USN NUMBER: 1RVU23CSE264
NAME: Meghana G

Ex No: 3 **Date: 10/9/25**	ETL Pipeline Implementation: Exporting Data from SQL Server to PostgreSQL

Objective:

This lab experiment provides practical experience in building a basic ETL (Extract, Transform, Load) data pipeline using Python. It guides participants through the core stages of the pipeline from extracting raw data, transforming it into a usable format, and loading it into a target system. The activity also simulates the roles and responsibilities of key stakeholders, such as data engineers, data scientists, and business analysts, to demonstrate their collaborative contributions in the data pipeline.

Outcomes:

1. Identify and describe the stages of the data engineering lifecycle.
2. Explain the roles and responsibilities of different stakeholders at each stage.
3. Perform basic data engineering tasks within a simulated environment.
4. Collaborate across simulated stakeholder roles to design and implement a data-driven solution.

Lab Procedure

STEP 1: Install the following software :

1. SQL Server Express + SQL Server Management Studio (SSMS)

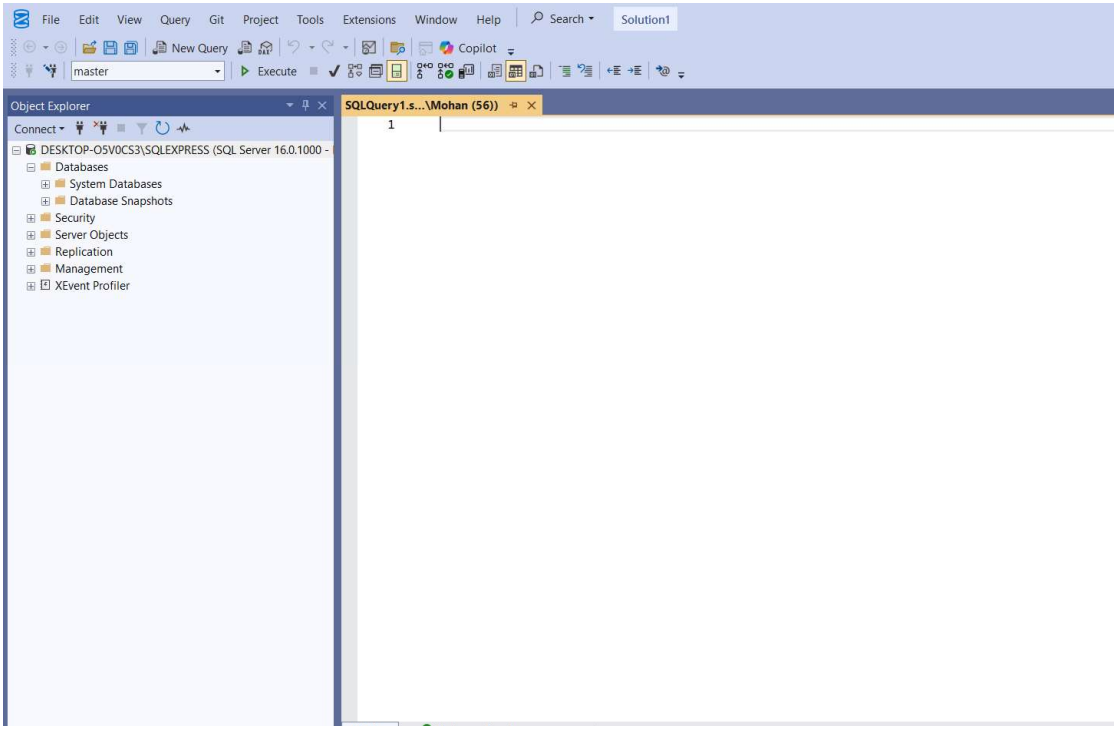
   - Links: SQL Server Downloads
   - Install SQL Server Express with default settings.
   - Install SSMS to manage the database.

