

Table of Contents

Executive Summary	2
Introduction	2
Key Findings	3
Recommendations	3
Introduction	4
Dataset Description	4
BI Requirements	5
Why Did I Choose This Dataset?	5
Does My Dataset Address the Big Data Problem?	5
Which specific features are you going to focus on?	5
Will this dataset help you in developing specific business skills?	5
What KPI questions do you seek to answer with your BI project?	5
Data Preprocessing	7
Download the Data Sets	7
Combine Multiple Data Sets	7
Create Single Table	7
Removing Duplicates	7
Removing Empty Rows	7
Removing Blank Rows and Null Data	7
Changing Data Types	7
Removing Columns	7
Renaming Columns	7
Column Manipulation	8
Calculated Columns	8
Calculated Measure	8
Data Model	8
Normalization	8
Create Relationship	8
Data Visualization	10
Dashboards	10
Crime Overview Dashboard	10

Crime Sporting Dashboard	10
Area wise Crime Dashboard	11
Crime Report Changes Dashboard	11
Appearance and navigation	12
Custom Theme Design	12
Slicer Panel	12
Navigation	12
References.....	12

Executive Summary

Introduction

The comprehensive dashboards created from UK Northeast Cities' crime data provide a detailed overview of the regional crime trends and hotspots. They offer critical insights into the types and distribution of criminal activities from January to October 2023. These visual analytics serve as a vital tool for informing law enforcement strategies and enhancing public safety initiatives.

Key Findings

Here are the key findings from the analysis of the UK Northeast Cities crime data dashboards:

High Incidence of Violence and Sexual Offenses: The most common crimes reported were related to violence and sexual offenses.

Concentration of Crimes in Urban Areas: Crime incidents are densely clustered within urban centers, indicating specific hotspots that require focused attention.

Variation in Crime Over Time: There appears to be a temporal pattern or trend in crime rates, with certain months showing higher incidences.

Localization of Crime Types: Certain areas are more prone to specific types of crime, as evidenced by the detailed mapping on the crime spotting dashboard.

Volume of Crime Reports: Some areas have a high number of repeat incidents, which may indicate persistent problems or effective crime reporting mechanisms.

Distribution of Crime by Type: The dashboards provide a clear hierarchy of the prevalence of different crime types.

Resolution of Crimes: There is an extensive breakdown of crime outcomes, which helps to understand the effectiveness of the criminal justice response.

Predictive Indicators for Law Enforcement: The month-wise changes in crime reports could serve as predictive indicators for law enforcement to anticipate and prepare for future trends.

These findings can help to guide policy, improve public awareness, and enhance the strategic deployment of law enforcement resources.

Recommendations

Based on the key findings from the UK Northeast Cities crime data dashboards, here are some recommendations:

Targeted Intervention for Violence and Sexual Offenses: Implement specialized campaigns and increase patrols in areas with high incidences of violence and sexual offenses. Consider community programs that educate on prevention and support victims.

Focused Resources on Urban Crime Hotspots: Allocate more resources to urban centers with high crime rates. This could include increased police presence, community policing initiatives, and the installation of surveillance equipment.

Seasonal Crime Prevention Strategies: Analyze the variation in crime over time to develop and implement seasonal or monthly crime prevention strategies that anticipate and mitigate potential increases in crime rates.

Localized Crime Type Response: Tailor law enforcement training and response plans to the specific types of crimes that are prevalent in certain areas to improve the effectiveness of crime handling.

Enhanced Reporting and Analysis of Repeat Incidents: Investigate areas with high numbers of repeat incidents to understand and address the root causes. Improve crime reporting mechanisms to ensure accurate and timely data collection.

Crime Type Prioritization: Use the distribution of crime by type to prioritize which crimes to address first based on their prevalence and impact on the community.

Review and Improvement of Crime Resolution Processes: Assess the current crime resolution outcomes and processes to identify areas for improvement, ensuring that cases are resolved efficiently and justly.

