



General Sir John Kotelawala Defense University
Faculty of Management, Social Science and Humanities
Department of Languages

BSc in Applied Data Science Communication

D/ADC/23/0001 – RDDU Damruwan

D/ADC/23/0005- GY Punchihewa

D/ADC/23/0006 – HTSH Samadhi

D/ADC/23/0042 – YML Kavindya

SQL For Data Science/MF1224

Year 1: Semester 2

Group Assignment1 – Task 01

10.10.2023

CONTENT

01. INTRODUCTION

(Analysis)

01.EXPLORATION OF DATA

2.1 Reviewing the data

(Methodology)

2.2 Cleaning the database using Excel Database

2.3 Exporting the database from Excel to SQL

2.4 Creating the View

03. EXPORTING THE SQL SERVER VIEW TO THE POWER BI

3.1 Creating the Line Graph in the dashboard using Power BI

3.2 Creating the stacked Bar Chart

3.3 Creating the Map using Power BI

3.4 Creating the KPI using the Power BI dashboard

04. Purpose of using dashboard to Visualize Data

4.1 Line Chart

4.2 Stacked Bar Chart

4.3 Maps

4.4 Key Performance Indicator Chart

4.5 Donut Chart

05. Conclusion

Metropolitan Police Service Monthly Crime Dashboard Data

01.INTRODUCTION

The aim of this report is to illustrate the UK Metropolitan police service monthly crime datasets, key metrics and insights into the crime statistics within the metropolitan police area.

The process includes extracting, analyzing, and presenting data through a dashboard using charts, tables, and graphs to finalize the statistics about the crimes in MPS area. The raw data has been taken from UK MPS monthly crime dashboard data.

The main steps of the process are importing the raw data to the MS Excel, then removing the outliers and find quartiles with logical operations, importing edited dataset to Microsoft SQL Server, exacting, and analyzing of raw data, connecting with Microsoft SQL Server, steps of Dashboard creating and presenting actionable information through the Dashboard.

The aim of designing a dashboard is to visualize data for the clients to give a clear insight and to get an idea about the crimes in MPS area effectively and efficiently. This report clearly and orderly describes all the steps in designing a dashboard.

02. EXPLORATION OF DATA

2.1 Reviewing the data

This data set contains both string and numeric data. Area Code, Area Name, and Crime type are some main columns included in data sets. Each column name in the Excel document was created using the raw data. Following those instructions ensured that the user could use a manageable and clean dataset. you can find more details about the Metropolitan Police on this web page.

You can download the data set using the link given below.

<https://data.london.gov.uk/dataset/mps-monthly-crime-dahboard-data>

And for more information of the above data can be referred through the link given below.

<https://www.met.police.uk/>

2.2 Cleaning the database using Excel

Cleaning the dataset includes,

- Replacing the blank spaces with the “Null” values

- Removing the outliers of the dataset.

In this dataset there weren't any blank cells they were renamed as “Not Known”, “Not Given”, “N/K”. Therefore, we replaced the given words with the “Null” and cleaned the data when necessary.

It is crucial to identify possible outliers for the reasons listed below. An outlier might be a sign of faulty data. For instance, it's possible that an experiment wasn't conducted properly or that the data was not coded correctly. The most accurate way to quantify variability in skewed distributions or outlier-filled data sets is to use the interquartile range. It is unlikely to be impacted by outliers because it is based on numbers from the middle half of the distribution. There are some functions that are used to find outliers and IQR.

The codes for the above procedure is given below.

- To get the Frist Quartile – (The range of the data by the cell name,1)
- To get the Third Quartile – (The range of the data by the cell name,3)
- To get the Inter Quartile Range- (Q3-Q1)
- To get the Lower limit - $Q1-(IQR*1.5)$
- To get the Upper limit - $Q1+(IQR*1.5)$

And after that we used a logical expression to remove the outliers.

[=OR(J2<lower limit,J2>upper limit)]

