

Data Analysis Project

Exploratory analysis of Chicago's crime data in 2021

Team Members

Lujain Mohamed El-said: 202002670

Mariam Hany Shehata: 202000596

Mary Nazeah Abdel-Nour: 202000724

Ranim Moustafa Abdelfattah: 202000515

John Moody Tawfik: 202001162

DR. Moustafa El-Attar

Table of Contents

Introduction	2
Observation.....	2
Description	3
Results	4
Conclusion	8

Introduction

In practically every nation, crime is rising quickly. Analyzing many aspects of crime and identifying criminal trends are highly important. Security agencies throughout the world are working hard to lower these crimes, but the volume of crime data is growing quickly, making it challenging to handle such a large amount of data and to maintain records of crimes that occur over a vast geographic area and over a variety of time periods. Therefore, having a criminal information system that can analyze a lot of data quickly is essential. Also, one of the most efficient techniques to explore, analyze, find patterns, and forecast future crimes in a vast quantity of data is data analysis by using clustering, classification, and association mining.

Observation

- Most occurring top 5 crimes in Chicago in 2021
- Modify the date column to a pythonic format
- Convert the date-time column to a known format
- Top crimes in Chicago in the year 2021
- Create a new column month
- The probability of arrest of Chicago
- Distribution of arrests across the months
- What are the most unsafe hours
- Is your house safe from burglary during the day
- Visualize a crime pattern for 24 hours
- Crime & time
- Crime & location
- Most common Occurrences per district
- Which region is given crime concentrated in?
- Visualizing Narcotics (to get an understanding of the most prevalent of drugs in the area)
- Visualizing with heatmaps
- Using Supervised machine learning to predict crime hotspots

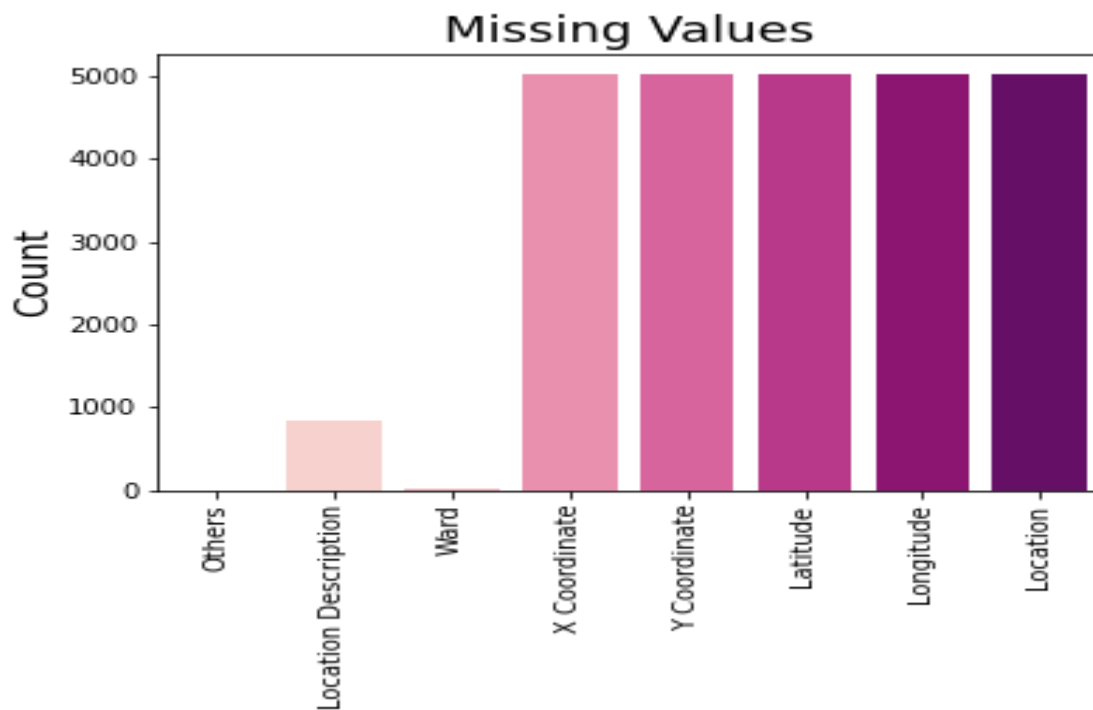
Description

We analyze the crime trend in the city of Chicago in 2021 in our case study which was uploaded from Kaggle https://www.kaggle.com/datasets/salikhussaini49/chicago-crimes?select=Chicago_Crimes_2021.csv. The dataset contains 22 columns and has 207915 cases. Every record identifies criminal activity using various features.

