General Sri John Kotelawala Defence University

Faculty Of Management, Social Sciences and Humanities

Department of Languages

Bcs in Applied Data Science Communication

Intake 40

A.G.A.U.S.Gunasekara - D/ADC/23/0034

Y.G.G.D. Ranasinghe - D/ADC/23/0025

M.L.B.T.S.Perera - D/ADC/23/0047

Year 02: Semester 04

Greater Manchester Domestic Energy Performance Analyzer

1) Introduction

2) Dataset Explanation

3) Power BI Dashboard

4) Conclusion

5) Recommendations

6) References

1) Introduction

The main objective of this assignment is to use Power BI to analyze and visualize Energy Performance Certificate (EPC) data for the local authorities in Manchester from 2013 to 2023. The objective is to develop an engaging and educational Power BI dashboard that offers significant insights into energy performance patterns over the past ten years to stakeholders in the energy industry. This research intends to demonstrate how EPC data may indicate areas for potential improvement while also revealing patterns in energy efficiency across different property types through data transformation, cleansing, and visualization.

The publicly accessible dataset for England and Wales contains extensive information on the energy efficiency of both business and residential buildings. Ranging from A (highest efficiency) to G (lowest efficiency), the data offers important insights into current energy use, emissions, and potential improvements that can improve energy efficiency.When making decisions to increase energy efficiency and lower carbon emissions, homeowners, landlords, property buyers, tenants, and government agencies can all benefit greatly from having access to EPC data.

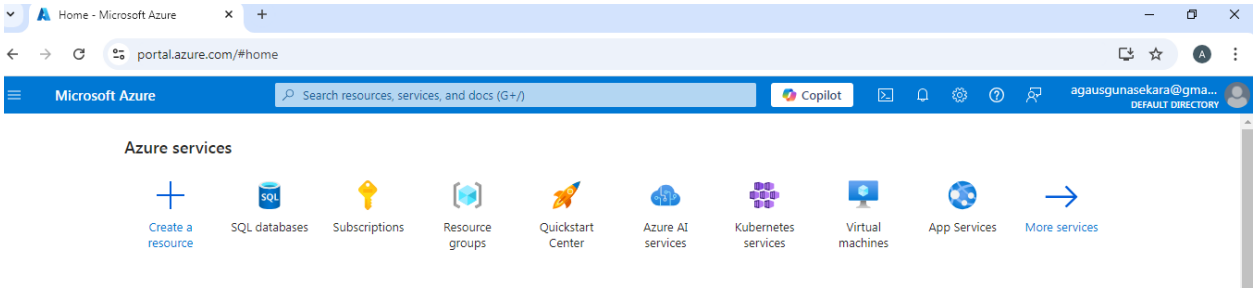
This project will produce a useful, user-friendly dashboard that assists the energy sector in meeting sustainability targets, identifying energy performance trends, and supporting long-term energy saving projects throughout Manchester by utilizing Power BI's sophisticated features.

2) Dataset Explanation

Creating a database in Azure cloud storage.

First you need to login to the Microsoft azure portal using your email.

In the Azure portal, select + Create a resource from the upper left-hand corner and search for Azure SQL. Then in the resulting Azure SQL page, select Create.

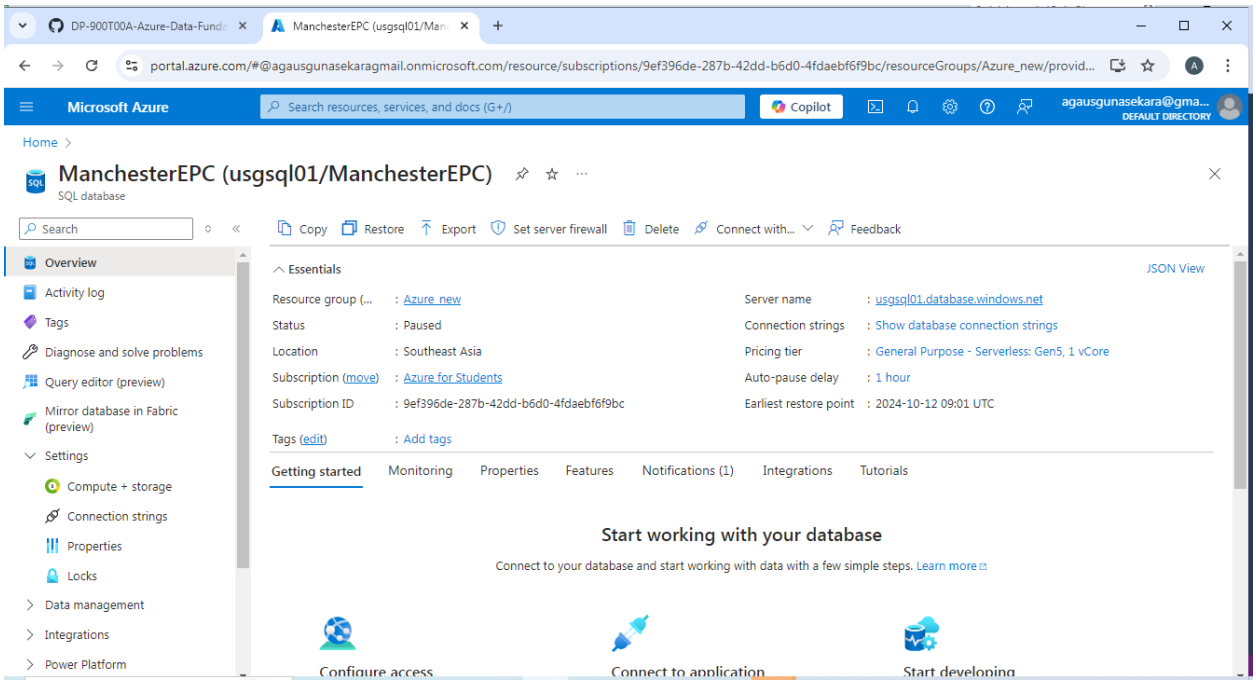


Enter the following values on the Create SQL Database page, and leave all other properties with their default setting:

- Subscription: Select your Azure subscription. (Azure for students)
- Resource group: Create a new resource group with a name of your choice. ([redacted])
- Database name: ManchersterEPC
- Server: Select Create new and create a new server with a unique name in any available location. (login : [redacted], password : [redacted])
- Want to use SQL elastic pool? : No
- Workload environment: Development
- Compute + storage: Leave unchanged
- Backup storage redundancy: Locally-redundant backup storage

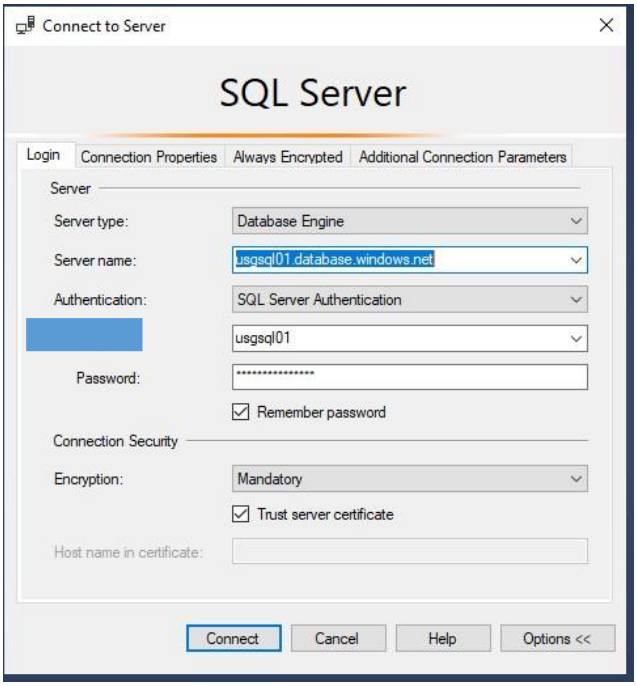
Select Review + Create, and then select Create to create your Azure SQL database.

Wait for process to complete. Then go to the resource that was deployed, which looks as follows.



Connecting the created Database to SQL Server Management Studio (SSMS).

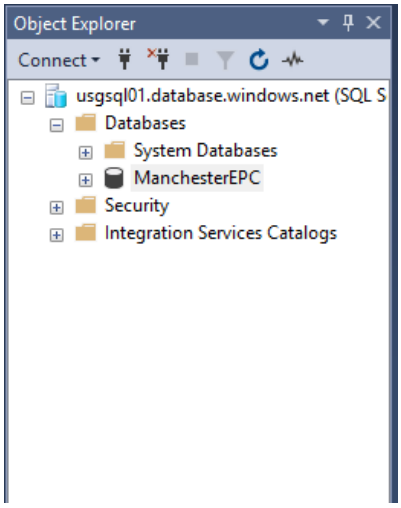
Open the SSMS and connect to server as follows.



Enter the server name correctly and select SQL Server Authentication in Authentication box.

Enter your login name and password there and connect to the server. (Make sure you have verify the trust server certificate box before connect.)

After connected to the server, you can see our created database “ManchesterEPC”.

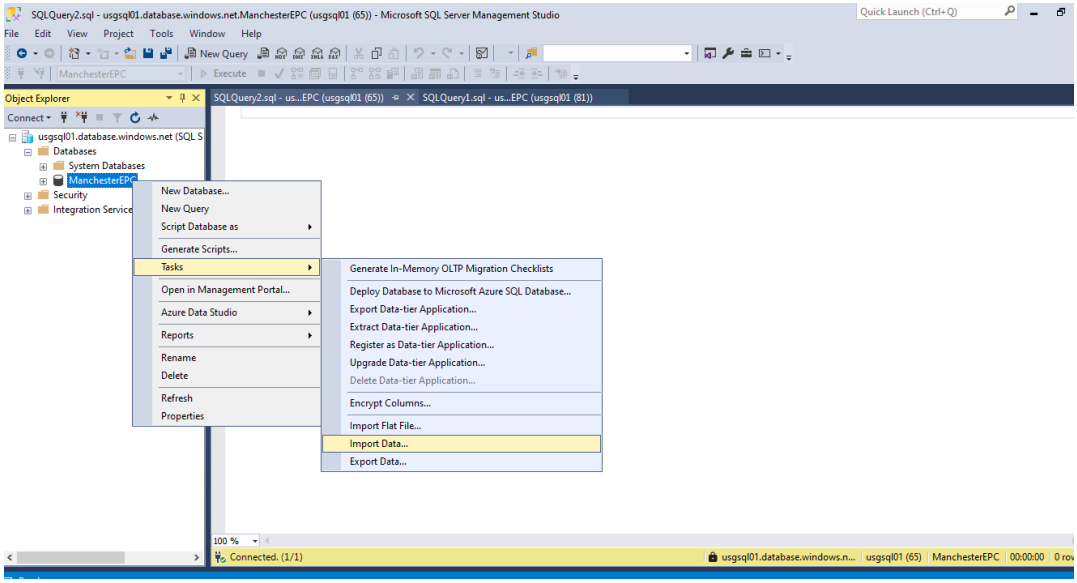


Importing Data from Excel to Microsoft SQL Server Database.

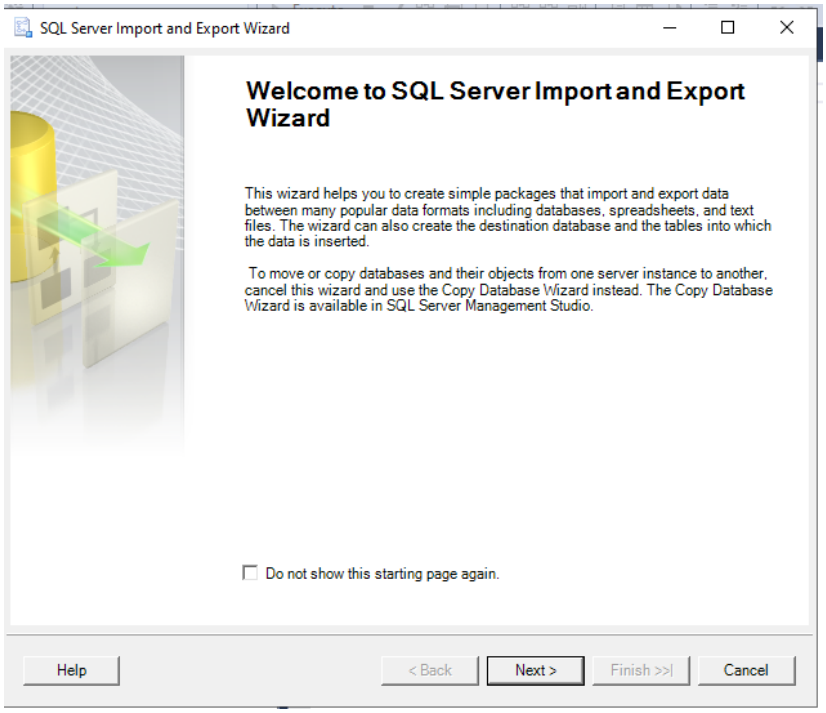
Import the CSV file to the SQL server database “ManchesterEPC”.

a) Right-click the ‘ManchesterEPC’ database.

