



GENERAL SIR JOHN KOTELAWALA DEFENCE UNIVERSITY

FACULTY OF MANAGEMENT SOCIAL SCIENCES & HUMANITIES

DEPARTMENT OF LANGUAGES

BACHELOR OF SCIENCE IN APPLIED DATA SCIENCE AND

COMMUNICATION

INTAKE 41

LB 2114: ADVANCED SQL AND CLOUD DATABASES

ASSIGNMENT

Group Members

D/ADC/24/0003 – Thesanya Lamahewa

D/ADC/24/0013 – Samoda De Silva

D/ADC/24/0028 – Thisari Perera

D/ADC/24/0048 – Dilakna Godagamage

Task 2

Energy Performance Trends in Manchester's Local Authorities (2014–2024)

Advanced SQL & Power BI Analysis



Content

1	Introduction	1
1.1	Background	1
1.2	Purpose of the report	1
1.3	Scope.....	1
1.4	Data Sources	2
2	SQL Data Processing.....	3
2.1	Importing Data to SQL server.....	3
2.2	Data cleaning with CTEs and Views.....	5
2.3	Challenges during importing the data	5
3	Exploratory Data Analysis.....	6
3.1	Create a view for cleaned data	6
3.2	Stored Procedure for Automated Cleaning	7
3.3	Use of system functions and Aggregates	8
3.4	Aggregate Summary by Property Type.....	11
4	Power BI Data Analysis & Visualization	12
4.1	Importing the Dataset to Power BI.....	12
4.2	DAX measures for Data Analyzing	13
4.3	Power BI Data Visualization.....	21
4.4	Explanation of Dashboard Visuals.....	22
4.5	The Connection of the SQL Server with the Power BI	27
4.6	Interpretation of the Dashboard.....	28
4.7	Application of Advanced Power BI Features.....	28
4.8	Key Insights and actionable recommendations	29
5	Conclusion.....	29

1 Introduction

1.1 Background

The energy performance certificates (EPC) are used to give details about the efficiency of energy of properties used in the UK, both residential and commercial. They display the current and potential energy rating of a property, A (most energy efficient) to G (least energy efficient). The use of the EPC data can be beneficial in researching the trends in the energy efficiency and in supporting the environmental policies as well as the housing standards. The local authorities of Manchester have experienced a great development during the last ten years and therefore an EPC analysis is useful in comprehending how property performance and sustainability objectives have changed.

1.2 Purpose of the report

This report is aimed at analyzing and visualizing EPC data of local authorities in Manchester, with the help of SQL advanced methods and using Power BI dashboards.

This report aims to:

- Prepare EPC datasets with SQL
- Determine energy performance trends 2014-2024
- Compare price of efficiency rates amongst local authorities

1.3 Scope

The current analysis is devoted to the EPC records that are to be issued within the Manchester Local Authority and in the period between 2014 and 2024.

The scope includes:

- Domestic Property only – Important variables included EPC rating, the property type, the age at which it was constructed, the postcode and the energy performance indicators.

1.4 Data Sources

The report uses publicly available EPC datasets provided through:

- UK Government Open Data (EPC Register)

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

The screenshot shows a list of local authorities with their names and download links. The list includes Lichfield, Lincoln, Liverpool, Luton, Maidstone, Maldon, Malvern Hills, and Manchester. Below these, there are sections for 'M' (highlighted in blue), 'Y' (highlighted in blue), and 'OTHER' (highlighted in dark blue). Under 'M', it lists Wrexham, Wychavon, Wycombe, Wyre, and Wyre Forest. Under 'Y', it lists York. Under 'OTHER', it lists '[Unknown Local Authority]'. Each item has a 'Download' button to its right.

3: Domestic EPCs by Lodgement Date

From 2025 we have bundled EPCs by month. For an earlier period, [filter the Domestic EPCs](#). These files all have the following structure:

Filename	Description
https://epc.opendatacommunities.org/files/domestic-E08000003-Manchester.zip	[a]

Figure 1: Dataset

2 SQL Data Processing

2.1 Importing Data to SQL server

All three EPC CSV datasets were successfully imported into SQL Server, forming the basis for subsequent data cleaning and transformation.

A new database was created in SQL Server to store the EPC datasets. Then the three CSV files were imported into the database using the SQL Server Import Wizard. During this process, the column mappings were checked to ensure that all fields matched correctly, and the data was successfully loaded into staging tables. This completed the initial setup required for further cleaning, transformation, and analysis.