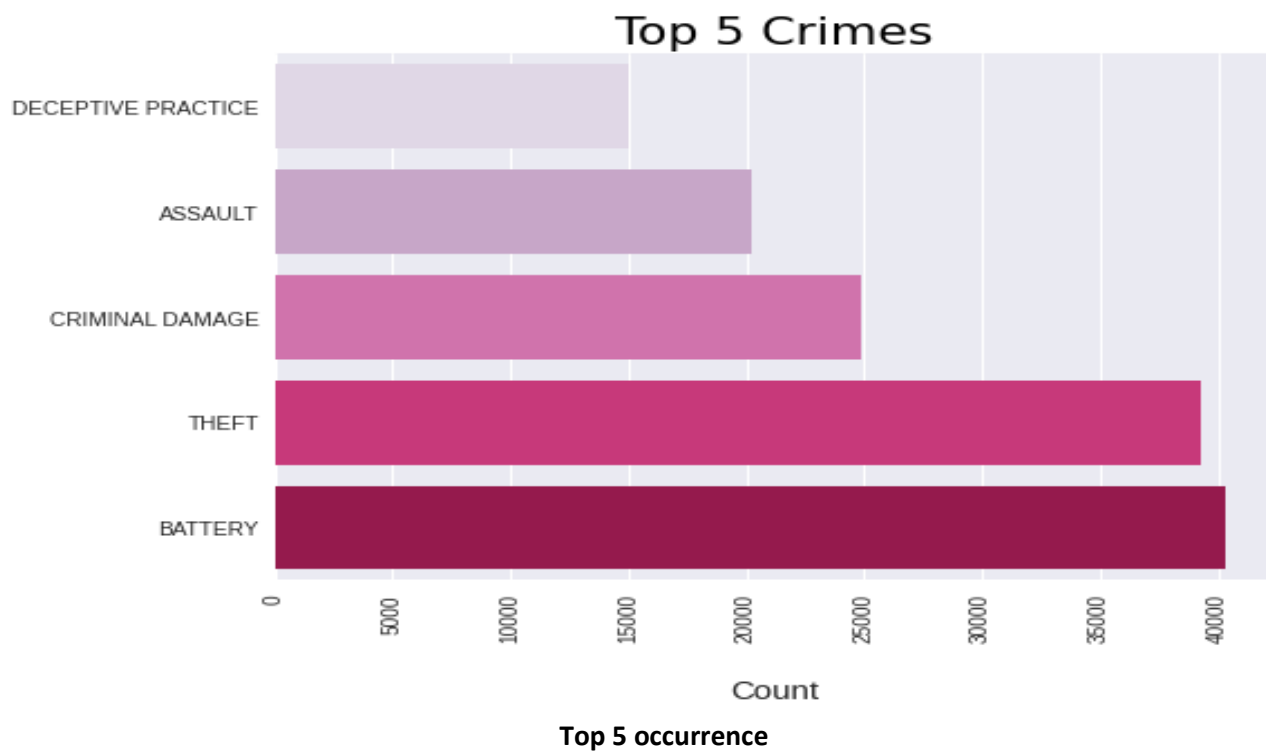
These attributes include:

1. **ID** → Unique identifier.
2. **Case Number**: Chicago Police Division Record Number.
3. **Date**: Date of the incident.
4. **Block**: The abbreviated address for criminal activity.
5. **IUCR**: Internal Uniform Chicago Crime Reporting Code.
6. **Primary Type**: Type of crime
7. **Description**: Little more details about the criminal activity
8. **Location Description**: Location where the crime occurred.
9. **Arrest**: Boolean indicating whether the arrest was made.
10. **Domestic**: Indicates whether the incident was domestic-related.
11. **Beat**: Indicates the Beat where the incident occurred.
12. **District**: Indicates Police District where the incident occurred.
13. **Ward**: The City Council District where the incident occurred.
14. **Community Area**: The Community where the incident occurred.
15. **FBI Code**: Indicates the crime classification code.
16. **X Coordinate**: The X coordinates the location of the incident.
17. **Y Coordinate**: The Y coordinate of the location of the incident
18. **Year**: Year the incident occurred.
19. **Updated On**: The date of the last update of the record.
20. **Latitude**: Latitude of the location of the incident
21. **Longitude**: Longitude of the location of the incident.
22. **Location**: Latitude and longitude saved as a tuple.

