

# Dokumentasi Simple ETL Dengan API ke PostgreSQL

## 1. Ambil data dengan API

```
etl-api.py - C:\Users\diana\OneDrive\Documents\File Script Postgres\DE\etl-api.py (3.9.7)
File Edit Format Run Options Window Help

import requests
import pandas as pd
import psycopg2

# 1. Extract: Ambil data dari API
response = requests.get("https://jsonplaceholder.typicode.com/users")
data = response.json()

# Convert data dari JSON ke pandas DataFrame
df = pd.DataFrame(data)

# Simpan data mentah ke file CSV sebagai "backup" atau bagian dari data lake
df.to_csv('data_lake/users_raw.csv', index=False)
print("Data mentah berhasil diambil dari API dan disimpan ke data_lake/users_raw")

# 2. Transform: Membersihkan dan memilih data
# Misalnya, kita hanya mengambil kolom name, email, dan city dari alamat pengguna
df_transformed = df[['name', 'email']]
df_transformed['city'] = df['address'].apply(lambda x: x['city'])

# Simpan data yang sudah ditransformasi
df_transformed.to_csv('data_lake/users_clean.csv', index=False)
print("Data berhasil ditransformasi dan disimpan di data_lake/users_clean.csv")

# 3. Load: Memuat data yang sudah bersih ke PostgreSQL
try:
    conn = psycopg2.connect(
        database="guest",
        user="postgres",
        password="postgres"
    )

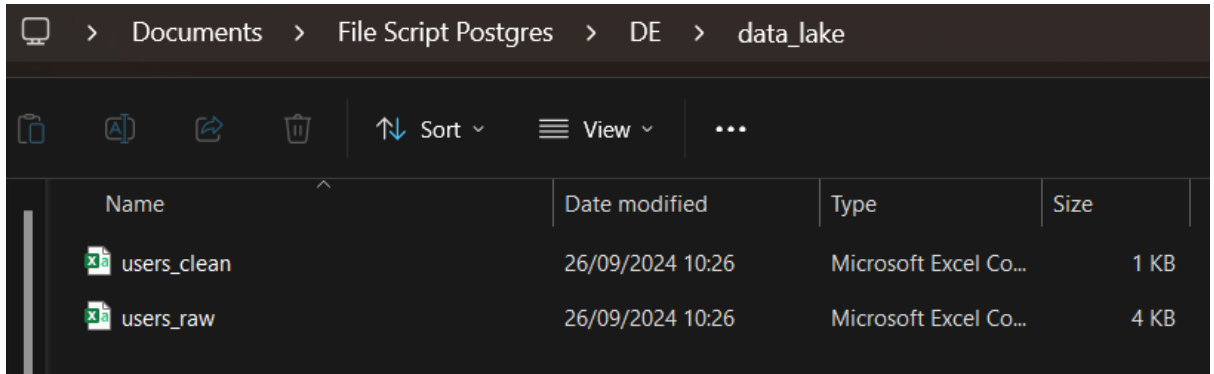
    cur = conn.cursor()

    # Buat tabel jika belum ada
    cur.execute('''
CREATE TABLE IF NOT EXISTS users (
    name VARCHAR(100),
    email VARCHAR(100),
    city VARCHAR(100)
)

Warning (from warnings module):
  File "C:\Users\diana\OneDrive\Documents\File Script Postgres\DE\etl-api.py", line 19
    df_transformed['city'] = df['address'].apply(lambda x: x['city'])
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

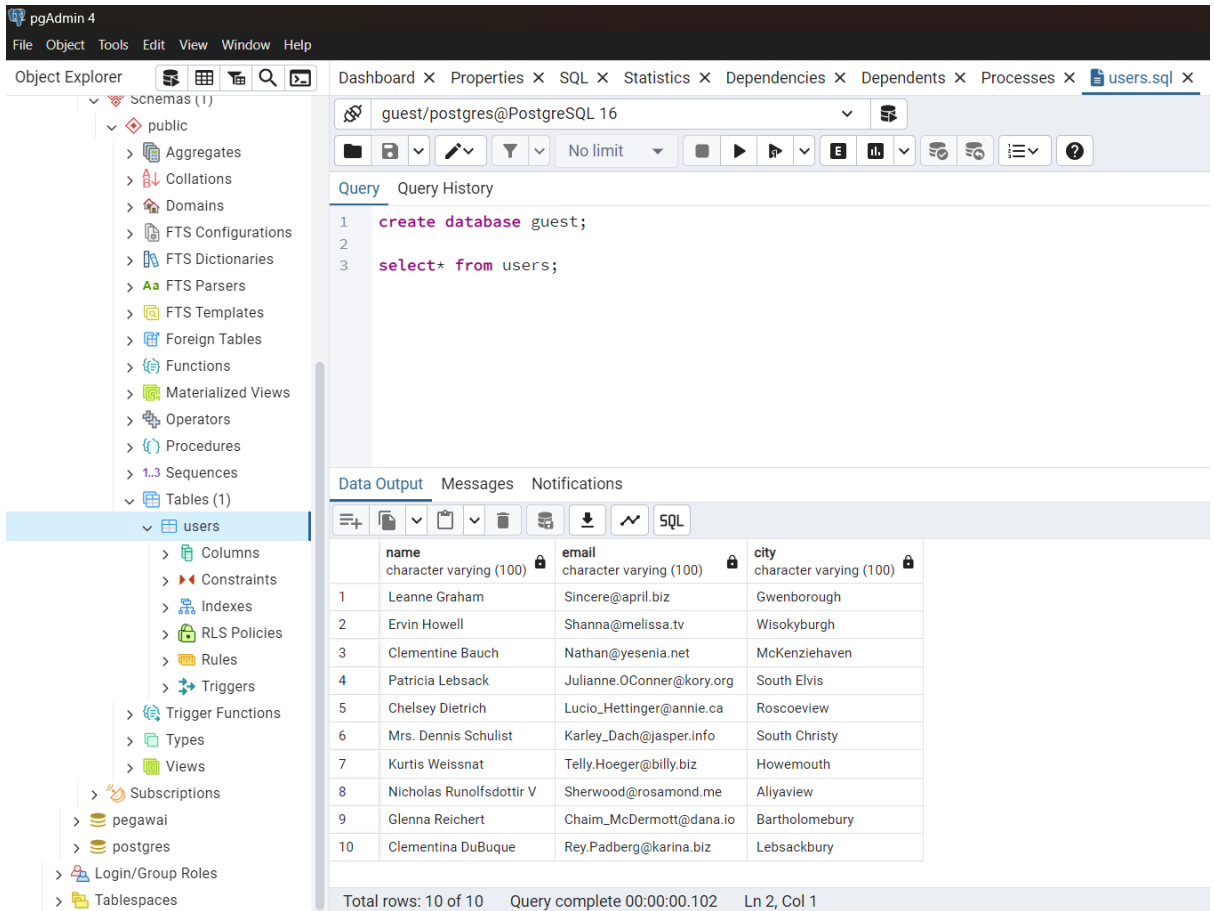
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
Data berhasil ditransformasi dan disimpan di data_lake/users_clean.csv
Data berhasil dimuat ke PostgreSQL
>>> |
```

