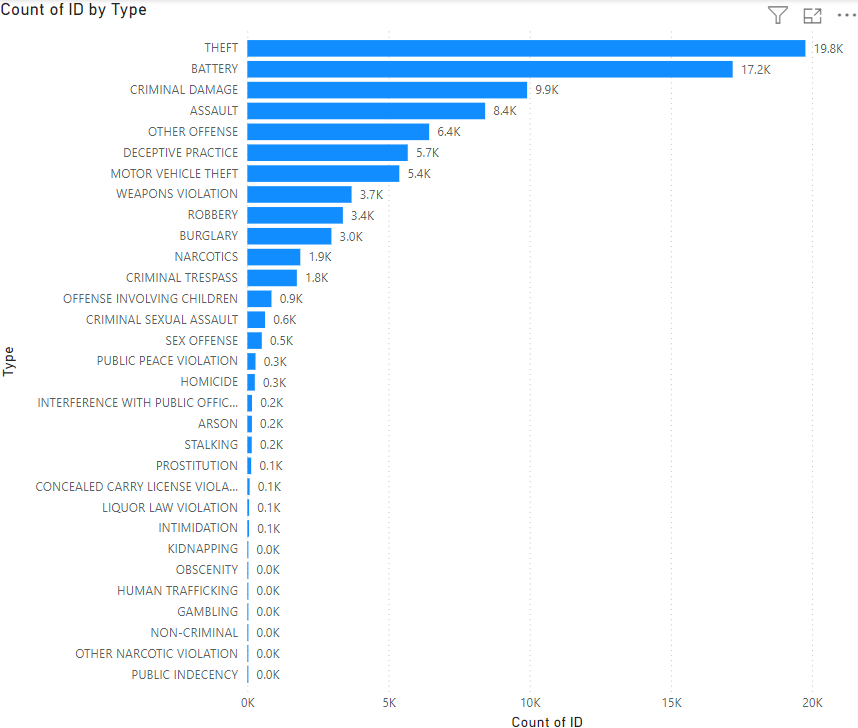
**Objective Questions**

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

* Loaded the data in to power query editor
* Removed the unwanted data and missing data rows from the dataset   
  • M column : Ward - 4 rows  
  • P column : X Coordinate - 943 rows   
  • Q column : Y Coordinate - 943 rows   
  • T column : Location - 943 rows
* Above mentioned inconsistencies and missing values are cleaned before analysis
* Changed the datatype of date and updated on column to date datatype

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

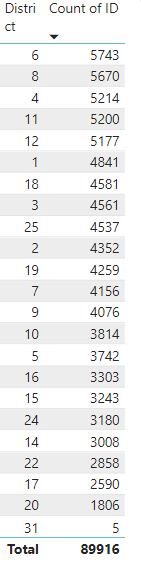
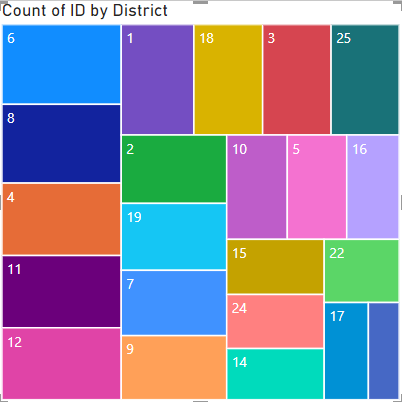
* By using Table and bar chart in power BI
* we can find the frequency of crime  
    
    
  

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

* Using measure we can find the arrest percentage
* **Arrest Percentage = CALCULATE(COUNTA(crimes\_data\_2022[Arrest]),crimes\_data\_2022[Arrest] = True)/COUNTA(crimes\_data\_2022[Arrest]) \* 100**



1. 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.