Results

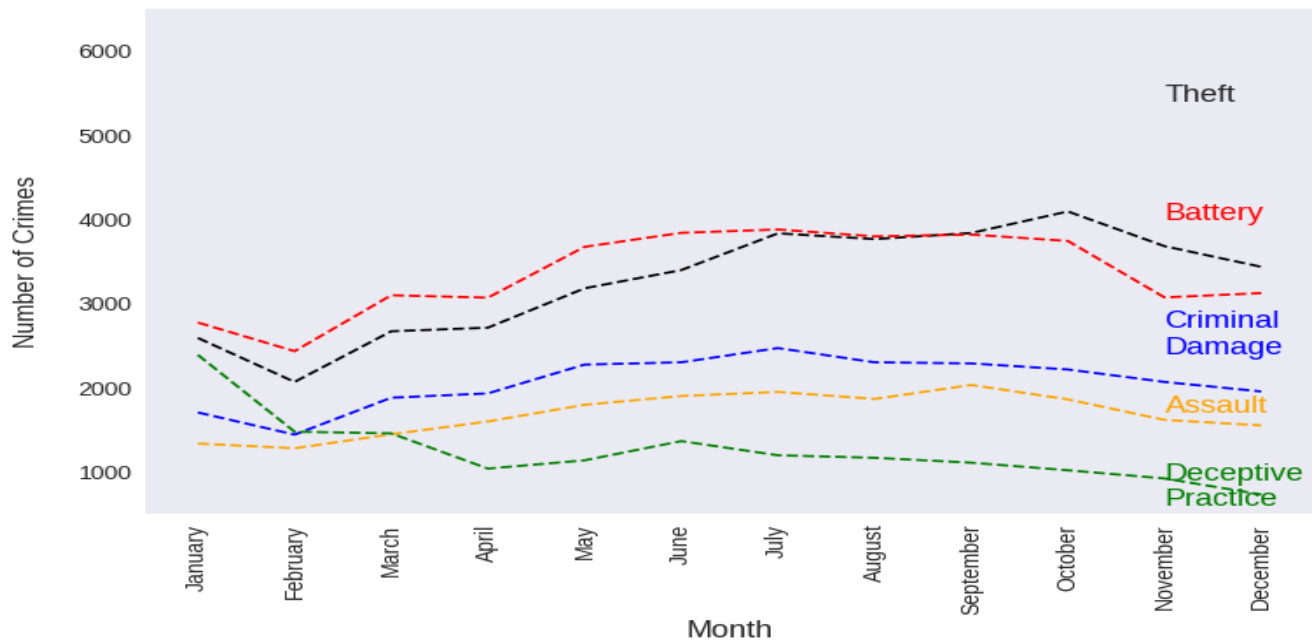


Dealing with missing values by dropping all rows that have at least one missing value.



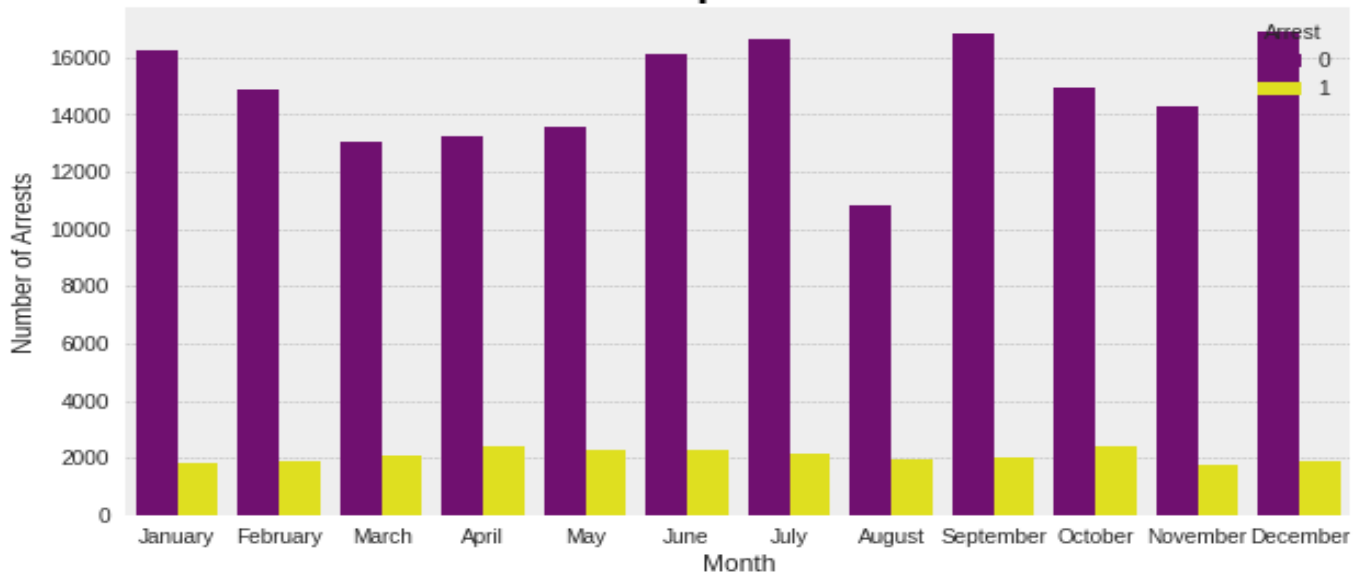
Battery is the most occurring value, and a high count of crimes indicates the presence of a physically violent community.

