General Sir John Kotelawala Defence University

Faculty of Management, Social Sciences and Humanities

Department of Languages

BSc in Applied Data Science Communication

Group - 09

D/ADC/23/0027 – PGD Rodrigo
D/ADC/23/0031 – PRM Perera
D/ADC/23/0034 – AGAUS Gunasekara
D/ADC/23/0040 – PMSS Rathnayaka

SQL for Data Science (LB1224)

Year 1: Semester 2

Group Assignment

30.11.2023

# The Report of

# Greater Manchester

# Property Price Monitor of

# UK property Market

# & Dashboard design

# Table of content

## Task 01

## Introduction

Greater Manchester Property Price Monitor is a simple and low-cost Power BI-based reporting solution that can be used to analyze the Greater Manchester area property market. It allows users to extract and present data in charts, tables, and other visualizations so users can find useful information about the different kinds of property types. The purpose of these tools is to translate data into actionable information to understand the local property market.

This report describes property details in the UK. The raw data has been taken from UK Price Paid Dataset which contains information about the UK property market from 2019 to 2022. This report describes how to import data for Microsoft SQL Server, how to import data from Microsoft SQL Server to Power BI, and how to develop a Power BI dashboard. This report also contains information about suggested designs and advice, as well as SQL statements with relevant comments and a full backup of the database and dashboard contents. This paper outlines all of the procedures involved in creating a dashboard in a straightforward and logical manner. The dashboard is intended to help clients understand data on UK property transactions. To detect the changes that occur during each year, the data is summarized and viewed through several types of charts.

## Exploration of data

### Data Set Review

This data set contains both string and numeric data under the column names; Transaction ID, Sales Price, Date of Transfer, Postcode, Property Type, Old/New, Duration, Primary Addressable Object Name , Secondary Addressable Object Name, Street, Locality, Town/City, District, County, Type of Price Paid, Record Status.

Below is a description of them;

1. Transaction ID - A reference number which is generated automatically recording each published sale. The number is unique and will change each time a sale is recorded.
2. Sales Price - The sale price is stated on the transfer deed.
3. Date of Transfer - Date when the sale was completed, as stated on the transfer deed.

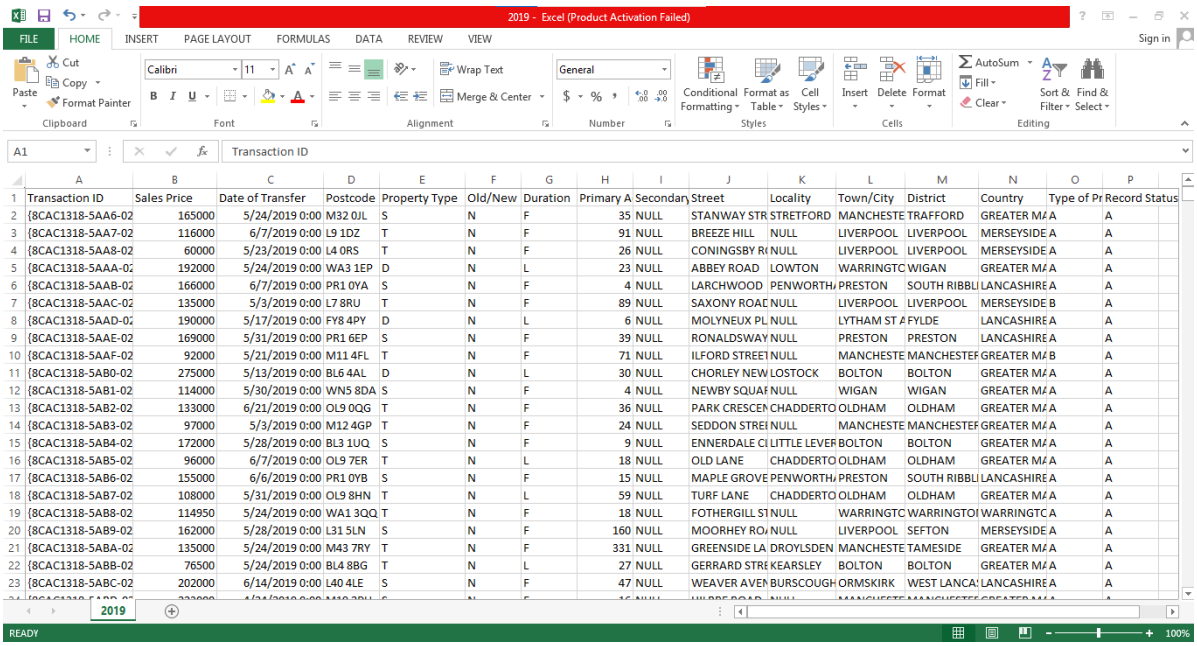
4. Postcode - This is the postcode used at the time of the original transaction. Note that postcodes can be reallocated, and these changes are not reflected in the Price Paid Dataset.
5. Property Type (D = Detached, S = Semi-Detached, T = Terraced, F = Flats/Maisonettes, O = Other)
6. Old/New - Indicates the age of the property and applies to all price-paid transactions, residential and non-residential. (Y = a newly built property, N = an established residential building)
7. Duration – Relates to the tenure. (F = Freehold, L= Leasehold etc.)
8. PAON (Primary Addressable Object Name) - Typically, the house number or name.
9. SAON (Secondary Addressable Object Name) - Where a property has been divided into separate units (for example, flats), the PAON (above) will identify the building and a SAON will be specified that identifies the separate unit/flat.
10. Street
11. Locality
12. Town/City
13. District
14. County
15. Type of Price Paid - Indicates the type of Price Paid transaction. (A = Additional Price paid entry includes single residential property sold for value, B = Additional Price paid entry including transfers under power of sale)
16. Record Status - Indicates additions, changes, and deletions to the records. (A = Addition, C = Change, D = Delete)

**Importing Data from Excel to a Microsoft SQL Server Database**

1. Download the 2019, 2020, 2021, 2022 UK property sales CSV files from the below link.

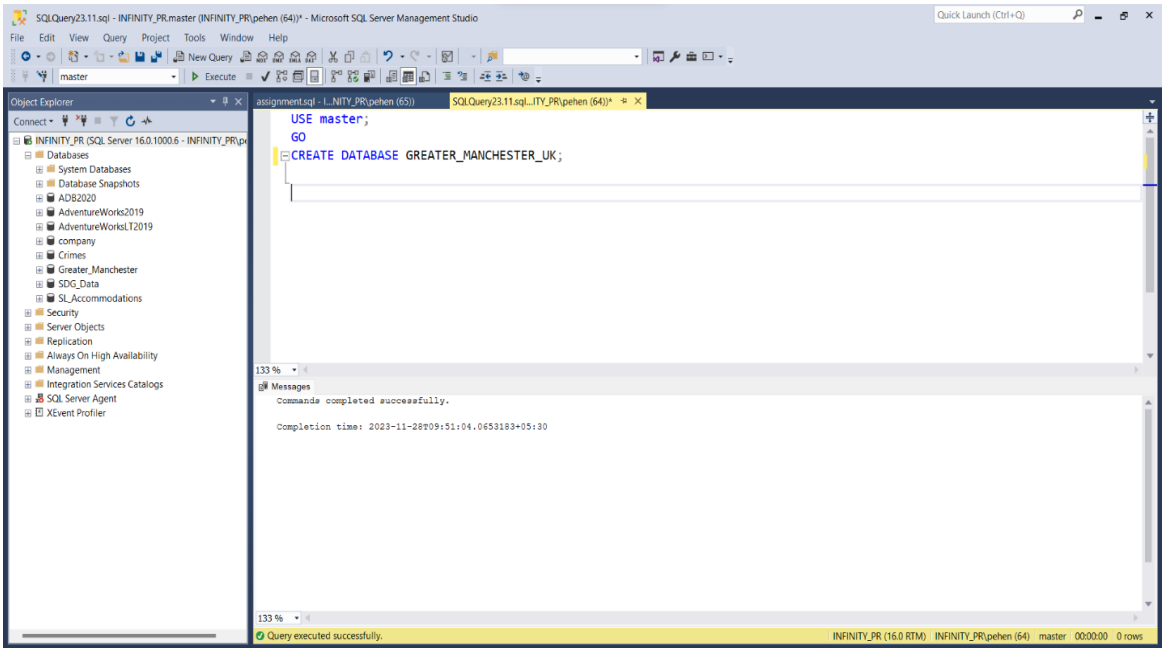
*(https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads)*

