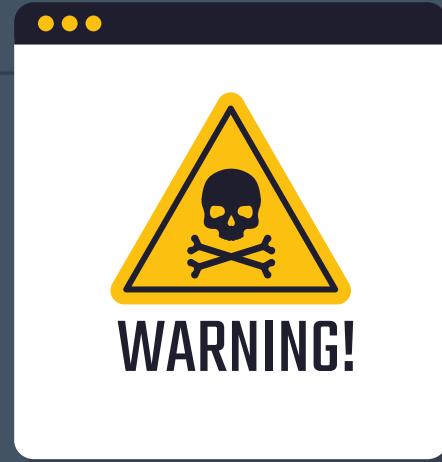


Safety Guide to Survive in Chicago

Yiyang Lu-501403



01

Data Description

Chicago Crimes dataset from
2001 to 2021

02

Why is it Big Data?

why choose this dataset

03

Problem Statement

The Goal

04

Research Questions

Questions to think about

05

Methods & Results

Integral & Detailed Analysis

06

Conclusion

Our Next Steps



Data Description

Chicago Crime 2001 to present dataset

Source:

Citizen Law Enforcement Analysis and Reporting ('CLEAR') System of **Chicago Police Department**.

Time span: 2001 to 2021

Size : 1.7 GB

Records: 7,448,538

Features: 22 columns

ID	Case number	Date	Block	IUCR	Primary Type
Description	Location Description	Arrest	Domestic	Beat	District
Ward	Community area	FBI Code	X coordinate	Y coordinate	Year
Updated on	Latitude	Longitude	Location		

Why is this Big Data ?



Why is it Big Data?

- Large data size with many features
- size grows as time goes on
- Hard to analyze on local systems

Why this dataset?

- Overwhelming news on crimes in Chicago

Problem Statement



Provide a Safety Guide to Newcomers to Chicago



Goal

Analyze the crime dataset to find out :

the overall crime distribution, arrest rate, frequent crimes, dangerous blocks
in Chicago.

Research Questions



Research Questions



1. Which times of the year do crimes occur the most?
2. Which location suffers the most from crime?
3. How has the arrest rate changed over the years?
4. Crime locations and their crime type versus arrest rate?
5. Average time of crimes occurred per month per year?
6. Crime distribution in Chicago.
7. In which year does the crime rate reach its peak ?
8. Average time took to make an arrest on a case.