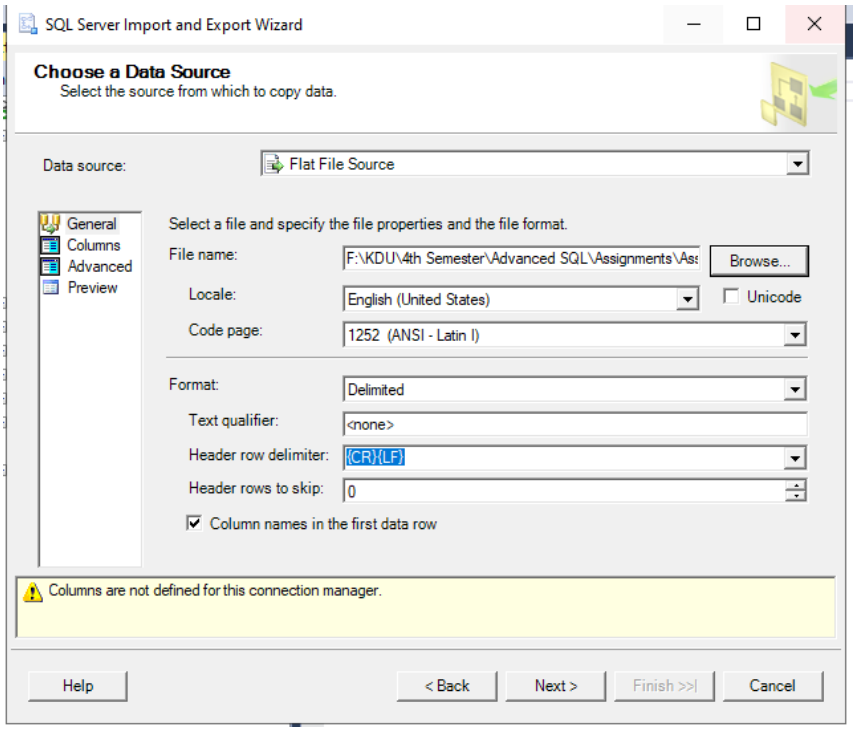
b) Select Tasks → Import Data



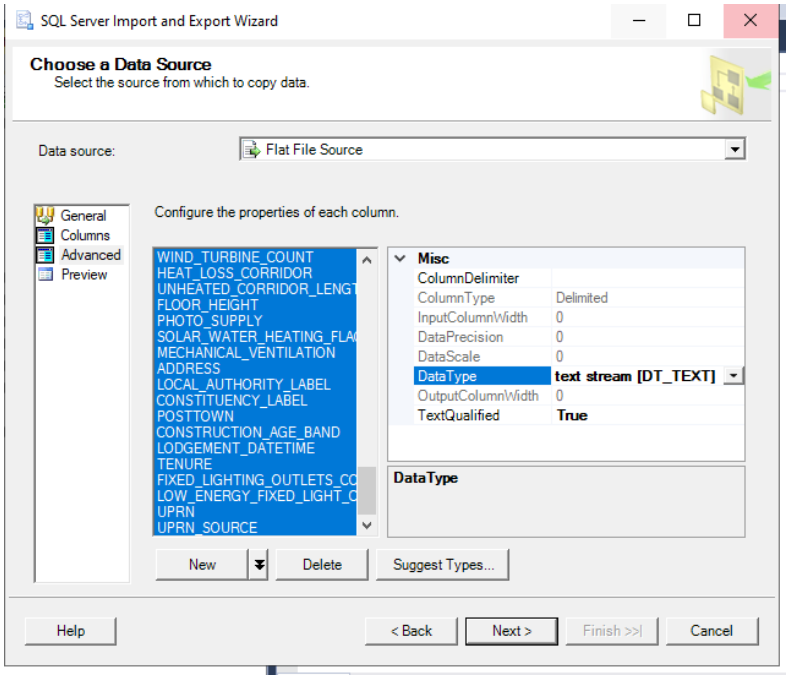
c) Click Next on the SQL Server Import and Export Wizard welcome page.



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



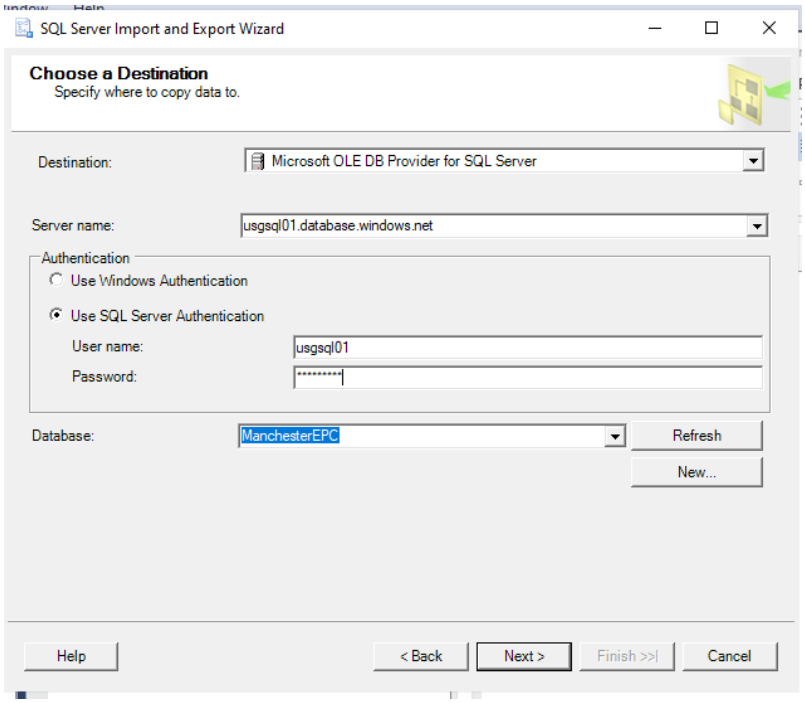
e) Go to the Advanced tab and change the datatype from string to text and click Next.



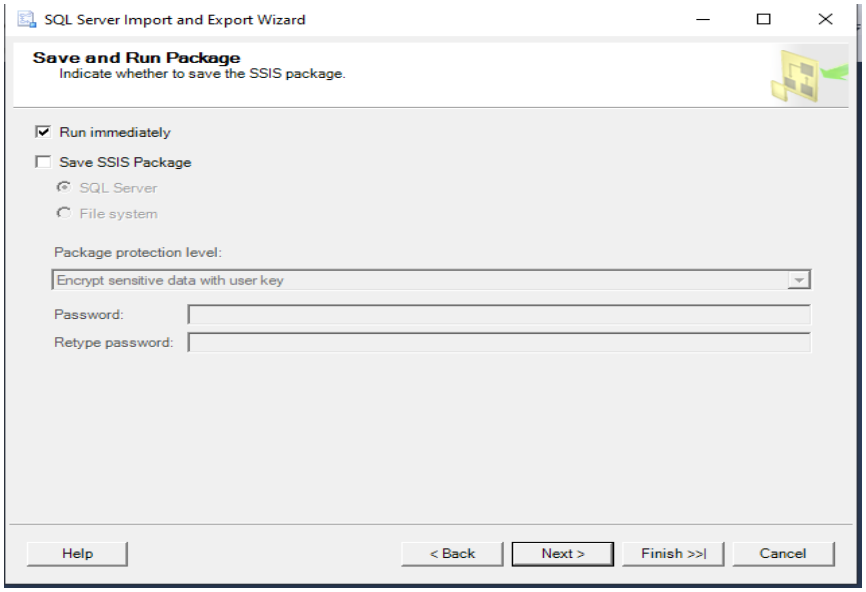
f) Select Microsoft OLE DB Provider for SQL Server and enter the server name there.

Select SQL Server Authentication and enter the username and password correctly.

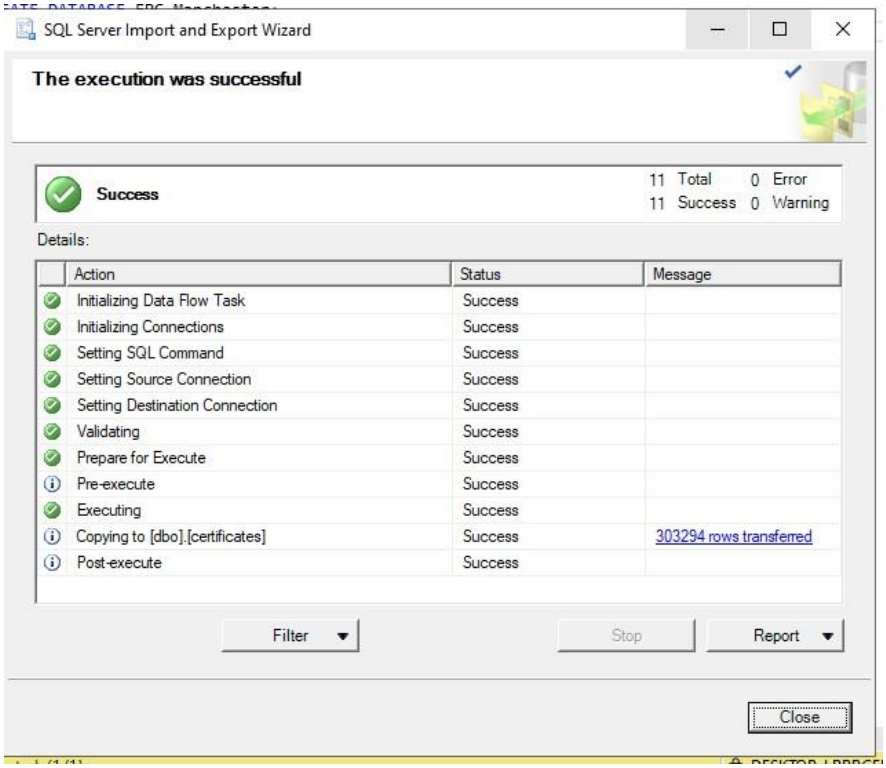
Then click next.



g) Accept the default and click next.

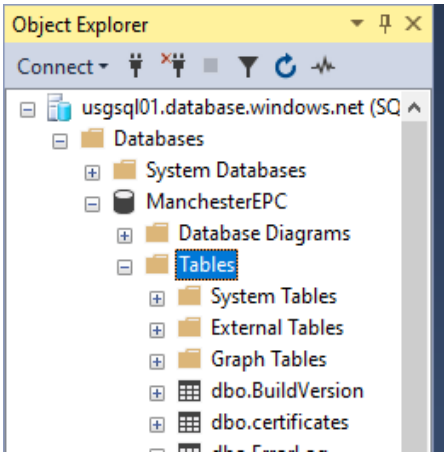


h) Click Finish.



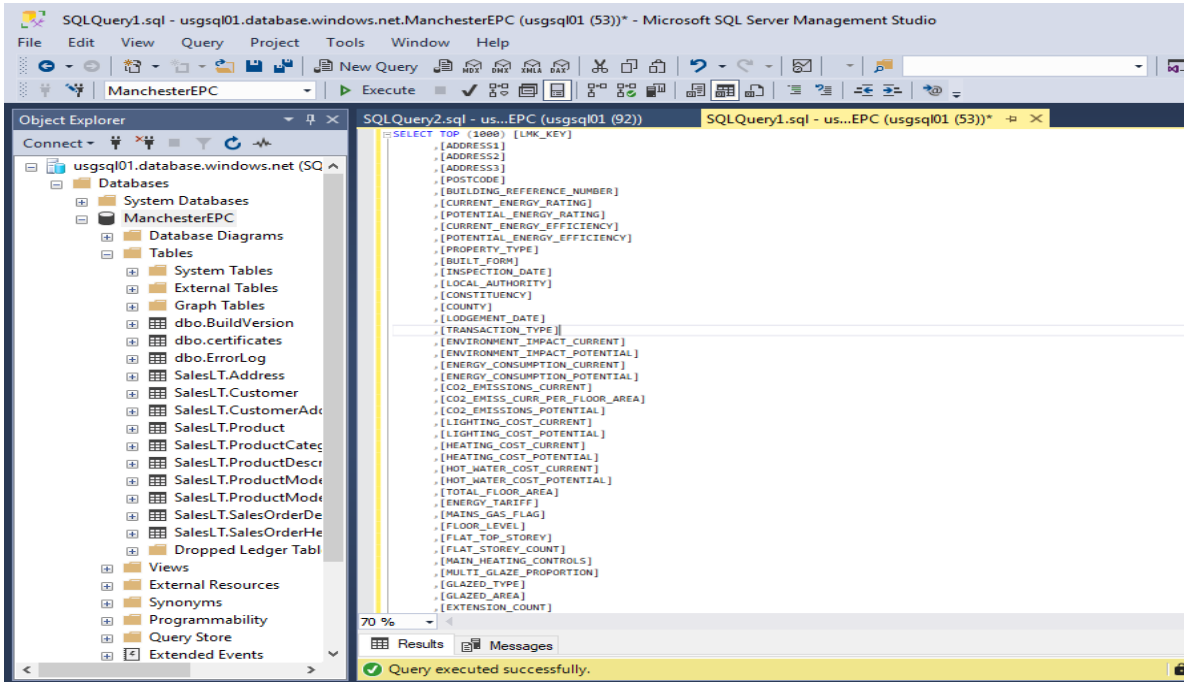
The Execution Results dialog box appears. Assuming that all went well, the data has loaded successfully.

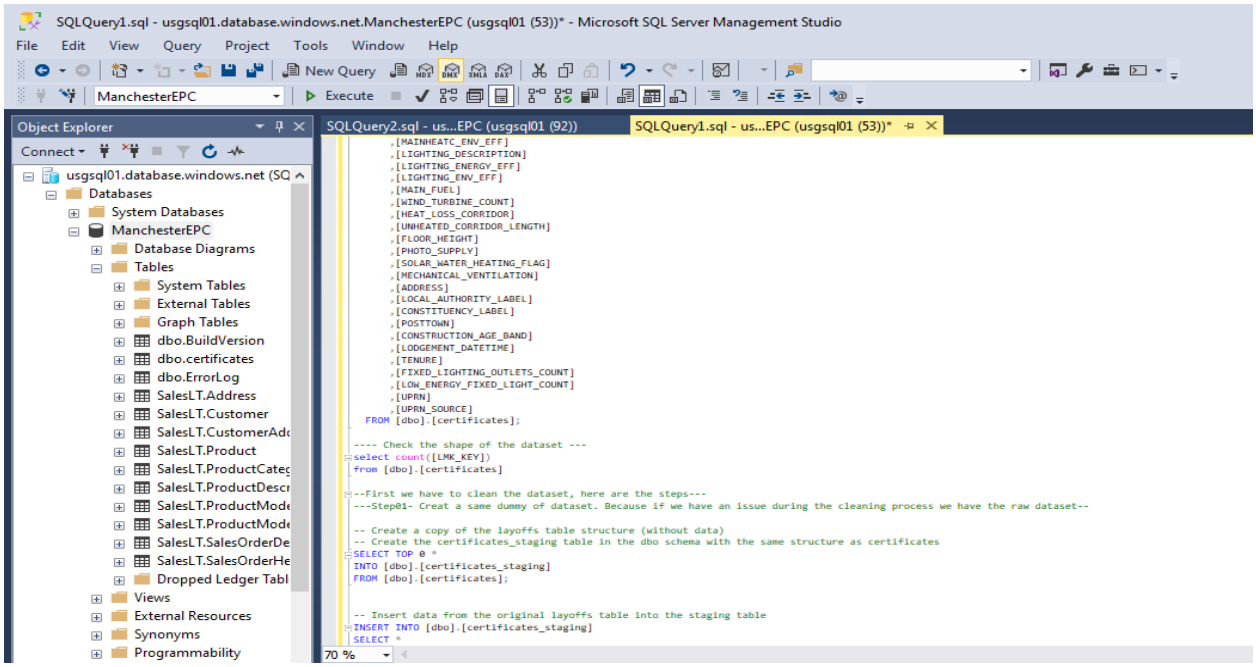
After we can see the new table ‘dbo.certificates’ in the ‘ManchesterEPC’ database.

