LAB- 3(FDE)

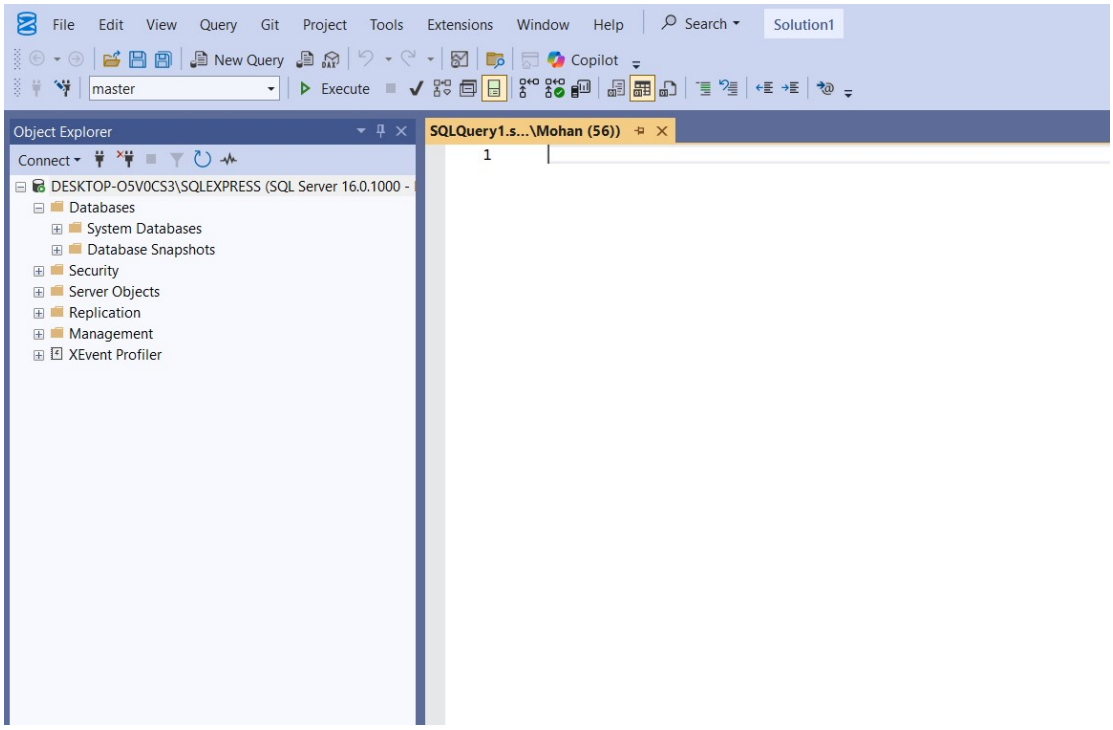
1. Install the following software :

SQL Server Express + SQL Server Management Studio (SSMS)

- Links: SQL Server Downloads
- Install SQL Server Express with default settings.
- Install SSMS to manage the database.

AdventureWorks Sample Database

- Use the link: AdventureWorks Installation
- Make sure it's installed in your SQL Server.



PostgreSQL
Download & install: PostgreSQL Downloads

- Default password: 12345 (you can use demopass for your ETL role as in your code).

- Also install pgAdmin (optional but helps to visualize database).