14-15 - Excel

gypunchihewa@gmail.com

Share

FileHomeInsertPage LayoutFormulasDataReviewViewHelp

Tell me what you want to do

CutCopyFormat Painter

Clipboard

Calibri11

BU*I*

Font

Wrap Text

Merge & Center

Alignment

General\$ % , ' 1/2 3/4 5/8 1/16 1/32

Number

Conditional FormattingTableStyles

Format as TableCell Styles

InsertDeleteFormatCells

Σ AutoSumFillSort & Find & Filter & Select

M12

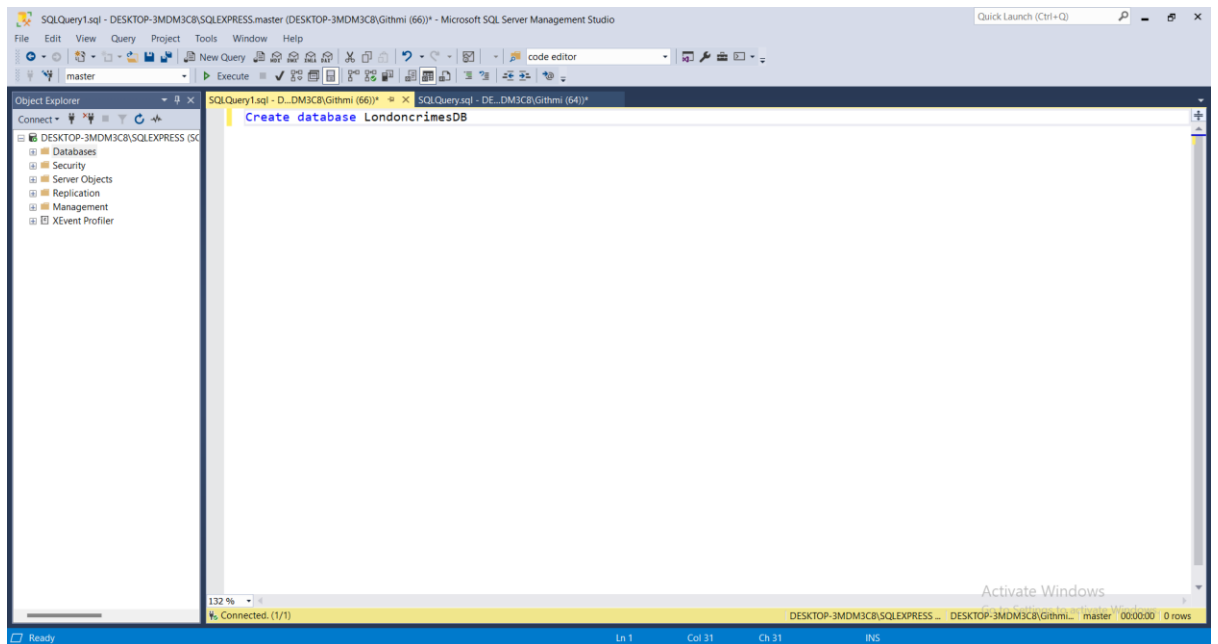
X Y Z

14-15

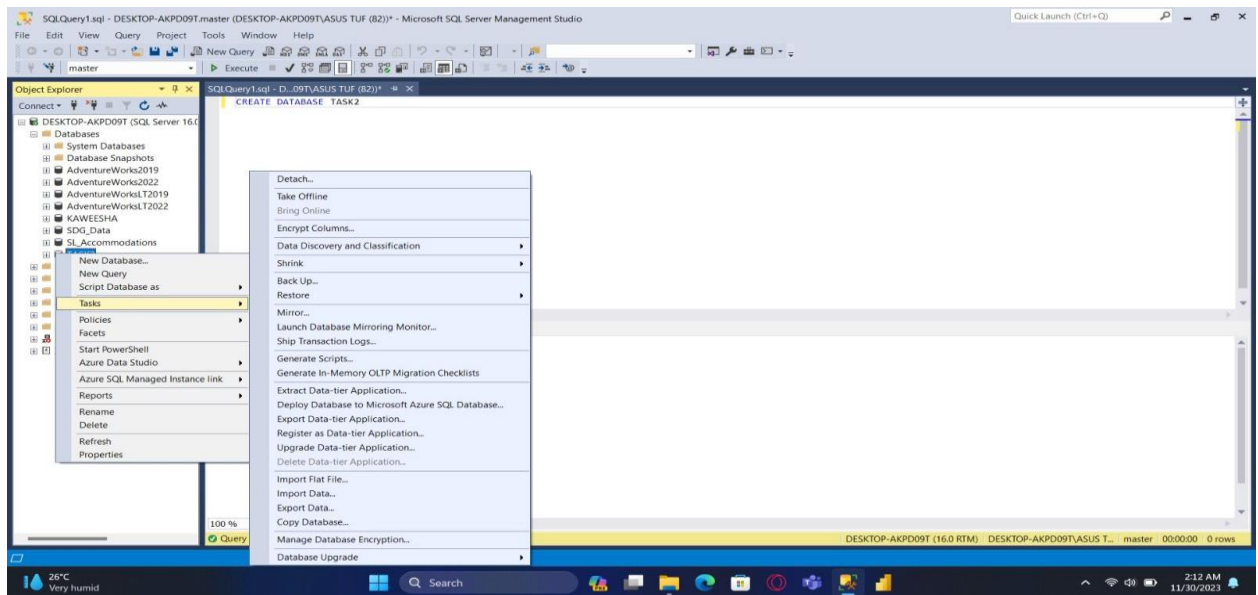
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Month_Ye	Area Type	Borough_5	Area Name	Area Code	Crime Typ	Crime Subl Measure	Financial Y	Count		outliers												
2	1/1/2015	Borough	Aviation St	Aviation St	SO18	Domestic / Domestic	Statistics	fy14-15	2		FALSE												
3	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Faith Crim Offences	fy14-15	2		FALSE												
4	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Faith Crim Outcomes	fy14-15	1		FALSE												
5	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Hate Crim Offences	fy14-15	2		FALSE												
6	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Hate Crim Outcomes	fy14-15	1		FALSE												
7	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Islamoph Offences	fy14-15	2		FALSE												
8	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Islamoph Outcomes	fy14-15	1		FALSE												
9	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Racist and Offences	fy14-15	2		FALSE												
10	1/1/2015	Borough	Aviation St	Aviation St	SO18	Hate crime	Racist and Outcomes	fy14-15	1		FALSE												
11	1/1/2015	Borough	Aviation St	Aviation St	SO18	Miscallene Theft	Pers Offences	fy14-15	3		FALSE												
12	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Gun crime	Gun Crime Offences	fy14-15	2		FALSE												
13	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Gun crime	Personal R Offences	fy14-15	1		FALSE												
14	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Hate crime	Hate Crim Outcomes	fy14-15	7		FALSE												
15	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Hate crime	Racist Crim Outcomes	fy14-15	7		FALSE												
16	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Hate crime	Racist and Outcomes	fy14-15	7		FALSE												
17	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Knife crime	Knife Injur Offences	fy14-15	6		FALSE												
18	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Knife crime	Knife Injur Offences	fy14-15	2		FALSE												
19	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Knife crime	Knife Injur Offences	fy14-15	6		FALSE												
20	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Knife crime	Knife crimi Outcomes	fy14-15	10		FALSE												
21	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Knife crime	Knife crimi Offences	fy14-15	10		FALSE												
22	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Knife crime	Knife crimi Outcomes	fy14-15	2		FALSE												
23	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Miscallene Dog Attac	Offences	fy14-15	4		FALSE												
24	1/1/2015	Borough	Barking & Barking	Barking & Barking	KG	Miscallene Robbery	m Outcomes	fy14-15	5		FALSE												
25	1/1/2015	Borough	Barnet	Barnet	SX	Gun crime	Gun Crime Offences	fy14-15	1		FALSE												
26	1/1/2015	Borough	Barnet	Barnet	SX	Gun crime	Gun Crime Outcomes	fy14-15	4		FALSE												
27	1/1/2015	Borough	Barnet	Barnet	SX	Gun crime	Personal R Offences	fy14-15	1		FALSE												
28	1/1/2015	Borough	Barnet	Barnet	SX	Hate crime	Anti-Semit Offences	fy14-15	7		FALSE												
29	1/1/2015	Borough	Barnet	Barnet	SX	Hate crime	Anti-Semit Outcomes	fy14-15	1		FALSE												

2.3 Exporting the database from Excel to SQL

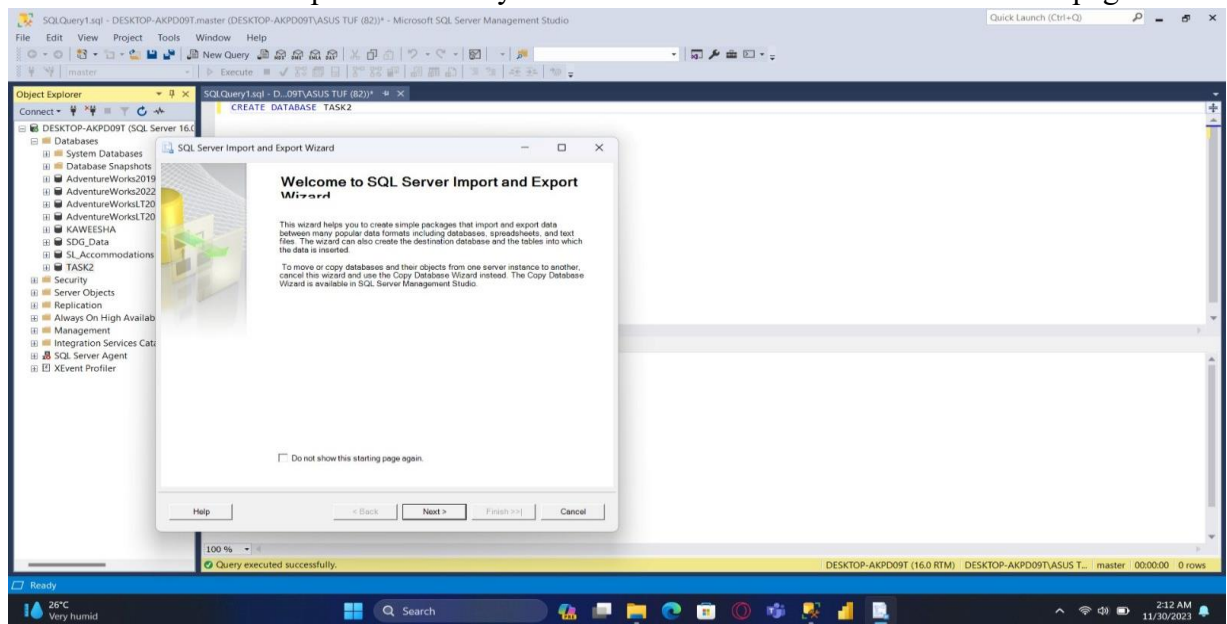
- Now Open SSMS and connect to SQL server instance.
- Create a database called 'LondonCrimesDB'.
- In the toolbar located and click the button labelled New Query.
- Now Enter the following code into a Query Editor window and execute.