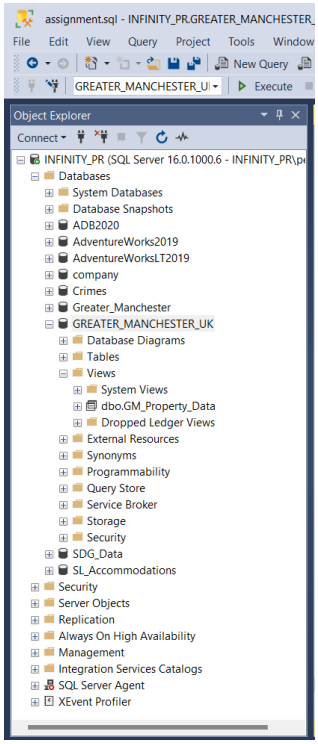
2. Insert column names to the datasets. (Transaction ID, Sales Price, Date of Transfer, Postcode, Property Type, Old/New, Duration, Primary Addressable Object Name , Secondary Addressable Object Name, Street, Locality, Town/City, District, County, Type of Price Paid, Record Status)
3. Fill the blank spaces in the datasets with NULL.




| Transaction ID | Sales Price | Date of Transfer | Postcode | Property Type | Old/New | Duration | Primary A | Secondary | Street | Locality | Town/City | District | Country | Type of Pr | Record Status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| {8CAC1318-5AA6-02 | 165000 | 5/24/2019 0:00 | M32 0JL | S | N | F | 35 | NULL | STANWAY STR | STRETFORD | MANCHESTE | TRAFFORD | GREATER M/ | A | A |
| {8CAC1318-5AA7-02 | 116000 | 6/7/2019 0:00 | L9 1DZ | T | N | F | 91 | NULL | BREEZE HILL | NULL | LIVERPOOL | LIVERPOOL | MERSEYSIDE | A | A |
| {8CAC1318-5AA8-02 | 60000 | 5/23/2019 0:00 | L4 0RS | T | N | F | 26 | NULL | CONINGSBY R | NULL | LIVERPOOL | LIVERPOOL | MERSEYSIDE | A | A |
| {8CAC1318-5AAA-02 | 192000 | 5/24/2019 0:00 | WA3 1EP | D | N | L | 23 | NULL | ABBEY ROAD | LOWTON | WARRINGTC | WIGAN | GREATER M/ | A | A |
| {8CAC1318-5AAB-02 | 166000 | 6/7/2019 0:00 | PR1 0YA | S | N | F | 4 | NULL | LARCHWOOD | PENWORTH/ | PRESTON | SOUTH RIBBL | LANCASHIRE | A | A |
| {8CAC1318-5AAC-02 | 135000 | 5/3/2019 0:00 | L7 8RU | T | N | F | 89 | NULL | SAXONY ROAD | NULL | LIVERPOOL | LIVERPOOL | MERSEYSIDE | B | A |
| {8CAC1318-5AAD-02 | 190000 | 5/17/2019 0:00 | FY8 4PY | D | N | L | 6 | NULL | MOLYNEUX PL | NULL | LYTHAM ST A | FYLDE | LANCASHIRE | A | A |
| {8CAC1318-5AAE-02 | 169000 | 5/31/2019 0:00 | PR1 6EP | S | N | F | 39 | NULL | RONALDSWAY | NULL | PRESTON | PRESTON | LANCASHIRE | A | A |
| {8CAC1318-5AAF-02 | 92000 | 5/21/2019 0:00 | M11 4FL | T | N | F | 71 | NULL | ILFORD STREE | NULL | MANCHESTE | MANCHESTEF | GREATER M/ | B | A |
| {8CAC1318-5AB0-02 | 275000 | 5/13/2019 0:00 | BL6 4AL | D | N | L | 30 | NULL | CHORLEY NEW | LOSTOCK | BOLTON | BOLTON | GREATER M/ | A | A |
| {8CAC1318-5AB1-02 | 114000 | 5/30/2019 0:00 | WN5 8DA | S | N | F | 4 | NULL | NEWBY SQUAI | NULL | WIGAN | WIGAN | GREATER M/ | A | A |
| {8CAC1318-5AB2-02 | 133000 | 6/21/2019 0:00 | OL9 0QG | T | N | F | 36 | NULL | PARK CRESCEI | CHADDERTO | OLDHAM | OLDHAM | GREATER M/ | A | A |
| {8CAC1318-5AB3-02 | 97000 | 5/3/2019 0:00 | M12 4GP | T | N | F | 24 | NULL | SEDDON STRE | NULL | MANCHESTE | MANCHESTEF | GREATER M/ | A | A |
| {8CAC1318-5AB4-02 | 172000 | 5/28/2019 0:00 | BL3 1UQ | S | N | F | 9 | NULL | ENNERDALE CI | LITTLE LEVEF | BOLTON | BOLTON | GREATER M/ | A | A |
| {8CAC1318-5AB5-02 | 96000 | 6/7/2019 0:00 | OL9 7ER | T | N | L | 18 | NULL | OLD LANE | CHADDERTO | OLDHAM | OLDHAM | GREATER M/ | A | A |
| {8CAC1318-5AB6-02 | 155000 | 6/6/2019 0:00 | PR1 0YB | S | N | F | 15 | NULL | MAPLE GROVE | PENWORTH/ | PRESTON | SOUTH RIBBL | LANCASHIRE | A | A |
| {8CAC1318-5AB7-02 | 108000 | 5/31/2019 0:00 | OL9 8HN | T | N | L | 59 | NULL | TURF LANE | CHADDERTO | OLDHAM | OLDHAM | GREATER M/ | A | A |
| {8CAC1318-5AB8-02 | 114950 | 5/24/2019 0:00 | WA1 3QQ | T | N | F | 18 | NULL | FOTHERGILL S | NULL | WARRINGTC | WARRINGTOI | WARRINGTC | A | A |
| {8CAC1318-5AB9-02 | 162000 | 5/28/2019 0:00 | L31 5LN | S | N | F | 160 | NULL | MOORHEY RO/ | NULL | LIVERPOOL | SEFTON | MERSEYSIDE | A | A |
| {8CAC1318-5ABA-02 | 135000 | 5/24/2019 0:00 | M43 7RY | T | N | F | 331 | NULL | GREENSIDE LA | DROYLSDEN | MANCHESTE | TAMESIDE | GREATER M/ | A | A |
| {8CAC1318-5ABB-02 | 76500 | 5/24/2019 0:00 | BL4 8BG | T | N | L | 27 | NULL | GERRARD STRE | KEARSLEY | BOLTON | BOLTON | GREATER M/ | A | A |
| {8CAC1318-5ABC-02 | 202000 | 6/14/2019 0:00 | L40 4LE | S | N | F | 47 | NULL | WEAVER AVEI | BURSCOUGH | ORMSKIRK | WEST LANCA | LANCASHIRE | A | A |

