer number of days), possibly due to lack of accessibly to other forms of enjoyment during cold Chicago winters.

The number of crimes per category are graphed in a horizontal trellis chart to show crime patterns throughout the week. Property crimes, burglary and theft are both dip drastically during Sundays and Saturdays, very likely due to the presence of homeowners being present. Battery suffers a significant increase on the weekends (from 149.2 to 195.5 average crimes per day); the following graph provide an explanation. Overall, weekdays do influence the rate of crimes, however based on the type of crime, this rate could increase or decrease depending on the presence of other humans.

Above is an area graph with the total number of battery crimes split into domestic and non-domestic cases. The sum of the two types of battery give us the U shape as seen in figure 5. However, this graph allows us to see the slight increase in the non-domestic cases of battery on and the dramatic surge of domestic battery cases on weekends. Unfortunately for victims of domestic abuse, the time spent at home is correlated with the likelihood of abuse.

Crime patterns are mapped by time of day. Property crimes peak during the middle of day, when homeowners are off at work while personal crimes peak at the end of the day when the work day is over and the sun has set. Narcotics experience two peaks, once at noon and once at 7:00pm perhaps when the work day is over.

Above are six trellis graphs that detail the location where each type of crimes is occurring. All crimes, regardless of frequency were mapped and the most insightful ones were chosen for sake of simplicity. The reason why certain graphs were eliminated is either due to too few data points or ubiquitous occurrence of crimes. Small colored marks are used to allow density of crimes to be visualized. Note that there is a limitation to the data as we know that theft is very common and around the west side of Chicago, there are straight lines to where theft no longer occurs: indicating the bounders of CPD.

Other notable insights are that deceptive practice such as fraud are concentrated in downtown Chicago on the middle right. Locations of prostitution highlight vertical and horizontal lines throughout Chicago, mapping out the streets of where prostitutes are most likely doing their business.

Finally, the last three graphs all show a very similar pattern. The next figure explains this in further detail.

By overlaying the crimes related with interference with public officer with narcotics, we see that interference with public officers almost always occur with the presence of narcotic crimes. We also know from figure 2 and 3 that narcotic crimes have approximately 99% arrest rate; confirming the presence of public officers. Here, it’s possible that the reason why interference with officers have high crime rates in areas where narcotic crimes are high is due to a high presence of police overall. The existence of one type of drug crimes could be the fuel to crimes with public officers and even robbery. However, whereas robbery is most likely a phenomenon due of having drugs in the area, interference with police may be a result of over policing in known drug zones. These drug zones might be known not because there are high number of criminals but because of over policing in the first place. We see a very likely existence of a cyclical pattern that damages these communities.

To further confirm this finding, linear regressions were made to visualize the correlation between Robbery and Narcotics and Interference with Public Officer. In both cases, the upper and lower 95% confidence bands are small and the p values for both are < 0.001, assurance the significance of the reliability of the linear regressions.

Above are box and whisker graphs for the number of crimes by type sorted by block. A logarithmic scale is used on the right to better see the medians of crime rates. To end on a positive note, despite all the crime that exists in Chicago, most of the blocks will only experience one or two crimes in their area throughout the course of a single year. There are dramatic outliers and residents of these areas should be aware and take the appropriate precautions.

1. **Overarching Question:** What factors influence crimes?
2. How does the time of year, week and day influence crime rates?
3. How do the crime patterns relate to the geography of Chicago? Are there areas that are known for specific kinds of crimes?
4. Are there specific crimes that correlate with one another? Is there potential for over policing biases to exist in these areas?

In summary, the time of year, week and day all have influences on crime rates. Most crime increases in the summer and dips in the winter times. Crime is constant throughout the workweek and depending on the type of crime, they increase or decrease on weekends. Time of day is also influenced based on the workday and when individuals are off work and at home.

The geography of Chicago primarily influences crimes like deceptive practice where they occur downtown, most likely in corporate businesses while prostitution occurs further away from the core of the city along main streets.

Crimes do correlate with one another, particularly if the crime committed is related to high levels of arrest rates. The presence of police officers will no doubt lead to more crime cases and arrests. This could be extremely problematic when there are systematic biases in policing which is aggregated by this phenomenon.