



General Sir John Kotelawala Defence University  
Faculty of Management, Social Sciences and Humanities  
Department of Languages

BSc in Applied Data Science Communication  
Group Assignment

- P Laksia – D/ADC/23/0013
- PRM Perera – D/ADC/23/0031
- MM Jayasinghe – D/ADC/23/0035
- RYN Sanduprabha – D/ADC/23/0046

Advanced SQL and Cloud Databases / LB2224

Year 2: Semester 2

Assignment number 2

13.10.2024

# **Content**

## **1. Task 01**

### **1.1. Introduction**

### **1.2. About the Dataset**

### **1.3. Database Utilization**

### **1.4. Key Patterns and Insights**

### **1.5. Recommendations**

### **1.6. Conclusion**

## **2. Task 02**

### **2.1. Introduction**

### **2.2. About the Dataset**

### **2.3. Database Utilization**

### **2.4. Key Patterns and Insights**

### **2.5. Recommendations**

### **2.6. Conclusion**

## **3. References**

## **4. Appendices**

# Task 01

# Child Wellbeing Monitor



## **1.1. Introduction**

No Poverty. According to Goal 1 of Sustainable Development Goals introduced by the United Nations, eradicating extreme poverty for all people everywhere all around the world in all forms by the end of the year 2023 is a pivotal goal (United Nations, 2023a). Poverty is the state of being extremely poor. It is the lack of financial and other essential resources not allowing them to satisfy their basic needs which ultimately reflects the living standards of the community.

Child poverty remains one of the world's most pressing concerns especially in low-income nations where economic instability, social inequality, and restricted access to resources have significant impacts on children's well-being. To address this issue, international organizations use powerful data-driven technologies to examine and assess the situation of child poverty.

The Child Wellbeing Monitor is a simple, low-cost reporting solution for analyzing and evaluating child poverty. It utilizes charts, tables and other visualization methods to analyze and measure the child poverty of two chosen lower income countries – in this scenario, Vietnam and Ethiopia. The report focuses on analyzing five poverty indicators namely - Income per capita, access to education, health expenditure per child, standard of living based on the access to pure clean water and the malnutrition rate of the two countries Vietnam and Ethiopia.

## **1.2. About the Dataset.**

The report considers of two datasets for the analysis where each dataset represents a specific country – Vietnam and Ethiopia. The datasets are extracted from the study **Young Lives: an International Study of Childhood Poverty: Rounds 1-5 Constructed Files, 2002-2016.**

The **first dataset** which provides details about the surveys carried out in **Vietnam** which reflects on a varied range of Vietnam's culture, geographical and social contexts which faces various issues and challenges such as large debt burdens, conflicts and exposure to environmental hazards such as floods and droughts. The dataset contains **213 variables** and **15,000 cases**. Whereas, the **second dataset** provides details on regards to the survey information collected from **Ethiopia**. The corresponding dataset contains **214 variables** and **14,995 cases**. The metadata and the relevant datasets are obtained from the following link: <https://doi.org/10.5255/UKDA-SN-7483-5>

A total of **23 variables** were extracted separately from each dataset to measure the poverty level of the two countries. The details of the variables are as follows;

1. **childid** = Child ID
2. **yc** = Cohort details of the participant. [Younger cohort=1; Older cohort=0]
3. **agemon** = Child's age (in months).
4. **wi\_new** = Wealth index.
5. **hq\_new** = Housing quality index.
6. **sv\_new** = Access to services index.
7. **ownlandhse** = Household owns land where house is on. [Yes = 1.0; No = 0.0]
8. **ownhouse** = Household own the house. [Yes = 1.0; No = 0.0]
9. **hschool** = Hours/day spent at school.
10. **caredu** = Caregiver's level of education.
11. **dadedu** = Father's level of education.
12. **momedu** = Mother's level of education.
13. **headedu** = Household head's level of education.
14. **chrephealth1** = Number of correct responses to 5 reproductive health statements.
15. **chrephealth2** = Child knows condom can prevent disease through sex. [Yes = 1.0; No = 0.0]
16. **chrephealth3** = Child knows healthy-looking person can pass on a disease sex. [Yes = 1.0; No = 0.0]
17. **chrephealth4** = Child's source of condom. [Shop or street vendor = 1.0; Family planning services or health facility = 2.0; Other = 3.0; I do not know what a condom is/I do not know where to get them = 4.0]
18. **chhealth** = Child's health in general. [Very poor = 1.0; Poor = 2.0; Average = 3.0; Good = 4.0; Very good = 5.0]
19. **drwaterq\_new** = Access to safe drinking water. [Yes = 1.0; No = 0.0]
20. **chweight** = Child's weight (kg).
21. **chheight** = Child's height (cm).
22. **bmi** = Calculated BMI=weight/squared(height).
23. **underweight** = Low weight for age. [Not underweight = 0.0; Moderately underweight = 1.0; Severely underweight = 2.0]

### 1.3. Database Utilization

The creation of the database, the data cleaning process and the creation of the views were carried out in the **SQL Server** where the creation of the five meaning summarized reports were carried out in the **Power BI Report Builder**. Initially, data cleaning and transformation was supposed to be carried out in **Power BI**, however since the utilization of the **Premium package** is essential for the task and due to the inability of the acquisition since the trial basis was already carried during the Workshop series, the data cleaning and transformation were all carried out in the SQL Server.

The steps taken for the modeling of the database for the analysis of the poverty index criteria are as follows;

#### *Steps to import the data from Excel to the SQL Server*

- After downloading the compressed zip file which consists of the metadata and the datasets, extract and preview the datasets and metadata separately to understand them.
- Choose the necessary datasets that will be considered for the analysis – Vietnam and Ethiopia datasets to be imported for the database in SQL server instance in SSMS.
- Create a database called **Child\_Welbeing\_Monitor**.
- Right click the database and select Tasks.
- Move to Import Data.

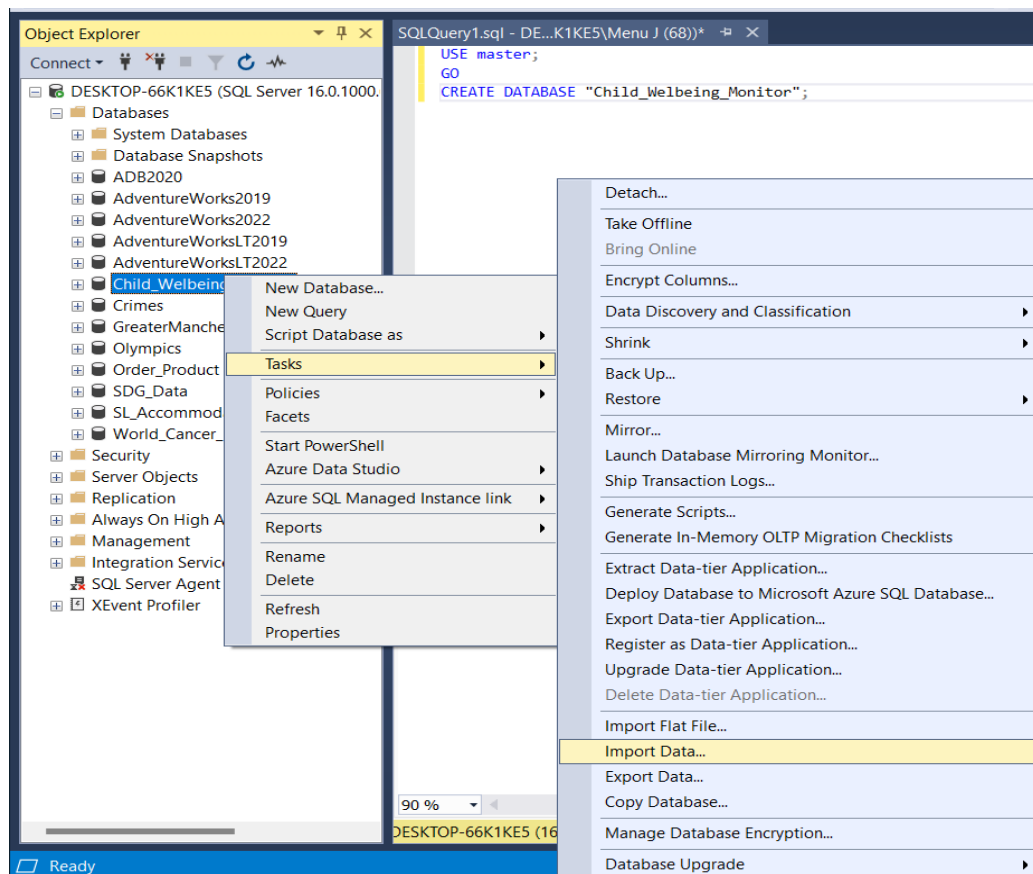


Figure 01

- Now Click next on Import data and Export Wizard welcome page.



Figure 02

- Select flat file source as the data source and enter or browse for the file.

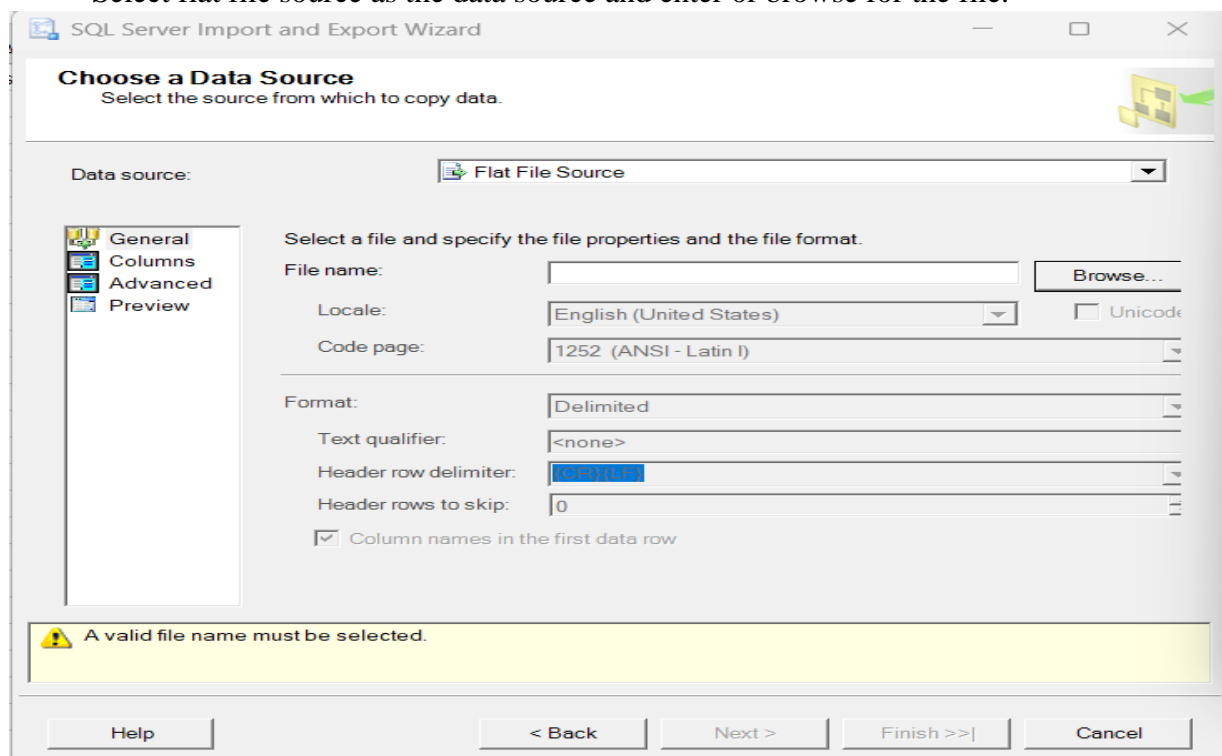


Figure 03

- Selecting the datasets to be imported.

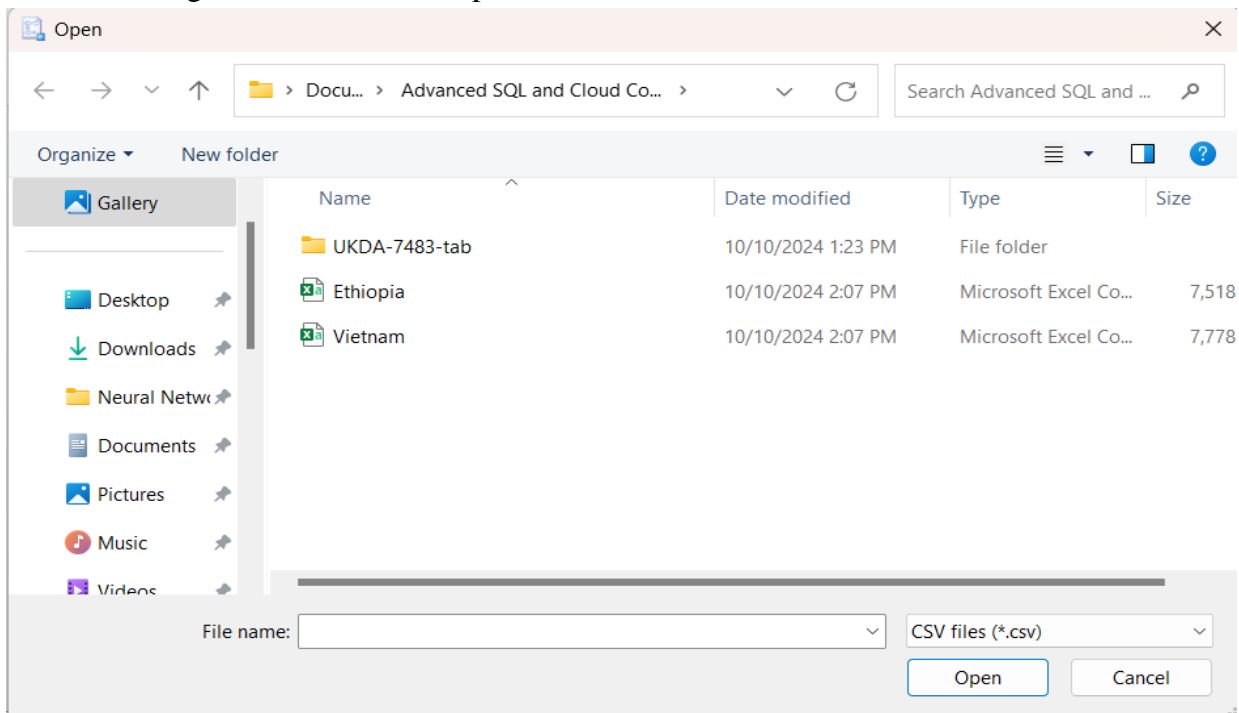


Figure 04

- Go to Advanced and change all the column widths from 50 to 100.