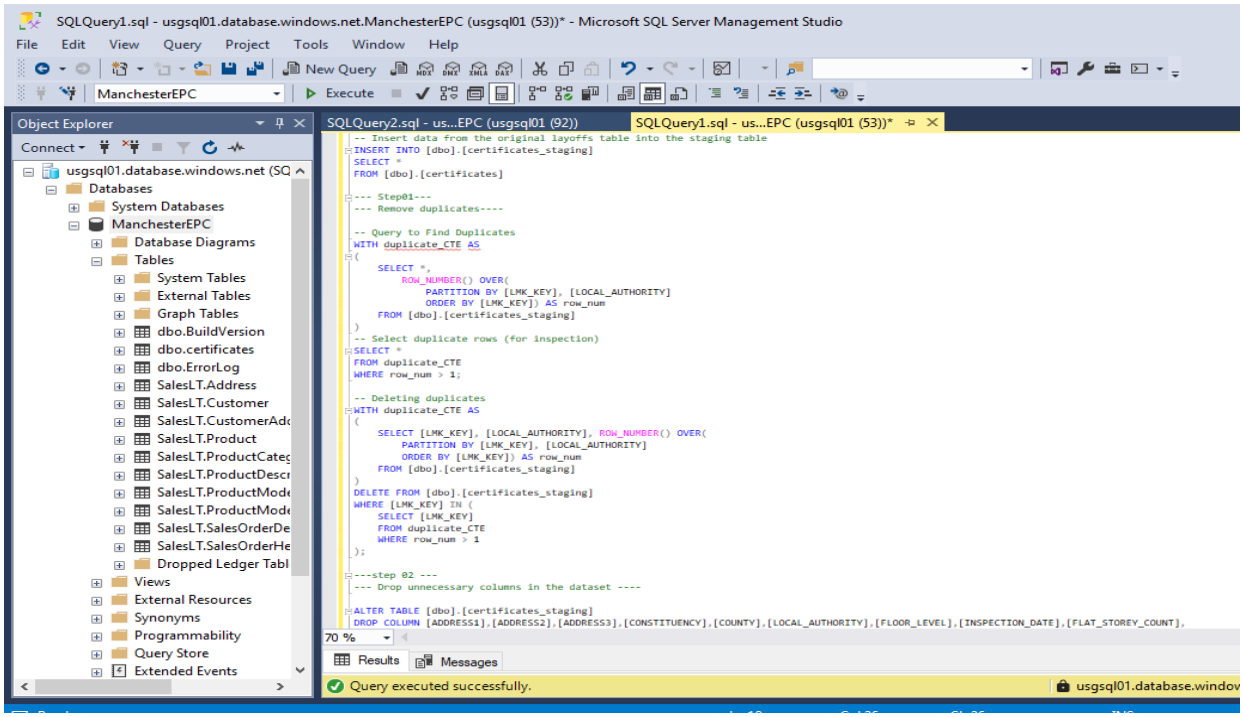


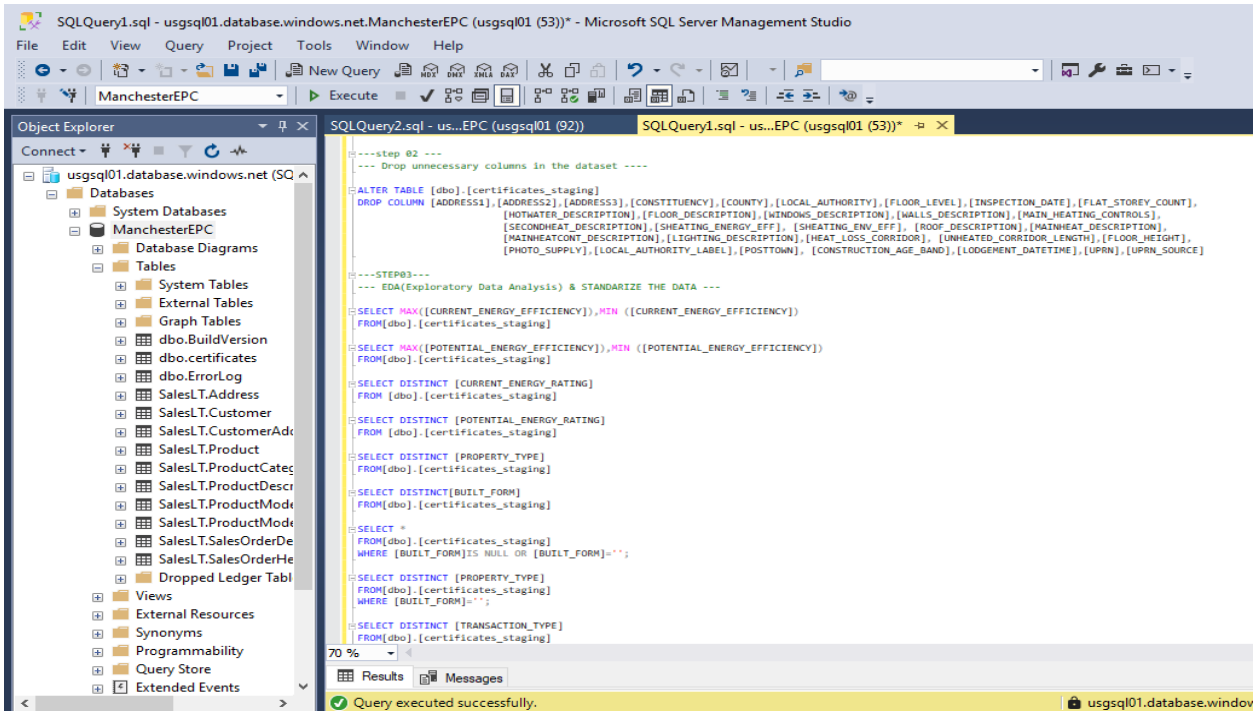
Cleaning the data

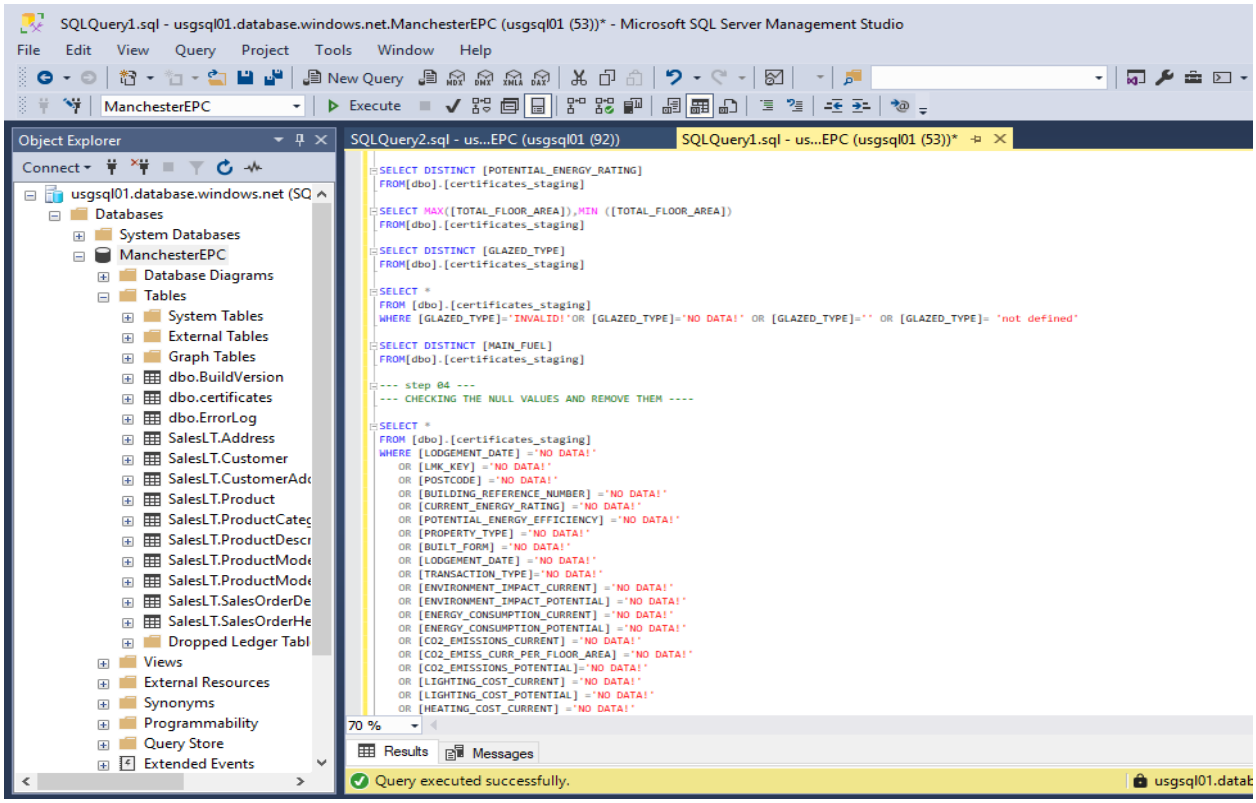
Under a new query, enter the following code to clean the ‘dbo.certificates’ dataset.

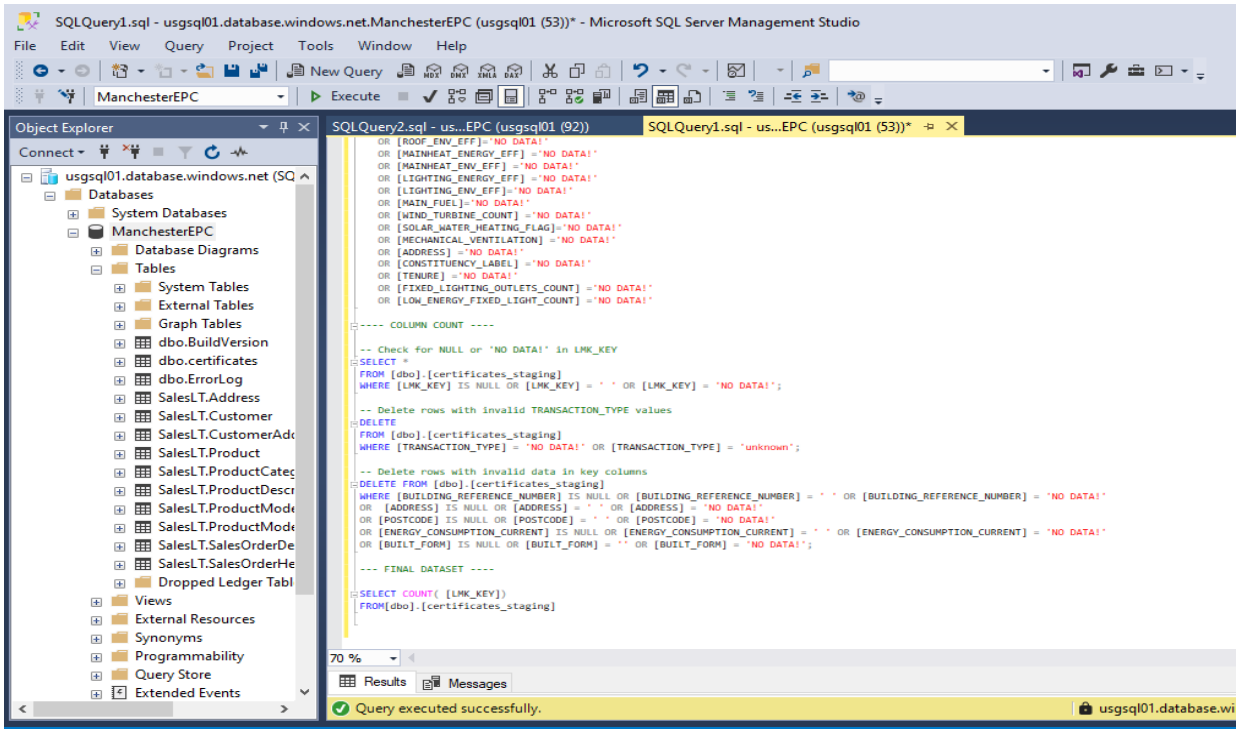





SQLQuery1.sql - usgsql01.database.windows.net.ManchesterEPC (usgsql01 (53))* - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
New Query
ManchesterEPC
Execute
Object Explorer
Connect
usgsql01.database.windows.net (SQ
Databases
System Databases
ManchesterEPC
Database Diagrams
Tables
System Tables
External Tables
Graph Tables
dbo.BuildVersion
dbo.certificates
dbo.ErrorLog
SalesLT.Address
SalesLT.Customer
SalesLT.CustomerAdd
SalesLT.Product
SalesLT.ProductCateg
SalesLT.ProductDescr
SalesLT.ProductMode
SalesLT.ProductMode
SalesLT.SalesOrderDe
SalesLT.SalesOrderHe
Dropped Ledger Tabl
Views
External Resources
Synonyms
Programmability
Query Store
Extended Events
SQLQuery2.sql - us...EPC (usgsql01 (92))
SQLQuery1.sql - us...EPC (usgsql01 (53))*
-- Insert data from the original layoffs table into the staging table
INSERT INTO [dbo].[certificates_staging]
SELECT *
FROM [dbo].[certificates]
--- Step01---
--- Remove duplicates----
-- Query to Find Duplicates
WITH duplicate_CTE AS
(
SELECT *,
ROW_NUMBER() OVER(
PARTITION BY [LMK_KEY], [LOCAL_AUTHORITY]
ORDER BY [LMK_KEY]) AS row_num
FROM [dbo].[certificates_staging]
)
-- Select duplicate rows (for inspection)
SELECT *
FROM duplicate_CTE
WHERE row_num > 1;
-- Deleting duplicates
WITH duplicate_CTE AS
(
SELECT [LMK_KEY], [LOCAL_AUTHORITY], ROW_NUMBER() OVER(
PARTITION BY [LMK_KEY], [LOCAL_AUTHORITY]
ORDER BY [LMK_KEY]) AS row_num
FROM [dbo].[certificates_staging]
)
DELETE FROM [dbo].[certificates_staging]
WHERE [LMK_KEY] IN (
SELECT [LMK_KEY]
FROM duplicate_CTE
WHERE row_num > 1
);
---step 02 ---
--- Drop unnecessary columns in the dataset ----
ALTER TABLE [dbo].[certificates_staging]
DROP COLUMN [ADDRESS1],[ADDRESS2],[ADDRESS3],[CONSTITUENCY],[COUNTY],[LOCAL_AUTHORITY],[FLOOR_LEVEL],[INSPECTION_DATE],[FLAT_STOREY_COUNT],
70 %
Results
Messages
Query executed successfully.
usgsql01.database.windov


