## **Objective Questions**

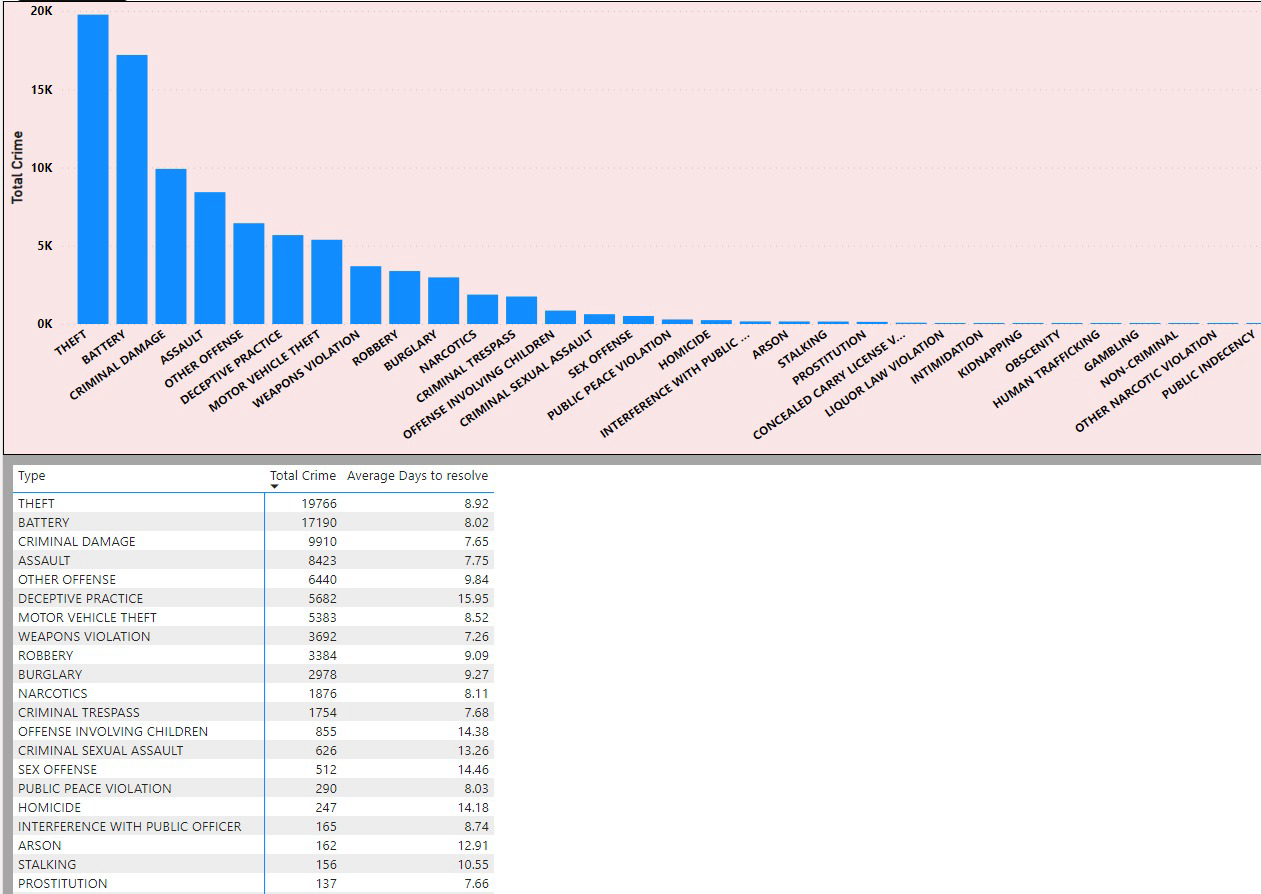
**1. In analyzing the provided dataset with Power BI, ensure data cleaning to address inconsistencies and missing values before further analysis.**

In analyzing the Crime dataset with Power BI, I applied following changes :

1. Removing null values from
   1. Ward
   2. X coordinate
   3. Y coordinate
2. Arrest Binary
   1. Assigned 1 for True and 0 for False from the Arrest Column
   2. Using the formula Arrest Binary = IF('crimes\_data\_2022 - crimes\_data\_2022 (1)'[Arrest],1,0)
3. Resolve Case Time
   1. Calculated the date difference between Date and Updated.
   2. Using the formula Resolve Case Time =DATEDIFF('crimes\_data\_2022 - crimes\_data\_2022 (1)'[Date],'crimes\_data\_2022 crimes\_data\_2022 (1)'[Updated On],DAY)
4. Month
   1. Extracted the month from the Date column using the formula Month= FORMAT('crimes\_data\_2022 - crimes\_data\_2022 (1)'[Date],"mmmm")
5. Latitude and Longitude
   1. Created using the Location column through the “Add Column by Example” feature.
6. Changed data types
   1. Date and Updated On to Date data type.
   2. Beat, District, Ward, Community Area, FBI Code to Text data type.

**2. Crime Type Analysis: Assess the frequency of each crime type to identify the most prevalent crimes occurring in the area.**

I examined various categories of criminal activities by devising a metric called Total Crime, and I depicted the findings through a clustered column chart. The analysis unveiled “theft” as the prevailing offense, with a recorded total of 19,766 incidents.



Location - Locality Tab

“Total Crime and Average days to resolve by Type” Visual.

**3. Arrest Rate Evaluation: Analyze the percentage of reported incidents that have resulted in an arrest to gauge law enforcement effectiveness.**

For evaluating law enforcement effectiveness, I delved into the arrest rate, calculated at 12.42%. This figure was derived by computing the average of the 'Arrest binary' column from the 'crimes data 2022' dataset.

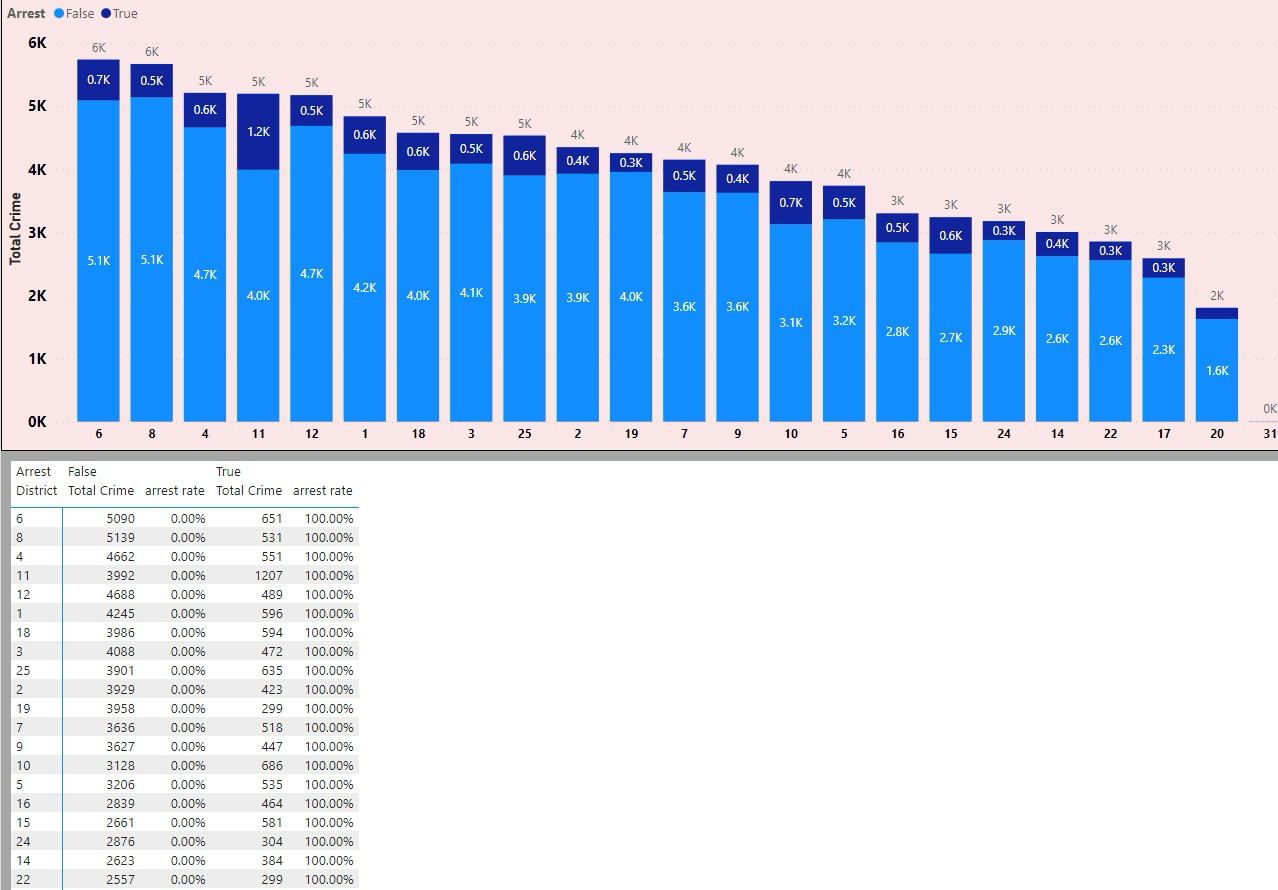


Location- Main Tab

“Arrest Rate” Visual.