4. Open SSMS and connect to SQL server instance.
5. Create a new database called “GREATER_MANCHESTER_UK”.

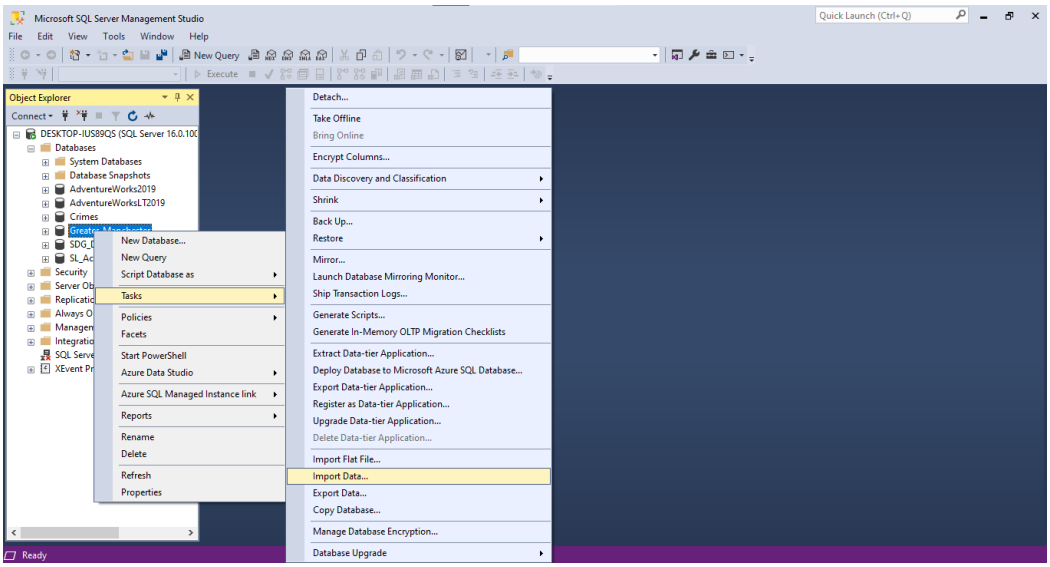


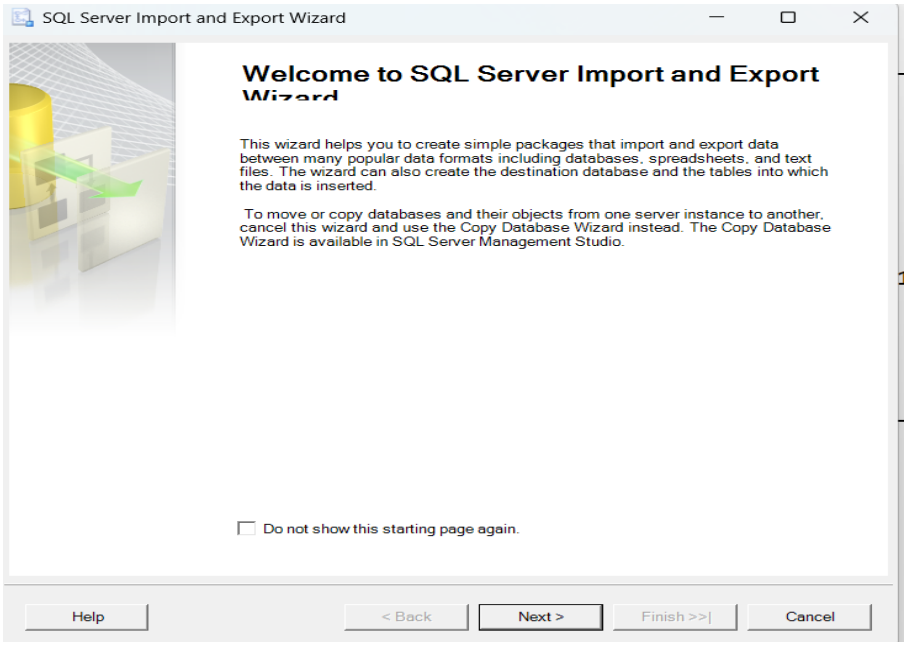

6. Import the CSV file year by year to the SQL server database ‘GREATER_MANCHESTER_UK’.

   a) Right-click the ‘GREATER_MANCHESTER_UK’ database.

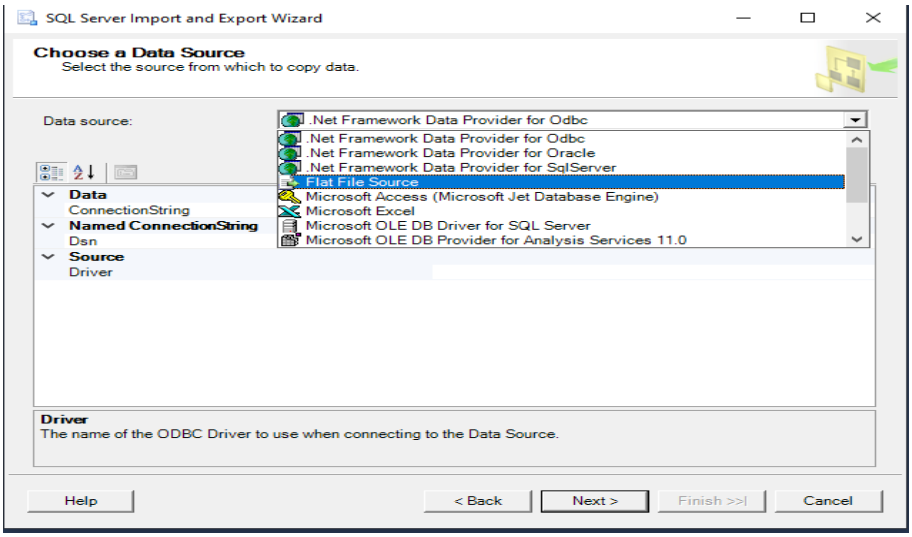
   b) Select Tasks → Import Data

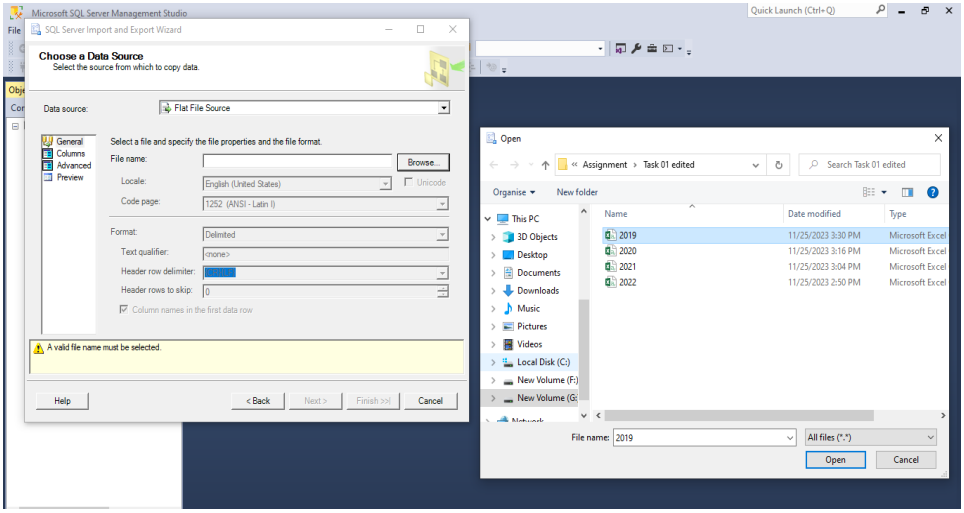



c) Click Next on the SQL Server Import and Export Wizard welcome page.

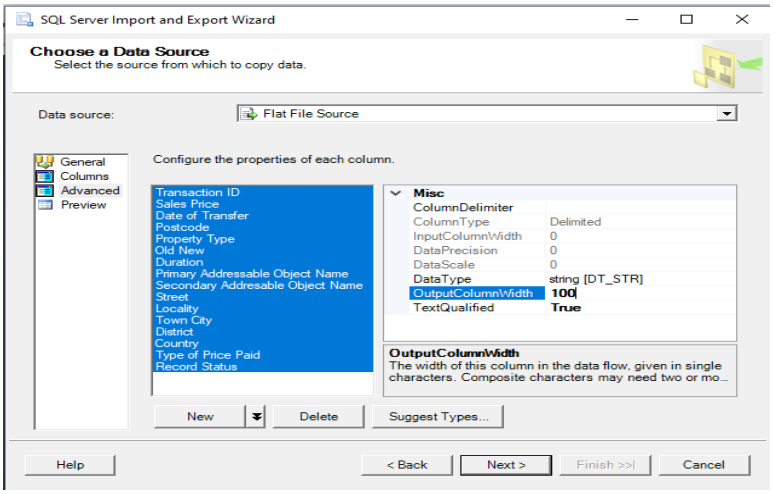



d) Select Flat File Source as the Data Source, and enter or browse for the file to import.

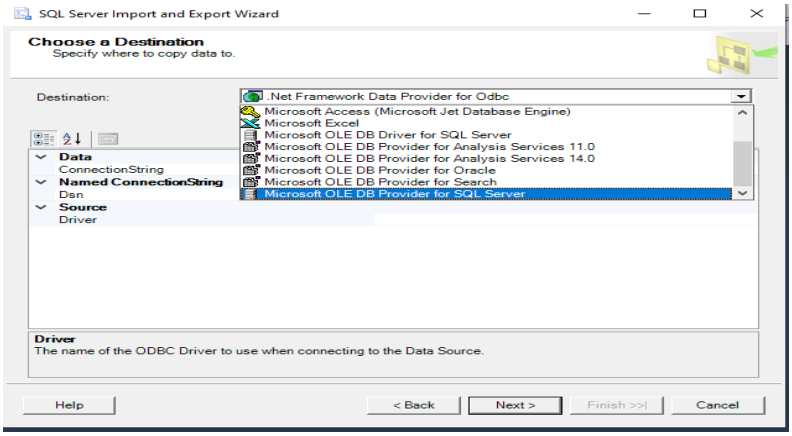


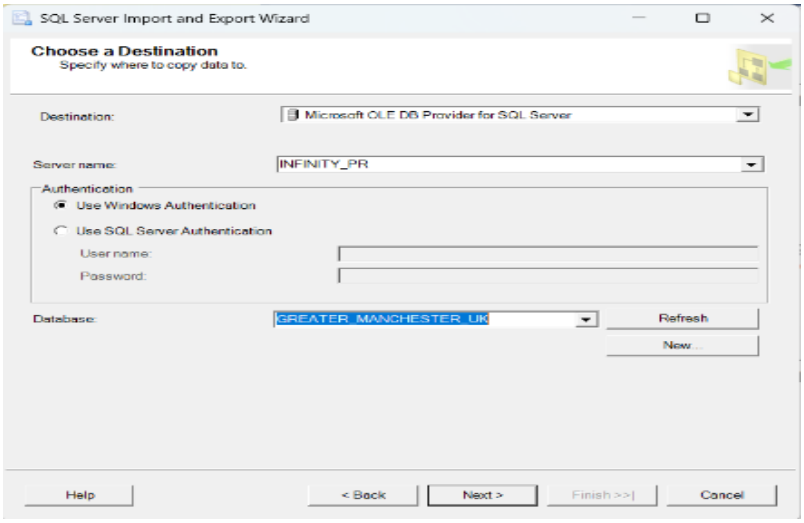




e) Go to the Advanced tab and change all the column widths from 50 to 100 and click Next.