SQLQuery1.sql - usgsql01.database.windows.net.ManchesterEPC (usgsql01 (53))* - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
New Query
ManchesterEPC
Execute
Object Explorer
Connect
usgsql01.database.windows.net (SQ
Databases
System Databases
ManchesterEPC
Database Diagrams
Tables
System Tables
External Tables
Graph Tables
dbo.BuildVersion
dbo.certificates
dbo.ErrorLog
SalesLT.Address
SalesLT.Customer
SalesLT.CustomerAdd
SalesLT.Product
SalesLT.ProductCateg
SalesLT.ProductDescr
SalesLT.ProductMode
SalesLT.ProductMode
SalesLT.SalesOrderDe
SalesLT.SalesOrderHe
Dropped Ledger Tabl
Views
External Resources
Synonyms
Programmability
Query Store
Extended Events
SQLQuery2.sql - us...EPC (usgsql01 (92))
SQLQuery1.sql - us...EPC (usgsql01 (53))*
---step 02 ---
--- Drop unnecessary columns in the dataset ----
ALTER TABLE [dbo].[certificates_staging]
DROP COLUMN [ADDRESS1],[ADDRESS2],[ADDRESS3],[CONSTITUENCY],[COUNTY],[LOCAL_AUTHORITY],[FLOOR_LEVEL],[INSPECTION_DATE],[FLAT_STOREY_COUNT],
[HOTWATER_DESCRIPTION],[FLOOR_DESCRIPTION],[WINDOWS_DESCRIPTION],[WALLS_DESCRIPTION],[MAIN_HEATING_CONTROLS],
[SECONDHEAT_DESCRIPTION],[SHEATING_ENERGY_EFF], [SHEATING_ENV_EFF], [ROOF_DESCRIPTION],[MAINHEAT_DESCRIPTION],
[MAINHEATCONT_DESCRIPTION],[LIGHTING_DESCRIPTION],[HEAT_LOSS_CORRIDOR], [UNHEATED_CORRIDOR_LENGTH],[FLOOR_HEIGHT],
[PHOTO_SUPPLY],[LOCAL_AUTHORITY_LABEL],[POSTTOWN], [CONSTRUCTION_AGE_BAND],[LODGEMENT_DATETIME],[UPRN],[UPRN_SOURCE]
---STEP03---
--- EDA(Exploratory Data Analysis) & STANDARIZE THE DATA ---
SELECT MAX([CURRENT_ENERGY_EFFICIENCY]),MIN ([CURRENT_ENERGY_EFFICIENCY])
FROM[dbo].[certificates_staging]
SELECT MAX([POTENTIAL_ENERGY_EFFICIENCY]),MIN ([POTENTIAL_ENERGY_EFFICIENCY])
FROM[dbo].[certificates_staging]
SELECT DISTINCT [CURRENT_ENERGY_RATING]
FROM [dbo].[certificates_staging]
SELECT DISTINCT [POTENTIAL_ENERGY_RATING]
FROM [dbo].[certificates_staging]
SELECT DISTINCT [PROPERTY_TYPE]
FROM[dbo].[certificates_staging]
SELECT DISTINCT[BUILT_FORM]
FROM[dbo].[certificates_staging]
SELECT *
FROM[dbo].[certificates_staging]
WHERE [BUILT_FORM]IS NULL OR [BUILT_FORM]='';
SELECT DISTINCT [PROPERTY_TYPE]
FROM[dbo].[certificates_staging]
WHERE [BUILT_FORM]='';
SELECT DISTINCT [TRANSACTION_TYPE]
FROM[dbo].[certificates_staging]
70 %
Results
Messages
Query executed successfully.
usgsql01.database.windov

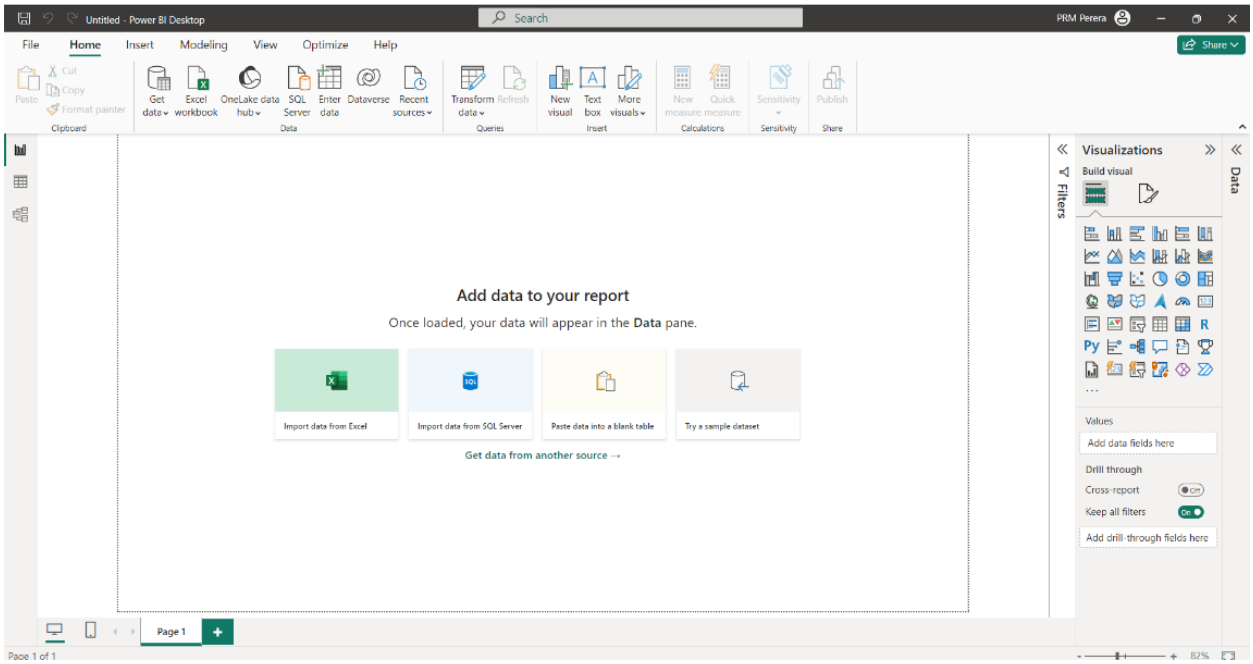

SQLQuery1.sql - usgsql01.database.windows.net.ManchesterEPC (usgsql01 (53))* - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
New Query
ManchesterEPC
Execute
Object Explorer
Connect
usgsql01.database.windows.net (SQ
Databases
System Databases
ManchesterEPC
Database Diagrams
Tables
System Tables
External Tables
Graph Tables
dbo.BuildVersion
dbo.certificates
dbo.ErrorLog
SalesLT.Address
SalesLT.Customer
SalesLT.Product
Dropped Ledger Tabl
Views
External Resources
Synonyms
Programmability
Query Store
Extended Events
SQLQuery2.sql - us...EPC (usgsql01 (92))
SQLQuery1.sql - us...EPC (usgsql01 (53))*
SELECT DISTINCT [POTENTIAL_ENERGY_RATING]
FROM[dbo].[certificates_staging]
SELECT MAX([TOTAL_FLOOR_AREA]),MIN ([TOTAL_FLOOR_AREA])
FROM[dbo].[certificates_staging]
SELECT DISTINCT [GLAZED_TYPE]
FROM[dbo].[certificates_staging]
SELECT *
FROM [dbo].[certificates_staging]
WHERE [GLAZED_TYPE]='INVALID!'OR [GLAZED_TYPE]='NO DATA!' OR [GLAZED_TYPE]='' OR [GLAZED_TYPE]= 'not defined'
SELECT DISTINCT [MAIN_FUEL]
FROM[dbo].[certificates_staging]
--- step 04 ---
--- CHECKING THE NULL VALUES AND REMOVE THEM ----
SELECT *
FROM [dbo].[certificates_staging]
WHERE [LODGEMENT_DATE] ='NO DATA!'
OR [LMK_KEY] ='NO DATA!'
OR [POSTCODE] ='NO DATA!'
OR [BUILDING_REFERENCE_NUMBER] ='NO DATA!'
OR [CURRENT_ENERGY_RATING] ='NO DATA!'
OR [POTENTIAL_ENERGY_EFFICIENCY] ='NO DATA!'
OR [PROPERTY_TYPE] ='NO DATA!'
OR [BUILT_FORM] ='NO DATA!'
OR [LODGEMENT_DATE] ='NO DATA!'
OR [TRANSACTION_TYPE]='NO DATA!'
OR [ENVIRONMENT_IMPACT_CURRENT] ='NO DATA!'
OR [ENVIRONMENT_IMPACT_POTENTIAL] ='NO DATA!'
OR [ENERGY_CONSUMPTION_CURRENT] ='NO DATA!'
OR [ENERGY_CONSUMPTION_POTENTIAL] ='NO DATA!'
OR [CO2_EMISSIONS_CURRENT] ='NO DATA!'
OR [CO2_EMISS_CURR_PER_FLOOR_AREA] ='NO DATA!'
OR [CO2_EMISSIONS_POTENTIAL]='NO DATA!'
OR [LIGHTING_COST_CURRENT] ='NO DATA!'
OR [LIGHTING_COST_POTENTIAL] ='NO DATA!'
OR [HEATING_COST_CURRENT] ='NO DATA!'
70 %
Results
Messages
Query executed successfully.
usgsql01.datab


SQLQuery1.sql - usgsql01.database.windows.net.ManchesterEPC (usgsql01 (53))* - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
New Query
ManchesterEPC
Execute
Object Explorer
Connect
usgsql01.database.windows.net (SQ
Databases
System Databases
ManchesterEPC
Database Diagrams
Tables
System Tables
External Tables
Graph Tables
dbo.BuildVersion
dbo.certificates
dbo.ErrorLog
SalesLT.Address
SalesLT.Customer
SalesLT.Product
Dropped Ledger Tabl
Views
External Resources
Synonyms
Programmability
Query Store
Extended Events
SQLQuery2.sql - us...EPC (usgsql01 (92))
SQLQuery1.sql - us...EPC (usgsql01 (53))*
OR [ROOF_ENV_EFF]='NO DATA!'
OR [MAINHEAT_ENERGY_EFF] ='NO DATA!'
OR [MAINHEAT_ENV_EFF] ='NO DATA!'
OR [LIGHTING_ENERGY_EFF] ='NO DATA!'
OR [LIGHTING_ENV_EFF]='NO DATA!'
OR [MAIN_FUEL]='NO DATA!'
OR [WIND_TURBINE_COUNT] ='NO DATA!'
OR [SOLAR_WATER_HEATING_FLAG]='NO DATA!'
OR [MECHANICAL_VENTILATION] ='NO DATA!'
OR [ADDRESS] ='NO DATA!'
OR [CONSTITUENCY_LABEL] ='NO DATA!'
OR [TENURE] ='NO DATA!'
OR [FIXED_LIGHTING_OUTLETS_COUNT] ='NO DATA!'
OR [LOW_ENERGY_FIXED_LIGHT_COUNT] ='NO DATA!'
---- COLUMN COUNT ----
-- Check for NULL or 'NO DATA!' in LMK_KEY
SELECT *
FROM [dbo].[certificates_staging]
WHERE [LMK_KEY] IS NULL OR [LMK_KEY] = ' ' OR [LMK_KEY] = 'NO DATA!';
-- Delete rows with invalid TRANSACTION_TYPE values
DELETE
FROM [dbo].[certificates_staging]
WHERE [TRANSACTION_TYPE] = 'NO DATA!' OR [TRANSACTION_TYPE] = 'unknown';
-- Delete rows with invalid data in key columns
DELETE FROM [dbo].[certificates_staging]
WHERE [BUILDING_REFERENCE_NUMBER] IS NULL OR [BUILDING_REFERENCE_NUMBER] = ' ' OR [BUILDING_REFERENCE_NUMBER] = 'NO DATA!'
OR [ADDRESS] IS NULL OR [ADDRESS] = ' ' OR [ADDRESS] = 'NO DATA!'
OR [POSTCODE] IS NULL OR [POSTCODE] = ' ' OR [POSTCODE] = 'NO DATA!'
OR [ENERGY_CONSUMPTION_CURRENT] IS NULL OR [ENERGY_CONSUMPTION_CURRENT] = ' ' OR [ENERGY_CONSUMPTION_CURRENT] = 'NO DATA!'
OR [BUILT_FORM] IS NULL OR [BUILT_FORM] = '' OR [BUILT_FORM] = 'NO DATA!';
--- FINAL DATASET ----
SELECT COUNT([LMK_KEY])
FROM[dbo].[certificates_staging]
70 %
Results
Messages
Query executed successfully.
usgsql01.database.wi

