

## LAB- 3(FDE)

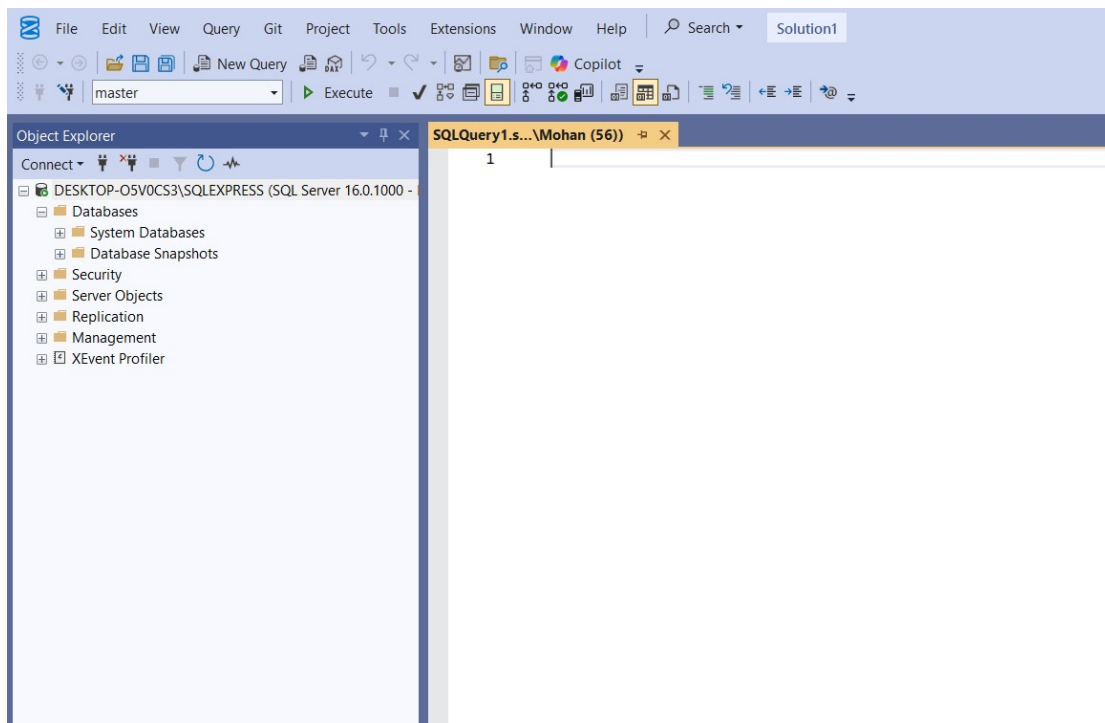
### 1. Install the following software :