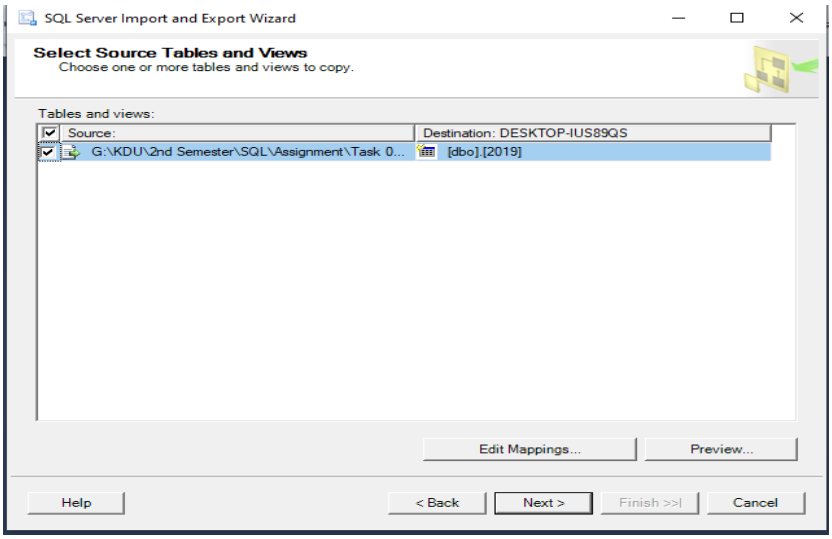


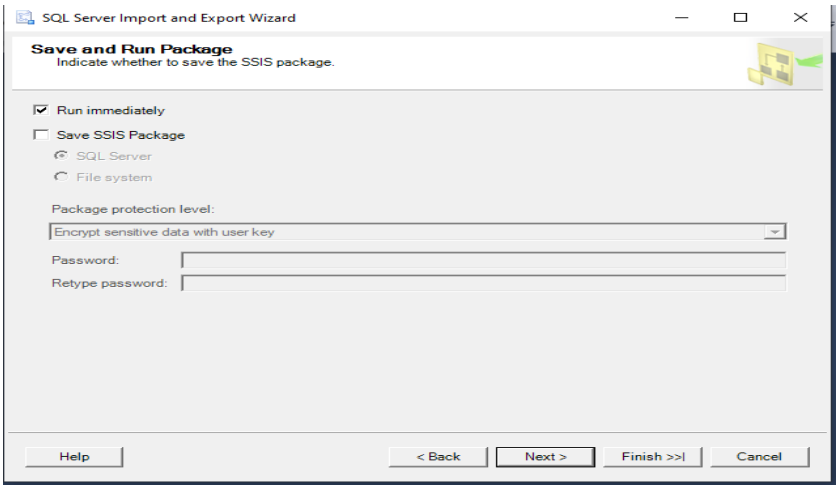
f) Select Microsoft OLE DB Provider for SQL Server and click next.



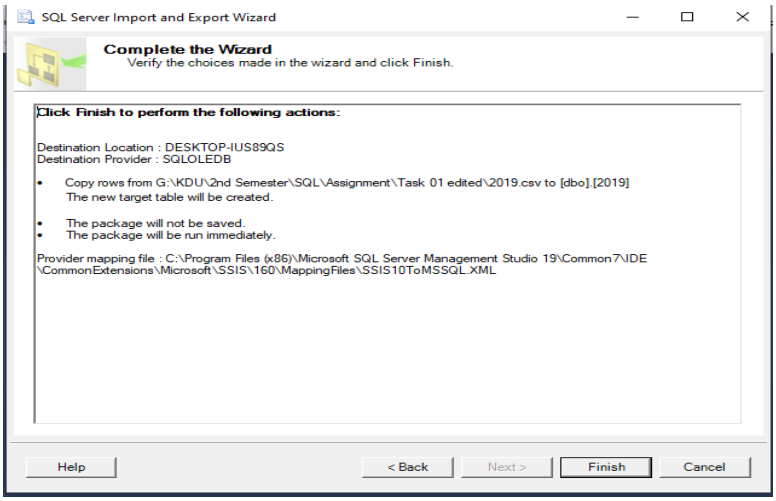




g) Accept the default and click next.

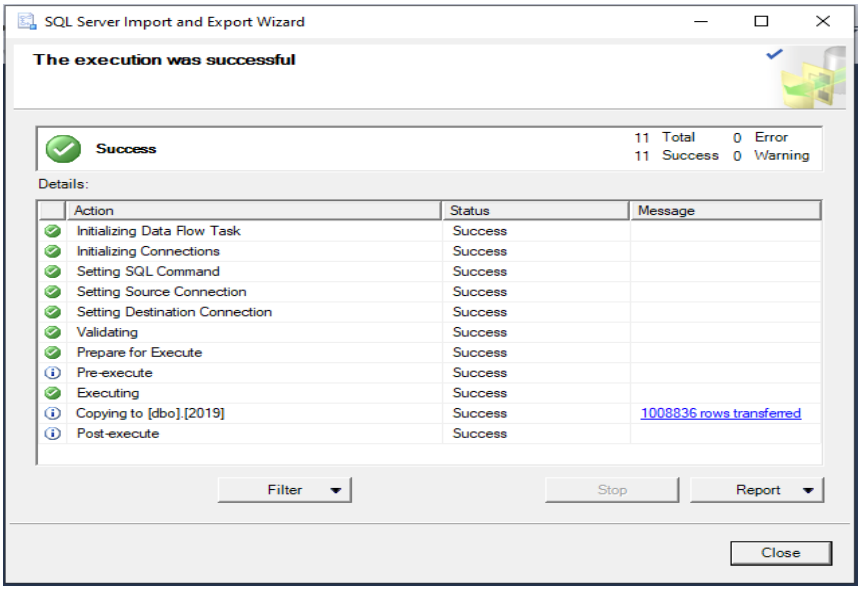




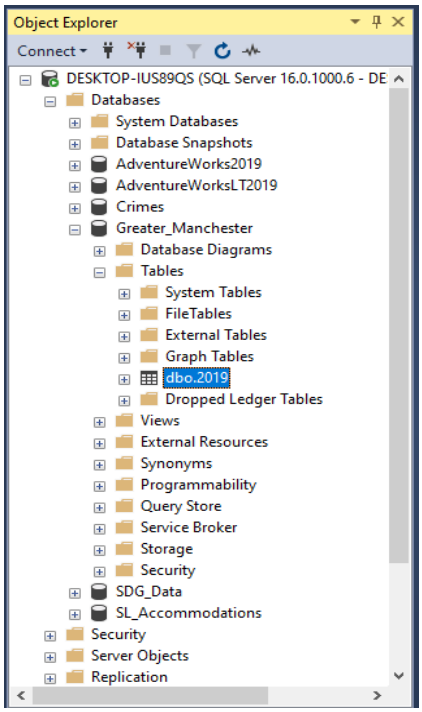

h) Click Finish.




i) The Execution Results dialog box appears. Assuming that all went well, the data has loaded successfully.



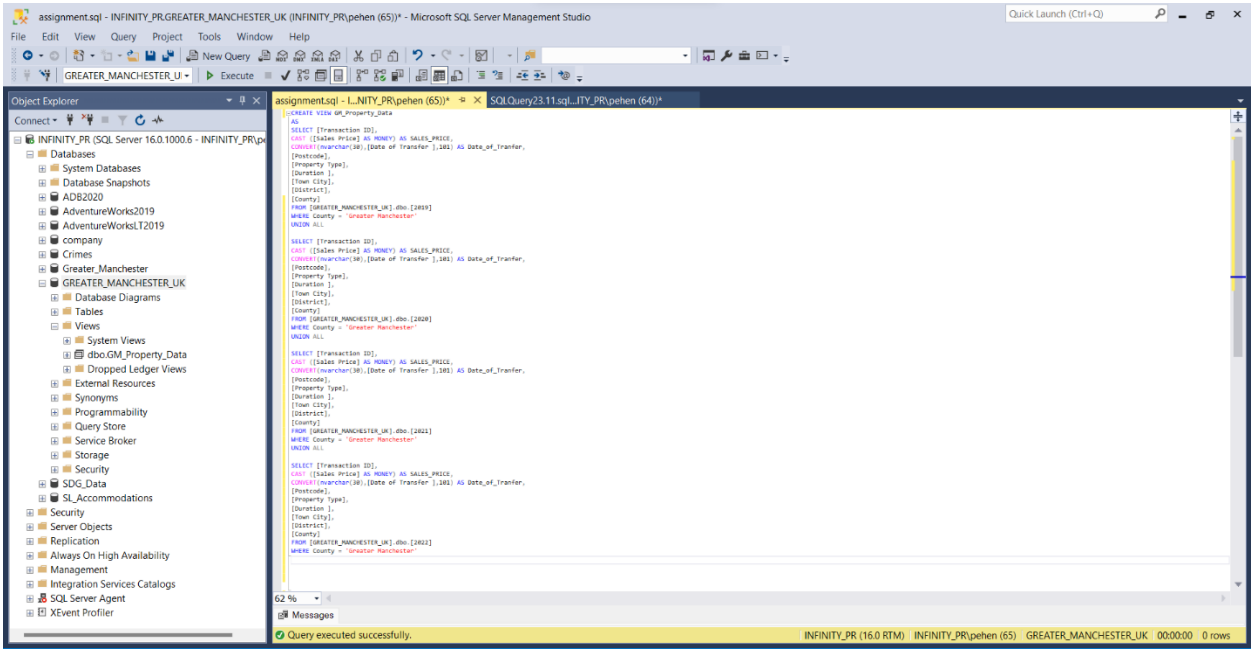

After we can see the new table in the 'GREATER_MANCHESTER_UK' database.