Method & Results



Tools used

Languages

SQL in Pandas SQL

Python

Linux

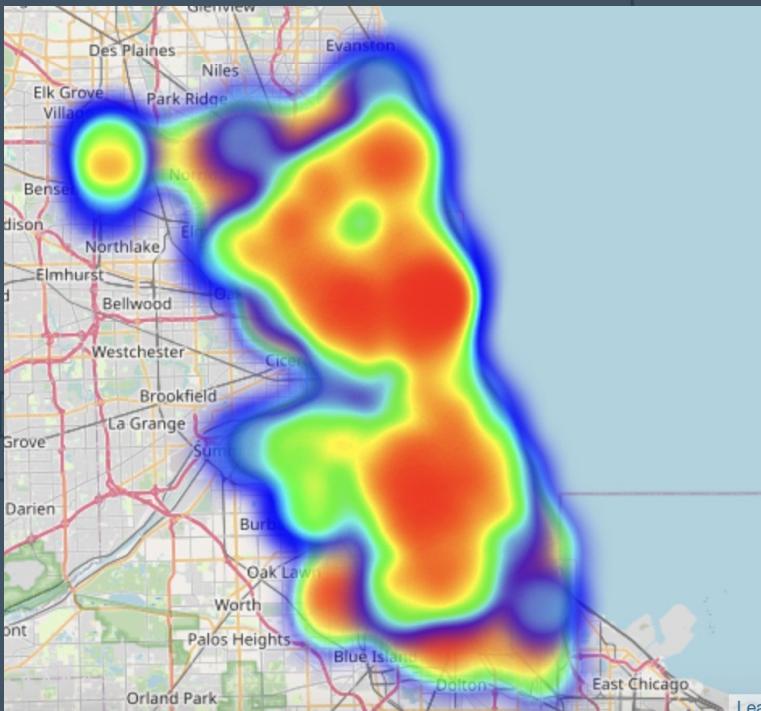
MapReduce

Visualizations

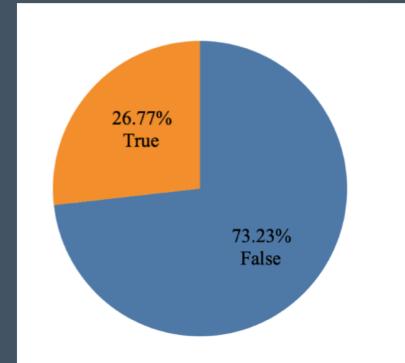
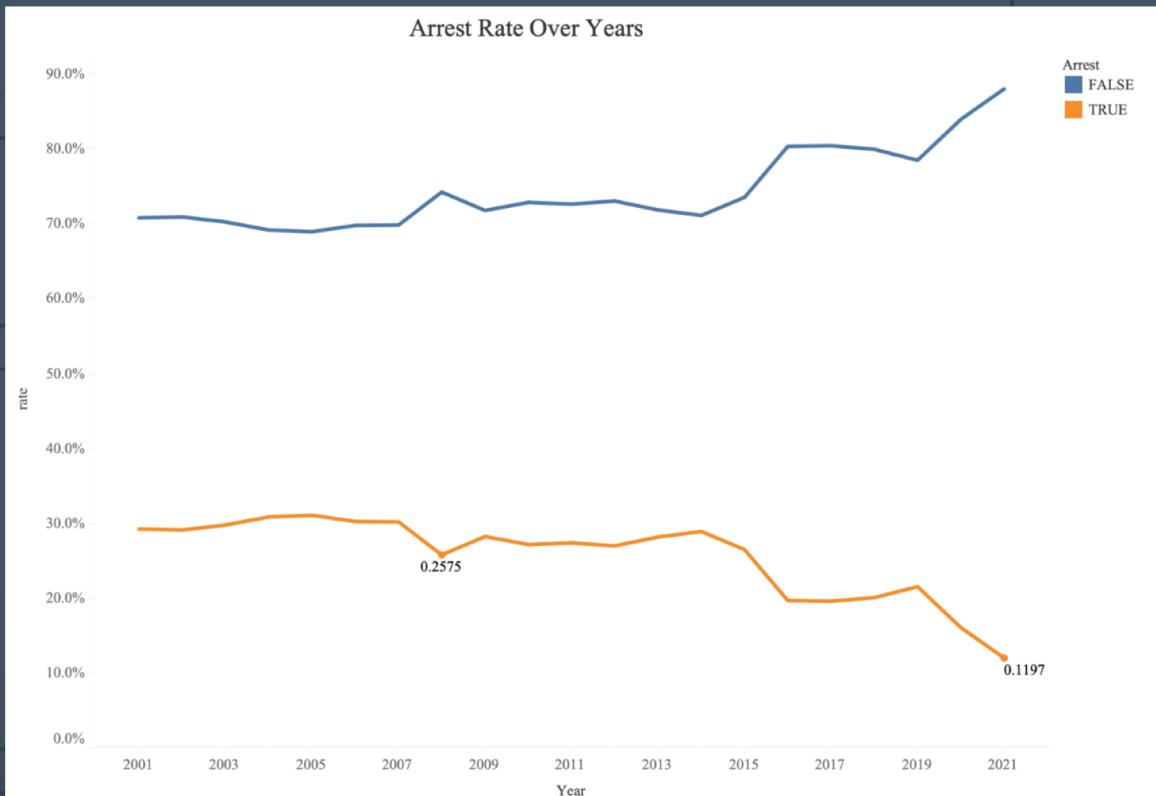
Tableau