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

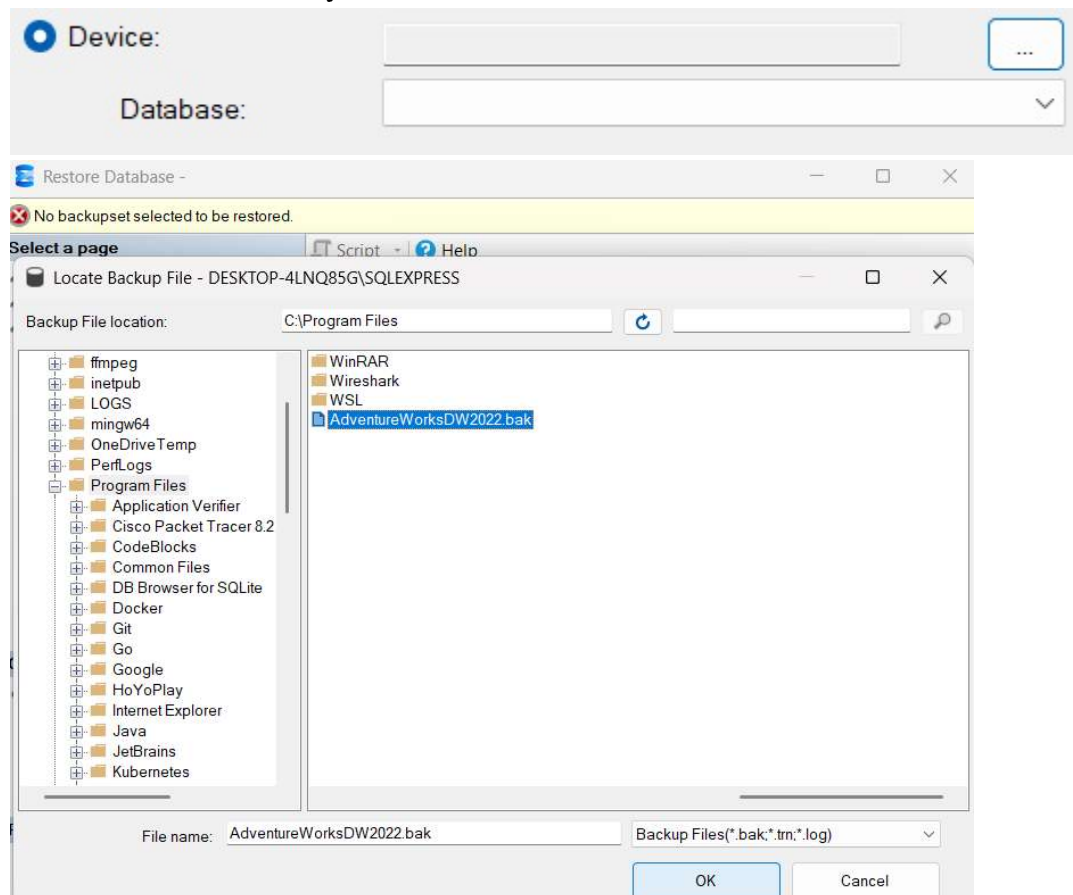
- Make sure you have Python 3.9+, [pandas](#), [sqlalchemy](#), and [psycopg2](#) installed.
- Install via pip