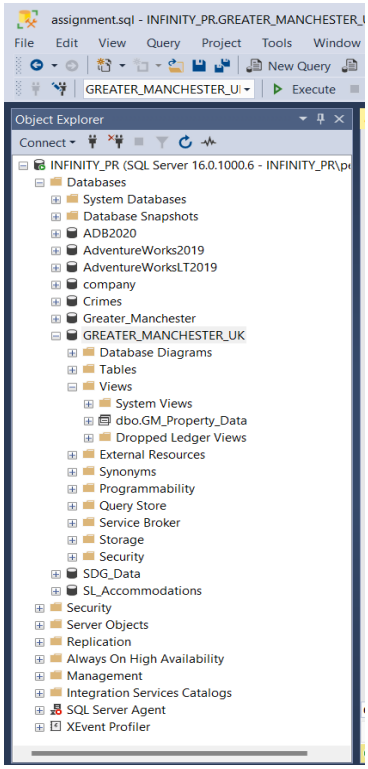

You must repeat the same procedure for three times to add the 2020, 2021, 2022 CSV files to the 'GREATER_MANCHESTER_UK' database.

## Create View

Under a new query, enter the following code to create view ‘GM_Property_Data’ by attaching all four tables.



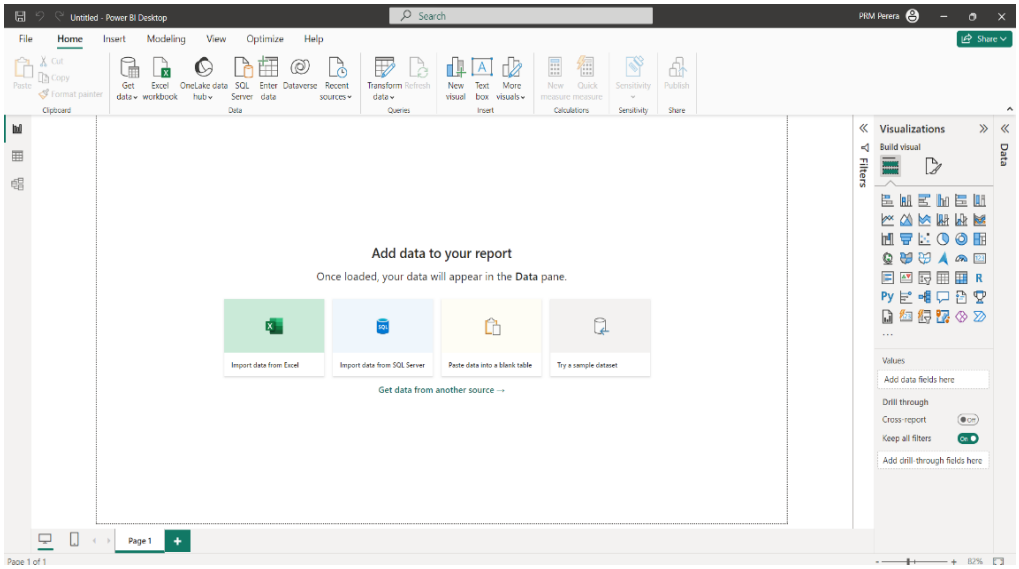
You can see the ‘dbo.GM_Property_Data’ view under the views, in ‘GREATER_MANCHESTER_UK’database.



