

Decision Support Systems Project

Chicago Crime Rate Analytical Report

Monday Class, Section 2

Authors

Shams Alasa’d 20190447

Saba Soufan 20200125

Nadine Sweidan 20200097

Talal Alkilani 20200782

Instructor

Dr. Noor Damer

**Table of Contents**

[Introduction 3](#_Toc123922509)

[1- Data Cleaning Phase 4](#_Toc123922510)

[2- The Analysis phase 6](#_Toc123922511)

[3- Recommendations and Outcomes 15](#_Toc123922512)

[References 17](#_Toc123922513)

**Table of Figures**

[Figure 1 4](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924620)

[Figure 2 5](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924621)

[Figure 3 6](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924622)

[Figure 4 6](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924623)

[Figure 5 6](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924624)

[Figure 6 7](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924625)

[Figure 7 8](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924626)

[Figure 8 9](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924627)

[Figure 9 10](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924628)

[Figure 10 10](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924629)

[Figure 11 11](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924630)

[Figure 12 11](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924631)

[Figure 13 12](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924632)

[Figure 14 12](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924633)

[Figure 15 12](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924634)

[Figure 16 13](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924635)

[Figure 17 14](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924636)

[Figure 18 15](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924637)

[Figure 19 15](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924638)

[Figure 20 16](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924639)

[Figure 21 16](file:///C:\Users\LEGENDA%20BOOKSTORE\Downloads\Decision%20Support%20Systems%20Project%20final.docx#_Toc123924640)

# Introduction

Chicago is a populous city located in The United States of America and it’s known for being diverse, how windy it gets, Oprah Winfrey, and its high crime rate from petty theft to full blown homicide.

To further investigate how true this reputation is, we’re going to analyse a set of data all about Chicago’s crime rate from the years 2012 to January 2017, the dataset houses information about the type of crimes committed from theft, battery to domestic violence, it has districts, wards, community areas and all the location points one would need to know, this information helps us understand what the dataset means and make educated assumptions on the crime rate and allow us to explore the reasons behind this perceived image of Chicago state.

The methods of analysis utilised in this case would be applied quantitative analysis consisting of the charts and graphs that will help understand the relationships between the variables available and qualitative analysis to improve overall understanding of the analysis and explain our findings in greater detail.

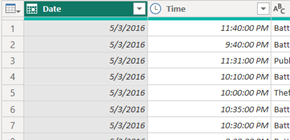
The goal of this analysis is to get a concise and definite conclusion on Chicago’s crime rate and the general safety surrounding it, also to find repetitive patterns in certain behaviours or actions that can help minimise crime and see the signs before it's too late. the analysis will help shine light on certain neglected factors such as most common crime type in specific areas, the relationships between certain variables and how they affect crimes,

Through our findings we’ll be able to suggest changes to police systems and give recommendations and ultimately enhance the decision-making process for the police departments and authoritative entities.

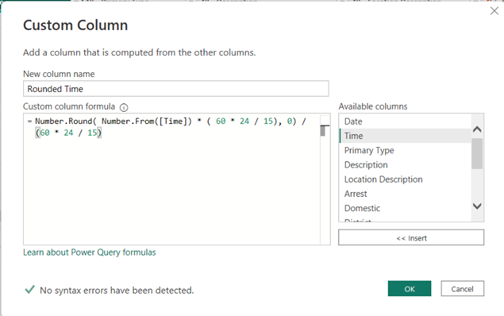
To begin faced issues regarding which tools and programs to use in-order to analyse the dataset, we used Microsoft Excel for some of it but soon enough was too complicated as we lacked certain skills to use effectively, overall it was inefficient and time-consuming, thus we switched to Power BI for its availability, ease of use and it had everything we needed to complete the analysis from tools to charts, we also struggled in the cleaning phase to determine which cells are relevant to our case and which isn't as most of our crime and justice system knowledge comes from NCIS: Los Angeles and Criminal Minds, not very useful yet entertaining! But eventually while going through the definitions and establishing progress, we bypassed this issue and moved on to the next step.

**The Process**

## 1- Data Cleaning Phase

We started the process by studying the dataset all about Chicago’s crime rate and understanding the definitions of all the variables given in the case in-order to have a clear, vision of what we needed to do in this analytical report, we faced a bit of trouble as we begun “cleaning” the data of any redundant irrelevant information that wouldn't serve the analysis such as the “X”, “Y” coordinates as the same information can be found in “longitude”, “latitude” and along with the location of the crimes we were able to establish relationships between the variables and a set of facts that will be explored later in the report, we also decided to separate the date and time as it helped us observe important dates and time frames.