* Using Table and Tree map we can analyse the District wise crime
* District 6,8,4,11,12 are the high – crime areas  
    
    
  

1. How many categorical attributes are there in the data?  
     
   There are 13 categorical attributes in the dataset:

* Block
* IUCR
* Type
* Description
* Location Description
* Arrest
* Domestic
* Beat
* District
* Ward
* Community Area
* FBI Code
* Year

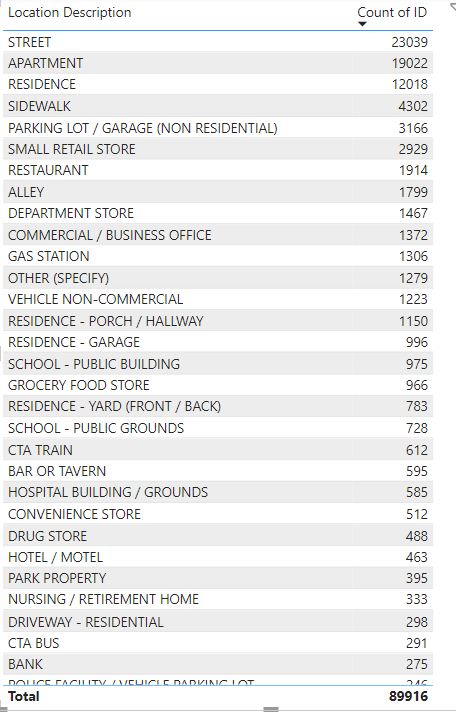
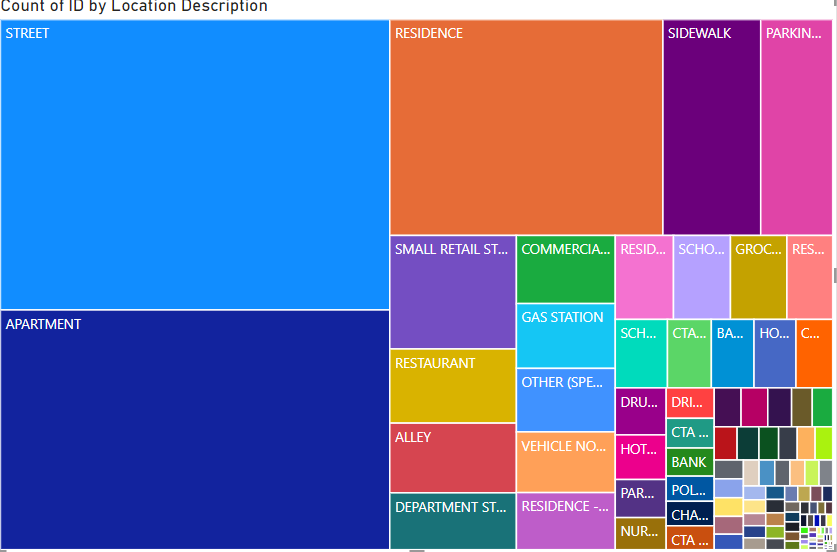
1. Were there any Null values in the data, if there were how did you handle them? What is the ideal way to handle Null values?

* I noticed some null values in the data while working in Power Query Editor.
* Initially, I filtered them out by unselecting null values, but the best approach would be to clean the data at the source level by removing these null values directly from the source file itself.
* This way, we ensure cleaner and more accurate data from the start.

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

* Using below measure formula we can calculate ratio
* **Domestic crime ratio = CALCULATE(COUNTA(crimes\_data\_2022[Domestic]),crimes\_data\_2022[Domestic] = True) / COUNTA(crimes\_data\_2022[ID])\*100**  
    
  
* **so the Ratio between domestic and other crime are – 1:5**

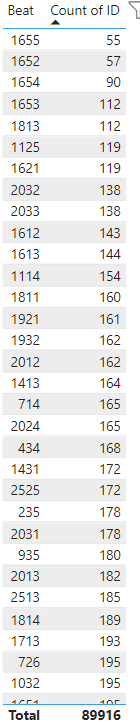
1. Is there any “Location Description” where the number of crimes is higher than expected? Come up with a table or visualization in which one can judge the frequency of crimes at each Location Description type.

* Using Table and Tree map we can analyse the frequency of crime location wise   
    
    
  

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

* Using below two measure formula’s we can find the average time
* **Time Taken = ABS(DATEDIFF(crimes\_data\_2022[Updated On],crimes\_data\_2022[Date],DAY))**
* **AVG TIME = ABS(AVERAGE(crimes\_data\_2022[Time Taken]))**  
  

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

* Using Table and sorting by ascending order we can find the
* Top patrol area’s with less crime reported  
    
  

