IS525: DWBI Final Project Report
Chicago Crime Analysis

Achyutha Sushanth Ariga | Jasmitha Duvvuru | Sahiti Sowmya Tadepally

Introduction:

The objective of this project is to analyze crime trends in the city of Chicago using publicly available crime data. The analysis focuses on understanding patterns of crime based on day, time, location, and type. By exploring these trends, we aim to uncover actionable insights that will help stakeholders identify critical areas of concern and take appropriate measures. The final outputs are interactive Power BI dashboards designed to provide a visual representation of crime patterns and arrest rates.

Problem Solved:

Crime remains one of the most pressing challenges in urban areas, including Chicago. The main questions we sought to answer in this project were: What times and days of the week witness the highest crime activity? Which types of crimes occur most frequently based on locations, and what is the overall arrest rate? Additionally, we wanted to observe how crime trends have changed over time and identify if certain crime types are more likely to lead to arrests. Answering these questions will allow decision-makers to allocate resources effectively and address the areas and times of heightened criminal activity.

Intended Audience:

This project is intended for several key stakeholders. Here are potential target audiences for crime analysis:

1. Law Enforcement Agencies:

Purpose:

- Identify crime hotspots and trends to allocate resources efficiently.
- Plan proactive measures during high-crime periods.
- Use predictive analysis to forecast and prevent potential crimes.

2. Residents and Neighborhood Associations:

Purpose:

- Understand local crime patterns for personal safety.
- Collaborate with law enforcement to address community-specific concerns.
- Advocate for better resources like lighting, security cameras, or police presence.

3. Policymakers and City Officials

Purpose:

- Shape crime prevention policies and urban planning strategies.
- Address socioeconomic factors contributing to crime in specific community areas or districts.
- Measure the impact of previous initiatives like increased patrols or surveillance.

4. Community Organizations
Purpose:

- Empower local groups to advocate for safer neighborhoods.
- Develop outreach programs to reduce crime in vulnerable areas.
- Track the effectiveness of community-led crime prevention campaigns.

Dataset Overview:

The data used for this project was sourced from the City of Chicago Open Data Portal, specifically from the dataset titled *"*Crimes - 2001 to Present*"*. This dataset contains detailed information about crimes that have occurred in Chicago, including crime type, location, arrest status, and timestamps. We have also got datasets for Community Areas and Police Districts to integrate with the area and district codes in the main crimes dataset. The use of such reliable and publicly available data ensures that the analysis is credible and up-to-date.

Client's Perspective:

Client Details : Anand Marepalli (anandm5@illinois.edu)

After reviewing the three dashboards, the visualizations provide valuable insights into crime patterns in Chicago over time. Anand is a student at UIUC who likes travelling and who often visits to chicago. Our friend Anand noted that the dashboards are intuitive and effectively communicate trends such as day-wise crime patterns, top crime types, and location-based data. The use of heatmaps, pie charts, and line graphs is appreciated, especially in breaking down complex datasets. However, he suggested a few areas of improvement to enhance clarity and visual appeal. For instance, the heatmap in the first dashboard could use a gradient color scale to better distinguish crime frequencies across hours. The gauge chart for solved crimes in one of the dashboards could benefit from an explicit goal or benchmark line for better comparison. Additionally, incorporating stacked bar charts or area charts to show arrests vs. non-arrests by crime type over time would add depth to the analysis.

Overall, the client found the dashboards informative but recommended minor visual tweaks and additional charts to enhance interactivity, readability, and predictive insights. We have incorporated all of these changes for our final output and enhanced our dashboards.

Steps Taken:

To prepare the dataset for meaningful visualization and analysis, we began by integrating the Chicago Crimes Dataset (2001-Present) with additional data sources to enhance its usability. This included adding community area names and police district names by mapping them to the corresponding codes present in the original dataset. These enrichments provided more context for geographic analysis, making it easier to interpret crime patterns in specific regions.