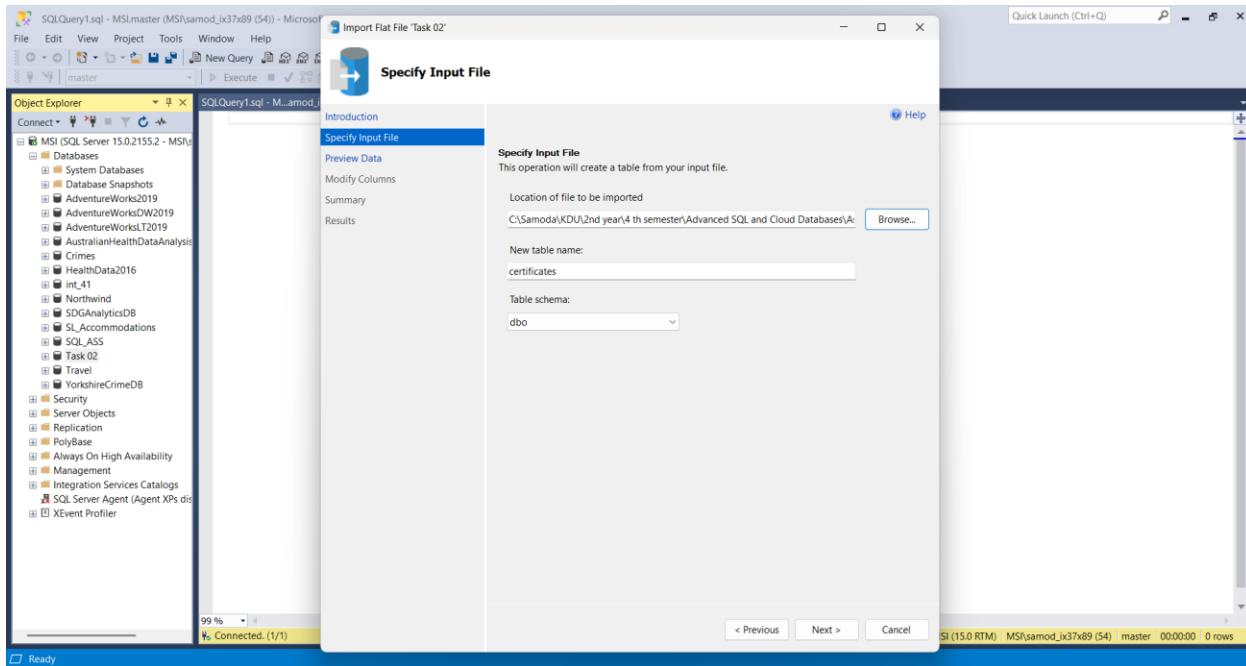


Figure 2: Specify Input File

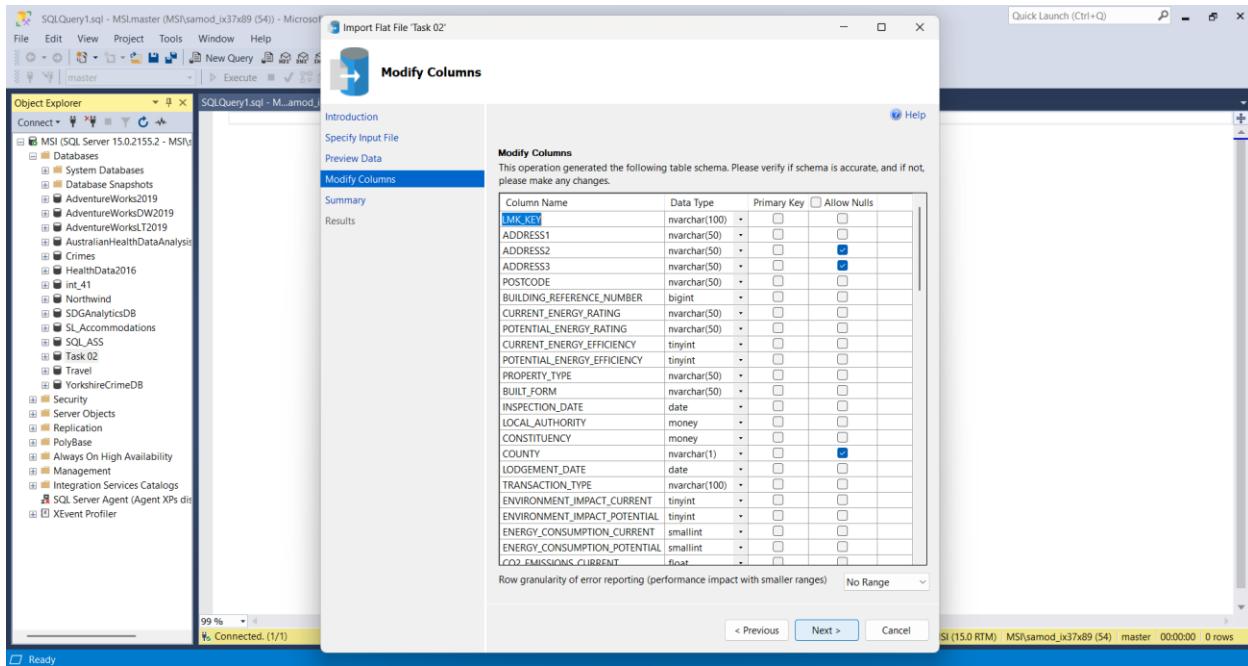


Figure 3: Modify Columns

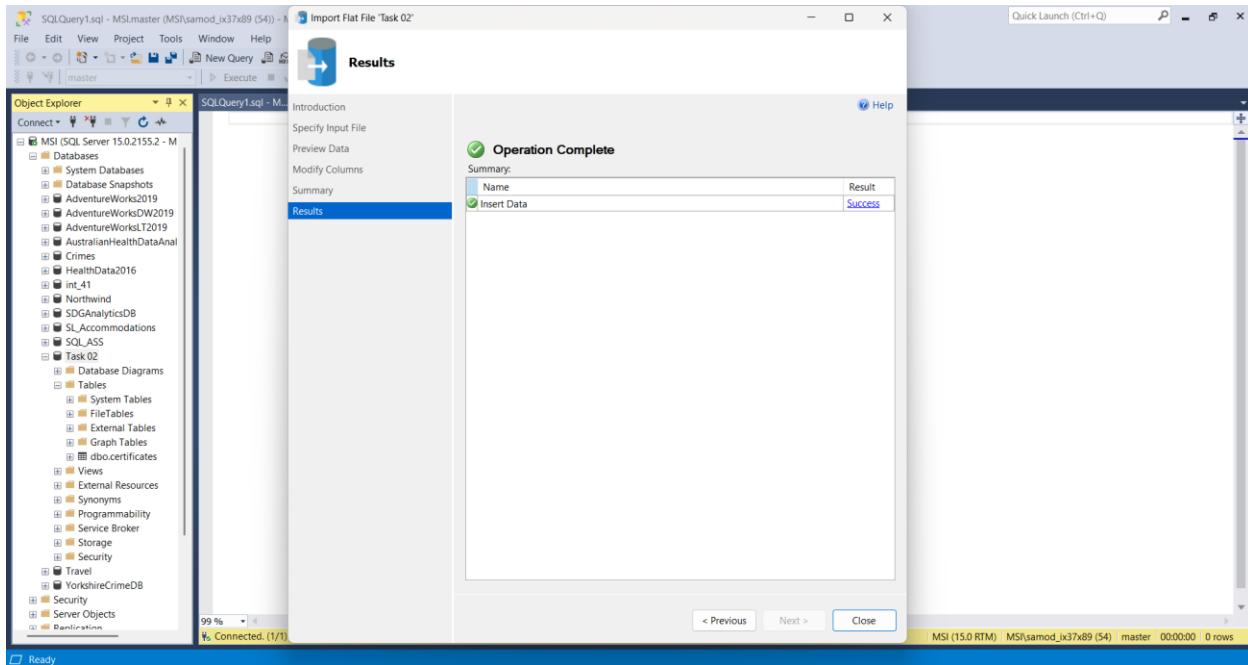
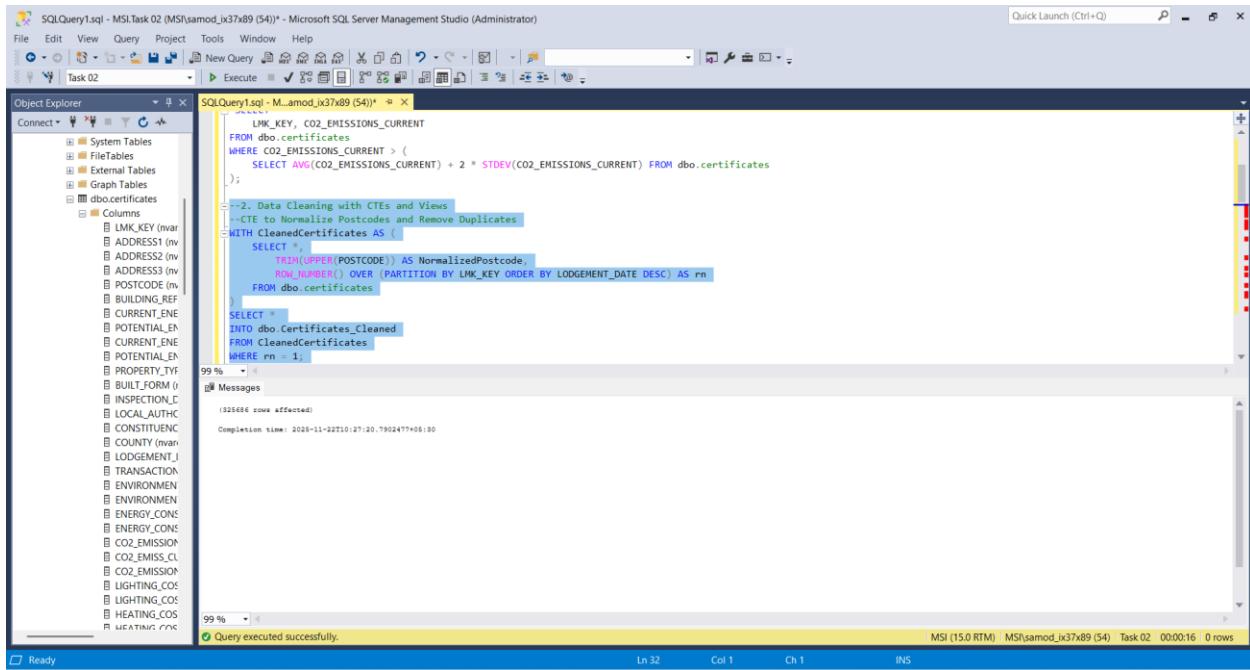


Figure 4: Operation complete

2.2 Data cleaning with CTEs and Views

The process of data cleaning was conducted with the help of Common Table Expressions (CTEs) and SQL Views that facilitated the process of arranging and making the transformation. Missing values, invalid records and the standardization of key fields were dealt with using CTEs and Views were generated to hold the cleaned and formatted datasets to facilitate easier querying and analysis in the Power BI.



The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left lists various database objects like System Tables, FileTables, External Tables, Graph Tables, and dbo.certificates. The central pane displays a query window titled 'SQLQuery1.sql - MSI.Task 02 (MSI\samod_1x37x89 (54))'. The query code is as follows:

```
--1. Data Cleaning with CTE View
--CTE to Normalize Postcodes and Remove Duplicates
WITH CleanedCertificates AS (
    SELECT *,
        TRIM(UPPER(POSTCODE)) AS NormalizedPostcode,
        ROW_NUMBER() OVER (PARTITION BY LMK_KEY ORDER BY LODGEMENT_DATE DESC) AS rn
    FROM dbo.certificates
)
SELECT *
INTO dbo.Certificates_Cleaned
FROM CleanedCertificates
WHERE rn = 1;
```

At the bottom of the query window, it says 'Query executed successfully.' and shows completion time: 2020-11-22T10:27:20.7902477+05:30. The status bar at the bottom indicates 'Ln 32 Col 1 Ch 1 INS'.

Figure5: Data Cleaning with CTE View

2.3 Challenges during importing the data

When handling EPC Manchester data, a number of challenges were encountered during importation and cleaning. There were irregular formatting, missing values and non- numeric entries in numeric key fields for an example, energy score and floor area in the raw files. These problems had to be carefully validated, invalid rows were to be eliminated and data types were to be standardized to make the data accurate.

3 Exploratory Data Analysis

3.1 Create a view for cleaned data

Check for missing values

The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left shows a database named 'Task 02' containing various tables and columns. The main window displays a T-SQL script to check for missing values:

```
--USE [Task 02]
--Exploratory Data Analysis (EDA)
--Check for Missing Values
SELECT
    COUNT(*) AS TotalRows,
    SUM(CASE WHEN ADDRESS1 IS NULL THEN 1 ELSE 0 END) AS Missing_ADDRESS1,
    SUM(CASE WHEN POSTCODE IS NULL THEN 1 ELSE 0 END) AS Missing_POSTCODE,
    SUM(CASE WHEN CURRENT_ENERGY_EFFICIENCY IS NULL THEN 1 ELSE 0 END) AS Missing_CURRENT_ENERGY
FROM dbo.certificates;
```

The results pane shows a single row of data:

TotalRows	Missing_ADDRESS1	Missing_POSTCODE	Missing_CURRENT_ENERGY
325686	1	0	0

At the bottom, a message indicates "Query executed successfully."

Figure 6: Checking Missing Values

Detect Outliers

The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left shows a database named 'Task 02' containing various tables and columns. The main window displays a T-SQL script to detect outliers in CO2 emissions:

```
--Detect Outliers (e.g., unusually high CO2 emissions
SELECT
    LMK_KEY, CO2_EMISSIONS_CURRENT
FROM dbo.certificates
WHERE CO2_EMISSIONS_CURRENT > (
    SELECT AVG(CO2_EMISSIONS_CURRENT) + 2 * STDEV(CO2_EMISSIONS_CURRENT) FROM dbo.certificates
);
```

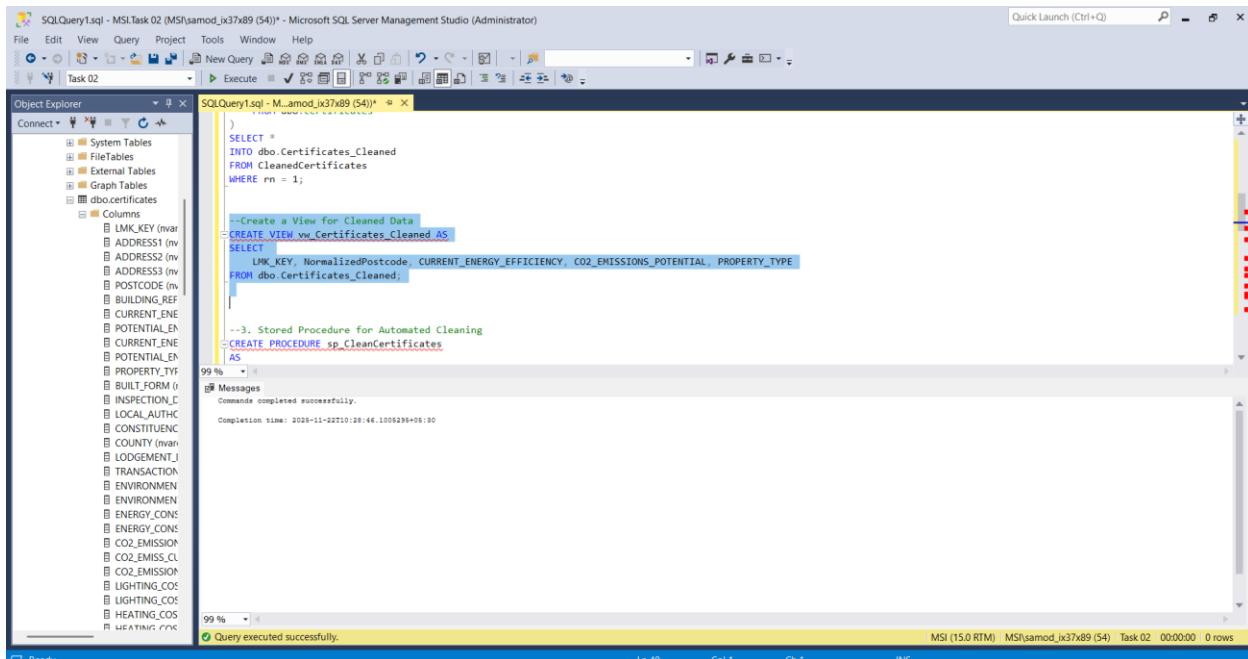
The results pane shows a list of certificates with their LMK_KEY and CO2_EMISSIONS_CURRENT values. The first few rows are:

LMK_KEY	CO2_EMISSIONS_CURRENT
1	13
2	8.3999991853027
3	11
4	7.6999990926514
5	7.6999990926514
6	10
7	7.8000019073486
8	7.8000019073486
9	8.5
10	11
11	12
12	13
13	8.10999938168973
14	9.3999991853027
15	7.8000019073486
16	7.900000958743
17	7.800000958743

At the bottom, a message indicates "Query executed successfully."