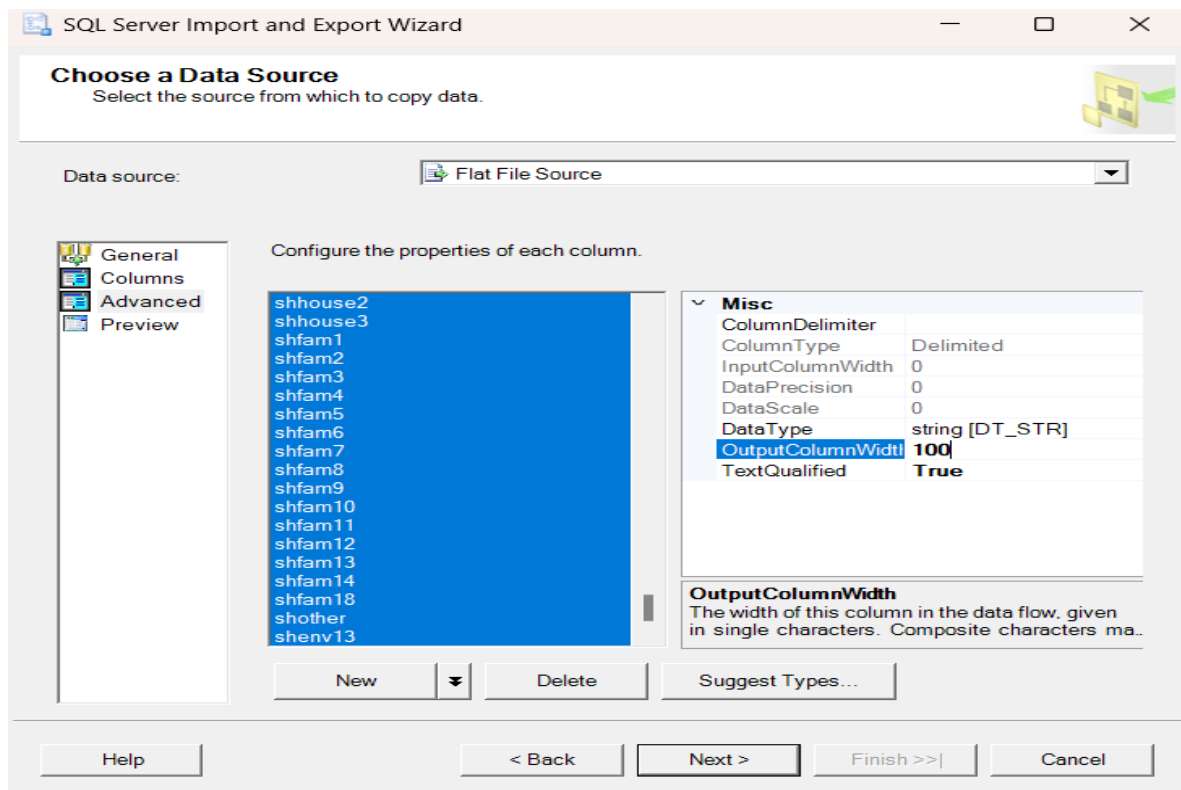
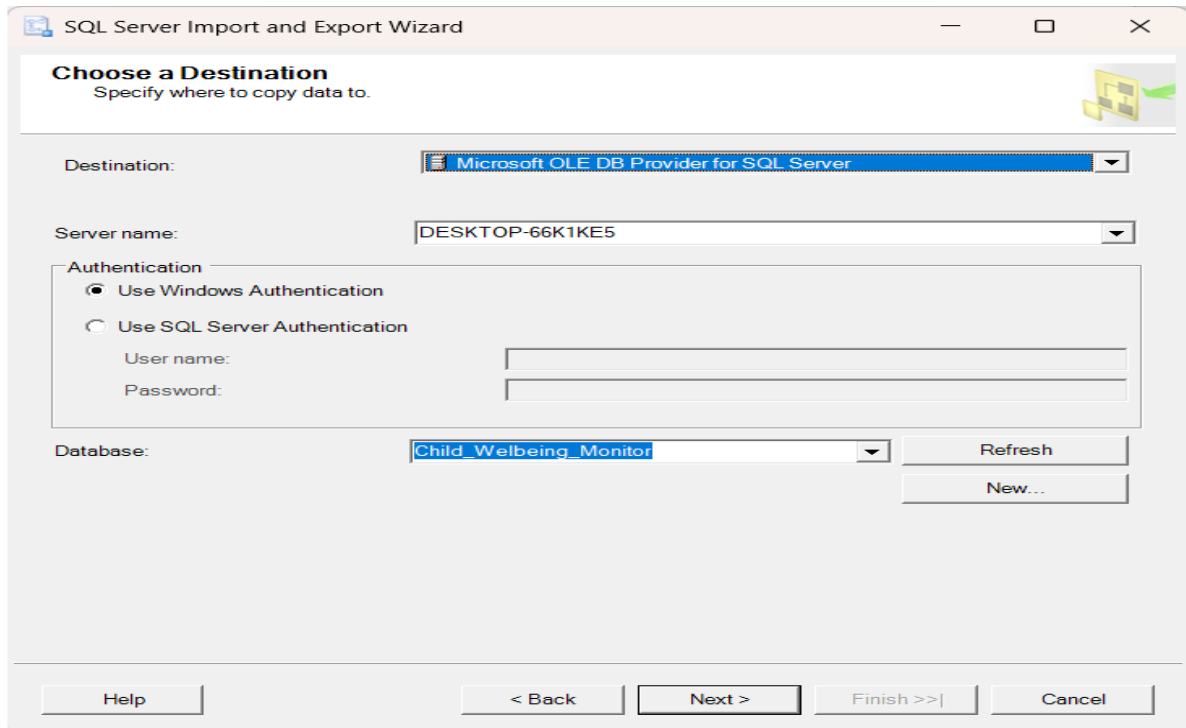


Figure 05



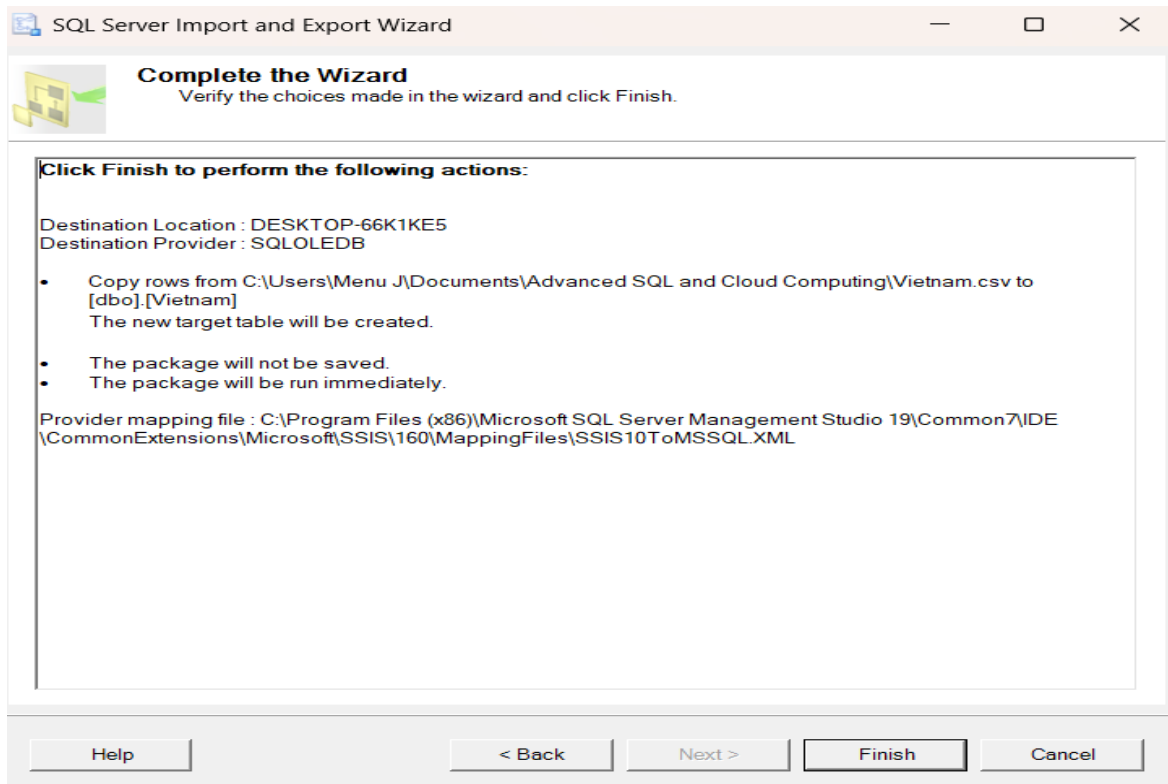
- Click Next to move forward.



The screenshot shows the 'Choose a Destination' step of the SQL Server Import and Export Wizard. The window title is 'SQL Server Import and Export Wizard'. The main heading is 'Choose a Destination' with the instruction 'Specify where to copy data to.' Below this, there are several fields: 'Destination' is set to 'Microsoft OLE DB Provider for SQL Server'; 'Server name' is 'DESKTOP-66K1KE5'; 'Authentication' has 'Use Windows Authentication' selected; 'Database' is 'Child\_Welbeing\_Monitor'. There are 'Refresh' and 'New...' buttons next to the database field. At the bottom, there are 'Help', '< Back', 'Next >', 'Finish >>|', and 'Cancel' buttons.

*Figure 06*

- Click Next and Accept the default.
- Click Finish.



The screenshot shows the 'Complete the Wizard' step of the SQL Server Import and Export Wizard. The window title is 'SQL Server Import and Export Wizard'. The main heading is 'Complete the Wizard' with the instruction 'Verify the choices made in the wizard and click Finish.' Below this, there is a box titled 'Click Finish to perform the following actions:' containing the following text: 'Destination Location : DESKTOP-66K1KE5', 'Destination Provider : SQLOLEDB', a list of actions (Copy rows from C:\Users\Menu J\Documents\Advanced SQL and Cloud Computing\Vietnam.csv to [dbo].[Vietnam], The new target table will be created., The package will not be saved., The package will be run immediately.), and 'Provider mapping file : C:\Program Files (x86)\Microsoft SQL Server Management Studio 19\Common7\IDE\Extensions\Microsoft\SSIS\160\MappingFiles\SSIS10ToMSSQL.XML'. At the bottom, there are 'Help', '< Back', 'Next >', 'Finish', and 'Cancel' buttons.

*Figure 07*

- The execution dialog box appears, if all the data has loaded and the execution was successful click close.

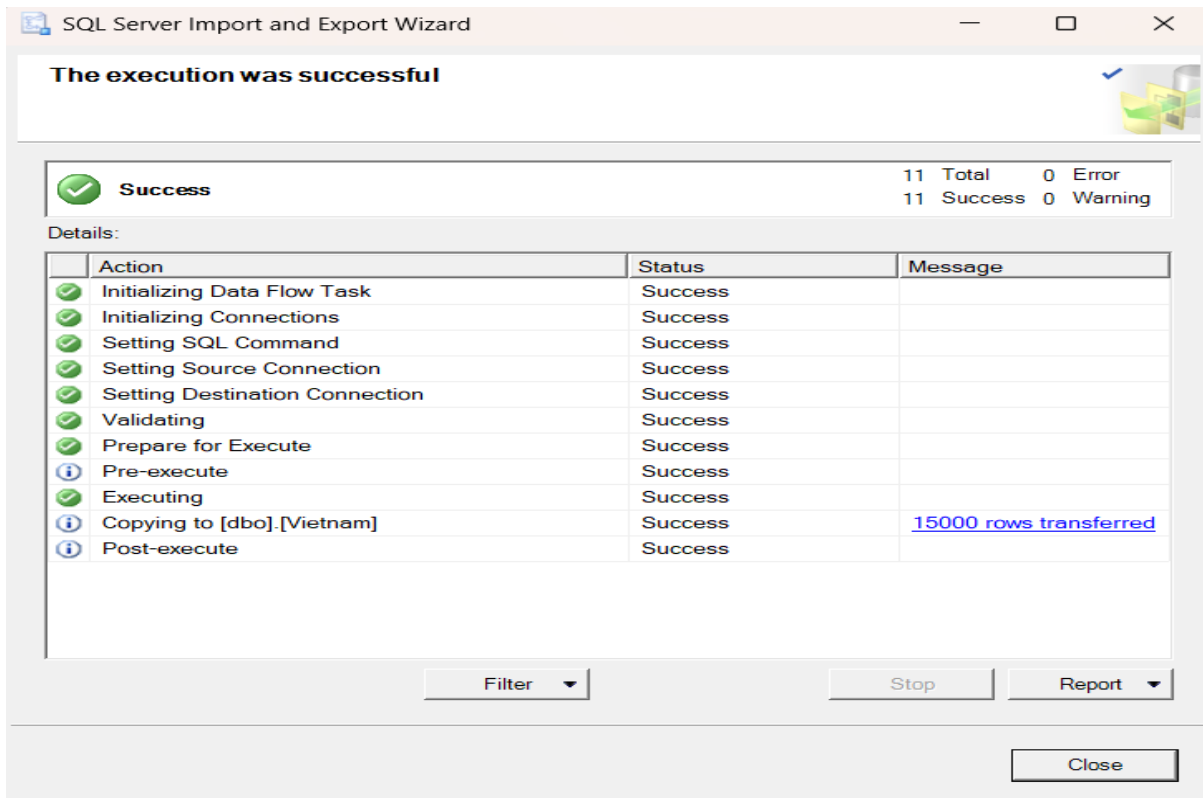


Figure 08

- Repeat the same procedure for the other dataset well.
- In a new query, type and execute the query to see the imported data for all the imported datasets.

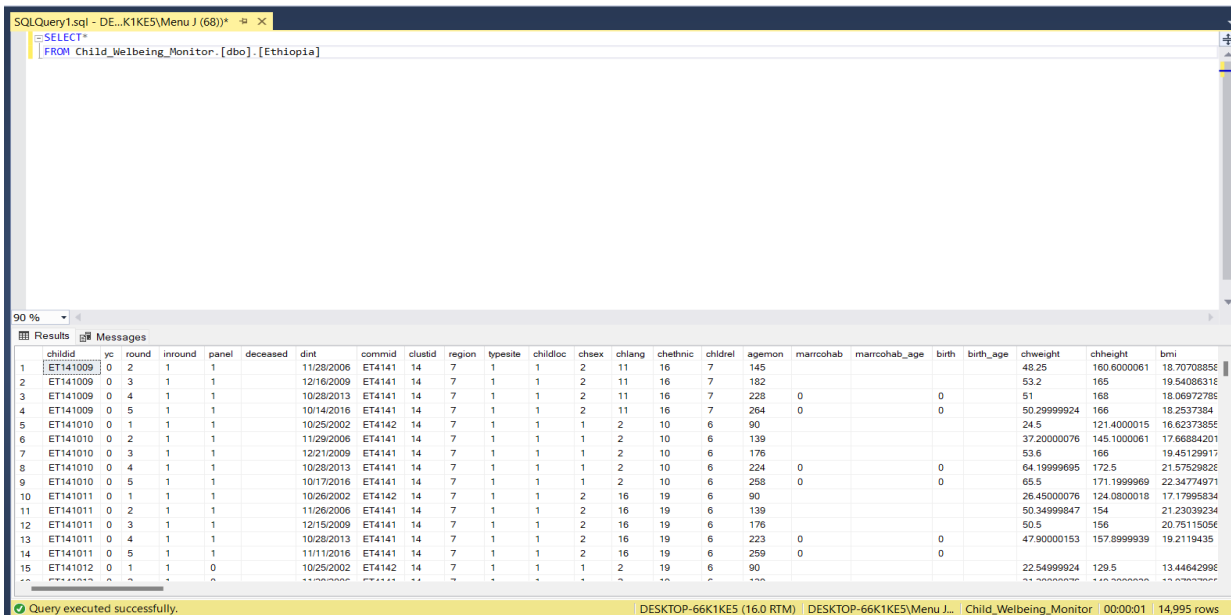


Figure 09

## Create View

- Before creating the view, the dataset should be understood and explored.

```
--Understanding and Exploring the datasets
SELECT TOP 100*
FROM Child_Welbeing_Monitor.dbo.Ethiopia;

SELECT TOP 100*
FROM Child_Welbeing_Monitor.dbo.Vietnam;
```

Figure 10

- The columns of the dataset and the data types should be checked so that the necessary variable columns can be selected for the further analysis of the dataset.