2. AdventureWorks Sample Database

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

PostgreSQL

USN NUMBER: 1RVU23CSE264
NAME: Meghana G



3. 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).
  Also install pgAdmin (optional but helps to visualize database).

4. Python environment

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

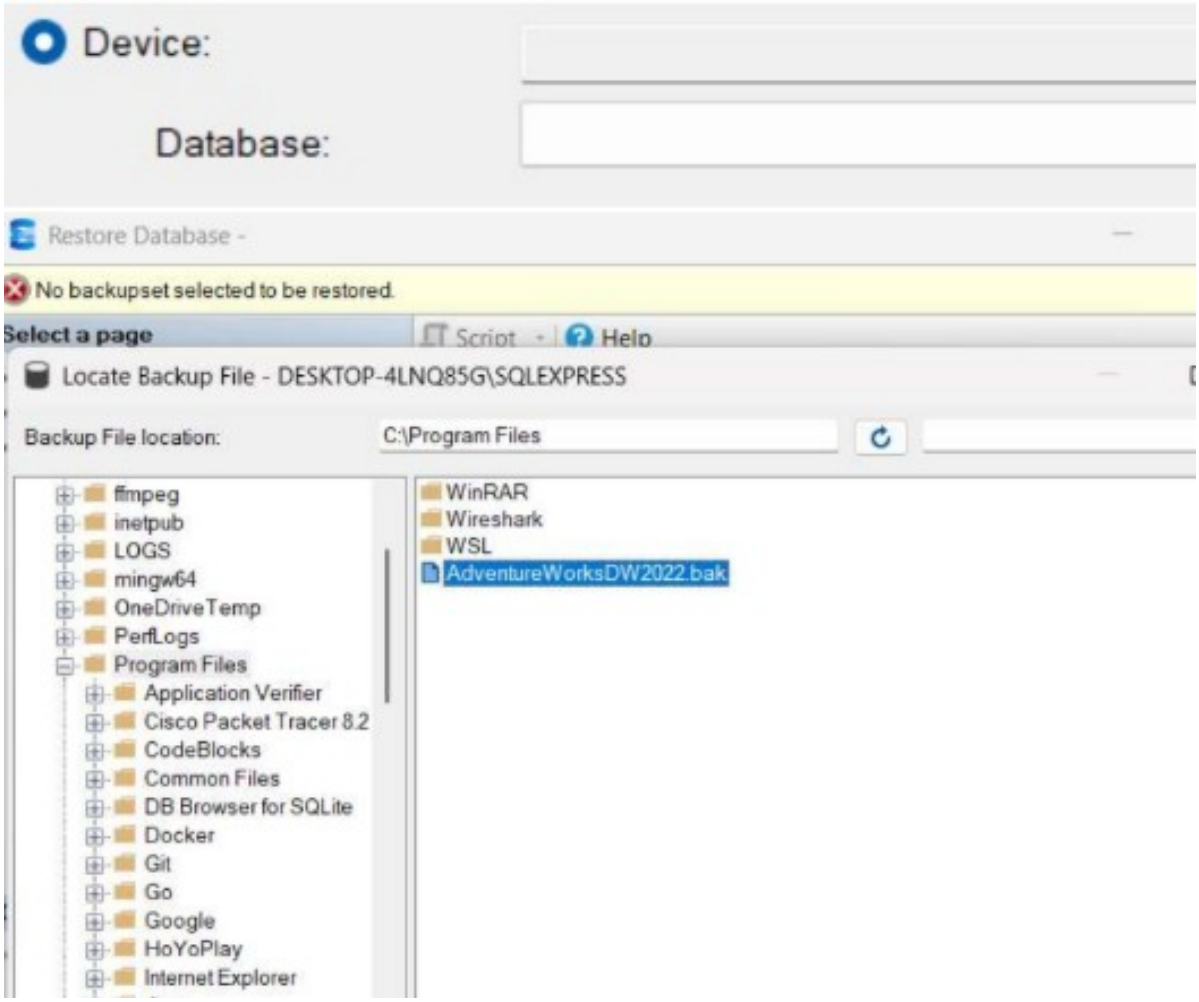
5. ODBC Driver for SQL Server

- Download ODBC Driver 17
  ODBC Driver 17

Step 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.