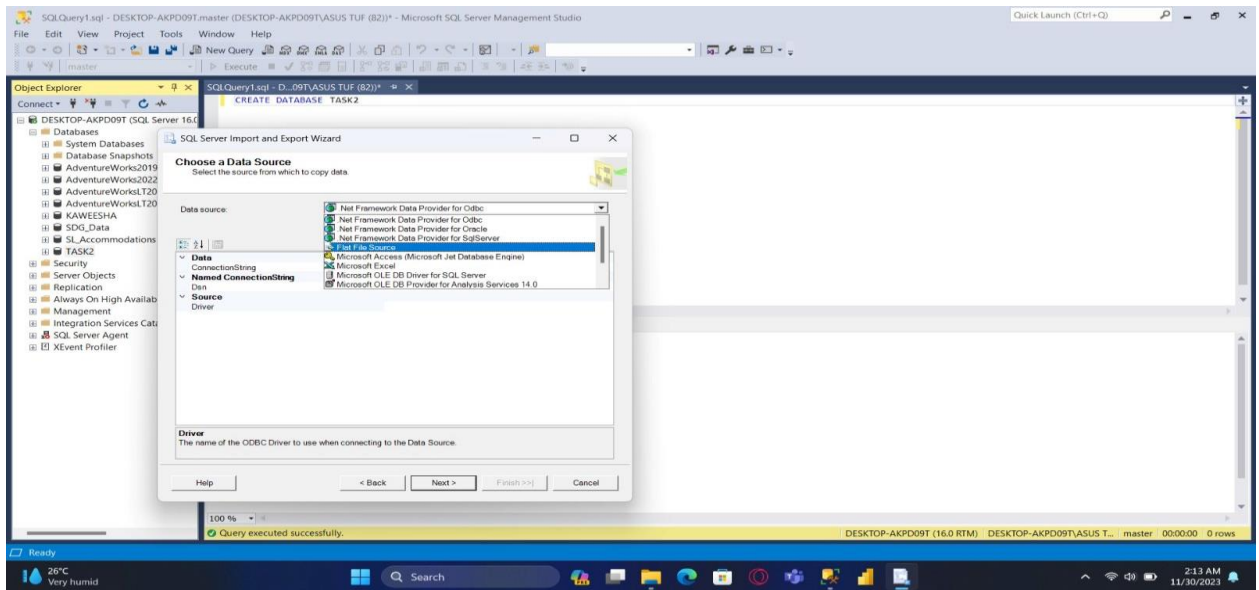
- When you refresh the Object Explorer you can see the newly created database in the dropdown of the database section.
- To import the data
- Expand the object explorer server tree.
- Expand the Database folder.
- Right click on 'LondonCrimesDB' Database.
- Select Tasks.
- Move to Import Data.



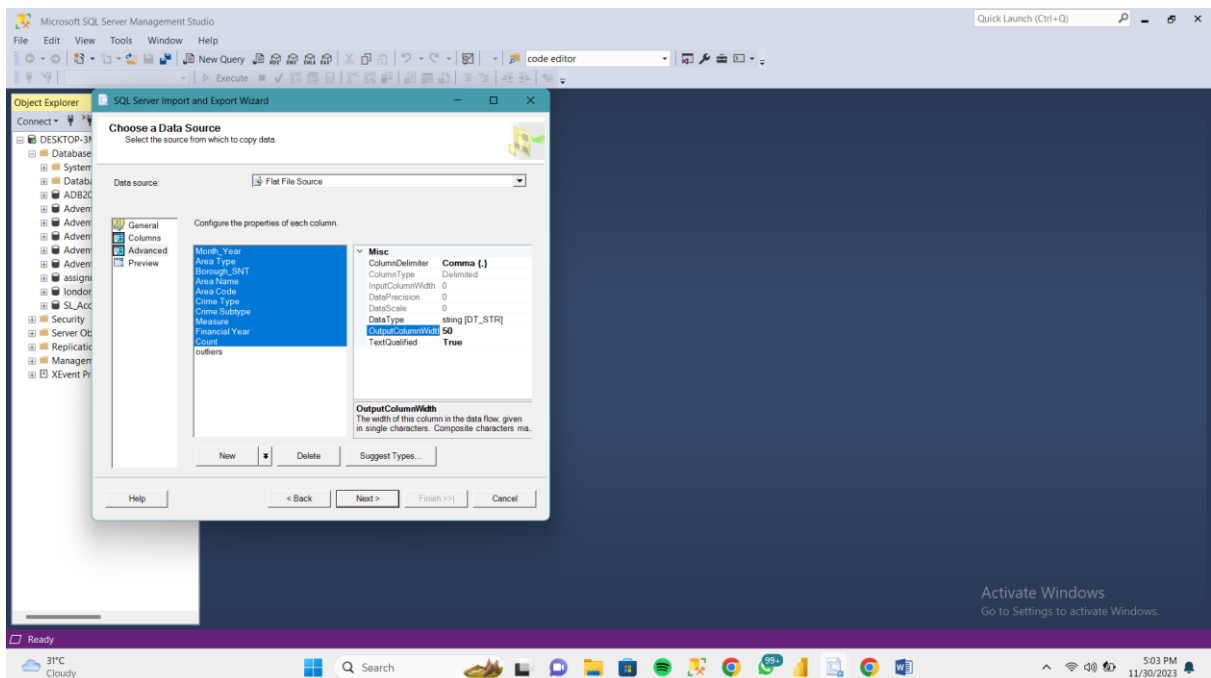
- Now select the 'Import data' and you will be directed to the Wizard Welcome page.