Development of Predictive Policing Capabilities: Leverage the data on month-wise changes in crime reports to develop predictive policing models that can forecast and prevent future crimes, optimizing resource deployment.

By implementing these recommendations, law enforcement and community leaders can work towards reducing crime rates, improving public safety, and fostering a more secure environment for all residents in the UK Northeast Cities.

Introduction

To understand the crime dynamics within the northeast of England, particularly in areas like Newcastle, Middlesbrough, Redcar and others, I have embarked on a Business Intelligence (BI) project using data obtained from the [data.Police.uk](https://data.police.uk) website. This project aims to delve into crime patterns in these regions, with a special focus on street crimes.

Dataset Description

The dataset spans from January to October 2023 and covers two police jurisdictions: **Cleveland and Northumbria Police**. It is comprehensive, including columns such as Crime ID, Month, Reporting Agency, Geographic Coordinates (Longitude, Latitude), Location, Lower Layer Super Output Area (LSOA) code and name, Crime Type, Last Outcome Category, and Context, and total number of **247,864** rows in data sets.

Column Name	Type	Description
Crime ID	String	Unique identifier for each crime incident
Month	Date	Month when the crime was reported
Reported by	String	Authority or department reporting the crime
Falls within	String	Jurisdiction or area of the crime
Longitude	float64	Longitude coordinates of the crime location
Latitude	float64	Latitude coordinates of the crime location

Location	String	Textual description or address of the crime location
LSOA code	String	Lower Layer Super Output Area code
LSOA name	String	Name of the Lower Layer Super Output Area
Crime type	String	Category of the crime
Last outcome category	String	Latest status or outcome of the crime
Context	float64	Numerical data linked to the crime's context

BI Requirements

Why Did I Choose This Dataset?

I chose this dataset primarily because of my personal interest in the safety of the area where I live. The Northeast, particularly cities like Middlesbrough and Redcar, has a reputation for street crimes. This dataset provides an opportunity to analyze crime trends and patterns in these areas.

Does My Dataset Address the Big Data Problem?

The dataset addresses the big data problem by providing a large volume of varied data over a significant time span. It allows for a detailed analysis of crime types, locations, and their changes over time.

Which specific features are you going to focus on?

I will focus on crime types, specifically street crimes, and their geographical distribution within the cities. The dataset's longitudinal aspect also allows for the examination of temporal trends in crime rates.

Will this dataset help you in developing specific business skills?

Yes, analyzing this dataset will enhance my skills in data analytics, pattern recognition, and geographical information systems (GIS). It will also improve my ability to draw actionable insights from complex data sets.

What KPI questions do you seek to answer with your BI project?

The dashboards designed to analyze the UK Northeast Cities crime data aim to answer several Key Performance Indicator (KPI) questions that can help in evaluating the performance of law enforcement and the effectiveness of crime prevention strategies. Here are some of the critical KPI questions these dashboards address:

1. What is the trend of overall crime rates over time in the region?

This KPI tracks the increase or decrease in crime over the months, providing insight into whether crime is becoming frequent.

2. Which types of crime are most prevalent, and how do they vary by area?

By breaking down crime by type and location, this KPI can help law enforcement prioritize resources and strategies for the most affected areas.

3. Are there specific hotspots where crime is significantly higher?

Identifying areas with high crime rates can help in deploying targeted interventions and preventative measures.

4. What are the patterns of violent and sexual offenses reported?

Given their severity, understanding the patterns of these offenses can aid in developing specialized responses and support services.

5. How effective are the crime resolution and outcome efforts?

This KPI assesses the success of the criminal justice system in resolving cases and can indicate areas for improvement.

6. What is the volume of repeat incidents in particular locations?

High volumes of repeat incidents can indicate areas where crime prevention measures may be failing, necessitating a review and adjustment of strategies.