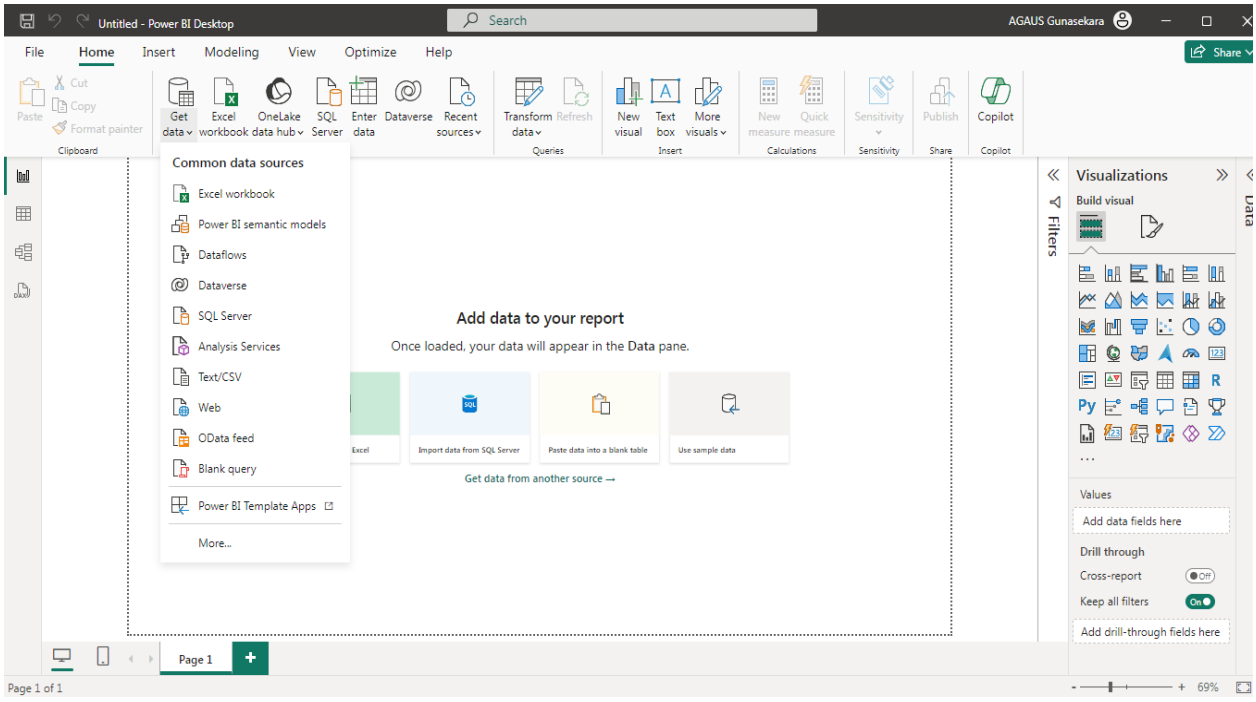
Dashboard design and Implementation

Bring Data into Power BI

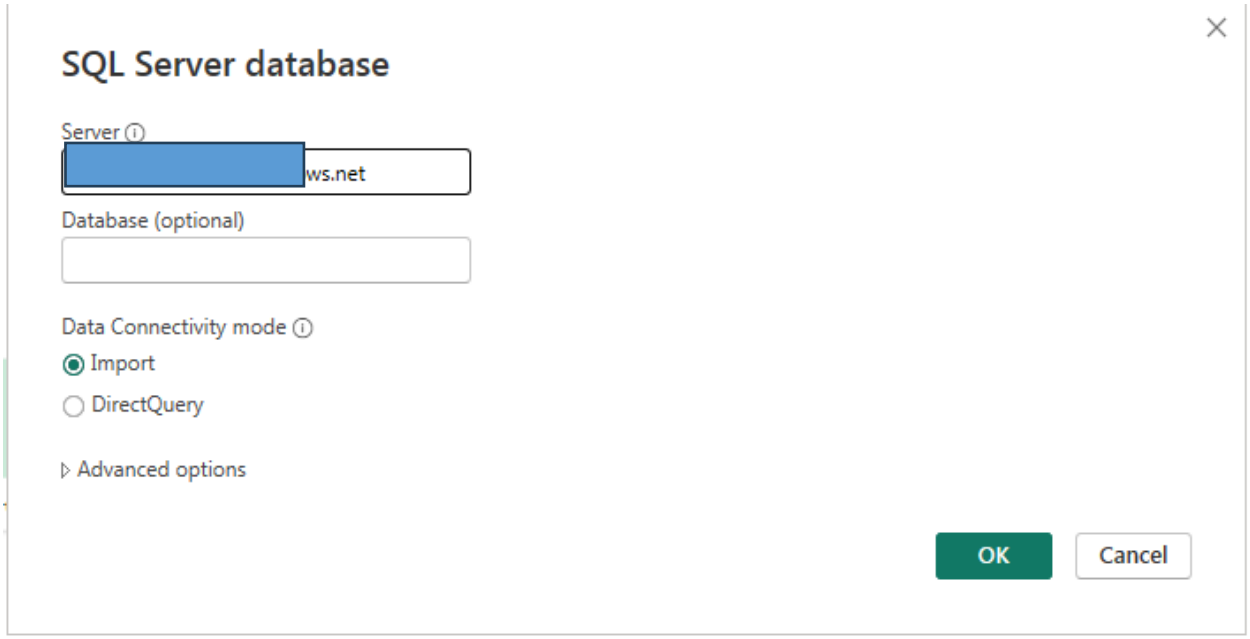
Open Power Bi desktop



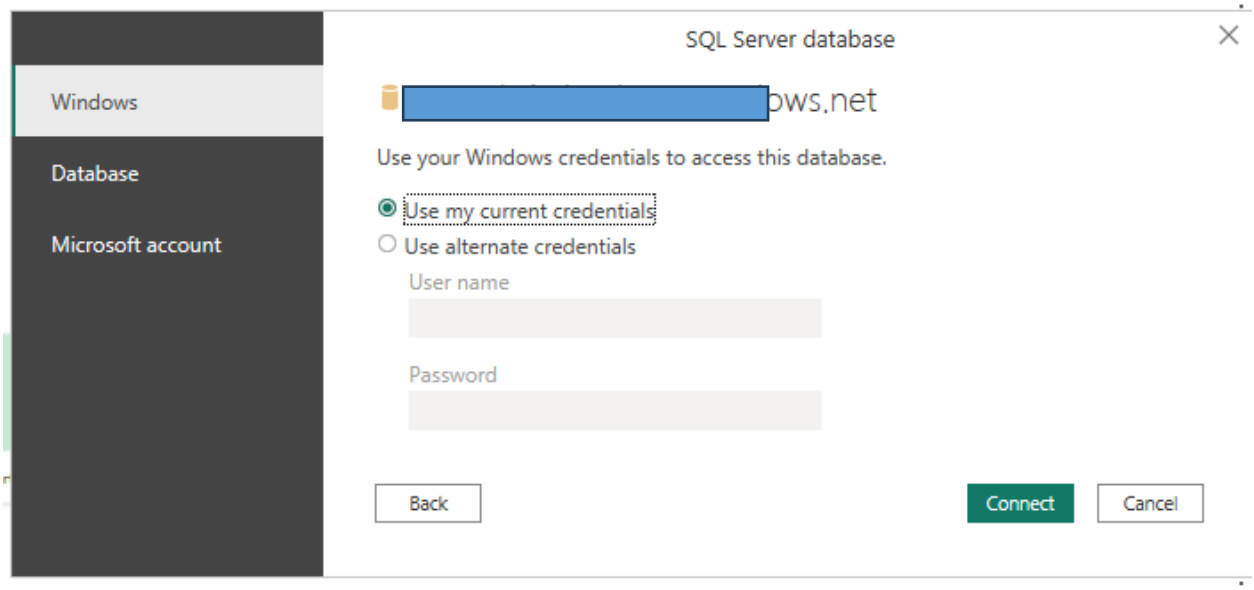
In Home tab, select Get Data → SQL Server



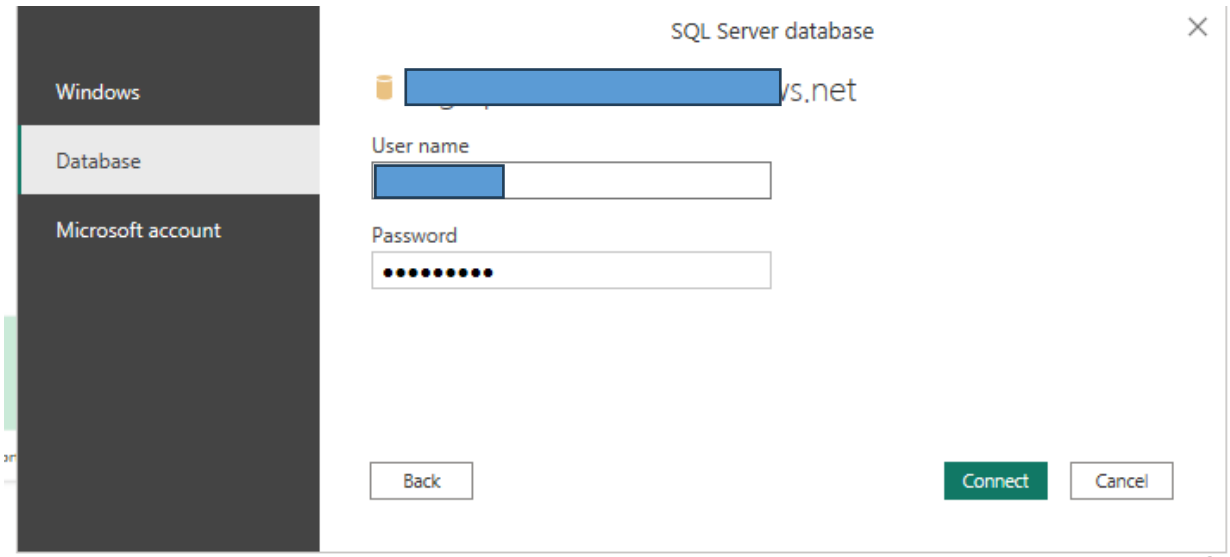
From SQL Server Database Server connect to your server. Enter the server name correctly in the dialogue box and click OK.



Enter the server username and password under the database section as follows.

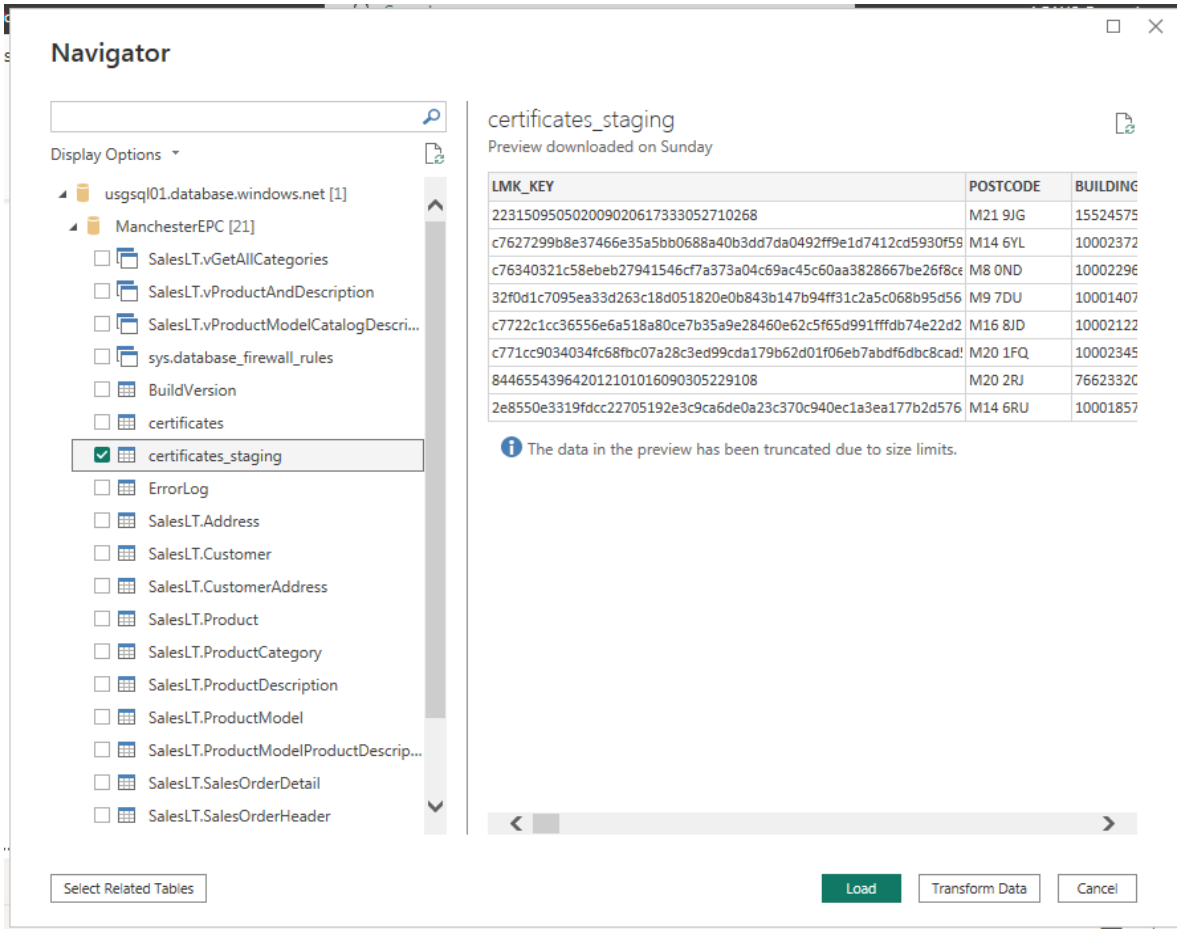


Then click on connect.

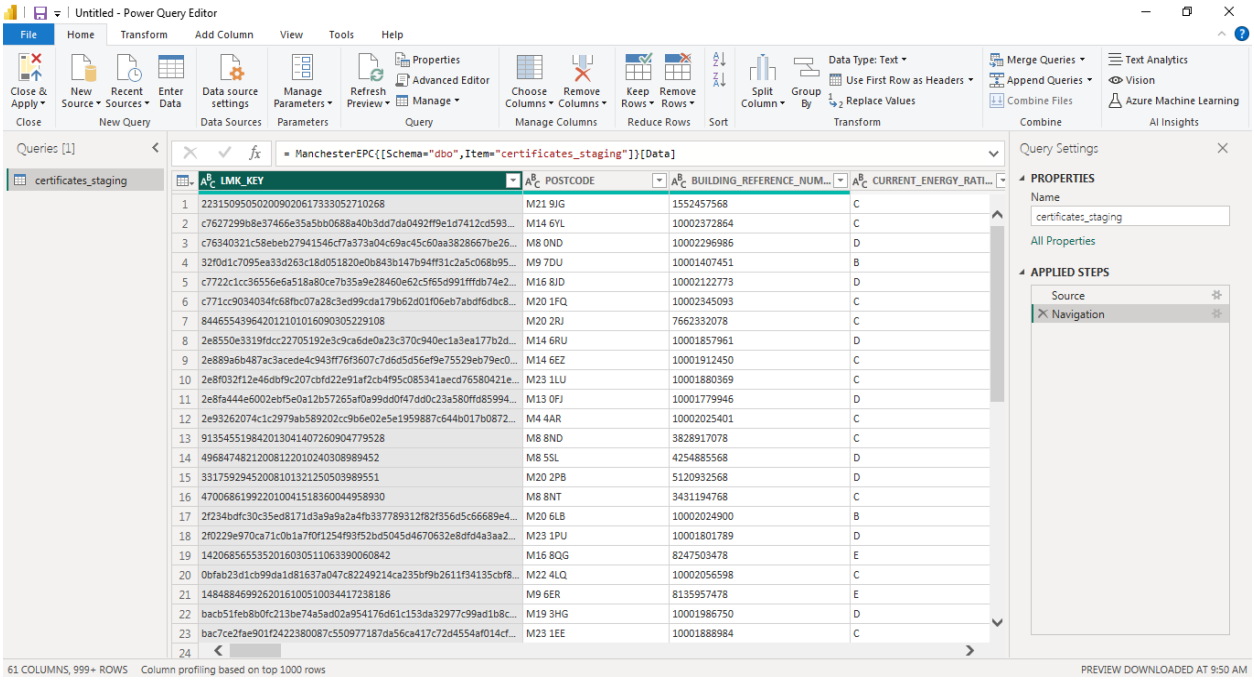


03. Power BI Dashboard

In Navigator tab, select 'certificates_staging' dataset from 'ManchesterEPC' Database which we have cleaned previously. Then select transform data.