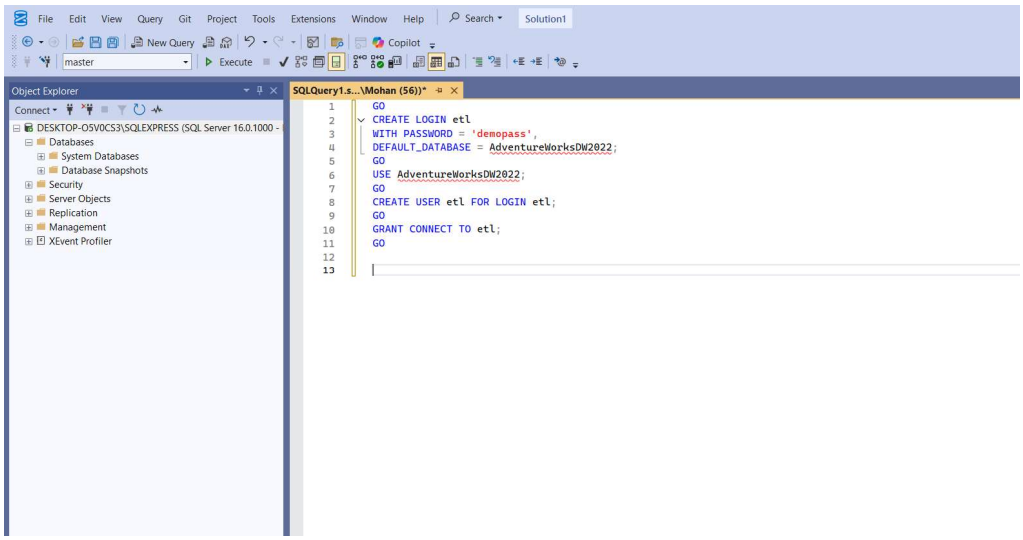
USN NUMBER: 1RVU23CSE264
NAME: Meghana G



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
```

USN NUMBER: 1RVU23CSE264
NAME: Meghana G



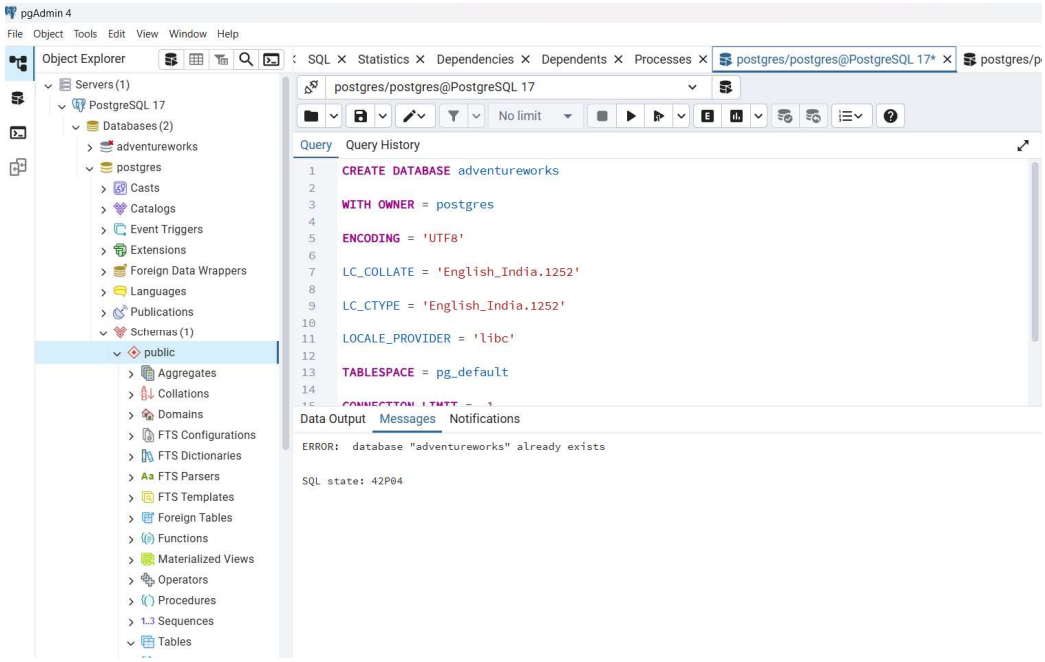
3.Open PGADMIN4:

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;
```

