

MySkill

#RintisKarirImpian

SQL Basics 2

Rizki Teguh Kurniawan

Data Scientist, eFishery

Youtube: Solving with Akoyo

Jumat, 28 Juli 2023



Jawaban Mini Task SQL Basic 1

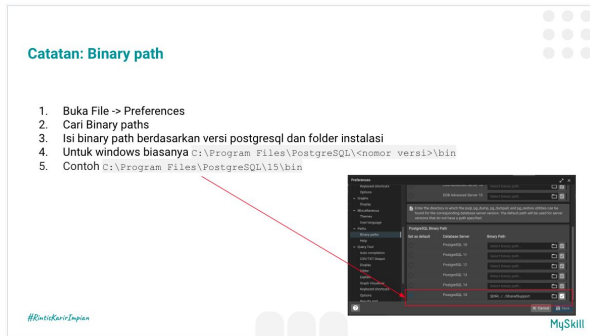
```
create table orders (  
    order_id varchar,  
    customer_id varchar,  
    postal_code varchar,  
    product_id varchar,  
    sales float,  
    quantity int,  
    discount float,  
    profit float,  
    category varchar,  
    subcategory varchar,  
    product_name varchar,  
    order_date date,  
    ship_date date,  
    ship_mode varchar,  
    customer_name varchar,  
    segment varchar,  
    country varchar,  
    city varchar,  
    state varchar,  
    region varchar  
)
```



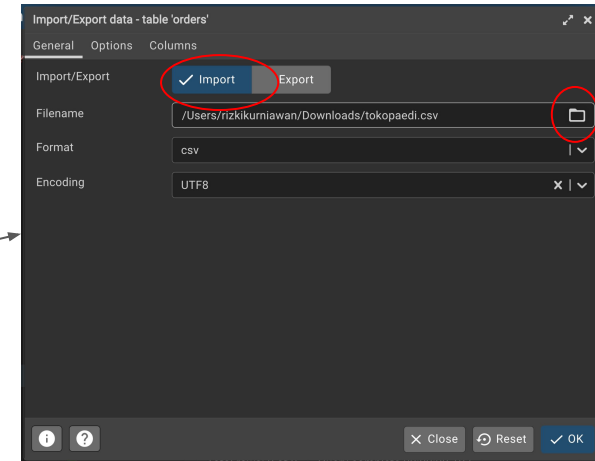
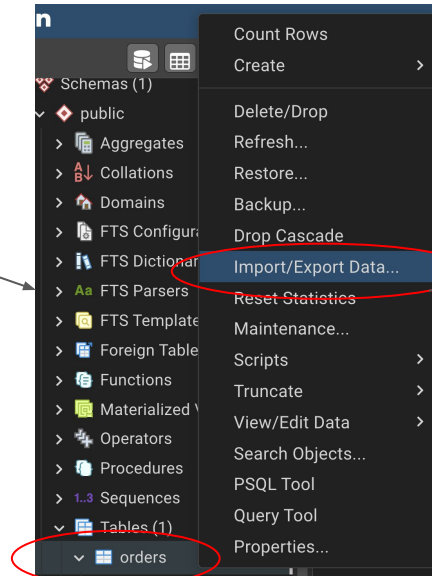
Silakan download
orders.csv

Import From CSV

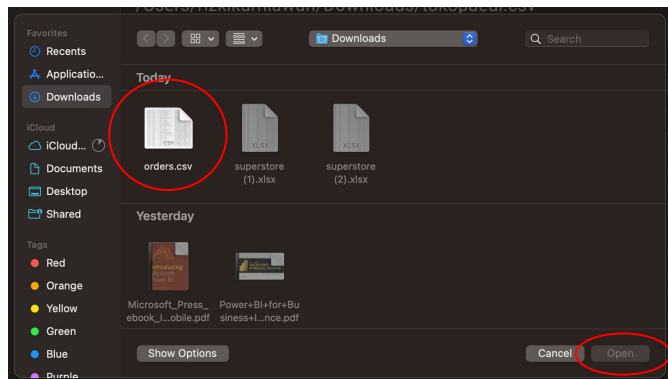
1. Dari mini task sebelumnya kita memiliki tabel orders
2. Selanjutnya kita akan membuat tabel dari file eksternal (csv)
3. perhatikan langkah berikut



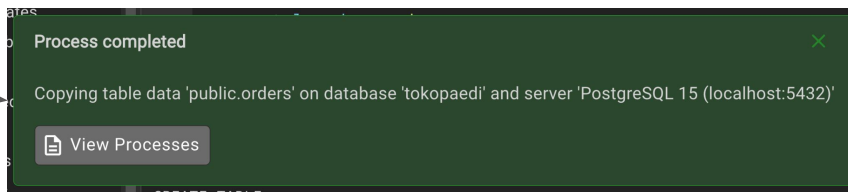
Mengatur binary path



Import From CSV



Klik pada berkas orders.csv lalu pilih Open



Tanda jika proses import data sudah selesai,
sekarang tabel Order memiliki nilai

Strategy change, agreement

1. Background check
2. Karena background peserta yang *highly diverse*, saya mengasumsikan pemahaman sql yang sangat dasar
3. QnA berjalan selama pemberian materi namun akan diprioritaskan untuk pertanyaan yang mendesak dan relevan dengan materi yg sedang berlangsung

Perbedaan Data Engineer, Data Analyst, dan Data Scientist

Data Engineer



Build and optimise the systems that allow data scientists and data analysts to perform their work

Requirements:

1. Strong programming skills
2. Cloud computing technologies
3. Big data

Tech stack: SQL, Python, Cloud, Distributed Computing

Data Analyst



Deliver value by analysing data, communicating the results to help make business decisions

Requirements:

1. Communication skills
2. Business savvy/domain knowledge

Tech stack: SQL, BI tools, Python, R

Data Scientist



Use data to solve business problems

Requirements:

1. Statistics
2. Mathematics
3. Programming skills
4. Big data

Tech stack: SQL, Python, R, Cloud

Agenda



- *Data preparation and Import data*
- *Data understanding*
- *Memilih kolom dengan SELECT FROM*
- *Memfilter data dengan WHERE*
- *Operator logika: AND, OR, NOT, dan IN*
- *Mengenal NULL values*



Memilih kolom dari tabel menggunakan **SELECT** dari **FROM**

```
SELECT
nama_kolom_1,
nama_kolom_2,
...
FROM
nama_tabel
```

Untuk nama kolom yang tidak mengandung spasi tidak perlu menggunakan tanda kutip

Tips: Untuk menampilkan semua kolom, gunakan * (asterisk)
Contoh: **SELECT * FROM** transactions

```
SELECT
    "Order ID",
    "Order Date",
    "Customer Name"
FROM
    transactions
```

Order ID character varying	Order Date date	Customer Name character varying
CA-2016-152156	2016-11-08	Claire Gute
CA-2016-152156	2016-11-08	Claire Gute
CA-2016-138688	2016-06-12	Darrin Van Huff
US-2015-108966	2015-10-11	Sean O'Donnell



Menyaring data menggunakan **WHERE**

```
SELECT
nama_kolom_1,
nama_kolom_2,
...
FROM
nama_tabel
WHERE
kondisi
```

Pada PostgreSQL nama kolom diapit oleh " sedangkan string diapit oleh '

```
SELECT
    "Order ID",
    "Order Date",
    "Customer Name"
FROM
    transactions
WHERE
    "Customer Name" = 'Claire Gute'
```

	Order ID character varying	Order Date date	Customer Name character varying
1	CA-2016-152156	2016-11-08	Claire Gute
2	CA-2016-152156	2016-11-08	Claire Gute
3	CA-2017-164098	2017-01-26	Claire Gute
4	US-2015-123918	2015-10-15	Claire Gute
5	US-2015-123918	2015-10-15	Claire Gute



Mengenal operator logika **AND**, **OR**, dan **NOT**

Klausula **WHERE** bisa digabungkan dengan operator **AND**, **OR**, dan **NOT**

Operator di atas digunakan untuk memfilter baris berdasarkan satu atau lebih kondisi

1. Operator **AND** akan menampilkan baris jika *semua* kondisi yang dipisahkan oleh **AND** bernilai **TRUE**
2. Operator **OR** akan menampilkan baris jika *salah satu* kondisi yang dipisahkan oleh **OR** bernilai **TRUE**
3. Operator **NOT** akan membalikkan nilai kebenaran dari suatu kondisi

p	q	p AND q	p OR q	NOT p	NOT q
T	T	T	T	F	F
T	F	F	T	F	T
F	T	F	T	T	F
F	F	F	F	T	T

Contoh

```
SELECT
  "Customer Name"
FROM
  transactions
WHERE
  "Segment" = 'Consumer' AND "City" = 'Kentucky'
```

Memunculkan nama konsumen dalam segmen Consumer yang berlokasi di Kota Kentucky

```
SELECT
  "Customer Name"
FROM
  transactions
WHERE
  "City" = 'California' OR "City" = 'Kentucky'
```

Memunculkan nama konsumen di Kota California atau Kota Kentucky

Operator **IN**

Bagaimana jika kondisinya seperti berikut:

Kamu diminta untuk memunculkan daftar transaksi di 5 kota: Los Angeles, Houston, Durham, San Francisco, Columbus

```
SELECT
    *
FROM
    transactions
WHERE
    "City" = 'Los Angeles' OR
    "City" = 'Houston' OR
    "City" = 'Durham' OR
    "City" = 'San Francisco' OR
    "City" = 'Columbus'
```

```
SELECT
    *
FROM
    transactions
WHERE
    "City" IN ('Los Angeles',
               'Houston',
               'Durham',
               'San Francisco',
               'Columbus')
```

*Operator **IN** digunakan untuk menyingkat beberapa operasi **OR** sekaligus*

Operator **NOT IN**

Operator **NOT IN** digunakan untuk memfilter baris yang dengan nilai yang tidak termuat di dalam operator **IN**

```
SELECT
    *
FROM
    transactions
WHERE
    "City" NOT IN ('Los Angeles', 'Houston', 'Durham', 'San Francisco', 'Columbus')
```

Query di atas akan memunculkan daftar transaksi di beberapa kota kecuali kota: Los Angeles, Houston, Durham, San Francisco, dan Columbus

Operator **IS NULL** dan **IS NOT NULL**

Entri dengan nilai **NULL** artinya entri tersebut tidak memiliki nilai (kosong)

Untuk entri dengan nilai **NULL** tidak berlaku operator perbandingan seperti =, <, atau >

Maka untuk menguji apakah suatu entri tersebut bernilai **NULL** atau tidak digunakan **IS NULL** atau **IS NOT NULL**

```
SELECT
    *
FROM
    transactions
WHERE
    "Postal Code" IS NULL
```

Menampilkan transaksi yang tidak memuat informasi kode pos

```
SELECT
    *
FROM
    transactions
WHERE
    "Postal Code" IS NOT NULL
```

Hanya menampilkan transaksi dengan informasi kode pos



Mini Task

Kamu sekarang memiliki tabel transaksi dari customer yang bertransaksi di Tokopaedi. Selanjutnya kamu diminta untuk

1. Tampilkan nama-nama konsumen segment Consumer yang pernah membeli meja
2. Tampilkan nama-nama konsumen dari segment Corporate dan Home Office yang berasal dari kota Los Angeles dan bertransaksi selama tahun 2018.