Creating a View for Cleaned Data

The Cleaned EPC Manchester data was visualized to give a simplified and reliable version of the data. It enables end users to view the corrected and standardized data without making any changes in the original tables. This simplifies analysis, provides uniformity in the reports, and facilitates effective data accessibility to proceed with additional Power BI visualization.



The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. The Object Explorer on the left shows a database structure with several tables under the 'cbo.certificates' schema, including LMK_KEY, ADDRESS1, ADDRESS2, ADDRESS3, POSTCODE, BUILDING_REF, CURRENT_ENE, POTENTIAL_EN, CURRENT_ALEN, POTENTIAL_ALEN, PROPERTY_TYF, BUILT_FORM, INSPECTION_LC, LOCAL_AUTHC, CONSTITUENC, COUNTY, LODGEMENT_J, TRANSACTION, ENVIRONMEN, ENERGY_CONS, CO2_EMISSION, CO2_EMSS_CL, CO2_EMSS, LIGHTING_COS, HEATING_COS, and HEATING_ALEN. The central Query Editor window contains T-SQL code for creating a view and a stored procedure:

```
SELECT *
INTO dbo.Certificates_Cleaned
FROM CleanedCertificates
WHERE rn = 1;

--Create a View for Cleaned Data
CREATE VIEW vw_Certificates_Cleaned AS
SELECT
    LMK_KEY, NormalizedPostcode, CURRENT_ENERGY EFFICIENCY, CO2_EMISSIONS_POTENTIAL, PROPERTY_TYPE
FROM dbo.Certificates_Cleaned;

--3. Stored Procedure for Automated Cleaning
CREATE PROCEDURE sp_CleanCertificates
AS
```

The status bar at the bottom indicates "Query executed successfully." and "Completion time: 2020-11-22T10:28:44 1000ms+00:00".

Figure 7: Create a view for cleaned data

3.2 Stored Procedure for Automated Cleaning

A Stored procedure was written to automatize the process of data cleaning, which is consistent and efficient when importing new EPC data. The process normalizes format, eliminates invalid or duplicate entries, manages missing data, and automatically updates the refined data, which saves manual effort and enhances the reliability of the data.

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar reads "SQLQuery1.sql - MSI.Task 02 (MSI\samod_ix37x89 (54)) - Microsoft SQL Server Management Studio (Administrator)". The left pane is the Object Explorer, showing a tree structure of database objects including System Tables, FileTables, External Tables, Graph Tables, and dbo.certificates. The right pane is the Query Editor window titled "SQLQuery1.sql - M...amod_ix37x89 (54)". It contains the following T-SQL code:

```
--3. Stored Procedure for Automated Cleaning
CREATE PROCEDURE sp_CleanCertificates
AS
BEGIN
    -- Remove rows with NULL critical fields
    DELETE FROM dbo.certificates
    WHERE CURRENT_ENERGY_RATING IS NULL OR POSTCODE IS NULL;

    -- Update inconsistent casing
    UPDATE dbo.certificates
    SET POSTCODE = UPPER(Trim(POSTCODE))
    WHERE POSTCODE IS NOT NULL;
END
```

The status bar at the bottom indicates "Query executed successfully." and "Completion time: 2026-11-22T10:29:07.0997839+00:00".

Figure 8: Stored Procedure for automated cleaning

3.3 Use of system functions and Aggregates

Major elements of the EPC data were summarized and analyzed using system functions and aggregate functions. System functions assisted in formatting dates, Standardization of therapeutic texts and extraction of particular data values and aggregate functions were used to compute rating distributions, discovering trends over times and comparing energy performance across the local authorities of Manchester as an example.

```

-- To run it:
EXEC sp_CleanCertificates;

-- 4. Use of System Functions and Aggregates
-- Ranking Properties by Energy Efficiency
SELECT
    LMK_KEY, CURRENT_ENERGY_EFFICIENCY,
    RANK() OVER (ORDER BY CURRENT_ENERGY_EFFICIENCY DESC) AS EnergyRank
FROM dbo.certificates;

-- Aggregate Summary by Property Type
SELECT
    PROPERTY_TYPE,
    COUNT(*) AS Total,
    AVG(CURRENT_ENERGY) AS AvgEnergy,
    ...

```

LMK_KEY	CURRENT_ENERGY_EFFICIENCY	EnergyRank
1	99950466608136a6d64924afdd42117044ed617ad1d5...	89
2	d98fe2e2a25e565605604a39d98300da4aefb43427f30...	89
3	12961012096201503051208114398782...	89
4	79c93a2a0...	89
5	a059419a19a0a0a0a0a0a0a0a0a0a0a0a0a0a0a0a0a0a0a0...	89
6	17809113112022020201133428887899...	89
7	dc0c936a92672a4ef55097717295a2899723b68f738d...	89
8	01fb2a2...	89
9	957b193aaac3e668a33a3a0a0a0a0a0a0a0a0a0a0a0a0a0...	90
10	158298555930201712131401205843927...	88
11	18484527928201012891982893968169	87
12	1802379032020151091151547404000...	87
13	79c93a2a0...	89
14	169464396902010911509184232758471	88
15	2578b90546511311e0de40030edea0a14d7e07a490...	89
16	95e6bd49707e26822a43a3330505321a7a16c32950b9c...	88
17	4673403480055181a1a1a1a1a1a1a1a1a1a1a1a1a1a1...	87

Figure 9: Use of system functions and aggregations

Using ISNUMERIC () to identify Non-numeric entries

The ISNUMERIC () is applied to assess whether a value could be treated as a number. This aids in the cleaning of data during the cleaning process since it allows one to detect rows that contain text or symbols or wrongly typed in values in fields which are only supposed to hold numeric values. Identifying nonnumeric values, it is possible to filter out or amend them in order to guarantee the accuracy of the data and preserve the quality of the cleaned EPC Manchester dataset.

SQLQuery1.sql - MSI Task 02 (MSI\samod_ix37x89 (54)) - Microsoft SQL Server Management Studio (Administrator)

File Edit View Query Project Tools Window Help

New Query Standard Object Explorer Task List Task 02

SQLQuery1.sql - M...\amod_ix37x89 (54)*

-- Ranking Properties by Energy Efficiency

```
SELECT
    LMK_KEY, CURRENT_ENERGY EFFICIENCY,
    RANK() OVER (ORDER BY CURRENT_ENERGY EFFICIENCY DESC) AS EnergyRank
FROM dbo.certificates;
```

-- Using ISNUMERIC() to Identify Non-Numeric Entries

```
SELECT
    LMK_KEY, CO2_EMISSIONS_CURRENT
FROM dbo.certificates
WHERE ISNUMERIC(CO2_EMISSIONS_CURRENT) = 0;
```

99 %

Results Messages

LMK_KEY	CO2_EMISSIONS_CURRENT
1	9897145950201200911600016377298
2	9897145950201200911600016377299
3	99911587920201309992021189202337
4	5455182725120110712161549969078
5	8895913295020110712161549969078
6	7259015196520111111510195299996
7	5998513298020110712161549969078
8	5998513298020110712161549969078
9	11022877520146910105911600016377290
10	9900252674201200919145401477228
11	110328516402014051610452825049218
12	99957339702013099191454016272428
13	599819629142011071215200482199028
14	59981962914201107121522338899328
15	99954661781201309919490795270013
16	1676811676167616761676167616761676
17	8895913295020110712161549969078

Query executed successfully.

Figure 10: Using Isnumeric ()

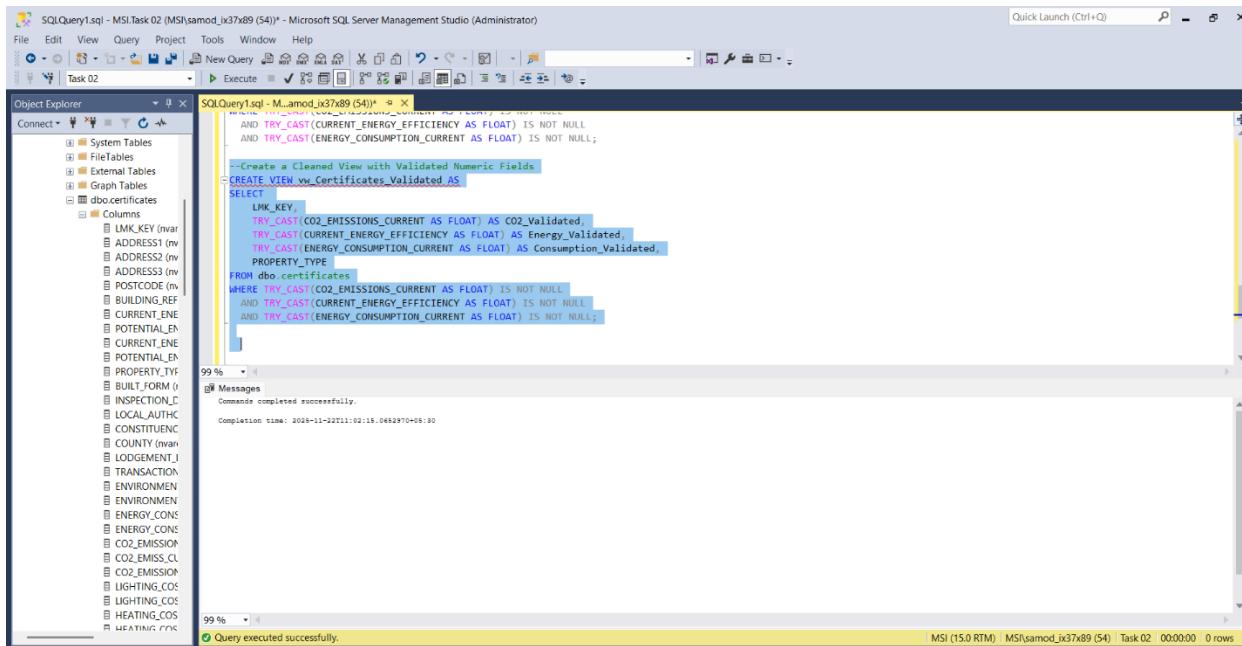
Using TRY_CAST() to safely convert and filter valid numbers

The TRY_CAST() function was adopted to convert the values safely into the numeric types eliminating invalid entries and only admitting the valid numbers into the analysis.