-- Get the column names and data types of the table

```
SELECT COLUMN_NAME, DATA_TYPE
FROM INFORMATION_SCHEMA.COLUMNS
WHERE TABLE_NAME = 'Ethiopia';

SELECT COLUMN_NAME, DATA_TYPE
FROM INFORMATION_SCHEMA.COLUMNS
WHERE TABLE_NAME = 'Vietnam';
```

	COLUMN_NAME	DATA_TYPE
1	childid	varchar
2	yc	varchar
3	round	varchar
4	inround	varchar
5	panel	varchar
6	deceased	varchar
7	dint	varchar
8	commid	varchar
9	clustid	varchar
10	region	varchar
11	typesite	varchar
12	childloc	varchar
13	chsex	varchar
14	chlang	varchar

	COLUMN_NAME	DATA_TYPE
1	childid	varchar
2	yc	varchar
3	round	varchar
4	inround	varchar
5	panel12345	varchar
6	deceased	varchar
7	dint	varchar
8	commid	varchar
9	clustid	varchar
10	typesite	varchar
11	region	varchar
12	childloc	varchar

Q | DESKTOP-66K1KE5 (16.0 RTM) | DESKTOP-66K1KE5\Menu J... | Child\_Welbeing\_Monitor | 00:00:00 | 427 rows

Figure 11

- After referring to the columns of the dataset a new column is added to both datasets called **Country\_Name** which will show the name of the country.

```
--Adding Country Column
--Ethiopia
ALTER TABLE [dbo].[Ethiopia]
ADD Country_Name Varchar(50);

--Vietnam
ALTER TABLE [dbo].[Vietnam]
ADD Country_Name Varchar(50);
```

*Figure 12*

- Create two table views for each country – Vietnam and Ethiopia, by extracting only the necessary columns for the analysis of the poverty criteria indices.
- The specific column variables which are used for the creation of the two table views are provided earlier. A total of **24 columns** were selected where the column variable **Country\_Name** was newly added to the datasets.
- During the creation of the views, for the data cleaning process the missing data points were replaced with NA.
- **Note that the data cleaning and data transformation processes were carried out in the SSMS SQL Server since there was an inability to acquire the Power BI Premium Package for the data transformation and data cleaning since the trial phase of the package was already utilized during the Workshop series.**

```
--Creating Table View
--Ethiopia
CREATE VIEW Country_Ethiopia AS
SELECT
    COALESCE(NULLIF([Country_Name], ''), 'Ethiopia') AS [Country],
    COALESCE(NULLIF([childid], ''), 'NA') AS [Child_ID],
    COALESCE(NULLIF([yc], ''), 'NA') AS [Cohort],
    COALESCE(NULLIF([agemon], ''), 'NA') AS [Child_Age],
    COALESCE(NULLIF([wi_new], ''), 'NA') AS [Wealth_Index],
    COALESCE(NULLIF([hq_new], ''), 'NA') AS [Housing_Quality_Index],
    COALESCE(NULLIF([sv_new], ''), 'NA') AS [Service_Index_Access],
    COALESCE(NULLIF([ownlandhse], ''), 'NA') AS [Own_Household_Land],
    COALESCE(NULLIF([ownhouse], ''), 'NA') AS [Own_House],
    COALESCE(NULLIF([hschool], ''), 'NA') AS [School_Hours],
    COALESCE(NULLIF([caredu], ''), 'NA') AS [Education_Level_Child],
    COALESCE(NULLIF([dadedu], ''), 'NA') AS [Education_Level_Dad],
    COALESCE(NULLIF([momedu], ''), 'NA') AS [Education_Level_Mom],
    COALESCE(NULLIF([headedu], ''), 'NA') AS [Education_Level_Head],
    COALESCE(NULLIF([chrephealth1], ''), 'NA') AS [Child_Health_1],
    COALESCE(NULLIF([chrephealth2], ''), 'NA') AS [Child_Health_2],
    COALESCE(NULLIF([chrephealth3], ''), 'NA') AS [Child_Health_3],
    COALESCE(NULLIF([chrephealth4], ''), 'NA') AS [Child_Health_4],
    COALESCE(NULLIF([chhealth], ''), 'NA') AS [Child_Health_General],
    COALESCE(NULLIF([drwaterq_new], ''), 'NA') AS [Water_Quality],
    COALESCE(NULLIF([chweight], ''), 'NA') AS [Child_Weight],
    COALESCE(NULLIF([chheight], ''), 'NA') AS [Child_Height],
    COALESCE(NULLIF([bmi], ''), 'NA') AS [Child_BMI],
    COALESCE(NULLIF([underweight], ''), 'NA') AS [Child_Underweight]
FROM Child_Welbeing_Monitor.dbo.Ethiopia;
```

*Figure 13 a*

```
--Vietnam
CREATE VIEW Country_Vietnam AS
SELECT
    COALESCE(NULLIF([Country_Name], ''), 'Vietnam') AS [Country],
    COALESCE(NULLIF([childid], ''), 'NA') AS [Child_ID],
    COALESCE(NULLIF([yc], ''), 'NA') AS [Cohort],
    COALESCE(NULLIF([agemon], ''), 'NA') AS [Child_Age],
    COALESCE(NULLIF([wi_new], ''), 'NA') AS [Wealth_Index],
    COALESCE(NULLIF([hq_new], ''), 'NA') AS [Housing_Quality_Index],
    COALESCE(NULLIF([sv_new], ''), 'NA') AS [Service_Index_Access],
    COALESCE(NULLIF([ownlandhse], ''), 'NA') AS [Own_Household_Land],
    COALESCE(NULLIF([ownhouse], ''), 'NA') AS [Own_House],
    COALESCE(NULLIF([hschool], ''), 'NA') AS [School_Hours],
    COALESCE(NULLIF([caredu], ''), 'NA') AS [Education_Level_Child],
    COALESCE(NULLIF([dadedu], ''), 'NA') AS [Education_Level_Dad],
    COALESCE(NULLIF([momedu], ''), 'NA') AS [Education_Level_Mom],
    COALESCE(NULLIF([headedu], ''), 'NA') AS [Education_Level_Head],
    COALESCE(NULLIF([chrephealth1], ''), 'NA') AS [Child_Health_1],
    COALESCE(NULLIF([chrephealth2], ''), 'NA') AS [Child_Health_2],
    COALESCE(NULLIF([chrephealth3], ''), 'NA') AS [Child_Health_3],
    COALESCE(NULLIF([chrephealth4], ''), 'NA') AS [Child_Health_4],
    COALESCE(NULLIF([chhealth], ''), 'NA') AS [Child_Health_General],
    COALESCE(NULLIF([drwaterq_new], ''), 'NA') AS [Water_Quality],
    COALESCE(NULLIF([chweight], ''), 'NA') AS [Child_Weight],
    COALESCE(NULLIF([chheight], ''), 'NA') AS [Child_Height],
    COALESCE(NULLIF([bmi], ''), 'NA') AS [Child_BMI],
    COALESCE(NULLIF([underweight], ''), 'NA') AS [Child_Underweight]
FROM Child_Welbeing_Monitor.dbo.Vietnam;
```

Figure 13 b

- Viewing the table views before creating separate views for each of the poverty index metrics.

```
--Viewing the Views created
--Ethiopia
SELECT*
FROM [dbo].[Country_Ethiopia];

--Vietnam
SELECT*
FROM [dbo].[Country_Vietnam];
```

Figure 14

- Create separate views for the five poverty measurement indices;
  - Income per capita.**
  - Access to proper education.**
  - Health expenditure per child.**
  - Living standards: access to pure clean water.**
  - Malnutrition rate.**

i. Measurement Index 1: Income per capita

```
--Creating separate views for each poverty index criteria
--Criteria 1: Income per capita
CREATE VIEW Income_Per_Capita AS
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Wealth_Index],
    [Housing_Quality_Index],
    [Service_Index_Access],
    [Own_Household_Land],
    [Own_House]
FROM [dbo].[Country_Ethiopia]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Wealth_Index] != 'NA' AND
    [Housing_Quality_Index] != 'NA' AND
    [Service_Index_Access] != 'NA' AND
    [Own_Household_Land] != 'NA' AND
    [Own_House] != 'NA'
UNION ALL
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Wealth_Index],
    [Housing_Quality_Index],
    [Service_Index_Access],
    [Own_Household_Land],
    [Own_House]
FROM [dbo].[Country_Vietnam]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Wealth_Index] != 'NA' AND
    [Housing_Quality_Index] != 'NA' AND
    [Service_Index_Access] != 'NA' AND
    [Own_Household_Land] != 'NA' AND
    [Own_House] != 'NA';
```

*Figure 15 a*

ii. Measurement Index 2: Access to proper education.

```
--Criteria 2: Education access
CREATE VIEW Education_Access_AS
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [School_Hours],
    [Eductaion_Level_Child],
    [Education_Level_Dad],
    [Education_Level_Mom],
    [Education_Level_Head]
FROM [dbo].[Country_Ethiopia]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [School_Hours] != 'NA' AND
    [Eductaion_Level_Child] != 'NA' AND
    [Education_Level_Dad] != 'NA' AND
    [Education_Level_Mom] != 'NA' AND
    [Education_Level_Head] != 'NA'
UNION ALL
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Wealth_Index],
    [Housing_Quality_Index],
    [Service_Index_Access],
    [Own_Household_Land],
    [Own_House]
FROM [dbo].[Country_Vietnam]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [School_Hours] != 'NA' AND
    [Eductaion_Level_Child] != 'NA' AND
    [Education_Level_Dad] != 'NA' AND
    [Education_Level_Mom] != 'NA' AND
    [Education_Level_Head] != 'NA';
```

*Figure 15 b*

iii. Measurement Index 3: Health expenditure per child.

```
--Criteria 3: Health Expenditure per child
CREATE VIEW Child_Health_Expenditure_AS
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Child_Health_1],
    [Child_Health_2],
    [Child_Health_3],
    [Child_Health_4],
    [Child_Health_General]
FROM [dbo].[Country_Ethiopia]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Child_Health_1] != 'NA' AND
    [Child_Health_2] != 'NA' AND
    [Child_Health_3] != 'NA' AND
    [Child_Health_4] != 'NA' AND
    [Child_Health_General] != 'NA'
UNION ALL
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Child_Health_1],
    [Child_Health_2],
    [Child_Health_3],
    [Child_Health_4],
    [Child_Health_General]
FROM [dbo].[Country_Vietnam]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Child_Health_1] != 'NA' AND
    [Child_Health_2] != 'NA' AND
    [Child_Health_3] != 'NA' AND
    [Child_Health_4] != 'NA' AND
    [Child_Health_General] != 'NA';
```

*Figure 15 c*



iv. Measurement Index 4: Living Standard (Access to pure clean water).

```
--Criteria 4: Living Standards: access to clean water
CREATE VIEW Living_Standard AS
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Water_Quality]
FROM [dbo].[Country_Ethiopia]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Water_Quality] != 'NA'
UNION ALL
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Water_Quality]
FROM [dbo].[Country_Vietnam]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Water_Quality] != 'NA';
```

*Figure 15 d*

v. Measurement Index 5: Malnutrition rate.

```
--Criteria 5: Malnutrition Rate
CREATE VIEW Malnutrition_Rate AS
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Child_Weight],
    [Child_Height],
    [Child_BMI],
    [Child_Underweight]
FROM [dbo].[Country_Ethiopia]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Child_Weight] != 'NA' AND
    [Child_Height] != 'NA' AND
    [Child_BMI] != 'NA' AND
    [Child_Underweight] != 'NA'
UNION ALL
SELECT
    [Country],
    [Child_ID],
    [Cohort],
    [Child_Age],
    [Child_Weight],
    [Child_Height],
    [Child_BMI],
    [Child_Underweight]
FROM [dbo].[Country_Vietnam]
WHERE
    [Country] != 'NA' AND
    [Child_ID] != 'NA' AND
    [Cohort] != 'NA' AND
    [Child_Age] != 'NA' AND
    [Child_Weight] != 'NA' AND
    [Child_Height] != 'NA' AND
    [Child_BMI] != 'NA' AND
    [Child_Underweight] != 'NA';
```

*Figure 15 e*

- After the creation of the relationships between the two countries – Vietnam and Ethiopia, five separate reports will design for the five child poverty measurement indices which will show visual comparisons between the two countries using **Power BI Report Builder**.

## 1.4. Key Patterns and Insights

The key patterns and insights of the five child poverty measurement indices which are used to compare the poverty levels of the two countries Vietnam and Ethiopia are as follows;

### i. Income per capita.

- The per capita income of the households of Vietnam and Ethiopia are measured based on the variables; **Wealth\_Index**, **Housing\_Quality\_Index**, **Service\_Index**, **Own\_Household\_Land** and **Own\_House**.

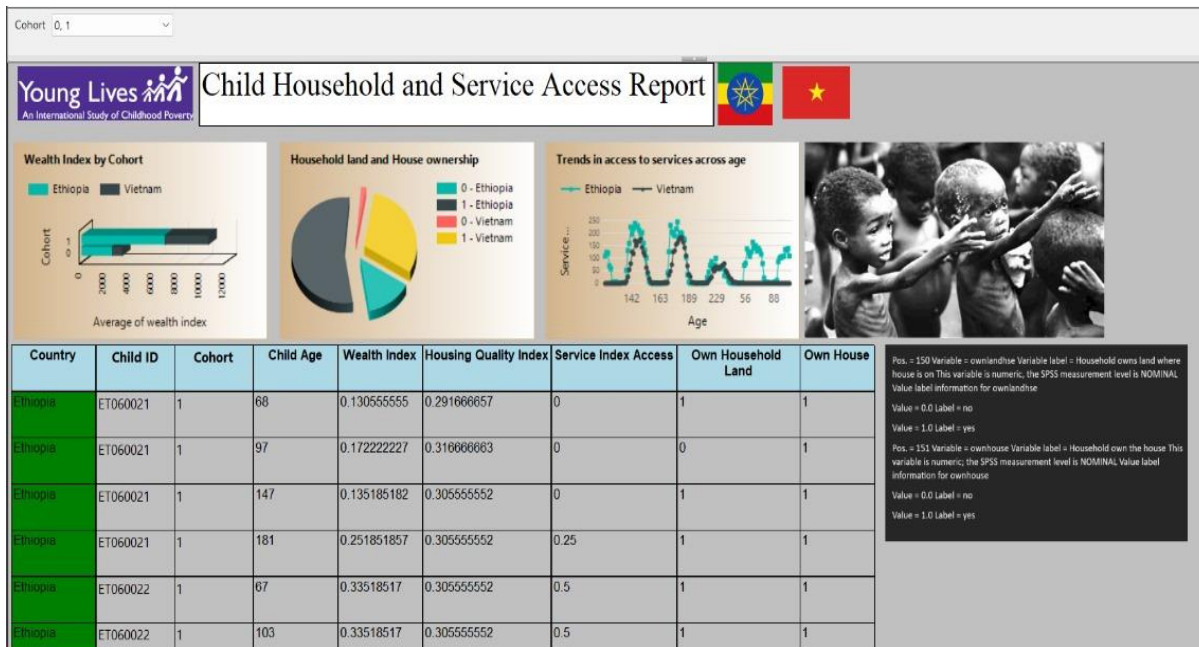


Figure 16

- The Child Household and Service Access report provides insights and key patterns towards the Income per capita between Vietnam and Ethiopia.
- Wealth Index by Cohort:** According to the bar chart, Vietnam shows a generally higher wealth index in comparison to Ethiopia. Generally, the younger cohort which is represented as Cohort 1 shows a greater wealth index in comparison to Cohort 0 which represents the older cohort in both countries. It shows that the children of Ethiopia cohorts are mostly clustered at the lower wealth levels while the children of Vietnam have spread more across the higher wealth brackets.
- Household Land and House Ownership:** According to the pie chart, Ethiopia has more households that has house and land ownerships in comparison to Vietnam.
- Trends in Access to Services Across Age:** The line chart shows that the service access of both countries has fluctuations with the age. Ethiopia has more inconsistent fluctuations in comparison to Vietnam which suggests that Ethiopia faces more challenges in regards to service accessibility.