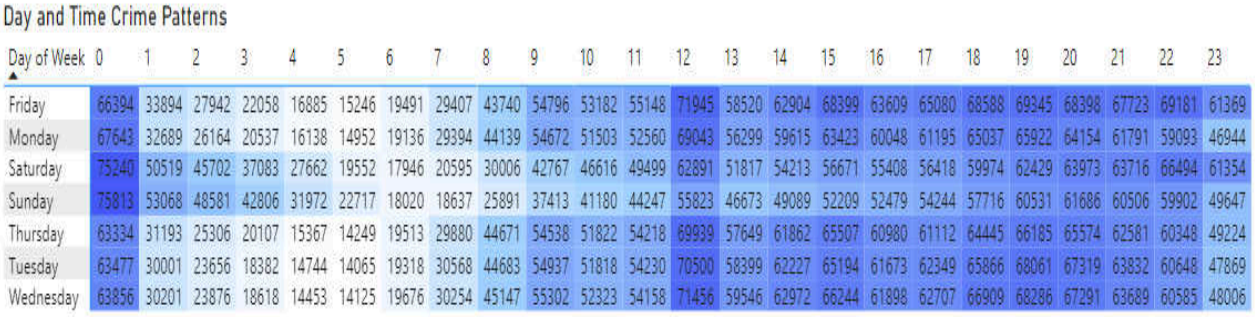
We then performed data cleaning to ensure the dataset was ready for visualization. This involved identifying and handling null values, particularly in critical fields like location, crime type, and date. Missing or incomplete entries were either imputed or excluded to maintain data accuracy. Formatting columns appropriately was another key step, such as converting

date-time fields into separate columns for month, time, and day to enable more granular filtering and analysis in dashboards.

Finally, we filtered out unnecessary information, such as redundant fields or records outside the desired analysis period, to reduce noise and improve processing efficiency. By organizing and structuring the data in this way, we created a clean, enriched dataset that supports intuitive filtering and enables deeper insights through visualizations like heatmaps, time-series charts, and geographic overlays in Power BI. Once the dashboard was completed, we shared it with the client for feedback and incorporated suggestions to make the insights more focused on arrest data and overall crime trends.

Analysis and Key Discoveries:

Below are few of the interesting visualizations covered as part of the three dashboards:



Day and Time Crime Patterns

Day of Week	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Friday	66394	33894	27942	22058	16885	15246	19491	29407	43740	54796	53182	55148	71945	58520	62904	68399	63609	65080	68588	69345	68398	67723	69181	61369
Monday	67643	32689	26164	20537	16138	14952	19136	29394	44139	54672	51503	52560	69043	56299	59615	63423	60048	61195	65037	65922	64154	61791	59093	46944
Saturday	75240	50519	45702	37083	27662	19552	17946	20595	30006	42767	46616	49499	62891	51817	54213	56671	55408	56418	59974	62429	63973	63716	66494	61354
Sunday	75813	53068	48581	42806	31972	22717	18020	18637	25891	37413	41180	44247	55823	46673	49089	52209	52479	54244	57716	60531	61686	60506	59902	49647
Thursday	63334	31193	25306	20107	15367	14249	19513	29880	44671	54538	51822	54218	69939	57649	61862	65507	60980	61112	64445	66185	65574	62581	60348	49224
Tuesday	63477	30001	23656	18382	14744	14065	19318	30568	44683	54937	51818	54230	70500	58399	62227	65194	61673	62349	65866	68061	67319	63832	60648	47869
Wednesday	63856	30201	23876	18618	14453	14125	19676	30254	45147	55302	52323	54158	71456	59546	62972	66244	61898	62707	66909	68286	67291	63689	60585	48006

Fig.1: Heat Map showing Day and Time Crime Patterns

This heat map shown in Fig.1 visualizes crime patterns in Chicago based on the day of the week and time of day, using data from the "Crimes - 2001 to Present" dataset. The rows represent days of the week from Friday to Wednesday, while the columns show hours of the day from 0 to 23. Each cell contains a number, representing the count of reported crimes, with color intensity indicating the frequency of incidents. Darker blue shades suggest higher crime rates, while lighter colors indicate fewer reported crimes. The heatmap reveals distinct temporal patterns in criminal activity across Chicago, with noticeable variations between different days and times. For instance, late night and early morning hours on weekends (particularly Friday and Saturday) appear to have higher crime rates, as indicated by the darker blue cells. This visualization allows for quick identification of peak crime times and potential trends in criminal activity throughout the week, which could be valuable for law enforcement resource allocation and crime prevention strategies.