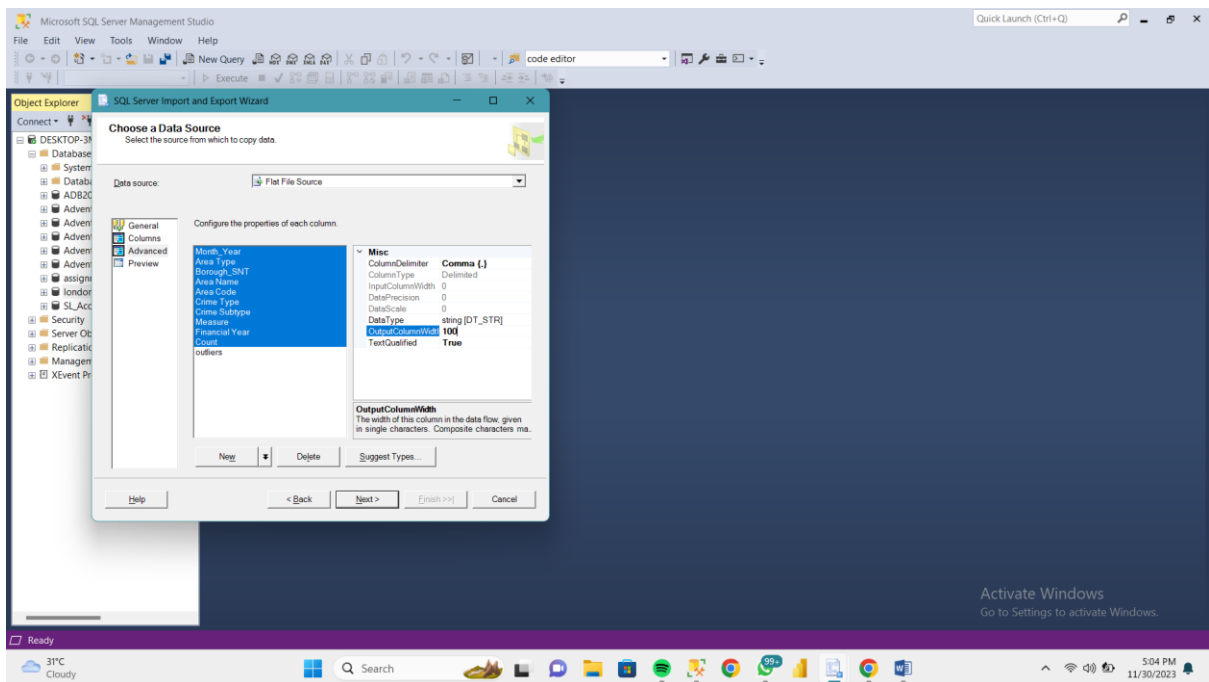
- Now you can select “Flat file” option as the data source and enter or browse the file in the device.



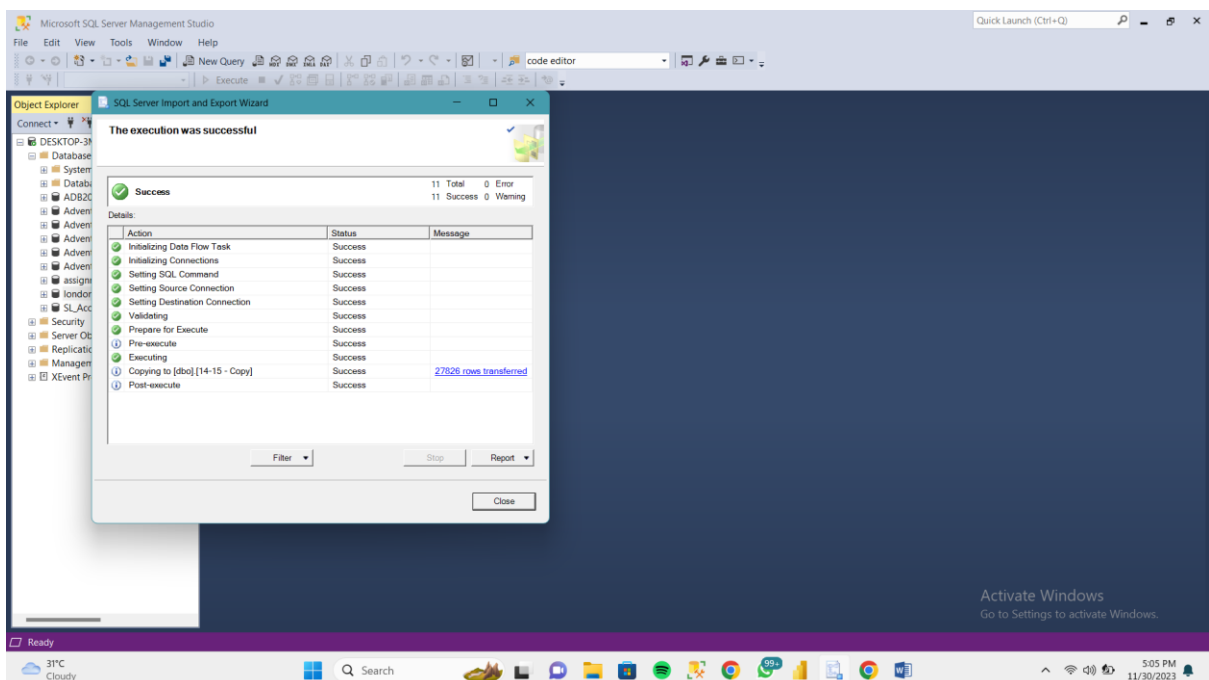
- After that go to the advance option and you will be able to see all the columns that are named in the database. And select all the columns that you want to have in the database.



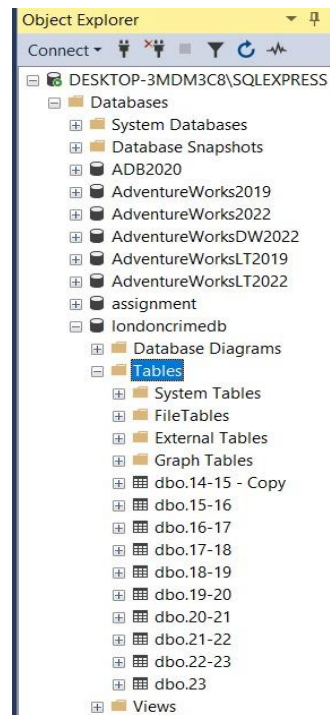
- Change the value in ‘Output Column Width’ from 50 to 100.



- Click on the next button until the process is done.

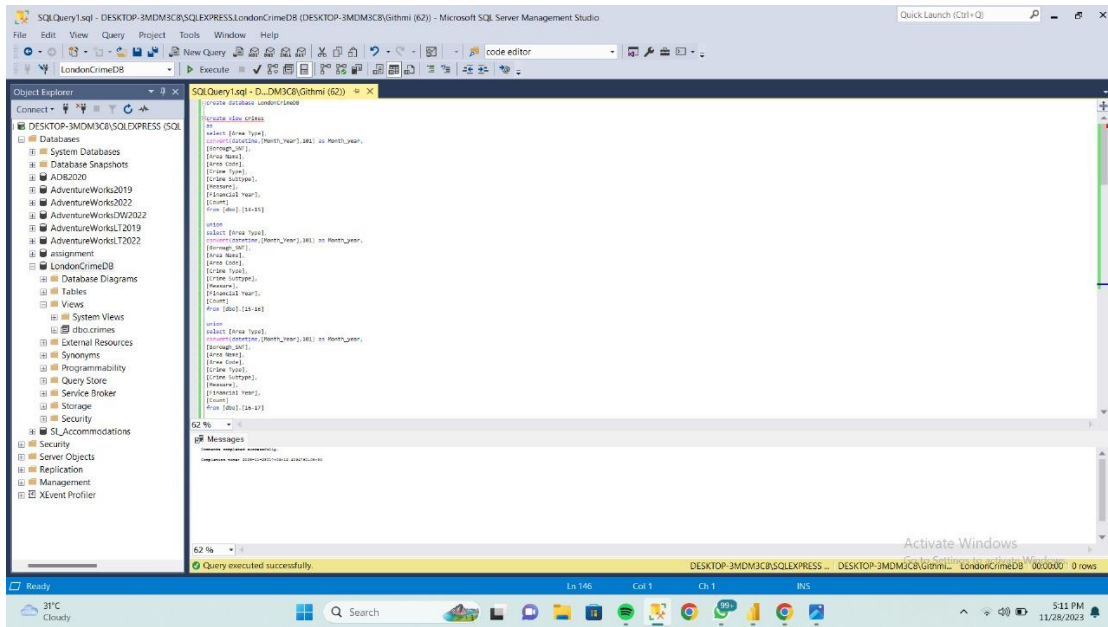


- After importing all the CSV files to the SQL Server, you can refresh the “Object Explorer” and see that the CSV files that have been imported successfully.