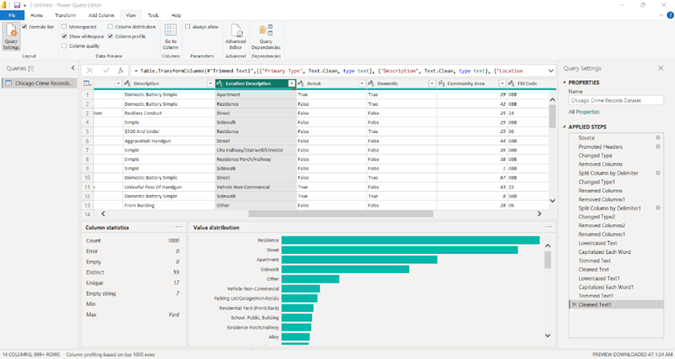
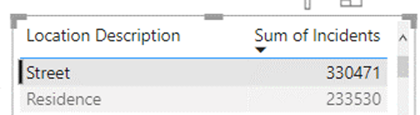
Figure

In order to have clearer and more accurate visualisation of the data we decided to round the time to the nearest 15 min using the following formula. 

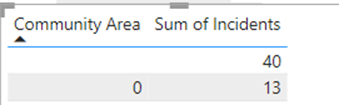
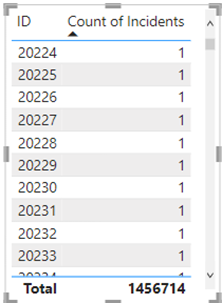
Figure

In-addition we worked on this phase throughout the process because as we moved forward with the analysis we discovered more inconsistencies and data errors like blank cells or null cells that we for most part easily filled in or removed with the =count IF blank function on excel although we were not sure of the rest to start but later on found no issues moving forward despite empty cells, we decided to keep the district as it reveals the area an correspondent police district, which with the location can also tell a lot about the demographics of the area residents which can correlate with the crimes or lack thereof and even more, we kept the obvious factors such as “ward” as it tells us the area where the crime took place, “description” of the crime as in type, if it’s domestic or not, the type of crime “arrests” record, “year” and “last updated on” as it might indicate the average timelines between when a person commits a crime and when they’re officially convicted (as in arrested), assuming that the date when the case file was last updated is the date of the arrest, then most of the arrests were made within a week from the incident, and in general but not necessarily the case, the more time passes the less the arrests.

Everything else we removed was due to the fact it fell short of what we needed and didn’t serve our analysis.

We also wanted to investigate whether or not there were crimes committed by the same person but all the IDs in the dataset were unique.

Figure

Some anomalies we noticed in the data, in addition to blanks, that the official count of community areas is 77, but we found community areas with the number 0. (Both were filled with the CA with the most crimes -25-)

Figure

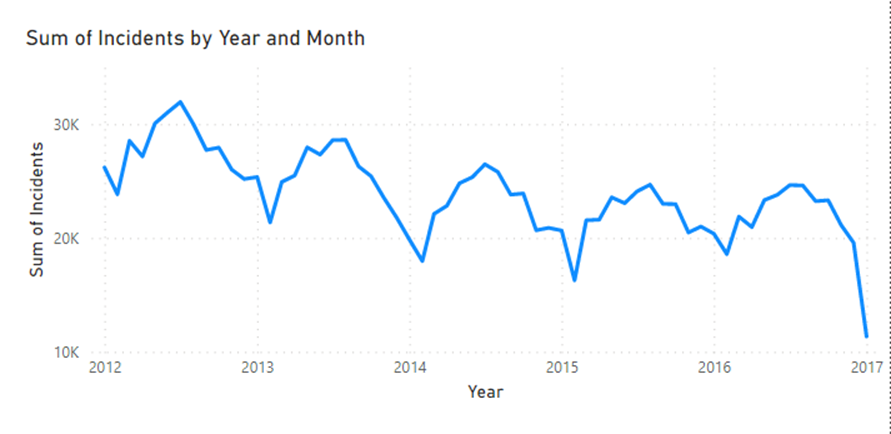
Figure

2- The Analysis phase

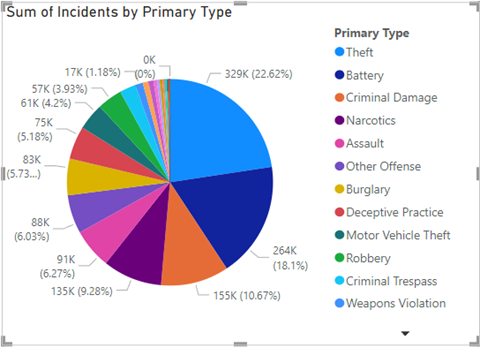
In this phase we were able to establish relationships between the variables in the dataset and the environmental factors surrounding the crimes and observe data in a way that gives it meaning and describe it, with the help of power BI tools we were able to assemble charts, graphs and build models to help describe and make the data make sense.