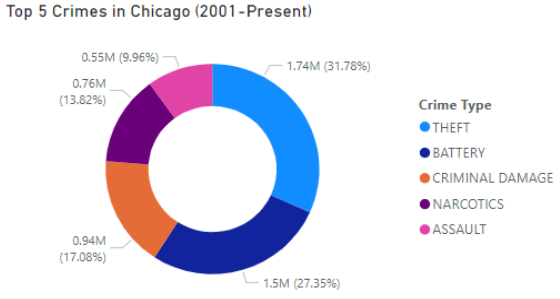


Fig.2 : Donut Chart with top 5 crimes in Chicago

The donut chart as shown in Fig.2 illustrates the distribution of the top five crime types reported in Chicago from 2001 to present, revealing a clear hierarchy in criminal activity. Theft dominates as the most prevalent crime, accounting for 31.78% of incidents with approximately 1.74 million cases, followed closely by battery at 27.35% with 1.5 million incidents. Criminal damage ranks third at 17.08% with 0.94 million cases, while narcotics-related crimes constitute 13.82% with 0.76 million incidents. Assault rounds out the top five at 9.98% with 0.55 million reported cases. This visualization effectively demonstrates that property crimes, particularly theft, and violent crimes like battery represent the majority of criminal activity in Chicago during this period, with these five categories collectively accounting for all major reported crimes in the dataset.

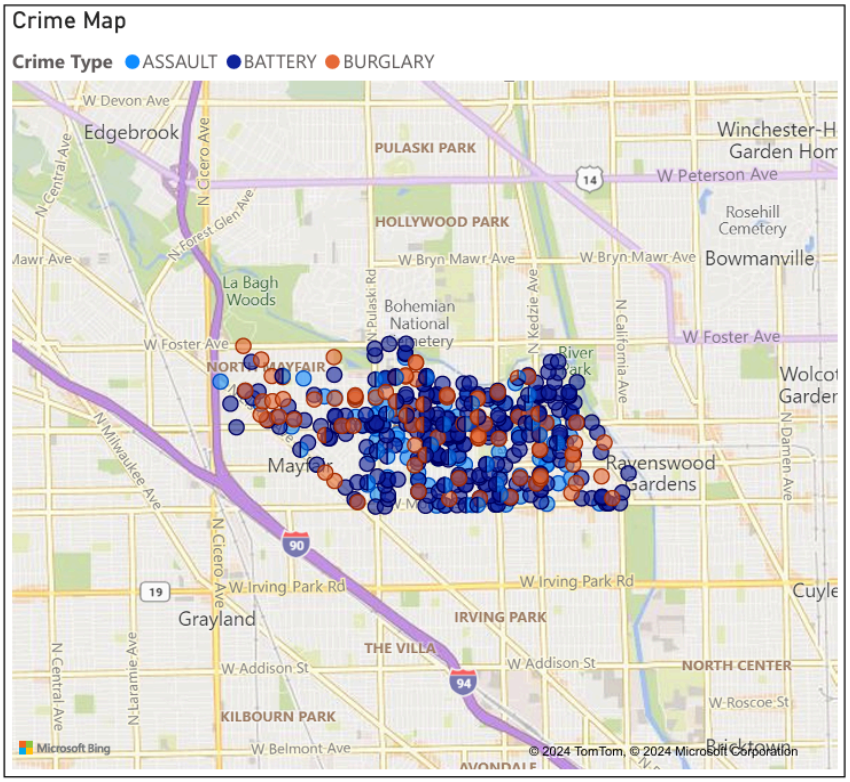


Fig.3 : Geographical Map showing type of crimes

The map visualization as seen in Fig.3 displays the geographic distribution of three specific crime types - assault, battery, and burglary - across Chicago's north side neighborhoods, particularly concentrated in the area between Irving Park and Edgebrook. The crimes are plotted as colored dots, with blue dots representing battery incidents, light blue for assaults, and orange for burglaries. The highest concentration of incidents appears in the Ravenswood Gardens area, with a particularly dense cluster of battery cases (blue dots) in this neighborhood. The visualization also shows a notable spread of burglaries (orange dots) throughout the Mayfair area, while criminal incidents appear to decrease in frequency towards the outer boundaries of the mapped region near LaBagh Woods and Pulaski Park. The map effectively illustrates how these three crime types are distributed spatially, revealing potential hotspots and patterns of criminal activity in these northern Chicago communities.

