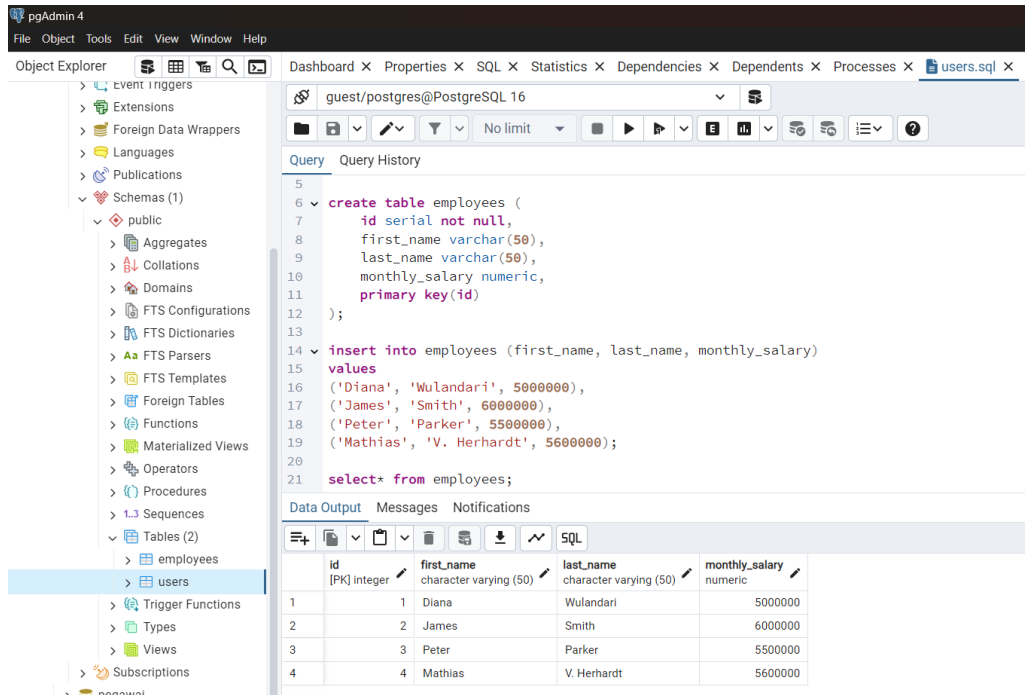


Dokumentasi Simple ETL, Data dari PostgreSQL dan Disimpan Kembali ke PostgreSQL

1. Rancangan sistem ETL

- Buat data dummy di PostgreSQL dengan nama tabel employees dan isi tabel terdiri atas first_name, last_name dan salary



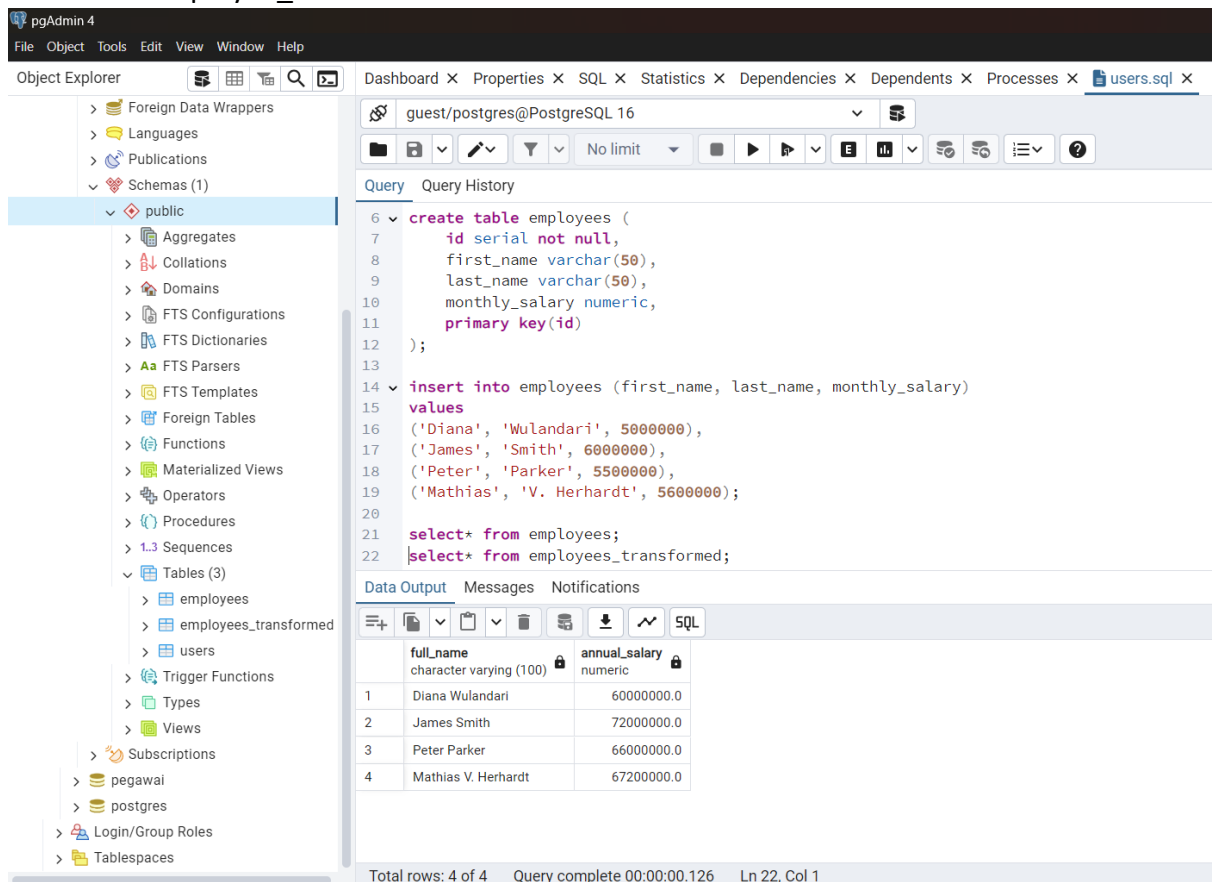
- Extract data, ambil data table employees dari PostgreSQL

```
# 1. Extract: Mengambil data dari tabel 'employees'  
def extract_data():  
    conn = connect_to_db()  
    query = "SELECT * FROM employees"  
    df = pd.read_sql(query, conn)  
    conn.close()  
    print("Data berhasil diekstrak dari PostgreSQL.")  
    return df
```