**4. District Crime Distribution Assessment: Calculate the number of crimes in each district to understand how crime is distributed across the city and identify high-crime areas.**

To comprehensively assess the distribution of crime throughout the city and pinpoint areas of heightened criminal activity, I conducted an analysis to determine the number of crimes occurring in each district. Employing a Stacked Column Chart, I revealed that District 6 exhibited the highest count of crime incidents, tallying up to 5,741 reported cases.



Location – Exploration Tab

“Total crime and Arrest rate by Description and Arrest” visual. Under District Field parameter.

**5. How many categorical attributes are there in the data?**

I identified several categorical columns in the dataset, including:-

1. Case Number.
2. Block.
3. IUCR.
4. Type.
5. Location Description.
6. Arrest.
7. Domestic.
8. Beat.
9. District.
10. Ward.
11. Community Area.
12. FBI Code.
13. Year.

**6. Were there any Null values in the data, if there were, how did you handle them? What is the ideal way to handle Null values?**

Upon scrutinizing the dataset, I identified null values present in the Ward, X-coordinate, and Y-coordinate columns, accounting for approximately 1% of the total dataset. In order to uphold data integrity and given the negligible proportion, I opted to eliminate these null values. Their removal is not anticipated to exert a substantial influence on the overall dataset.

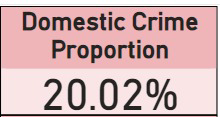
**7. Domestic Crime Proportion Analysis: Analyze the ratio of domestic-related crimes to other types of crimes to understand the prevalence of domestic incidents.**

I determined that Domestic Crime accounts for 20.02% of the total crimes in Chicago.

To showcase this insight, I incorporated a card visual in the Main tab, specifically featuring the Domestic Crime Proportion with the following DAX formula:

Domestic Crime Proportion = DIVIDE(SUM(‘crimes data 2022’,[Domestic binary]),COUNT(‘crimes data 2022’,[Domestic binary]))

This visual provides a clear representation of the proportion of domestic crimes in the overall dataset.

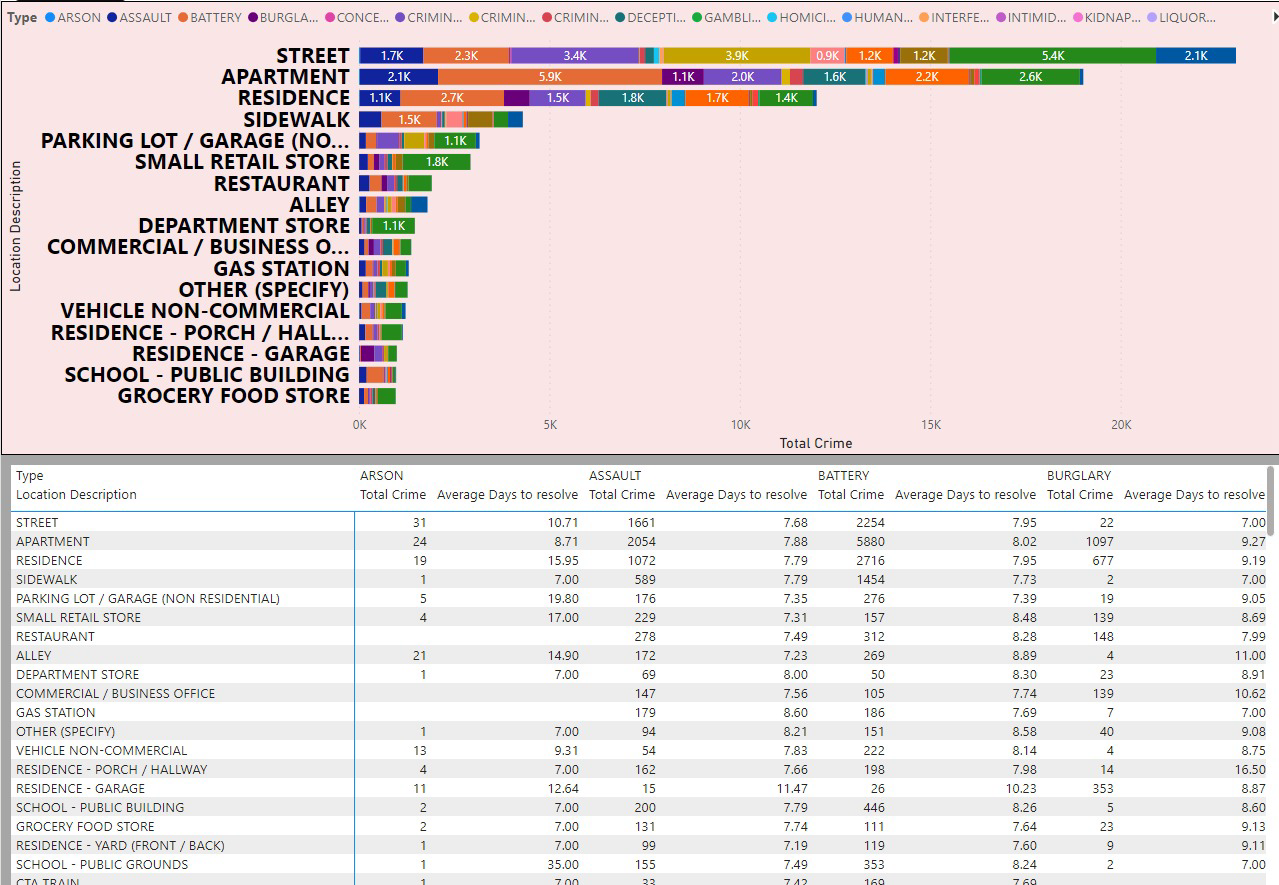


Location – Main Tab

“Domestic Crime Proportion” Visual.

**8. Is there any “Location Description” where the number of crimes is higher than expected? Come up with a table or visualization in which one can judge the frequency of crimes at each Location Description type.**

Street location highest crime frequency reaching 18.8K incidents.



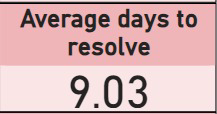
Location – Locality Tab.

“Total Crime and average days to resolve by Location Description and Type” Visual.

**9. What is the average time between reporting and solving a case as per the data?**

The average time to solve a crime case is calculated to be 9.03 days, determined using the DAX formula: Average days to resolve = AVERAGE (‘crimes data 2022’[Resolving case time])

This metric provides insights into the typical duration it takes to resolve cases based on the available data.