## ii. Access to Education.

- The access to proper education of the members in the household of Vietnam and Ethiopia are measured based on the variables; **School\_Hours**, **Education\_Level\_Child**, **Education\_Level\_Dad**, **Education\_Level\_Mom** and **Education\_Level\_Head**.

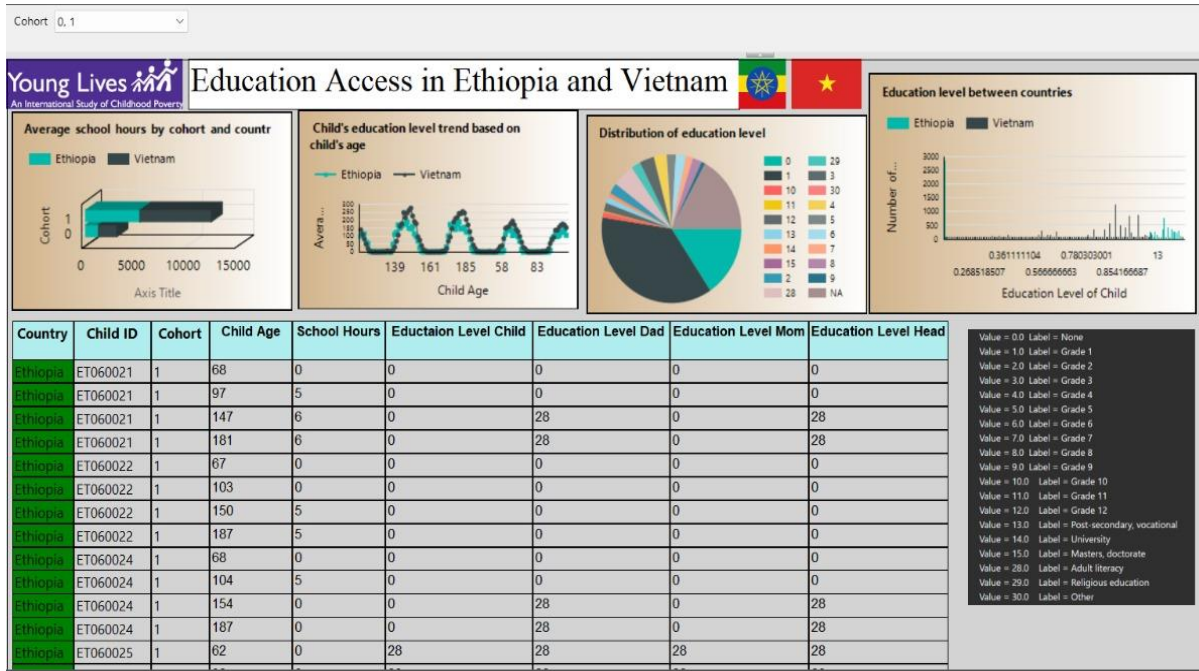


Figure 17

- The Education Access in Ethiopia and Vietnam report provides key patterns and insights towards the access of proper education by the residents and households.
- Average School Hours by Cohort and Country:** According to the bar chart, Vietnam has higher average school hours in comparison among the two cohorts in comparison to Ethiopia. Generally, Cohort 1 which represents the younger cohort displays a higher average school hours spent when compared to the Cohort 0 which represents the older cohort. This suggests that Vietnam has better access to access to school resources and educational systems in comparison to Ethiopia.
- Child's Education Level Trend Based on Child's Age:** The line shows the educational level trends across the age in both countries. Accordingly, Vietnam shows peaks at the early ages indicating that the children of Vietnam attain higher education levels earlier in comparison to the children of Ethiopia. This displays the progression of the children of Ethiopia attain higher education levels at a much slower rate in comparison to Vietnam.
- Distribution of Education Levels:** The pie chart displays the distribution of the education levels of both countries. There is a large segment of levels of education are concentrated at level 1 which is related to the Grade 1 level. This indicates that a significant portion of the population have received primary education while the least level distribution goes to Grade 9 which suggests that children might have dropped out at grade 9 or there might have had limited education resources.

- **Education Level Between Countries:** This shows the levels of education achieved by the children in relation to their countries. Accordingly, more children in Vietnam complete higher school levels which are closer to Grade 12 and University. This indicates that more children retain in the education system towards higher education. While, the children in Ethiopia are highly concentrated towards the lower education levels (from Grade 1 to Grade 8) with a low number of children are concentrated towards the advancement in the higher education levels signifying the difficulties and challenges the children in Ethiopia face towards achieving higher education.

### iii. Health Expenditure per Child.

- The health expenditure incurred per child in Vietnam and Ethiopia are measured based on the variables; **Child\_Health\_1, Child\_Health\_2, Child\_Health\_3, Child\_Health\_4** and **Child\_Health\_General**.

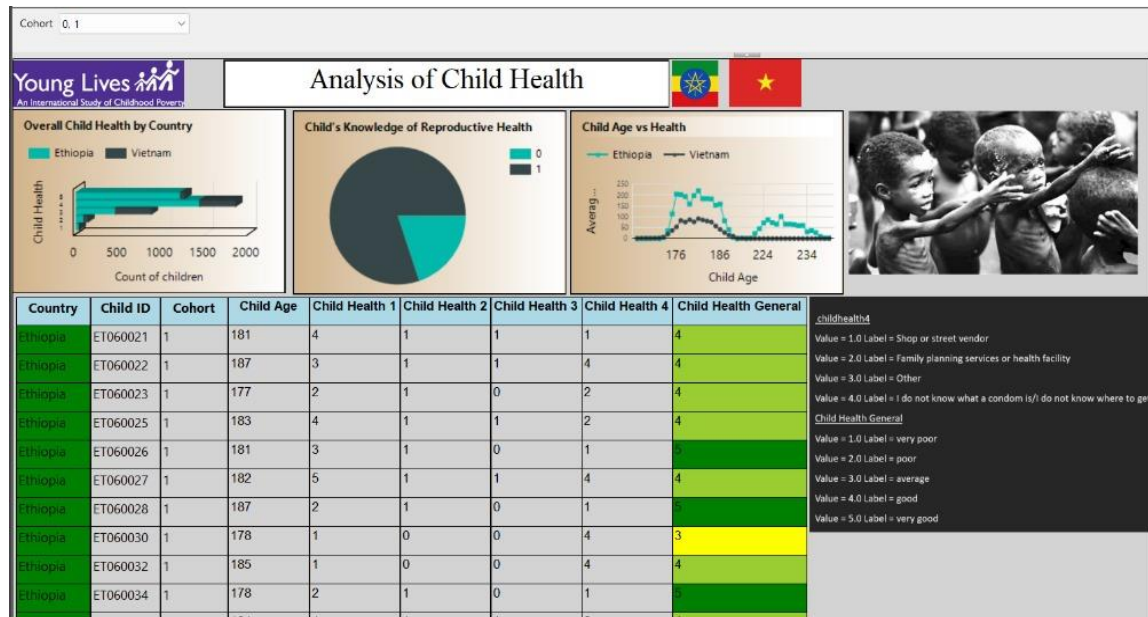


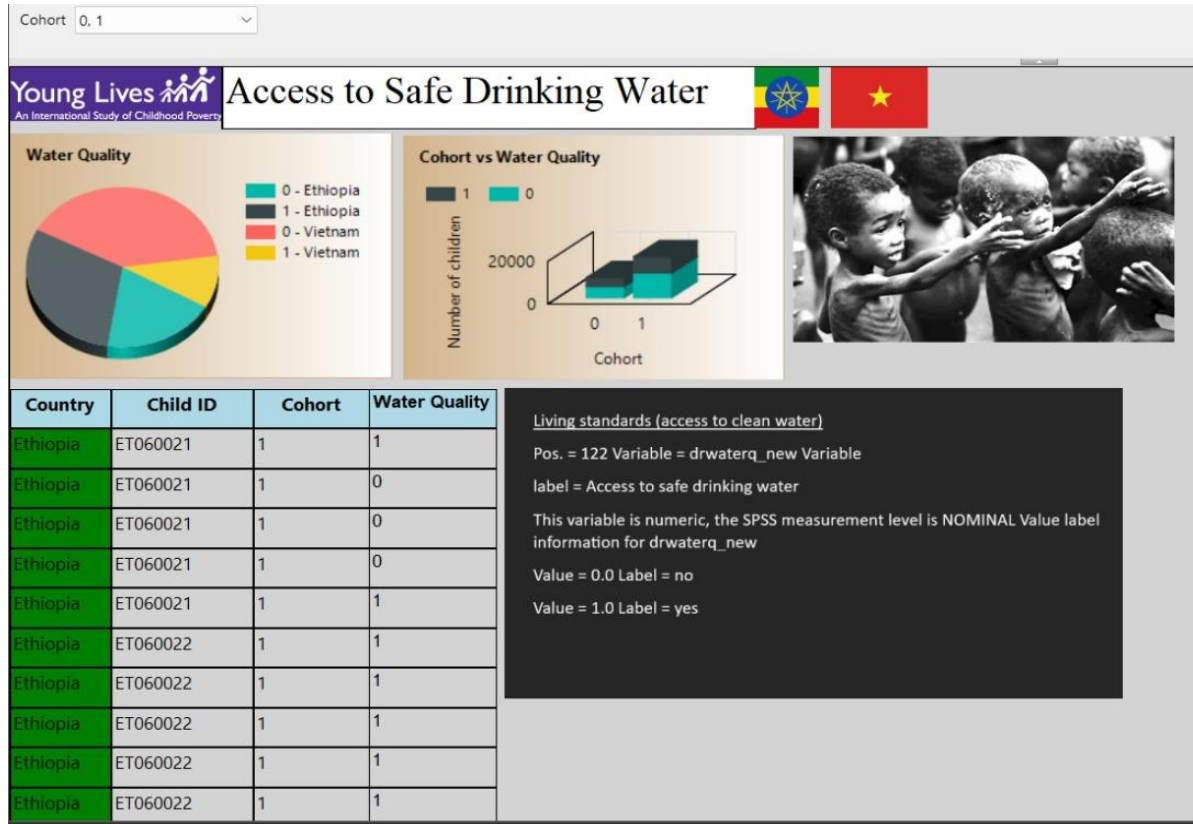
Figure 18

- The analysis of child health report provides key patterns and insights towards the health expenditure spend per child by the countries – Vietnam and Ethiopia.
- **Overall Child Health by Country:** According to the bar chart, Ethiopia shows a higher number of children reports in comparison to Vietnam where, the highest number of children in Ethiopia have reported that their health is good in general while the most of the children in Vietnam have reported that the children either have average health or are good health conditions.
- **Child's Knowledge of Reproductive Health:** According to the pie chart, it shows a significant number of children are aware of the knowledge in regards to the Reproductive system while a least number of children has a lack of knowledge on regards to the Reproductive system.

- **Child Age vs Health Trends:** According to the line chart, both countries show consistent fluctuations in the average health score across the child's age in both countries. Notably, while Ethiopia shows a more gradual distribution in the average health trendline, Vietnam shows spikes around the child age 178 months to 186 months.

#### iv. Living Standards: Access to Clean Water.

- The ability of the children to access clean and pure water which shows its impact on the living standard of the two countries are measured based on the variable **Water\_Quality**.



*Figure 19*

- The analysis of the safe drinking water report provides the key patterns and insights towards the level of living standards based on the availability of clean safe drinking water.
- **Water Quality:** The pie chart shows the distribution of the water quality among the two countries. Accordingly, a large number of the children of Ethiopia have access to pure and clean drinking water while a large number of the children of Vietnam do not have access to pure and clean drinking water. This shows that the children of Ethiopia have better living standards in regards to the availability of pure quality drinking water in comparison to Vietnam.
- **Cohort vs Water Quality:** According to the bar chart the younger cohort which is represented by Cohort 1 has higher access to poor quality drinking water in comparison to the older cohort which is represented by Cohort 0.



v. **Malnutrition Rate.**

- The malnutrition rate of the children living in Vietnam and Ethiopia are measured based on the variable; **Child\_Weight**, **Child\_Height**, **Child\_BMI** and **Child\_Underweight**.

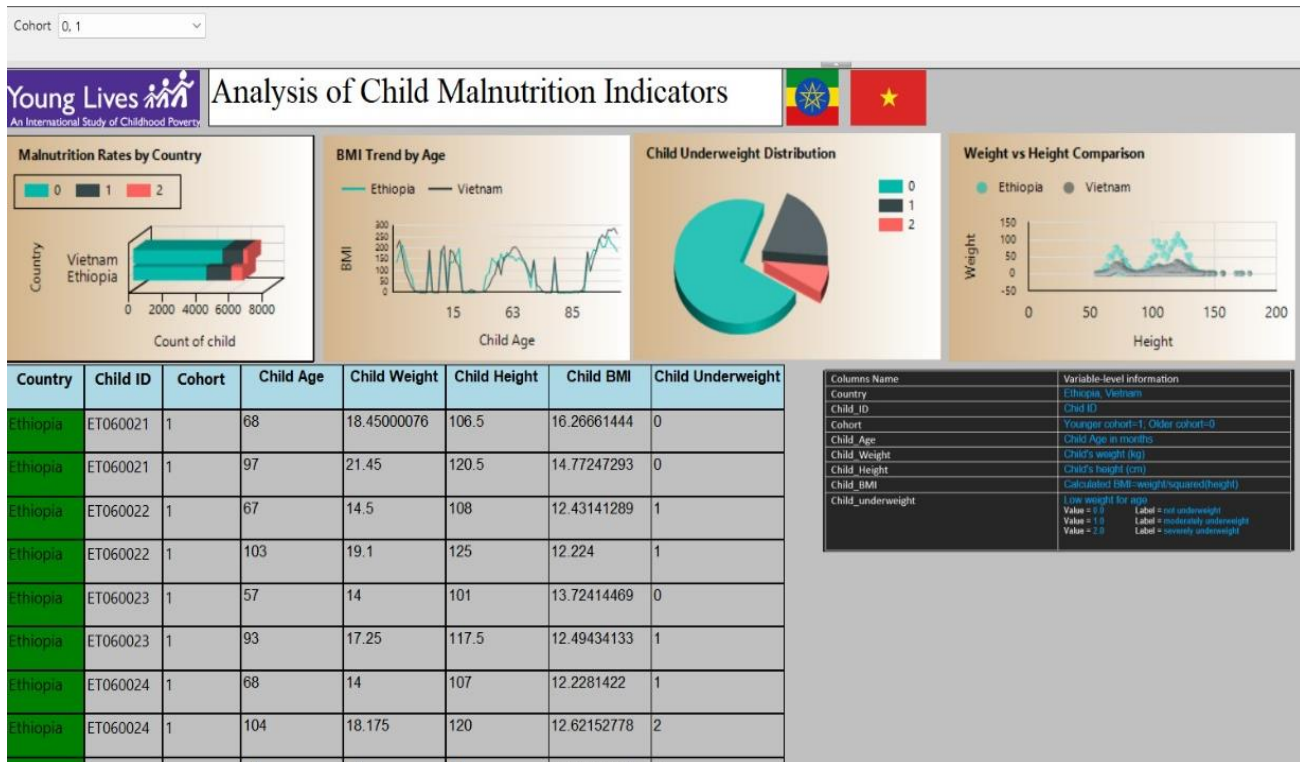


Figure 20

- The analysis of child malnutrition indicators provides key patterns and insights towards the malnutrition rate of the children living in Vietnam and Ethiopia.
- Malnutrition Rates by Country:** According to the bar chart, Ethiopia has a high number of malnutrition children in comparison to the children of Vietnam. Based on the rates, a large number of children are under the category of not underweight which is represented by 0 while the lowest rate goes to severely underweight category which is 2.
- BMI Trend by Age:** According to the line chart, Vietnam shows a smooth increase in the BMI values which indicates a relatively stable growth and nutrition availability to the children. Whereas, Ethiopia displays more fluctuations and lower BMIs along the trend which indicates inconsistencies in the food securities and child health during their growth.
- Child Underweight Distribution:** According to the pie chart, a large distribution rate falls under not underweight while a small distribution rate falls under severely underweight category.
- Weight vs Height Comparison:** The scatter plot shows that the children in Vietnam shows more consistency in their growth patterns while the children in Ethiopia displays more variance indicating poor growth metrics and undernourishment rates.

