Data Visualization Competition

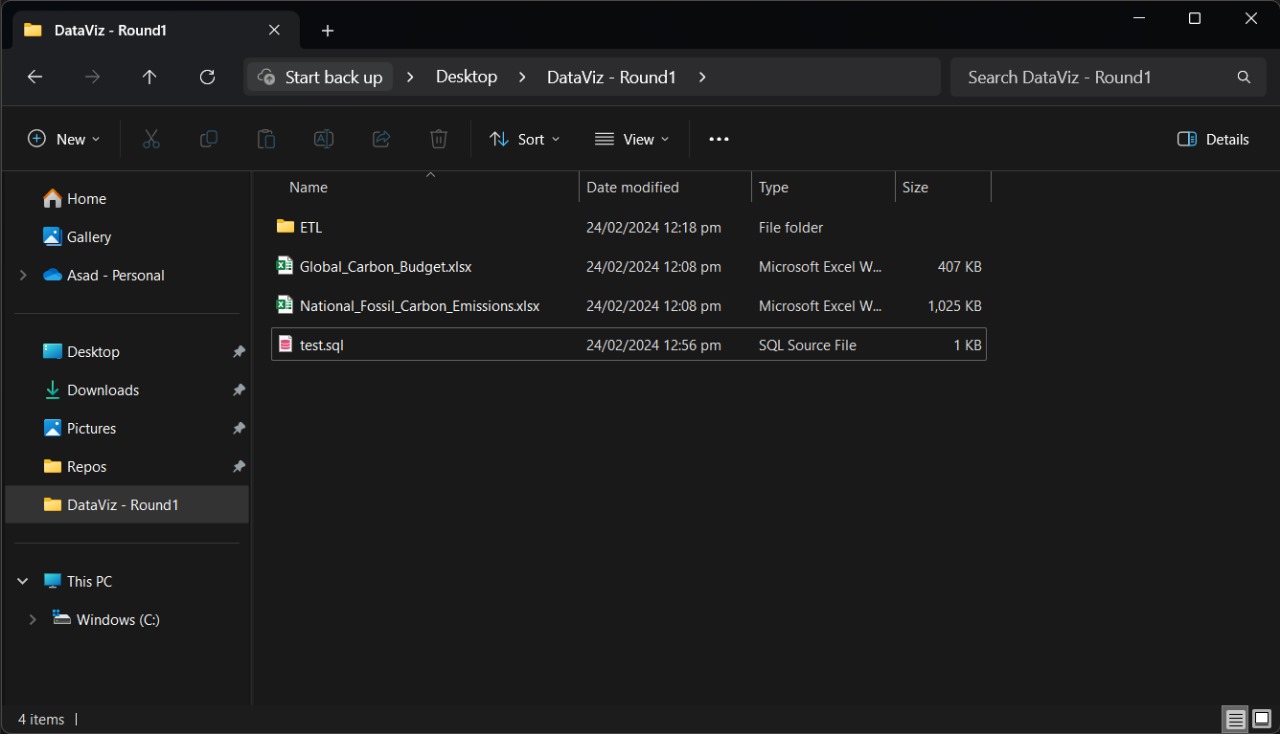
Team: Visualizers

Members: Asad Ali, M Haroon Shahzad

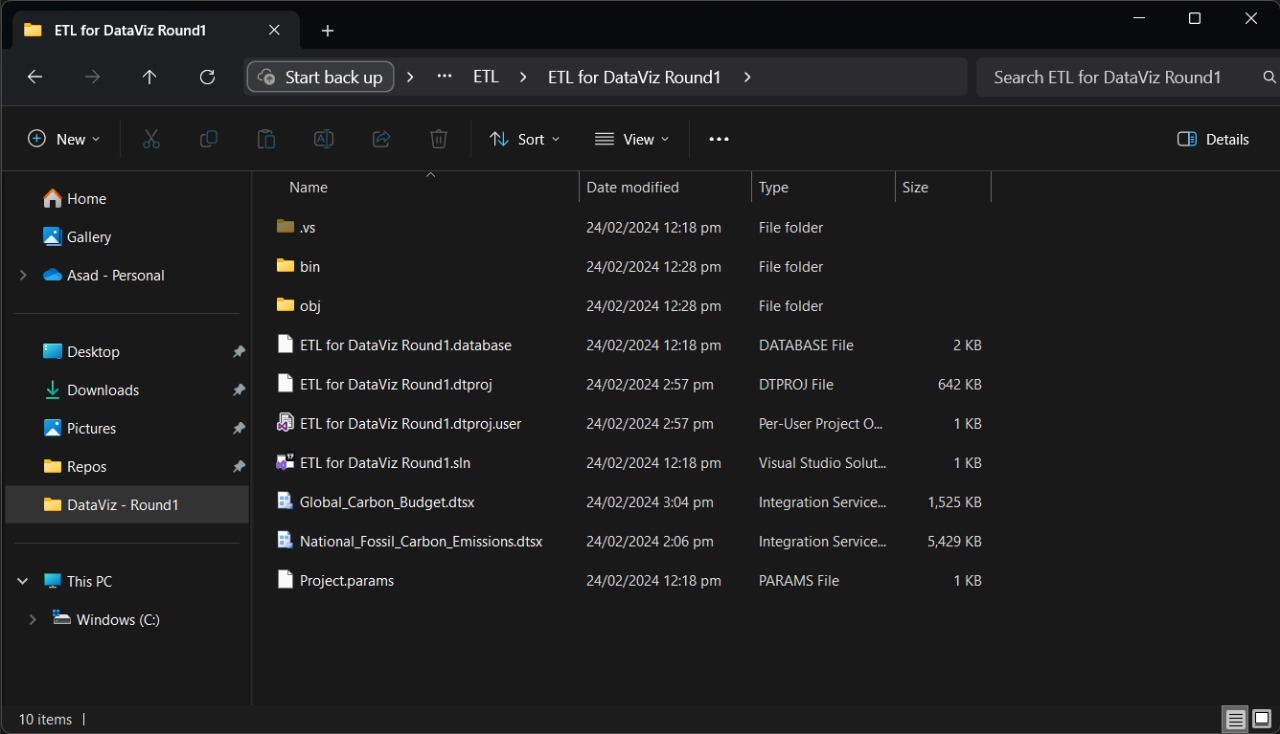