### 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.

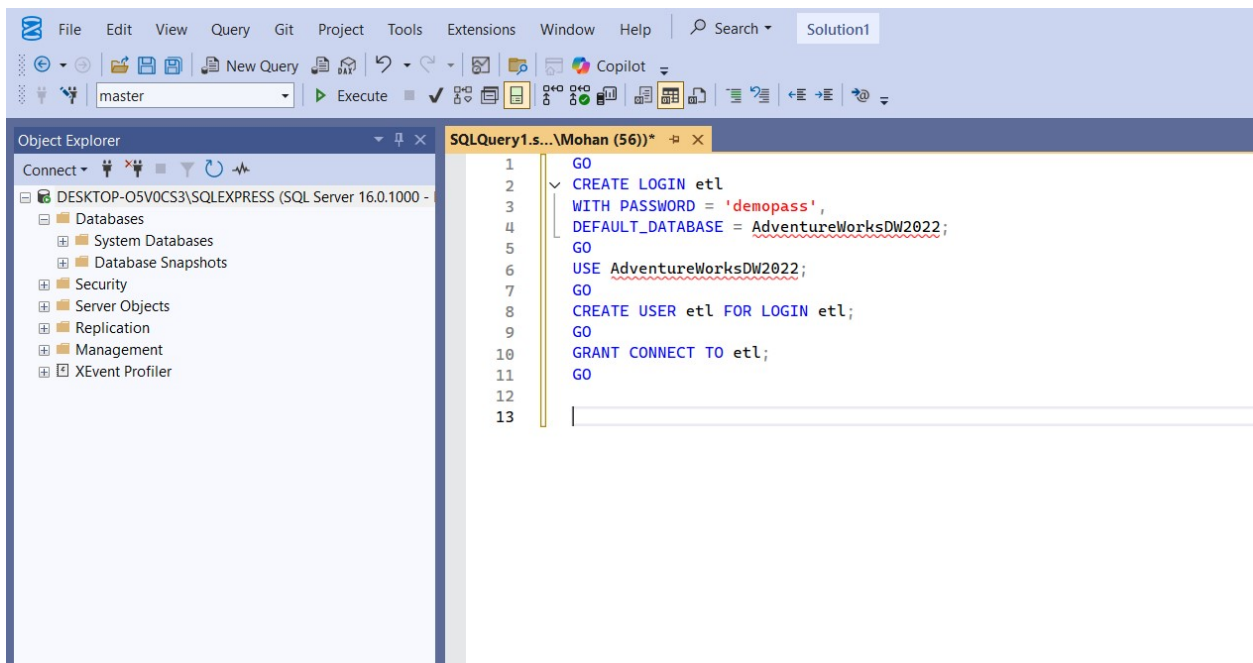


3. Now open 'New Query' and run the following code:  
USE master;  
GO  
CREATE LOGIN etl

```

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;
```