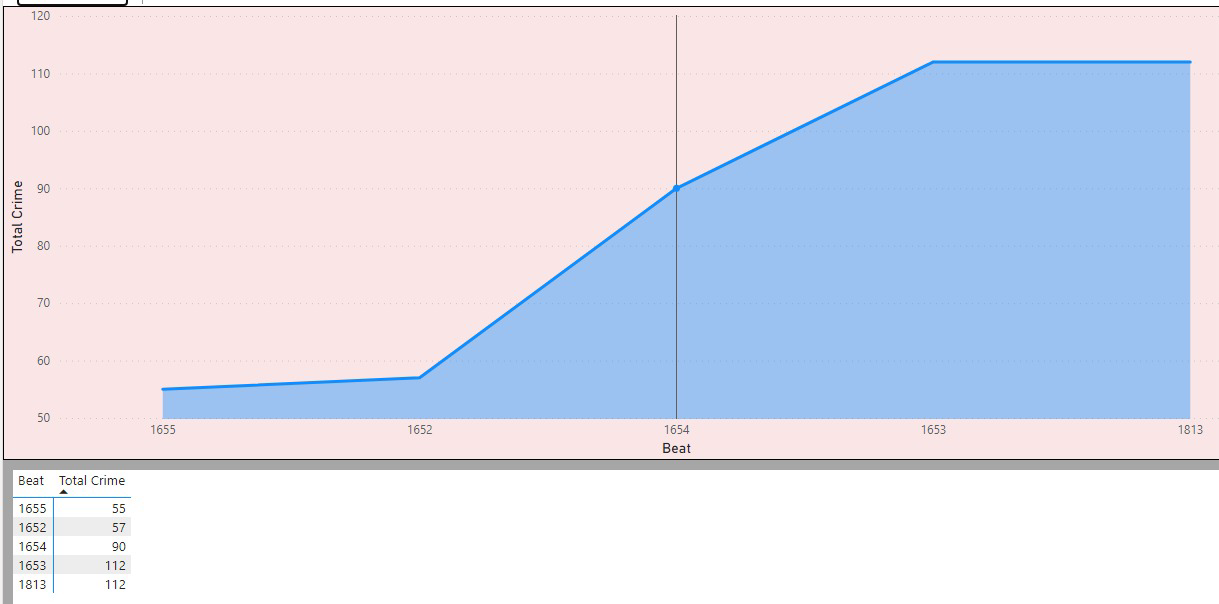
Location – Main Tab.

“Average days to resolve” Visual.

**10. To reward the patrol officers, find the patrol area where the crimes reported were under control.**

I employed an area chart to examine the correlation between beats and total reported crimes. Among these I suggested the following area, managed by patrol officers, experiences a relatively low incidence of criminal activities.

Beat 1655 is notable for its low count of reported crimes, standing at just 55 incidents. This suggests that this particular area, under the supervision of patrol officers, experiences relatively infrequent criminal activity.



Location – Locality Tab.

“Petrol area with low crime” Visual.

**11. Did you create any calculated columns in this project? What is the difference between the ‘calculated column’ and ‘add column’ functions?**

Yes, I did create calculated columns in this project. I utilized the Add Column from example feature in Power BI Query to create three additional columns:

Longitude and Latitude: I generated these columns by using the Column from Example feature, where I provided sample data separated by a comma as the delimiter. This allowed Power BI to inter the pattern and create the Longitude and Latitude columns accordingly.

The key difference between the calculated column and add column functions is in their application and functionality:

1. Calculated Column: Created in the Power BI data model using DAX formulas, calculated columns are computed and stored during data loading. They can be used for various calculations and operations within the dataset and are available for use in visualizations and analyses.
2. Add Column (Custom Column): Utilized in Power Query Editor, the Add Column function allows for the creation of custom columns during data transformation. These columns are dynamically generated based on transformations applied to existing columns and are not stored in the data model. They are primarily used for data preparation and enrichment tasks.

In summary, while both calculated columns and custom columns serve similar purposes, their creation methods and functionalities differ, with calculated columns being stored in the data.

**12. Using ‘Calculate’ and a row iteration DAX function calculate the number of crimes which are of type ‘theft’ and happened in ‘District 8’.**

Theft in District 8 is reported at 1143 incidents. This calculation is derived using the formula:

Theft crime in District 8 = CALCULATE([Total crime], AND(‘crimes data 2022’,[Type]=”theft”,’crimes data 2022’,[District]= “8”))

This formula specifically targets “Theft” type crimes within District 8, providing a precise count of 1143.



Location – Main Tab.

“Theft Crime in District 8” Visual.

**13. Using PowerBI can you separate the Longitude and Latitude from the Locations Column (Longitude, Latitude)? Which feature will you use?**

I employed the "Add Column with Example" functionality within the Power Query Editor to effectively separate values using a comma delimiter. This streamlined the process of data manipulation and organization, facilitating enhanced analysis and visualization capabilities.

**14. When we add a column in Power Query what’s the code that comes in M language in the formula bar? What do you know about M-query?**

When we add a column in Power Query using the Column feature, the M language code that appears in the formula bar depends on the specific transformations applied to the data.

Let's examine two examples:

1. Splitting Location Column: The M query for splitting the “Location” column into columns based on the comma delimiter would look like this:
   1. = Table.SplitColumn(#”Filtered Rows” , “Location”, Splitter.Split TextByDelimiter,”,”(QuoteStyle.None), {“Location.1”, “Location.2” })In this code:
   2. TableSplitColumn function is used to split the “Location” column into two separate columns, “Location.1” and “Location.2”, based on the comma delimiter.
   3. “#”Filtered Rows” refers to the previous step in the query where rows might have been filtered.
2. Extracting Last Characters for Moment of Register Case: The M query for adding a column that extracts the last two characters from the “Date” column to determine the moment of registering each case would look like this:
   1. = Table.AddColumn(Renamed Columns, Last Characters, each Text.End(Text.From([Date], “en-US”), 2), type text)

In this code:

Table.AddColumn function is used to add a new column named “Last”

“Characters”to the table each Text.End(Text.From([Date], “en-US”), 2) extracts the last two characters from the Date column for each row type text specifies the data type of the new column as text.

M-query, also known as the Power Query Formula Language, is a functional language used for data transformation in Power Query. It allows users to define step-by-step transformations on their data within Power Query Editor. M-query supports a wide range of operations such as data type conversion, splitting columns, merging tables, filtering rows, and more.

With M-query, users can perform complex data transformations and data cleaning tasks to prepare their data for analysis. It provides a powerful and flexible framework for manipulating data within Power Query, enabling users to create custom data transformation processes tailored to their specific requirements.

## **Subjective Questions**

**1. Is there any month-wise change in crime rates? If not, what could be the mistake in that operation?**

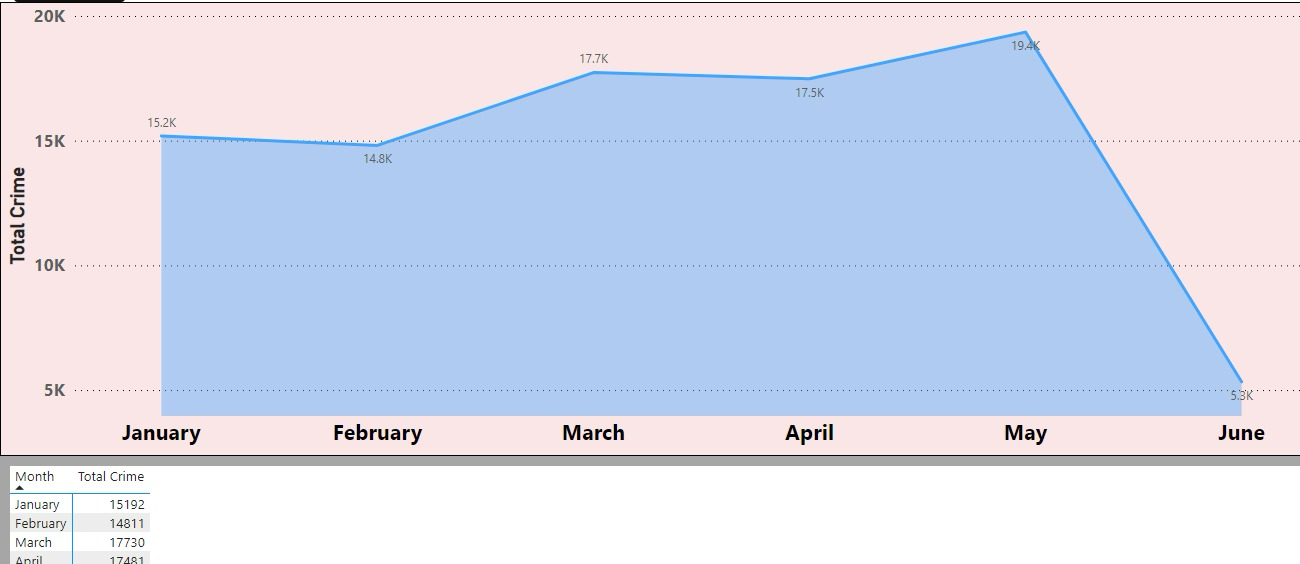
Upon scrutinizing the crime rates month by month from January to June 2022, a discernible pattern emerges. The data indicates a steady escalation in crime rates until May, succeeded by a notable decline in June. However, it's crucial to acknowledge that the sharp decrease in June might be attributed to the limited dataset, extending only until the 9th of June.