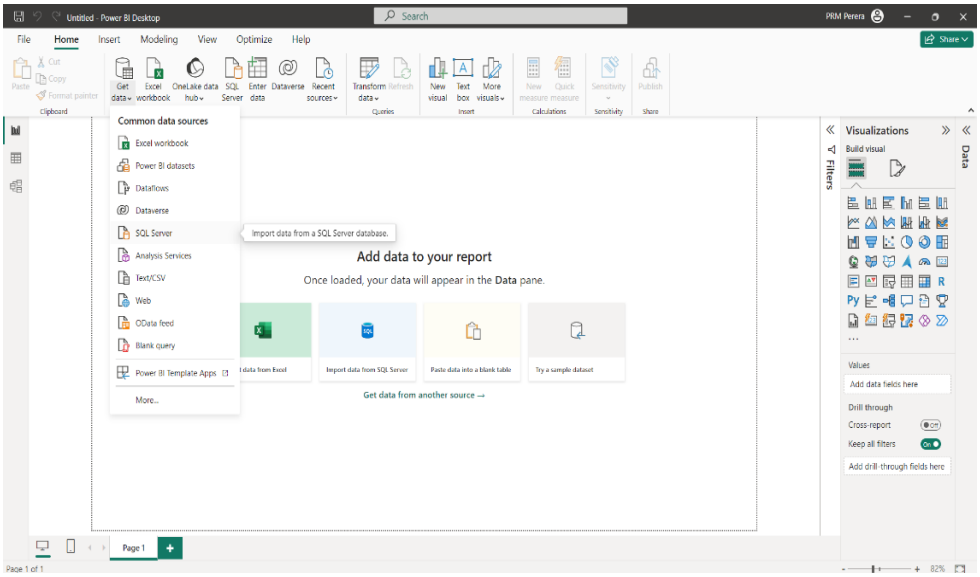
# Dashboard design and Implementation

## Bring Data into Power BI

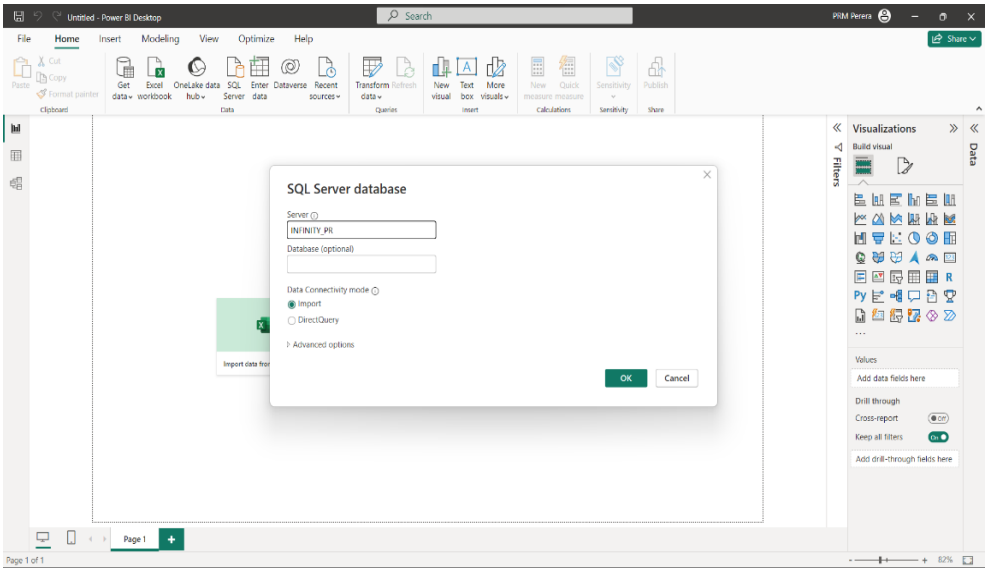
1. Open Power Bi desktop



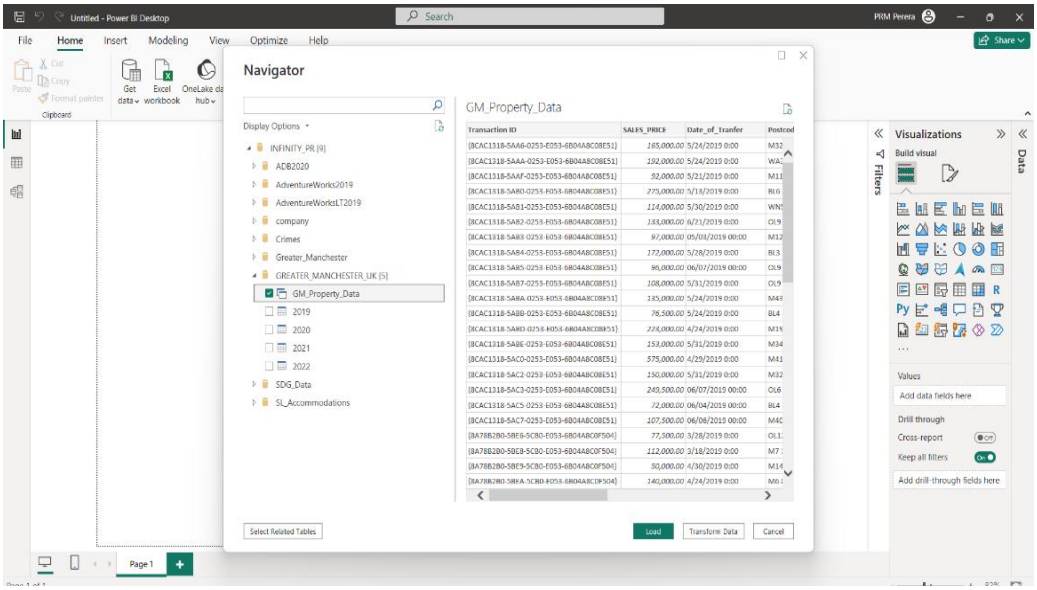

2. In Home tab, select Get Data → SQL Server




3. From SQL Server Database Server connect to your server. Then select Import and OK.




4. In Navigator tab, select ‘GM_Property_Data’ view from ‘GREATER_MANCHESTER_UK’ Database.



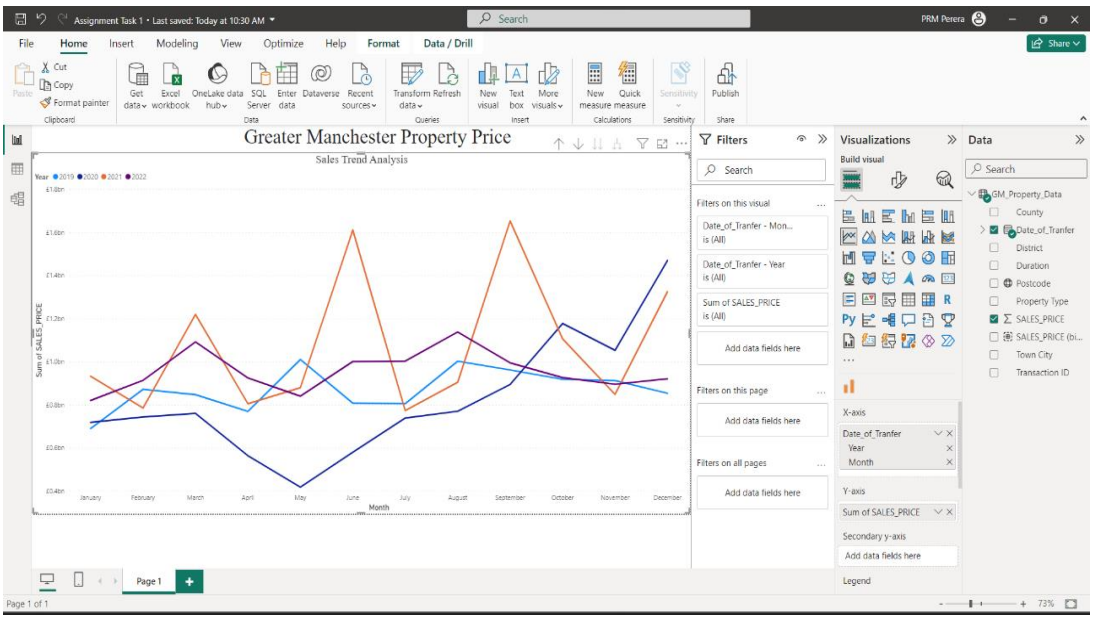
**Dashboard draw up**

1. Creating the Line chart.

First go to visualization & select line chart.

Then, go to values and add column ‘Date_of_transfer’ into X-axis. Change the name into ‘Date Hierarchy’. Add column ‘Sales_Price’ into Y-axis and change the name into ‘Sum of sales price’. Add legends as ‘Date of Transfer’ & ‘Date Hierarchy’.

The line chart views as follows.



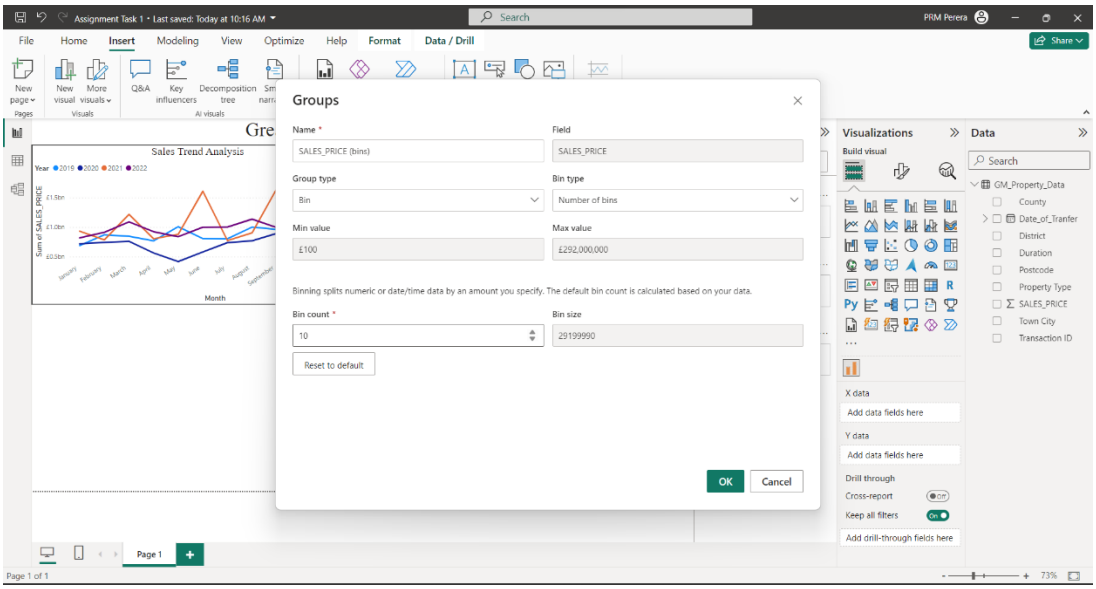

The above line chart shows monthly property sales over time in Greater Manchester.

2. Creating the Histogram.

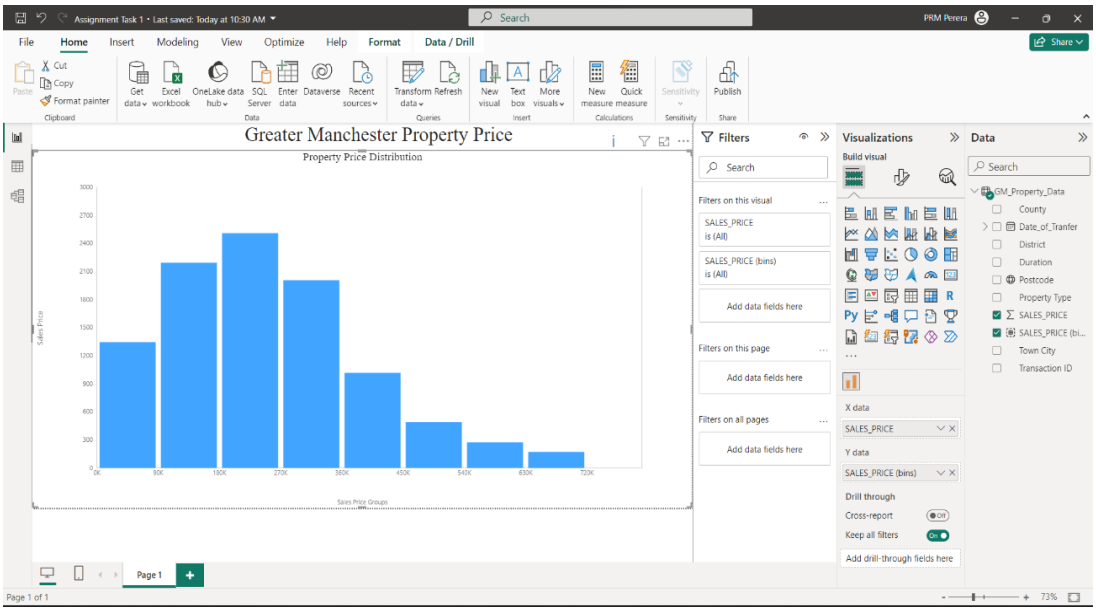
First go to visualizations & select a histogram.