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

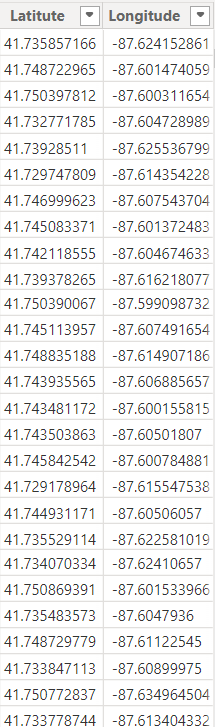
* + **Yes , Created few calculated column in this project**
  + **In Power BI, both calculated columns and custom columns (Add Columns) are used to create new columns in your data model, but they are created in different ways and serve slightly different purposes.**
  + **Calculated Columns are created in the Data or Model view using DAX (Data Analysis Expressions).**
  + **Custom Columns** are created in Power Query Editor using the **M language (Power Query Formula Language)**.
  + **Calculated Columns** are computed **after** the data is loaded into the data model. They are evaluated when the data model is refreshed.
  + **Custom Columns** are computed **during** the data load process (ETL) in Power Query, before the data enters the data model.

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

* Using Below measue formula we can find
* **Total\_Theft\_District8 = CALCULATE(COUNTROWS(crimes\_data\_2022),crimes\_data\_2022[Type] = "THEFT" , crimes\_data\_2022[District]= 8)**



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

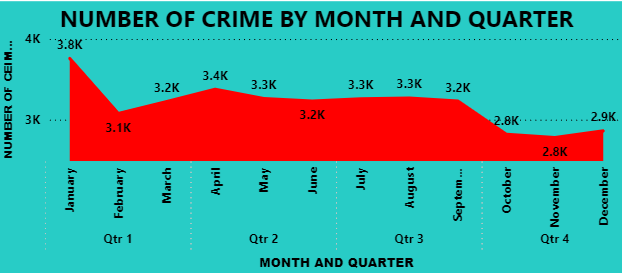
* In Tranfrom data – move to power query editor 🡪 split column option we can split the column by using delimiter  
  Here (,) is the delimiter used to separate and column has to be renamed
* using replace value – “(“ , “)” brackets will be removed and latitude and longitude data will be extracted from location column   
    
  

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

* When you add a column in Power Query, the code that appears in the M language formula bar typically looks like this:
* **Table.AddColumn(#"PreviousStepName", "NewColumnName", each [ExistingColumn] + 1)**
* M-query, or Power Query Formula Language, is a functional language used in Power BI .
* Power Query Editor for data transformation.
* It allows you to clean, reshape, and manipulate data through a series of steps.
* M-query is case-sensitive, relies on functions, and supports operations like adding/removing columns, filtering rows, and more.

**Subjective Question**

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

* Using line chart we can see the month and quarter wise crime rate   
    
  
* Jan Month the crime rate is very high comparing to other months
* Nov Month the crime rate is less comparing to other months

1. 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?  
   * **THEFT (19766 incidents):** Theft has the highest occurrence. Prioritizing measures to reduce theft could have a significant impact on the overall crime rate.

Increase community awareness and vigilance.  
Improve surveillance in high-risk areas.  
Implement more robust security measures in retail and residential areas.

* + **BATTERY (17,190 incidents):** Battery is the second most common crime, and addressing it could significantly lower crime rates.  
      
    Increase police presence in areas with high rates of battery.  
    Implement conflict resolution programs.  
    Enhance support services for victims of domestic violence.
* **CRIMINAL DAMAGE (9,910 incidents):** Criminal damage, while less frequent than theft and battery, can have a large impact on the community's sense of safety and economic well-being.

Improve lighting and surveillance in public spaces.  
Engage in community beautification projects to foster a sense of ownership and responsibility.

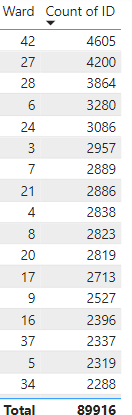
* **ASSAULT (8,423 incidents):** Assault, like battery, directly affects individuals and can contribute to a perception of lawlessness if not adequately addressed.

Increase law enforcement in known hotspots.  
Offer anger management programs and counselling services.  
Strengthen legal consequences for repeat offenders.

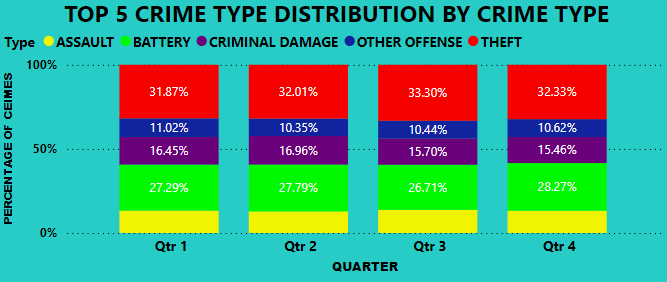
1. Which localities experience higher crime rates, and what measures can we ensure to reduce these numbers?

* Street: 23,039 incidents
* Apartment: 19,022 incidents
* Residence: 12,088 incidents
* Street-Level Crime Reduction:
* Increased Surveillance: Install more CCTV cameras, especially in areas with high foot traffic and known crime hotspots. Ensure these cameras are monitored in real-time.
* Improved Lighting: Enhance street lighting in poorly lit areas, as well-lit environments can deter criminal activity.
* Police Patrols: Increase police presence and patrols in areas with high crime rates. This could include foot patrols, bike patrols, and community policing initiatives.
* Apartment Complexes:
* Security Systems: Encourage or mandate the installation of secure entry systems, such as key fobs, video intercoms, and secure mailrooms.
* Resident Awareness: Conduct safety workshops or distribute materials to residents on how to safeguard their homes (e.g., locking doors, using window bars, reporting suspicious behavior).
* Apartment Patrols: Implement regular security patrols in larger apartment complexes, especially during night hours.
* Residential Areas:
* Neighborhood Watch Programs: Establish or strengthen neighborhood watch groups to encourage residents to look out for each other’s properties.
* Home Security Systems: Promote the use of home security systems, such as alarms, security cameras, and smart locks. Some municipalities may offer subsidies or tax breaks for installing such systems.
* Public Education: Run campaigns to educate homeowners about common home safety practices, such as keeping doors and windows locked, not advertising when they’ll be away on social media, and keeping the area around the home well-lit.

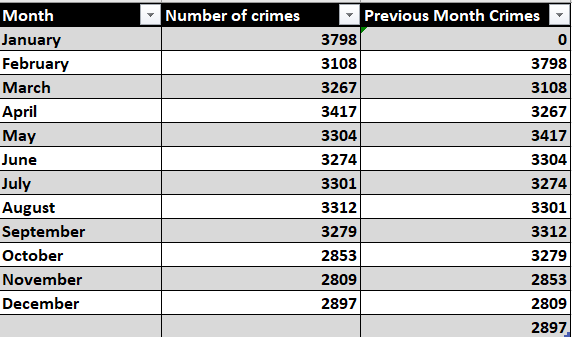
1. Can you suggest wards where security improvements should be made to reduce crime?

* 42 , 27,28 , 6,24 These are wards where security improvements should be made to reduce crime rate   
    
    
   **Suggestion to reduce the crime rate**
  + - Install more CCTV cameras and improve real-time monitoring.
    - Enhance street lighting in poorly lit areas.
    - Increase police patrols, including foot and bike patrols.
    - Install secure entry systems like key fobs and video intercoms.
    - Implement regular security patrols, especially at night.

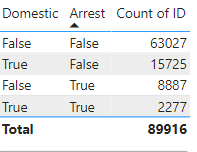
1. Crime Rate Trend Analysis: Monitor changes in crime rates over time to detect any discernible patterns or trends.

* Using Stacked column chart
* Giving type in legend we can analysis the crime rate by over time  
    
  

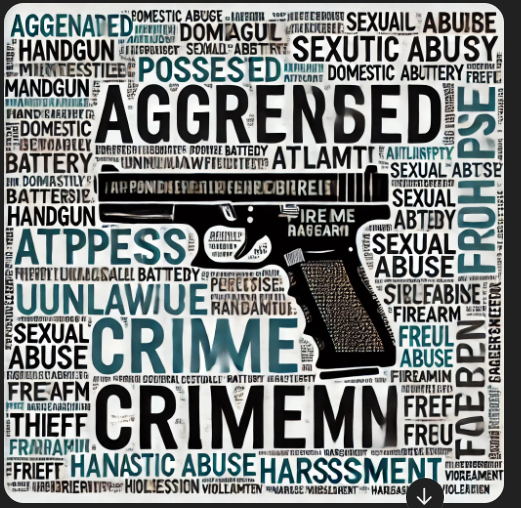
1. Create a monthwise tabular data consisting of two columns, month and total no. of crimes in that month. Also, add one more column where each row of the column contains the total no. of crimes for the previous month. Do we need to use any filter-based DAX function here (All, All except, etc)?