The frequency of Top 5 Crimes



The plot shows that battery is rising during summer. All types of crime were at their lowest in February. Assault and deceptive practice are smoothies that don't have ups and downs.

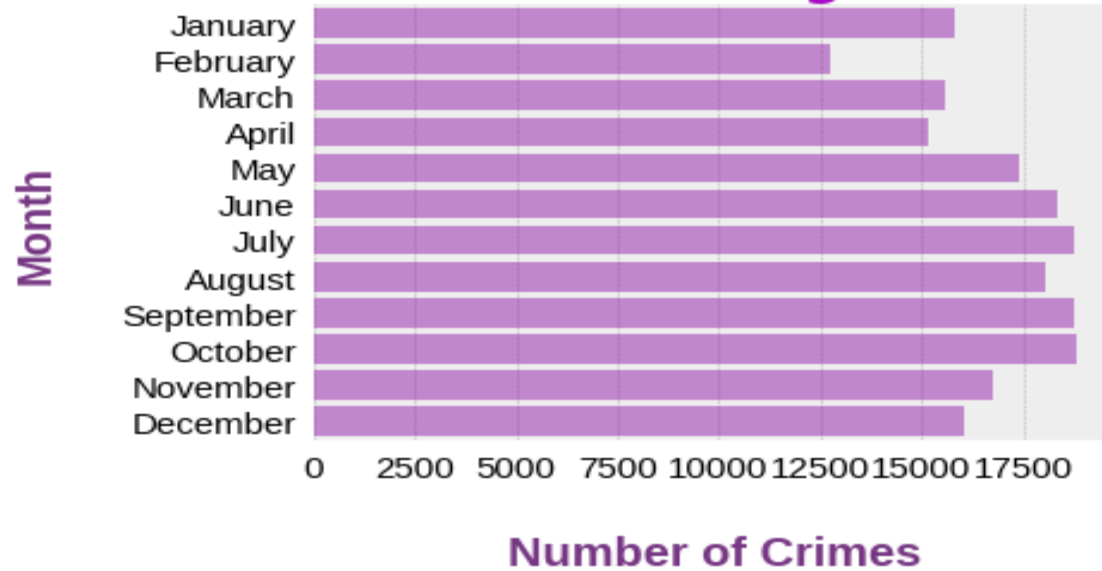
Arrests Made per Month in 2021



87 of the crimes saw no arrests of all reported crimes.