Figure 11: Using TRY_CAST()

Create a Cleaned View with Validated Numeric Fields

An create view was made to records that had valid numeric values were incorporated in the final dataset. The invalid or improperly formatted numbers were filtered out using the check functions like ISNUMERIC () or TRY_CAST. The view offers a reliable analysis ready presentation of the EPC Manchester data and assist in preserving the accuracy of the data and ensure all calculations, visualizations and reporting use confirmed numeric fields.



The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left lists various database objects like System Tables, External Tables, and the dbo.certificates table, which contains columns such as LMK_KEY, CO2_EMISSIONS_CURRENT, CURRENT_ENERGY_EFFICIENCY, and ENERGY_CONSUMPTION_CURRENT. The central Query Editor window displays T-SQL code for creating a view named vu_Certificates_Validated. The code uses TRY_CAST to validate numeric values for these columns before inclusion in the view. The status bar at the bottom indicates the command was completed successfully.

```
--Create a Cleaned View with Validated Numeric Fields
CREATE VIEW vu_Certificates_Validated AS
SELECT
    LMK_KEY,
    TRY_CAST([CO2_EMISSIONS_CURRENT] AS FLOAT) AS CO2_Validated,
    TRY_CAST([CURRENT_ENERGY_EFFICIENCY] AS FLOAT) AS Energy_Validated,
    TRY_CAST([ENERGY_CONSUMPTION_CURRENT] AS FLOAT) AS Consumption_Validated,
    PROPERTY_TYPE
FROM dbo.certificates
WHERE TRY_CAST([CO2_EMISSIONS_CURRENT] AS FLOAT) IS NOT NULL
    AND TRY_CAST([CURRENT_ENERGY_EFFICIENCY] AS FLOAT) IS NOT NULL
    AND TRY_CAST([ENERGY_CONSUMPTION_CURRENT] AS FLOAT) IS NOT NULL;
```

Figure 12: Create a Cleaned View

3.4 Aggregate Summary by Property Type

The number of averages of totals of or other indicators of interest summarized by each property type. This allows comparing and analyzing trends, performance level, or distribution within various categories fast to make informed decisions and based on facts.

```

-- 4. Use of System Functions and Aggregates
--Ranking Properties by Energy Efficiency
SELECT
    LNK_KEY, CURRENT_ENERGY EFFICIENCY,
    RANK() OVER (ORDER BY CURRENT_ENERGY EFFICIENCY DESC) AS EnergyRank
FROM dbo.certificates;

--Aggregate Summary by Property Type
SELECT
    PROPERTY_TYPE,
    COUNT(*) AS Total,
    AVG(CURRENT_ENERGY EFFICIENCY) AS AvgEnergy,
    AVG(CO2 EMISSIONS POTENTIAL) AS AvgCO2
FROM dbo.certificates
GROUP BY PROPERTY_TYPE;

```

PROPERTY_TYPE	Total	AvgEnergy	AvgCO2
Bungalow	4007	88	1.36819903996148
Flat	143040	71	1.82015987461
House	173995	53	2.181472958895
Maisonette	4843	86	2.6004307062955
Park home	1	42	3.0999999048325

Figure 12: Aggregate Summary

4 Power BI Data Analysis & Visualization

4.1 Importing the Dataset to Power BI

Once the data was cleaned and ready in SQL Server, the processed data was added to power BI with the inbuilt SQL Server connector. Power BI loaded the cleaned tables by connecting directly to the database and was able to model and visualize the data.

The Certificates, Columns, and Recommendations CSV files were successfully imported into Power BI using a structured data-loading method, ensuring that all datasets were properly transferred and fully accessible for additional cleaning, modelling and analysis.

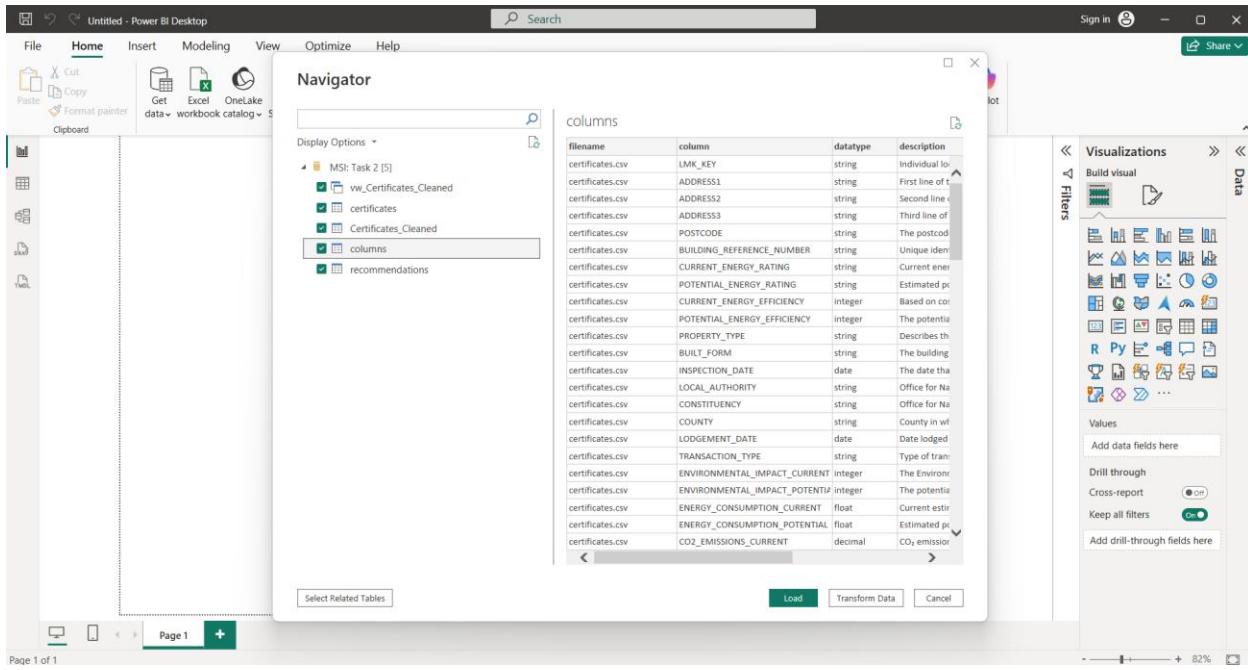


Figure 13: Importing Data to Power BI

4.2 DAX measures for Data Analyzing

The measures were developed in Power BI to have dynamic calculations like totals, averages and performance comparisons with the help of DAX, the dashboard is capable of generating correct real time insights in cases of user selection and filtering. These actions assist in the conversion of raw data to meaningful metrics which enhances the depth of analysis and clarity of the visualizations in general.

Power Query Editor

Power Query Editor cleans, transforms, and prepares raw data and then loads it in the model. It enables to eliminate the errors, divide or combine columns, alter data types, filter rows and carry out the transformation without coding. All the actions are automatically registered, which makes the cleaning process visible and repeatable. Reliance on the Power Query Editor makes the EPC Manchester data is arranged well, correct, and can be analyzed in the dashboard.

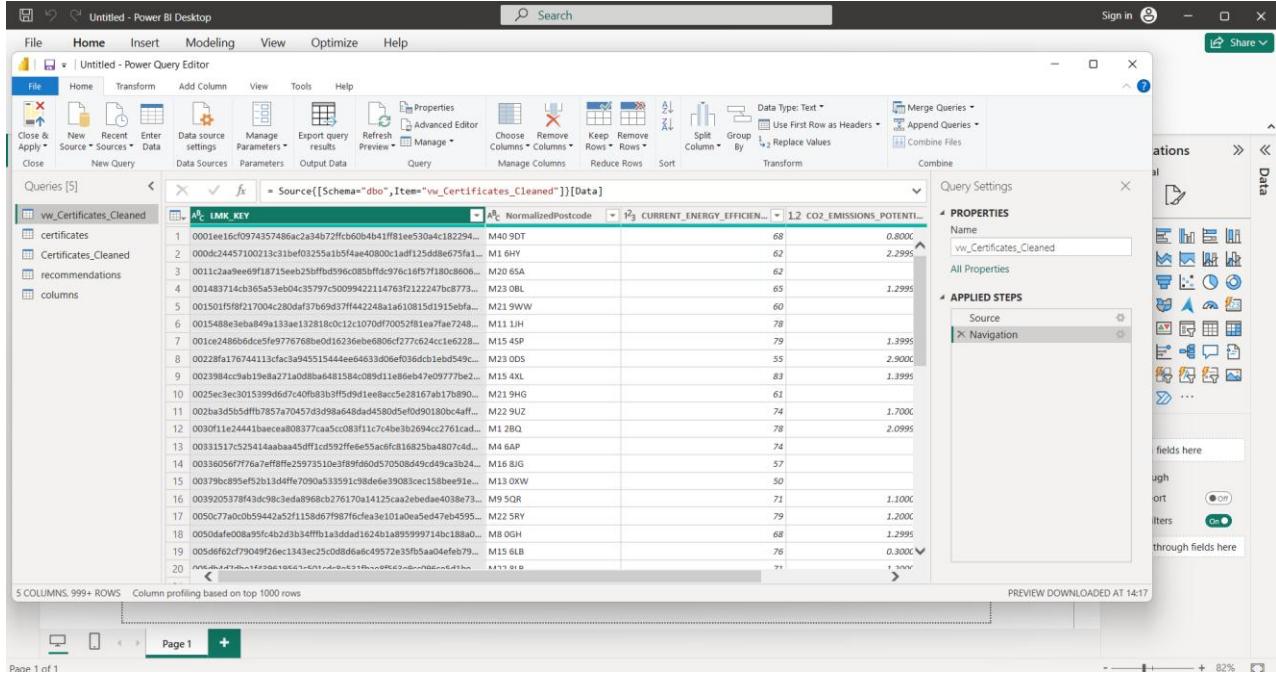


Figure 14: Power Query Editor

AVERAGE DAX FUNCTION

The mean value of chosen fields was computed with the AVERAGE DAX measure that gave a clear picture of the average level of performance.