- Transformasi data, ambil full name employees dan hitung gaji tahunan

```
# 2. Transform: Menghitung gaji tahunan dan membuat kolom baru  
def transform_data(df):  
    df['annual_salary'] = df['monthly_salary'] * 12  
    df['full_name'] = df['first_name'] + ' ' + df['last_name']  
    df_transformed = df[['full_name', 'annual_salary']]  
    print("Data berhasil ditransformasi.")  
    return df_transformed
```

- d. Load data, simpan data yang telah di transformasi kembali ke PostgreSQL dengan tabel baru `employee_transformed`



The screenshot shows the pgAdmin 4 interface. On the left, the Object Explorer shows the database structure, including the 'public' schema and the 'employees' and 'employees_transformed' tables. The main pane displays a SQL query and its results.

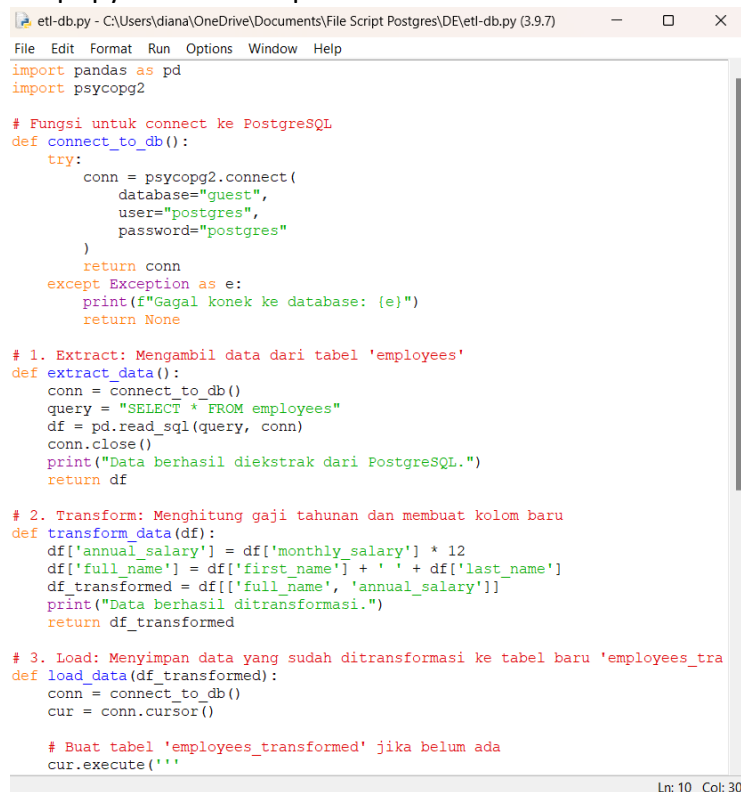
```
6 create table employees (  
7     id serial not null,  
8     first_name varchar(50),  
9     last_name varchar(50),  
10    monthly_salary numeric,  
11    primary key(id)  
12 );  
13  
14 insert into employees (first_name, last_name, monthly_salary)  
15 values  
16 ('Diana', 'Wulandari', 5000000),  
17 ('James', 'Smith', 6000000),  
18 ('Peter', 'Parker', 5500000),  
19 ('Mathias', 'V. Herhardt', 5600000);  
20  
21 select* from employees;  
22 select* from employees_transformed;
```