Specifically, in May, the recorded crime rate reached 19,353 incidents, marking it as a month with a higher-than-average crime rate during the observed period. This insight highlights the importance of delving deeper into the factors contributing to this surge in criminal activities during May.

Furthermore, the analysis points out that February witnessed a higher arrest rate of 13.83%.

While this might be an encouraging aspect in terms of law enforcement efficiency, it's crucial to understand the dynamics contributing to this peak in arrests during February.

Focusing on mitigating factors contributing to May’s surge can be a key strategy for reducing overall crime in Chicago.



Location – Main Tab.

“Total crime and Arrest rate by Month” Visual.

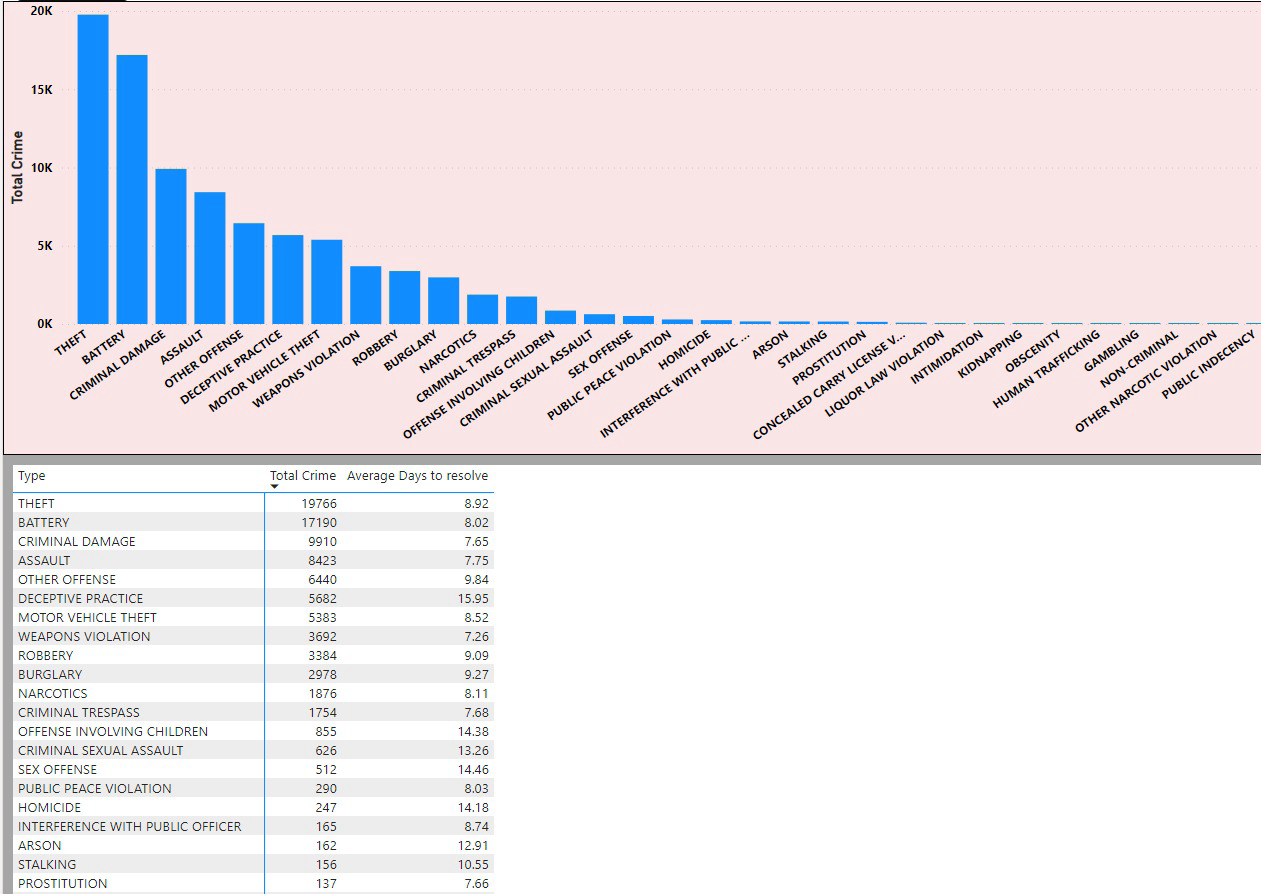
**2. How can we reduce the no. of crimes, and which types of crime should we focus on to achieve improvement in the overall number of crimes?**

Analysis of the crime data highlights theft and battery as the most predominant crime types, with theft constituting 19,776 incidents and battery accounting for 17,190. Upon closer inspection, a notable disparity in arrest rates between the two categories emerges—4.08% for theft and 13.92% for battery.

Additionally, domestic crime plays a significant role, constituting 5.40% of theft incidents and a notably higher 54.64% of battery incidents.

To effectively reduce the overall crime rate, a nuanced approach is required, acknowledging the complexity of factors contributing to criminal activities. Key observations point towards the following strategic areas:

1. Policing Initiatives:
   1. Strengthen and optimize policing initiatives, focusing on areas with high incidents of theft and battery.
   2. Implement community policing strategies to enhance trust and collaboration between law enforcement and the community.
2. Education and Awareness:
   1. Conduct public education campaigns to raise awareness about laws and their consequences.
   2. Emphasize the importance of community members reporting suspicious activities to aid law enforcement.
3. Domestic Crime Prevention:
   1. Devise targeted programs to address the underlying causes of domestic crimes, particularly in cases of battery.
   2. Collaborate with social services to provide support and intervention in situations prone to domestic violence.
4. Surveillance and Technology:
   1. Increase surveillance in areas with high theft rates, utilizing technology such as CCTV cameras and smart city initiatives.
   2. Implement advanced analytics for predictive policing to identify potential crime hotspots.



Location – Locality Tab

“Crime by Type” Visual.

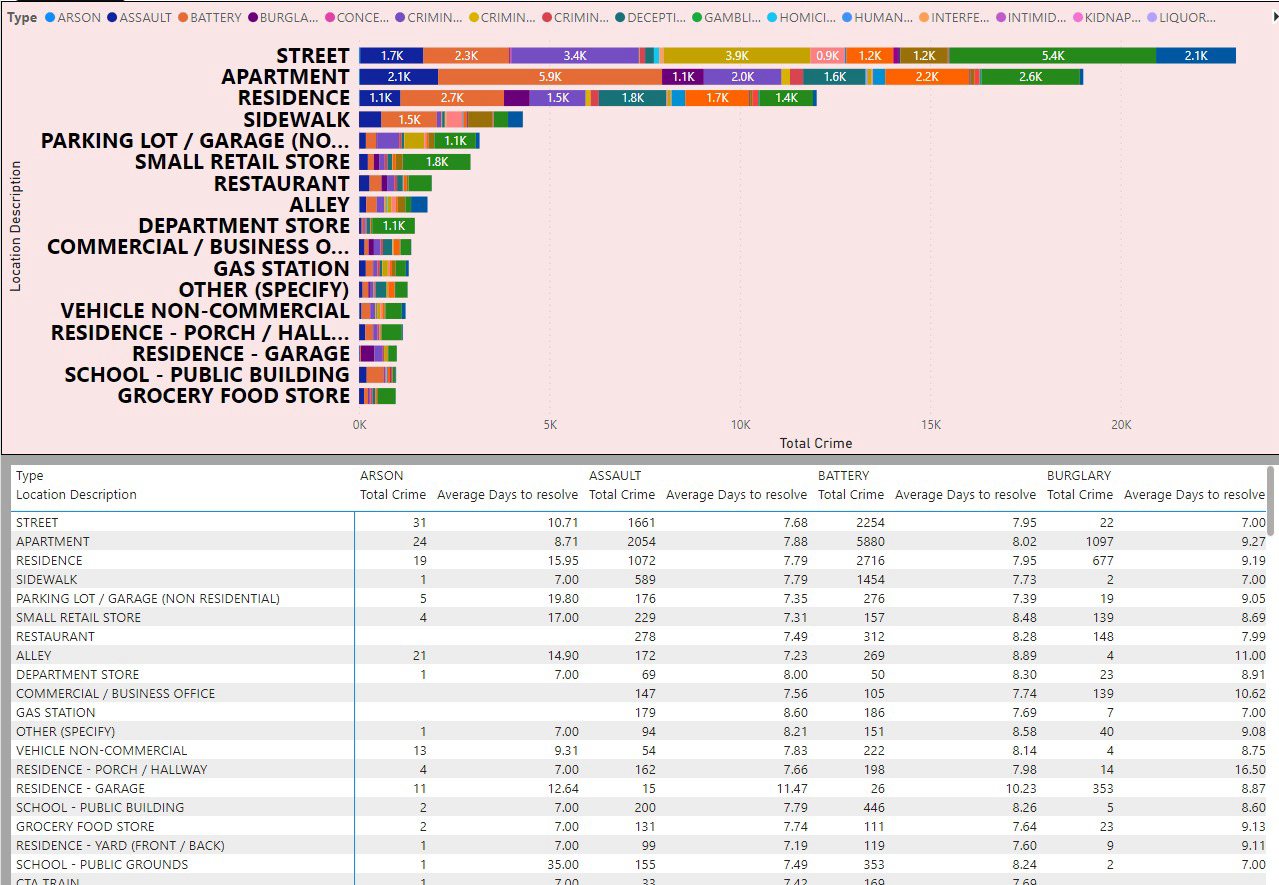
**3. Which localities experience higher crime rates, and what measures can we ensure to reduce these numbers?**