### **1.5.Recommendations**

The recommendations to reduce the poverty levels of low-level income earning countries such as Vietnam and Ethiopia (in this scenario) are as follows.

- **Improving the access to education:** According to the report, Vietnam demonstrates better access to educational resources compared to Ethiopia. Improving programs that are aimed at expanding the school's infrastructure, providing educational materials and addressing issues such as the long distances that the students have to children especially to children who live in rural areas. Encouraging online learning facilities for educational purposes is another way to improve the access to education.
- **Improving health outcomes:** Health expenditure spent on a child is a critical concern in terms of investment and raises the child's health concerns. Investing towards healthcare facilities, increasing awareness towards health education, ensuring a constant provision of medical supplies and ensuring a well-maintained nutritional diet are some ways that can improve the health outcomes reducing the concerns on regards to the health.
- **Securing safe drinking water facilities:** Accordingly, Ethiopia shows a better access to pure and clean drinking water in comparison to Vietnam. Investing towards water purification infrastructure facilities, promoting awareness towards policies that contribute towards sustainable water resource management, promoting awareness on regards to water sanitation methods and practices are some practices that can be carried out to secure safe drinking water facilities.
- **Addressing malnutrition:** Activities such as increasing awareness towards malnutrition and it's causes, introducing healthy nutritional meal programs from schools, improving awareness about having a nutritional diet, promoting and investing on sustainable agricultural activities – can ensure the reducing of the malnutrition rates.
- **Targeted poverty reduction programs:** Implementing programs such as building sustainable agricultural programs, improving employment opportunities, improving the access to education, raising awareness on regards to health and safety, increasing the support which allows women to work in the workforce are some activities that can be carried out in order to improve the income gap and improving the overall economic growth.



## **1.6. Conclusion**

The comparative analysis between the low-income earning countries – Vietnam and Ethiopia using the child poverty indices reveals that both nations face unique challenges in maintaining the wellbeing of the children. While Vietnam excels in having a better access to education and having a better growth in the per capita income, it lacks in providing a better access to pure clean drinking water which indicates their need to improve the water infrastructure facilities. On the other hand, Ethiopia faces challenges with having inconsistent health outcomes while having high malnutrition rates which targets the food security.

Overall, improving the child well-being in these countries is extremely essential and it takes multifaceted approaches to improve the child well-being. These approaches include the involvement of the local governments, international organizations and community stakeholders. By addressing the issues in relation to these gaps such as health, education and the gaps in the living standards and promoting awareness towards these issues and implementing action and policies in order to reduce these gaps, countries such as Vietnam and Ethiopia will be able to make significant changes towards alleviating child poverty further improving the future of the younger generations.

# **Task 02**

## **Greater Manchester Domestic Energy Performance Analyzer**



## **2.1. Introduction**

This report presents a comprehensive analysis of the energy performance of domestic buildings across Greater Manchester's local authorities between 2013 and 2023. Leveraging publicly available Energy Performance Certificate (EPC) data, we aim to uncover key trends, identify influential factors affecting energy consumption, and provide actionable insights for stakeholders involved in energy efficiency initiatives.

By examining the distribution of energy efficiency ratings across various property types and localities, we seek to understand the current state of energy performance in Greater Manchester. Additionally, we will investigate the relationship between property size, age, and energy consumption to identify potential correlations and inform targeted interventions.

Furthermore, the report will explore the impact of different heating systems and insulation types on energy efficiency. This analysis will help to identify areas where improvements in heating and insulation can significantly reduce energy consumption and emissions.

Finally, we will identify common areas where properties can benefit from upgrades to enhance their energy performance. These recommendations will provide valuable guidance for property owners and policymakers alike, enabling them to make informed decisions about investments in energy efficiency measures.

In conclusion, this report aims to contribute to a more sustainable and energy-efficient Greater Manchester by providing evidence-based insights into the region's energy performance. The findings presented in this report will inform policy development, support targeted interventions, and encourage property owners to adopt energy-efficient practices.

## 2.2. About the Dataset.

AutoSave

certificates.csv • Saved to this PC

Search

FileHomeInsertPage LayoutFormulasDataReviewViewAutomateDeveloperHelp

Clipboard

Font

Alignment

Number

Styles

Cells

Editing

Add-ins

AutoSum

Sort & Filter

Fill

Clear

Comments

Share

Post

Clipboard

Font

Font Face: Aptos Narrow, Size: 11, Bold, Italic, Underline, Paragraph: Bullets, Numbering, Indentation, Styles: Merge & Center

Alignment

Number

Styles

Cells

Editing

Add-ins

Wrap Text

General

Conditional Formatting

Format as Table

Cell Styles

Insert

Delete

Format

AutoSum

Sort & Filter

Fill

Clear

Comments

Share

A1

LMK KEY

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W						
LMK KEY	ADDRESS1	ADDRESS2	ADDRESS3	POSTCODE	BUILDING	CURRENT	POTENTIAL	CURRENT	POTENTIAL	PROPERTY	BUILT	FOF	INSPECT	LOCAL	AU	CONSTITU	COUNTY	LODGEME	TRANSACT	ENVIRONN	ENVIRONN	ENERGY	C	ENERGY	C	CO2	EMIS	CC
20c4b426	428 St. Marys Road			M40 0DE	1E+10 C	B	70	85	House	Mid-Terrac	#####	E08000000	E14000807	#####	marketed :	67	83	209	103	3								
5.45E+32	20 Parkleigh Drive			M40 3RY	8.95E+08 D	D	64	68	House	Semi-Deta	#####	E08000000	E14000807	#####	marketed :	58	63	279	250	4.1								
2.27E+32	22 Silverwell Street			M40 1PA	5.99E+09 D	B	55	84	House	Mid-Terrac	#####	E08000000	E14000807	#####	none of the	52	84	322	104	3.6								
1.41E+33	Apartment 1 Pennington Street			M12 4QJ	3.75E+09 G	D	11	56	Flat	Enclosed E	#####	E08000000	E14000808	#####	none of the	24	33	505	409	13								
2.88E+32	10 Morna Walk			M12 6WP	7.87E+09 C	C	74	75	House	End-Terrac	#####	E08000000	E14000807	#####	rental (soc	70	71	182	179	3.7								
25c0064a	42 GRINDLEY AVENUE			M21 7NF	1E+10 C	B	72	86	House	Mid-Terrac	#####	E08000000	E14000809	#####	rental	70	84	189	93	2.4								
25d0064a	APARTME LOCKES YF 4 GREAT M			M1 5AL	1E+10 D	C	66	72	Flat	Mid-Terrac	#####	E08000000	E14000807	#####	rental	62	61	281	286	2.7								
1.07E+33	14 Montpelior Road			M22 0DZ	6.33E+09 D	C	63	80	House	Semi-Deta	#####	E08000000	E14001059	#####	marketed :	64	81	217	110	3.3								
1.64E+33	1678 Fog Lane			M20 6FJ	9.87E+09 C	C	69	72	Flat	Mid-Terrac	#####	E08000000	E14000809	#####	rental (priv	71	75	273	236	1.6								
1.26E+33	24 Carrgreen Close			M19 1LT	5.18E+08 D	B	56	86	House	Semi-Deta	#####	E08000000	E14000809	#####	none of the	50	85	363	97	3.7								
1.84E+32	14 Chervil Close			M14 7DP	6.94E+09 C	B	78	81	Flat	Enclosed E	#####	E08000000	E14000808	#####	marketed :	72	74	243	229	2.2								
1.22E+33	27 Riverstone Drive			M23 9QW	7.25E+09 D	C	67	75	Flat	Semi-Deta	#####	E08000000	E14001059	#####	none of the	48	60	425	321	3.6								
1.48E+33	Flat 88 1 Cambridge Street			M1 5GB	2.55E+09 B	B	83	83	Flat	Detached	#####	E08000000	E14000807	#####	new dwelli	85	85	120	120	1.1								
1.47E+32	Flat A/1 Etr Barlow Moor Road			M20 2QQ	6.45E+08 D	B	65	81	Flat	Semi-Deta	#####	E08000000	E14000809	#####	marketed :	63	79	533	288	1.8								
2.20E+32	Apartment 108 High Street			M4 1HT	4.95E+09 D	C	64	80	Flat	Enclosed h	#####	E08000000	E14000807	#####	rental (priv	68	68	216	218	2.6								
2.29E+32	80 The Boulevard			M20 2EU	3.89E+09 B	B	82	86	Flat	Semi-Deta	#####	E08000000	E14000809	#####	rental (priv	81	81	188	191	1.3								
5.10E+32	81 Ravenscar Crescent			M22 0HP	1.98E+09 F	D	37	65	House	Semi-Deta	#####	E08000000	E14001059	#####	marketed :	32	59	533	281	7.5								
5.41E+32	17 Goodworth Walk			M40 5NW	2.76E+09 D	C	62	72	House	End-Terrac	#####	E08000000	E14000807	#####	rental (soc	55	68	303	219	4.4								
27c46237f	4 Bates Str Longsight			M13 0WL	1E+10 C	B	71	86	House	End-Terrac	#####	E08000000	E14000808	#####	rental	69	84	191	92	2.8								
259c3810f	FLAT 159 70 OLD MILL STREET			M4 6LY	1E+10 B	B	83	83	Flat	Detached	#####	E08000000	E14000807	#####	new dwelli	85	85	122	122	1								
4.21E+32	52 Painswick Road			M22 1GG	9.52E+09 C	C	73	73	Flat	Detached	#####	E08000000	E14001059	#####	rental (soc	74	74	182	182	2.1								
2.75E+32	Apartment 5 Ludgate Hill			M4 4TG	9.29E+09 C	B	77	85	Flat	Mid-Terrac	#####	E08000000	E14000807	#####	rental (priv	79	79	222	217	1.3								
1.16E+33	32 Chapel Road			M22 4JW	5.28E+09 E	C	45	78	House	Semi-Deta	#####	E08000000	E14001059	#####	none of the	40	75	338	131	7.2								
1.70E+32	65 Rockhampton Street			M18 8UP	40282568 E	B	42	86	House	Mid-Terrac	#####	E08000000	E14000808	#####	marketed :	36	85	518	102	5.1								

certificates

Accessibility: Unavailable

Figure 21

The above dataset used in this analysis is the publicly available in Energy Performance Certificates (EPC) dataset for properties in England and Wales ([Energy Performance of Buildings Data England and Wales \(opendatacommunities.org\)](https://www.opendatacommunities.org/energy-performance-of-buildings)).

The dataset contains **300888 rows** and total of **92 columns**. An EPC is required when properties are constructed, sold, or rented, and provides detailed information about the energy use and efficiency of a building, including:

- **Energy Efficiency Rating:** Ranges from A (most efficient) to G (least efficient).
- **Property Information:** Type, size, and age.
- **Heating, Insulation, and Emissions:** Data on heating systems, insulation type, CO2 emissions, and potential areas for improvement.

This dataset includes fields such as property type, energy cost, potential improvements, and the energy ratings for various local authorities in Greater Manchester. For the creation of the database the reports from the year 2013 to 2023 are considered.



## 2.3. Database Utilization

The importing and the process taken for the creation of the database were implemented on a similar basis as the Database Utilization of section 1.1 of Task 1. However, for the creation of the database which is created in the SQL Server, the server connection is build up using the Microsoft Azure Portal.

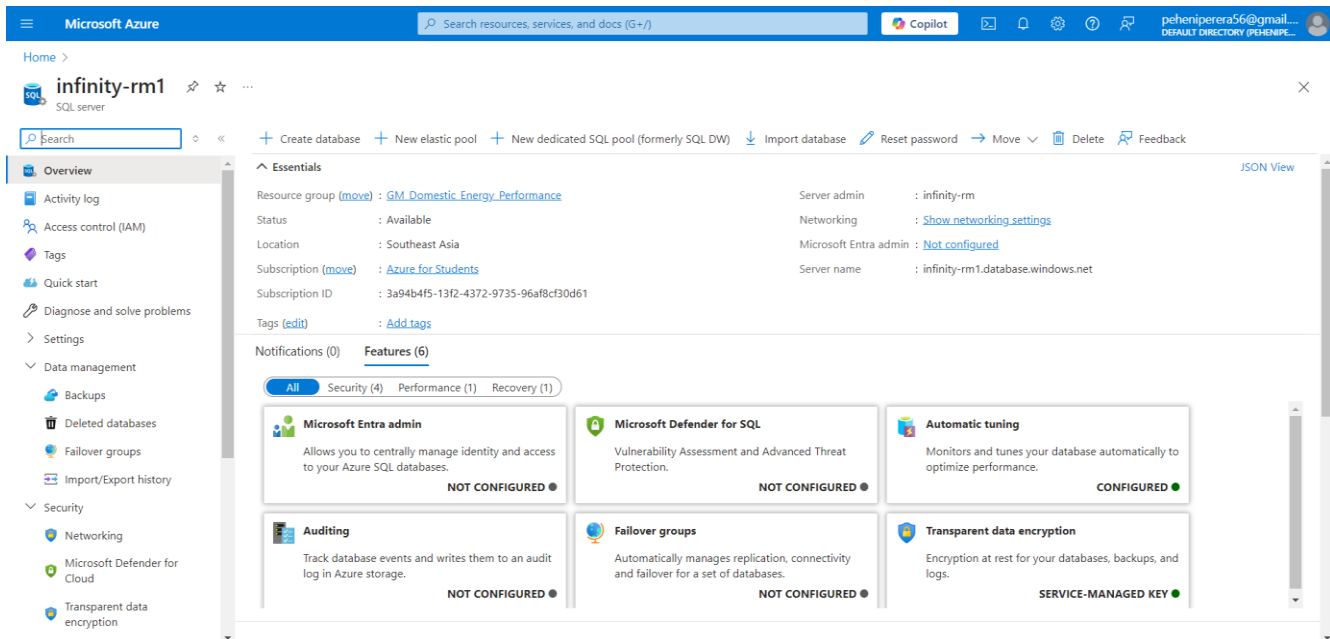


Figure 22

### Steps to import and setting up the database.

- After downloading the compressed zip file which consists of the metadata and the datasets, extract and preview the datasets and metadata separately to understand them.
- Choose the necessary dataset (Certificate dataset) that will be considered for the analysis.
- Connecting to the SQL Server using azure login credentials to the authentication of the server.

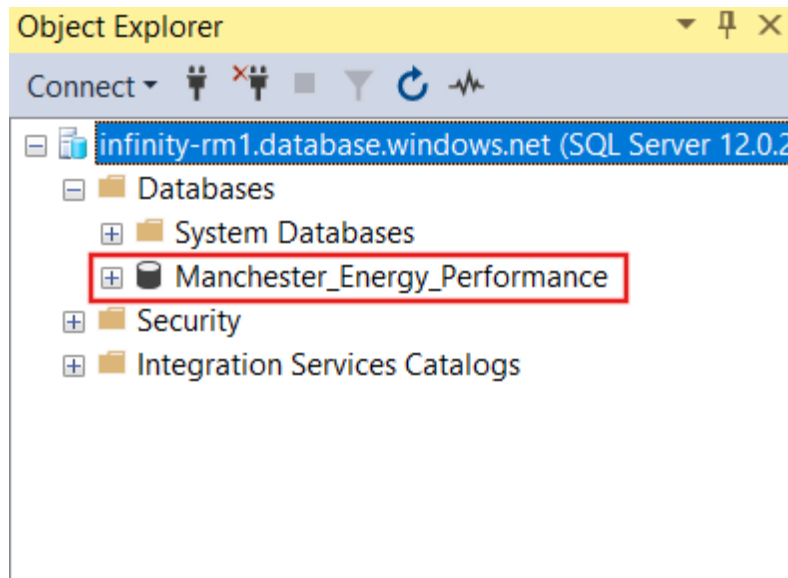


Figure 23

- Create a database called **Manchester\_Energy\_Performance**.

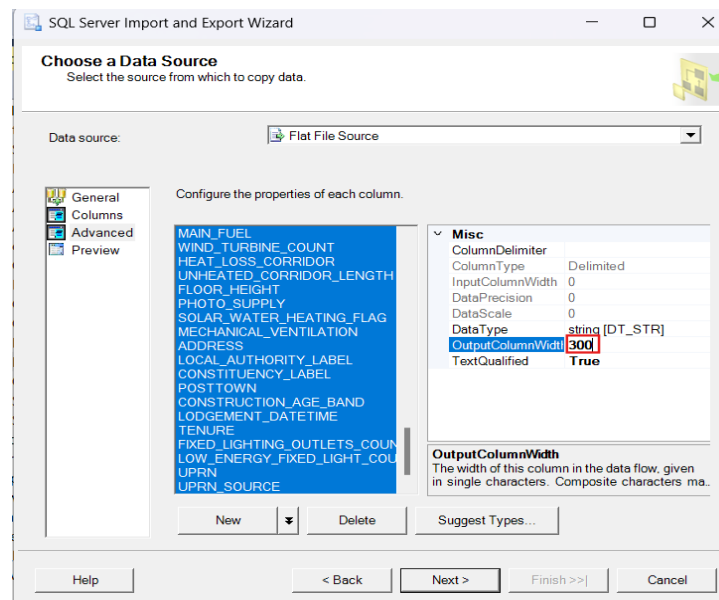
```
USE master;
GO
CREATE DATABASE Manchester_Energy_Performance
```

*Figure 24*



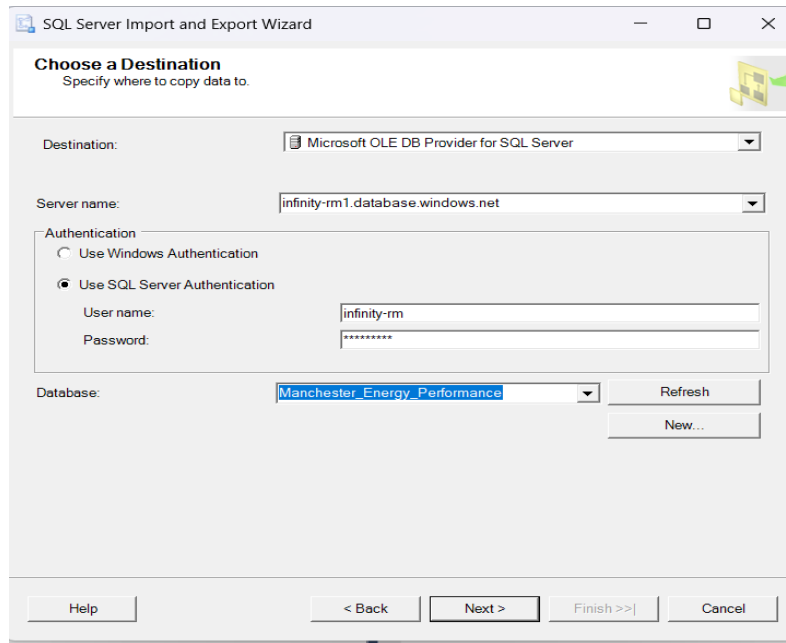
*Figure 25*

- The steps in importing the datasets using the SQL Server Import and Export wizard is similar to *Figure 1* to *Figure 4* in Task 1.
- After selecting all the columns from the dataset, the **Output Column Width** is changed to 300.



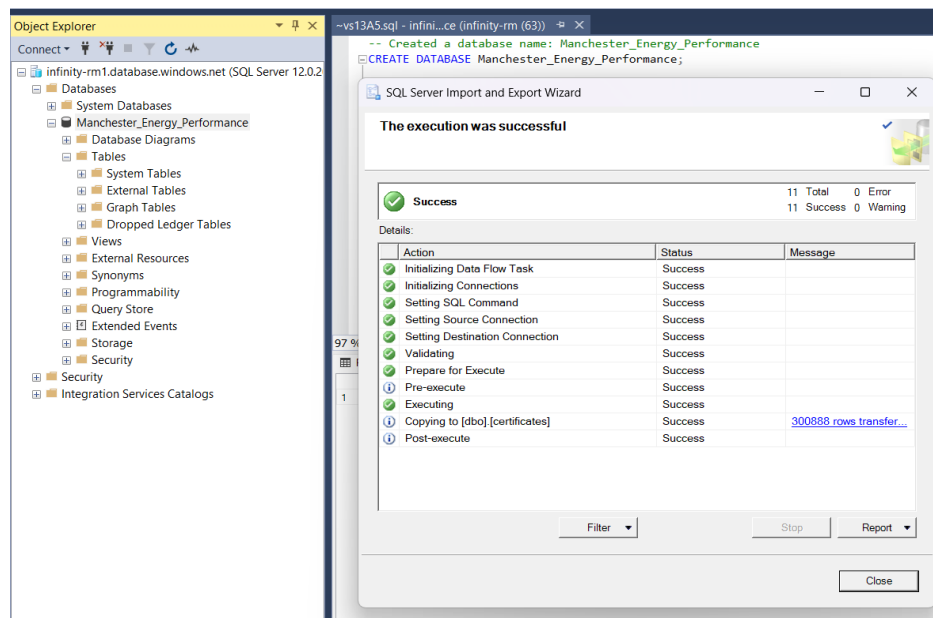
*Figure 26*

- Changing the destination of the of SQL Server Import and Export Wizard to Microsoft OLE DB Provider for SQL Server and changing the authentication to SQL Server Authentication while providing azure login credentials.



*Figure 27*

- The next steps of the Importing process SQL Server are similar to **Figure 6** and **Figure 7** of Task 1.
- The execution dialog box appears, if all the data has loaded and the execution was successful click close.



*Figure 28*

- Creating 6 View Queries which are based on the data from the year 2013 to 2023. During the process **13 columns** were removed.
- The views are created for the ease in creating the dashboard for the visual analysis of the energy performance of Greater Manchester. For the modification of each Views the essential variable columns are considered. The SQL queries for the 6 views are as follows;

**i. View 1: Current energy levels of Greater Manchester.**

```
CREATE VIEW Manchester_Current_Energy1 AS
SELECT -- Selecting the necessary columns needed for the analysis
    COALESCE(NULLIF([LMK_KEY], ''), 'NA') AS [LMK_KEY], -- Each columns Missing values are replaced as 'NA'
    COALESCE(NULLIF([CURRENT_ENERGY_RATING], ''), 'NA') AS [CURRENT_ENERGY_RATING],
    COALESCE(NULLIF([CURRENT_ENERGY_EFFICIENCY], ''), 'NA') AS [CURRENT_ENERGY_EFFICIENCY],
    COALESCE(NULLIF([ENVIRONMENT_IMPACT_CURRENT], ''), 'NA') AS [ENVIRONMENT_IMPACT_CURRENT],
    COALESCE(NULLIF([ENERGY_CONSUMPTION_CURRENT], ''), 'NA') AS [ENERGY_CONSUMPTION_CURRENT],
    COALESCE(NULLIF([CO2_EMISSIONS_CURRENT], ''), 'NA') AS [CO2_EMISSIONS_CURRENT],
    COALESCE(NULLIF([CO2_EMISS_CURR_PER_FLOOR_AREA], ''), 'NA') AS [CO2_EMISS_CURR_PER_FLOOR_AREA],
    COALESCE(NULLIF([LIGHTING_COST_CURRENT], ''), 'NA') AS [LIGHTING_COST_CURRENT],
    COALESCE(NULLIF([HEATING_COST_CURRENT], ''), 'NA') AS [HEATING_COST_CURRENT],
    COALESCE(NULLIF([HOT_WATER_COST_CURRENT], ''), 'NA') AS [HOT_WATER_COST_CURRENT]
FROM [dbo].[certificates]
WHERE CONVERT(DATE, [INSPECTION_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31'
AND CONVERT(DATE, [LODGEMENT_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31';
```

*Figure 29*

**ii. View 2: Potential energy levels of Greater Manchester.**

```
CREATE VIEW Manchester_Potential_Energy1 AS
SELECT -- Selecting the necessary columns needed for the analysis
    COALESCE(NULLIF([LMK_KEY], ''), 'NA') AS [LMK_KEY], -- Each columns Missing values are replaced as 'NA'
    COALESCE(NULLIF([POTENTIAL_ENERGY_RATING], ''), 'NA') AS [POTENTIAL_ENERGY_RATING],
    COALESCE(NULLIF([POTENTIAL_ENERGY_EFFICIENCY], ''), 'NA') AS [POTENTIAL_ENERGY_EFFICIENCY],
    COALESCE(NULLIF([ENVIRONMENT_IMPACT_POTENTIAL], ''), 'NA') AS [ENVIRONMENT_IMPACT_POTENTIAL],
    COALESCE(NULLIF([ENERGY_CONSUMPTION_POTENTIAL], ''), 'NA') AS [ENERGY_CONSUMPTION_POTENTIAL],
    COALESCE(NULLIF([CO2_EMISSIONS_POTENTIAL], ''), 'NA') AS [CO2_EMISSIONS_POTENTIAL],
    COALESCE(NULLIF([LIGHTING_COST_POTENTIAL], ''), 'NA') AS [LIGHTING_COST_POTENTIAL],
    COALESCE(NULLIF([HEATING_COST_POTENTIAL], ''), 'NA') AS [HEATING_COST_POTENTIAL],
    COALESCE(NULLIF([HOT_WATER_COST_POTENTIAL], ''), 'NA') AS [HOT_WATER_COST_POTENTIAL]
FROM [dbo].[certificates]
WHERE CONVERT(DATE, [INSPECTION_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31'
AND CONVERT(DATE, [LODGEMENT_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31';
```

*Figure 30*

**iii. View 3: Energy efficiency rates of Greater Manchester.**

```
CREATE VIEW Energy_Efficiency_Rating1 AS
SELECT -- Selecting the necessary columns needed for the analysis
    COALESCE(NULLIF([LMK_KEY], ''), 'NA') AS [LMK_KEY], -- Each columns Missing values are replaced as 'NA'
    COALESCE(NULLIF([ENERGY_CONSUMPTION_CURRENT], ''), 'NA') AS [ENERGY_CONSUMPTION_CURRENT],
    COALESCE(NULLIF([ENERGY_CONSUMPTION_POTENTIAL], ''), 'NA') AS [ENERGY_CONSUMPTION_POTENTIAL],
    COALESCE(NULLIF([HOT_WATER_ENERGY_EFF], ''), 'NA') AS [HOT_WATER_ENERGY_EFF],
    COALESCE(NULLIF([FLOOR_ENERGY_EFF], ''), 'NA') AS [FLOOR_ENERGY_EFF],
    COALESCE(NULLIF([WINDOWS_ENERGY_EFF], ''), 'NA') AS [WINDOWS_ENERGY_EFF],
    COALESCE(NULLIF([WALLS_ENERGY_EFF], ''), 'NA') AS [WALLS_ENERGY_EFF],
    COALESCE(NULLIF([SHEATING_ENERGY_EFF], ''), 'NA') AS [SHEATING_ENERGY_EFF],
    COALESCE(NULLIF([ROOF_ENERGY_EFF], ''), 'NA') AS [ROOF_ENERGY_EFF],
    COALESCE(NULLIF([MAINHEAT_ENERGY_EFF], ''), 'NA') AS [MAINHEAT_ENERGY_EFF],
    COALESCE(NULLIF([MAINHEATC_ENERGY_EFF], ''), 'NA') AS [MAINHEATC_ENERGY_EFF]
FROM [dbo].[certificates]
WHERE CONVERT(DATE, [INSPECTION_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31'
AND CONVERT(DATE, [LODGEMENT_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31';
```

*Figure 31*



**iv. View 4: Environmental efficiency rates of Greater Manchester.**

```
CREATE VIEW Environmental_Efficiency_Rating1 AS
SELECT -- Selecting the necessary columns needed for the analysis
    COALESCE(NULLIF([LMK_KEY], ''), 'NA') AS [LMK_KEY], -- Each columns Missing values are replaced as 'NA'
    COALESCE(NULLIF([ENVIRONMENT_IMPACT_CURRENT], ''), 'NA') AS [ENVIRONMENT_IMPACT_CURRENT],
    COALESCE(NULLIF([ENVIRONMENT_IMPACT_POTENTIAL], ''), 'NA') AS [ENVIRONMENT_IMPACT_POTENTIAL],
    COALESCE(NULLIF([HOT_WATER_ENV_EFF], ''), 'NA') AS [HOT_WATER_ENV_EFF],
    COALESCE(NULLIF([FLOOR_ENV_EFF], ''), 'NA') AS [FLOOR_ENV_EFF],
    COALESCE(NULLIF([WINDOWS_ENV_EFF], ''), 'NA') AS [WINDOWS_ENV_EFF],
    COALESCE(NULLIF([WALLS_ENV_EFF], ''), 'NA') AS [WALLS_ENV_EFF],
    COALESCE(NULLIF([SHEATING_ENV_EFF], ''), 'NA') AS [SHEATING_ENV_EFF],
    COALESCE(NULLIF([ROOF_ENV_EFF], ''), 'NA') AS [ROOF_ENV_EFF],
    COALESCE(NULLIF([MAINHEAT_ENV_EFF], ''), 'NA') AS [MAINHEAT_ENV_EFF],
    COALESCE(NULLIF([MAINHEATC_ENV_EFF], ''), 'NA') AS [MAINHEATC_ENV_EFF]
FROM [dbo].[certificates]
WHERE CONVERT(DATE, [INSPECTION_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31'
AND CONVERT(DATE, [LODGEMENT_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31';
```