* Using tables we can see the month wise crime
* And adding one more column we can also find the previous month crimes  
    
  

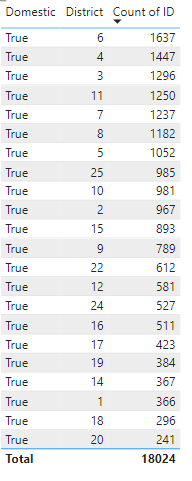
1. As per the previous reports, most domestic crimes do not result in arrest due to public hesitation and family pressure, is this trend also visible in our data?

* Yes in our dataset also most domestic crimes do not result in arrest due to public hesitation and family pressure  
    
  
* Totally 18,002 cases are domestic in that 2277 cases are only arrested   
  Which Means 12.64 % of cases only arrested in Domestic cases

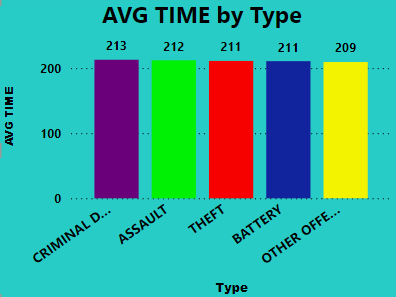
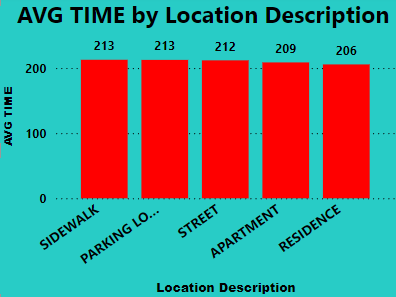
1. Could you generate a visual representation that emphasizes the frequently occurring terms within the "Description" column?

* Totally there are 259 Different Descriptions in the dataset  
    
  
* Here's the visual representation you requested, highlighting the frequently occurring terms within the "Description" column. The prominent terms like "AGGRAVATED," "POSSESS," and "FIREARM" are emphasized in the word cloud.

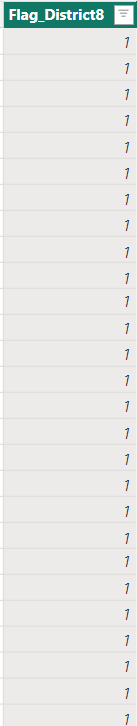
1. Are there any particular regions as per the data where the number of domestic crimes reported is very high?

* Region wise Domestic Crime
* District 6,4,3,11,7 these places the domestic crimes are very high  
  

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

* Yes the solving time of cases is also depend on the type of crime and the location where it is happening
* AVG TIME by TYPE  
    
  
* AVG TIME by Location  
    
  

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

* Using below measure formula we can find
* **Flag\_District8 = if(crimes\_data\_2022[District] = 8 && crimes\_data\_2022[Domestic] = True,1,0)**  
    
  

1. 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?  
     
   Ranked List of Crimes by Occurrences:
2. **THEFT** - 20,348 occurrences
3. **BATTERY** - 17,211 occurrences
4. **CRIMINAL DAMAGE** - 9,988 occurrences
5. **ASSAULT** - 8,443 occurrences
6. **OTHER OFFENSE** - 6,466 occurrences
7. **DECEPTIVE PRACTICE** - 5,848 occurrences
8. **MOTOR VEHICLE THEFT** - 5,390 occurrences
9. **WEAPONS VIOLATION** - 3,694 occurrences
10. **ROBBERY** - 3,390 occurrences
11. **BURGLARY** - 2,985 occurrences
12. **NARCOTICS** - 1,877 occurrences
13. **CRIMINAL TRESPASS** - 1,764 occurrences
14. **OFFENSE INVOLVING CHILDREN** - 864 occurrences
15. **CRIMINAL SEXUAL ASSAULT** - 632 occurrences
16. **SEX OFFENSE** - 514 occurrences
17. **PUBLIC PEACE VIOLATION** - 291 occurrences
18. **HOMICIDE** - 265 occurrences
19. **INTERFERENCE WITH PUBLIC OFFICER** - 165 occurrences
20. **ARSON** - 162 occurrences
21. **STALKING** - 157 occurrences
22. **PROSTITUTION** - 138 occurrences
23. **CONCEALED CARRY LICENSE VIOLATION** - 82 occurrences
24. **LIQUOR LAW VIOLATION** - 65 occurrences
25. **INTIMIDATION** - 59 occurrences
26. **KIDNAPPING** - 35 occurrences
27. **OBSCENITY** - 14 occurrences
28. **HUMAN TRAFFICKING** - 10 occurrences
29. **GAMBLING** - 4 occurrences
30. **NON-CRIMINAL** - 2 occurrences
31. **PUBLIC INDECENCY** - 1 occurrence
32. **OTHER NARCOTIC VIOLATION** - 1 occurrence
33. What do you understand by Power BI gateway? What are its use cases?