The examination of crime data underscores particular localities, specifically Streets, Apartments, and Residence locations, as encountering elevated crime rates, with incidents totaling 14,768, 13,910 and 8,345 respectively. Significantly, these areas manifest a prevalence of theft and battery crimes, coupled with arrest rates below the desired threshold. Furthermore, the analysis hints at a potential societal influence on arrest rates, particularly evident in domestic crime instances.

To address and reduce crime in these localities, several targeted measures can be implemented:

1. Enhanced Surveillance:
   1. Increase police presence and surveillance in high-crime areas, especially Street, Apartment and Residence Location.
   2. Utilize advanced surveillance technology, such as CCTV cameras, to monitor and deter criminal activities.
2. Community Engagement:
   1. Establish and encourage neighborhood watch programs to involve residents in crime prevention efforts.
   2. Foster a sense of community and cooperation to deter criminal behavior through collective vigilance.
3. Infrastructure Improvement:
   1. Improve street lighting to enhance visibility during nighttime, reducing opportunities for criminal activities.
   2. Focus on urban planning strategies that discourage crime, such as well-lit public spaces.
4. Address Root Causes:
   1. Invest in education and job training programs to provide economic opportunities for residents, addressing underlying causes of crime.
   2. Ensure accessible mental health and substance abuse services to support individuals in need.
5. Youth Engagement Programs:
   1. Implement proactive youth engagement programs to steer young individuals away from criminal activities.
   2. Advocate for and implement effective gun control measures to prevent unauthorized access to firearms.

By adopting a comprehensive, community-oriented approach that combines law enforcement strategies with social and economic interventions, these localities can experience a significant reduction in crime rates. It is crucial to address both the immediate factors contributing to crime and the underlying root causes to create lasting positive change in these areas.

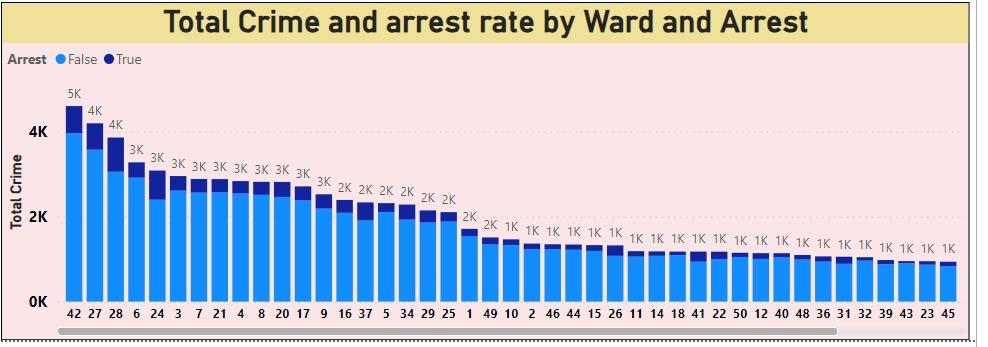


Location – Locality Tab.

“Total Crime and Average Days to Resolve by Location Description and Type” Visual.

**4. Can you suggest wards where security improvements should be made to reduce crime?**

The examination of the stacked column chart reveals that Wards 42, 27, 28, 6, and 24 demonstrate elevated crime rates, each surpassing 3000 incidents. Furthermore, Ward 7 stands out due to a notable percentage of domestic crime. To bolster security and mitigate crime in these wards, several measures can be implemented:



1. Increased Police Presence:
   1. Deploy additional law enforcement officers and resources to actively patrol and deter criminal activities in the identified high-crime wards.
2. Community Policing:
   1. Foster strong community-police partnerships in the highlighted wards to build trust, improve communication, and encourage residents to report suspicious activities.
3. CCTV Surveillance:
   1. Install surveillance cameras strategically to monitor and record activities, acting as a deterrent and aiding in investigations.
4. Lighting Upgrades:
   1. Improve street lighting in crime-prone areas to enhance visibility and reduce opportunities for criminal behavior.
5. Youth Engagement Programs:
   1. Invest in youth programs and educational initiatives to provide positive alternatives for young residents and reduce youth involvement in crime.
6. Cybersecurity:
   1. Address the modern threat landscape by investing in cybersecurity infrastructure, raising awareness about online safety, and providing resources for cybercrime prevention.
7. By implementing this comprehensive set of measures, a more effective and nuanced approach can be taken to reduce crime rates in the identified wards, fostering a safer and more secure environment for the residents.

Location – Exploration Tab.

“Total Crime and Arrest Rate by Ward and Arrest ” Visual.

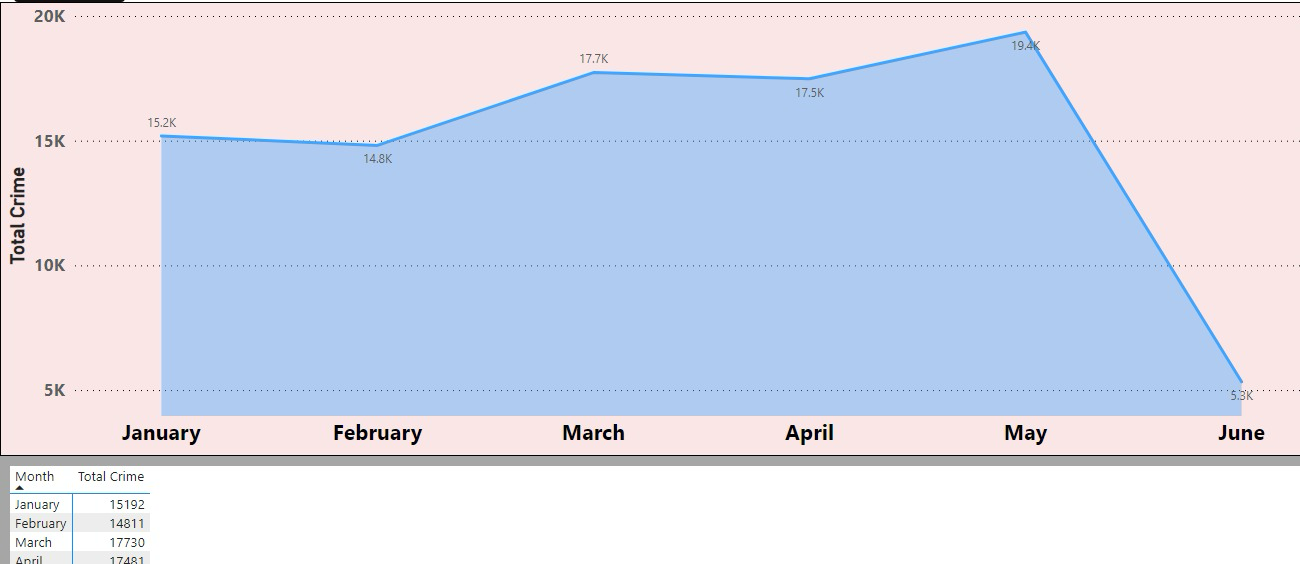
**5. Crime Rate Trend Analysis: Monitor changes in crime rates over time to detect any discernible patterns or trends.**

The analysis of crime data for the year 2022 reveals variations in crime rates across different months, with May showing a notable peak.