The Data Output pane shows the results of the query:

	full_name	annual_salary
1	Diana Wulandari	60000000.0
2	James Smith	72000000.0
3	Peter Parker	66000000.0
4	Mathias V. Herhardt	67200000.0

Total rows: 4 of 4 Query complete 00:00:00.126 Ln 22, Col 1

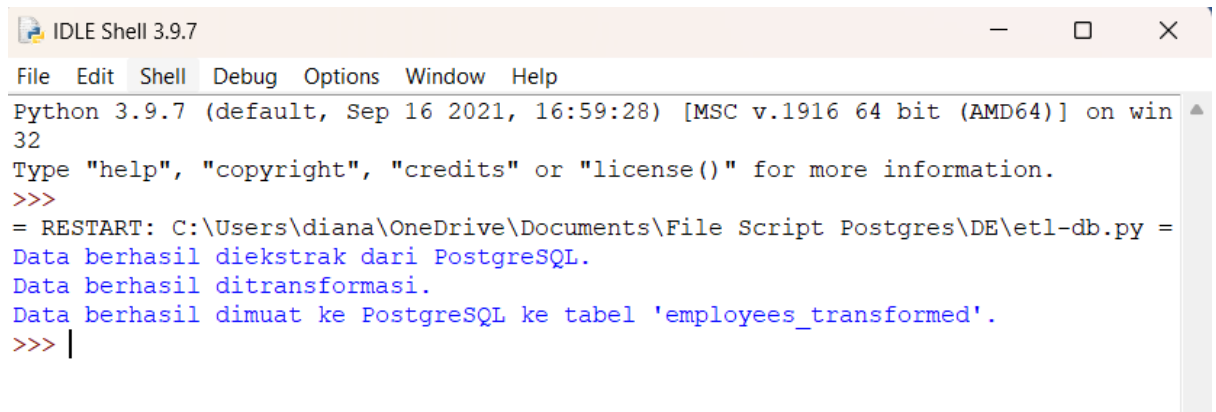
2. Script python dan output



The screenshot shows a Python script in a text editor. The script is titled 'etl-db.py' and is located at 'C:\Users\diana\OneDrive\Documents\File Script Postgres\DE\etl-db.py (3.9.7)'. The script contains the following code:

```
import pandas as pd  
import psycopg2  
  
# Fungsi untuk connect ke PostgreSQL  
def connect_to_db():  
    try:  
        conn = psycopg2.connect(  
            database="guest",  
            user="postgres",  
            password="postgres"  
        )  
        return conn  
    except Exception as e:  
        print(f"Gagal konek ke database: {e}")  
        return None  
  
# 1. Extract: Mengambil data dari tabel 'employees'  
def extract_data():  
    conn = connect_to_db()  
    query = "SELECT * FROM employees"  
    df = pd.read_sql(query, conn)  
    conn.close()  
    print("Data berhasil diekstrak dari PostgreSQL.")  
    return df  
  
# 2. Transform: Menghitung gaji tahunan dan membuat kolom baru  
def transform_data(df):  
    df['annual_salary'] = df['monthly_salary'] * 12  
    df['full_name'] = df['first_name'] + ' ' + df['last_name']  
    df_transformed = df[['full_name', 'annual_salary']]  
    print("Data berhasil ditransformasi.")  
    return df_transformed  
  
# 3. Load: Menyimpan data yang sudah ditransformasi ke tabel baru 'employees_tra'  
def load_data(df_transformed):  
    conn = connect_to_db()  
    cur = conn.cursor()  
  
    # Buat tabel 'employees_transformed' jika belum ada  
    cur.execute(''
```

Ln: 10 Col: 30



The image shows a screenshot of the IDLE Shell 3.9.7 window. The title bar at the top reads "IDLE Shell 3.9.7" and includes standard window controls (minimize, maximize, close). Below the title bar is a menu bar with the following options: File, Edit, Shell, Debug, Options, Window, and Help. The main text area displays the output of a Python script. The first line is the Python version and system information: "Python 3.9.7 (default, Sep 16 2021, 16:59:28) [MSC v.1916 64 bit (AMD64)] on win32". The second line is a prompt: "Type 'help', 'copyright', 'credits' or 'license()' for more information." This is followed by three lines of output, each preceded by a red prompt ">>>":
1. "= RESTART: C:\Users\diana\OneDrive\Documents\File Script Postgres\DE\etl-db.py ="
2. "Data berhasil diekstrak dari PostgreSQL."
3. "Data berhasil ditransformasi."
The final line shows the data being loaded into a table: "Data berhasil dimuat ke PostgreSQL ke tabel 'employees_transformed'." The prompt ">>>" is followed by a vertical cursor bar.

```
Python 3.9.7 (default, Sep 16 2021, 16:59:28) [MSC v.1916 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:\Users\diana\OneDrive\Documents\File Script Postgres\DE\etl-db.py =
Data berhasil diekstrak dari PostgreSQL.
Data berhasil ditransformasi.
Data berhasil dimuat ke PostgreSQL ke tabel 'employees_transformed'.
>>> |
```