We identified the following observations:

* The information provided crime rates of the years from the Mayan year 2012 till early 2017, thus we established that the year with the highest crime rate was 2012 with theft being the main crime committed and reported, now that was interesting to us as we know back in 2012 everyone thought the world was going to end according to the Mayan calendar (the maya being an old civilization with its own systems and set of beliefs) and we wondered if there is a correlation between that event and the crime rate like the events of Y2K or was it just a mere coincidence, we decided to work that angle and see what comes up after a bit of digging we concluded that although certain events and apocalyptic scares do in-fact inspire crimes as there was a plethora of sources highlighting crimes related to sports, concerts, holidays of all kinds seems to produce more crime of all kinds, 2013 came at a close second with similar numbers as you can see below.

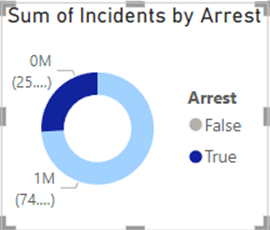
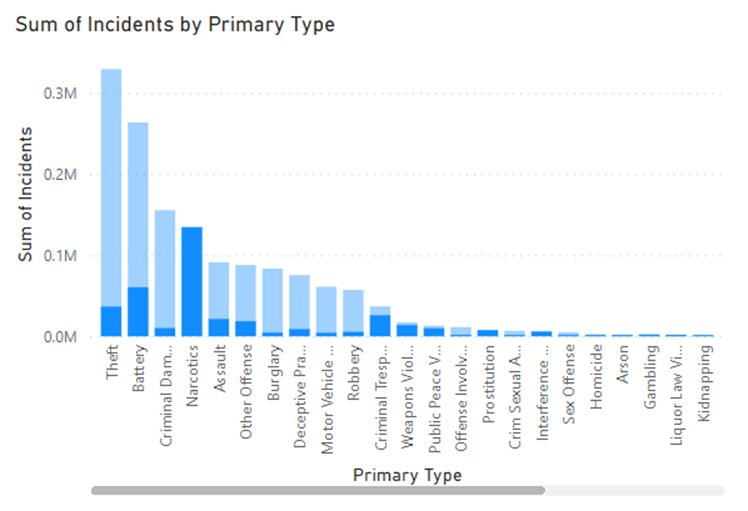


Figure

* On the other hand The year with the least crime reported in the dataset was 2017 that’s because the dataset only provided information up to January of that year, for that we attempted to make a predictive analysis through the historical data we have at our disposal after analysing the dataset thus we could predict how the crime rate would be in the up-coming months for example we could forecast that as temperatures rise in the summer-time crime will most likely increase compared to other months in the year, we could also anticipate the crimes being majority theft and battery more than anything else with narcotics being the crime most convicted and arrested for, in-addition we have a hunch district 11 and 8 won’t be on the top ten safest areas in Chicago list anytime soon.

Figure

* The most common arrest type as in the crime that garners the most arrest records reported in our analysis was for narcotics, a narcotics arrest involves the offender either being in-possession of illegal substances or being caught for or while using, selling, owning illegal drugs, throughout the years in the data-set arrest for narcotics has been consistent, that highlights the drug issue Chicago seems to be dealing with and after further investigating we came to the conclusion that Chicago is in-fact considered one of the biggest drug hubs in the united states with its large volume of legit trade and highways, trains and air transportation along with thousands of warehouses it does indeed pose a threat, the area most notorious for drug abuse is district 11 which we will delve deeper into in the next point. Arrest for Theft comes in second despite being the most committed crime with about 22.6% of all crimes and battery with 18.1% combined they make up 40.7% of all crime. *(The Drug Situation in the Chicago Field Division DEA Intelligence Report, 2019)*



Figure

In the above visualisation you can see the difference between the number of times the incident/crimes were committed (light blue) and the number of arrests (dark blue) theft is the primary crime Chicago is facing yet narcotics crimes get the most arrest