While the insights derived from this snapshot are valuable, it's imperative to acknowledge the limitations inherent in analyzing data from a single year. To accurately detect discernible patterns or trends, it is advisable to gather and analyze data spanning multiple years. Longitudinal data would facilitate a more comprehensive examination of fluctuations, seasonal variations, and long-term trends in crime rates, thereby providing a more nuanced understanding of the dynamics at play.

From a location perspective, we observe that crime is evenly distributed across various areas, making it challenging to discern any specific trends or patterns. However, a significant disparity emerges when analyzing crime occurrences during different times of the day. Crime rates notably increase during nighttime hours, indicating a need for more proactive measures during this period.

Addressing nighttime crime requires proactive efforts from authorities, such as increased patrolling, enhanced lighting in vulnerable areas, and community engagement initiatives tailored to nocturnal activities. By focusing resources and attention on nighttime crime prevention strategies, authorities can effectively reduce overall crime rates in the city and enhance public safety.



Location – Main Tab.

“Total Crime and Arrest Rate by Month ” Visual.

**6.Create a monthwise tabular data consisting of two columns, month and total no. of crimes in that month. Also, add one more column where each row of the column contains the total no. of crimes for the previous month. Do we need to use any filter-based DAX function here (All, All except, etc)?**

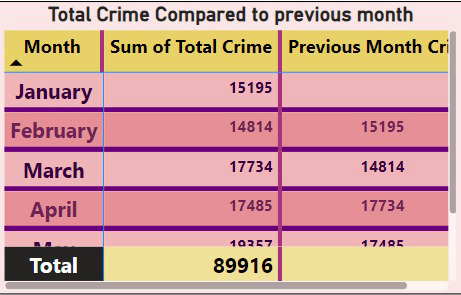
In my Power BI project, I faced a scenario where I had to analyze crime data on a month-by-month basis. To tackle this, I initially crafted a continuous date table to ensure the effective utilization of DAX functions. However, upon encountering the necessity to calculate the total number of crimes for the preceding month, I recognized the imperative for adopting an alternative approach.

I subsequently created a new table named "Dates\_crime," which contains all distinct dates associated with the crimes that occurred on those specific dates. Then, I established a one-to-one relationship between this new table and my main date table. This relationship allowed me to accurately calculate the total number of crimes for the previous month using the PREVIOUSMONTH function in DAX.

To visualize the relationship between the current and previous months' crime data, I deployed a table in Power BI. This table enabled me to compare the total number of crimes for each month with those of the preceding month. Through this analysis, I identified significant differences in crime rates between certain months.

Overall, by structuring my data model in this manner and utilizing DAX functions effectively, I was able to gain valuable insights into month-to-month variations in crime occurrences, which are essential for making informed decisions and implementing targeted interventions.

Dax Formula -> previous month = CALCULATE(SUM(Dates\_crime[Total Crime]),PREVIOUSMONTH('Dates Table'[Date]))



Location – Main Tab.

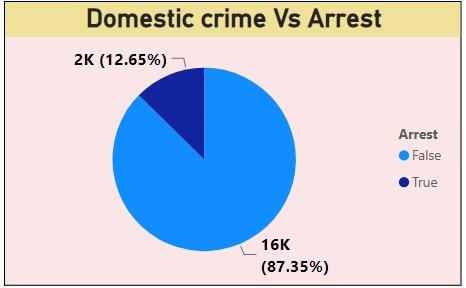
“Total Crime Compared to Previous Month” Visual.

**7. As per the previous reports, most domestic crimes do not result in arrest due to public hesitation and family pressure, is this trend also visible in our data?**

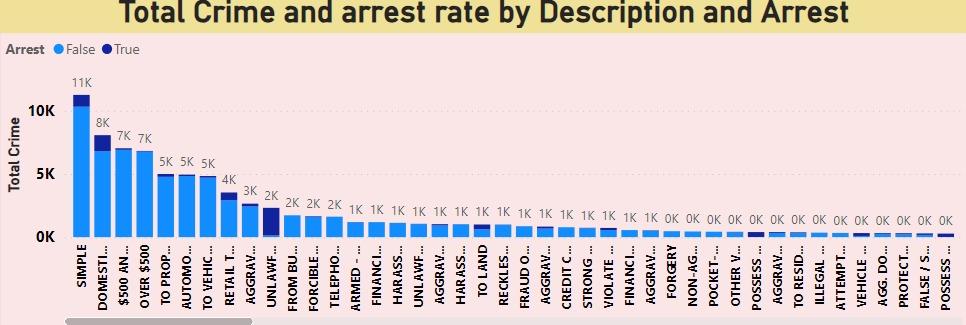
The analysis of the data reveals a concerning trend regarding domestic crimes and arrest rates. According to the pie chart visualization, only 12.65% of domestic crimes result in arrest. This statistic aligns with the observation that in previous reports, many domestic crimes do not lead to arrests due to public hesitation and family pressure.

The low arrest rate for domestic crimes indicates a significant gap in addressing this particular category of offenses. Public reluctance to report domestic incidents and familial pressure may contribute to underreporting and the subsequent lack of arrests.

To address this trend effectively, it is imperative to implement measures to encourage reporting of domestic crimes, provide support for victims, and ensure that law enforcement agencies are adequately trained and equipped to handle such cases sensitively and effectively. Additionally, raising awareness about the importance of reporting domestic incidents and breaking the cycle of silence within families is crucial for improving arrest rates and ensuring justice for victims of domestic violence and abuse.



**8. Could you generate a visual representation that emphasizes the frequently occurring terms within the Description column?**



Analysis of the description column uncovers several frequently occurring terms, notably 'Simple' with 11,244 instances, 'Domestic Battery Simple' with 8,060 occurrences, '$500 And Under' with 7,024 instances, and 'Over $500' with 6,813 occurrences, among others. This visualization effectively showcases these common terms, providing insight into prevalent types of incidents.

By referencing this visualization, stakeholders can gain detailed insights into the distribution of frequently occurring terms. It’s important to note that the visualization includes a filter, allowing users to focus on specific categories. Removing the filter expands the view, providing a comprehensive overview of term frequencies across all categories.

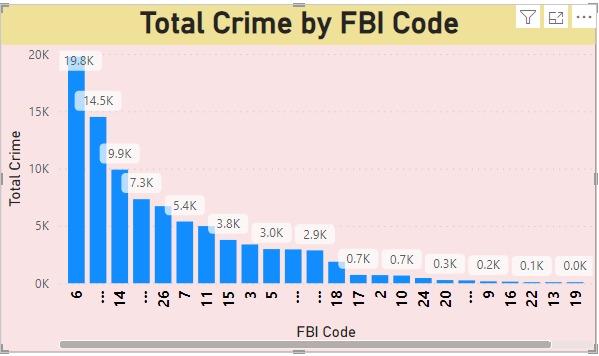
**9. Are there any particular regions as per the data where the number of domestic crimes reported is very high?**

The data reveals that there are no specific regions or geographic areas explicitly designated as domestic within the dataset. Domestic crimes are typically categorized based on the nature of the offense (e.g., domestic violence, domestic disturbance) rather than specific geographic regions. As a result, it's not feasible to identify particular regions solely based on the dataset where the number of reported domestic crimes is exceptionally high.