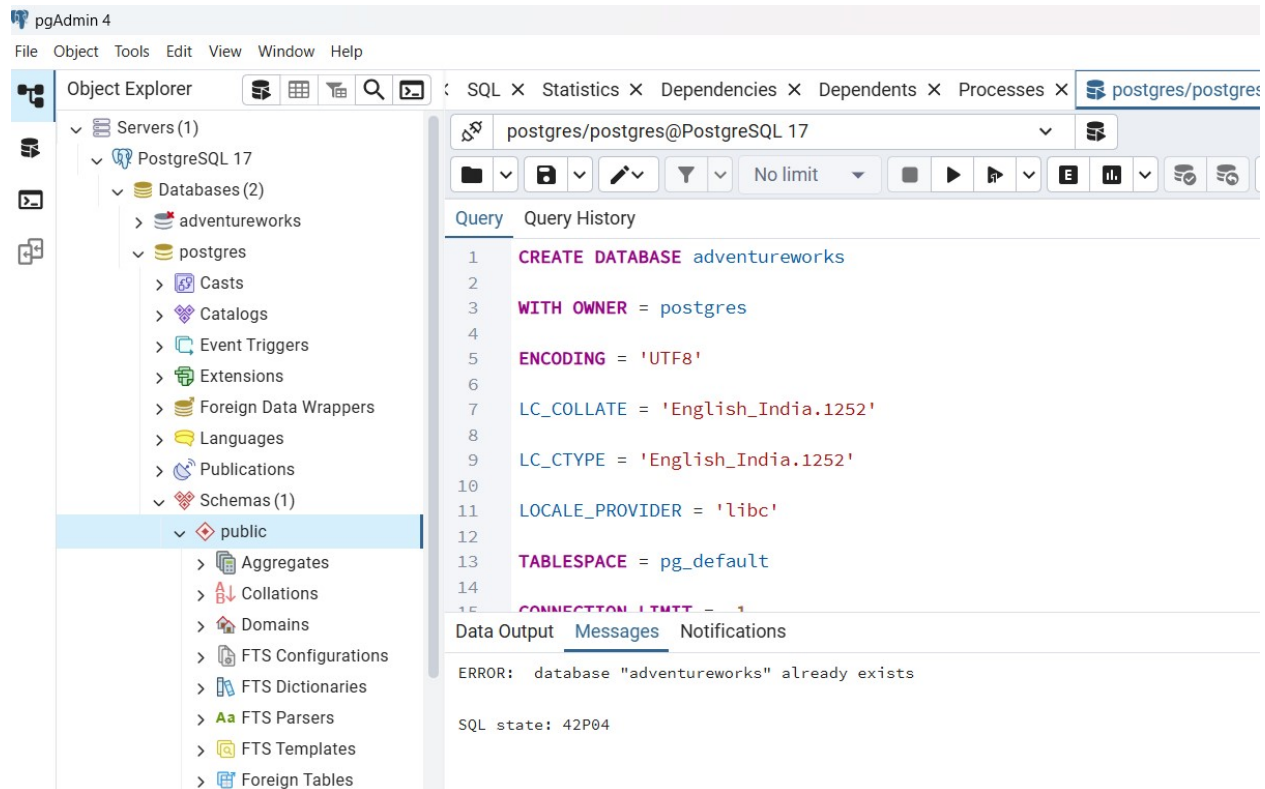
-- Run this while connected to adventureworks

```
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:

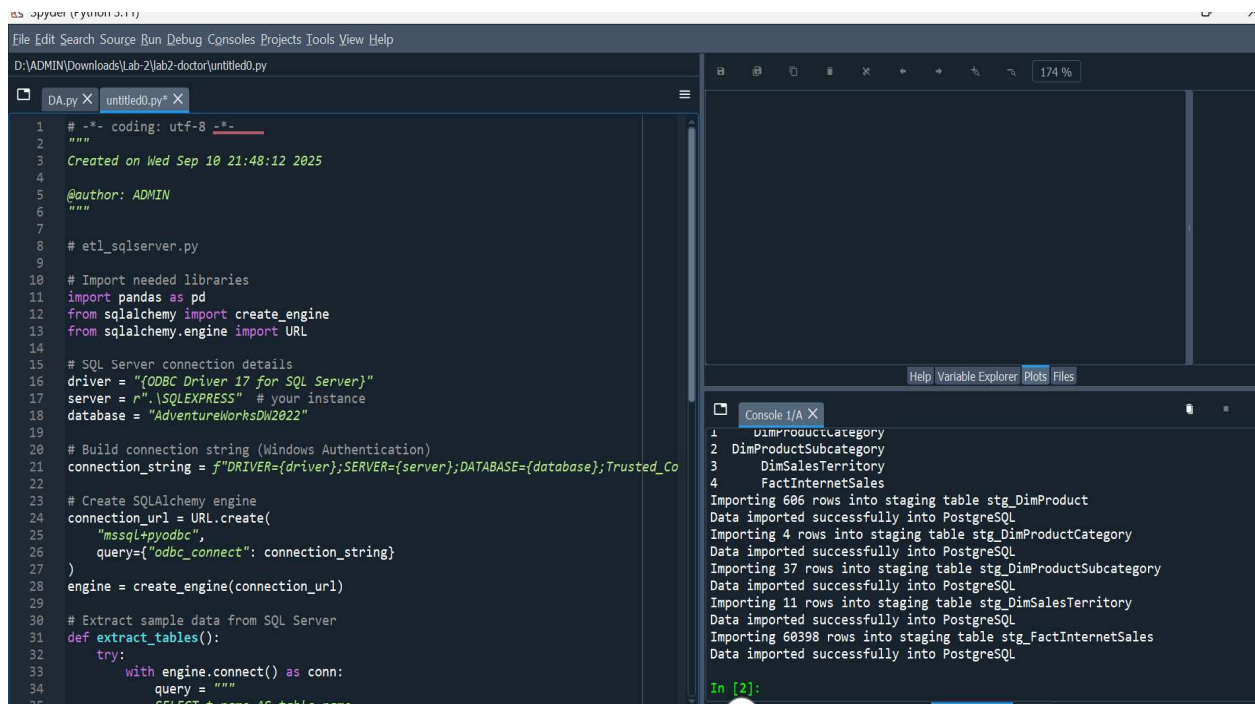
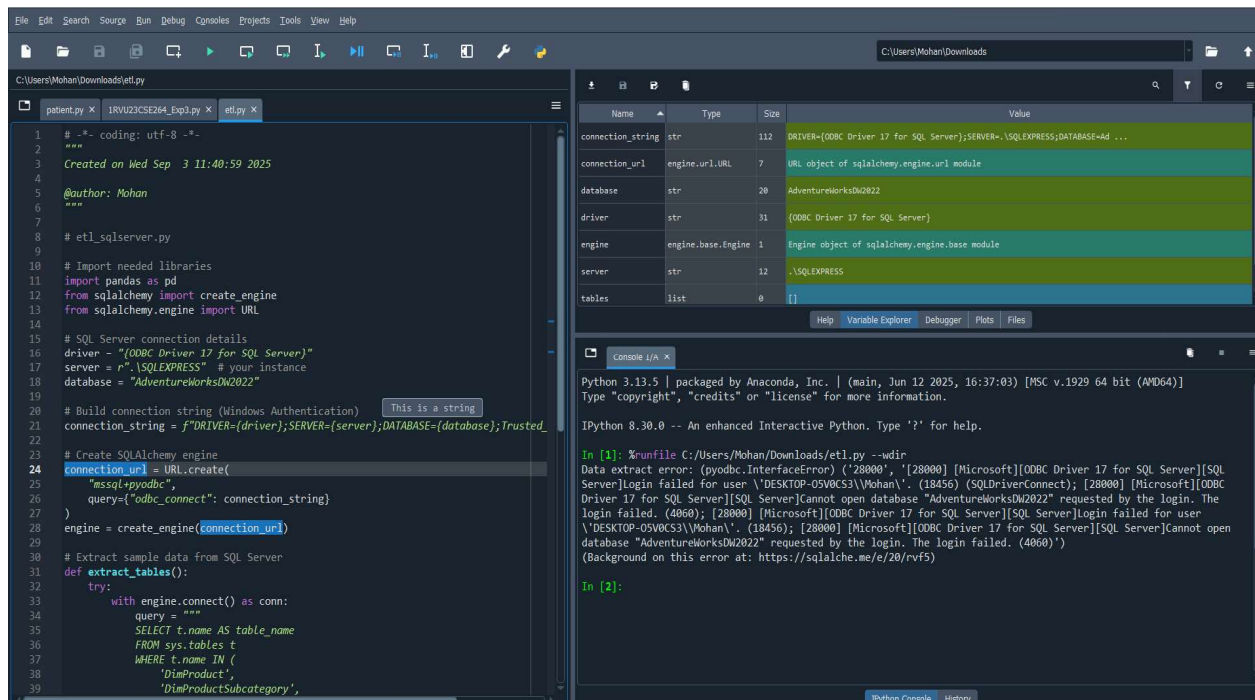
1. Run the following code to install dependencies:  