*Figure 32*

**v. View 5: Building Locations of Greater Manchester.**

```
CREATE VIEW Building_Location1 AS
SELECT -- Selecting the necessary columns needed for the analysis
    COALESCE(NULLIF([LMK_KEY], ''), 'NA') AS [LMK_KEY], -- Each columns Missing values are replaced as 'NA'
    COALESCE(NULLIF([POSTCODE], ''), 'NA') AS [POSTCODE],
    COALESCE(NULLIF([BUILDING_REFERENCE_NUMBER], ''), 'NA') AS [BUILDING_REFERENCE_NUMBER],
    COALESCE(NULLIF([PROPERTY_TYPE], ''), 'NA') AS [PROPERTY_TYPE],
    COALESCE(NULLIF([BUILT_FORM], ''), 'NA') AS [BUILT_FORM],
    COALESCE(NULLIF([INSPECTION_DATE], ''), 'NA') AS [INSPECTION_DATE],
    COALESCE(NULLIF([LOCAL_AUTHORITY], ''), 'NA') AS [LOCAL_AUTHORITY],
    COALESCE(NULLIF([LODGEMENT_DATE], ''), 'NA') AS [LODGEMENT_DATE],
    COALESCE(NULLIF([TRANSACTION_TYPE], ''), 'NA') AS [TRANSACTION_TYPE],
    COALESCE(NULLIF([CONSTITUENCY], ''), 'NA') AS [CONSTITUENCY],
    COALESCE(NULLIF([ADDRESS], ''), 'NA') AS [ADDRESS],
    COALESCE(NULLIF([LOCAL_AUTHORITY_LABEL], ''), 'NA') AS [LOCAL_AUTHORITY_LABEL],
    COALESCE(NULLIF([CONSTITUENCY_LABEL], ''), 'NA') AS [CONSTITUENCY_LABEL],
    COALESCE(NULLIF([POSTTOWN], ''), 'NA') AS [POSTTOWN],
    COALESCE(NULLIF([CONSTRUCTION_AGE_BAND], ''), 'NA') AS [CONSTRUCTION_AGE_BAND],
    COALESCE(NULLIF([LODGEMENT_DATETIME], ''), 'NA') AS [LODGEMENT_DATETIME],
    COALESCE(NULLIF([TENURE], ''), 'NA') AS [TENURE],
    COALESCE(NULLIF([UPRN], ''), 'NA') AS [UPRN],
    COALESCE(NULLIF([UPRN_SOURCE], ''), 'NA') AS [UPRN_SOURCE]
FROM [dbo].[certificates]
WHERE CONVERT(DATE, [INSPECTION_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31'
AND CONVERT(DATE, [LODGEMENT_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31';
```

*Figure 33*

## vi. View 6: Building energy rates of Greater Manchester.

```

CREATE VIEW Building_Energy1 AS
SELECT
    -- Selecting the necessary columns needed for the analysis
    COALESCE(NULLIF([LMK_KEY], ''), 'NA') AS [LMK_KEY], -- Each columns Missing values are replaced as 'NA'
    COALESCE(NULLIF([TOTAL_FLOOR_AREA], ''), 'NA') AS [TOTAL_FLOOR_AREA],
    COALESCE(NULLIF([ENERGY_TARIFF], ''), 'NA') AS [ENERGY_TARIFF],
    COALESCE(NULLIF([MAINS_GAS_FLAG], ''), 'NA') AS [MAINS_GAS_FLAG],
    COALESCE(NULLIF([FLOOR_LEVEL], ''), 'NA') AS [FLOOR_LEVEL],
    COALESCE(NULLIF([FLAT_TOP_STOREY], ''), 'NA') AS [FLAT_TOP_STOREY],
    COALESCE(NULLIF([FLAT_STOREY_COUNT], ''), 'NA') AS [FLAT_STOREY_COUNT],
    COALESCE(NULLIF([MAIN_HEATING_CONTROLS], ''), 'NA') AS [MAIN_HEATING_CONTROLS],
    COALESCE(NULLIF([MULTI_GLAZE_PROPORTION], ''), 'NA') AS [MULTI_GLAZE_PROPORTION],
    COALESCE(NULLIF([GLAZED_TYPE], ''), 'NA') AS [GLAZED_TYPE],
    COALESCE(NULLIF([GLAZED_AREA], ''), 'NA') AS [GLAZED_AREA],
    COALESCE(NULLIF([EXTENSION_COUNT], ''), 'NA') AS [EXTENSION_COUNT],
    COALESCE(NULLIF([NUMBER_HABITABLE_ROOMS], ''), 'NA') AS [NUMBER_HABITABLE_ROOMS],
    COALESCE(NULLIF([NUMBER_HEATED_ROOMS], ''), 'NA') AS [NUMBER_HEATED_ROOMS],
    COALESCE(NULLIF([LOW_ENERGY_LIGHTING], ''), 'NA') AS [LOW_ENERGY_LIGHTING],
    COALESCE(NULLIF([NUMBER_OPEN_FIREPLACES], ''), 'NA') AS [NUMBER_OPEN_FIREPLACES],
    COALESCE(NULLIF([MAIN_FUEL], ''), 'NA') AS [MAIN_FUEL],
    COALESCE(NULLIF([WIND_TURBINE_COUNT], ''), 'NA') AS [WIND_TURBINE_COUNT],
    COALESCE(NULLIF([HEAT_LOSS_CORRIDOR], ''), 'NA') AS [HEAT_LOSS_CORRIDOR],
    COALESCE(NULLIF([UNHEATED_CORRIDOR_LENGTH], ''), 'NA') AS [UNHEATED_CORRIDOR_LENGTH],
    COALESCE(NULLIF([FLOOR_HEIGHT], ''), 'NA') AS [FLOOR_HEIGHT],
    COALESCE(NULLIF([PHOTO_SUPPLY], ''), 'NA') AS [PHOTO_SUPPLY],
    COALESCE(NULLIF([SOLAR_WATER_HEATING_FLAG], ''), 'NA') AS [SOLAR_WATER_HEATING_FLAG],
    COALESCE(NULLIF([MECHANICAL_VENTILATION], ''), 'NA') AS [MECHANICAL_VENTILATION],
    COALESCE(NULLIF([FIXED_LIGHTING_OUTLETS_COUNT], ''), 'NA') AS [FIXED_LIGHTING_OUTLETS_COUNT],
    COALESCE(NULLIF([LOW_ENERGY_FIXED_LIGHT_COUNT], ''), 'NA') AS [LOW_ENERGY_FIXED_LIGHT_COUNT]
FROM [dbo].[certificates]
WHERE CONVERT(DATE, [INSPECTION_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31'
AND CONVERT(DATE, [LODGE_MENT_DATE], 103) BETWEEN '2013-01-01' AND '2023-12-31';

```

Figure 34

- Query which provides the use of system functions, including aggregate and ranking functions.

```

-- Number of rows after creating the view
SELECT COUNT(*)
FROM [dbo].[Building_Energy1]

-- Total number of records and the average energy efficiency.
SELECT
    COUNT(*) AS TotalRecords,
    AVG(CAST([CURRENT_ENERGY_EFFICIENCY] AS FLOAT)) AS AvgEnergyEfficiency
FROM [dbo].[Manchester_Current_Energy1];

-- Ranking the records by their current energy efficiency in descending order.
SELECT
    [POSTCODE],
    [CURRENT_ENERGY_EFFICIENCY],
    ROW_NUMBER() OVER (ORDER BY [CURRENT_ENERGY_EFFICIENCY] DESC) AS Rank
FROM [dbo].[Building_Location1], [dbo].[Manchester_Current_Energy1];

-- Calculates the length of each postcode and also returns the current date.
SELECT
    [POSTCODE],
    LEN([POSTCODE]) AS PostcodeLength,
    GETDATE() AS CurrentDate
FROM [dbo].[Building_Location1];

```

Figure 35

- After the creation of the Views. the views will be imported to Microsoft Power BI for the visual analysis of the Domestic Energy Performance of Greater Manchester.

## 2.4. Key Patterns and Insights.

- The dashboard illustration of the Greater Manchester Domestic Energy Performance from the year 2013 to 2023. The dashboard consists of a Home page, Overview page, a page dedicated to the Energy Efficiency and another page dedicated for the Environmental Impact.
- The dashboard consists of charts, graphs, cards and slicers.

### Overview page

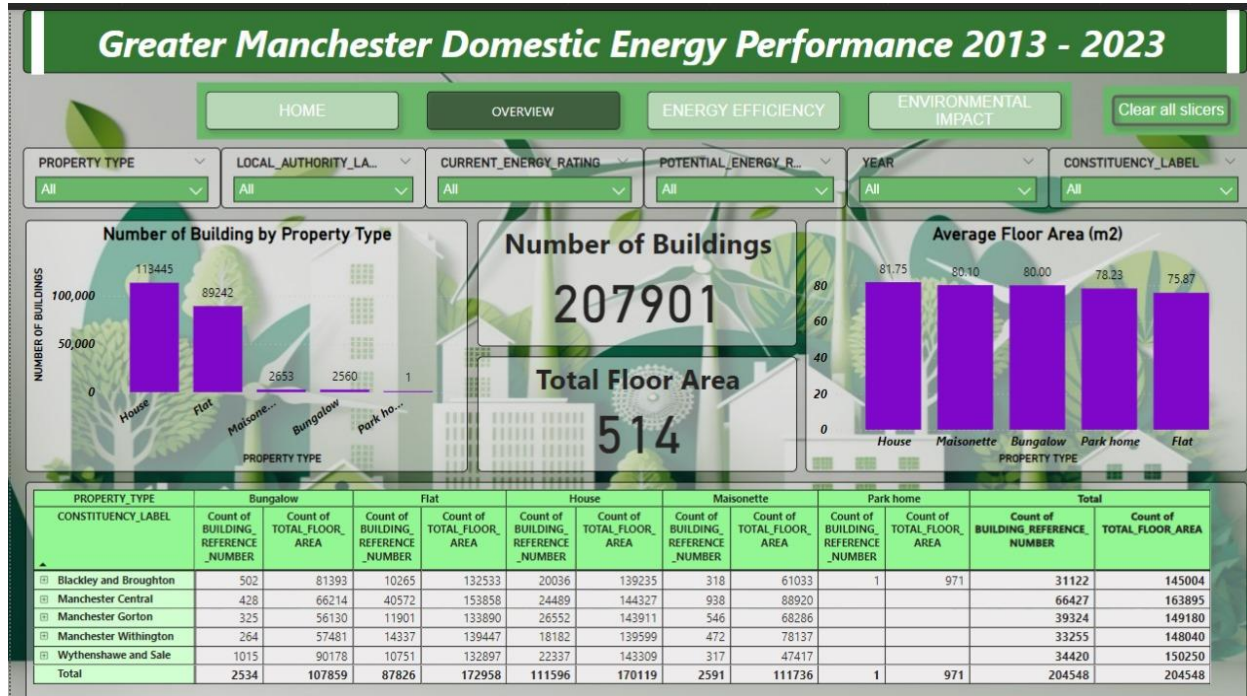


Figure 36

- The overview page provides visual analysis about the buildings, the property types, building types and the floor area taken by each building property.
- The page consists of a matrix, bar charts, cards and slicers.
- Number of Building by Property Type:** The bar chart shows the number of buildings over the property type. Accordingly, the property type house has the highest number of buildings all over the Greater Manchester which is 113,445 while property type park home has the least number of buildings all over the Greater Manchester which is 1.
- Average Floor Area (m2):** the bar chart represents the average floor area of each property area. Accordingly, the property type house has largest average floor area which is 81.75m2 while property type flats have the smallest average floor area which is 75.87m2.
- Property Type Table:** the matrix table breakdowns the number of buildings, the total floor area of each property type within different constituencies in Greater Manchester. According to the matrix, Manchester Central has the highest total number of property types and the highest total property floor area.



## Energy Efficiency

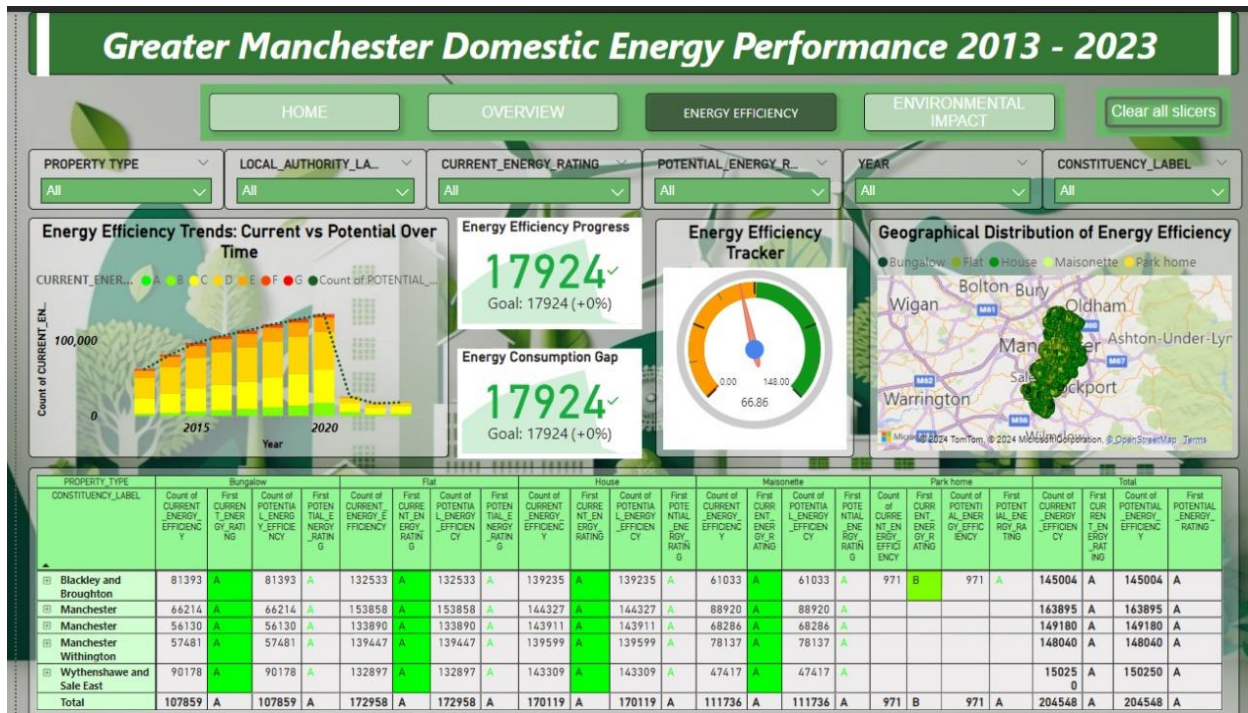


Figure 37

- The energy efficiency page provides visual analysis about the energy consumption, energy efficiency, and geographical distribution of energy efficiency across Greater Manchester.
- The page consists of a line and stacked column chart, a dial gauge, a bubble map, cards and slicers.
- **Energy Efficiency Trends: Current vs Potential Over Time:** The line and stacked column chart provides a visual analysis about the current energy efficiency ratings over time. The ratings are shown from A to G where rating A is the most energy efficient while G is the least energy efficient. The trend shows a steady improvement from the least rates – G, F, E and D to most energy rates – A, B, C which is peaking around 2020. The potential energy rates suggest a target of achievement around A or B if improvements are made.
- **Energy Efficiency Tracker:** The tracker gauge measures the current energy efficiency rate which is visualized using a dial gauge. According to the tracker, it indicates a rate of 66.86% which indicates an average energy efficiency performance showing room for more improvement.
- **Geographical Distribution of Energy Efficiency:** The bubble map visualization technique is used to show the distribution of energy-efficient properties across Greater Manchester. The colored clusters represent the different property types where the clusters suggests that the majority of energy-efficient properties are concentrated in Central Manchester and surrounding urban areas.
- **Property Type Table:** The matrix table provides a detailed count of current and potential energy efficiency ratings for different property types and constituencies.