Right click on ‘Sales_Price’ and select New Group.

Change the values as follows and select OK.




The histogram views as follows.




The above histogram shows the distribution of property prices in Greater Manchester.

3. <u>Creating the Map.</u>

   First go to visualizations & select a map.

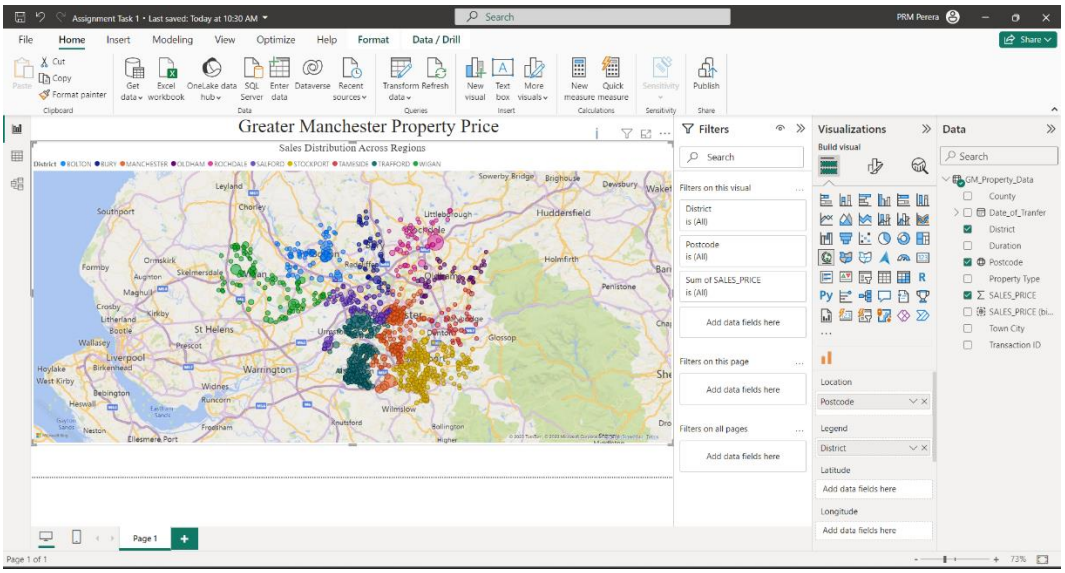
   Then, go to values and add ‘Postcode’ column to the visual. Then, change the data category into ‘Postal code’.

   Select ‘Postcode’ from Location.

Select ‘District’ from Legend.

Select ‘Sum_of sales price’ from Bubble size.

The map view as follows.



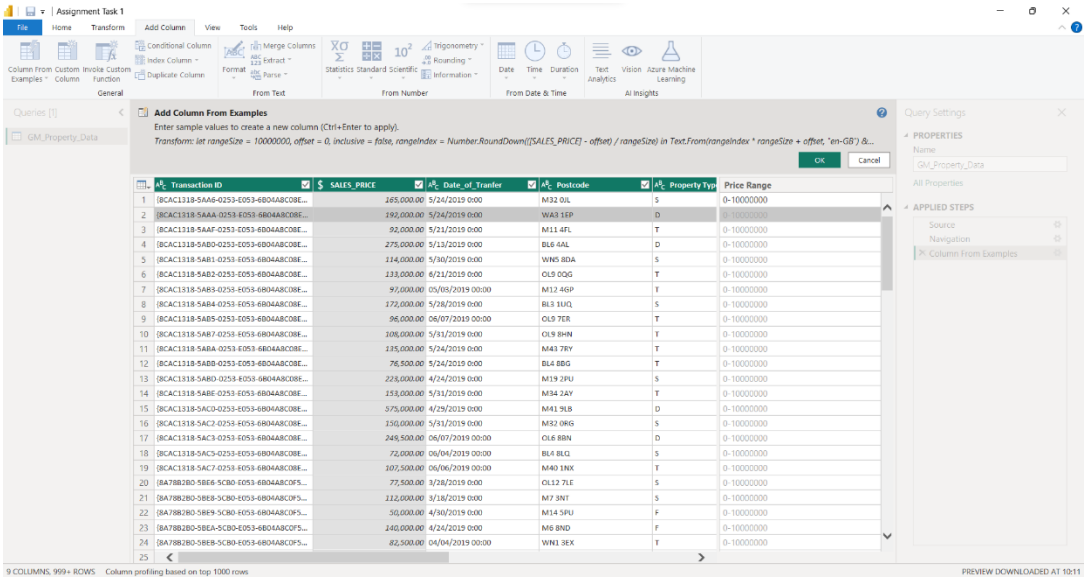

This map shows the sales distribution across all 10 regions.

4. <u>Creating the Column chart.</u>

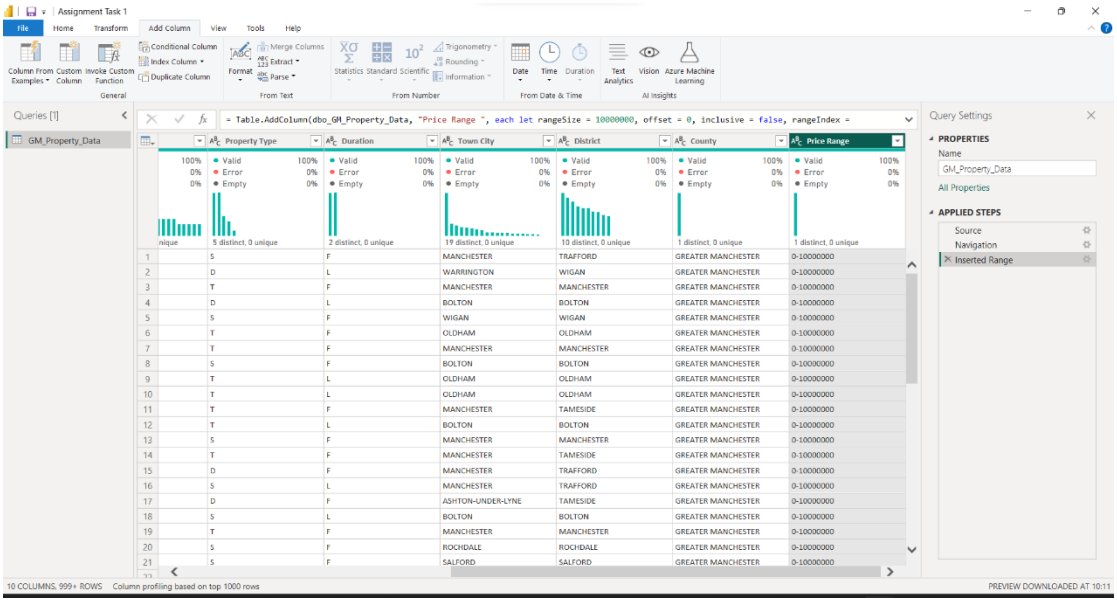
First, follow these steps to range the price.

Transform data → Add column → Column from Examples

Add a column named ‘Price Range’ and insert value 0-10000000 as the range.

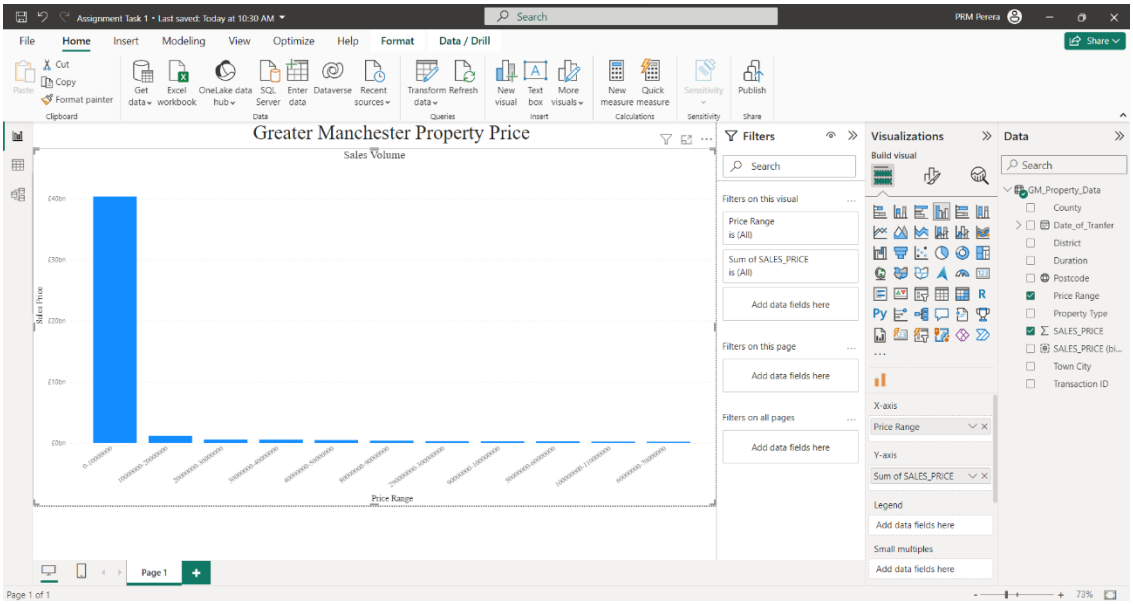


Then close the transform tab and select Apply.