Folium python package

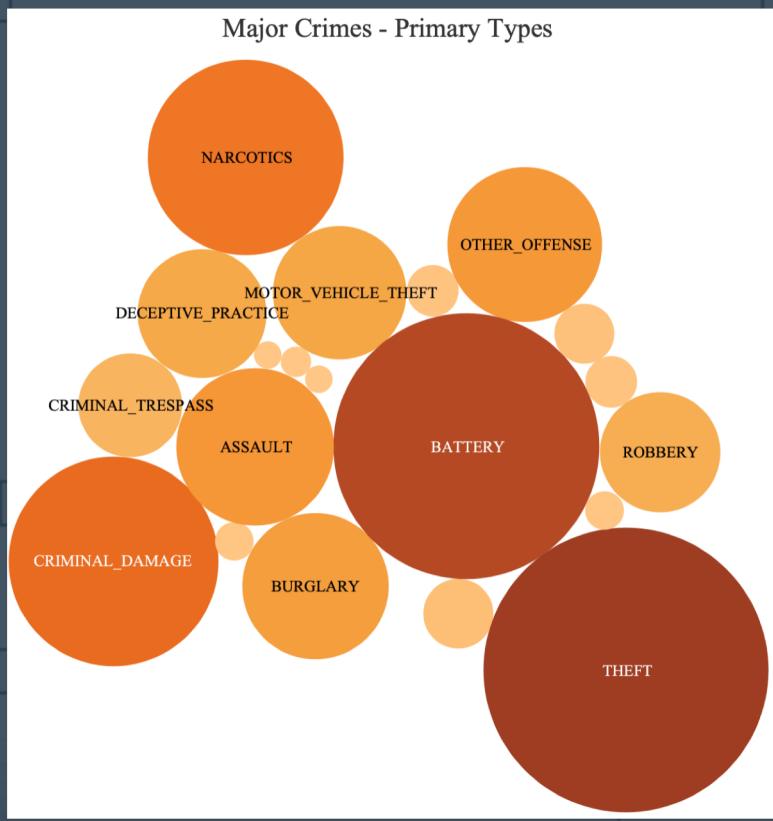
Integral Analysis-Overall Crimes



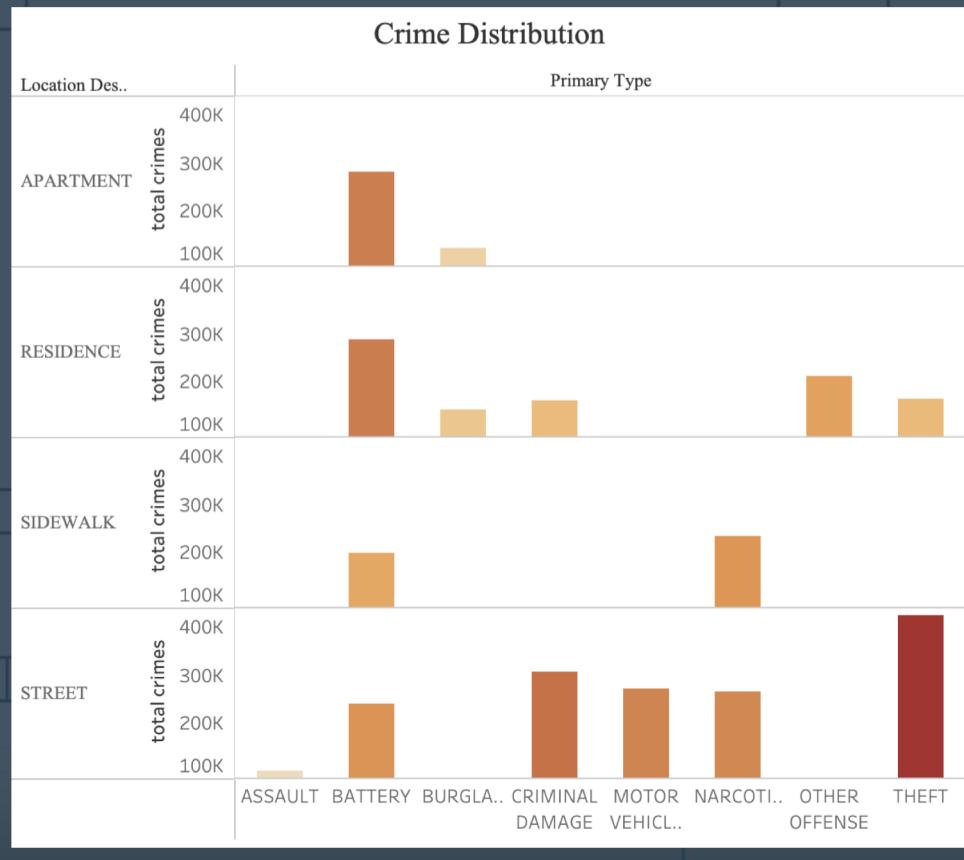
Integral Analysis- Arrest Rate lower than 21-year average