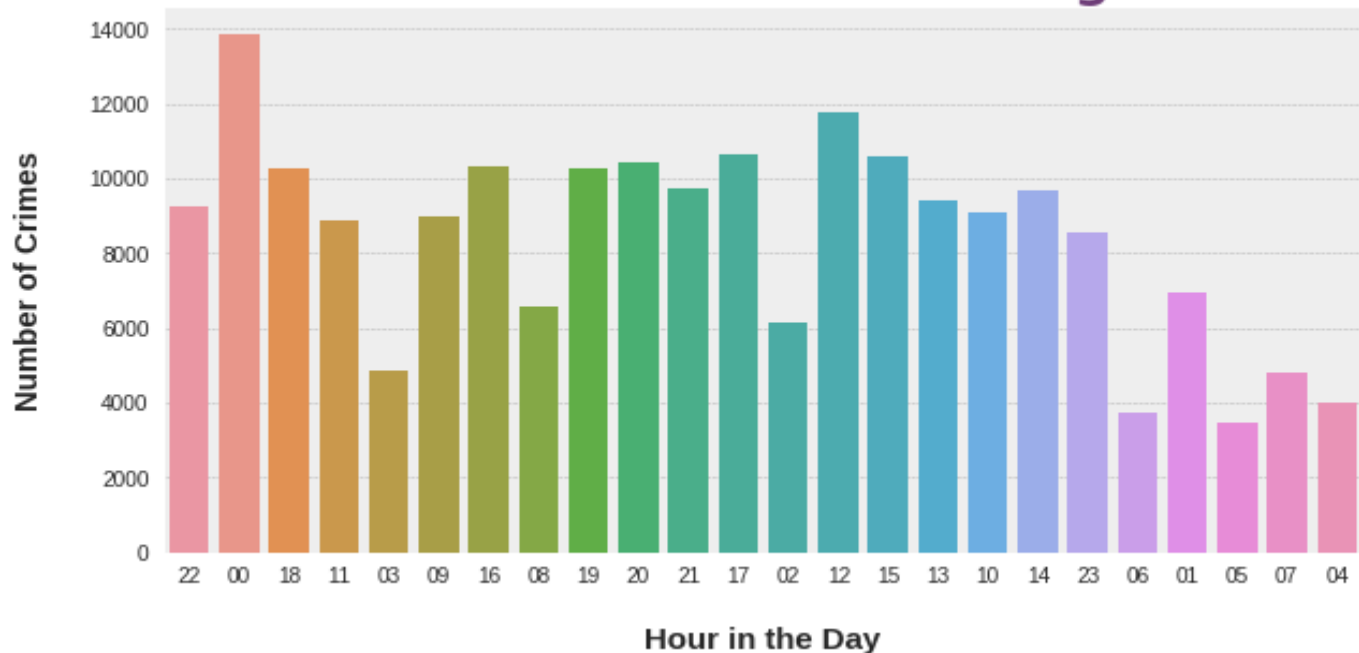
The total number of crimes with no arrest → is greater than crimes with arrest.

Crimes rise during Summer !

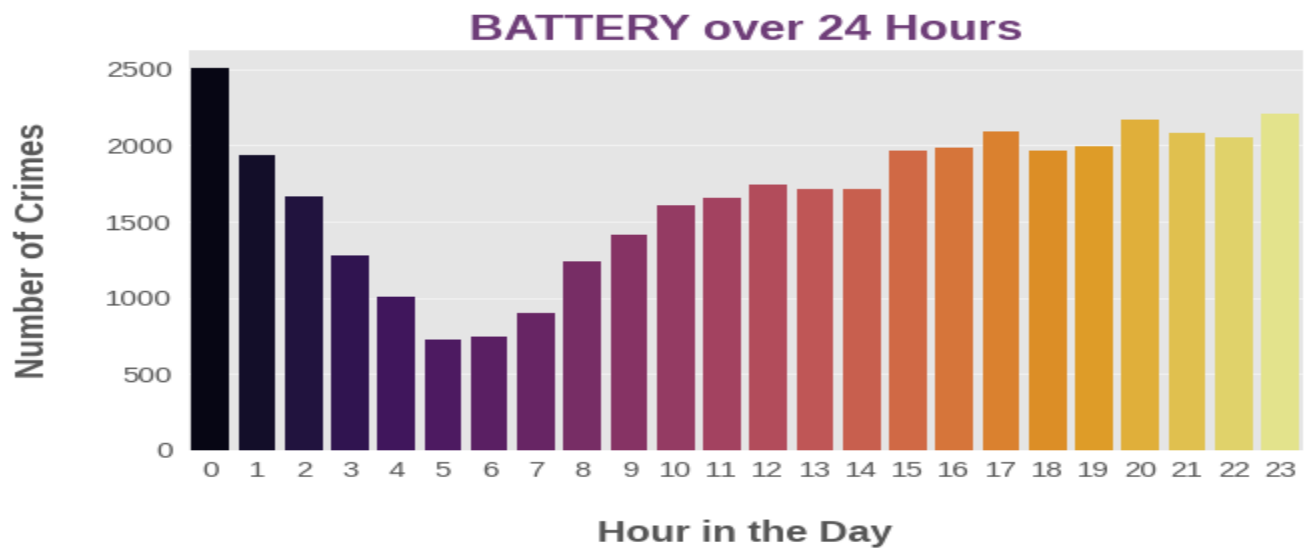


The months of July, June, September, and October → record the greatest number of crimes in the city. So, crime rises in summer.