7. How do crime types and rates correlate with socio-demographic and economic factors in various neighborhoods?

Analyzing crime data alongside socio-demographic information can reveal underlying factors that contribute to crime, guiding community development efforts.

8. What is the response time and effectiveness of law enforcement interventions?

This KPI measures how quickly and effectively law enforcement responds to and deals with reported crimes.

9. What seasonal or temporal factors affect crime rates, and how can this information improve policing schedules?

Understanding how crime rates change with seasons or times of the day can help in optimizing patrol schedules and resource allocation.

10. Can emerging trends in the data be used to predict and prevent future crimes?

This KPI explores the potential of using historical data to forecast crime trends and implement preventative measures proactively.

These KPI questions, when answered, can provide actionable insights for law enforcement agencies, policymakers, and community leaders to enhance public safety and crime prevention efforts.

Data Preprocessing

Download the Data Sets

The initial step involved downloading separate files for each month from the Police.co.uk website, for both Cleveland and Northumbria Police departments.

Combine Multiple Data Sets

To streamline the analysis, these monthly files were combined into a single dataset. This was accomplished using Power BI's Combine Load to select a folder of files and combine them upon loading, other option I have gone through merge and append functionalities, but I choose to select the combine and select folder option, one more thing we can achieve it by placing the upcoming data sets file that are not published yet, once it will published we can place the file in data sets folder and we can Scheduled Refresh In Power BI Service, we can set up a schedule to refresh the dataset at specific intervals. This is useful for keeping our data up to date without manual intervention. The frequency of refresh can vary depending on your Power BI service plan.

Create Single Table

After combining the files, a single table was created, named "Street Crime," to facilitate a unified analysis.

Removing Duplicates

Duplicates in the dataset were identified and removed. This step is crucial to ensure the accuracy and reliability of the analysis.

Removing Empty Rows

Empty rows, which can skew results and analyses, were identified and eliminated from the dataset.

Removing Blank Rows and Null Data

Further data cleaning involved the removal of blank rows and null data across all columns, enhancing the dataset's integrity.

Changing Data Types

Each column's data type was examined and adjusted as necessary to align with the nature of the data it contained.

Removing Columns

Columns that were deemed irrelevant, such as 'Source.Name' and 'Context', were removed. This step streamlined the dataset for more focused analysis.

Renaming Columns

The 'Month' column was renamed to 'Date' for better clarity and to reflect the nature of the data more accurately.

Having meticulously completed the extensive process of data processing and cleaning, we have successfully refined our dataset to a total of **194,923** rows. This comprehensive effort ensures the accuracy and reliability of the data, paving the way for more effective and insightful analysis.

Column Manipulation

Calculated Columns

Calculated Measure

Several calculated measures were introduced for enhanced analysis:

For the 'CrimeType' table: 'Total Crime by Type' and 'Unique Crime Types'.

For the 'Location' table: 'Highest Crime Type and Count', 'Least Crime Type and Count', 'Total Crime by Location', and 'Unique Locations'. For the 'Report' table: 'Total Report'. For the 'Street Crime' table: 'Count by Crime', 'Crime by Type', 'Percentage', 'Total Crime', and 'Total Outcomes'.

These calculated columns and measures are instrumental in providing a nuanced understanding of the data, allowing for a detailed analysis of street crime trends and patterns in the covered areas.

Data Model

The process of normalizing the data involved breaking down the single 'Street Crime' table into three separate tables: 'Location', 'Crime Type', and 'Report'. This approach helps in organizing the data into a more manageable and efficient structure.

Normalization

Normalization was achieved by following these steps:

Street Crime: Retained essential columns like 'Crime ID', 'Date', 'Falls within', and 'Last outcome category', making it the central fact table that holds the key data points for each crime incident.

Location: This dimension table includes 'Crime ID' (as a foreign key), 'Latitude', 'Longitude', 'Location', 'LSOA code', and 'LSOA name'. These columns contain geographical data that describe where each crime occurred.