Python environment

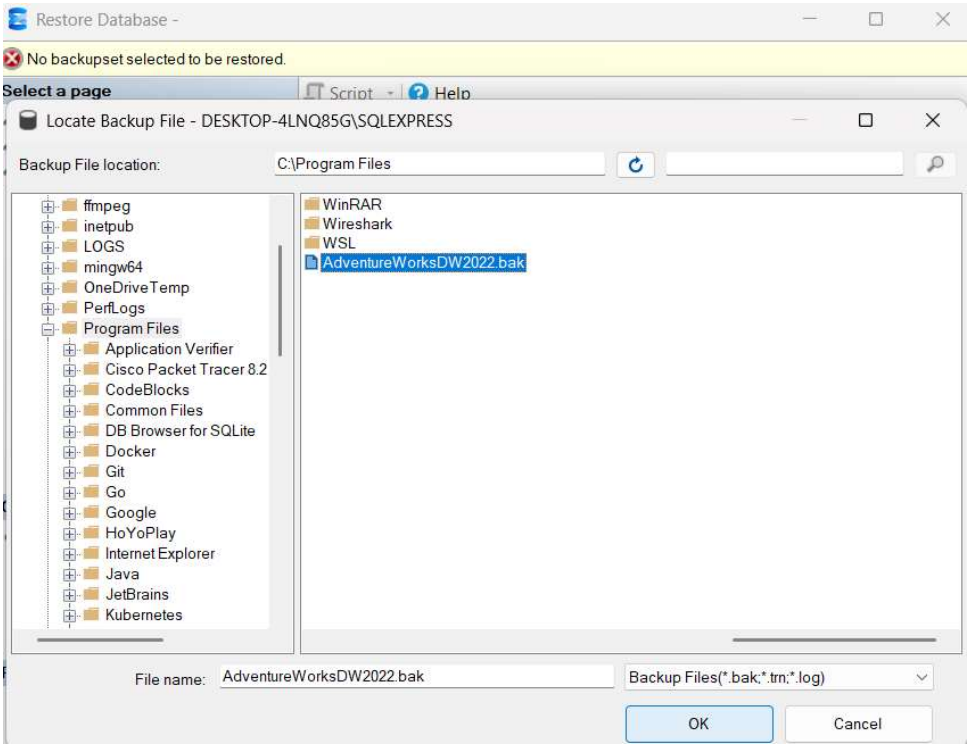
- Make sure you have Python 3.9+, pandas, sqlalchemy, and psycopg2 installed.
- Install via pip

ODBC Driver for SQL Server

- Download ODBC Driver 17

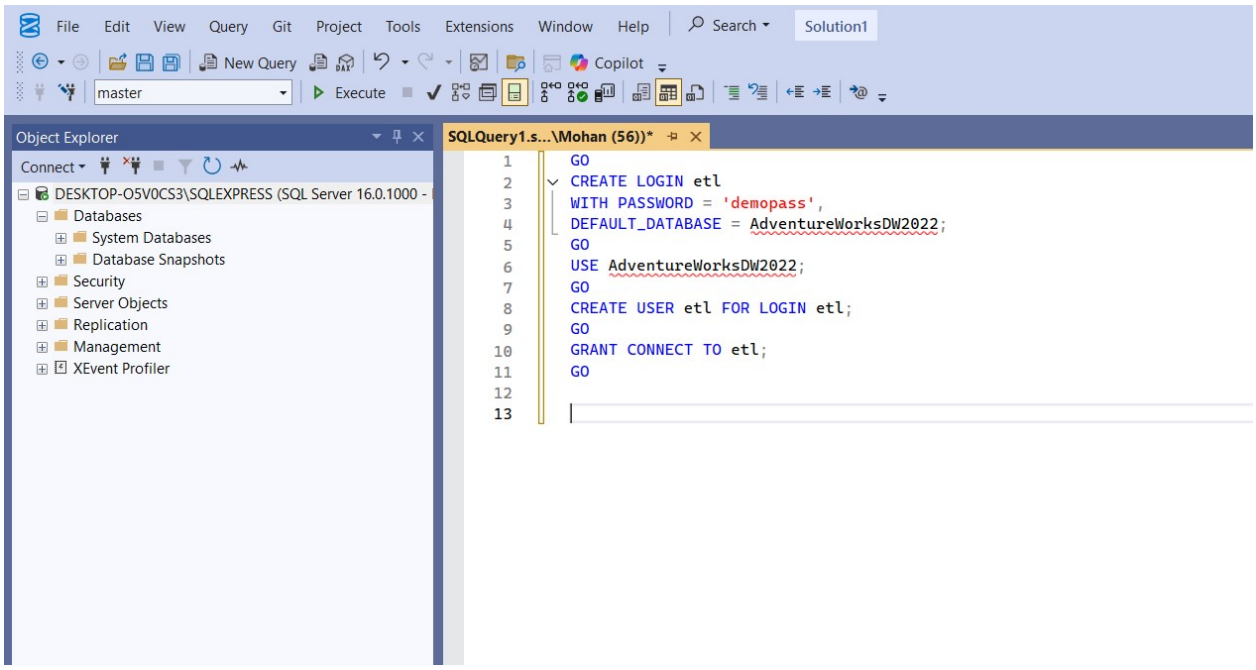
2. Open SSMS(Through Visual Studio):