2.4 Creating the View

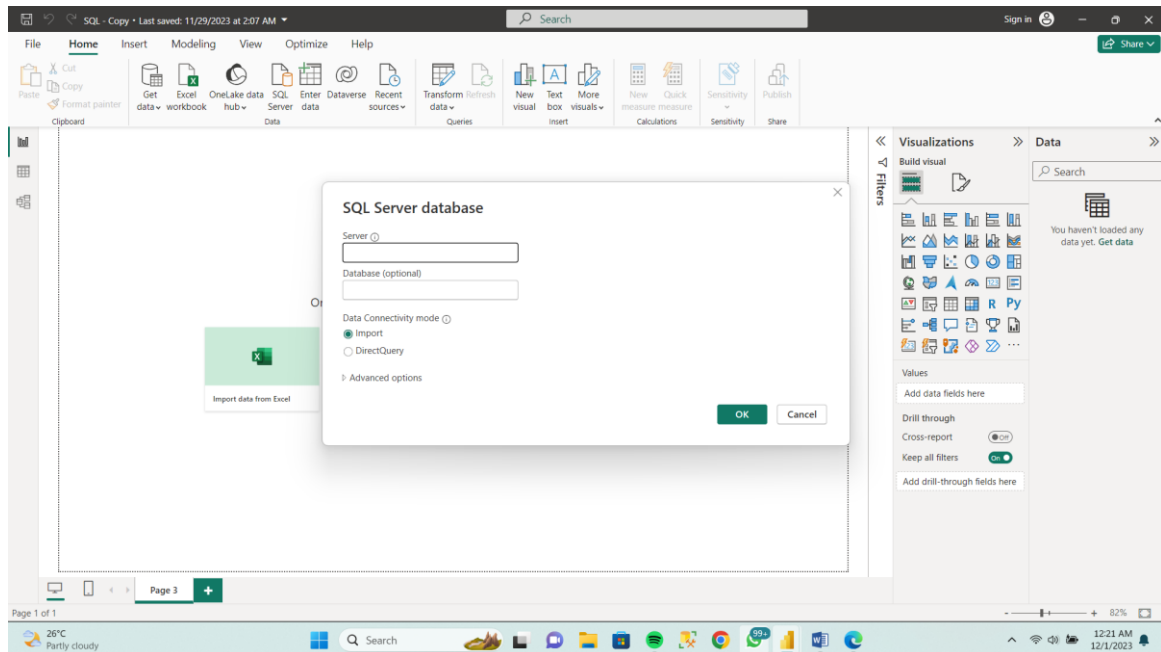
Views are typically used to help users focus, simplify, and personalize their experience with the database. By allowing users to access data through the view without offering them direct access to the view's underlying base tables, views can be utilized as security measures.



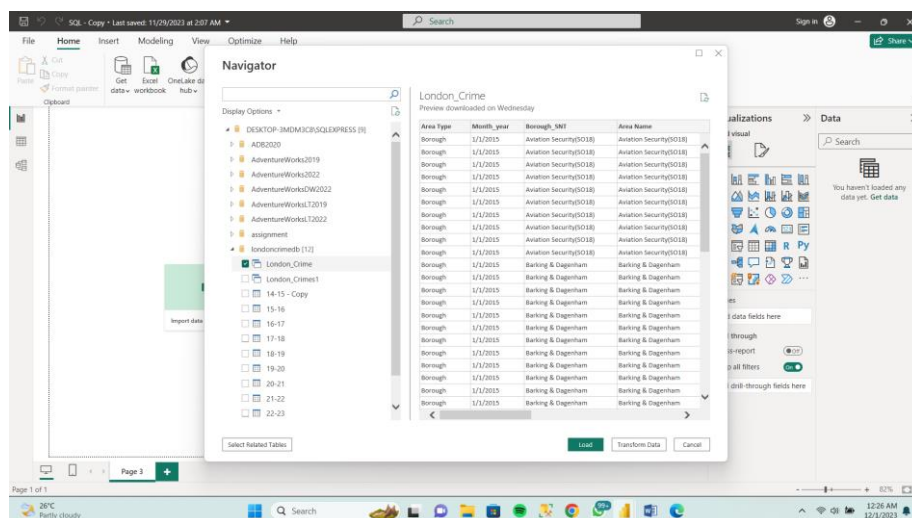
- After creating the view, you can refresh the “Object Explore” and see the view table being created under the “Views” under the “Database” section.
- Now you can export the view table to Power BI and create the dashboard.

03. EXPORTING THE SQL SERVER VIEW TO THE POWER BI

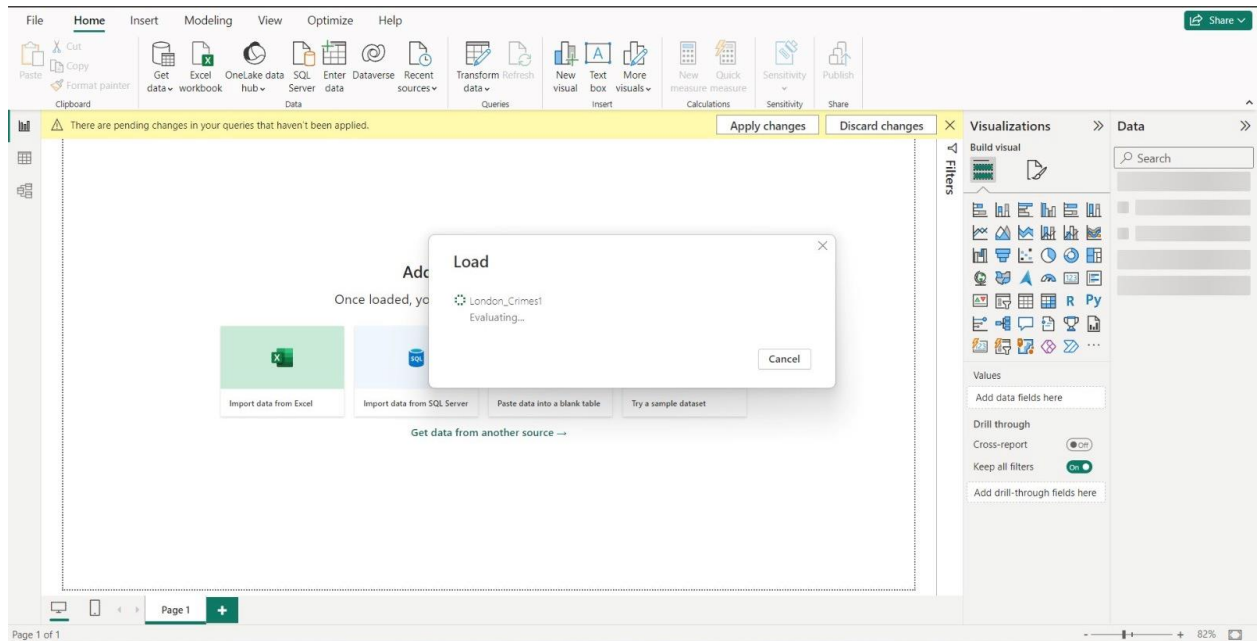
- Open the Power BI in your device. And login using the SQL Server name from the properties part SQL.