ETL Pipeline

**Tools:** Microsoft SQL Integration Servies Extension, Visual Studio

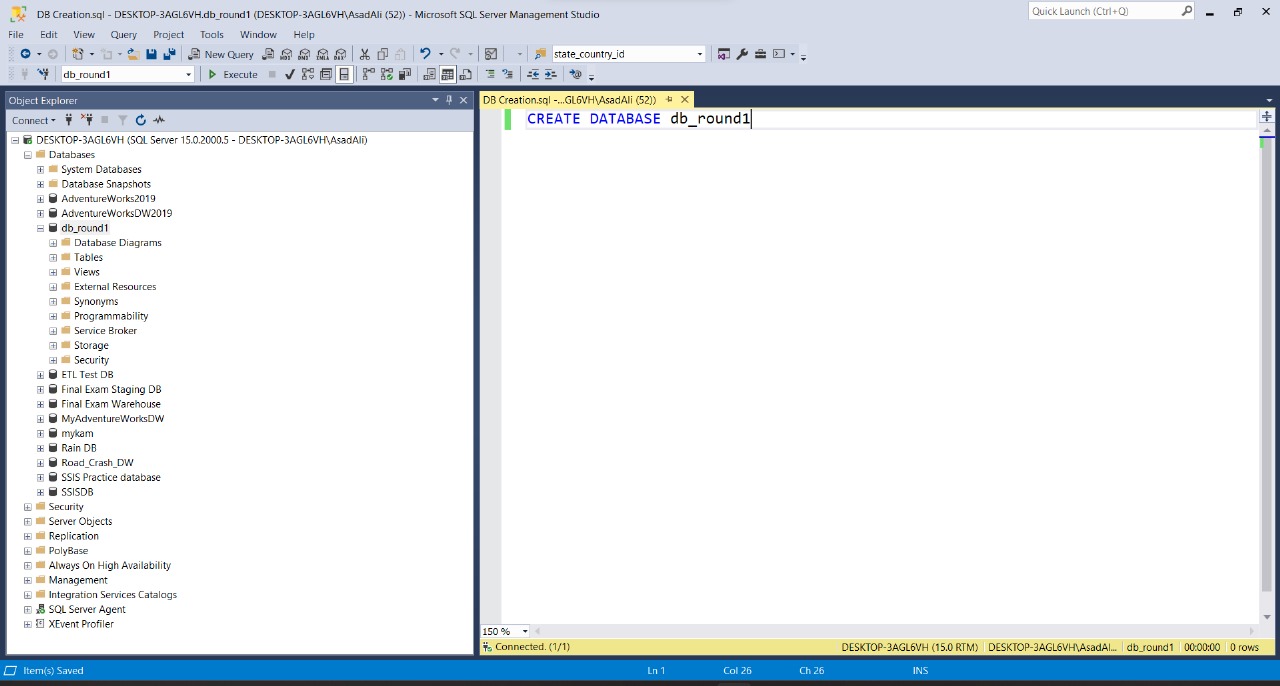
This is our working project folder:



This is our ETL project in Visual Studio using SSIS extension provided by Microsoft to design ETL pipelines.



On the other hand, create a DB in MS SQL server using the query in DB Creation.sql file:



Created ETL 2 packages for 2 input excel files:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

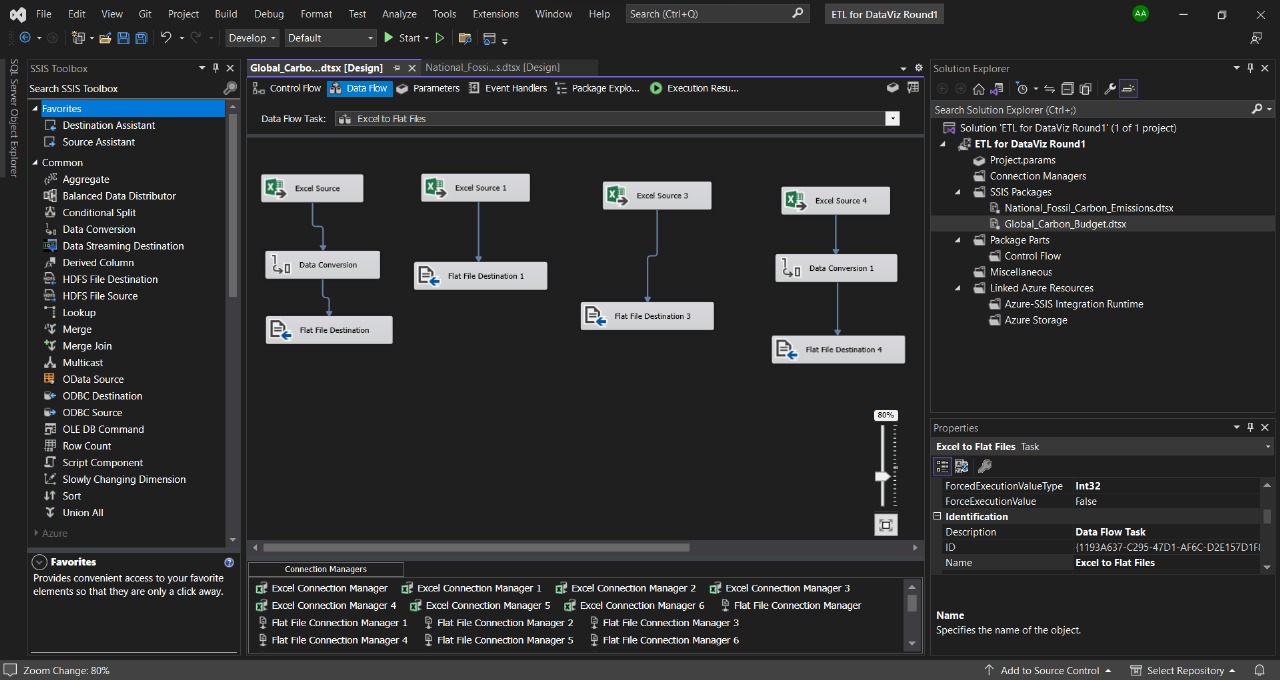
Description automatically generated

First of all use execute sql component to execute sql queries that first truncate our tables if the already exists other wise new will be created automatically:

A screenshot of a computer

Description automatically generated

Now load excel data to flat files to conveniently load in DB:

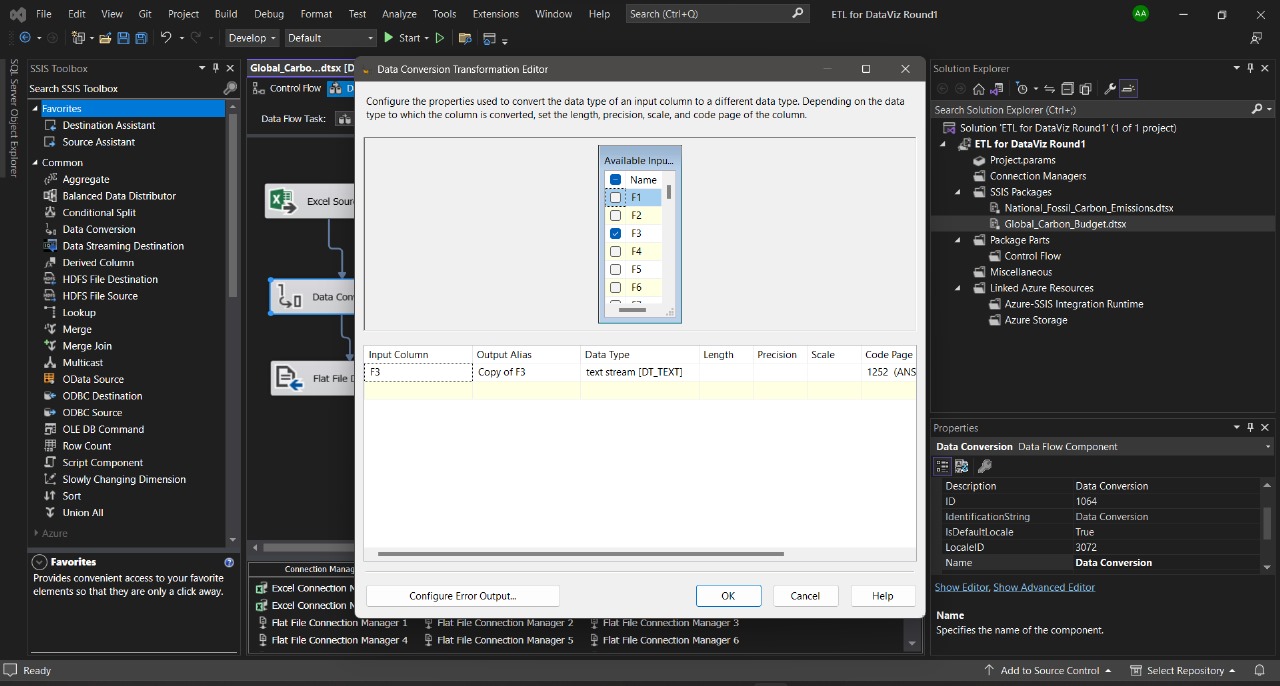


Connecting to Excel file and selecting a sheet and this step will be performed on all sheets:

A screenshot of a computer

Description automatically generated

While converting excel file to flat files there was an error of type mis match. We used data conversion component to convert data to NTEXT to TEXT format:



Save data from excel to Flat .txt files:

A screenshot of a computer

Description automatically generated

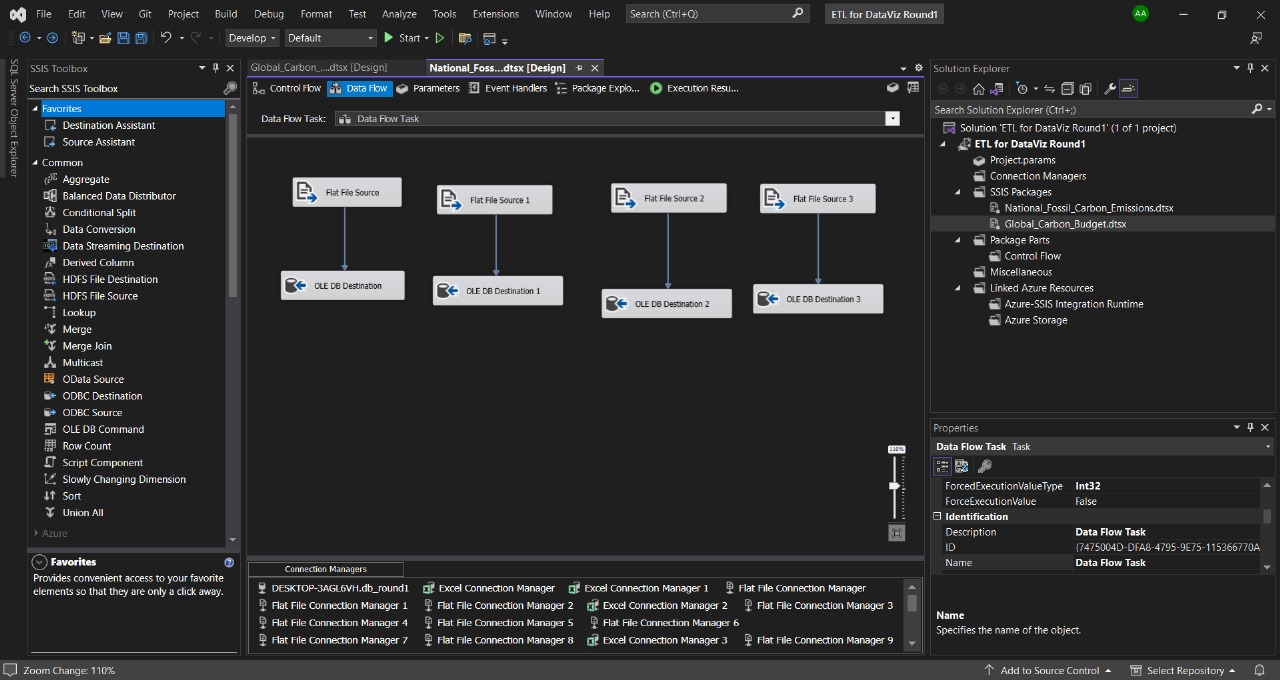
Now there will be text files for each excel sheet given in excel file:

A screenshot of a computer

Description automatically generated

This intermediate format conversion was used to skip extra text at the top of each excel sheet.

Now use these flat files to load data into DB:



Check column names before loading data to DB:

A screenshot of a computer

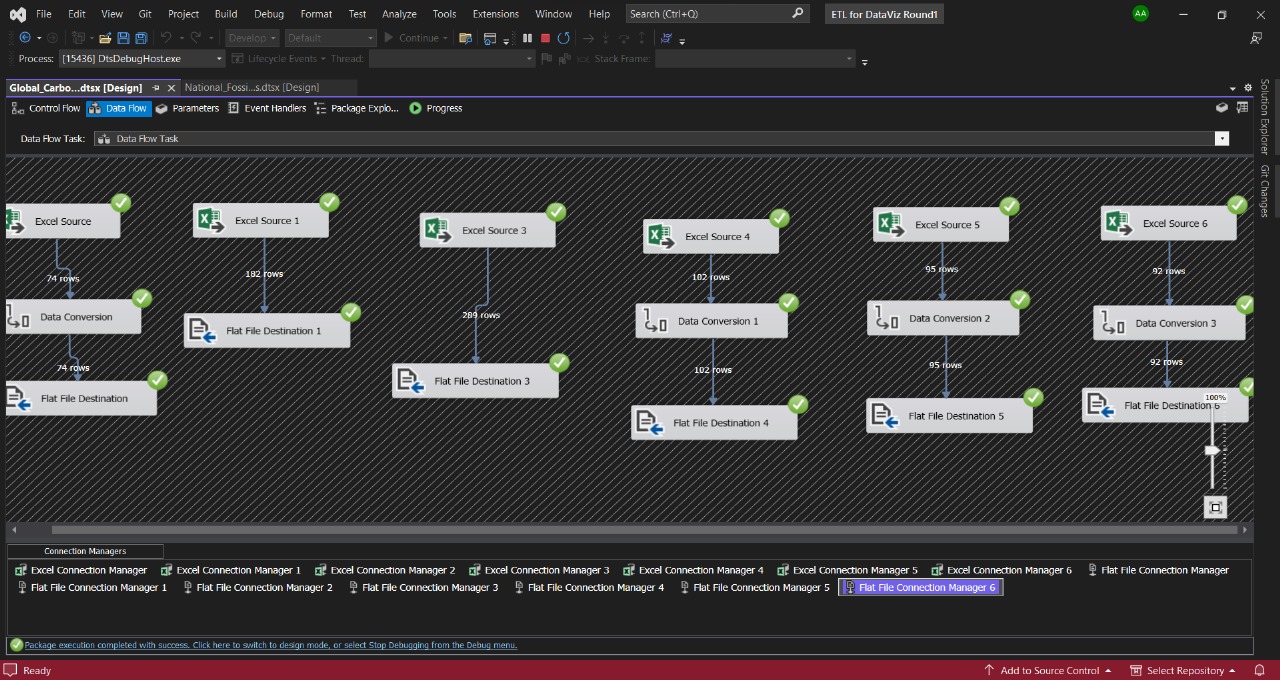
Description automatically generated

Create tables from SSIS project using SQL query that will create a table:

A screenshot of a computer

Description automatically generated

After successful execution of all SSIS packages:



After that all the data from each excel sheet there will be a table in DB:

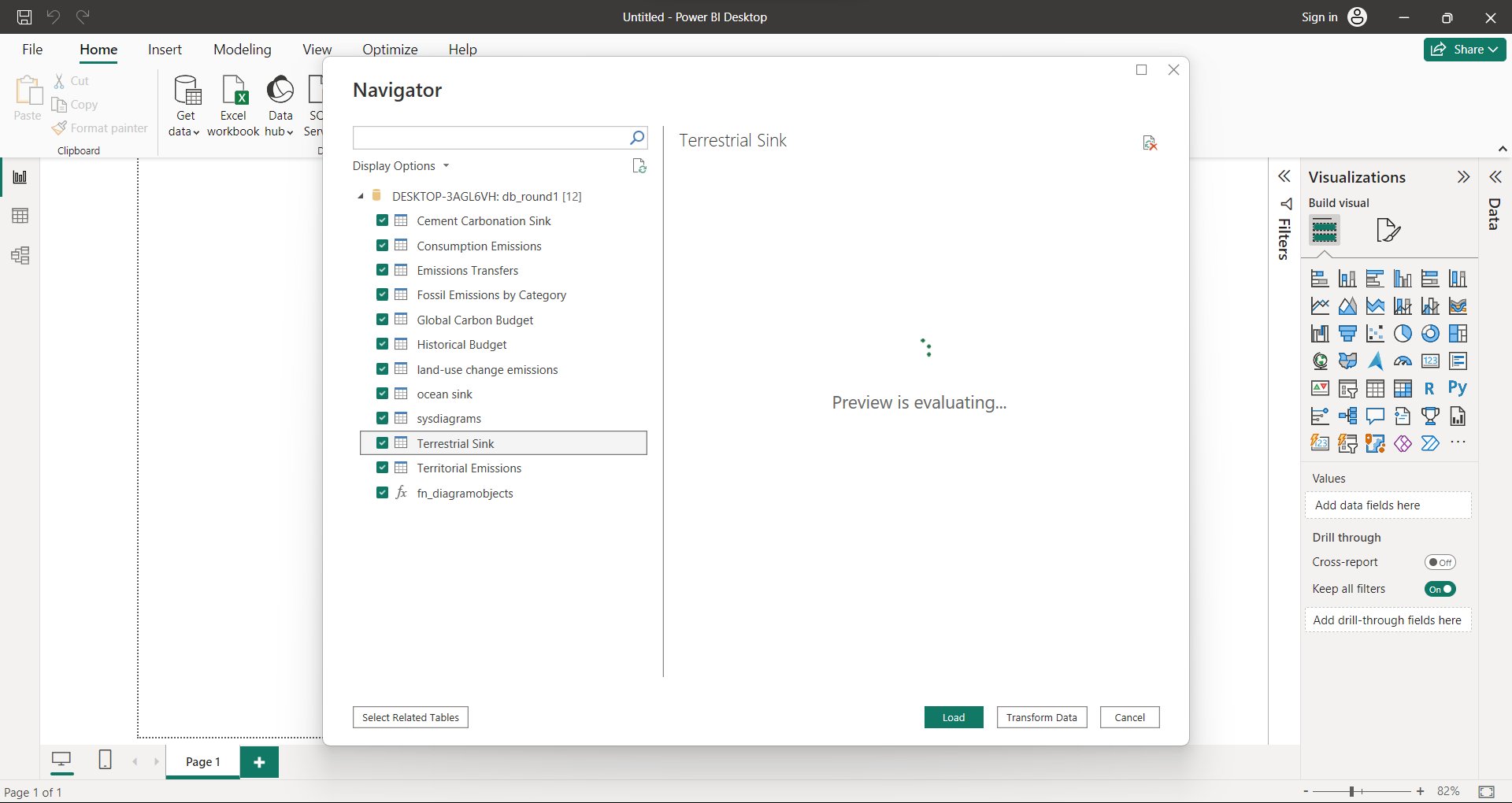