1. Connect to the sql server that you downloaded(usually ends with \SQLEXPRESS)
2. After connecting, Databases->restore databases->device, click on the three dots and click on add, then select your adventureworks.bak file.



3. Now open ‘New Query’ and run the following code:
   USE master;
   GO
   CREATE LOGIN etl

```
WITH PASSWORD = 'demopass',
DEFAULT_DATABASE = AdventureWorksDW2022;
GO
USE AdventureWorksDW2022;
GO
CREATE USER etl FOR LOGIN etl;
GO
GRANT CONNECT TO etl;
GO
```



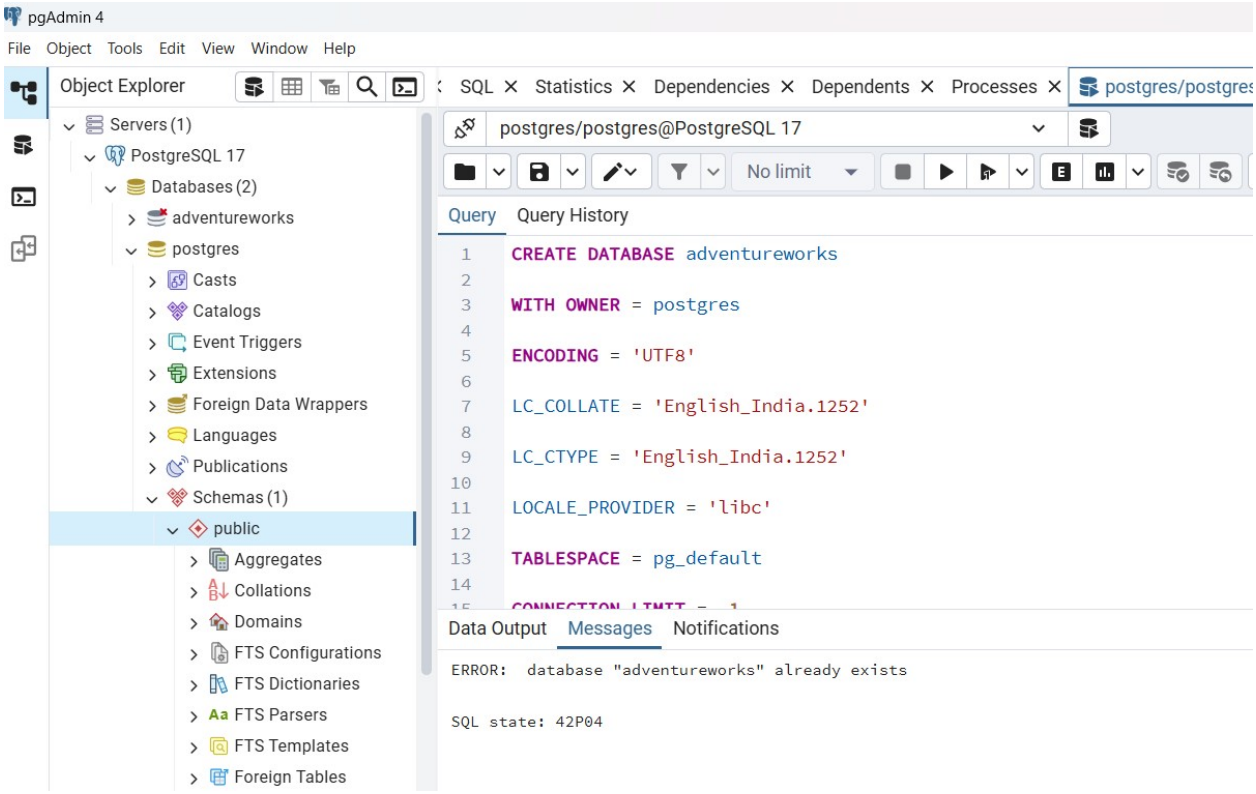
3.Open PGADMIN4:

1. Click on ‘Databases’->Create-> Database-> name it ‘adventureworks’->save
2. Open Query Tool and run the following code :

```
-- Creating the Role and Granting the privileges
CREATE ROLE etl WITH
LOGIN
PASSWORD 'demopass';
GRANT ALL PRIVILEGES ON DATABASE adventureworks TO etl;

-- Run this while connected to adventureworks
GRANT USAGE, CREATE ON SCHEMA public TO etl;
GRANT ALL PRIVILEGES ON ALL TABLES IN SCHEMA public TO etl;
GRANT ALL PRIVILEGES ON ALL SEQUENCES IN SCHEMA public TO
etl;
```

ALTER DEFAULT PRIVILEGES IN SCHEMA public
GRANT ALL PRIVILEGES ON TABLES TO etl;
ALTER DEFAULT PRIVILEGES IN SCHEMA public
GRANT ALL PRIVILEGES ON SEQUENCES TO etl;



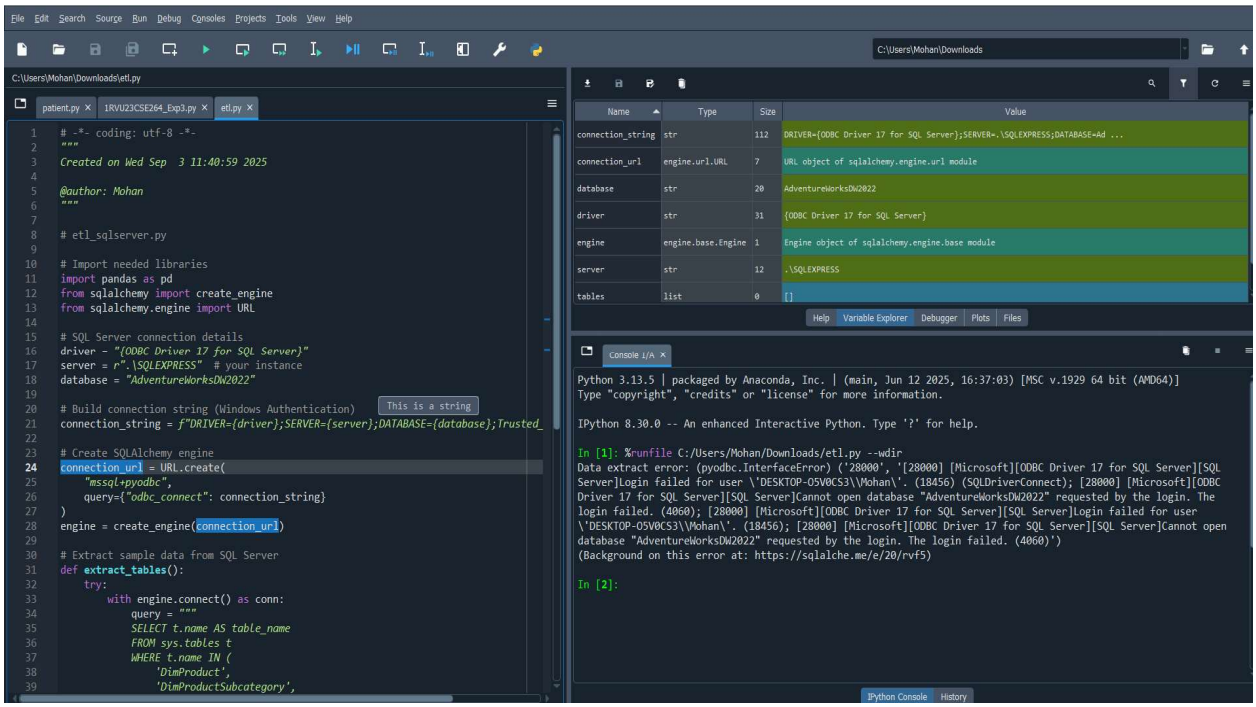
4.Open cmd:

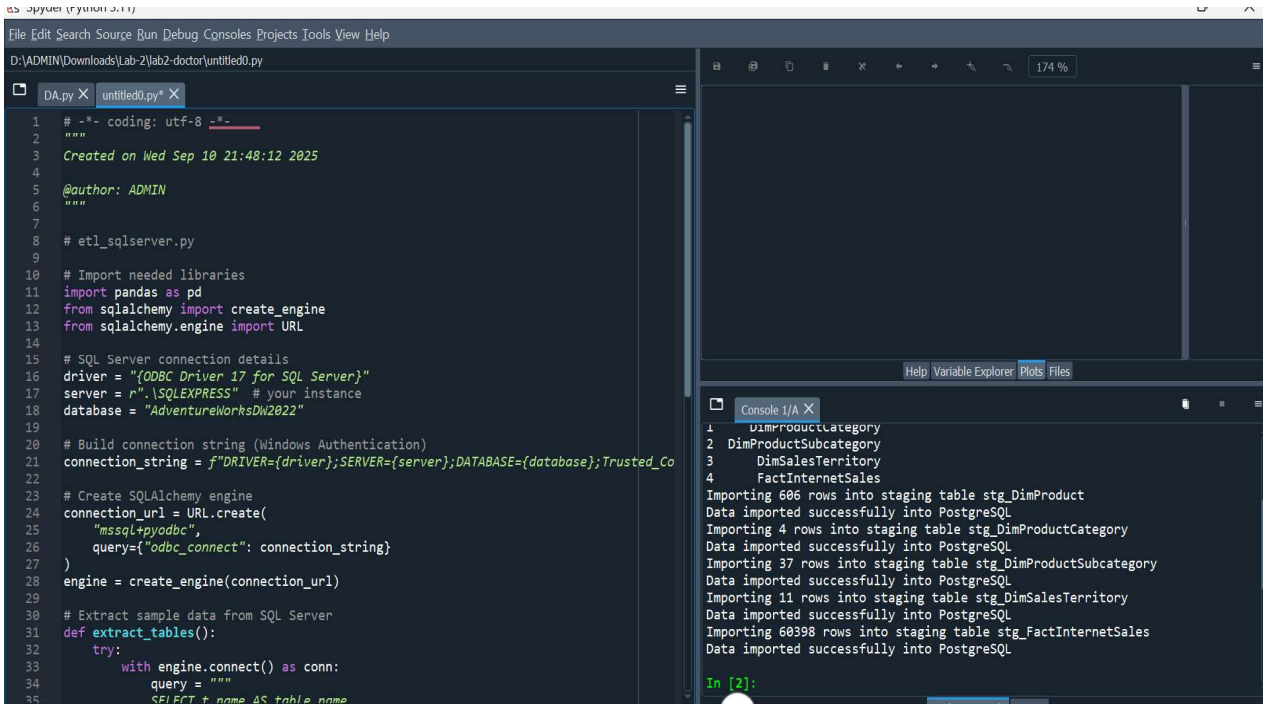
1. Run the following code to install dependencies:
   pip install psycopg2-binary pandas sqlalchemy pyodbc

5.Running Python Script:

In the python script, on line 10, **'server = r".\SQLEXPRESS" # your instance'** , you need to replace it with your SQL server name. For eg: if you sql server name is DESKTOP-4LNQ85G\SQLEXPRESS then your line of code will be: **server = r"DESKTOP-4LNQ85G\SQLEXPRESS". THIS IS THE SAME SERVER THAT YOU CONNECTED TO SSMS.**

Now run the script and you should see the output in the terminal.