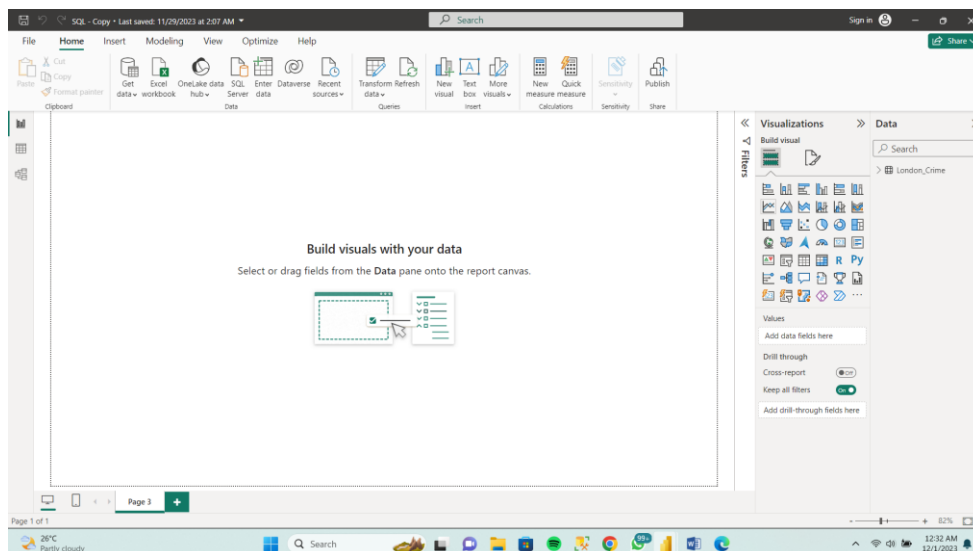
- Select the dropdown of the dataset and select the view table and load the data to Power BI.



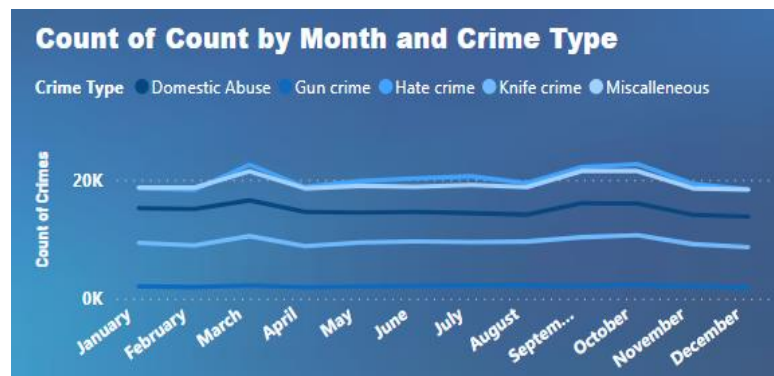
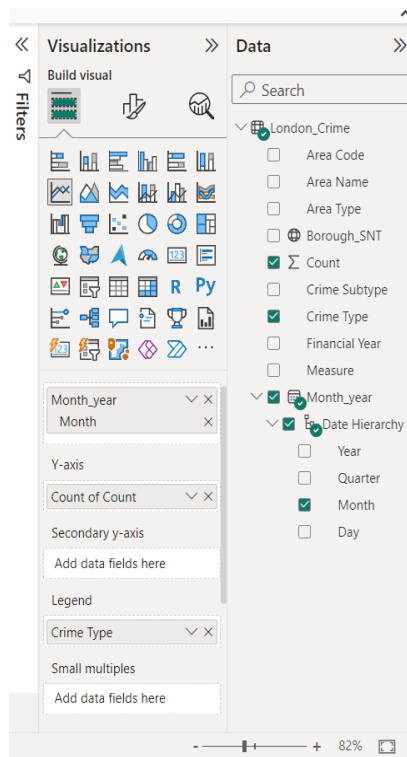
- After that this message will be visible in the desktop.



3.1 Creating the Line Graph in the dashboard using Power BI

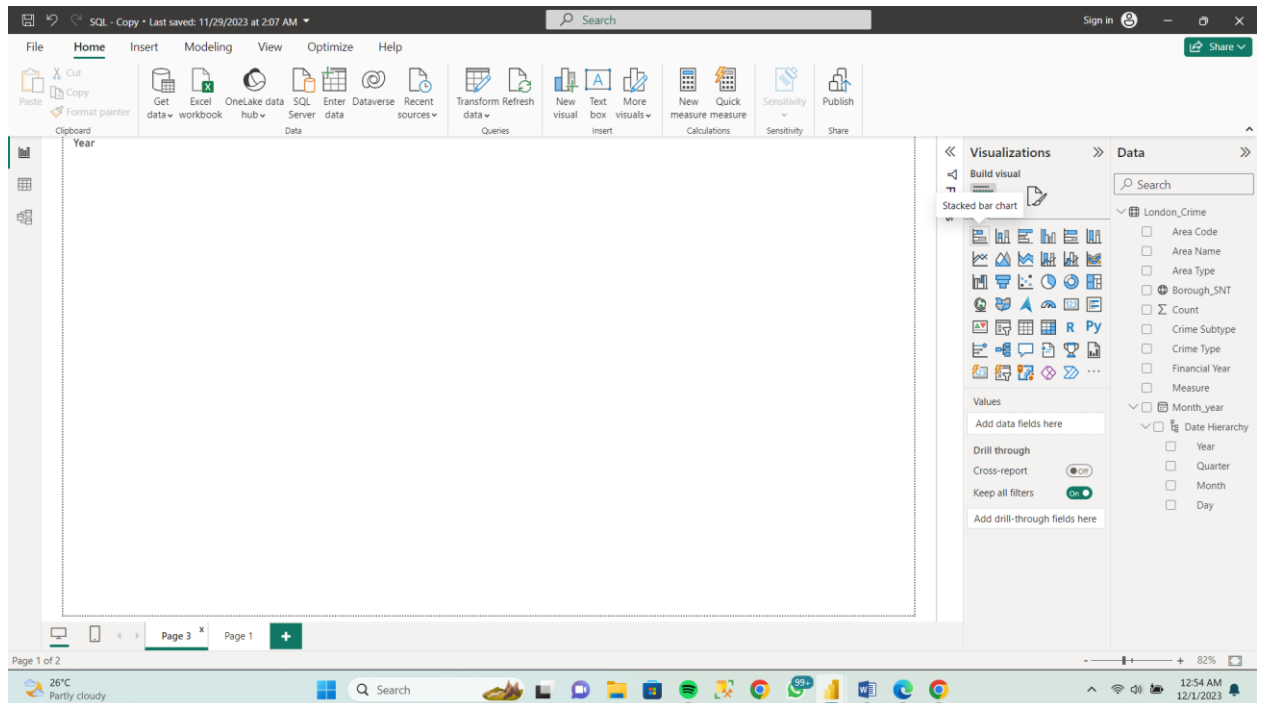


- Select the values to create the line chart as below.
X axis- Year
Y axis- Count of crimes
Legend- The Crime Types
- And you can customize the graph for your preference in the “Visual” and the “General” sections.