* A Power BI Gateway is a bridge that facilitates secure data transfer between on-premises data sources and Power BI service, Power Automate, Azure Analysis Services, and Power Apps.
* It allows organizations to keep their data on-premises while leveraging cloud-based services for analytics, automation, and reporting.
* Types of Power BI Gateways:
* On-premises data gateway (Standard Mode):

Used for connecting multiple users to multiple on-premises data sources.

Supports a wide range of data sources like SQL Server, Oracle, SharePoint, etc.

Allows for scheduled data refreshes and real-time data updates in Power BI reports and dashboards.

Can be used by Power BI, Power Apps, Power Automate, Azure Logic Apps, and Azure Analysis Services.

* On-premises data gateway (Personal Mode):

Used by individual users who only need to connect to on-premises data sources for their own reports and dashboards.

Supports Power BI only.

Ideal for single-user scenarios, where data is refreshed based on the user’s credentials

* Key Use Cases for Power BI Gateway:

Connecting On-Premises Data to the Cloud

Scheduled Data Refresh

Real-Time Data Analytics

Hybrid Scenarios

Security and Compliance

Multi-User Scenarios

Power BI Dataflows

1. How would you approach this problem, if the objective and subjective questions weren't given?

* Clarify the goal to assess the impact of the government’s operation on reducing crime rates in Chicago, and identify key metrics like crime rate trends, case-solving efficiency, and high-crime areas.
  + Gather and clean data on crime incidents before and after the operation, including crime locations, types, and case-solving times. Collect external factors such as population density and economic conditions.
  + Analyze crime rate trends, evaluate changes in case-solving times, and use geospatial techniques to identify high-crime areas.
  + Conduct statistical tests to determine if the observed changes in crime rates and case-solving times are significant, and explore correlations with external factors.
  + Summarize key findings, identify root causes of high crime areas, propose actionable recommendations, and prepare a comprehensive report with visualizations to communicate the results effectively.

1. 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 new table?

* The type of relationship between the **district** field in our existing data and the **district\_id** in the new table will be **many-to-one** relationship.
* **District (our data):** This is likely a field that identifies the district for various records
* **District\_ID (in the new table):** This is a unique identifier for each district in the new table.  
    
  **Relationship Type: Many-to-One**

**Report:**

The department has asked for a dashboard with three tabs:

1. Main Tab
2. Locality Tab
3. Type Exploration Tab

* **Using the Main tab in the report,** the government should be able to review the decrease in crime numbers and the improvements resulting from special police operations. This tab should include a date slicer and a filter for the primary type of crime.
* **Using the Locality Tab,** police management and the head of operations should be able to identify the most common types of crime and the localities/wards where these crimes occurred. This tab should include slicers forward, date, and primary crime type.
* **Using the Type Exploration Tab,** the Special Cops team aims to examine the total number of domestic crimes and, from those, how many arrests were made by the district. Additionally, they want to identify the crime type with the highest number of arrests. Essentially, this tab will provide metrics to summarize the department's actions and the number of tasks pending. The tab should include slicers for a month and the primary crime type.

**Make sure that all the visualizations look decent and are placed in a proper order. There are different POCs (Point Of Contact) for each tab, so make sure you involve all the metrics that POC may look at in that tab along with those mentioned in the tab description.**

**​​After making the report on the Desktop ensure that it is published on the PowerBI service and use the hosted link for submission of the dashboard and mentioning on the resume.**