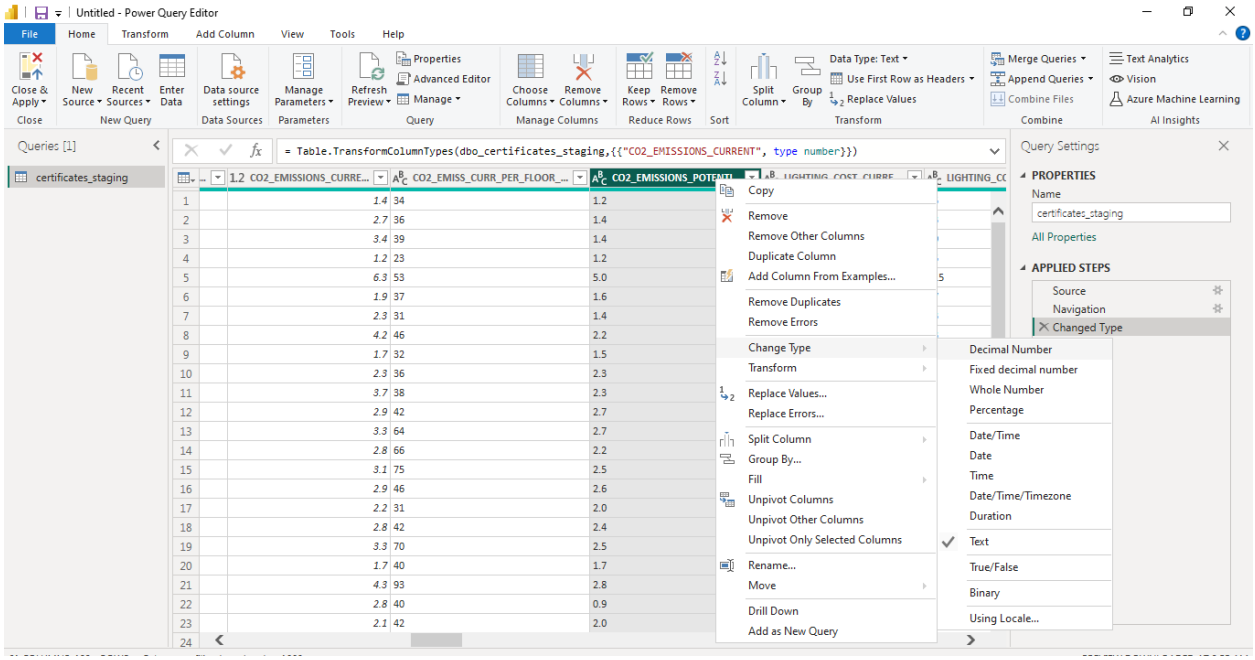
Then your dataset will be transformed into Power Query Editor as follows.



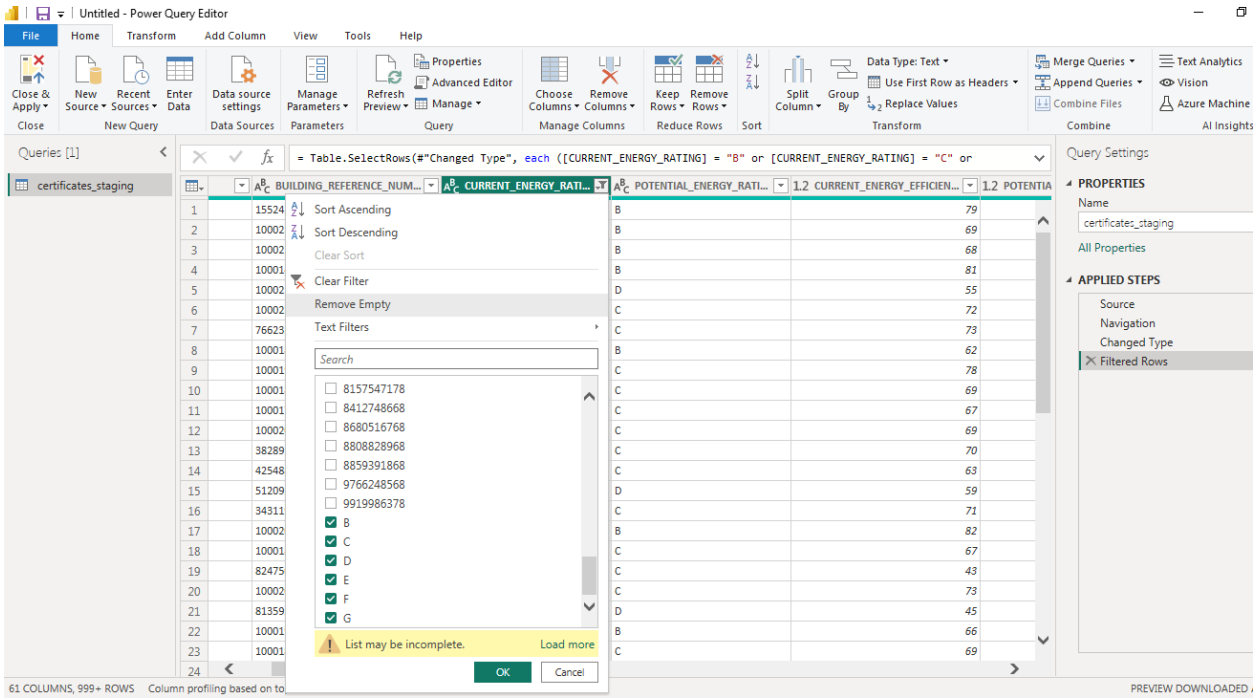
To create relevant data models using DAX, the data type of some columns have changed as follows.

- CO2_EMISSIONS_CURRENT
- CO2_EMISSIONS_POTENTIAL
- CURRENT_ENERGY_EFFICIENCY
- POTENTIAL_ ENERGY_EFFICIENCY

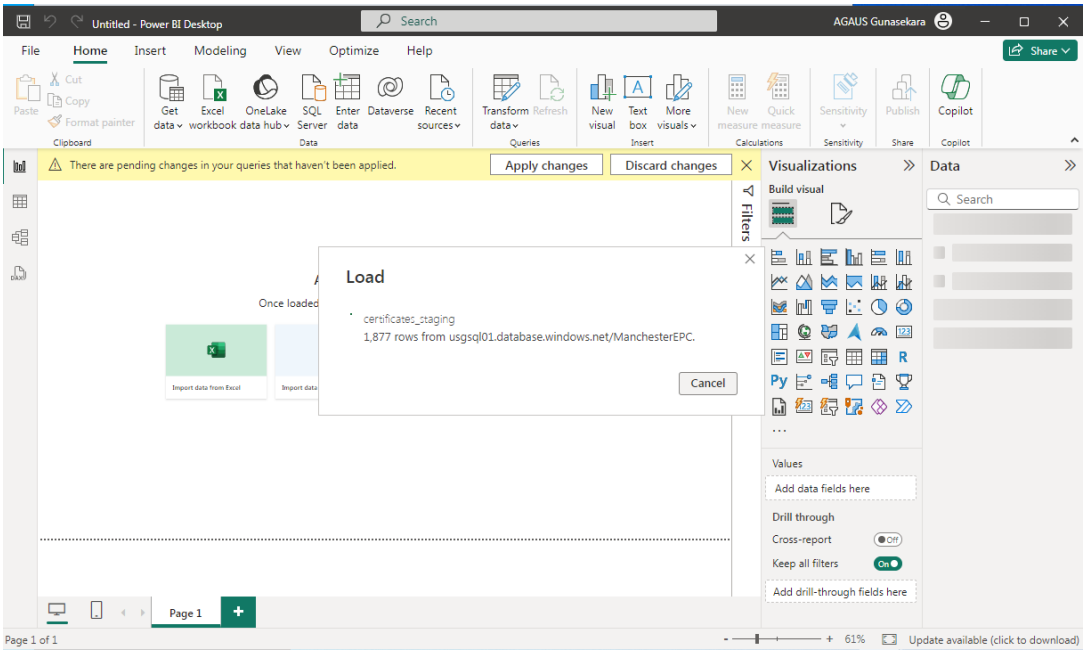
The datatype of the above columns have changed from string to numeric (Text → Decimal Number).



Further, some rows are filtered as follows.

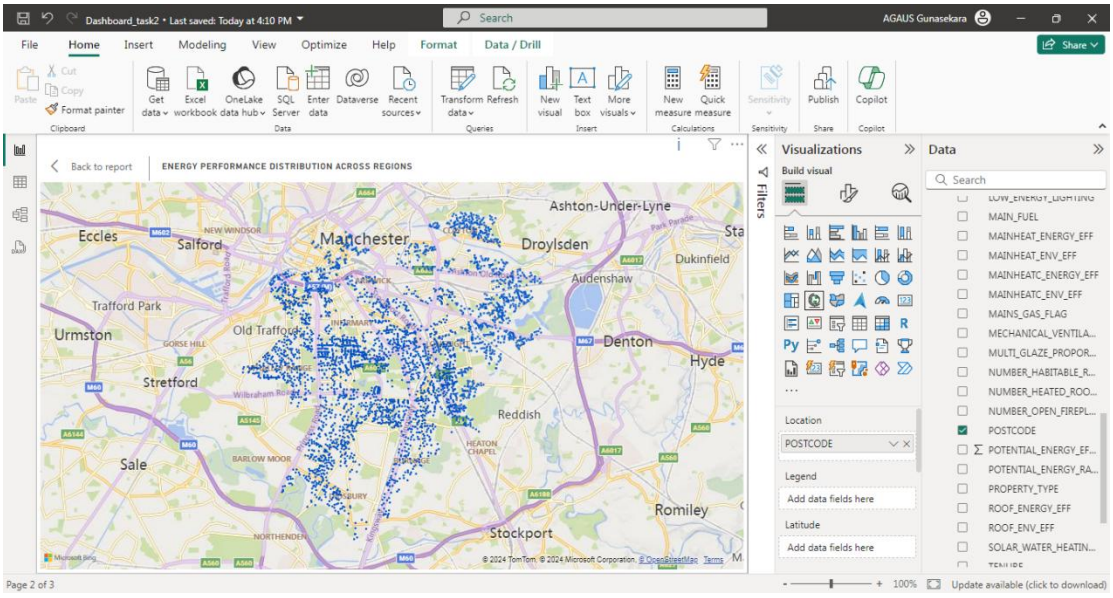


Then select on Close and Apply to do create the dashboard. It will load the finalized dataset to the Power BI desktop.



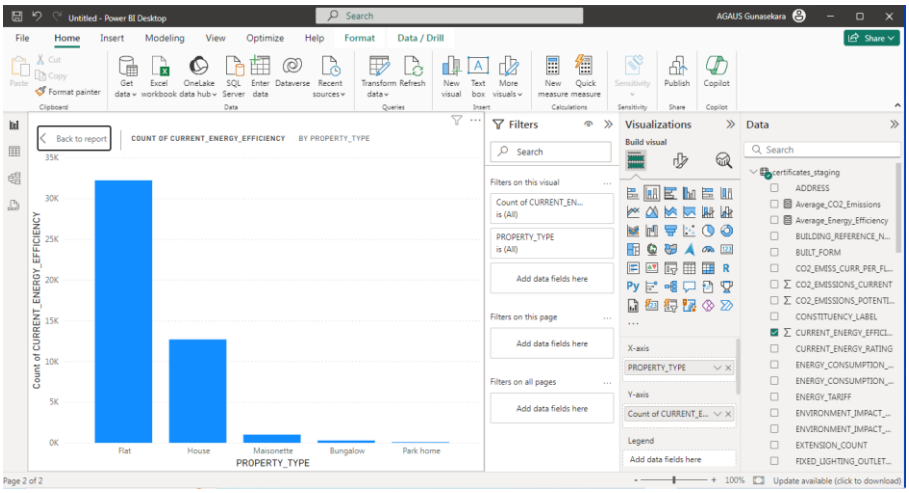
Dashboard Draw up

Creating the Map



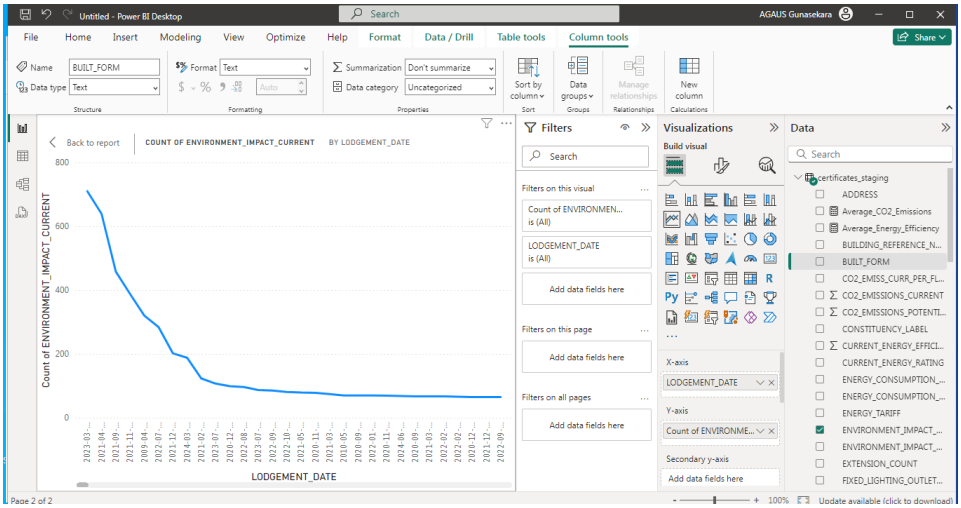
The map visualizes a geospatial representation of Greater Manchester properties, within the United Kingdom. It displays the distribution of energy performance across different regions within Greater Manchester. The dots represents individual data points, potentially corresponding to locations within the region. This map visualization helps to identify areas with higher or lower energy efficiency in Greater Manchester.