3.2 Creating the stacked Bar Chart

- You can create the Stacked bar chart using the “stacked bar chart” function i=given in the visualizations part in the desktop.

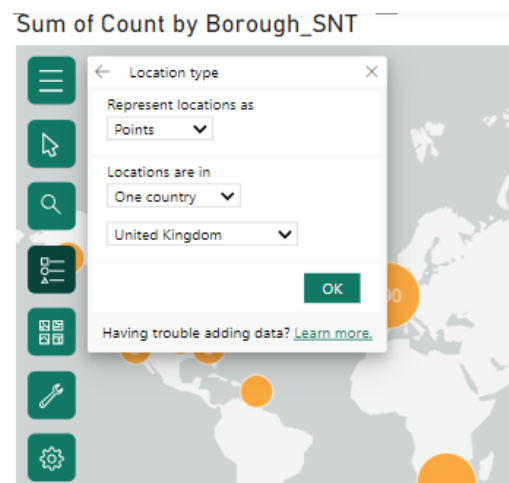
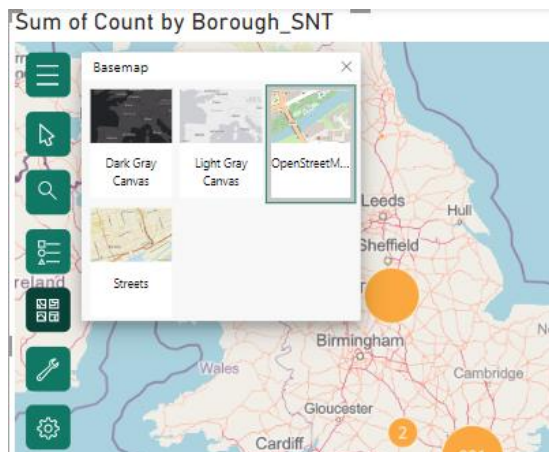
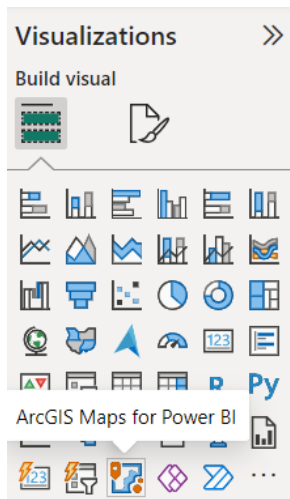


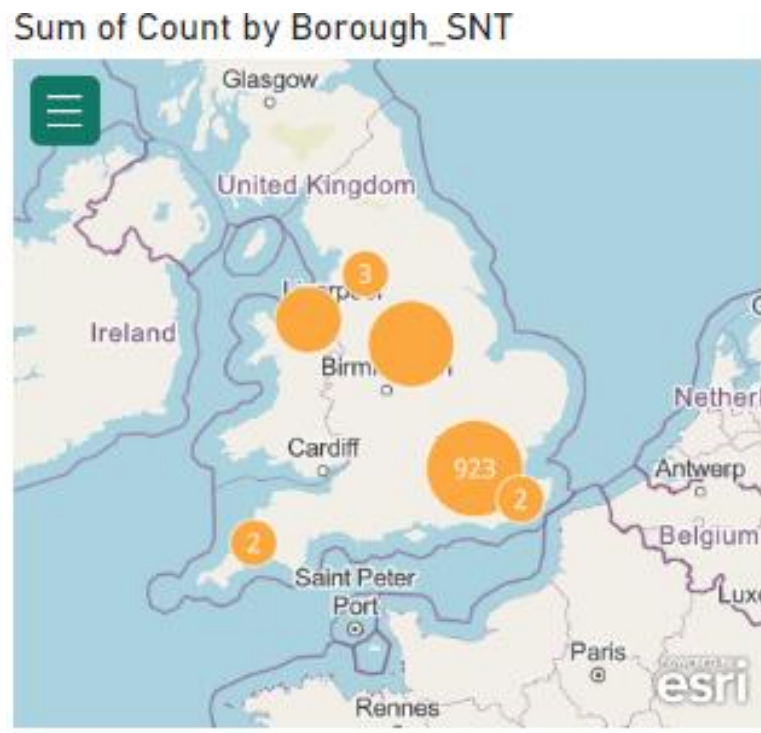
- Select the values to create the lie chart as below.
 X axis- Year
 Y axis- Count of crimes
 Legend- The Crime Types
- And you can customize the graph for your preference in the “Visual” and the “General” sections.



16 | Page

3.3 Creating the Map using Power BI

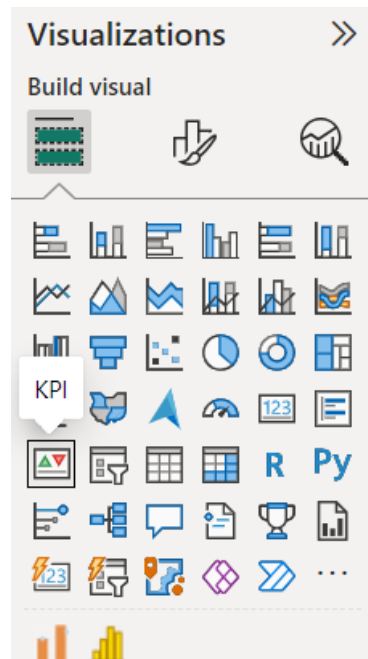




Maps offer a visually intuitive way to display geographic data, aiding in spatial analysis and comprehension. They enhance storytelling by presenting complex information in an easily digestible format, enabling effective communication and decision-making.

3.4 Creating the KPI using the Power BI dashboard

- To create a Key Performance Indicator using power BI you have to use the “KPI” chart in the visualization section.



- After that you have to assign values as follows,

Values – Total Crimes by Year

Trend axis- Crime Type

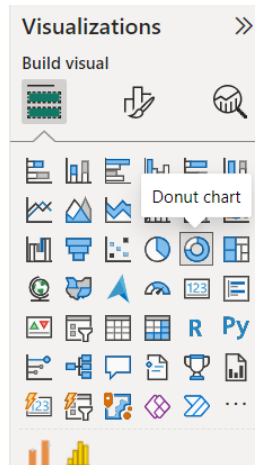
Target- Sum of crimes



3.5 Creating a Donut Chart using Power BI dashboard

Donut charts effectively illustrate the composition of data by displaying proportions of a whole, making it easy to grasp relative sizes of categories. Their ring-shaped design allows for better utilization of space, enabling clear visualization of multiple categories while maintaining a focus on comparisons and emphasizing percentages or proportions within the data.