```
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

pgAdmin 4 interface showing the table `public.stg_DimProduct`. The table contains 15 rows of product data. The columns are: `ProductKey` (bigint), `ProductAlternateKey` (text), `ProductSubcategoryKey` (double precision), `WeightUnitMeasureCode` (text), `SizeUnitMeasureCode` (text), and `EnglishProductName` (text).

	ProductKey bigint	ProductAlternateKey text	ProductSubcategoryKey double precision	WeightUnitMeasureCode text	SizeUnitMeasureCode text	EnglishProductName text
1	1	AR-5381	[null]	[null]	[null]	Adjustable Race
2	2	BA-8327	[null]	[null]	[null]	Bearing Ball
3	3	BE-2349	[null]	[null]	[null]	BB Ball Bearing
4	4	BE-2908	[null]	[null]	[null]	Headset Ball Bearings
5	5	BL-2036	[null]	[null]	[null]	Blade
6	6	CA-5965	[null]	[null]	[null]	LL Crankarm
7	7	CA-6738	[null]	[null]	[null]	ML Crankarm
8	8	CA-7457	[null]	[null]	[null]	HL Crankarm
9	9	CB-2903	[null]	[null]	[null]	Chainring Bolts
10	10	CN-6137	[null]	[null]	[null]	Chainring Nut
11	11	CR-7833	[null]	[null]	[null]	Chainring
12	12	CR-9981	[null]	[null]	[null]	Crown Race
13	13	CS-2812	[null]	[null]	[null]	Chain Stays
14	14	DC-8732	[null]	[null]	[null]	Decal 1
15	15	DC-9824	[null]	[null]	[null]	Decal 2

pgAdmin 4 interface showing the table `public.stg_DimProductCategory`. The table contains 4 rows of product category data. The columns are: `ProductCategoryKey` (bigint), `ProductCategoryAlternateKey` (bigint), `EnglishProductCategoryName` (text), `SpanishProductCategoryName` (text), and `FrenchProductCategoryName` (text).