Then go to visualizations & select clustered column chart. Select ‘Price Range’ as X-axis and select ‘Sum of SALES_PRICE’ as Y-axis.

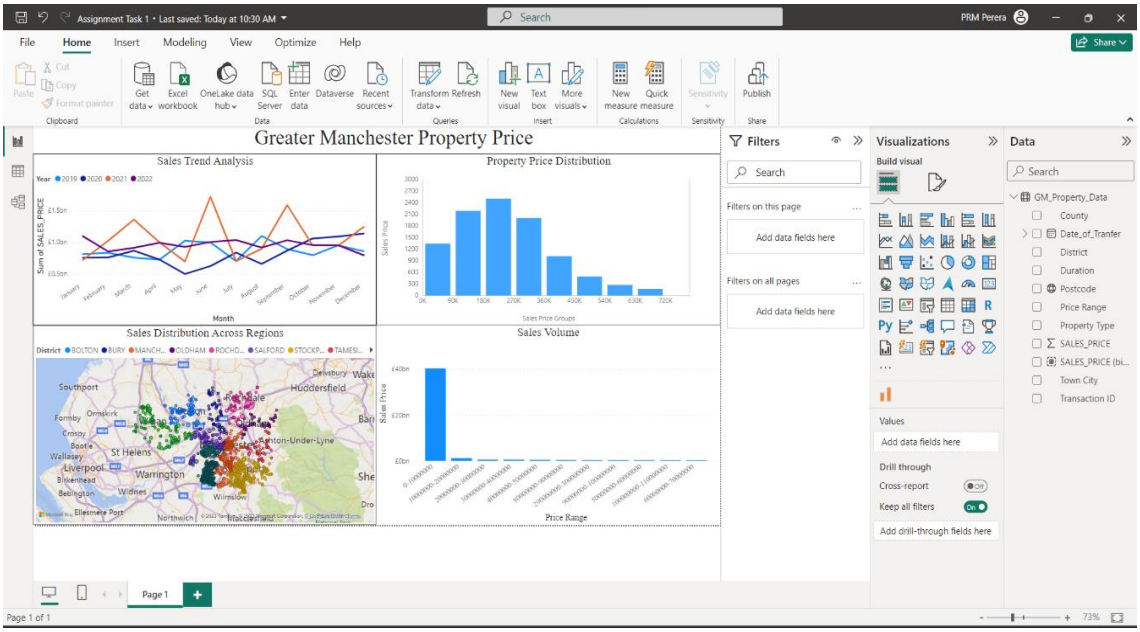
The column chart views as follows.



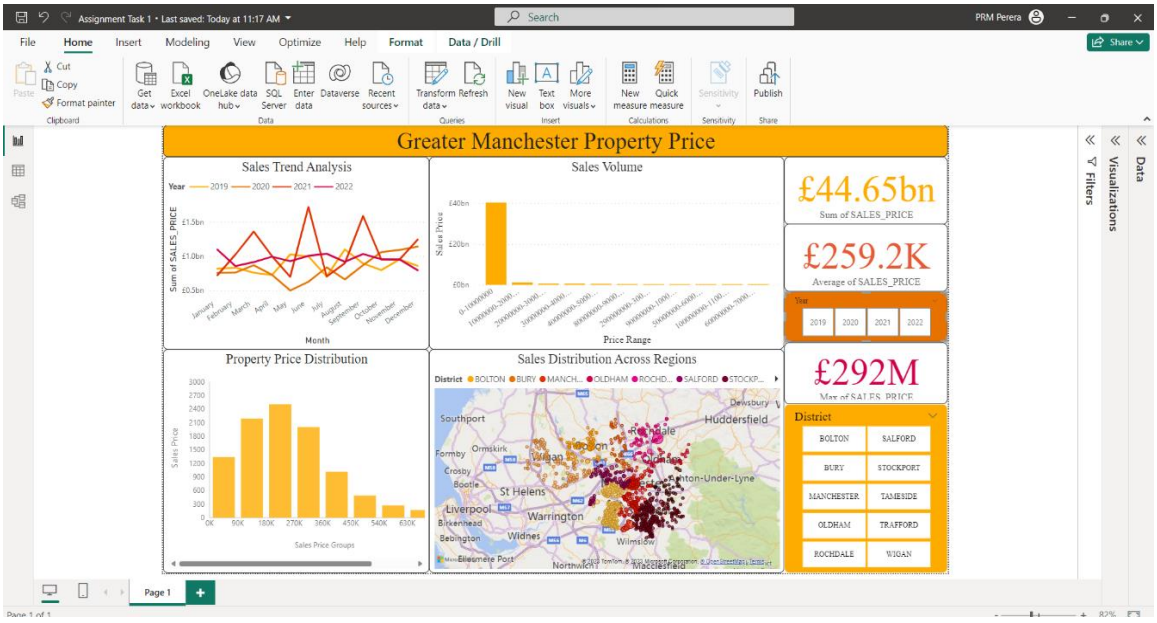

The above column chart represents the number of sales in different price brackets.

## Dashboard Illustration

The above four charts combination shows as follows.




The final dashboard views as follows.



## Implications & Limitations

### Implications

The following are some of the implications of the property price paid dataset analysis:

1. Market Trends: Identify trends in property prices over time, providing insights into the real estate market's dynamics.
2. Investment judgments: Assist investors in making educated judgments by recognizing which industries have demonstrated consistent growth or have the potential for rewards.
3. Policy Development: The data can be used by the government and regulatory organizations to develop housing laws, zoning restrictions, and taxing plans.
4. Real Estate Development: Inform developers on high-demand areas, allowing for smart judgments on new development projects.
5. Urban Planning: Help with urban planning activities by emphasizing regions of growth or decline and directing infrastructure development.

### Limitations

The following are some of the limitations of the property price paid dataset analysis:

1. Inadequate Information: Certain characteristics, such as property status, renovations, or specific features, may be missing from the dataset, limiting the depth of research.
2. Time Factors: Economic conditions can influence property values, and the dataset may not capture all significant temporal elements influencing pricing variations.
3. Geographical Variability: Prices can vary greatly depending on location-specific factors such as local amenities or development plans that are not represented in the dataset.
4. Property Categories: The dataset may not distinguish between different types of properties (e.g., residential versus commercial), reducing the accuracy of insights for specific segments.
5. Data Accuracy: Errors or inconsistencies in the dataset may result in incorrect findings, either due to recording errors or outdated information.
6. Market Dynamics: Changes in housing demand, mortgage rates, or government policies may not be fully reflected, impacting the overall predictability.

## Recommendations

**For Property Investors:**

1. Diversification: Use the information to uncover diverse investment options across locations and property kinds, spreading risk and maximizing possible returns for property investors.
2. Trend Analysis: Use historical data to detect trends in property appreciation, which can assist investors in making informed decisions about when and where to invest.
3. Risk Mitigation: When assessing the risk associated with certain investments, consider the dataset, taking into account elements such as market volatility and economic indicators.
4. Due Diligence: To make well-informed investment decisions, conduct complete due diligence by combining dataset insights with on-the-ground research.

**For Policy Makers:**

1. Housing Affordability: Use the dataset to better understand housing affordability concerns, hence assisting in the development of policies that address the requirements of people of different income levels.
2. Urban Planning: Inform urban planning decisions by identifying high-demand locations, allowing for targeted infrastructure development and city planning.
3. Zoning restrictions: Based on the dataset, adjust zoning restrictions to account for changes in property demand and usage trends.
4. Taxation Policies: Design taxation policies that support prudent property development while discouraging speculative behaviors, resulting in a more balanced real estate market.

## Conclusion

This report provides information on property in the United Kingdom. The raw data was obtained from the UK Price Paid Dataset, which covers information about the UK property market from 2019 to 2022. Greater Manchester Property Price Monitor is a simple and low-cost Power BI-based reporting solution for analyzing the Greater Manchester area property market. It enables users to extract and show data in charts, tables, and other visualizations in order to uncover relevant information about various property types. The goal of these tools is to turn data into usable information in order to better understand the local real estate market. According to the dashboard results, the following information was obtained. The line chart of Sales Trend Analysis shows the maximum sales were in June 2021 when the minimum sales were in May 2020. The histogram shows positive skewness with a long right-tail depicts Property Price Distribution. The bar graph of Sales Volume shows maximum price range of properties is between 0-10000000. According to the dashboard, the sum of sales price £44.65bn, the average of sales price £25.2K and the maximum of sales price £292M.