* Subsequently the least crime recorded for arrests was human trafficking with 4 arrests

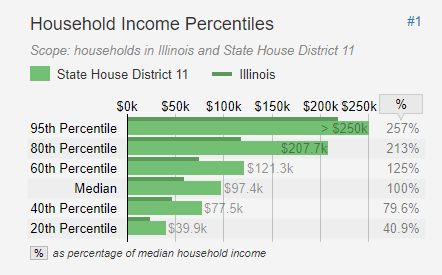
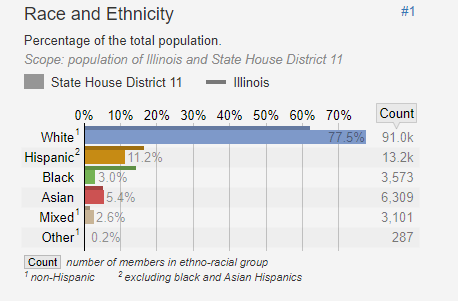
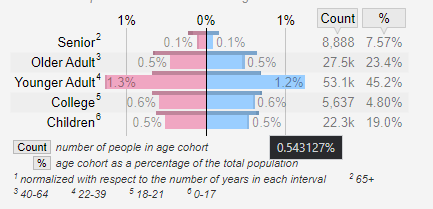
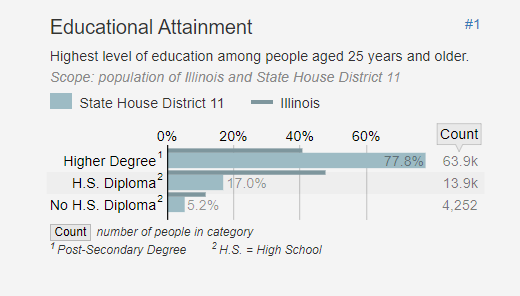
But overall, they were only 28 crimes throughout the years according to our dataset, we were apprehensive though as it made no sense as Chicago does have a high human trafficking rate and after quick research our suspensions were confirmed, Chicago is a centre for human trafficking and this could be a reporting issue or an inconsistency the datasheet. *(Human Trafficking: Chicago Is One of the Leading Cities in This Global Epidemic, 2012)*

* District 11 was the area with the highest crime rate in Chicago, the district falls under cook county, Chicago area to be exact, the crime type being narcotics use, the most dominate demographic factors in district 11 consist of young people between the ages of 25-30, predominantly Caucasian, 39% married and with an unemployment rate as low as 3% in that area.

The average income level of a household in district 11 is around 97k a year which is relatively high (measured against poverty-line being 27k a year).

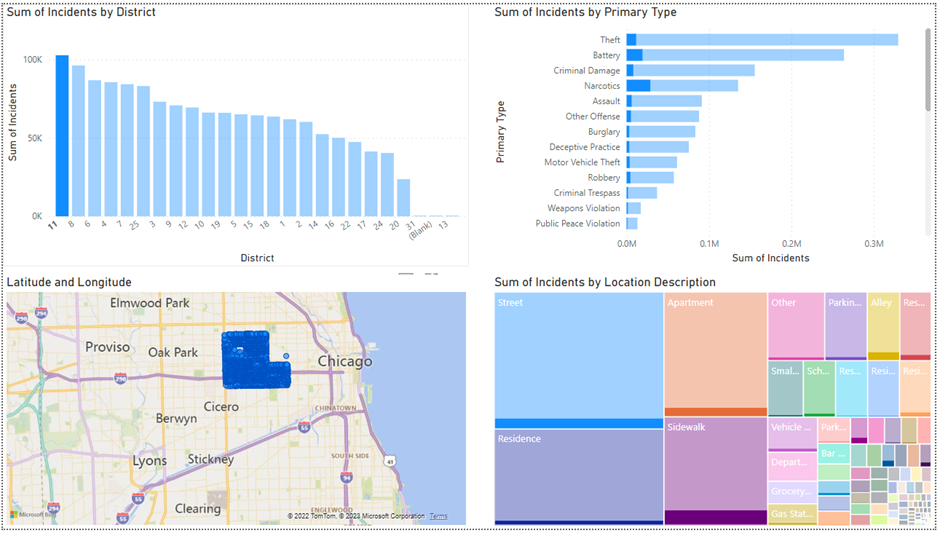
These numbers were all extracted from the district and area community which provided information on demographics and the description of the crimes.