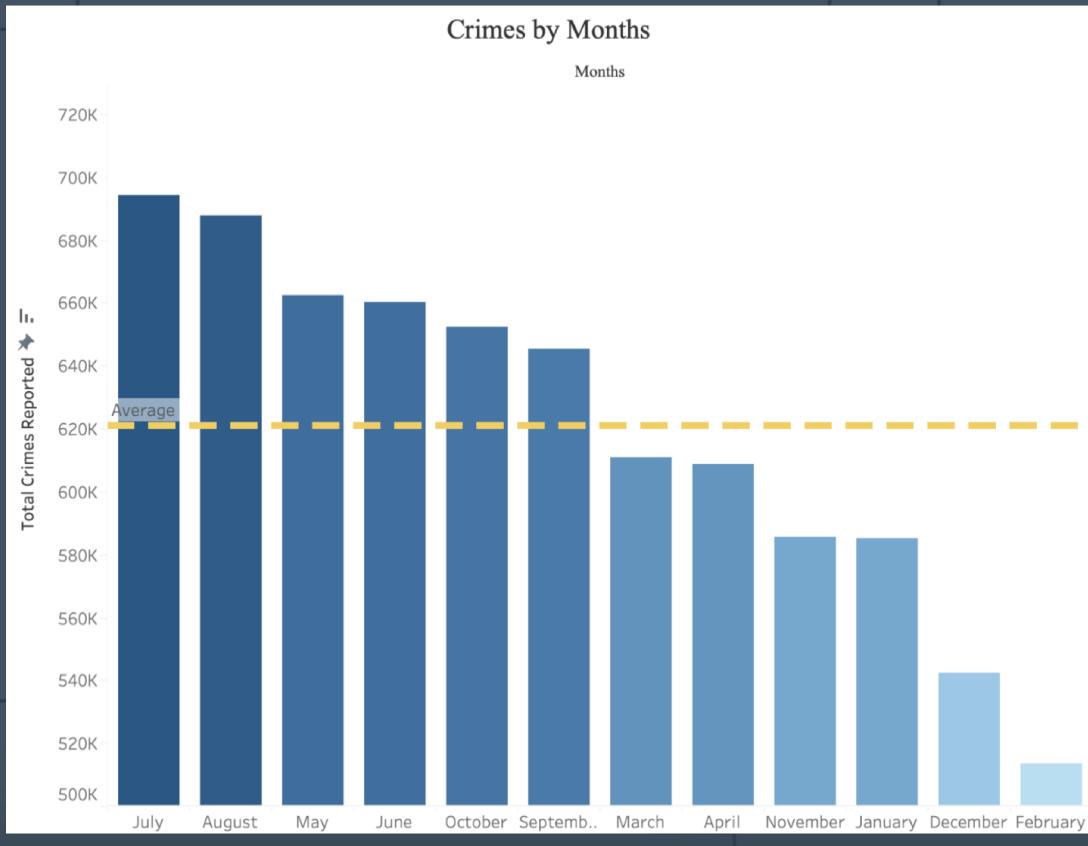
Integral Analysis – Be aware of Theft and Battery