This measure is used to compute the average of a numeric column. It helps in interpretation of key trends in the EPC Manchester data, including the mean energy rating, mean Co2 emissions or mean potential improvement score. This measure will assist in the identification of overall levels of performance and the comparison of energy efficiency in various groups of properties as a result of summarizing large volumes of data in to one representative value.

The screenshot shows the Power BI Desktop interface with the 'Measure tools' ribbon selected. A DAX formula is being typed into the formula bar:

```
Avg_CO2_Emissions = AVERAGE(Certificates_Cleaned[CO2_EMISI...]
```

The formula bar includes dropdowns for 'Name' (set to 'Avg_CO2_Emissions'), 'Format' (set to 'General'), 'Data category' (set to 'Uncategorized'), and 'Calculated column' (set to 'Auto'). Below the formula bar are sections for 'Structure', 'Formatting', 'Properties', and 'Calculations'. The main area displays a table with columns: LMK_KEY, ADDRESS1, ADDRESS2, ADDRESS3, POSTCODE, BUILDING_REFERENCE_NUMBER, CURRENT_ENERGY_RATING, POTENTIAL_ENERGY_RATING, and CURRI. The table contains several rows of data. To the right of the table, there is a vertical pane showing the context of the current measure, listing various table names and their corresponding columns.

Figure 15: AVERAGE DAX FUNCTION

SWITCH DAX FUNCTION

The assigning labels or categories according to certain conditions was performed by the SWITCH DAX function. Simplifies complex nested IF statements and simplifies the logic to manage which, through the simplification of these statements, aids in creating more accurate and clearer classifications in the dashboard.

This measure classifies or tags information according to certain requirements. In the EPC Manchester dataset, as an example a switch-based measurement can simply give each property an energy efficiency band based on its numeric EPC score. The measure compares every row to specific ranges and the relevant category is provided and enables the dashboard to group the properties, compare performance and view the results in a meaningful manner.

The screenshot shows the Power BI Desktop interface with the 'Measure tools' ribbon selected. In the formula bar, a DAX SWITCH function is being defined:

```
=SWITCH(TRUE(),
    'Certificates_Cleaned'[Avg_CO2_Emissions] >= 80, "Excellent",
    'Certificates_Cleaned'[Avg_CO2_Emissions] >= 60, "Good",
    'Certificates_Cleaned'[Avg_CO2_Emissions] >= 40, "Average",
    "Poor"
)
```

The Data pane on the right lists various columns and measures from the EPC Manchester dataset, such as LMK_KEY, ADDRESS1, ADDRESS2, ADDRESS3, POSTCODE, BUILDING_REFERENCE_NUMBER, CURRENT_ENERGY_RATING, POTENTIAL_ENERGY_RATING, and FLOOR_DESCRIPTION.

Figure 16: SWITCH DAX FUNCTION

DIVIDE DAX FUNCTION

To safely calculate ratios and percentages while avoiding division by zero errors, the DIVIDE DAX function was used. When compared to the conventional division operator, it yields cleaner, more dependable results, increasing the precision of dashboard metrics.

This measure can divide two numbers and in case the denominator maybe zero. Ratios that can be calculated using this based on DIVIDE in the EPC Manchester dataset include average Co2 emissions per property or energy improvement per unit. This makes there are no errors when dividing by zero and results are accurate thus reliable comparisons and calculations in the dashboard.

The screenshot shows the Power BI Desktop interface with the 'Measure tools' tab selected. A calculated column named 'Improvement_Potential_Percent' is being created. The formula is:

```

1 Improvement_Potential_Percent =
2 DIVIDE(
3     COUNTROWS(FILTER(Certificates_Cleaned, Certificates_Cleaned[CURRENT_ENERGY EFFICIENCY] < Certificates_Cleaned[POTENTIAL_ENERGY EFFICIENCY]))),
4     COUNTROWS(Certificates_Cleaned),
5     0
6 )

```

The 'Data' pane on the right lists various columns from the 'Certificates_Cleaned' table, including FLOOR_DESCRIPTION, FLOOR_ENERGY_EFF, and HEATING_COST_CURRENT.

Figure 17: DIVIDE DAX FUNCTION

This picture shows how a calculated column called Co2_Reduction_Potential was created in the Certificates_Cleaned table in Power BI Desktop.

The screenshot shows the Power BI Desktop interface with the 'Measure tools' tab selected. A calculated column named 'Cost_Per_Property' is being created. The formula is:

```

1 Cost_Per_Property =
2 DIVIDE(
3     SUM(recommendations[INDICATIVE_COST]),
4     DISTINCTCOUNT(recommendations[LMK_KEY]),
5     0
6 )

```

The 'Data' pane on the right lists various columns from the 'Certificates_Cleaned' table, including ADDRESS, ADDRESS1, ADDRESS2, ADDRESS3, and CO2_EMISSIONS_POTENTIAL.

Figure 18: Co2_Reduction_Potential

Table: Certificates_Cleaned (325,686 rows) Column: CO2_Reduction_Potential (0 distinct values)

Figure 19: Cleaned Co2_Reduction_Potential

COUNT_ROW DAX FUNCTION

The number of records in a table or filtered dataset was determined using the COUNT_ROW DAX function, this metric aids in measuring the amount of data and supports dashboard insights.

This measure titled Solar_Heating_Count counts the amount of property certificates with the Solar Water Heating Flag set to “Yes”. This measure helps in measuring the adoption of solar water-heating systems and allows comparison across regions, time periods, or other dimensions in the EPC Manchester analysis by limiting the cleaned dataset for solar equipped properties and counting each valid record.

Figure 19: COUNT_ROW DAX FUNCTION

DISTINCTCOUNT DAX FUNCTION

The number of distinct values in a column was counted using the DISTINCTCOUNT DAX function. The DISTINCTCOUNT function, which reduces duplicates before finding the total number of unique values in the selected column, is applied in the EPC Manchester dataset to measure this.

It makes it easier to identify how many distinct properties, certificates, or categories actually exist and presents a clear and more accurate picture of the data without repetition.

Untitled - Power BI Desktop

File Home Help Table tools Measure tools

Name: Solar_Heating_Target Format: General Data category: Uncategorized

Home table: Certificates_Cleaned

Structure: Formatted Properties: Calculations

Data

1 Solar_Heating_Target = 0.2 * DISTINCTCOUNT(Certificates_Cleaned[LMK_KEY])

LMK_KEY	ADDRESS1	ADDRESS2	ADDRESS3	POSTCODE	BUILDING_REFERENCE_NUMBER	CURRENT_ENERGY_RATING	POTENTIAL_ENERGY_RATING	CURR
1002441598552017121514255994939713	38, Newhey Avenue			M22 8LQ	7320843178	D	B	
1014532939324201809181719571489688	2, Kemmel Avenue			M22 4QX	4229434178	D	B	
101421995902201711081638160488403	74, Rushey Road			M22 8ER	64534178	D	B	
1022925071332017113016504873978401	33, Gorsey Road			M22 9HQ	7153494178	D	B	
1031662659302018031316494819589878	142, Gladstone Road			M22 9GS	353955178	D	B	
102944203233201712061335619078198	19, Benchill Road			M22 8LB	3879145178	C	B	
102969879181201803081710594080710	58, Gladstone Road			M22 9GG	7044145178	C	B	
1037103579922017113017062975298063	48, Panfield Road			M22 9NQ	6796795178	D	B	
1038469909492015102116532311652198	45, Broadoak Road			M22 9NE	5912406178	D	B	
1042706635652017113017133599390813	44, Shayfield Drive			M22 8FA	8321536178	D	B	
1043386039302018032116492417682498	6, Brookfield Garden			M22 9PX	5311836178	C	B	
1038092809502017091317140811639278	113, Baxter Gardens			M23 1BH	594306178	C	B	
1042544759962017113016515716228263	17, Nearbrook Road			M22 9WT	6120536178	D	B	
1042952403832018073014521680078205	9, Merewood Avenue			M22 8DP	560246178	C	B	
1045926459002018012518161917682158	111, Broadoak Road			M22 9PW	4088856178	D	B	
106257722521201802171309059780918	29, Benchill Road			M22 8FJ	9859577178	D	B	
10521618703320171023120587328994	6, Brookfield Garden			M22 8BE	7403996178	D	B	
1080627888752019110514484090019218	50, Seabrook Road			M40 2RZ	2864809178	C	B	
1085451169262018032116441899189878	11, Chesham Avenue			M22 9PX	7900639178	C	B	
1050822255652018030616272299080112	54, Gladstone Road			M22 9GG	1261596178	C	B	
1051636342552018022117115090280016	111, Woodhouse Lane			M22 8JX	619207178	D	B	
10497769263201602147111282978204	31, Lawton Moor Road			M23 0QA	9412286178	D	B	
1070168219762017113015541678108643	12, Barncroft Garden			M22 8ET	7501228178	C	B	
1087559264752018013163439508017	42, Newhey Road			M22 8LZ	4787138178	D	B	
108331402451201711301701399203014	109, Hollyhedge Road			M22 8LJ	1733459178	C	B	
				M22 8LR	1425429178	D	B	

Figure 20: DISTINCTCOUNT DAX FUNCTION

Untitled - Power BI Desktop

File Home Help Table tools Measure tools

Name: Cost_Per_Property Format: Whole number Data category: Uncategorized

Home table: Certificates_Cleaned

Structure: Formatted Properties: Calculations

Data

1 Cost_Per_Property =

```

2 DIVIDE(
3 | SUM('recommendations'[INDICATIVE_COST]),
4 | DISTINCTCOUNT('recommendations'[LMK_KEY]),
5 | 0
6 )