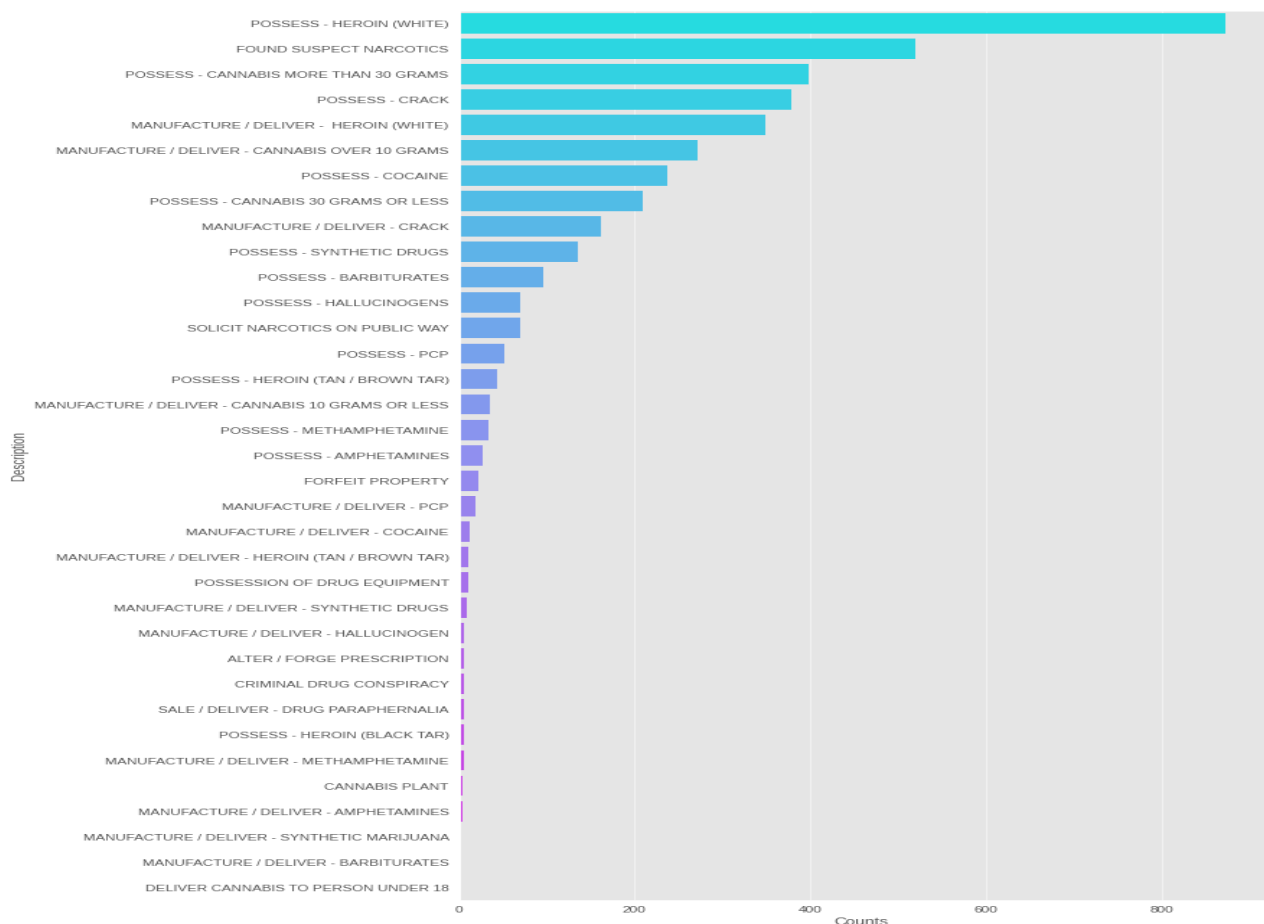
Unsafest Hours in Chicago



Peak hour for the rate of crimes in the city.



Visualization of NARCOTICS shows that POSS: HEROIN(WHITE) is the most prevalent type of drug in the city.





Plotting for District vs Month in Gang crimes

Conclusion

By looking at crime statistics, we were able to find answers to several questions we had concerning crimes in Chicago. We used a methodical approach to achieve our objectives. We found responses to questions like which crimes were committed most frequently in 2021, the riskiest times of day, the riskiest places, the most frequent gang-related activity in Chicago, and more. We also learned the startling fact that no arrests were made in 87% of the occurrences. However, we may be more optimistic and state that 13% of the crimes committed resulted in arrests within the first 24 hours of the incident being reported because the information acquired was based on first-hand accounts.