This analysis helps us paint a picture of what the area with the highest reported crime looks like and could help end prejudice as well as bias as most common assumptions about crime related discourse centre around poverty, non-white under-educated neighbourhoods, and as you can see (and will see) the below statistics express a different story.[*(The Demographic Statistical Atlas of the United States - Statistical Atlas, 2018.)*](#_jb1bcpmptsar)

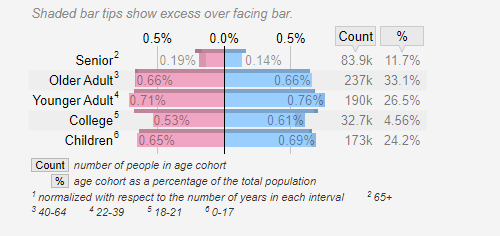
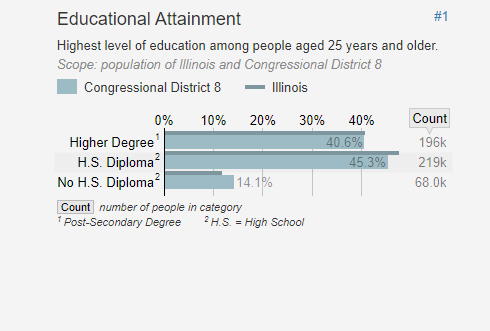
These statistics are taken from the demographic statistical atlas of USA to make the above make more sense

Figure

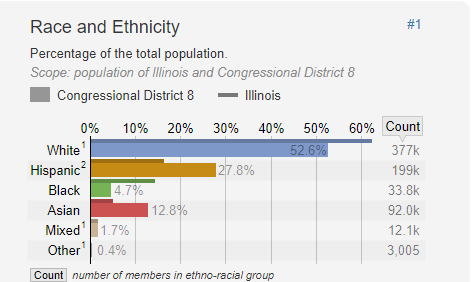
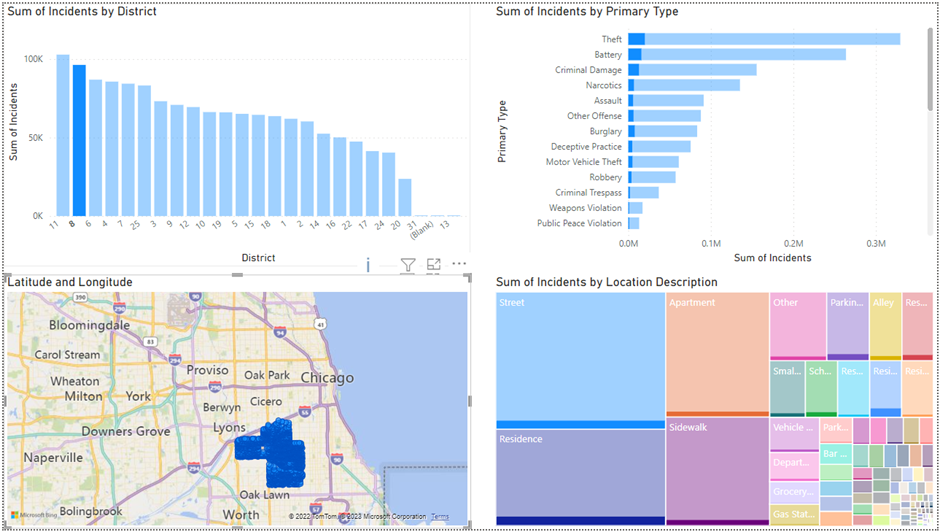
Figure



Figure

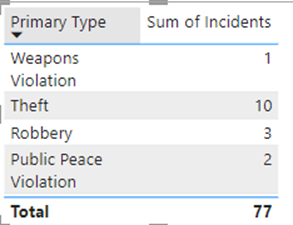
* District 8 came at a close second with the crime type being predominantly theft and battery, through knowing the community area and location we were able to work-out the demographic of district 8 which were mainly Caucasian people, the households consisted of majority married individuals and nuclear families taking up 74.8% of the area, 40.6% of district 8 residents go on to attain higher education degrees while 45.3% are sufficient enough with just a high-school diploma and the rest of the 14.1% didn’t finish high-school for unknown reasons, the average income of these households ranging between 75k-100k a year what that tells us is this district is above poverty line, well-educated neighbourhood, the theft could be explained by the area having commercial entities as well as residential.[*(The Demographic Statistical Atlas of the United States - Statistical Atlas, 2018)*](#_jb1bcpmptsar).