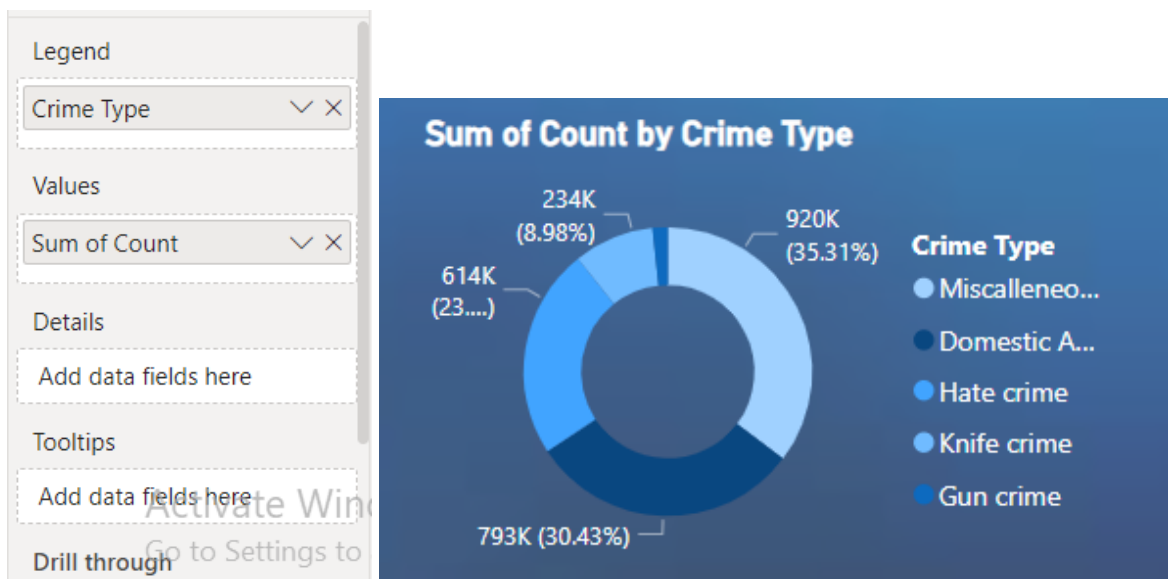
- To create this chart, select the donut icon in the visualization section.



- After selecting the donut chart, you have to assign values to indicate the summarize the data.

Legend- Crime Type

Values- Sum of Count



04. Purpose of using dashboard to Visualize Data

Line Chart

Advantages- Line charts are ideal for displaying trends and patterns over time, offering a clear visualization of relationships, identifying correlations, and highlighting changes or fluctuations within data sets.

Disadvantages- Line charts might oversimplify complex relationships, lack clarity when displaying a vast array of categories, and could mislead if data points are sparse or irregularly distributed.

Stacked Bar Chart

Advantages- Stacked bar charts display multiple categories within each bar, facilitating easy comparison of both individual segments and the total while showcasing the composition and trends across different groups.

Disadvantages- Stacked bar charts might obscure precise values within individual categories, especially in taller stacks, making it challenging to compare elements accurately across different groups or segments.

Maps

Advantages- Geographical maps offer a concise and intuitive method to display spatial data, aiding in pattern recognition, spatial analysis, and effective communication of location-based information.

Disadvantages- Geographical maps can sometimes oversimplify complex data, leading to misinterpretation or neglect of non-spatial correlations, and might present challenges in accurately representing data with irregular geographic boundaries.

Key Performance Indicator Chart

Advantages- KPI charts offer a concise snapshot of key performance indicators, facilitating quick insights into business metrics, aiding in goal tracking, and enabling swift decision-making based on crucial data trends.

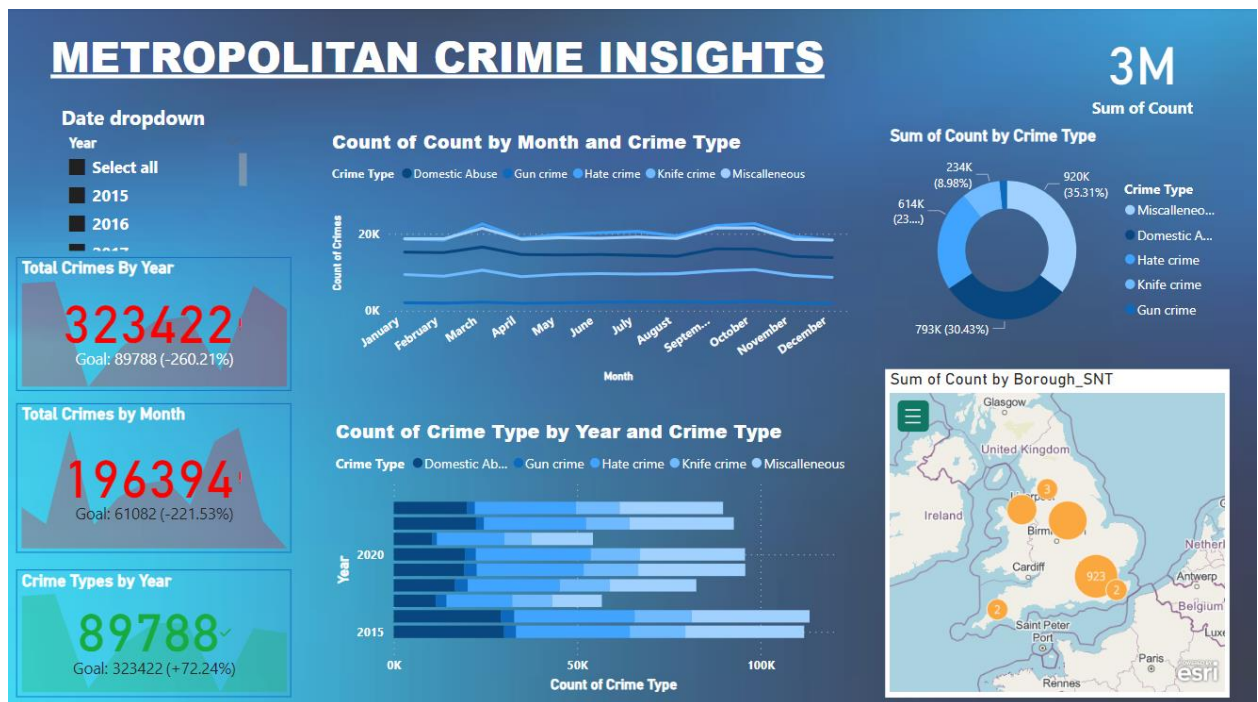
Disadvantages- KPI charts can oversimplify complex data, potentially neglecting crucial context or nuances. They might lead to misinterpretation when used in isolation without supporting information or benchmarks.

Donut Chart

Advantages- Donut charts present data as parts of a whole, aiding in visualizing proportions and comparisons succinctly, allowing for a clear understanding of relative sizes within categories.

Disadvantages- Donut charts might be less effective in displaying precise values, especially when dealing with numerous categories, and can potentially make comparisons between segments challenging

05. Conclusion



London crimes from 2019 to 2022. This database was constructed under the term “Metropolitan Crime Insights”.

The dashboard was created using data from the London Crimes that happened in the Metropolitan police district. The dashboard is created by using Excel to clean the, SQL to combine and recreate the database and Power BI to create the dashboard given above.

The dashboard provided a summary of how different the crimes that have occurred yearly from 2015 to 2022 October. By referring to the map in the dashboard you can get data driven insight about which police division has faced the highest number of crimes. There are total Five different types of crimes. Out of all of them Domestic abuses are the mostly occurred crime in the past few years it can be showed as 30% of the total crime types and the mostly occurred area is London City. The crime which has the least occurring is Gun crimes which takes only 1.71% as of the total crimes.

People may gain a data driven clear insight of how crimes has varied throughout the seven years of time by referring to the given dashboard.