```

LMK_KEY	ADDRESS1	ADDRESS2	ADDRESS3	POSTCODE	BUILDING_REFERENCE_NUMBER	CURRENT_ENERGY_RATING	POTENTIAL_ENERGY_RATING	CURR
1002441598552017121514255994939713	38, Newhey Avenue			M22 8LQ	7320843178	D	B	
1014532939324201809181719571489688	2, Kemmel Avenue			M22 4QX	4229434178	D	B	
101421995902201711081638160488403	74, Rushey Road			M22 8ER	64534178	D	B	
1022925071332017113016504873978401	33, Gorsey Road			M22 9HQ	7153494178	D	B	
1031662659302018031316494819589878	142, Gladstone Road			M22 9GS	353955178	D	B	
102944203233201712061335619078198	19, Benchill Road			M22 8LB	3879145178	C	B	
102969879181201803081710594080710	58, Gladstone Road			M22 9GG	7044145178	C	B	
1037103579922017113017062975298063	48, Panfield Road			M22 9NQ	6796795178	D	B	
1038469909492015102116532311652198	45, Broadoak Road			M22 9NE	5912406178	D	B	
1042706635652017113017133599390813	44, Shayfield Drive			M22 8FA	8321536178	D	B	
1043386039302018032116492417682498	6, Brookfield Garden			M22 9PX	5311836178	C	B	
1038092809502017091317140811639278	113, Baxter Gardens			M23 1BH	594306178	C	B	
1042544759962017113016515716228263	17, Nearbrook Road			M22 9WT	6120536178	D	B	
1042952403832018073014521680078205	9, Merewood Avenue			M22 8DP	560246178	C	B	
1045926459002018012518161917682158	111, Broadoak Road			M22 9PW	4088856178	D	B	
106257722521201802171309059780918	29, Benchill Road			M22 8FJ	9859577178	D	B	
10521618703320171023120587328994	6, Brookfield Garden			M22 8BE	7403996178	D	B	
1080627888752019110514484090019218	50, Seabrook Road			M40 2RZ	2864809178	C	B	
1085451169262018032116441899189878	11, Chesham Avenue			M22 9PX	7900639178	C	B	
1050822255652018030616272299080112	54, Gladstone Road			M22 9GG	1261596178	C	B	
1051636342552018022117115090280016	111, Woodhouse Lane			M22 8JX	619207178	D	B	
10497769263201602147111282978204	31, Lawton Moor Road			M23 0QA	9412286178	D	B	
105552458421201711301438379270615	54, Hayley Road			M22 8ET	6565427178	C	B	

Figure 21: DISTINCTCOUNT DAX FUNCTION 2

4.3 Power BI Data Visualization

Key insights from the cleaned EPC Manchester dataset were visually presented in the Power BI dashboard through the use of interactive charts, cards, and slicers. By emphasizing significant trends, comparisons, and patterns among Manchester properties, each visual aids user in rapidly comprehending energy performance data.

Clearer and more informed decision-making is supported by the dashboard's interactive features, which include drilldowns and filters that make it simple to delve deeper into the data and explore particular areas.

FINAL DASHBOARD

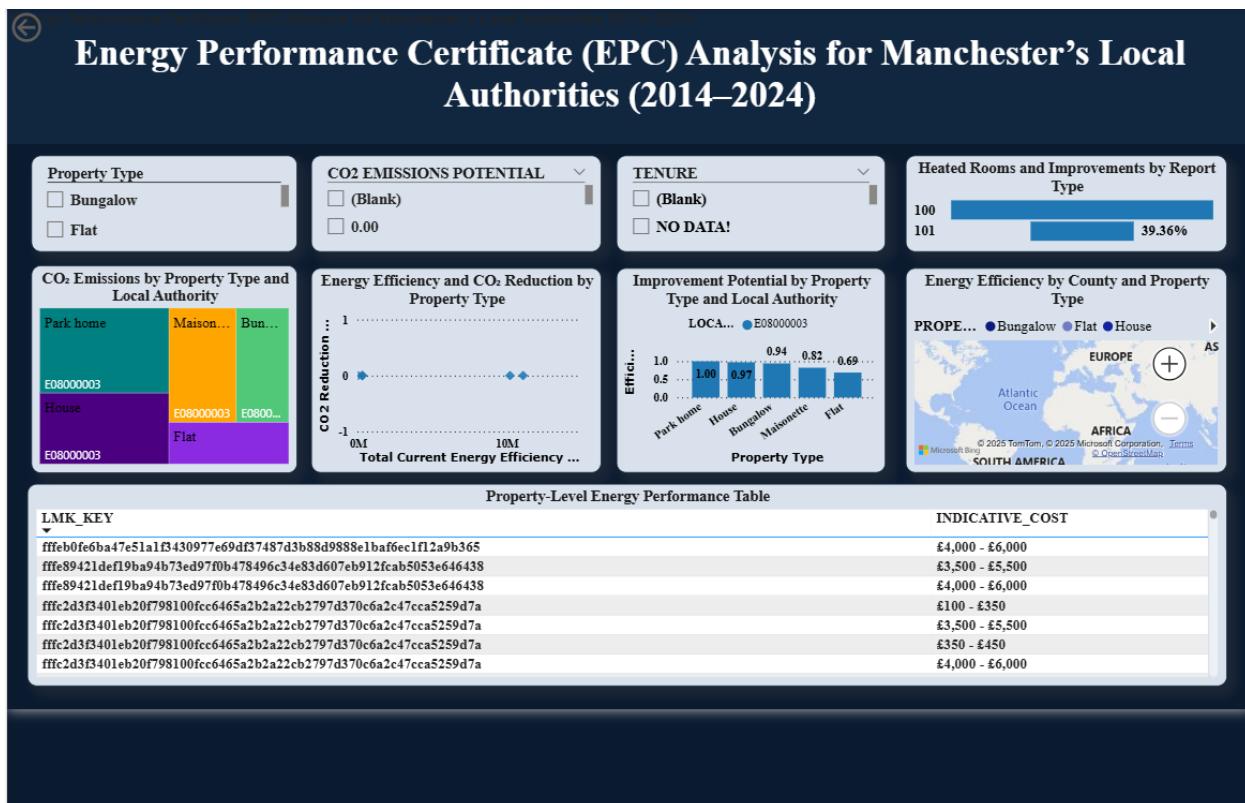


Figure 22: Dashboard

4.4 Explanation of Dashboard Visuals

Filled Map

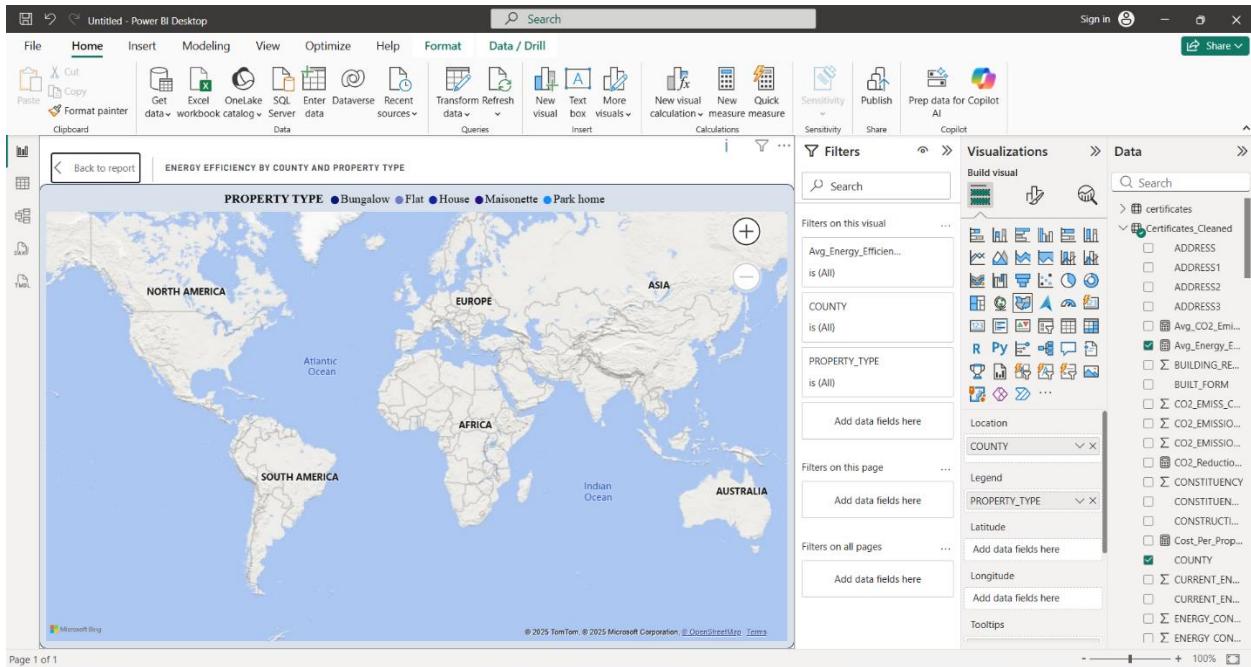


Figure 23: Filled Map

A data foundation is created using the certificates_cleaned table, which contains property level energy ratings. A crucial dax calculation that calculates the maximum savings per property is the Co2 Reduction Potential. By aggregating this metric and displaying it on a Filled Map that is divided by COUNTY and PROPERTY_TYPE, the analysis then directly identifies the specific geographic regions and housing types with the highest potential for reduction efforts.