Crime Type: Contains 'Crime ID' (as a foreign key) and 'Crime type'. This table categorizes each crime incident, allowing for analysis based on the type of crime.

Report: Includes 'Crime ID' and 'Reported by', linking each crime to the entity that reported it.

By duplicating the original 'Street Crime' table and retaining only the necessary columns for each of the new tables, a clean and normalized set of tables was created.

Create Relationship

Power BI's auto-detection feature recognized the relationships between these tables based on the 'Crime ID' column, which acts as a primary key in the 'Street Crime' table and a foreign key in the other tables. The relationships are likely one-to-many, with 'Street Crime' being the one side, indicating that each crime ID is unique to the 'Street Crime' table and can be

related to multiple entries in the other tables if there are multiple reports, locations, or types associated with a single crime.

The data model diagram shows how the data is connected. Having such a normalized data model is beneficial for several reasons:

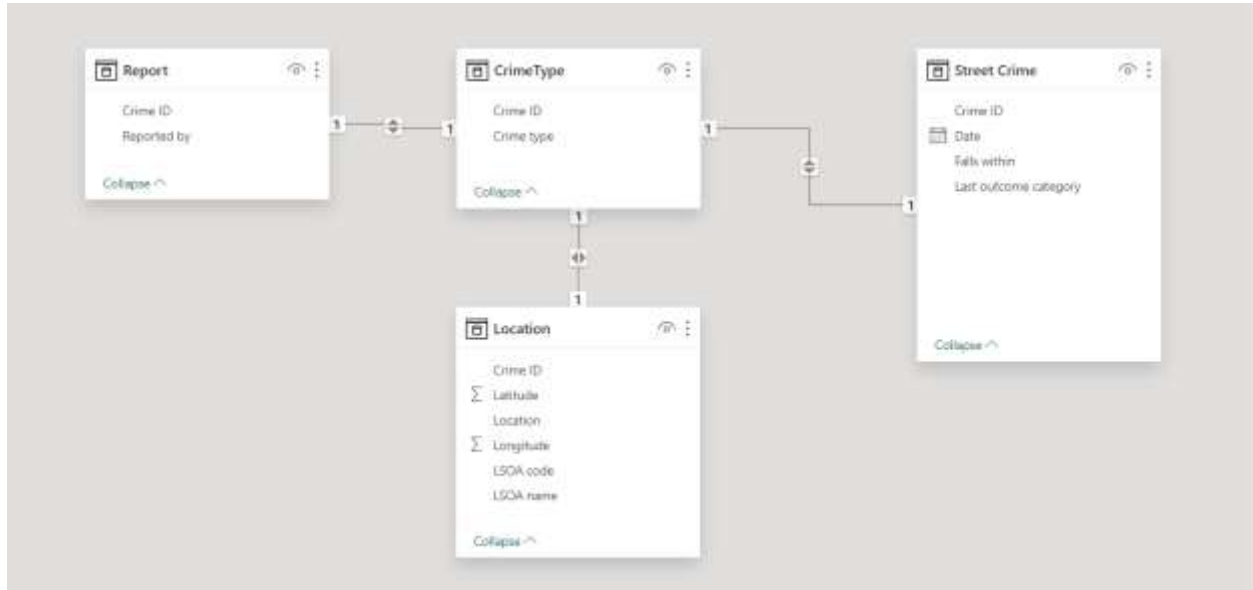


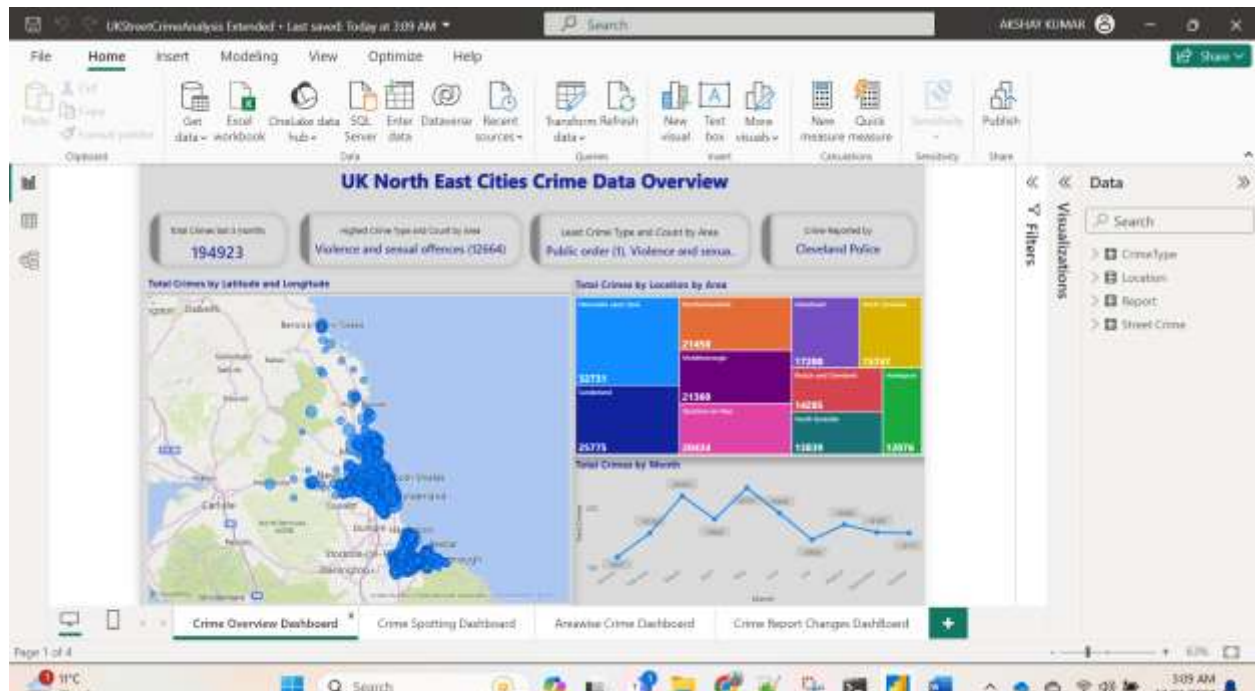
Figure 2: Data model diagram after normalization

It reduces redundancy and inconsistency, ensuring that each data element is stored only once. It simplifies the maintenance of the data because updates, deletions, and insertions are made in just one place. It enhances data integrity and accuracy, which are crucial for reliable analysis and reporting. With the relationships established, we can now perform a wide range of analyses using Power BI's powerful data visualization tools. We can explore crime trends over time, analyze crime by location, and compare the frequency of different crime types, among other insights. This normalized data model is a strong foundation for a comprehensive BI solution.