Creating the Column chart



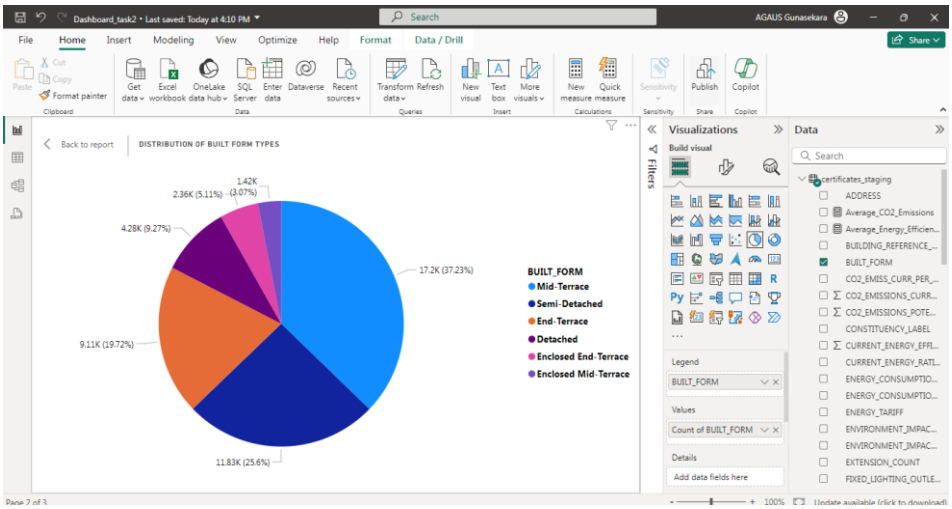
The above graph presents a comparison of the "Count of CURRENT_ENERGY_EFFICIENCY" across different "PROPERTY TYPE" categories. The x-axis represents the "PROPERTY TYPE" categories (Flat, House, Maisonette, Bungalow, Park home), while the y-axis displays the "Count of CURRENT_ENERGY_EFFICIENCY." The height of each bar corresponds to the number of properties in that particular "PROPERTY TYPE" category with a recorded "CURRENT_ENERGY_EFFICIENCY." The bar graph effectively illustrates the distribution of "CURRENT_ENERGY_EFFICIENCY" among different property types, allowing for easy comparison and identification of trends.

Creating the Line chart



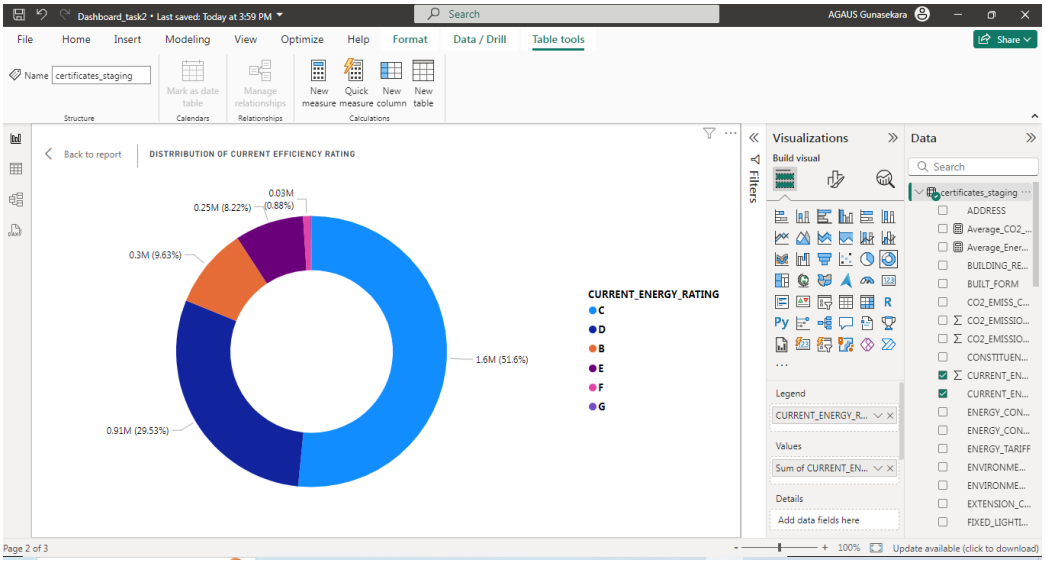
The line chart represents a time series analysis of the "Count of ENVIRONMENT_IMPACT_CURRENT" over a period spanning from 2013 to 2023. The x-axis represents "LODGEMENT_DATE," indicating the time progression of the data. The y-axis displays the "Count of ENVIRONMENT_IMPACT_CURRENT," suggesting the number of environmental impact occurrences over time. The line chart shows a general downward trend, indicating a decrease in the number of environmental impact cases over the years

Creating the Pie chart



The pie chart is a representation of the distribution of "BUILT FORM" categories in Greater Manchester. The pie chart is divided into segments, each representing a different "BUILT FORM" type. The chart includes a legend that identifies the different "BUILT FORM" types and their corresponding percentages. This visualization provides a clear and concise overview of the relative prevalence of each "BUILT FORM" category in the data.

Creating the Donut chart

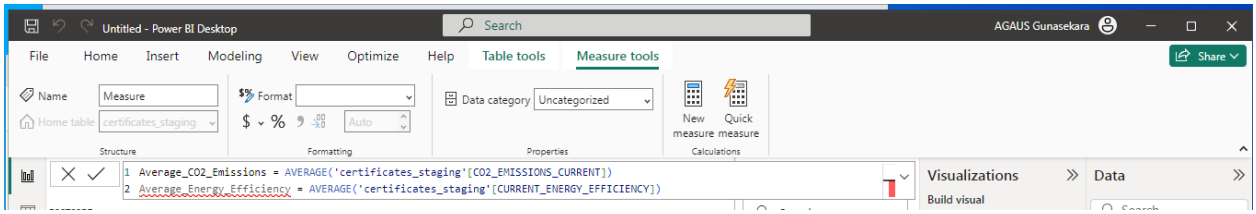


The donut chart shows the distribution of current energy ratings among domestics in Greater Manchester. The categories represent the different energy ratings, ranging from A to G, with A indicating the highest efficiency and G the lowest. The chart reveals that the majority of properties fall within the D category, followed by the C category. The remaining categories, from B to G, have relatively smaller proportions.

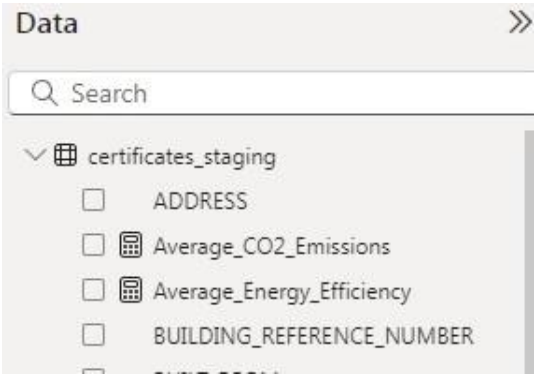
Creating the KPI

Use of DAX to calculate average values.

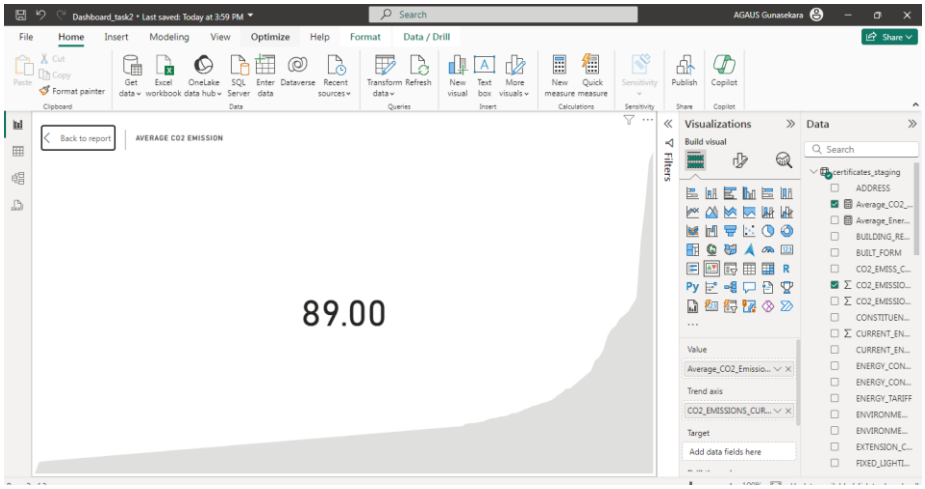
Before create KPIs we need to find the average values for current CO2 Emission and current energy efficiency. For that we can click on ‘New measure’ on the toolbar tab and enter the code as follows.

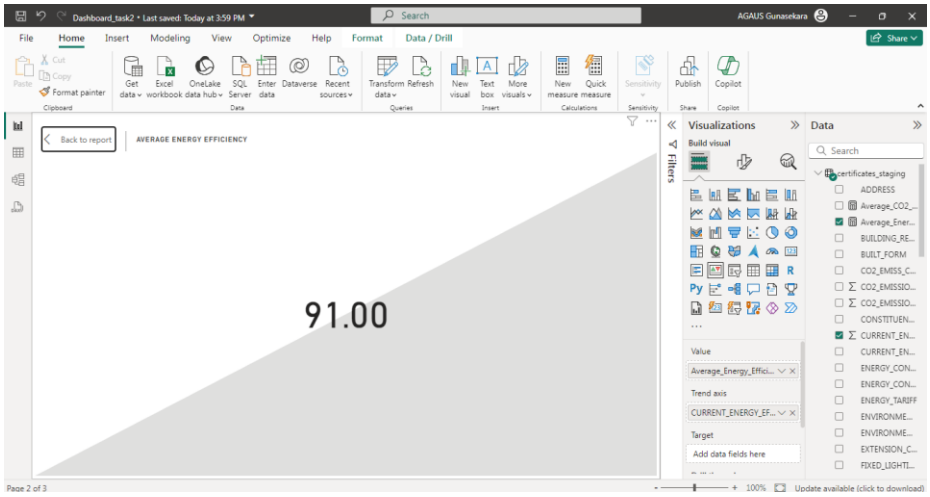


This will create you 2 new measures ‘Average_CO2_Emissions’ and ‘Average_Energy_Efficiency’.



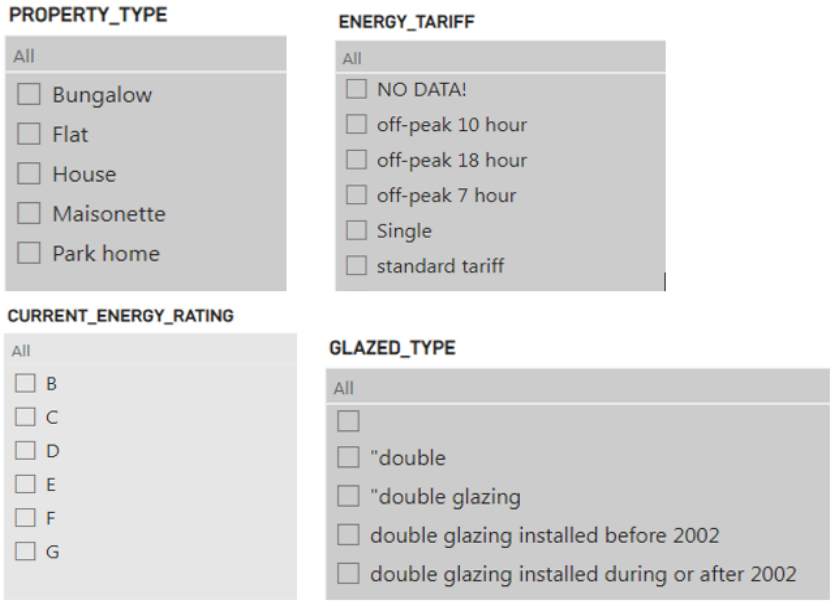
Creating KPI





The above KPIs highlights two important metrics related to domestic energy performance in Greater Manchester. The Average CO2 Emission is measured at 89.00, indicating the typical carbon emissions generated by residential properties, which is a key factor in assessing environmental impact. The Average Energy Efficiency is 91.00, reflecting the overall efficiency of energy use within these homes. These metrics provide an essential overview of the region's energy performance and serve as benchmarks for evaluating future improvements in energy efficiency and emission reduction.