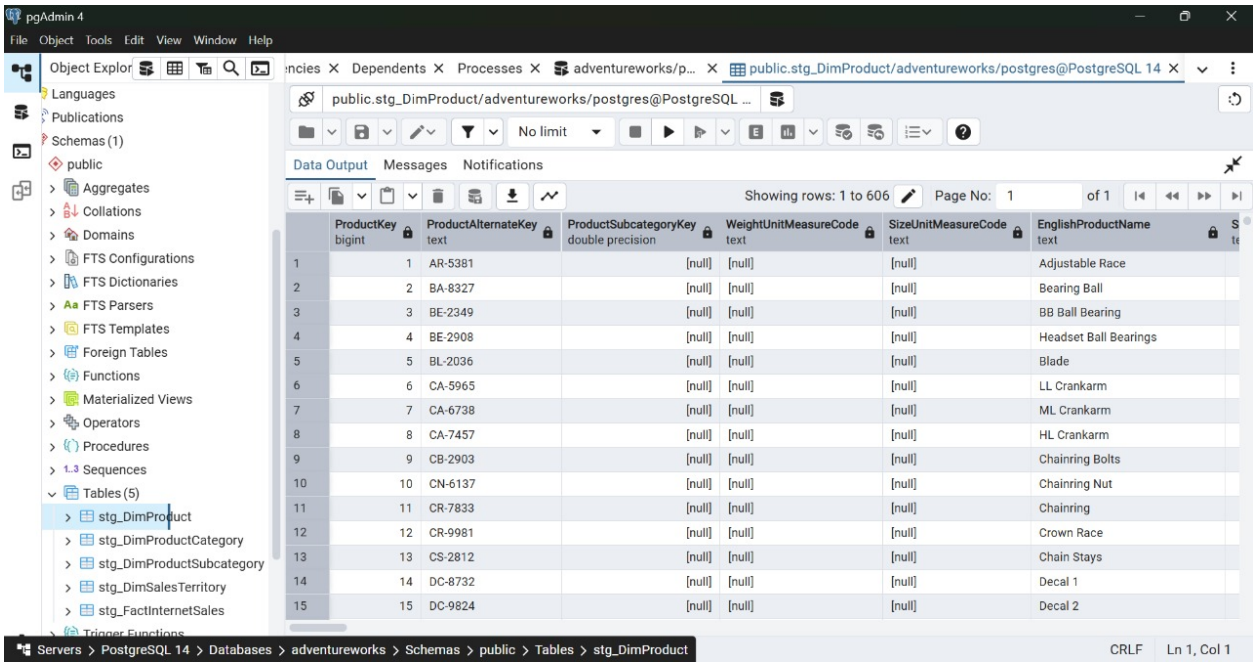


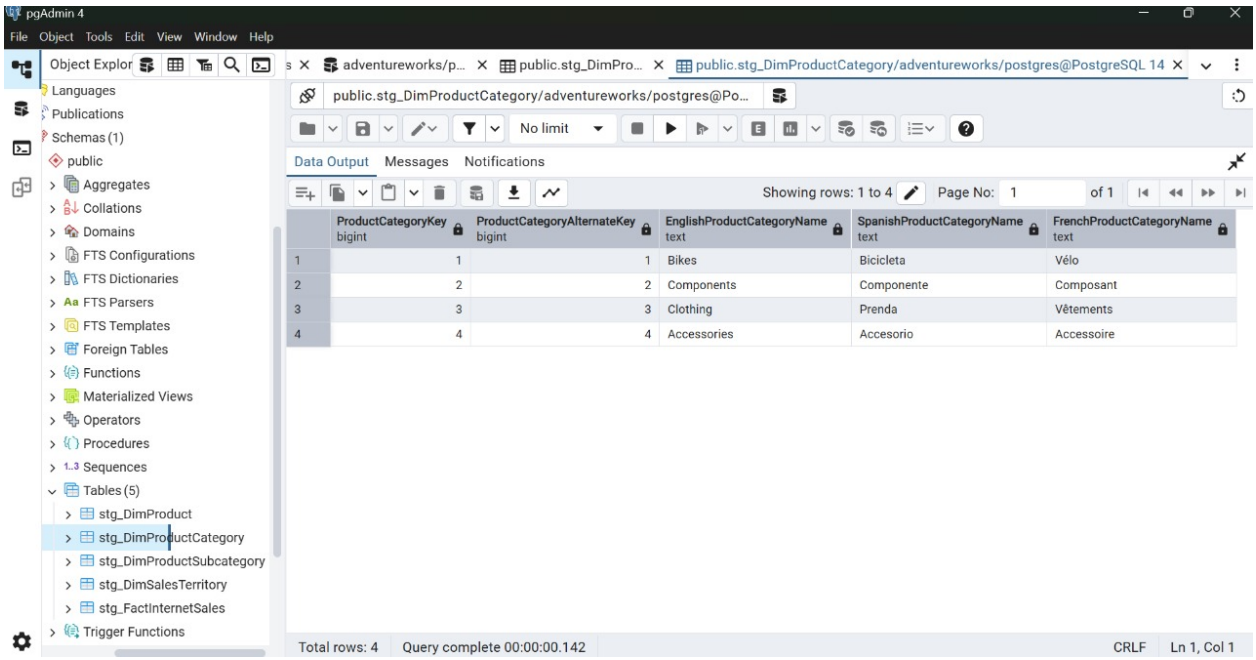


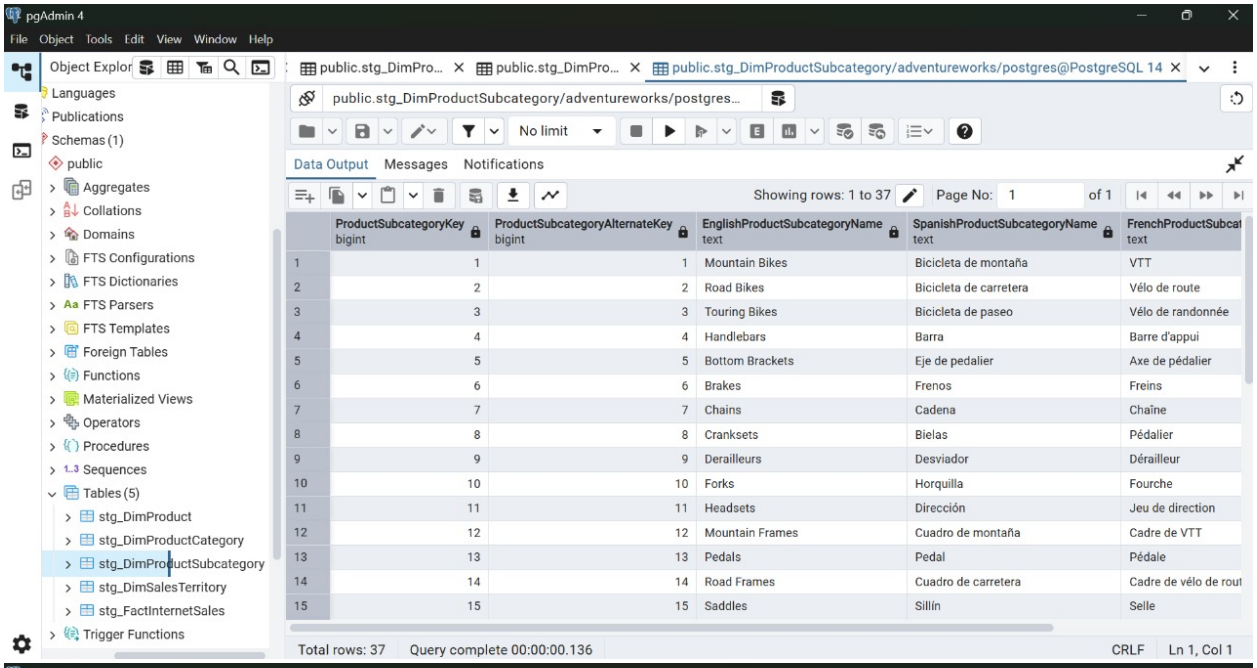
6. OUTPUT

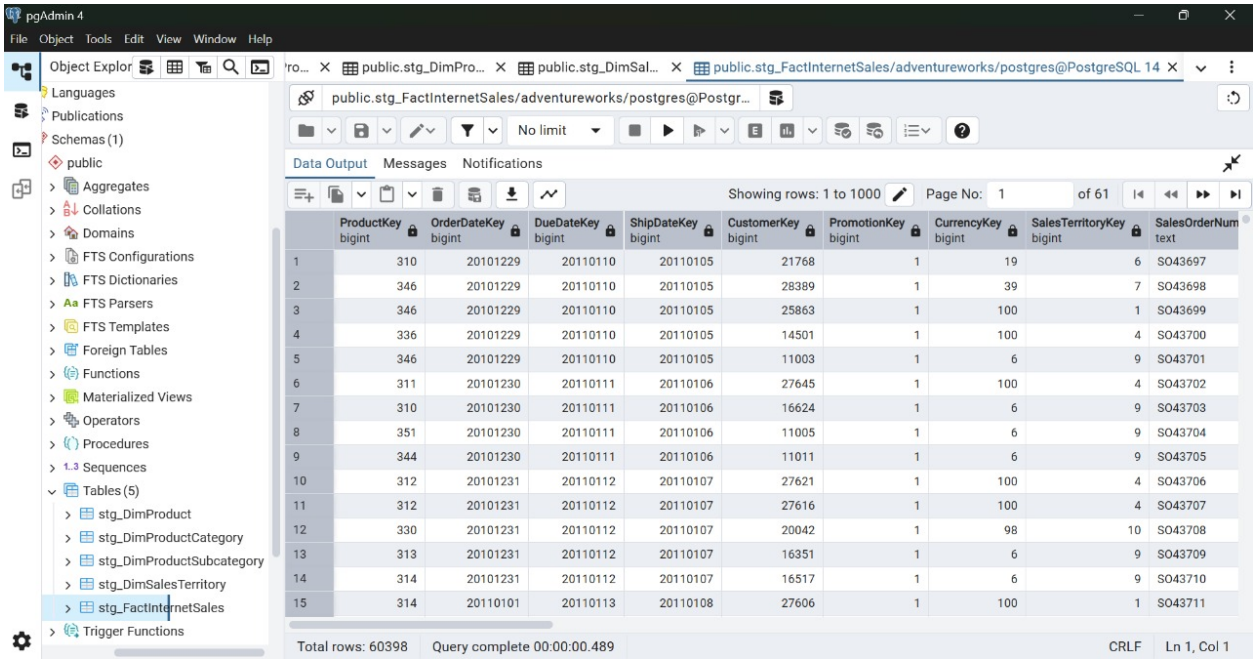
Open PgAdmin->Databases->adventureworks->schemas->public->tables-> right click on any one(stg_dimproduct
stg_dimproductcategory

stg_dimproductsubcategory
stg_dimsalesterritory
stg_factinternetsales) ->View and Edit-> Take an ss and submit









GITHUB LINK:
https://github.com/meghana1653/DataEngineering/blob/main/LAB3%20FDE.pdf