Figure



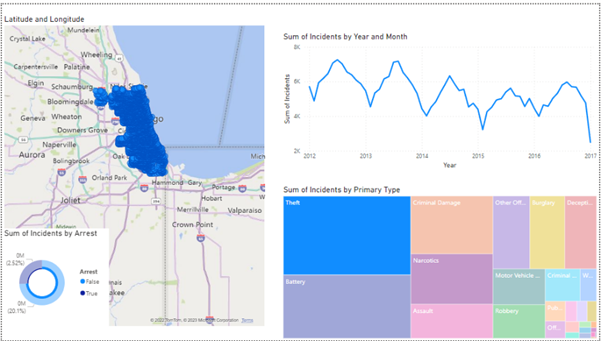
Figure

Figure

* Theft is the most committed crime throughout the years given in the dataset making it the most common crime type in Chicago state, the place most common is theft on the street in all hours of the day this crime is the most constant, and after taking a quick look at theft crime reports we gathered that the theft in Chicago is borderline ridiculous the most reported stolen items are cars! Kia and Hyundai to be exact, thousands of dollars lost every day due to theft, the number of retail shop-lifting, robberies at gun-point and even looting during protests all of which contribute to Chicago's increasing theft problem.

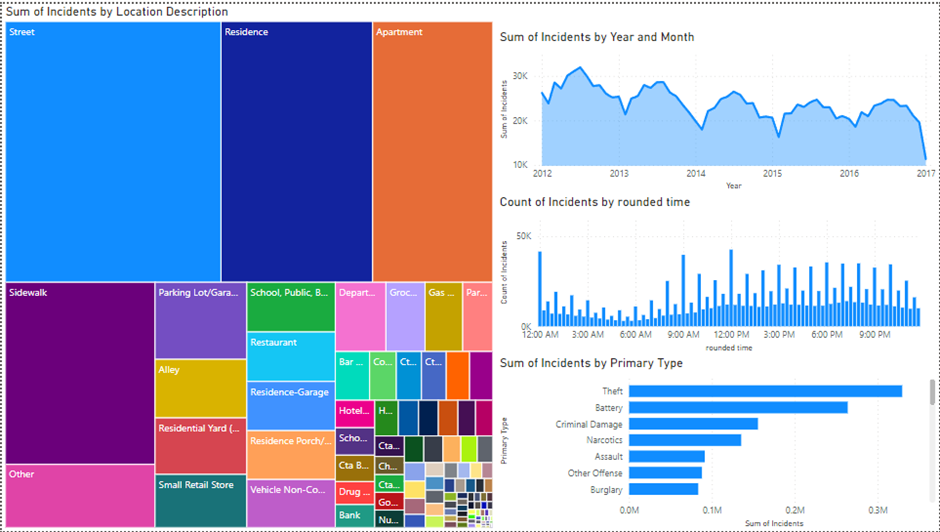
Figure

* The peak period of time for crimes was during the summer time more specifically June and July of each year documented, that’s a trend we observed as each documented year come summertime the crime rate spikes, which lets us know that in the up-coming years come summertime crimes will increase based on historical data.

The reasoning behind crimes spiking in the summer could be phonological, the discomfort caused by high outdoor temperatures acts as a stressor and alters the brain’s chemistry to make people more aggressive and irritable. As the days become longer and hotter, tempers become short and violent crimes increase, another reason is that a number of college kids come back home for the summer holidays such as 4th of July, parties and alcohol consumption also are higher during the summer-time, somewhat creating more crime opportunities and a bigger pool of possible victims. [(Times, 2022)](#_jb1bcpmptsar)