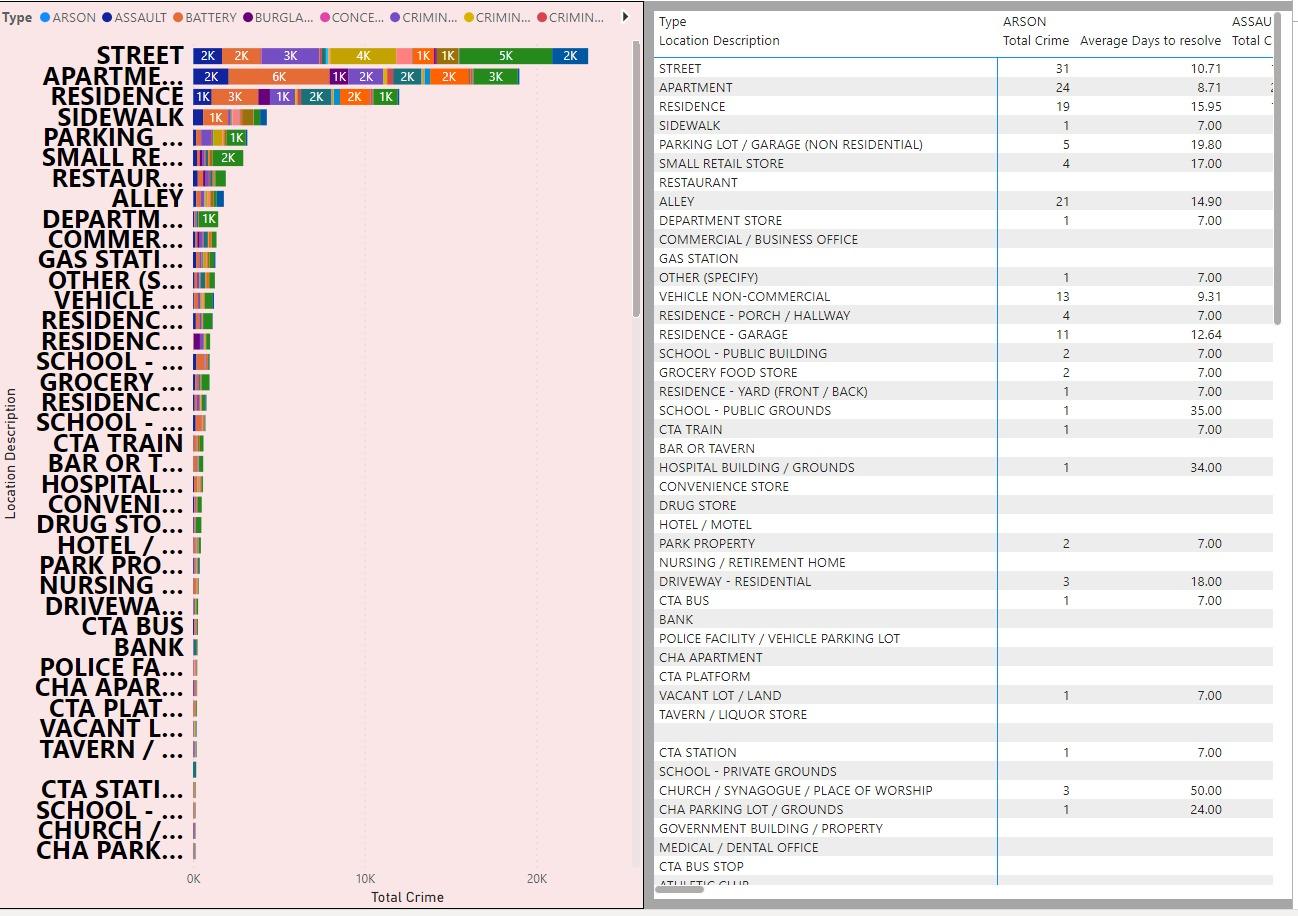
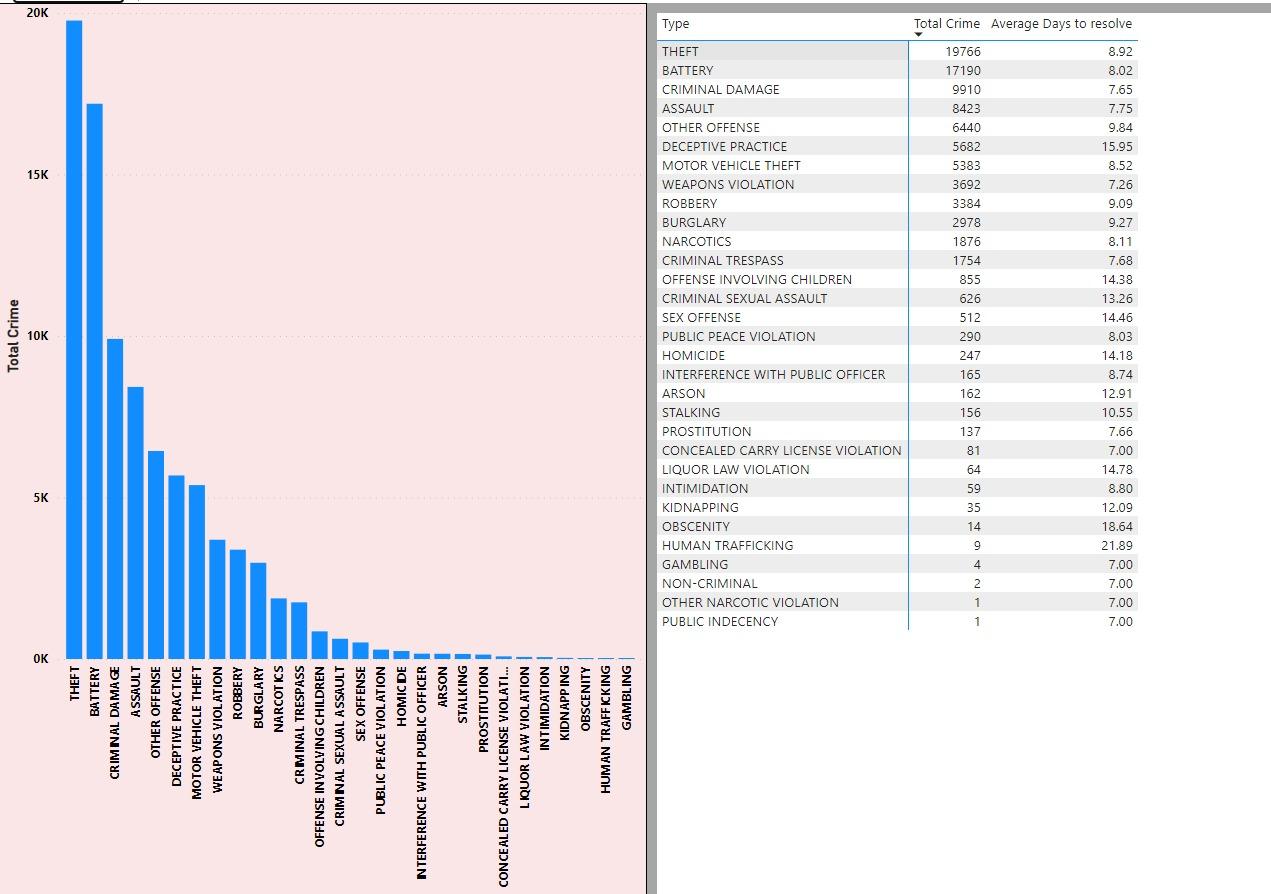
To pinpoint regions with high domestic crime rates, it would require a more detailed analysis that considers the location descriptions, addresses, or community areas where domestic incidents occurred. Such an analysis would involve filtering and categorizing incidents related to domestic crimes and then assessing their prevalence in specific areas.

It's crucial to recognize that domestic crimes can transpire across diverse neighborhoods and communities. Effectively addressing this issue frequently entails collaboration between law enforcement, social services, and community organizations. This collaboration aims to offer support and intervention for victims while holding offenders accountable for their actions.

**10. Is the solving time of cases also dependent upon the type and locality of crime?**



The provided visual representation illustrates the relationship between crime type, locality, FBI codes, and the average time of case resolution. Crimes related to license violation and gambling exhibit shorter resolution times, whereas cases involving human trafficking and obscenity tend to have longer resolution times. Similarly, cases occurring at airport terminals and on boats or watercraft are resolved relatively quickly, whereas cases associated with pawn shops and colleges experience longer resolution times.