## 2. Backup data mentah



Name	Date modified	Type	Size
users_clean	26/09/2024 10:26	Microsoft Excel Co...	1 KB
users_raw	26/09/2024 10:26	Microsoft Excel Co...	4 KB

## 3. Transform data dan Load data ke PostgreSQL



pgAdmin 4

File Object Tools Edit View Window Help

Object Explorer

- Schemas (1)
  - public
    - Aggregates
    - Collations
    - Domains
    - FTS Configurations
    - FTS Dictionaries
    - FTS Parsers
    - FTS Templates
    - Foreign Tables
    - Functions
    - Materialized Views
    - Operators
    - Procedures
    - Sequences
    - Tables (1)
      - users
        - Columns
        - Constraints
        - Indexes
        - RLS Policies
        - Rules
        - Triggers

- Subscriptions
- pegawai
- postgres
- Login/Group Roles
- Tablespaces

Dashboard x Properties x SQL x Statistics x Dependencies x Dependents x Processes x users.sql x

guest/postgres@PostgreSQL 16

No limit

Query Query History

```
1 create database guest;
2
3 select* from users;
```

Data Output Messages Notifications

	name	email	city
	character varying (100)	character varying (100)	character varying (100)
1	Leanne Graham	Sincere@april.biz	Gwenborough
2	Ervin Howell	Shanna@melissa.tv	Wisokyburgh
3	Clementine Bauch	Nathan@yesenia.net	McKenziehaven
4	Patricia Lebsack	Julianne.OConner@kory.org	South Elvis
5	Chelsey Dietrich	Lucio_Hettinger@annie.ca	Roscoeview
6	Mrs. Dennis Schulist	Karley_Dach@jasper.info	South Christy
7	Kurtis Weissnat	Telly.Hoeger@billy.biz	Howemouth
8	Nicholas Runolfsdottir V	Sherwood@rosamond.me	Aliyaview
9	Glenna Reichert	Chaim_McDermott@dana.io	Bartholomebury
10	Clementina DuBuque	Rey.Padberg@karina.biz	Lebsackbury

Total rows: 10 of 10 Query complete 00:00:00.102 Ln 2, Col 1