Figure

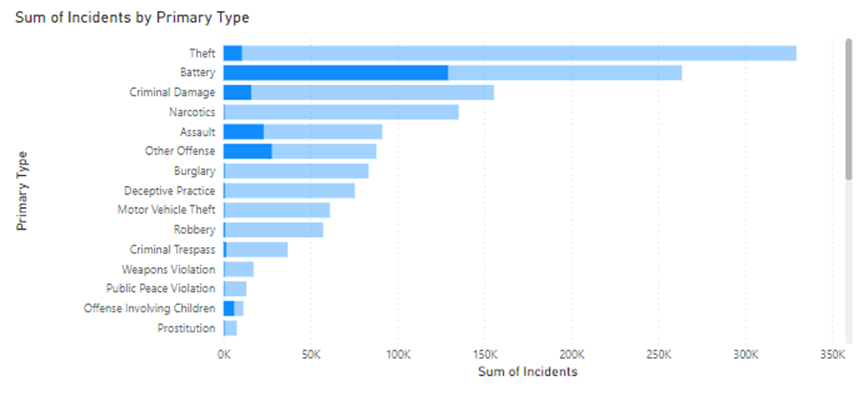
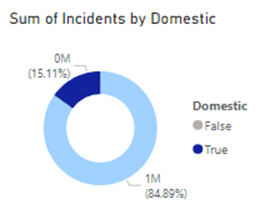
* Peak hours were the as early as 9am, as busy as 6pm and as expected as 12am, at first we were a bit apprehensive of this fact because like most people we assumed that primarily all crimes must happen in the hidden shadows of the night away from watchful eyes, 9am-12pm are typically busy hours of the day and in broad day-light, which in our analysis after further investigations made sense as most recurring crimes in the dataset are theft on the street which includes burglary, misdemeanours, assault and robberies as well as battery, these types crimes happen routinely in the earlier hours of the day and night and while more violent crimes take place at night-time after midnight.



Figure

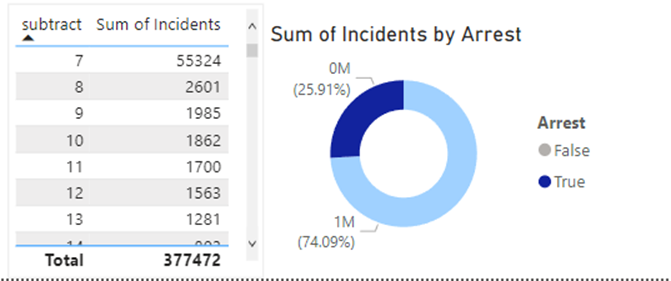
* Domestic crimes as defined by the Chicago police department is a unique crime in

which the offender lives in the same household as the victim, and according to our analysis domestic crimes contribute to 15% of crimes in Chicago consisting of domestic disturbances, batteries and protection orders, although we'd like to state the fact that most domestic crimes go unreported or get dropped by the victim due to numerous factors and we believe that percentage is in actuality higher, according to the national crime victimisation surveys only half of all domestic abuse claims are reported to law enforcements, why’s that? Well for starters domestic violence more often than not occurs in rather intimate and familial matrimonial relationships away from watchful eyes that only makes it harder for the victims to come forward out of fear, embarrassment and another consideration to keep in-mind is the victims lack of financial power and moral support all of these reasons prevent a true consensus on domestic crime rate. (Talking Parents, 2021)



Figure

* Last updated on information helped us know just how long it normally takes for a crime to get a verdict in and in our analysis we found that the average time between date of crime committed and “last updated on date” in our dataset would range from one week to five years (depending on the crime type and other undetermined variables), to get more details we compared the last update on date with arrest and found that most arrests took place after one week of the perpetrator being caught ( this doesn't include the trial or court time as we lack the information).



Figure

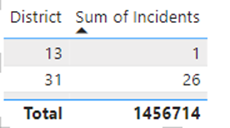
Through-out the analysis process we encountered some inconsistencies in the data provided and or rather unusual outcomes such as the examples below:

*  We found a crime linked to Missouri which could've been either an error or the case of dual jurisdiction, that basically means when an offender commits a crime in one state then moves on to commit other crimes in different states more than one police department can be involved and have jurisdictional power as they cross each-other and they can register the crime/offender in their systems and prosecute the offender with what they deem is suitable.

Figure

* Another curiosity is district 13 has one crime reported throughout the years which again could be a complete error or for reasons that area has a very low to no crime/violence well explore possible reasons, district 13 demographics show a Caucasian majority, well-educated households all similar statistics to other districts we mentioned one difference is that district 13 quite affluent at-least in-comparison to other districts as some household earn above 200k a year all these factors could indicate gated communities or increased security measures that turn off offenders and criminals. [*(The Demographic Statistical*](#_jb1bcpmptsar)

[*Atlas of the United States - Statistical Atlas, 2018)*](#_jb1bcpmptsar)