Funnel

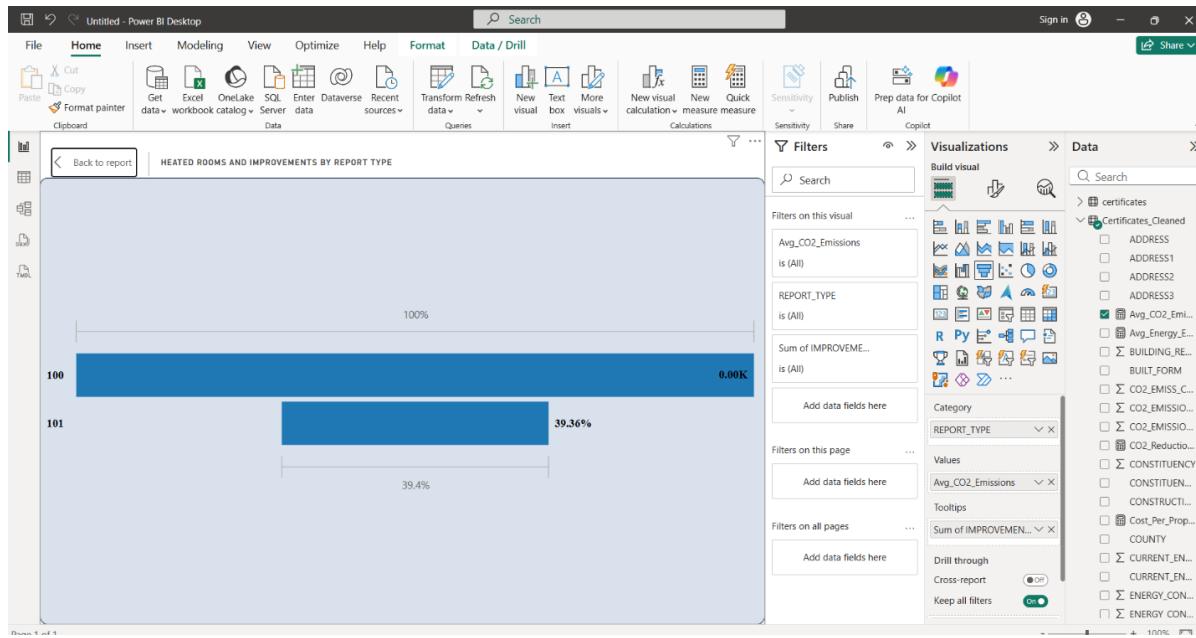
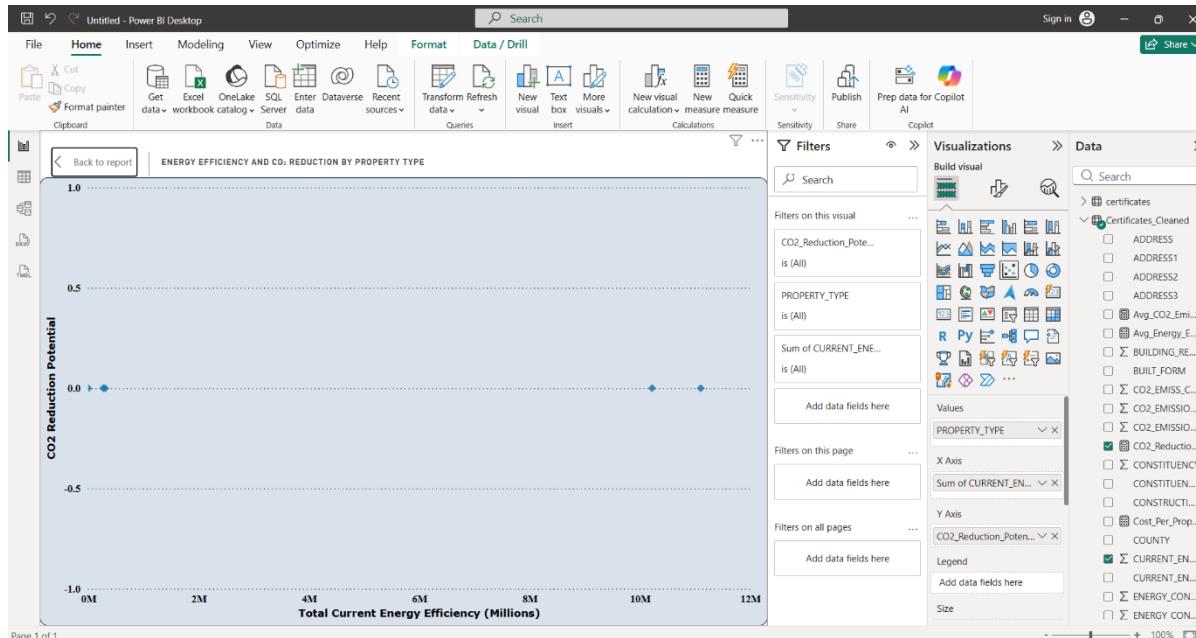


Figure 24: Funnel

This graph is used as a visual aid for qualification or conversion. Using the raw data to the report it reveals that various report categories (defined by REPORT_TYPE) add to the overall improvements. When the criteria move from the general baseline (stage 100) to the specific group (stage 101), the graphic emphasizes a significant limitation or feature.

Scatter Chart



The overall conclusion is that, despite significant differences in the total energy efficiency reported across property types, the immediate remaining reduction opportunity for these aggregated groups appears marginal.

Slicers

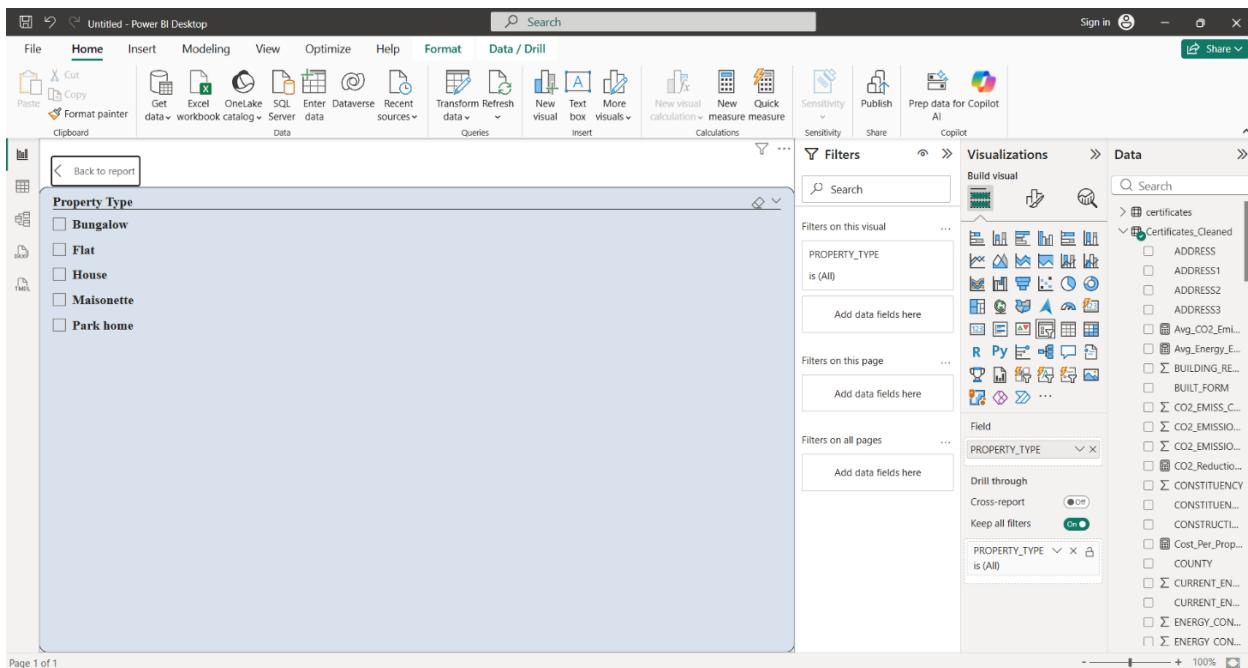


Figure 26: Property Slicer

The guidance of this slicer, users can narrow down the dashboard according to different types of properties found in the EPC Manchester dataset, such as Park Home, Bungalow, Flat, House, and Maisonette. By choosing one or more categories, users may concentrate the analysis on particular asset types. This makes it possible to get specific insights into energy performance and other important statistics connected to each kind of property.

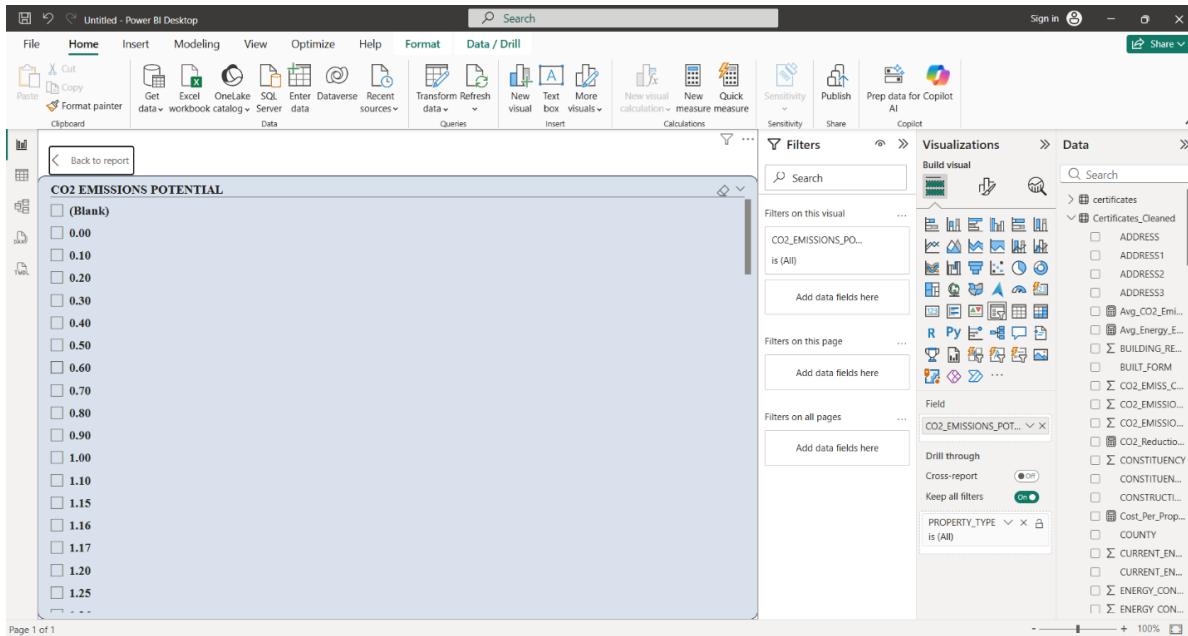


Figure 27: Co2 Emissions Slicer

To filter the dataset using the estimated Co2 emissions value of each property, can use the Co2 Emissions potential slicer. It makes environmental analysis much more comprehensive by assisting in comparing the degree of energy efficiency in different ranges of emissions. This is an instrument that enables decision makers, analysts and homeowners to easily calculate the properties that have a higher or lesser environmental impact.

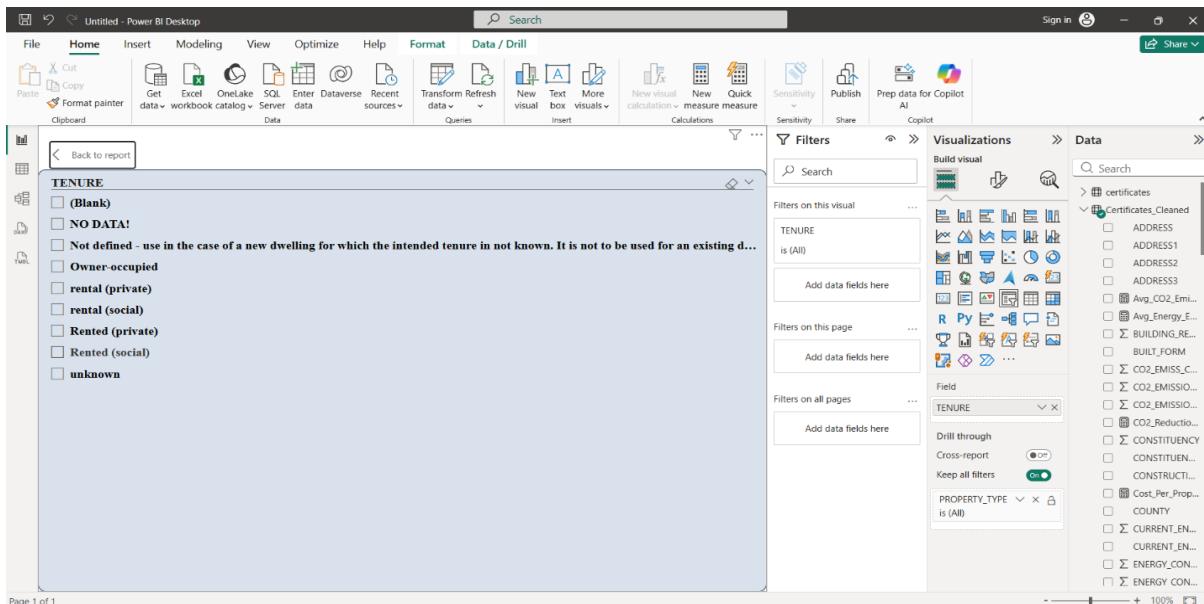


Figure 28: Tenure Slicer

The Tenure Slicer offers filtering options according to each property's occupancy status, including owner -occupied, private/social rental, and unknown categories. This enables users to investigate how different living arrangements affect a property's energy performance. By exposing trends related to ownership, rental markets and unclear tenure types, it also aids in socioeconomic analysis.