**11. Create a calculated column to flag the domestic crimes that took place in District 8.**

I've successfully created a calculated column using the following formula:

Domestic Crime in District 8 = IF( AND(‘crimes data 2022’,[Domestic],’crimes data 2022’,[District] = 8), “Yes”, “No” )

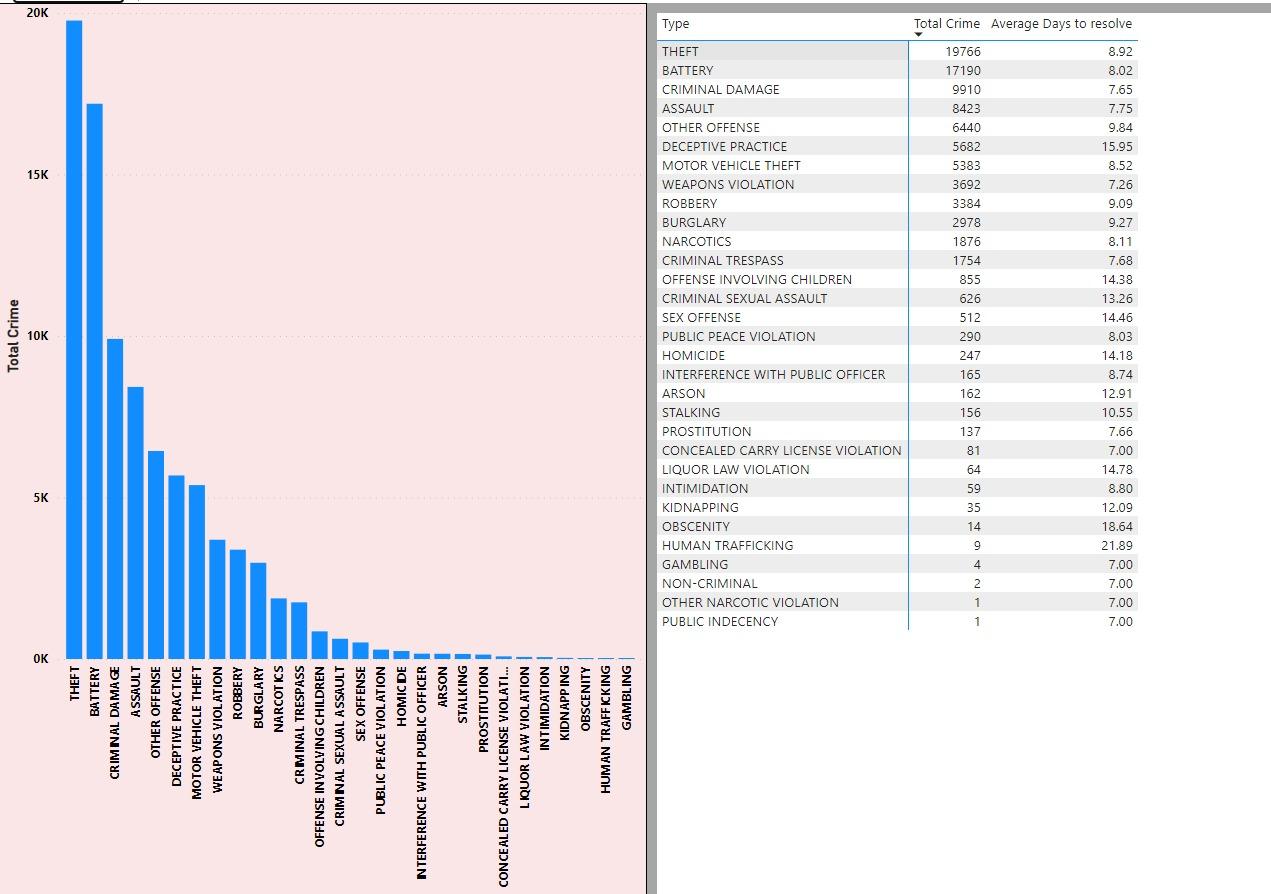
This formula accurately flags domestic crimes occurring in District 8 within our dataset. By employing the IF function in conjunction with AND, it efficiently evaluates if a crime is domestic (crimes data 2022,[Domestic]) and occurred in District 8 (crimes data 2022,[District] = 8),

assigning a value of Yes, if both conditions are met, and No otherwise. This calculated column provides a valuable categorization for further analysis, facilitating the identification and examination of domestic incidents specifically within District 8.

**12. Out of all the types of crimes which do you think is the most dangerous one and rank the type of crimes according to their no. of occurrences?**

After scrutinizing the dataset, it's apparent that theft, battery, and criminal damage are among the crime types with the highest occurrence rates. Despite their prevalence, these crimes also demonstrate notably low arrest rates.

Based on the frequency of occurrences and the potential impact on individuals and communities, theft emerges as particularly concerning. Not only does it affect victims financially, but it also contributes to a sense of insecurity and mistrust within society.



Ranking the types of crimes according to their number of occurrences, theft would likely hold the top position, followed by battery and criminal damage. This ranking reflects the gravity of these crimes in terms of their prevalence and societal impact. Furthermore, the observed low arrest rates for these crimes underscore the challenges faced in effectively addressing and combating them.

**13. What do you understand about the PowerBI gateway? What are its use cases?**

Power BI Gateway is a tool provided by Microsoft that serves as a bridge between on-premises data sources and the cloud-based Power BI service. It enables Power BI to access and refresh data from on-premises sources securely. Here’s a breakdown of its key aspects and use cases:

**Data Connectivity:** Power BI Gateway allows you to connect to various on-premises data sources such as SQL Server, Oracle, SharePoint, and more. It establishes a secure connection between these data sources and the Power BI service in the cloud.

**Data Refresh:** One of the primary use cases of Power BI Gateway is to enable scheduled data refresh for on-premises data sources. It allows Power BI reports and dashboards to stay up-to-date by automatically refreshing data at regular intervals.

**Direct Query:** Power BI Gateway supports Direct Query mode, where queries from Power BI reports are sent directly to the on-premises data source in real-time. This enables users to analyze the most current data without storing it in the cloud.

**Personal and Enterprise Gateway:** Power BI Gateway comes in two editions: Personal Gateway, designed for individual users or small teams, and Enterprise Gateway, suitable for organizations with multiple users and complex data sources.

**Use Cases:**

**Hybrid Deployments:** Power BI Gateway facilitates hybrid deployments where organizations have a mix of cloud-based and on-premises data sources. It allows seamless integration of data from both environments.

**Scheduled Data Refresh:** Power BI Gateway enables scheduled data refresh for reports and dashboards connected to on-premises data sources. This ensures that insights are based on the most recent data.

**Real-time Analytics:** With Direct Query mode, Power BI Gateway supports real-time analytics by connecting directly to on-premises data sources without data replication.

**Data Security:** By keeping sensitive data on-premises, Power BI Gateway helps organizations maintain data security and compliance with regulatory requirements.

**Large-scale Deployments:** Enterprise Gateway is suitable for large-scale deployments in organizations with multiple users and data sources. It provides centralized management and monitoring capabilities for IT administrators.

Overall, Power BI Gateway plays a critical role in enabling organizations to leverage the power of Power BI while securely accessing and refreshing data from on-premises sources. It enhances data connectivity, enables real-time analytics, and ensures data security and

compliance.

**14. How would you approach this problem, if the objective and subjective questions weren’t given?**

My approach on this project :

1. Aggregate Criminal Incidents and Arrest Rate: Start by calculating the total number of crimes and the arrest percentage. This provides an overview of the crime situation and the effectiveness of law enforcement in addressing these crimes.
2. Crime Location BreakdownUtilize a map visualization to understand the distribution of crimes across different locations. This helps identify areas with high crime rates and may inform resource allocation and law enforcement strategies.
3. Percentage of Crime Occuring during Day and Night : Analyze the percentage of crimes occurring during the day and night. This information is valuable for planning patrols and allocating resources based on the time of day when crimes are most likely to occur.
4. Crime Type and Location Examination : Explore the data further by analyzing the distribution of crime types across different locations. This can help identify patterns and hotspots for specific types of crimes, guiding targeted interventions and prevention efforts.
5. Determining Resolution Time: Calculate the resolution time for each case by subtracting the date of occurrence from the date updated. This metric provides insight into the efficiency of law enforcement in resolving cases and may highlight areas for improvement.
6. Domestic Crime Evaluation : Investigate the proportion of domestic crimes and the factors influencing their resolution. Understanding the dynamics of domestic crimes, including family pressure, is crucial for developing effective intervention and support mechanisms.

**15. If you are also given a table of districts-states with state\_id, district\_id and name, what would be the type of relationship between district of our data and district\_id of the new table?**

If the district\_id is incorporated into the new table, the relationship between the current table and the new table would be many-to-one. This is because in my dataset, each district\_id is repeated multiple times as each district has numerous associated cases. Consequently, each district in my data can only relate to one district\_id in the new table, establishing a many-to-one relationship.