Figure

## 3- Recommendations and Outcomes

After the initial analysis phase ended we can look back study our findings and the reasons behind such outcomes better in-order for us to conclude this analysis and give recommendations based on the insights, to reiterate briefly we were able to use and analyse the variables in the data set through power BI tools to effectively establish a descriptive model that breaks down the analysis results and as we can see a lot of it is consistent each year, like crime rates have been mulling around the same range, narcotics have been the most arrested for crime, theft seems to be the favourite past-time of criminals and the timeline of crime each year is almost identical (summer-time), were able to forecast the crimes in the year 2017 all due to the consistent outcomes of previous years, without further ado here are our suggestions and recommendations to authoritative figures/entities:

* we suggest police departments and authoritative entities should focus on creating a program that can accurately build predictive models based on historical data and use data analysis to serve their purpose that they can go back to in terms of decision making, for example through data analysis they can track people better for things such as trace weapons back to offenders through purchase records, they can cross reference (using the data systems they have) individuals with specific keywords which can help link them to certain locations, stores, other people and in the end the crime.
* Another suggestion on using historical data which highlights recurring issues to improve and try and lower the crime rates by initiating programs that focus on certain crimes to bring awareness on why they happen, conduct conferences all over the areas mentioned like district 8-11 on the consequences of such actions, build programs aimed to rehabilitate first time offenders so they break out of the mould and go on to be up-standing citizens rather than just adding them to the system and moving on as these records show a massive disconnect between law enforcers and the state of crimes in Chicago as the crimes not only have been constant but they are persistent, and using predictive models can help police know where and when to deploy police support, helps identify offenders quicker and even prevent such crimes as with predictive models and training they can anticipate such things better.
* Data analysis can help the criminal justice systems and police implement crime prevention measures such as build profiles on offenders, a profile is a description consisting of relevant attributes and clues of potential offenders and it’s used typically to detain them, creating patterns can also help narrow down the suspect list and find connections between cases more efficiently than starting from zero, having data supported patterns that police departments can share the pattern with all police stations in different districts instantaneously will result in lower crime rates, using machine learning will be more cost effective and even sustainable as police departments won’t be running around wasting resources, time and money when data analysis can go through all available variables in a matter of seconds and reach the a more precise outcome. *(The Role of Data Analytics in Crime Prediction & Prevention | Classes near Me Blog, 2022.)*

Overall, we can confidently say after conducting this analysis that Chicago’s reputation does in-fact precede it, it’s muddled with crimes and chaos as the tabloids suggest but on the bright-side there are common denominators and clear red flags in this case that they can work on and with our suggestions and technology they can take the steps necessary towards a safer future.

# References

1*-* The Demographic Statistical Atlas of the United States - Statistical Atlas. (n.d.). Statisticalatlas.com. Retrieved December 29, 2022, from <https://statisticalatlas.com/state-lower-legislative-district/Illinois/State-House-District-11/Overview>

2-The Demographic Statistical Atlas of the United States - Statistical Atlas. (n.d.). Statisticalatlas.com. Retrieved December 29, 2022, from <https://statisticalatlas.com/congressional-district/Illinois/Congressional-District-8/Household-Typesv>

3- The Demographic Statistical Atlas of the United States - Statistical Atlas. (n.d.). Statisticalatlas.com. Retrieved December 29, 2022, from <https://statisticalatlas.com/state-lower-legislative-district/Illinois/State-House-District-13/Household-Income>

‌4- Times, N. (July. 2022). Why does crime increase in the summer? | Natchitoches Times. Retrieved December 29, 2022, from <https://www.natchitochestimes.com/2022/07/19/why-does-crime-increase-in-the-summer/>

*5- The Drug Situation in the Chicago Field Division DEA Intelligence Report*. (2019). <https://www.dea.gov/sites/default/files/2019-11/PRB%20FINAL%20--%20DIR-004-20%20The%20Drug%20Situation%20in%20the%20Chicago%20Field%20Division.pdf>

6- Talking Parents. (2021). *Why Domestic Violence Goes Unreported*. Talking Parents. <https://talkingparents.com/parenting-resources/unreported-domestic-violence>

7- *Human Trafficking: Chicago is One of the Leading Cities in This Global Epidemic*. (2012, June 1). NPR Illinois. <https://www.nprillinois.org/statehouse/2012-06-01/human-trafficking-chicago-is-one-of-the-leading-cities-in-this-global-epidemic>

*8- The Role of Data Analytics in Crime Prediction & Prevention | Classes Near Me Blog*. (2022.). [Www.nobledesktop.com. https://www.nobledesktop.com/classes-near-me/blog/data-analytics-in-crime-prediction-and-prevention](http://www.nobledesktop.com.)