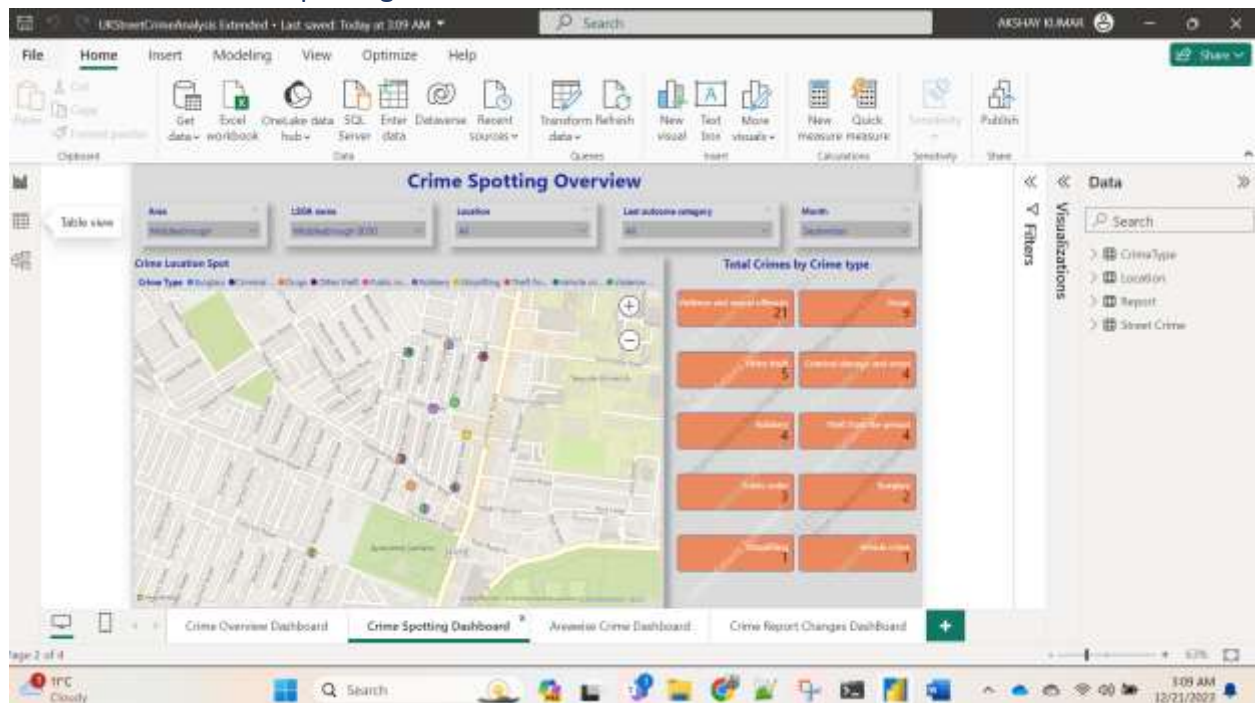
Data Visualization

Dashboards

Crime Overview Dashboard



Crime Spotting Dashboard



The screenshot displays a Tableau Desktop interface with a crime analysis dashboard. The dashboard is titled "Crime Area wise Analysis" and contains two main visualizations:

Total Reports and Unique Locations by Area

This bar chart shows the total reports and unique locations for various areas. The Y-axis represents "Total Reports" (0 to 2500) and the X-axis represents "Area". The data is as follows:

Area	Total Reports
Area 1	2400
Area 2	2200
Area 3	2000
Area 4	1800
Area 5	1600
Area 6	1400
Area 7	1200
Area 8	1000
Area 9	800
Area 10	600
Area 11	400
Area 12	200
Area 13	100
Area 14	50
Area 15	20

Area wise Crime Analysis

This table provides a detailed breakdown of crime data by area, location, count, and category.