Detailed Analysis – Conclusion



Detailed Analysis – Good Months to travel

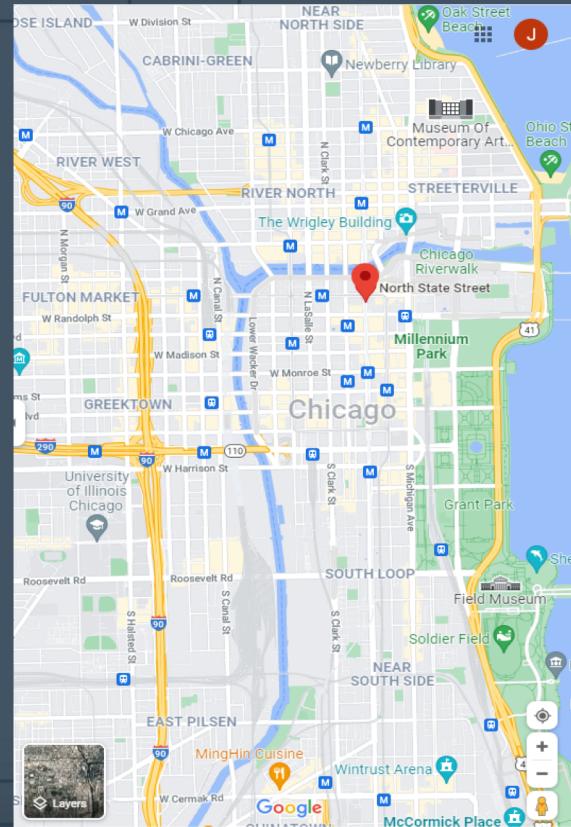
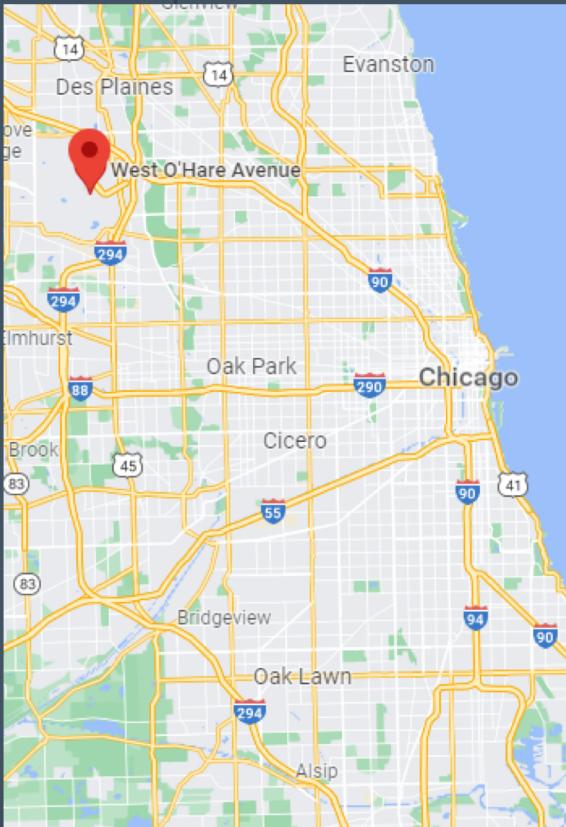