Table

PROPERTY-LEVEL ENERGY PERFORMANCE TABLE

LMK_KEY	INDICATIVE_COST
fffe0f6e9a47e51a153430977e69df37487d3b88d9888e1bafe6ee1f12a9b365	£4,000 - £6,000
fffe89421de19ba94b73e97f0b478496c34e83d607eb912ca5b053e646438	£3,500 - £5,500
fffe89421de19ba94b73e97f0b478496c34e83d607eb912ca5b053e646438	£4,000 - £6,000
fffc2d3f3401eb20f798100fc6e465a2b2a22cb2797d3706e6a2e47cc5259d7a	£100 - £350
fffc2d3f3401eb20f798100fc6e465a2b2a22cb2797d3706e6a2e47cc5259d7a	£3,500 - £5,500
fffc2d3f3401eb20f798100fc6e465a2b2a22cb2797d3706e6a2e47cc5259d7a	£550 - £450
fffc2d3f3401eb20f798100fc6e465a2b2a22cb2797d3706e6a2e47cc5259d7a	£4,000 - £6,000
fffc2d3f3401eb20f798100fc6e465a2b2a22cb2797d3706e6a2e47cc5259d7a	£500 - £1,500
fffc2d3f3401eb20f798100fc6e465a2b2a22cb2797d3706e6a2e47cc5259d7a	£800 - £1,200
fffaeaddcdce4729d1f76895ea9722bd3dfe4f53ca47221d07a67d3e6309b5	£4,000 - £6,000
ff991c4f433d33a54c19191f104008a7fbad1919da1ade48ee1577d7ca	£1,200 - £1,800
ff9903a3c952004c4954cea844378059fd5b7dec28393192a6da333537a95	£3,500 - £5,500
ff9903a3c952004c4954cea844378059fd5b7dec28393192a6da333537a95	£4,000 - £6,000
ff8bedbe3f3d02b835435e4356580549a93dc8d3a699782508e90959ee32a	£1,200 - £1,800
ff8bedbe3f3d02b835435e4356580549a93dc8d3a699782508e90959ee32a	£30
ff82d931d0ce9bcb59da3e6a2ae1c12562443310a2eb8963a8a410daa7	£3,500 - £5,500
ff82d931d0ce9bcb59da3e6a2ae1c12562443310a2eb8963a8a410daa7	£4,000 - £6,000
ff433c977c71debf86e52c169e23e9eb85817377eade62e805b4e94137e26f8	£10
ff433c977c71debf86e52c169e23e9eb85817377eade62e805b4e94137e26f8	£4,000 - £6,000
ff37456e672de6603a40f07e3fcfa0124cb4e4f92d7c1fb0be5c5840838ef0	£4,000 - £6,000
ff37456e672de6603a40f07e3fcfa0124cb4e4f92d7c1fb0be5c5840838ef0	£5,000 - £8,000
ff32b8e10779a8287000bda38e6b7aeecc75d1506761ce3a5b4134920c7e	£4,000 - £6,000
ff316c6daaa7a0073c36b7a68572d3ce9b08deceb2d1d99662d8c7400ee5564	£5,000 - £10,000
ff316c6daaa7a0073c36b7a68572d3ce9b08deceb2d1d99662d8c7400ee5564	£8,000 - £10,000
ff1fa335799f0fa1c048686c1e4f8fb30b0eb4eac92253d79374616099c	£3,500 - £5,500
ff1fa335799f0fa1c048686c1e4f8fb30b0eb4eac92253d79374616099c	£30
ff1fa335799f0fa1c048686c1e4f8fb30b0eb4eac92253d79374616099c	£4,000 - £6,000

Figure 29: Table

This table shows the possible cost for energy efficiency upgrades by connecting each property's LMK_KEY to its estimated INDICATIVE COST range. It provides a financial viewpoint that aids in evaluating both the potential for Co2 reduction and cost effectiveness.

Tree Map

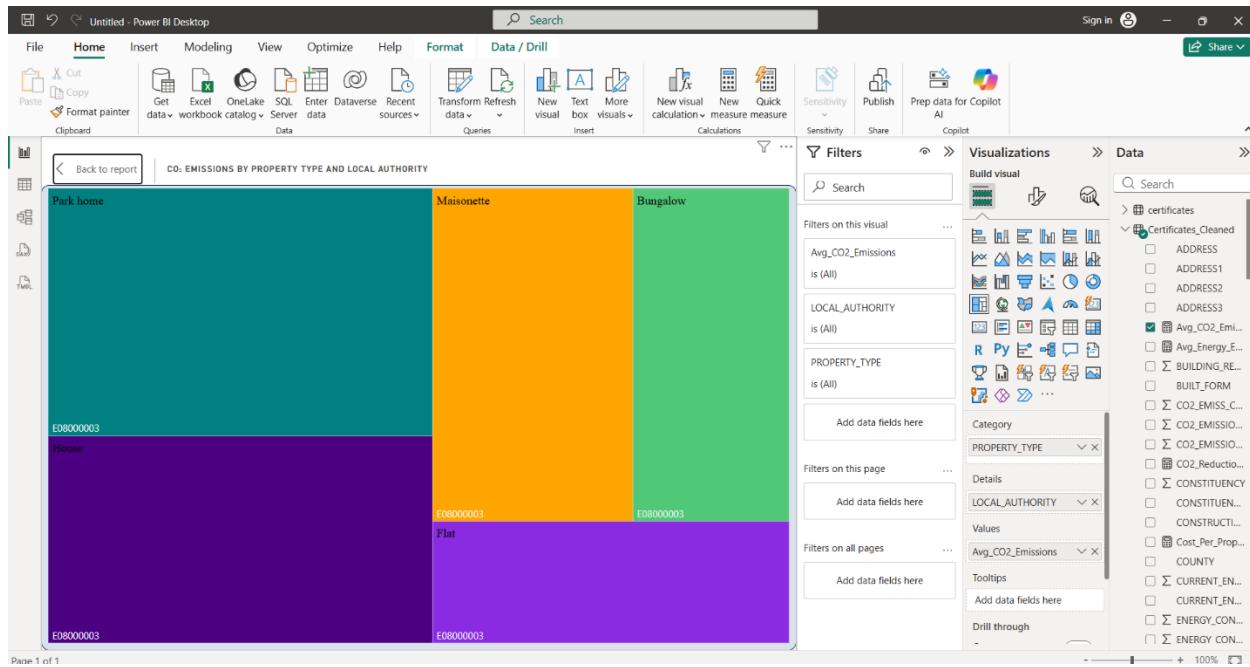


Figure 30: Tree Map

By clearly displaying which property segments generates the highest emissions, this tree map helps in resource prioritization. It recognizes important goals for energy efficiency measures and directs the most efficient use of funds based on the estimated INDICATIVE COST.

4.5 The Connection of the SQL Server with the Power BI

It is easier to analyze the data as validated and it minimized the manual cleaning in Power BI, Provided the steady and updated data by automatic refreshing and enhanced the overall performance.

4.6 Interpretation of the Dashboard

This dashboard allows seeing a comprehensive overview of trends in Energy Performance Certificate (EPC) costs in local authorities throughout Manchester in 2014 – 2024, showing the role of various types of property in the energy consumption and carbon dioxide emissions. The visualizations demonstrate that some types of properties (flats and maisonettes) are more associated with a greater number of emissions in particular local authority areas, whereas the scatter and bar graphs demonstrate significant differences in the current energy efficiency and the potential of its improvement. The interactive filters enable the user to inquire the patterns by property type, tenure and Co2 emissions, giving information on where the targeted inventions may offer the most significant carbon reduction benefits. In general, the dashboard is a visible tool to detect performance gaps and make data-driven decisions to improve the housing conditions to be more energy efficient throughout the region.

4.7 Application of Advanced Power BI Features

Application of advanced Power BI were used to improve the analysis of EPC Manchester Data. DAX was used to create measures and calculated columns to derive insights as average energy scores, comparison of emissions and rating classification. The data was cleaned and structured using Power query transformations thus the errors were removed, the fields were split, and the formats standardized. Interactive capabilities that included slicers, drill through pages were also added to enable users to take a closer look at the performance of the property.

4.8 Key Insights and actionable recommendations

According to the EPC Manchester analysis, a sizable percentage of properties are in the mid-to-lower energy efficiency bands, suggesting that there is a lot of room for improvement. Due in large part to ineffective insulation and heating systems, older properties and some types of properties consistently have lower EPC ratings. Stakeholders are urged to give priority to specific retrofit projects like enhancing roof and wall insulation, switching to energy-efficient heating systems, and encouraging the use of renewable technologies. While landlords and homeowners can use the results to lower energy costs, increase compliance, and boost long-term property value, policymakers and local authorities can use these insights to create incentive programs and awareness campaigns.

5 Conclusion

This comparison of EPC records on local authorities within Manchester between 2014 to 2024 reveals that there are evident differences in the property-level energy efficiency, Co2 emissions and potential improvements. The dashboard indicates that although certain types of properties perform fairly, others are highly effective in the domain of carbon reduced with the help of specific upgrades and policy solutions. The visualization will be an effective decision-making tool as it can be used to identify which areas the company should focus on to implement sustainability strategies in the future by filtering, geographical mapping, and performance indicators. In general, the results support the necessity of the data-driven planning to enhance the housing energy standards, decrease emissions, and facilitate the long term environmental and energy effectiveness objectives of Manchester.