Area	Location	Count of Crime Count	Last crime category
Area 1	Location 1	100	Category 1
Area 1	Location 2	150	Category 2
Area 1	Location 3	200	Category 3
Area 1	Location 4	250	Category 4
Area 1	Location 5	300	Category 5
Area 1	Location 6	350	Category 6
Area 1	Location 7	400	Category 7
Area 1	Location 8	450	Category 8
Area 1	Location 9	500	Category 9
Area 1	Location 10	550	Category 10
Area 1	Location 11	600	Category 11
Area 1	Location 12	650	Category 12
Area 1	Location 13	700	Category 13
Area 1	Location 14	750	Category 14
Area 1	Location 15	800	Category 15
Area 1	Location 16	850	Category 16
Area 1	Location 17	900	Category 17
Area 1	Location 18	950	Category 18
Area 1	Location 19	1000	Category 19
Area 1	Location 20	1050	Category 20
Area 1	Location 21	1100	Category 21
Area 1	Location 22	1150	Category 22
Area 1	Location 23	1200	Category 23
Area 1	Location 24	1250	Category 24
Area 1	Location 25	1300	Category 25
Area 1	Location 26	1350	Category 26
Area 1	Location 27	1400	Category 27
Area 1	Location 28	1450	Category 28
Area 1	Location 29	1500	Category 29
Area 1	Location 30	1550	Category 30
Area 1	Location 31	1600	Category 31
Area 1	Location 32	1650	Category 32
Area 1	Location 33	1700	Category 33
Area 1	Location 34	1750	Category 34
Area 1	Location 35	1800	Category 35
Area 1	Location 36	1850	Category 36
Area 1	Location 37	1900	Category 37
Area 1	Location 38	1950	Category 38
Area 1	Location 39	2000	Category 39
Area 1	Location 40	2050	Category 40
Area 1	Location 41	2100	Category 41
Area 1	Location 42	2150	Category 42
Area 1	Location 43	2200	Category 43
Area 1	Location 44	2250	Category 44
Area 1	Location 45	2300	Category 45
Area 1	Location 46	2350	Category 46
Area 1	Location 47	2400	Category 47
Area 1	Location 48	2450	Category 48
Area 1	Location 49	2500	Category 49
Area 1	Location 50	2550	Category 50
Area 1	Location 51	2600	Category 51
Area 1	Location 52	2650	Category 52
Area 1	Location 53	2700	Category 53
Area 1	Location 54	2750	Category 54
Area 1	Location 55	2800	Category 55
Area 1	Location 56	2850	Category 56
Area 1	Location 57	2900	Category 57
Area 1	Location 58	2950	Category 58
Area 1	Location 59	3000	Category 59
Area 1	Location 60	3050	Category 60
Area 1	Location 61	3100	Category 61
Area 1	Location 62	3150	Category 62
Area 1	Location 63	3200	Category 63
Area 1	Location 64	3250	Category 64
Area 1	Location 65	3300	Category 65
Area 1	Location 66	3350	Category 66
Area 1	Location 67	3400	Category 67
Area 1	Location 68	3450	Category 68
Area 1	Location 69	3500	Category 69
Area 1	Location 70	3550	Category 70
Area 1	Location 71	3600	Category 71
Area 1	Location 72	3650	Category 72
Area 1	Location 73	3700	Category 73
Area 1	Location 74	3750	Category 74
Area 1	Location 75	3800	Category 75
Area 1	Location 76	3850	Category 76
Area 1	Location 77	3900	Category 77
Area 1	Location 78	3950	Category 78
Area 1	Location 79	4000	Category 79
Area 1	Location 80	4050	Category 80
Area 1	Location 81	4100	Category 81
Area 1	Location 82	4150	Category 82
Area 1	Location 83	4200	Category 83
Area 1	Location 84	4250	Category 84
Area 1	Location 85	4300	Category 85
Area 1	Location 86	4350	Category 86
Area 1	Location 87	4400	Category 87
Area 1	Location 88		

The screenshot displays a Power BI dashboard titled "Crime Report Changes by Month". The dashboard is divided into four main sections:

- Percentage by Crime type:** A grid of 16 small bar charts showing the percentage distribution of various crime types. The crime types listed are: Rape, Sexual Assault, Sexual Offense, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape.
- Total Outcomes by Crime type:** A funnel chart showing the total outcomes for various crime types. The crime types listed are: Rape, Sexual Assault, Sexual Offense, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape, Sexual Offense - Rape.
- Total Crimes Report Changes Monthwise:** A line chart with area fill showing the total crimes report changes monthwise. The x-axis represents months from January to December, and the y-axis represents the total crimes report changes.
- Sum of Crime Count by Month:** A donut chart showing the sum of crime count by month. The months listed are: January, February, March, April, May, June, July, August, September, October, November, and December.

A sidebar on the right shows filters for "CrimeType", "Location", "Report", and "Street Crime". The top of the image shows the Power BI interface with the "Home" tab selected and various toolbars.

Appearance and navigation

- Custom Theme Design

- Slicer Panel

- Navigation

References