Detailed Analysis – Avoid these blocks

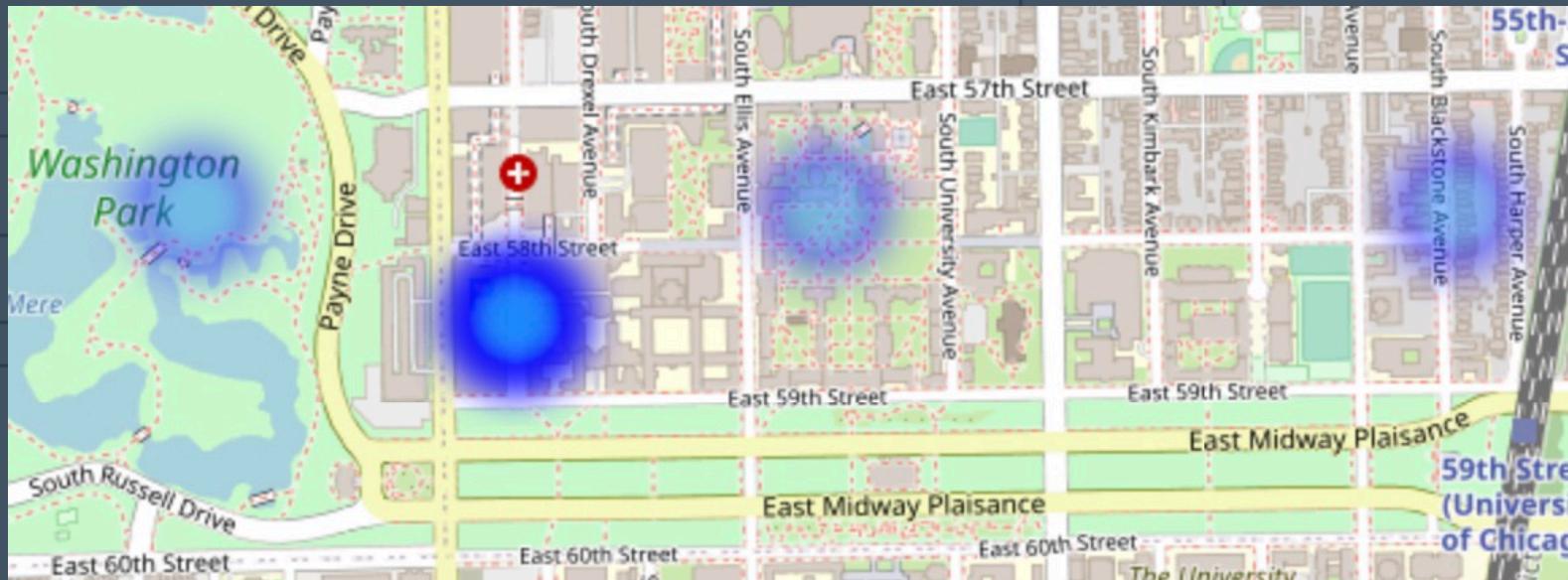
Most Dangerous Blocks in Chicago	
Block	Crimes reported
100XX W OHARE ST	16,009
001XX N STATE ST	14,303
076XX S CICERO AVE	9,821
008XX N MICHIGAN AVE	9,232
0000X N STATE ST	8,602
0000X W TERMINAL ST	5,744
064XX S DR MARTIN LUTH..	5,686
063XX S DR MARTIN LUTH..	5,403
023XX S STATE ST	5,227
001XX W 87TH ST	4,477
008XX N STATE ST	4,288
006XX N MICHIGAN AVE	4,131
0000X S STATE ST	4,080
012XX S WABASH AVE	4,079
022XX S STATE ST	4,005
009XX W BELMONT AVE	3,913
057XX S CICERO AVE	3,851
038XX W ROOSEVELT RD	3,717
075XX S STONY ISLAND A..	3,642
002XX W 87TH ST	3,627