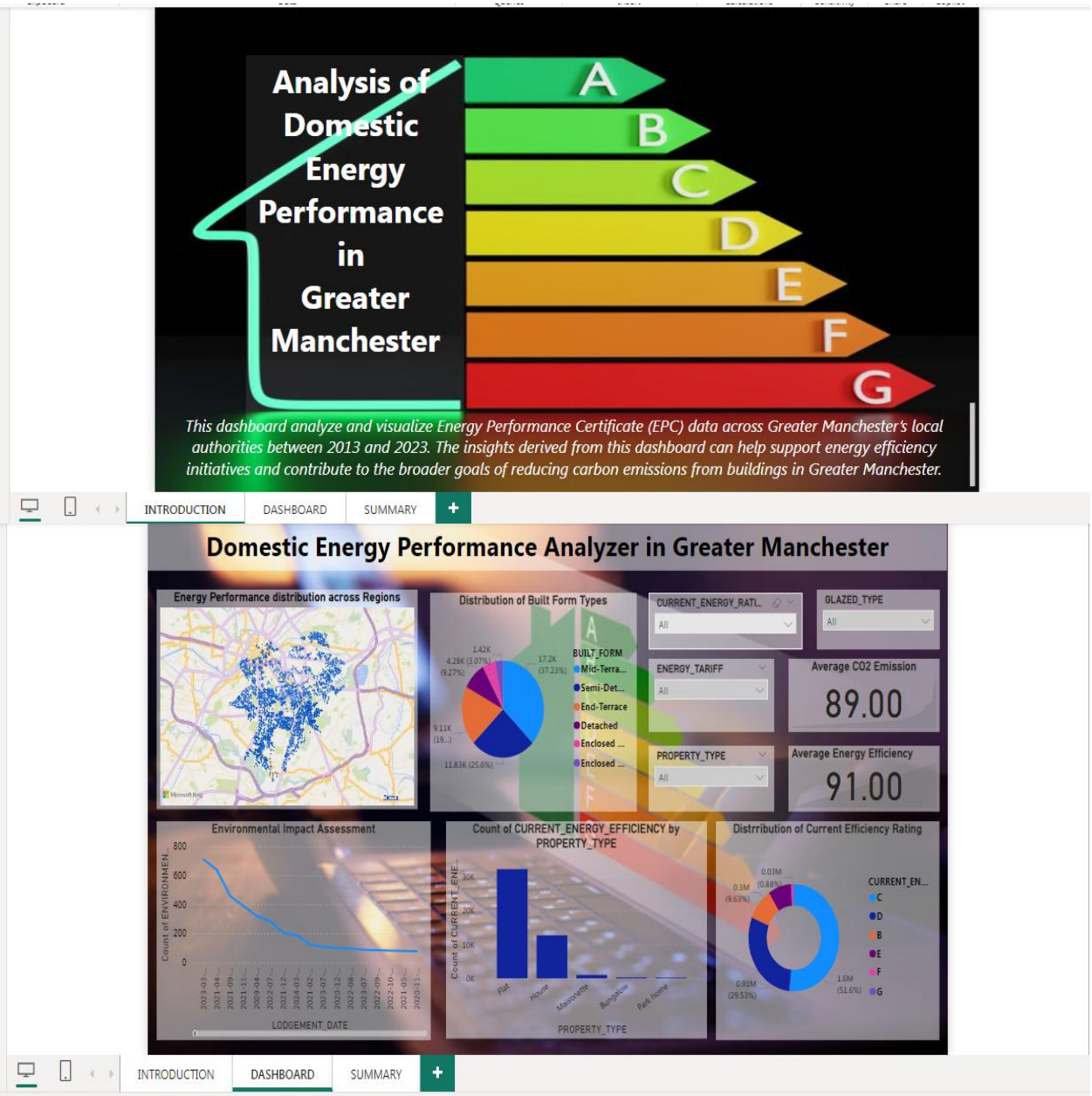
Creating Filters

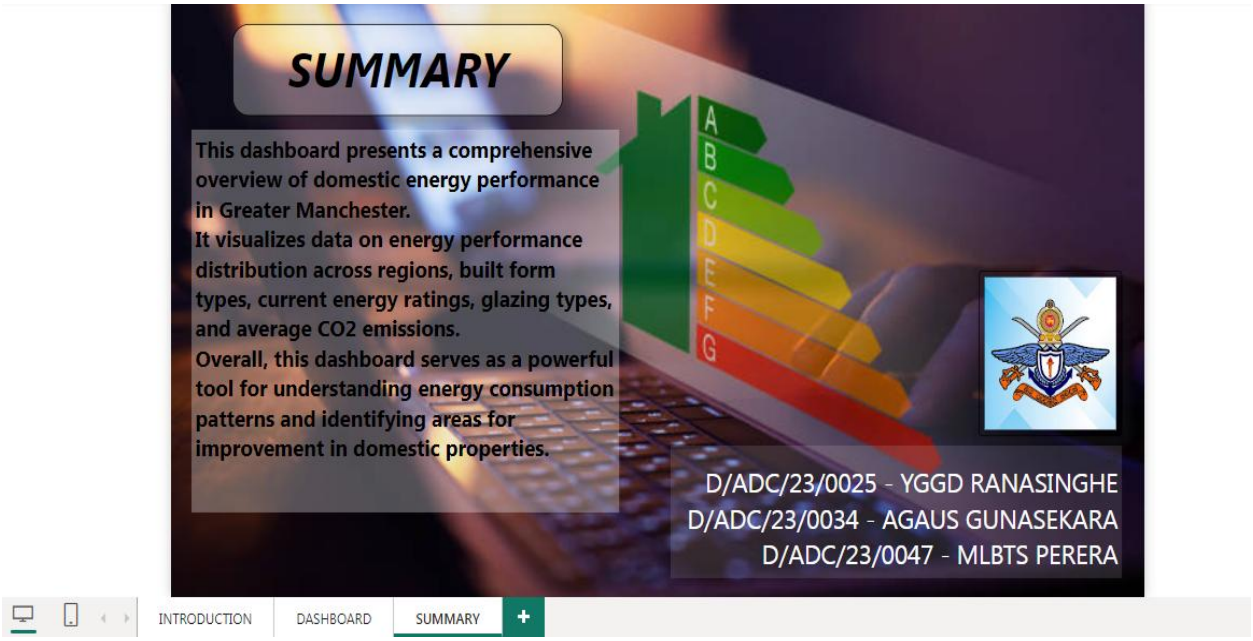


The above filters allow users to filter out the analysis by selecting specific property types, energy tariffs, current energy ratings, and glazing types. By applying these filters, users can focus on particular segments of the data and gain deeper insights into the energy performance characteristics of different property types, energy usage patterns, and glazing types.

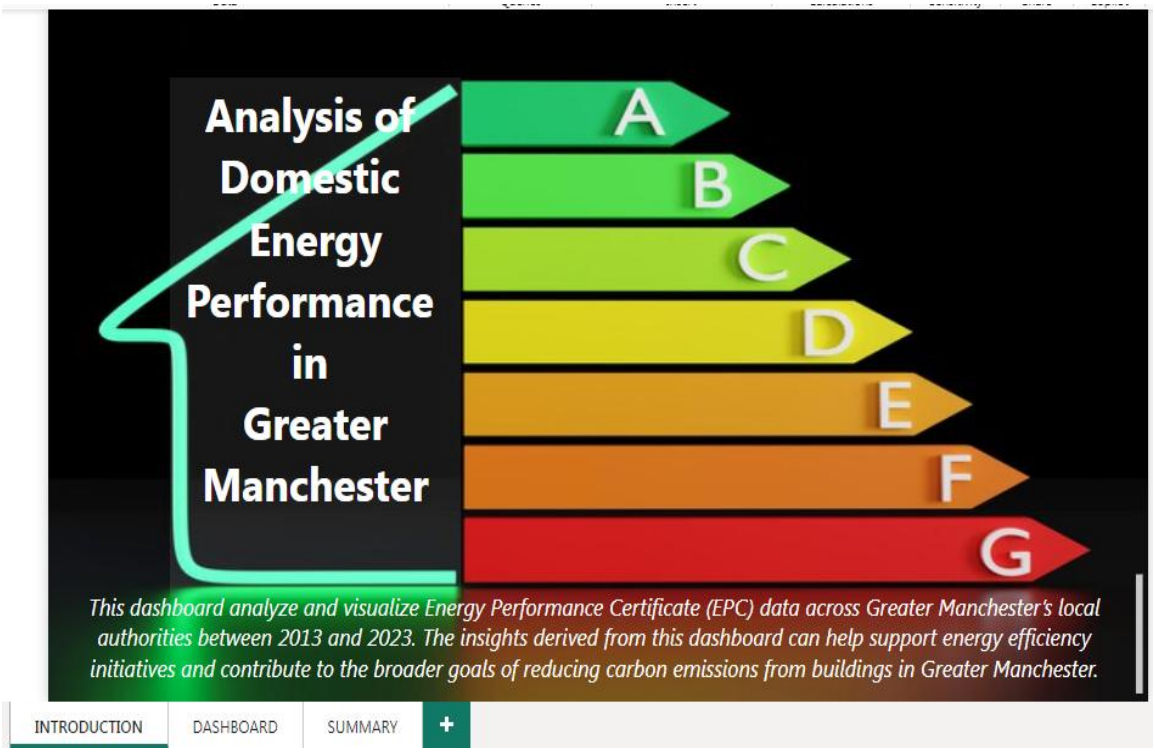
Edit the Dashboard

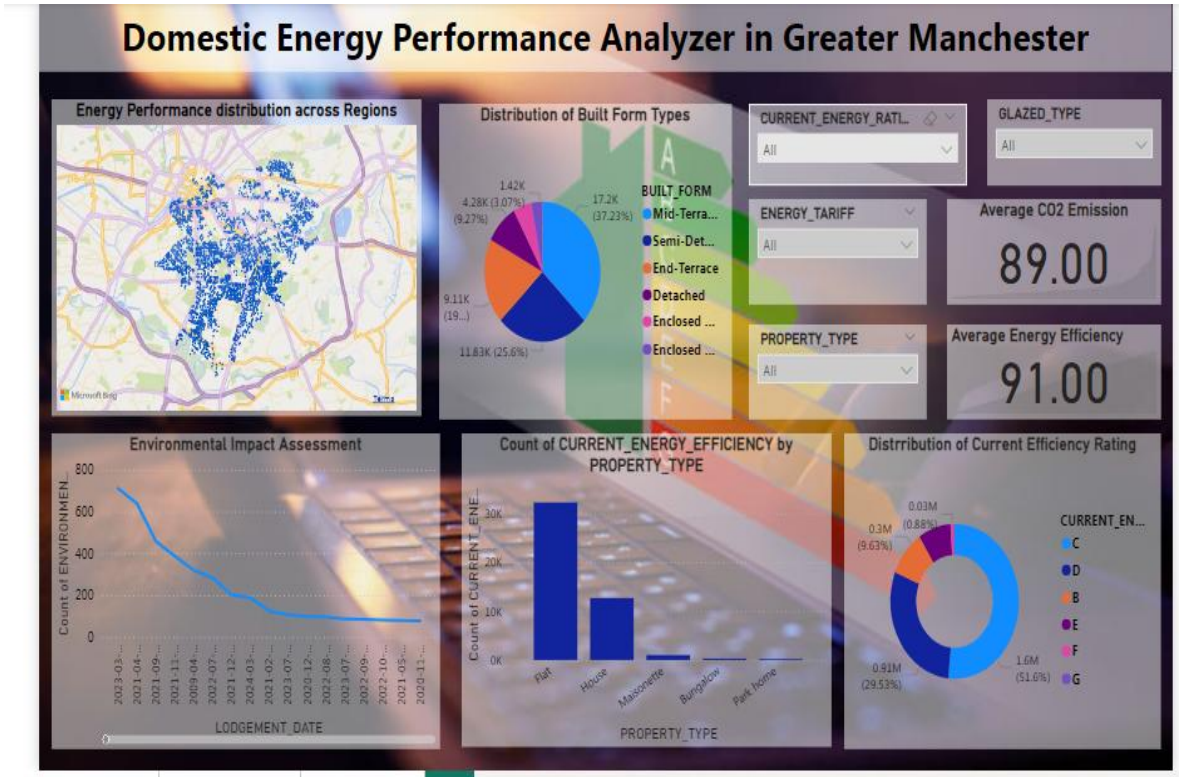
We edited the font sizes, font colors, backgrounds, images in the visualization parts.
As the next step, we arranged three pages in the whole Dashboard Document.
The 1st page is for the introduction, 2nd page is for the dashboard and 3rd page is for the summary.





f) Finalized Dashboard.




Domestic Energy Performance Analyzer in Greater Manchester
Energy Performance distribution across Regions
Distribution of Built Form Types
BUILT_FORM
Mid-Terra...
Semi-Det...
End-Terrace
Detached
Enclosed ...
Enclosed ...
CURRENT_ENERGY_RATI...
GLAZED_TYPE
ENERGY_TARIFF
PROPERTY_TYPE
Average CO2 Emission
89.00
Average Energy Efficiency
91.00
Environmental Impact Assessment
Count of ENVIRONMEN...
LODGEMENT_DATE
Count of CURRENT_ENERGY_EFFICIENCY by PROPERTY_TYPE
Count of CURRENT_ENE...
Flat
House
Maisonette
Bungalow
Park home
PROPERTY_TYPE
Distrribution of Current Efficiency Rating
CURRENT_EN...
C
D
B
E
F
G
INTRODUCTION
DASHBOARD
SUMMARY

SUMMARY
This dashboard presents a comprehensive overview of domestic energy performance in Greater Manchester.
It visualizes data on energy performance distribution across regions, built form types, current energy ratings, glazing types, and average CO2 emissions.
Overall, this dashboard serves as a powerful tool for understanding energy consumption patterns and identifying areas for improvement in domestic properties.
A
B
C
D
E
F
G
D/ADC/23/0025 - YGGD RANASINGHE
D/ADC/23/0034 - AGAUS GUNASEKARA
D/ADC/23/0047 - MLBTS PERERA
INTRODUCTION
DASHBOARD
SUMMARY

4) Conclusion

A dashboard provides a comprehensive and easily digestible overview of data, fostering better decision-making, collaboration, and performance monitoring to the viewers. The above dashboard provides a comprehensive overview of domestic energy performance in Greater Manchester. It includes several visualizations that present various aspects of energy efficiency, built form types, and environmental impact. Therefore, by looking at the above dashboard viewers can get an idea about the domestic energy performance in Greater Manchester. The map displays the distribution of energy performance across different regions within Greater Manchester while Pie chart shows the distribution of built form types and current efficiency ratings. The line chart Illustrates the trend of environmental impact assessments over time. The bar chart compares the count of current energy efficiency by property type. Overall, the dashboard offers a valuable tool for analyzing domestic energy performance in Greater Manchester. It enables stakeholders to identify areas for improvement, understand the impact of different building types, and track progress in environmental sustainability efforts.

5) Recommendations

Prioritize energy efficiency in new developments: The distribution of built form types shows a prevalence of older, less energy-efficient structures. Encouraging the construction of new homes and buildings with high energy efficiency standards can significantly reduce overall energy consumption in the region.

Encourage the adoption of renewable energy: The region can benefit from investing in renewable energy sources like solar panels and wind turbines. Providing financial incentives and simplifying the installation process can encourage homeowners and businesses to adopt these sustainable energy solutions.

Implement targeted energy efficiency programs: Developing and implementing tailored energy efficiency programs for specific demographics, such as low-income households and small businesses, can help address energy poverty and improve energy efficiency across the region.

Promote energy awareness and education: Raising public awareness about energy efficiency and implementing educational campaigns, providing energy-saving tips, and organizing community events can help empower residents and businesses to make informed decisions about energy consumption.

6) References

https://epc.opendatacommunities.org/downloads/domestic#local-authority