USN NUMBER: 1RVU23CSE264
NAME: Meghana G



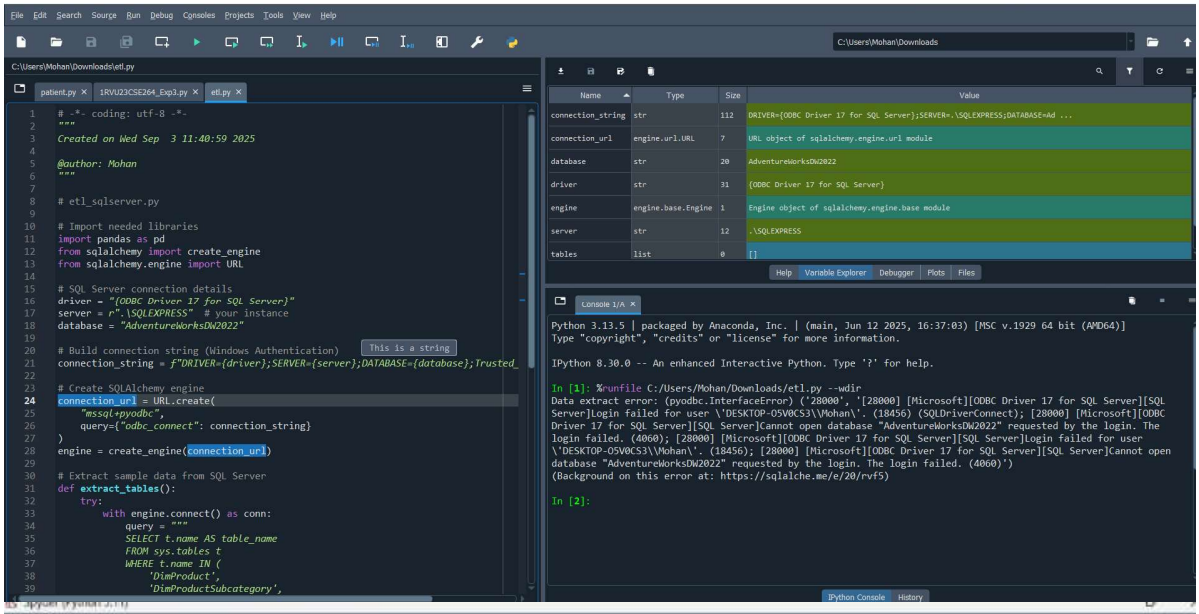
4.Open cmd:

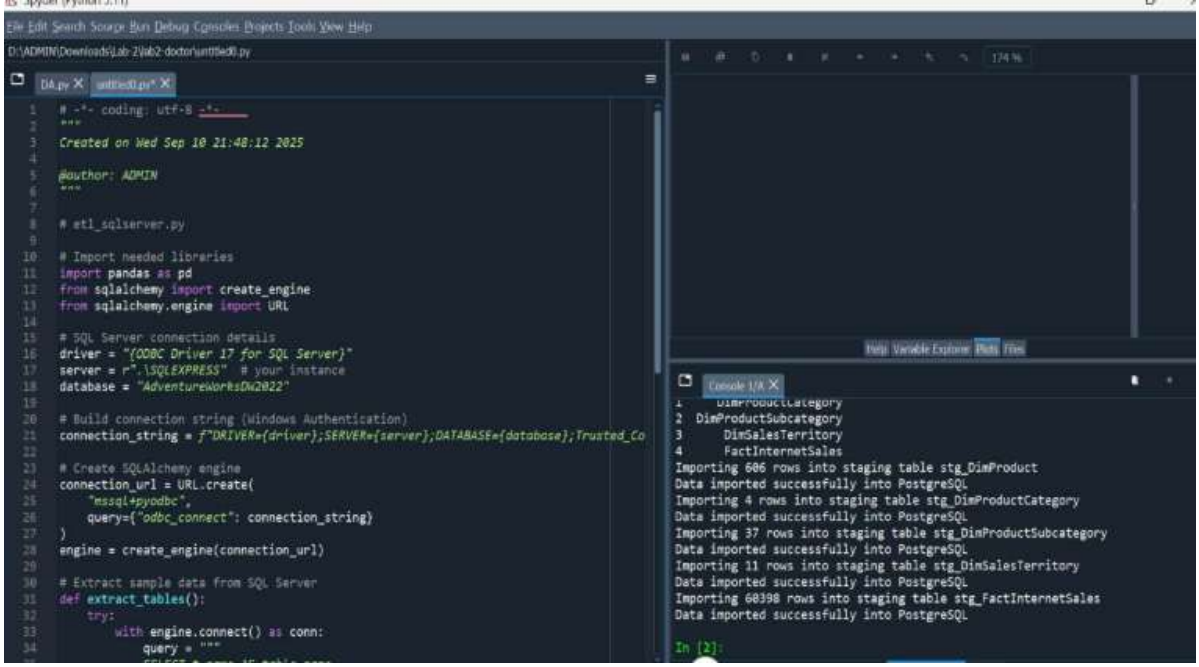
1. Run the following code to install dependencies:
   pip install psycopg2-binary pandas sqlalchemy pyodbc

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.