	ProductCategoryKey bigint	ProductCategoryAlternateKey bigint	EnglishProductCategoryName text	SpanishProductCategoryName text	FrenchProductCategoryName text
1	1	1	Bikes	Bicicleta	Vélo
2	2	2	Components	Componente	Composant
3	3	3	Clothing	Prenda	Vêtements
4	4	4	Accessories	Accesorio	Accessoire



**Top Screenshot: stg\_DimProductSubcategory**

	ProductSubcategoryKey bigint	ProductSubcategoryAlternateKey bigint	EnglishProductSubcategoryName text	SpanishProductSubcategoryName text	FrenchProductSubcategoryName text
1	1	1	Mountain Bikes	Bicicleta de montaña	VTT
2	2	2	Road Bikes	Bicicleta de carretera	Vélo de route
3	3	3	Touring Bikes	Bicicleta de paseo	Vélo de randonnée
4	4	4	Handlebars	Barra	Barre d'appui
5	5	5	Bottom Brackets	Eje de pedalier	Axe de pédalier
6	6	6	Brakes	Frenos	Freins
7	7	7	Chains	Cadena	Chaîne
8	8	8	Cranksets	Bielas	Pédalier
9	9	9	Derailleurs	Desviador	Dérailleur
10	10	10	Forks	Horquilla	Fourche
11	11	11	Headsets	Dirección	Jeu de direction
12	12	12	Mountain Frames	Cuadro de montaña	Cadre de VTT
13	13	13	Pedals	Pedal	Pédale
14	14	14	Road Frames	Cuadro de carretera	Cadre de vélo de route
15	15	15	Saddles	Sillín	Selle

Total rows: 37    Query complete 00:00:00.136    CRLF    Ln 1, Col 1

**Bottom Screenshot: stg\_FactInternetSales**

	ProductKey bigint	OrderDateKey bigint	DueDateKey bigint	ShipDateKey bigint	CustomerKey bigint	PromotionKey bigint	CurrencyKey bigint	SalesTerritoryKey bigint	SalesOrderNumber text
1	310	20101229	20110110	20110105	21768	1	19	6	SO43697
2	346	20101229	20110110	20110105	28389	1	39	7	SO43698
3	346	20101229	20110110	20110105	25863	1	100	1	SO43699
4	336	20101229	20110110	20110105	14501	1	100	4	SO43700
5	346	20101229	20110110	20110105	11003	1	6	9	SO43701
6	311	20101230	20110111	20110106	27645	1	100	4	SO43702
7	310	20101230	20110111	20110106	16624	1	6	9	SO43703
8	351	20101230	20110111	20110106	11005	1	6	9	SO43704
9	344	20101230	20110111	20110106	11011	1	6	9	SO43705
10	312	20101231	20110112	20110107	27621	1	100	4	SO43706
11	312	20101231	20110112	20110107	27616	1	100	4	SO43707
12	330	20101231	20110112	20110107	20042	1	98	10	SO43708
13	313	20101231	20110112	20110107	16351	1	6	9	SO43709
14	314	20101231	20110112	20110107	16517	1	6	9	SO43710
15	314	20110101	20110113	20110108	27606	1	100	1	SO43711

Total rows: 60398    Query complete 00:00:00.489    CRLF    Ln 1, Col 1

GITHUB LINK:

<https://github.com/meghana1653/DataEngineering/blob/main/LAB3%20FDE.pdf>