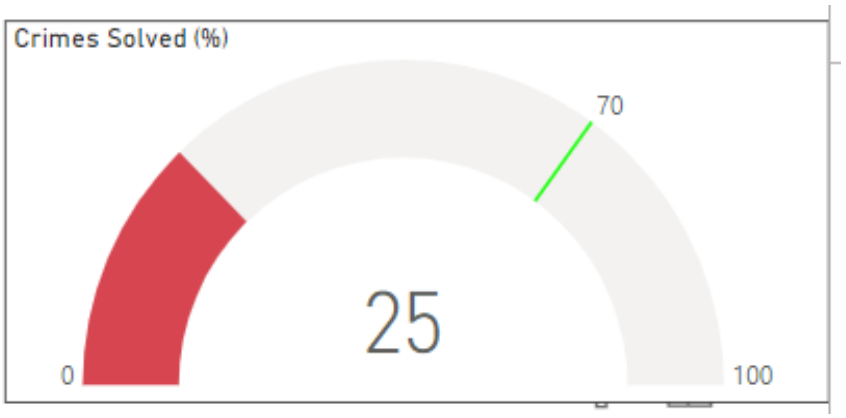


Fig.4 : Gauge Visualization with Crimes Solved%

The gauge visualization as shown in Fig.4 presents a stark picture of crime resolution in Chicago, displaying a 25% solve rate for reported crimes. The semicircular gauge uses a color gradient from red to white to green, with the needle pointing to 25 on a scale of 0 to 100, indicating that only a quarter of all reported crimes in the dataset reach resolution. The red section on the left side of the gauge emphasizes the concerning nature of this low clearance rate, while the predominantly white and green sections to the right represent the aspirational higher solve rates that remain unachieved.

Challenges and Resolutions:

During the project, we faced a few challenges. One of the major issues was managing the large size of the dataset, which initially caused performance lags in Power BI. To address this, we filtered the data to focus on relevant time periods and aggregated metrics for analysis. Another challenge was visualizing the day and time crime patterns in a way that was easy to interpret. We resolved this by creating a heatmap that displays crime counts by day of the week and hour of the day, which effectively highlights patterns of criminal activity.

Adjustments to the Original Plan:

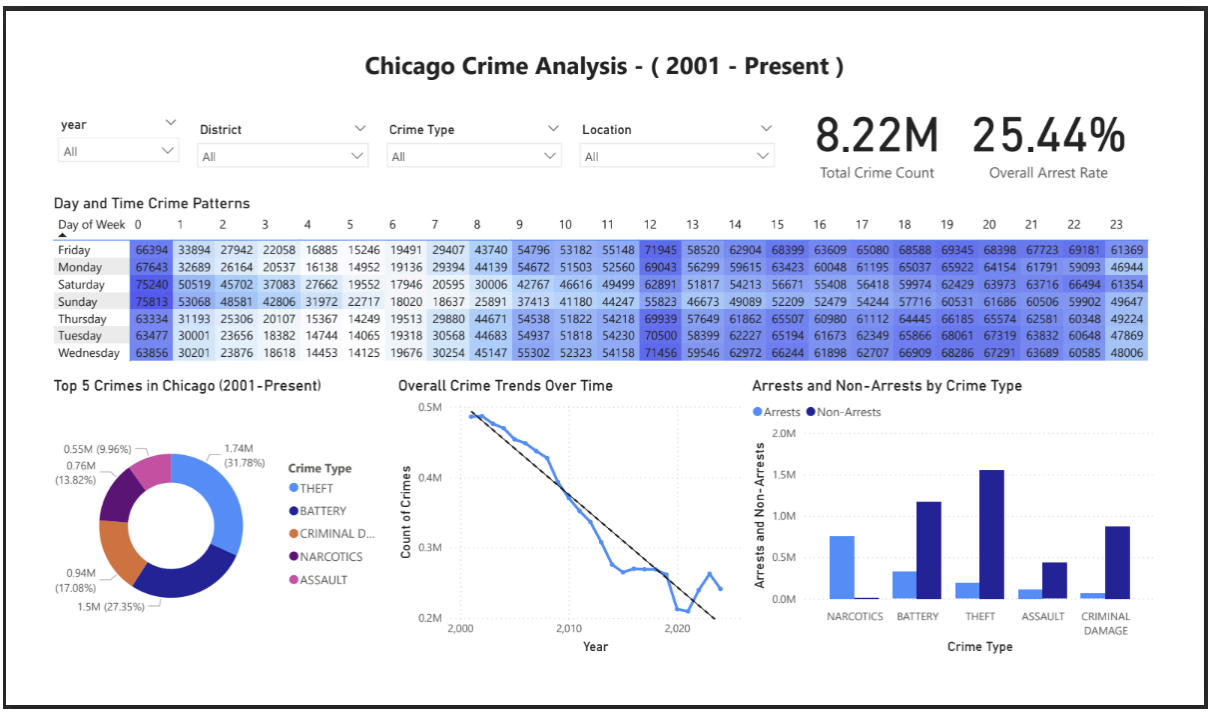
Originally, we intended to analyze crime trends at a granular level, focusing on individual neighborhoods which made us integrate the main dataset with community area names and police district names too. However, due to the large size of the dataset and performance considerations, the scope was adjusted to analyze citywide trends, overall patterns, and arrest data. This adjustment allowed for a more focused and efficient analysis without compromising the key objectives of the project.