USN NUMBER: 1RVU23CSE264
NAME: Meghana G

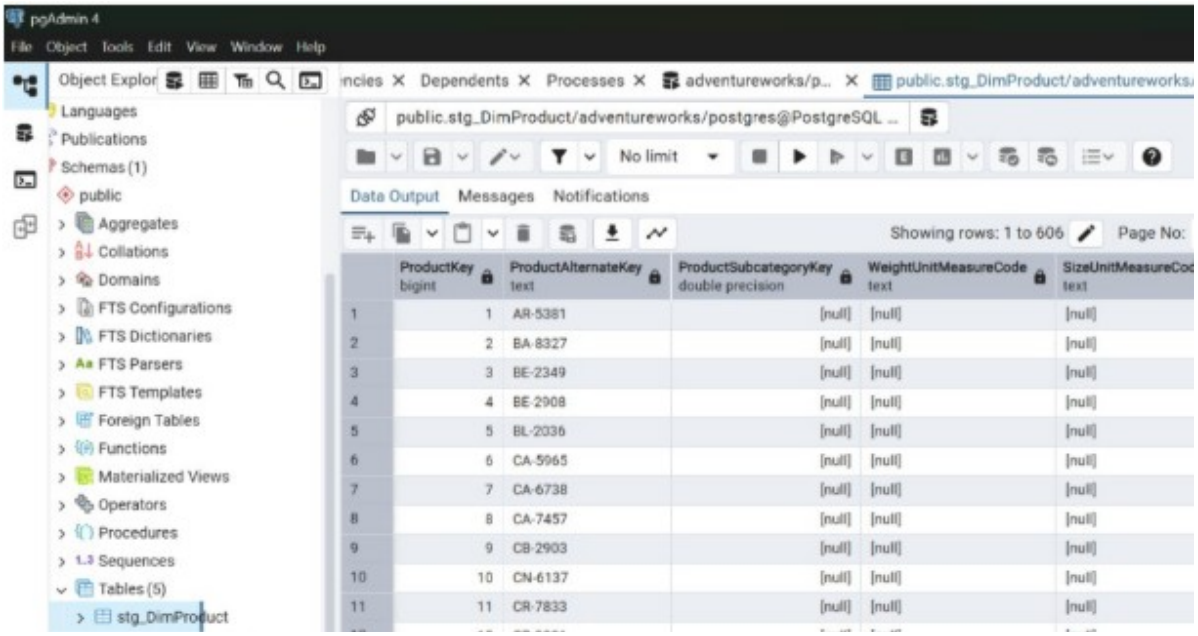


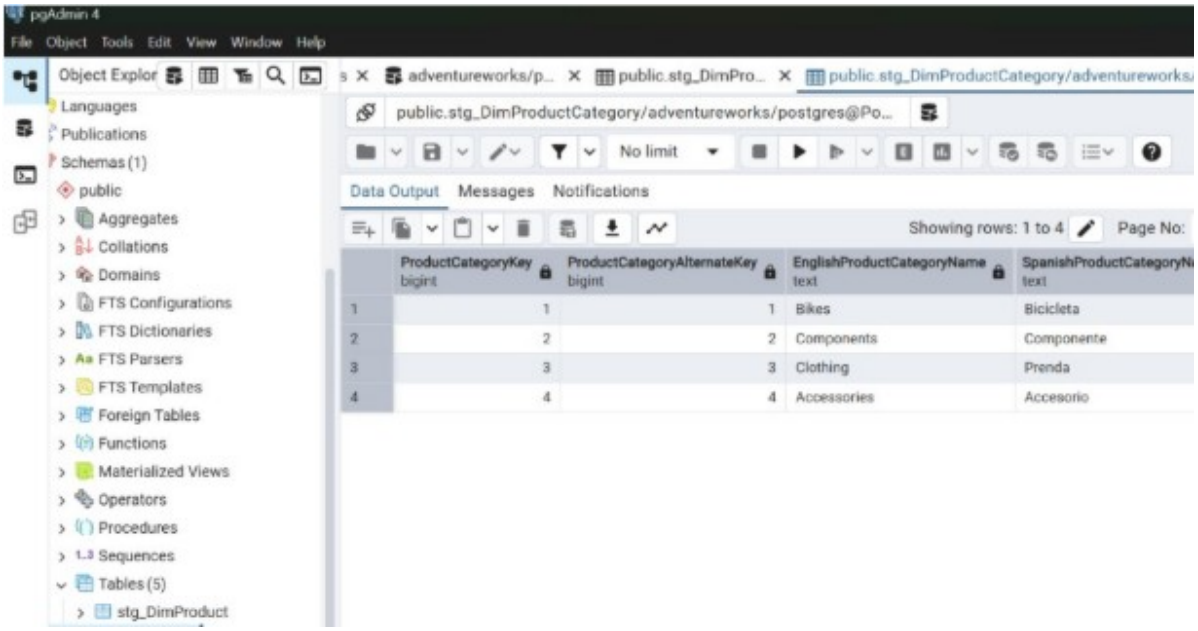


6. OUTPUT

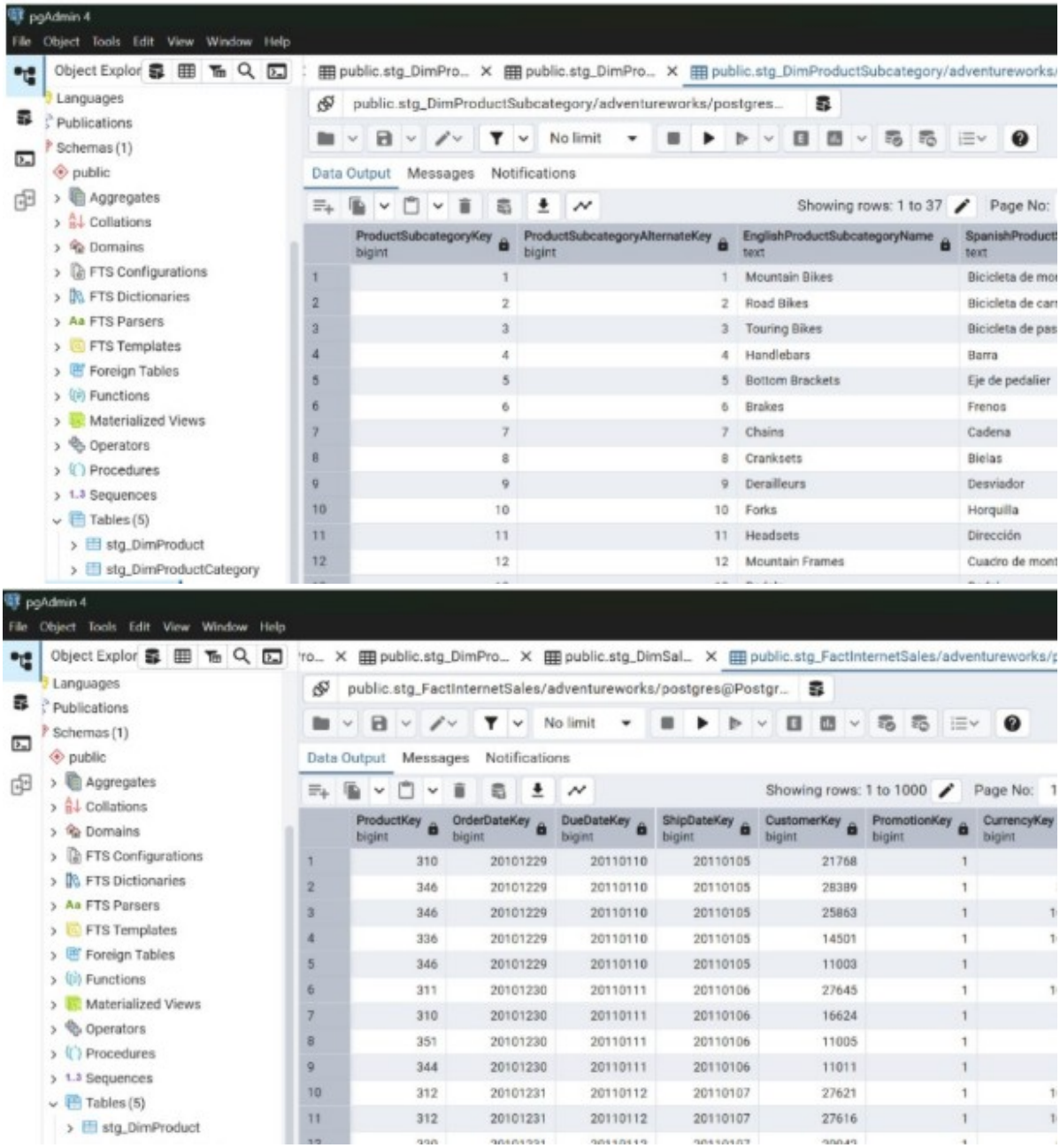
Open PgAdmin->Databases->adventureworks ->schemas->public->tables-> right click on any
one(stg_dimprodu
ct stg_dimproduct
category
stg_dimproductsubcategory
stg_dimsalesterritory
stg_factinternetsales) ->View and Edit-> Take an ss and submit

USN NUMBER: 1RVU23CSE264
NAME: Meghana G





USN NUMBER: 1RVU23CSE264
NAME: Meghana G



Github link:
https://github.com/meghana1653/Data-Engineering/blob/main/LAB-3.zip