## Environmental Impact

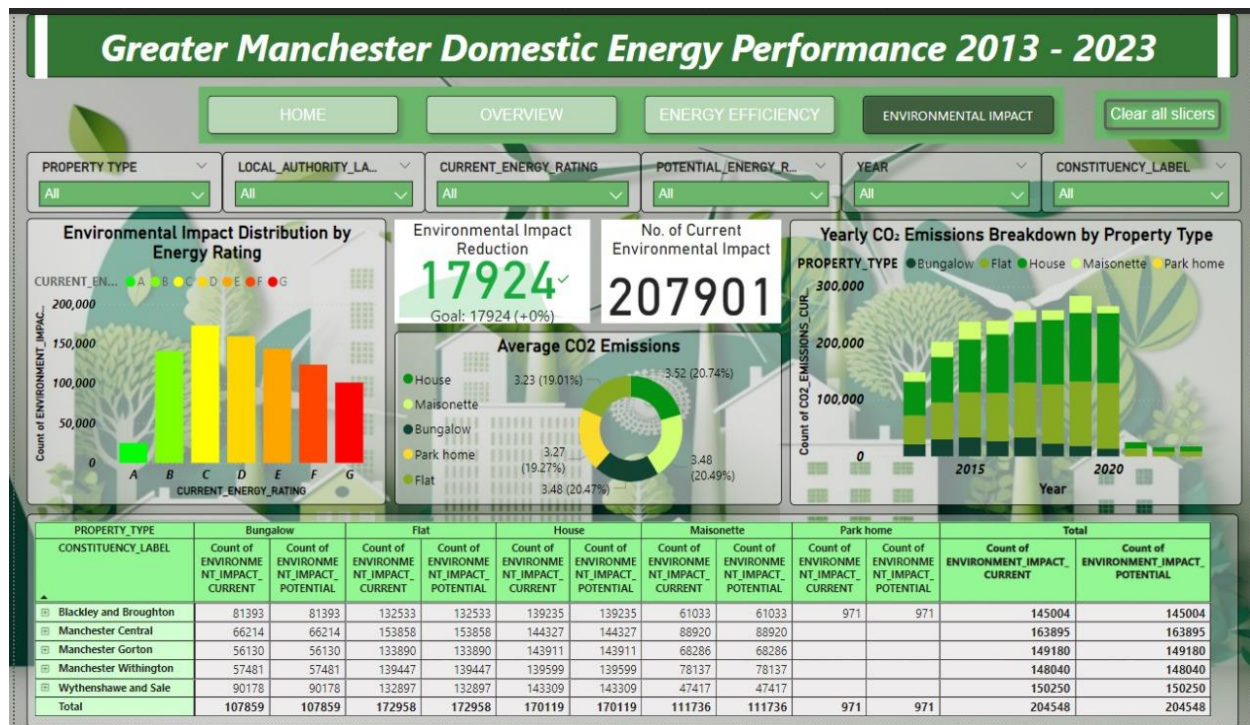


Figure 38

- The environmental impact page provides visual analysis about the distribution of the environmental impact, the distributions annual and average carbon emission.
- The page consists of a column chart, a donut pie chart, a stacked column chart, cards and slicers.
- Environmental Impact Distribution by Energy Rating:** The column chart is used to show the distribution of the environmental impact towards the energy ratings. According to the distribution, when the environmental impact increases, the energy ratings decrease. Accordingly, most properties fall under the low energy rates C and D which are associated with higher environmental impacts around 150,000+. Properties with A and B ratings contribute the least to environmental damage, indicating the need for energy upgrades to improve the distribution toward these categories.
- Average CO2 Emissions:** The donut pie chart shows the average distribution of CO2 emissions of each property type. Accordingly, house property type has the highest average CO2 emissions which is 3.52 tons with a percentage of 20.74% while flat property type has the least average CO2 emissions which is 3.23 tons with a percentage of 19.01%.
- Yearly CO2 Emissions Breakdown by Property Type:** The stacked column chart shows the yearly distribution of the CO2 emission breakdowns. A visible peak in the emissions between the years 2015 to 2020 which is followed by a reduction which is most likely due to the improvements in the energy efficiency.
- Property Type Table:** The matrix table provides details about each property type and constituency. The matrix covers both current and potential environmental impacts.

## 2.5. Recommendations

The datasets acquired from the Energy Performance of Buildings Data England and Wales for readily provide recommendations for the energy efficiency improvements in Greater Manchester which covers the improvement in fields such as type, cost and description.

LMK_KEY	IMPROVEMENT_ITEM	IMPROVEMENT_SUMMARY_TEXT	IMPROVEMENT_DESCR_TEXT
20e4f9426b3cec3b7a4ba942	1	Cavity wall insulation	Cavity wall insulation
20e4f9426b3cec3b7a4ba942	2	Internal or external wall insulation	Internal or external wall insulation
20e4f9426b3cec3b7a4ba942	3	Floor insulation (suspended floor)	Floor insulation (suspended floor)
20e4f9426b3cec3b7a4ba942	4	Solar water heating	Solar water heating
20e4f9426b3cec3b7a4ba942	5	Solar photovoltaic panels, 2.5 kWp	Solar photovoltaic panels
5.45384E+32	1	Low energy lighting for all fixed outlets	Low energy lighting
5.45384E+32	2	Replace boiler with new condensing boiler	Replace boiler with new condensing boiler
5.45384E+32	3	Solar water heating	Solar water heating
5.45384E+32	4	Solar photovoltaic panels, 2.5 kWp	Solar photovoltaic panels
2.27229E+32	1	Cavity wall insulation	Cavity wall insulation
2.27229E+32	2	Floor insulation	Floor insulation
2.27229E+32	3	Low energy lighting for all fixed outlets	Low energy lighting
2.27229E+32	4	Heating controls (room thermostat and TRVs)	Heating controls (room thermostat and TRVs)
2.27229E+32	5	Replace boiler with new condensing boiler	Condensing boiler
2.27229E+32	6	Flue gas heat recovery device in conjunction with boiler	Flue gas heat recovery
2.27229E+32	7	Solar photovoltaic panels, 2.5 kWp	Solar photovoltaic panels
2.27229E+32	8	Wind turbine	Wind turbine
1.40568E+33	1	Cavity wall insulation	Cavity wall insulation
1.40568E+33	2	High heat retention storage heaters and dual immersion cylinder	High heat retention storage heaters
2.88425E+32	2		A solar water heating panel, usually fixed to the roof, uses the sun to pre-heat the hot water supply. This will significantly reduce the d
2.88425E+32	1		Replacement of traditional light bulbs with energy saving recommended ones will reduce lighting costs over the lifetime of the bulb, an
2.88425E+32	3		A solar PV system is one which converts light directly into electricity via panels placed on the roof with no waste and no emissions. Thi
25cb44ed6c67d568b7de44f0	1	Floor insulation (suspended floor)	Floor insulation (suspended floor)
25cb44ed6c67d568b7de44f0	2	Solar water heating	Solar water heating
25cb44ed6c67d568b7de44f0	3	Solar photovoltaic panels, 2.5 kWp	Solar photovoltaic panels
25d00f64a40b898bea970f5di	1	High heat retention storage heaters	High heat retention storage heaters
1.0726E+33	1	Draught proofing	Draught proofing
1.0726E+33	2	Low energy lighting for all fixed outlets	Low energy lighting

Figure 39

Some of the recommendations provided by the Energy Performance of Buildings Data England and Wales are as follows.

- Upgrading Lighting to LED Bulbs.
- Installing a smart thermostat to better control heating.
- Sealing the gaps around windows, doors, and chimneys to prevent heat loss.
- Installing water-efficient showerheads.
- Adding an insulation to attic hatches to prevent heat from escaping.

Other additional recommendations that would show significant improvements in the domestic energy performance of Greater Manchester are as follows.

- Installing an air-source or ground-source heat pumps to replace the traditional heating systems. This will provide renewable energy to provide heating and cooling which would significantly reduce the energy bills and carbon footprint.
- Installing a living green roof to improve installation and to manage rainwater. This will enhance thermal insulation while reducing stormwater runoff.
- Installing internal shutters or thermal-lined heavy curtains to improve window insulation. This will help trap heat inside the home especially during the winter season which reduces the heating requirements.

## **2.6. Conclusion**

The trends on the energy performance of residential buildings in Greater Manchester are from the years 2013 to 2023. They reveal new data about the energy efficiency situation in the region. This report draws from Energy Performance Certificates (EPCs) databases and evaluates the effectiveness of insulation and heating system types, the impact of energy rating on the environment, and essential patterns of energy use of properties.

The findings indicate that energy performance ratings have been on upward trend; a shift from low to high energy performance ratings (A-C) was noticed about 2020. Still, the current average energy efficiency of 66.86% shows that there is much room for improvement. The highest average CO2 emissions are from households located in the Central Manchester area, while buildings in that area seem to exhibit a high level of energy efficiency, which calls for targeted measures.

The report further mentions other installations too like modern heating systems such as air source or ground source heat pumps fittings, retrofitting lighting, improving insulation, among other activities that can improve energy efficiency and the implementation of these measures would enable Greater Manchester to be more environmentally friendly and have low carbon by reducing energy use and lower emissions.

Consequently, the current analysis is not only informative for the city authority and property owners, but it is also important for establishing energy efficiency policy for Greater Manchester in the near future. As a result of continuous improvements in energy performance, the region will be able to achieve its goal of minimizing carbon output and reach its sustainability objectives.



#### 4. References

1. app.powerbi.com. (n.d.). *Microsoft Power BI*. [online] Available at: <https://app.powerbi.com/view?r=eyJrIjojNTI3ODI0ODktMDAxMS00NGQ2LWJmYTItMTA2MzA4YjZjMjBjIiwidCI6ImJmMzQ2ODEwLTljN2QtNDNkZS1hODcyLTl0YTJlZjM5O TVhOCJ9&pageName=ReportSectionabd88355d2a923eacb50>.
2. beta.ukdataservice.ac.uk. (n.d.). *UK Data Service > Series*. [online] Available at: <https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000060#>.
3. DiscoverEI (2022). *#15 PowerBI4Enviro's - Global CO2 Emissions: Power BI Story Telling with Pragati Jain*. [online] YouTube. Available at: <https://www.youtube.com/watch?v=y8uJxzEiuo> [Accessed 13 Oct. 2024].
4. epc.opendatacommunities.org. (n.d.). *Energy Performance of Buildings Data England and Wales*. [online] Available at: <https://epc.opendatacommunities.org/>.
5. jusTech talk (2023). *Installing Microsoft SQL Server Reporting Services (SSRS) / Step-by-Step Guide*. [online] YouTube. Available at: <https://www.youtube.com/watch?v=sELwsF4WsdM> [Accessed 25 Aug. 2024].
6. maggiesMSFT (n.d.). *What is SQL Server Reporting Services - SQL Server Reporting Services (SSRS)*. [online] learn.microsoft.com. Available at: <https://learn.microsoft.com/en-us/sql/reporting-services/create-deploy-and-manage-mobile-and-paginated-reports?view=sql-server-ver16>.
7. PK: An Excel Expert (2022). *Energy Consumptions Dashboard in Power BI / Step by Step complete tutorial*. [online] YouTube. Available at: <https://www.youtube.com/watch?v=CEDzvrUsSg8> [Accessed 13 Oct. 2024].
8. Sql Training Sessions (2020). *Drill Through Reports in SSRS / BIDS / SQL Training Sessions*. [online] YouTube. Available at: [https://www.youtube.com/watch?v=M\\_BdXjf4aJk](https://www.youtube.com/watch?v=M_BdXjf4aJk) [Accessed 25 Aug. 2024].
9. United Nations (2023a). *Goal 1: End Poverty in All Its Forms Everywhere*. [online] United Nations Sustainable Development. Available at: <https://www.un.org/sustainabledevelopment/poverty/>.
10. United Nations (2023b). *Poverty / UN Global Compact*. [online] unglobalcompact.org. Available at: <https://unglobalcompact.org/what-is-gc/our-work/social/poverty>.
11. WiseOwlTutorials (2021). *SSRS Report Builder Part 12.1 - Drill Through Reports using Tables*. [online] YouTube. Available at: <https://www.youtube.com/watch?v=7Qk8BWLwGPI> [Accessed 25 Aug. 2024].



## **5.Appendices**

**Figure 01:** Importing the datasets to the database of Task 1.

**Figure 02:** Import data and Export Wizard welcome page.

**Figure 03:** Selecting the flat file source as the data source.

**Figure 04:** Selecting the datasets that are needed to be imported.

**Figure 05:** Changing the column width of the variable columns of the datasets.

**Figure 06:** Moving onwards towards the next step.

**Figure 07:** Finishing the wizard.

**Figure 08:** Completion of the successful execution.

**Figure 09:** Query to check the data of the imported datasets.

**Figure 10:** Query to understand the dataset by exploring the first 100 data points of the dataset.

**Figure 11:** Query to identify the column variable names and their data types.

**Figure12:** Adding a new column variable to the two datasets called **Country\_Name**.

**Figure 13 a:** Query to view the table for country Ethiopia for the analysis.

**Figure 13 b:** Query to view the table for country Vietnam for the analysis.

**Figure 14:** Query to view the table views created.

**Figure 15 a:** Query to view the for the poverty measurement index: **Income per capita**.

**Figure 15 b:** Query to view the for the poverty measurement index: **Education access**.

**Figure 15 c:** Query to view the for the poverty measurement index: **Health expenditure per child**.

**Figure 15 d:** Query to view the for the poverty measurement index: **Living Standards (access to clean pure water)**.

**Figure 15 e:** Query to view the for the poverty measurement index: **Malnutrition rate**.

**Figure 16:** Report design for the Income per capita.

**Figure 17:** Report design for the Access to proper Education.

**Figure 18:** Report design for the Health expenditure per child.

**Figure 19:** Report design for the Living Standards: Access to clean and pure water.

**Figure 20:** Report design for the Malnutrition Rate.

**Figure 21:** The dataset for Task 2: Energy Performance of Buildings at Greater Manchester.

**Figure 22:** Final view of the created database in the Microsoft Azure Portal.

**Figure 23:** Azure login credentials for the SQL Server Authentication.

**Figure 24:** SQL Query to create the database **Manchester\_Energy\_Performance**.

**Figure 25:** Database for Task 2: **Manchester\_Energy\_Performance**.

**Figure 26:** Changing the column width of the variable columns of the datasets.

**Figure 27:** Changing the authentication server and the server destination of the SQL Server.

**Figure 28:** Completion of the successful execution.

**Figure 29:** View query for the **Current energy levels of Greater Manchester**.

**Figure 30:** View query for the **Potential energy levels of Greater Manchester**.

**Figure 31:** View query for the **Energy efficiency rates of Greater Manchester**.

**Figure 32:** View query for the **Environmental efficiency rates of Greater Manchester**.

**Figure 33:** View query for the **Building locations of Greater Manchester**.

**Figure 34:** View query for the **Building energy rates of Greater Manchester**.

**Figure 35:** Query for the use of system functions, including aggregate and ranking functions of the created views.

**Figure 36:** Dashboard illustration of the Overview page.

**Figure 37:** Dashboard illustration of the Energy efficiency page.

**Figure 38:** Dashboard illustration of the Environmental impact page.

**Figure 39:** Recommendation dataset provided by the Energy Performance of Buildings Data England and Wales of Greater Manchester.