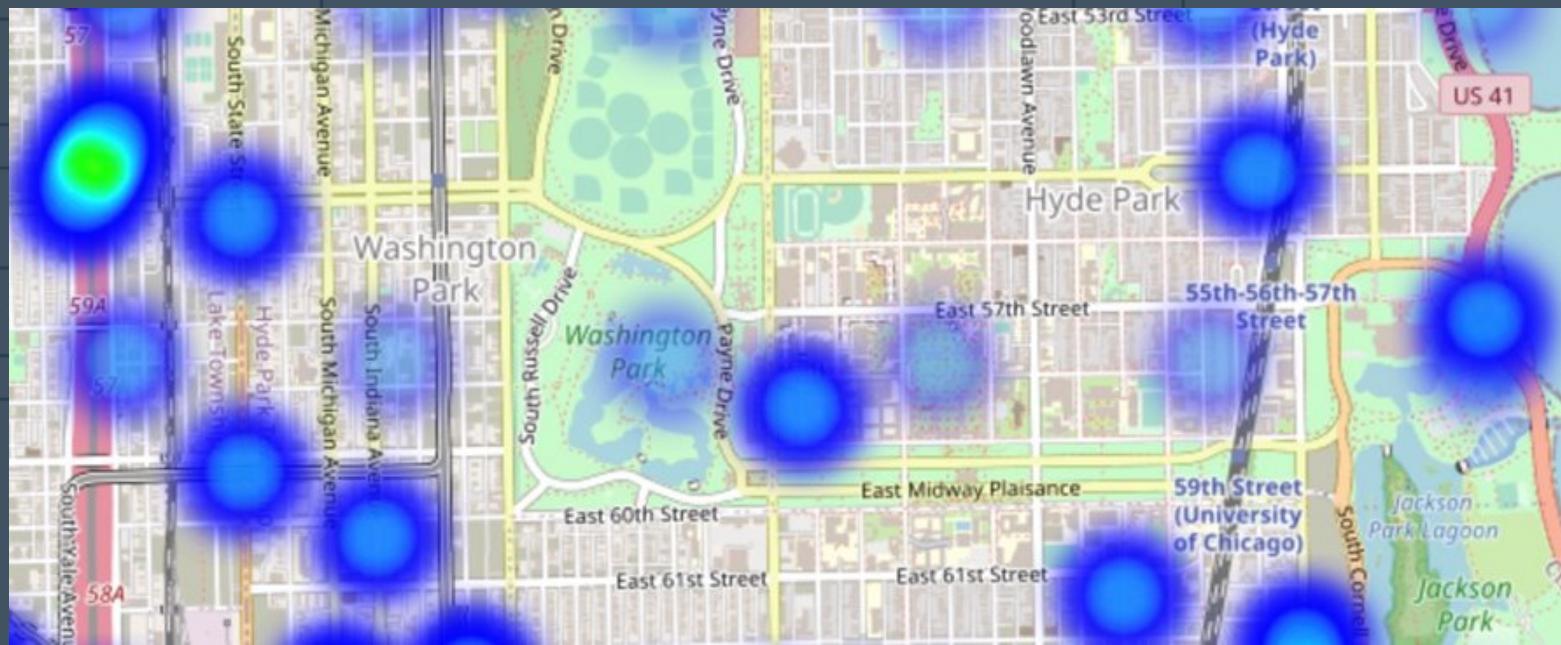
Crimes reported
3,627 16,009

Detailed Analysis – Avoid these blocks



Detailed Analysis – U of Chicago surroundings





Detailed Analysis – case stacking ratio

Year	Average Time taken to Close the Case	Case Stacking Ratio
2001	5451.89	0.71
2002	5668.3	0.78
2003	5351.43	0.77
...		
2018	39.25	0.03
2019	25.61	0.02
2020	22.4	0.03
2021	14.23	0.04

Conclusion

Future steps



Key takeaways from our research

- ⦿ A city with high crime rate.
- ⦿ safer between November and April.
- ⦿ The area around O'Hare airport dangerous(high volume).
- ⦿ Decreasing case stacking ratio(higher case process efficiency)

Recommendations:

- ⦿ Travelers had better visit between November and April and bring a safety map.
- ⦿ Will contact the police department for further clarification and corporate with them to come up with ideas to fight crime.



Thanks!

Questions?



CREDITS: This presentation template was created
by [Slidesgo](#), including icons by [Flaticon](#), and
infographics & images by [Freepik](#)