Client Feedback and Incorporation:

After presenting the initial version of the dashboard, the client provided valuable feedback. One of the key requests was to include a more detailed analysis of arrests across different crime types. To address this, we changed heat map and gauge chart characteristics and also added a bar chart comparing arrests and non-arrests for the top crime categories. This addition provided the client with a clearer understanding of law enforcement efforts and arrest rates for specific crimes.

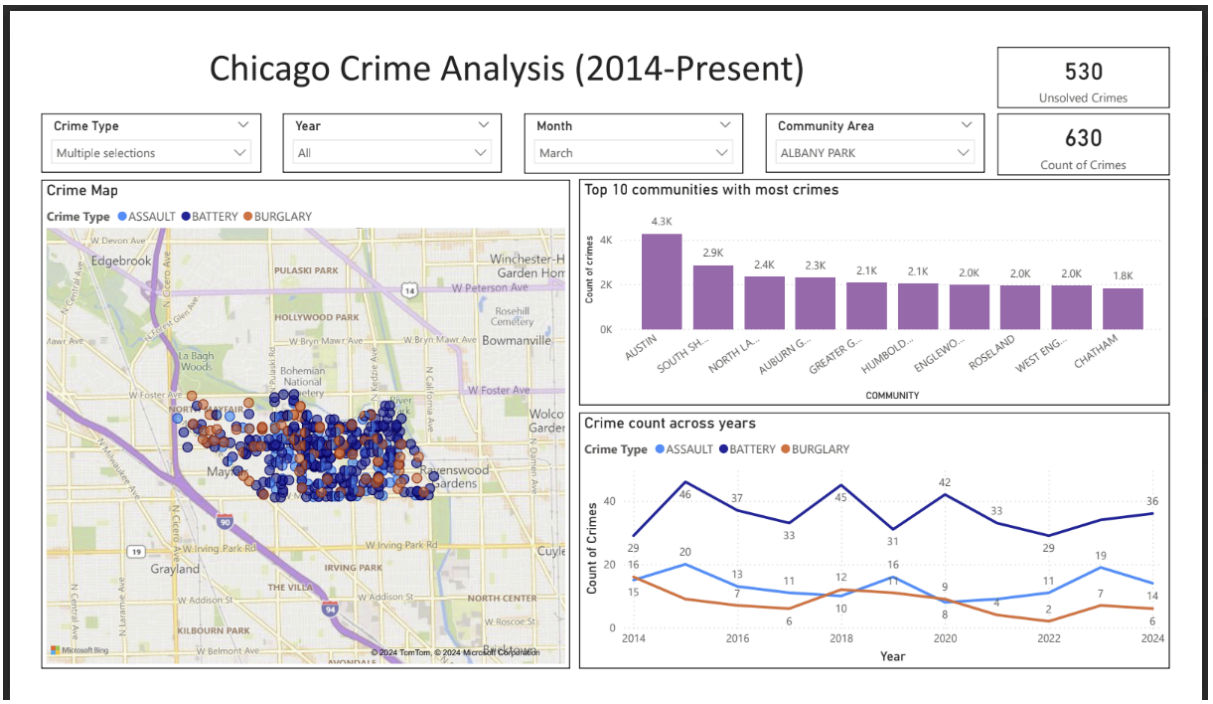
Final Product Screenshots:

Dashboard 1:

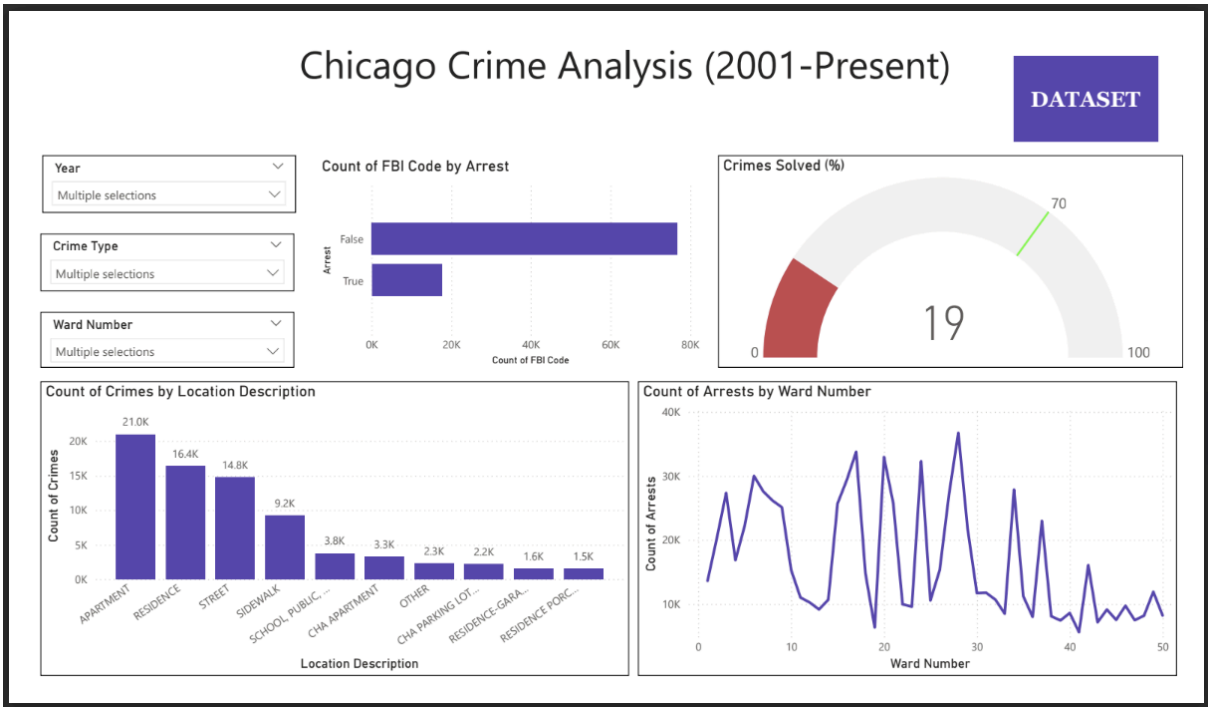


Day of Week	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Friday	66394	33894	27942	22058	16885	15246	19491	29407	43740	54796	53182	55148	71945	58520	62904	68399	63609	65080	68588	69345	68398	67723	69181	61369
Monday	67643	32689	26164	20537	16138	14952	19136	29394	44139	54672	51503	52560	69043	56299	59615	63423	60048	61195	65037	65922	64154	61791	59093	46944
Saturday	75240	50519	45702	37083	27662	19552	17946	20595	30006	42767	46616	49499	62891	51817	54213	56671	55408	56418	59974	62429	63973	63716	66494	61354
Sunday	75813	53068	48581	42806	31972	22717	18020	18637	25891	37413	41180	44247	55823	46673	49089	52209	52479	54244	57716	60531	61686	60506	59902	49647
Thursday	63334	31193	25306	20107	15367	14249	19513	29880	44671	54538	51822	54218	69939	57649	61862	65507	60980	61112	64445	66185	65574	62581	60348	49224
Tuesday	63477	30001	23656	18382	14744	14065	19318	30568	44683	54937	51818	54230	70500	58399	62227	65194	61673	62349	65866	68061	67319	63832	60648	47869
Wednesday	63856	30201	23876	18618	14453	14125	19676	30254	45147	55302	52323	54158	71456	59546	62972	66244	61898	62707	66909	68286	67291	63689	60585	48006

Dashboard 2:



Dashboard 3:



Conclusion:

This analysis successfully uncovered critical insights into Chicago's crime patterns. The findings revealed when crimes are most likely to occur, which crime types are most prevalent, and how arrest rates differ across categories. By visualizing these trends in a user-friendly Power BI dashboard, the project equips stakeholders such as law enforcement, city planners, and policymakers with valuable information to improve public safety. The client, Anand, was able to gain a better understanding of crime trends in his city and make informed decisions about his daily routines and safety.