A screenshot of a computer

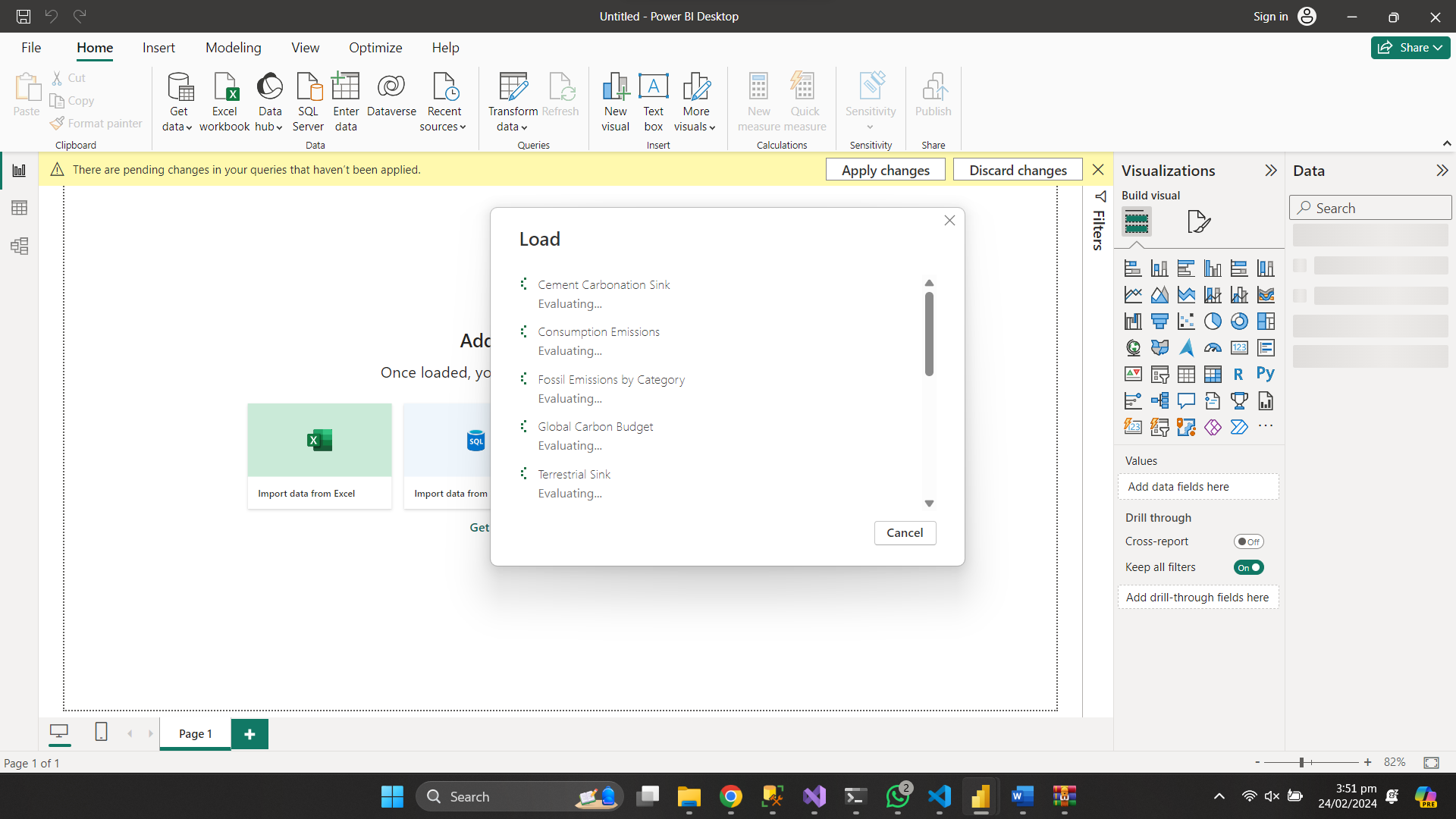
Description automatically generated

This is how all the data will be loaded in DB.

**Feature Engineering**

Now connect the DB in PowerBI:





Transform this table:

