

LAPORAN PRAKTIKUM
JOBSHEET 03
DATABASE ANALYTICAL (A dan B)

Disusun untuk memenuhi nilai tugas
Mata Kuliah : Data Warehouse



Oleh :
Aqueena Regita Hapsari
2341760096
SIB-2B - 03

PROGRAM STUDI D-IV SISTEM INFORMASI BISNIS
POLITEKNIK NEGERI MALANG
TAHUN AJARAN 2024/2025

STUDI KASUS

1. Buatlah sebuah database yang digunakan sebagai OLAP dengan nama dw_LegendVehicle.

```
CREATE DATABASES dw_LegendVehicle;
```

```
use dw_LegendVehicle;
```

```
✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.2254 seconds.)

CREATE DATABASE dw_LegendVehicle;

[ Edit inline ] [ Edit ] [ Create PHP code ]
```

```
✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0003 seconds.)

use dw_LegendVehicle;

[ Edit inline ] [ Edit ] [ Create PHP code ]
```

2. Buatlah table untuk menyimpan data master waktu atau yang disebut dengan tabel dimensi. Beri nama table tersebut dengan nama dimDate.

```
CREATE TABLE dimDate(
```

```
id_dimDate int not null AUTO_INCREMENT PRIMARY KEY
```

```
date date, year int, month int, day int );
```

```
✓ MySQL returned an empty result set (i.e. zero rows). (Query took 1.1353 seconds.)

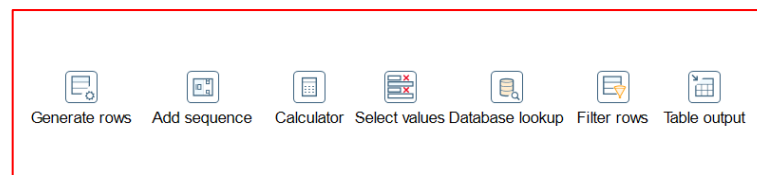
CREATE TABLE dimDate ( id_dimDate INT NOT NULL AUTO_INCREMENT
PRIMARY KEY, date DATE, year INT, month INT, day INT );

[ Edit inline ] [ Edit ] [ Create PHP code ]
```

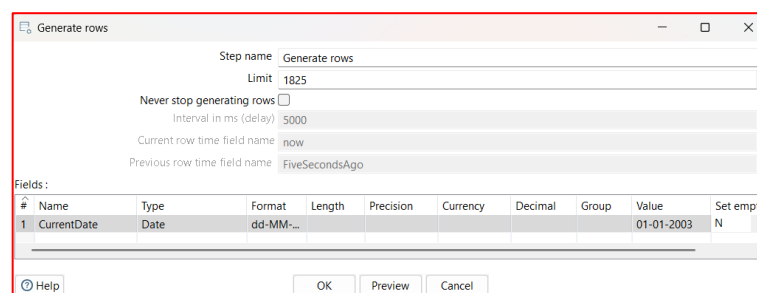
Pada tahapan selanjutnya, untuk membuat tabel dimensi dimDate , maka diperlukan generate data tanggal. Data tanggal yang disiapkan pada tabel dimDate menyesuaikan dengan proses bisnis yang berjalan.

1. Proses bisnis pada LegendVehicle adalah 5 tahun. Sehingga data pada tabel dimdate yang harus tersedia adalah tanggal selama 5 tahun. Mulai dari 1 Januari 2023
2. Buka PDI Spoon. Buat Transformation baru -> File - New - Transformation.
3. Drag and Drop beberapa objek yaitu:
 - a) Generate Rows: digunakan untuk membuat baris data baru.

- b) Add Sequence: digunakan untuk membuat sequence, dalam hal ini membuat data di setiap harinya.
- c) Calculator: digunakan untuk menjumlahkan hari dan mengambil data tahun, bulan dan hari.
- d) Select Values: digunakan untuk memilih field yang digunakan.
- e) Database Lookup: digunakan untuk melihat dan memastikan bahwa data yang akan dimasukkan kedalam tabel dimDate tidak kembar atau sama dengan data yang ada pada tabel dimDate itu sendiri.
- f) Filter Rows: digunakan untuk mengambil data yang belum ada pada table dimDate setelah dicek sebelumnya.
- g) Table Output: digunakan untuk menyimpan data pada tabel tujuan (dimDate).



4. Konfigurasi pada Generate Rows adalah merubah limit menjadi 1825 dimana memiliki arti bahwa data yang akan dibuat sebanyak 1825 data. 1825 merupakan jumlah hari dalam 5 tahun (365 hari x 5 tahun).
5. Membuat fields baru bernama CurrentDate dengan type data Date dan format dd-MM-yyyy serta value awal 01-01-2003.



Gambar konfigurasi generate rows

6. Hubungkan output dari Generate Rows menuju Add Sequence.
7. Konfigurasi pada Add Sequences adalah merubah Name of value menjadi incrementDay dengan start value bernilai 0 dan increment by bernilai 1

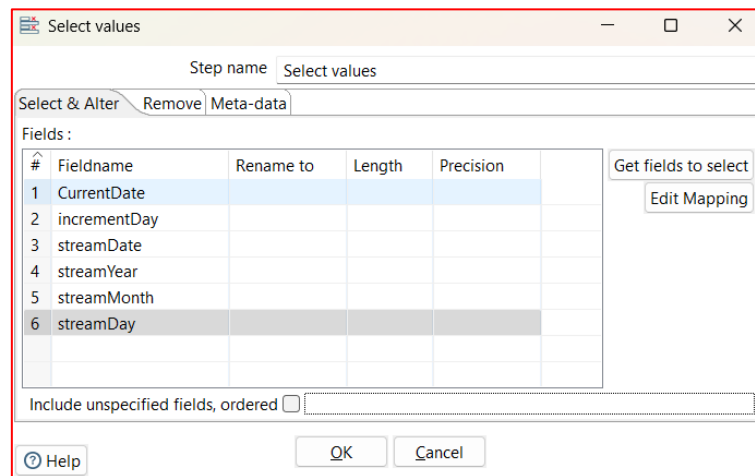
Gambar konfigurasi add sequences

8. Hubungkan output dari add sequences menuju calculator.
9. Konfigurasi pada calculator dengan membuat fields baru sebagai berikut:
 - a) streamDate merupakan kalkulasi dari CurrentDate + incrementDay
 - b) streamYear merupakan Year dari streamDate
 - c) streamMonth merupakan Month dari streamDate
 - d) streamDay merupakan Day of month dari streamDate

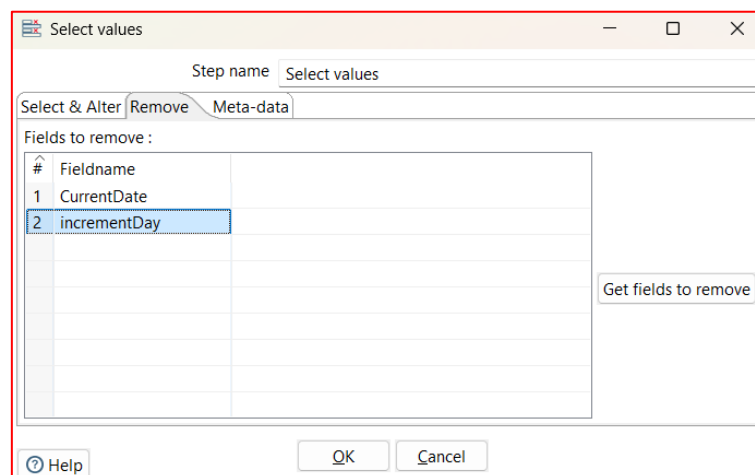
Gambar konfigurasi kalkulator

10. Hubungkan output dari calculator menuju Select values
11. Konfigurasi pada select values adalah dengan menekan tombol Get fields to select pada tab Select & Alter. Secara otomatis semua fields dari data input akan muncul.

Dikarenakan tidak semua fields digunakan, maka pada tab Remove diisikan fields CurrentDate dan incrementDay dikarenakan kedua fields tersebut tidak digunakan.

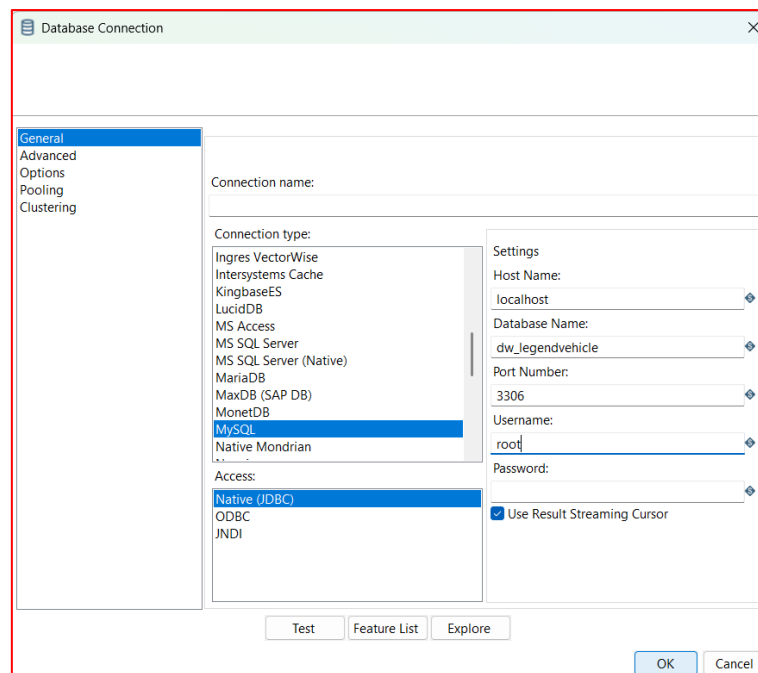


Gambar konfigurasi tab select & alter pada select values



Gambar konfigurasi tab remove pada select values

12. Hubungkan output select values menuju database lookup.
13. Sebelum melakukan konfigurasi pada database lookup, buatlah koneksi terlebih dahulu pada database melalui File - New - Database Connection. Gunakan Connection type MySQL dengan host name, database name, port number, username dan password sesuai konfigurasi MySQL pada device masing-masing. beri nama connection name tersebut dengan nama conn_dw_destination.



Gambar konfigurasi database connection

14. Konfigurasi pada database lookup adalah dengan memberikan connection dengan koneksi yang sudah dibuat pada step sebelumnya. dengan schema nama database yang digunakan dan tabel dimdate yang telah dibuat pada langkah pertama.
 - a) Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak kembar adalah:
 - b) field date pada table dimdate dengan field streamDate
 - c) field year pada table dimdate dengan field streamYear
 - d) field month pada table dimdate dengan field streamMonth
 - e) field day pada table dimdate dengan field streamDay
 - f) Field yang akan di retrieve adalah field yang ada pada table dimDate yaitu date, year, month, dan day.

Database lookup

Step name: Database lookup

Connection: conn_dw_destination [Edit...] [New...] [Wizard...]

Lookup schema: dw_legendvehicle [Browse...]

Lookup table: dimdate [Browse...]

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	date	=	stream...	
2	year	=	stream...	
3	month	=	stream...	
4	day	=	stream...	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	date			None
2	year			None
3	month			None
4	day			None

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

[?] Help [OK] [Cancel] [Get Fields] [Get lookup fields]

Gambar konfigurasi database lookup

15. Hubungkan output dari database lookup dengan filter rows
16. Konfigurasi pada filter rows adalah dengan melakukan konfigurasi output true data pada table output. Pada bagian ini data yang tidak memiliki kesamaan pada tahapan sebelumnya akan dicek dimana jika fields Stream tidak memiliki kesamaan dengan field dimDate, maka field dimDate tersebut akan bernilai null. Pada pernyataan kondisi tuliskan (date is null and year is null and month is null and day is null)

Filter rows

Step name: Filter rows

Send 'true' data to step: Table output

Send 'false' data to step:

The condition:

☐ To edit a subcondition, simply click on it +

date IS NULL

AND

year IS NULL

AND

month IS NULL

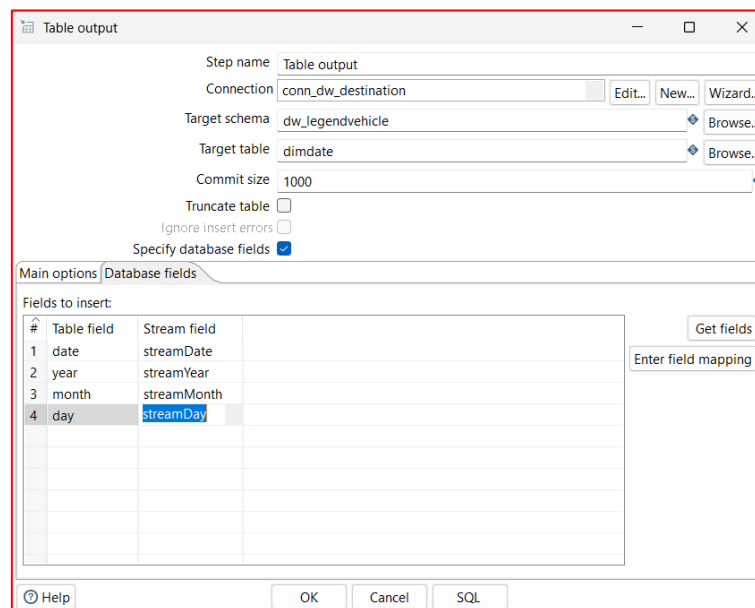
AND

day IS NULL

[?] Help [OK] [Cancel]




Gambar konfigurasi filter rows

17. Hubungkan output dari filter rows menuju table output.
18. Konfigurasi pada table output adalah memberikan koneksi pada conn_dw_destination dengan schema dw_legendvehicle dan table dimdate.
19. Aktifkan specify database fields.
20. Pada tab Database fields, mapping data input streamDate, streamYear, streamMonth dan streamDay dengan fields yang ada pada dimDate. Pada tahapan ini akan dilakukan insert data menuju tabel dimDate.



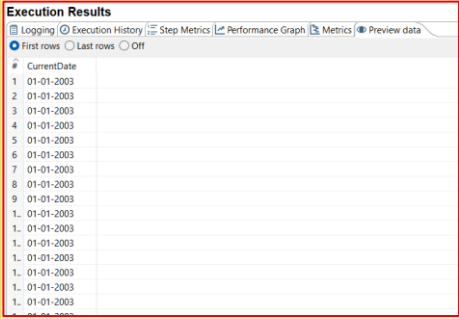
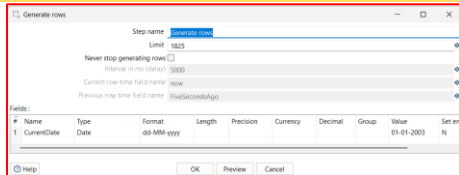
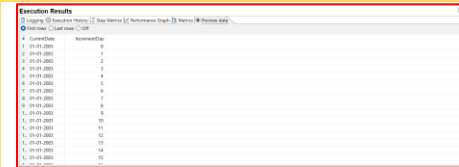
Gambar Konfigurasi table output

21. cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.

<div><div><div></div><div></div></div><div><div></div><div></div></div></div>							Id_dimDate	date	year	month	day
<input type="checkbox"/>		Edit		Copy		Delete	1	2003-01-01	2003	1	1
<input type="checkbox"/>		Edit		Copy		Delete	2	2003-01-02	2003	1	2
<input type="checkbox"/>		Edit		Copy		Delete	3	2003-01-03	2003	1	3
<input type="checkbox"/>		Edit		Copy		Delete	4	2003-01-04	2003	1	4
<input type="checkbox"/>		Edit		Copy		Delete	5	2003-01-05	2003	1	5
<input type="checkbox"/>		Edit		Copy		Delete	6	2003-01-06	2003	1	6
<input type="checkbox"/>		Edit		Copy		Delete	7	2003-01-07	2003	1	7
<input type="checkbox"/>		Edit		Copy		Delete	8	2003-01-08	2003	1	8
<input type="checkbox"/>		Edit		Copy		Delete	9	2003-01-09	2003	1	9
<input type="checkbox"/>		Edit		Copy		Delete	10	2003-01-10	2003	1	10
<input type="checkbox"/>		Edit		Copy		Delete	11	2003-01-11	2003	1	11
<input type="checkbox"/>		Edit		Copy		Delete	12	2003-01-12	2003	1	12
<input type="checkbox"/>		Edit		Copy		Delete	13	2003-01-13	2003	1	13
<input type="checkbox"/>		Edit		Copy		Delete	14	2003-01-14	2003	1	14
<input type="checkbox"/>		Edit		Copy		Delete	15	2003-01-15	2003	1	15
<input type="checkbox"/>		Edit		Copy		Delete	16	2003-01-16	2003	1	16
<input type="checkbox"/>		Edit		Copy		Delete	17	2003-01-17	2003	1	17
<input type="checkbox"/>		Edit		Copy		Delete	18	2003-01-18	2003	1	18
<input type="checkbox"/>		Edit		Copy		Delete	19	2003-01-19	2003	1	19
<input type="checkbox"/>		Edit		Copy		Delete	20	2003-01-20	2003	1	20
<input type="checkbox"/>		Edit		Copy		Delete	21	2003-01-21	2003	1	21
<input type="checkbox"/>		Edit		Copy		Delete	22	2003-01-22	2003	1	22
<input type="checkbox"/>		Edit		Copy		Delete	23	2003-01-23	2003	1	23
<input type="checkbox"/>		Edit		Copy		Delete	24	2003-01-24	2003	1	24
<input type="checkbox"/>		Edit		Copy		Delete	25	2003-01-25	2003	1	25

TUGAS 1

1. Buka preview tab pada execution result area di setiap proses object. amati input dan output data yang ada. bandingkan di setiap prosesnya. jelaskan perbedaan di setiap prosesnya.

No.	Proses Objek	SS Data Input	SS Data Output	Keterangan
1.	Generate Rows	(tidak ada karena ini adalah tahap awal)		<ul style="list-style-type: none"> • Output Data: Menghasilkan sejumlah baris sesuai dengan limit yang ditentukan (misalnya 1825 untuk 5 tahun). • Perbedaan: Tahap ini hanya membuat baris kosong yang nantinya akan diisi data.
2.	Add Sequences			<ul style="list-style-type: none"> • Input Data: Data kosong dari Generate Rows. • Output Data: Menambahkan kolom sequence (id_dimDate) yang berisi angka bertambah mulai dari 1 hingga jumlah baris yang dihasilkan. • Perbedaan: Sekarang setiap baris memiliki ID unik yang dapat digunakan sebagai primary key.

3. Calculator

#	CurrentDate	IncrementDay	streamNumber	streamMonth	streamDay
1	2003-01-01 00:00:00.000	1	2003	1	1
2	2003-01-02 00:00:00.000	2	2003	1	2
3	2003-01-03 00:00:00.000	3	2003	1	3
4	2003-01-04 00:00:00.000	4	2003	1	4
5	2003-01-05 00:00:00.000	5	2003	1	5
6	2003-01-06 00:00:00.000	6	2003	1	6
7	2003-01-07 00:00:00.000	7	2003	1	7
8	2003-01-08 00:00:00.000	8	2003	1	8
9	2003-01-09 00:00:00.000	9	2003	1	9
10	2003-01-10 00:00:00.000	10	2003	1	10
11	2003-01-11 00:00:00.000	11	2003	1	11
12	2003-01-12 00:00:00.000	12	2003	1	12
13	2003-01-13 00:00:00.000	13	2003	1	13
14	2003-01-14 00:00:00.000	14	2003	1	14
15	2003-01-15 00:00:00.000	15	2003	1	15

- **Input Data:** Data dari Add Sequence, hanya berisi ID.
- **Output Data:** Menambahkan kolom tanggal (CurrentDate) yang dihitung dari tanggal awal (Start Date) dengan penambahan sequence sebagai offset hari.
- **Perbedaan:** Sekarang setiap baris memiliki tanggal yang berbeda, biasanya dimulai dari 01-01-2023.

4. Select Values

#	CurrentDate	streamNumber	streamMonth	streamDay
1	2003-01-01 00:00:00.000	2003	1	1
2	2003-01-02 00:00:00.000	2003	1	2
3	2003-01-03 00:00:00.000	2003	1	3
4	2003-01-04 00:00:00.000	2003	1	4
5	2003-01-05 00:00:00.000	2003	1	5
6	2003-01-06 00:00:00.000	2003	1	6
7	2003-01-07 00:00:00.000	2003	1	7
8	2003-01-08 00:00:00.000	2003	1	8
9	2003-01-09 00:00:00.000	2003	1	9
10	2003-01-10 00:00:00.000	2003	1	10
11	2003-01-11 00:00:00.000	2003	1	11
12	2003-01-12 00:00:00.000	2003	1	12
13	2003-01-13 00:00:00.000	2003	1	13
14	2003-01-14 00:00:00.000	2003	1	14
15	2003-01-15 00:00:00.000	2003	1	15

- **Input Data:** Semua kolom yang dihasilkan oleh Calculator.
- **Output Data:** Hanya kolom yang dipilih, misalnya CurrentDate, Year, Month, dan Day.
- **Perbedaan:** Menghapus kolom yang tidak diperlukan agar output lebih bersih.

5. Database Lookup

#	CurrentDate	streamNumber	streamMonth	streamDay	year	month	day
1	2003-01-01 00:00:00.000	2003	1	1	2003	1	1
2	2003-01-02 00:00:00.000	2003	1	2	2003	1	2
3	2003-01-03 00:00:00.000	2003	1	3	2003	1	3
4	2003-01-04 00:00:00.000	2003	1	4	2003	1	4
5	2003-01-05 00:00:00.000	2003	1	5	2003	1	5
6	2003-01-06 00:00:00.000	2003	1	6	2003	1	6
7	2003-01-07 00:00:00.000	2003	1	7	2003	1	7
8	2003-01-08 00:00:00.000	2003	1	8	2003	1	8
9	2003-01-09 00:00:00.000	2003	1	9	2003	1	9
10	2003-01-10 00:00:00.000	2003	1	10	2003	1	10
11	2003-01-11 00:00:00.000	2003	1	11	2003	1	11
12	2003-01-12 00:00:00.000	2003	1	12	2003	1	12
13	2003-01-13 00:00:00.000	2003	1	13	2003	1	13
14	2003-01-14 00:00:00.000	2003	1	14	2003	1	14
15	2003-01-15 00:00:00.000	2003	1	15	2003	1	15

- **Input Data:** Data dari Select Values.
- **Output Data:** Data yang telah dicek apakah sudah ada di database (dimDate).
- **Perbedaan:** Menambahkan kolom hasil pencocokan dengan

6. Filter Rows

database, misalnya ID yang sudah ada.

- **Input Data:** Data dari Database Lookup.
- **Output Data:** Hanya menyimpan data yang belum ada dalam tabel dimDate untuk menghindari duplikasi.
- **Perbedaan:** Menghapus baris yang sudah ada dalam database.

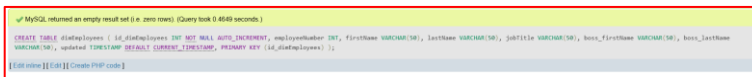
7. Table Output

- **Input Data:** Data yang lolos filter (belum ada dalam database).
- **Output Data:** Data ditulis ke tabel dimDate dalam MySQL.
- **Perbedaan:** Data sudah tersimpan dalam database dan bisa diakses melalui phpMyAdmin.

B. DIMENSI PEGAWAI

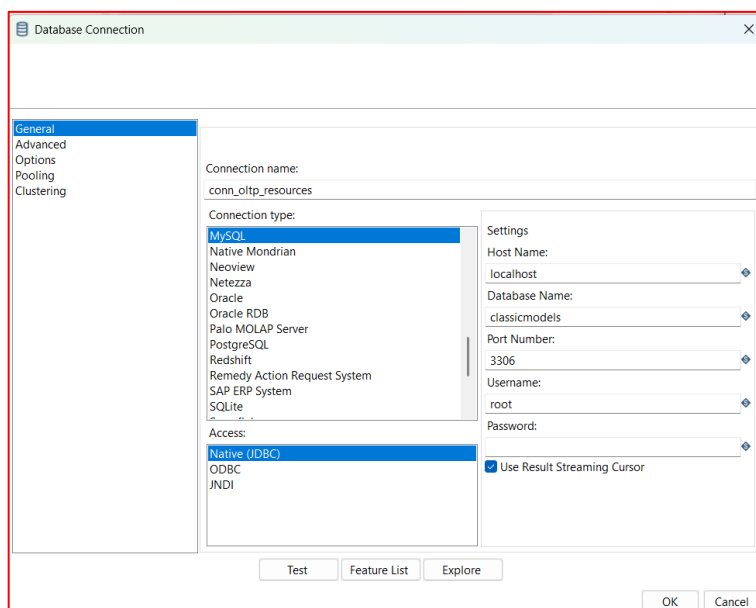
1. Buatlah tabel **dimPegawai** pada **dw_legendVehicle**.

```
CREATE TABLE dimEmployees(
  id_dimEmployees int not null auto increment primary,
  employeeNumber int(11),
  firstName varchar(50),
  lastname varchar(50),
  jobtitle varchar(50),
  boss_firstname varchar(50),
  boss_lastname varchar(50),
  updated date DEFAULT CURRENT_DATE
);
```



2. Pada PDI Spoon buatlah koneksi baru dengan nama **conn_oltp_resources** yang menghubungkan dengan database oltp. sesuaikan **hostname**, **database name**, **port number**, **username** dan **password** dengan keadaan pada device masing-masing.

(CONNECTION 1 KE DATABASE CLASSICMODELS)

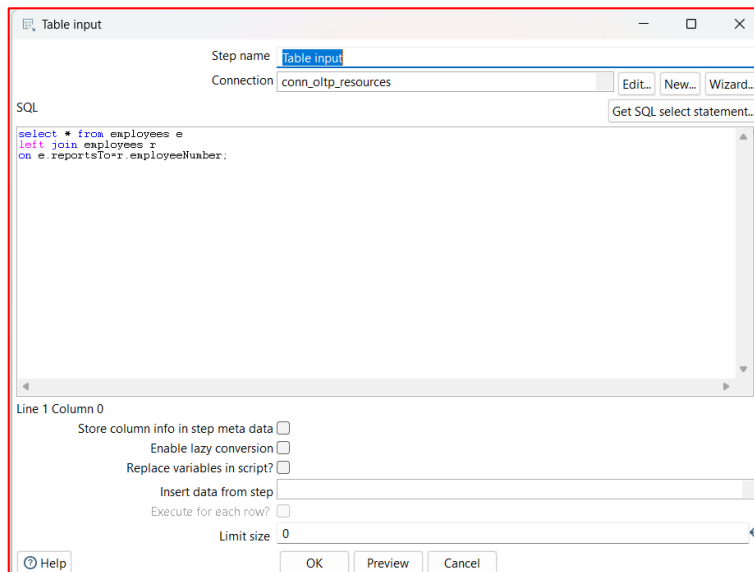


3. Drag and drop beberapa objek sebagai berikut:
 - a) **Table input:** digunakan mengambil data dari database OLTP.
 - b) **Select values:** memilih field yang digunakan untuk proses Transform dan Load.
 - c) **Database lookup:** digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar
 - d) **Filter rows:** digunakan untuk memilih data stream yang masih belum ada apada tabel dimEmployees.

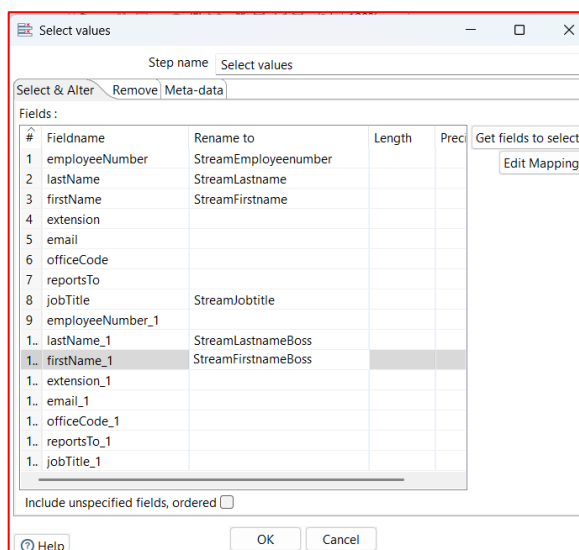
e) **Table output:** Memasukkan data ke dalam table dimEmployees

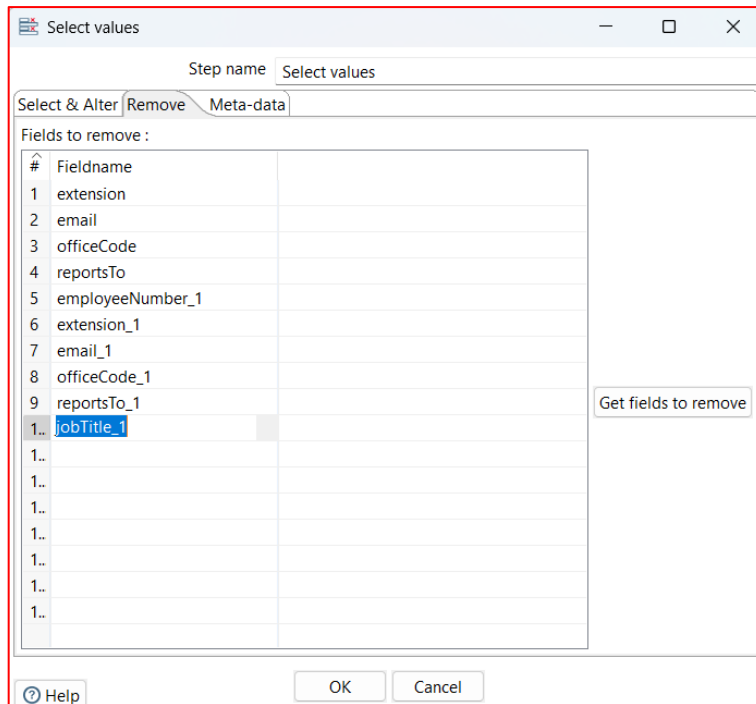
- Konfigurasi pada table input dengan menghubungkan **Connection** pada koneksi **conn_oltp_resources**. Untuk mengambil data sumber menggunakan query dibawah ini.

```
select * from employees e
left join employees r
on e.reportsTo=r.employeeNumber;
```

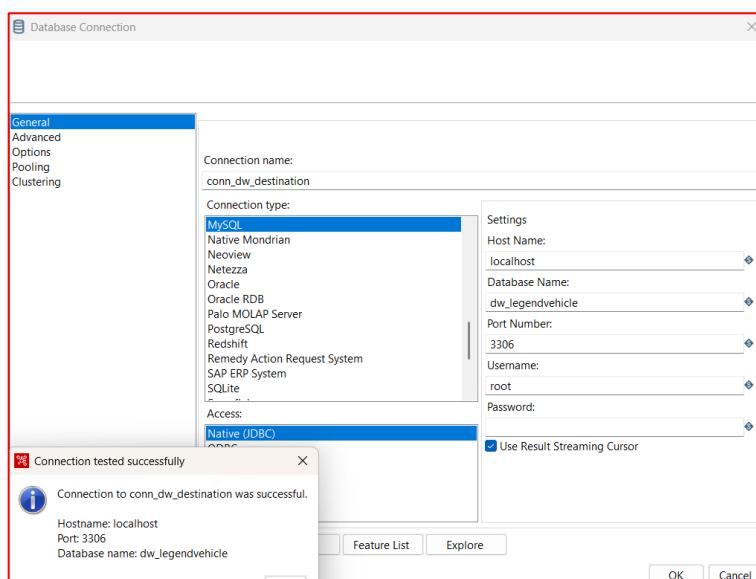


- Hubungkan output **table input** pada **select values**.
- Konfigurasi pada **Select values** yaitu mengambil data dari field **employeenumber**, **lastname**, **firstname**, **jobtitle**, **lastname_1** dan **firstname_1** sebagai **data stream** yang digunakan pada proses ETL pada tab **select & alter**.
- Hilangkan field lain yang tidak digunakan pada tab **remove**. (CONNECTION 2 KE DATABASE DW_LEGENDVEHICLE)





8. Hubungkan output **select values** pada **database lookup**
9. Konfigurasi pada database lookup adalah dengan menghubungkan koneksi pada **conn_dw_destination** dengan table lookup **dimEmployees** yang telah dibuat pada tahap pertama.
10. Field yang di lookup adalah field pada tabel **dimEmployees** dengan **field stream input** dari OLTP. sedangkan field yang di **retrieve** adalah field dari **dimEmployees** itu sendiri. Jika tidak ada data yang sama maka akan muncul null.



Database lookup

Step name: Database lookup

Connection: conn_dw_destination [Edit...] [New...] [Wizard...]

Lookup schema: dw_legendvehicle [Browse...]

Lookup table: dimemployees [Browse...]

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
2	firstName	=	StreamFirstname	
3	lastName	=	StreamLastname	
4	jobTitle	=	StreamJobtitle	
5	boss_firstName	=	StreamFirstnameBoss	
6	boss_lastName	=	StreamLastnameBoss	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	employeeNumber	employeeNumber		Integer
2	firstName	firstName		String
3	lastName	lastname		String
4	jobTitle	jobtitle		String
5	boss_firstName	boss_firstname		String
6	boss_lastName	boss_lastname		String

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

[Help] [OK] [Cancel] [Get Fields] [Get lookup fields]

11. Hubungkan **output database lookup** dengan **filter rows**.
12. Pada **filter rows** berikan kondisi field yang **null** pada field **dimemployees** untuk dimasukkan pada proses selanjutnya. Hal itu menandakan bahwa **data stream** belum memiliki kesamaan pada data di **dimemployees**.

Filter rows

Step name: Filter rows

Send 'true' data to step: Table output

Send 'false' data to step:

The condition:

☐ To edit a subcondition, simply click on it +

employeeNumber IS NULL

AND

firstName IS NULL

AND

lastname IS NULL

AND

jobtitle IS NULL

AND

boss_firstname IS NULL

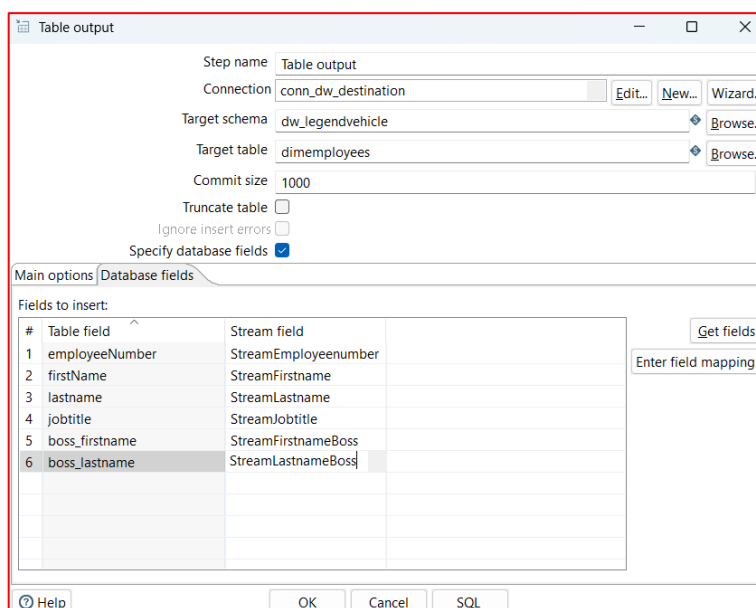
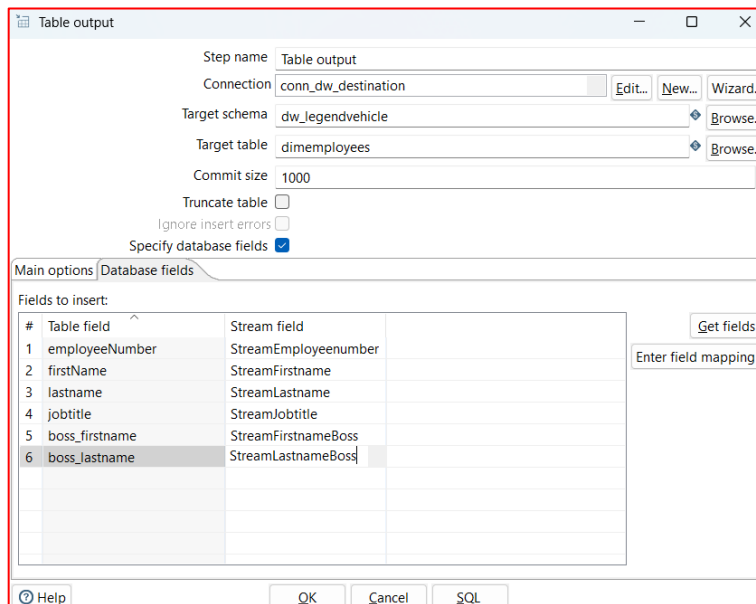
AND

boss_lastname IS NULL

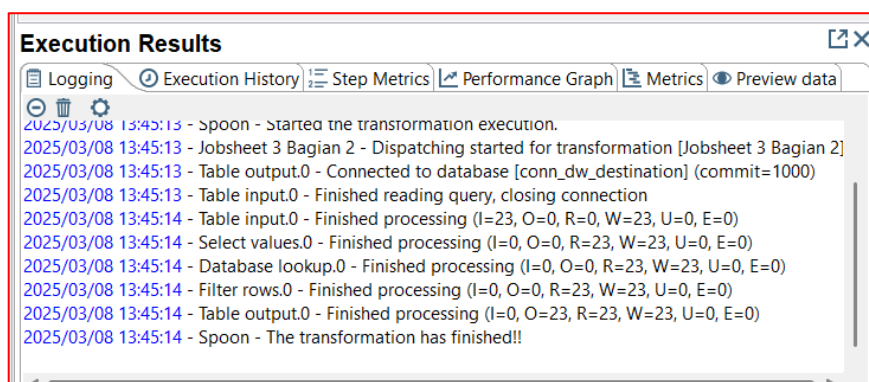
[Help] [OK] [Cancel]

13. Hubungkan output dari **filter rows** dengan **table output**.
14. Pada **table output**, gunakan **connection conn_dw_destination** untuk memasukkan data pada tabel **dimemployees**.

15. Aktifkan **specify database fields**, dan mapping **data stream input** dari oltp terhadap field yang ada pada **dimemployees**.



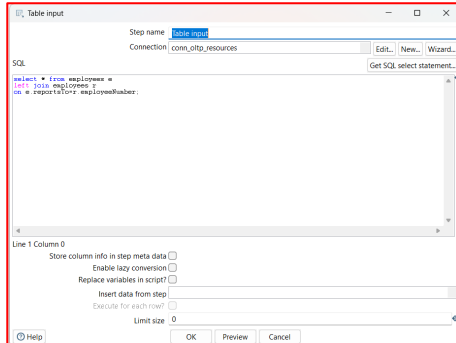
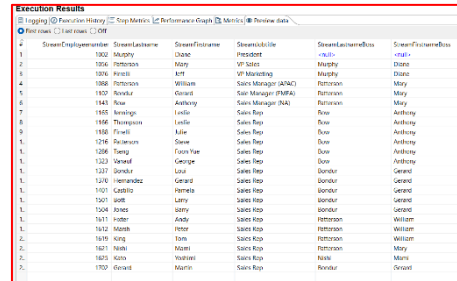
16. jika proses keseluruhan berhasil maka tabel **dimemployees** akan terisi data pegawai dari database OLTP.

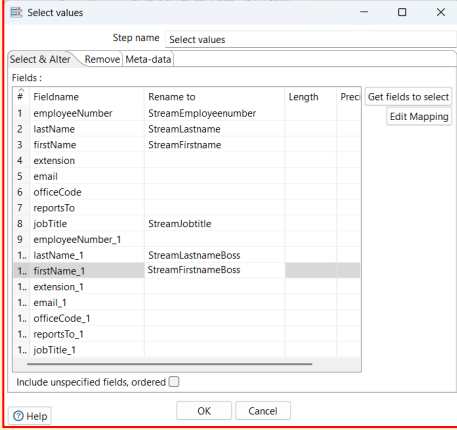
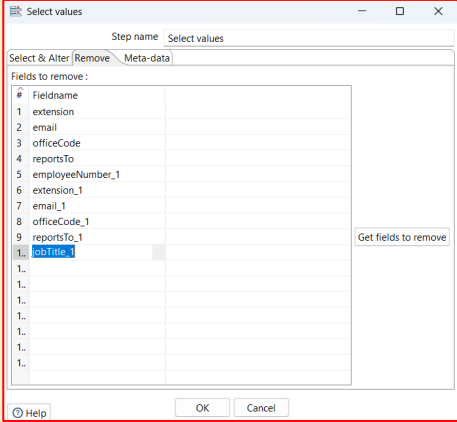
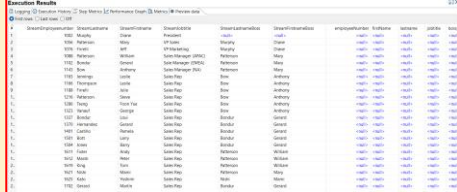


id_dimEmployees								
		employeeNumber	firstName	lastName	jobTitle	boss_firstName	boss_lastName	updated
<input type="checkbox"/>				1	1002 Diane Murphy	President	NULL	2025-03-08 13:45:14
<input type="checkbox"/>				2	1056 Mary Patterson	VP Sales	Diane Murphy	2025-03-08 13:45:14
<input type="checkbox"/>				3	1076 Jeff Firrelli	VP Marketing	Diane Murphy	2025-03-08 13:45:14
<input type="checkbox"/>				4	1088 William Patterson	Sales Manager (APAC)	Mary Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				5	1102 Gerard Bondur	Sale Manager (EMEA)	Mary Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				6	1143 Anthony Bow	Sales Manager (NA)	Mary Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				7	1165 Leslie Jennings	Sales Rep	Anthony Bow	2025-03-08 13:45:14
<input type="checkbox"/>				8	1166 Leslie Thompson	Sales Rep	Anthony Bow	2025-03-08 13:45:14
<input type="checkbox"/>				9	1188 Julie Firrelli	Sales Rep	Anthony Bow	2025-03-08 13:45:14
<input type="checkbox"/>				10	1216 Steve Patterson	Sales Rep	Anthony Bow	2025-03-08 13:45:14
<input type="checkbox"/>				11	1286 Foon Yue Tseng	Sales Rep	Anthony Bow	2025-03-08 13:45:14
<input type="checkbox"/>				12	1323 George Vanauf	Sales Rep	Anthony Bow	2025-03-08 13:45:14
<input type="checkbox"/>				13	1337 Loui Bondur	Sales Rep	Gerard Bondur	2025-03-08 13:45:14
<input type="checkbox"/>				14	1370 Gerard Hernandez	Sales Rep	Gerard Bondur	2025-03-08 13:45:14
<input type="checkbox"/>				15	1401 Pamela Castillo	Sales Rep	Gerard Bondur	2025-03-08 13:45:14
<input type="checkbox"/>				16	1501 Larry Bott	Sales Rep	Gerard Bondur	2025-03-08 13:45:14
<input type="checkbox"/>				17	1504 Barry Jones	Sales Rep	Gerard Bondur	2025-03-08 13:45:14
<input type="checkbox"/>				18	1611 Andy Fixter	Sales Rep	William Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				19	1612 Peter Marsh	Sales Rep	William Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				20	1619 Tom King	Sales Rep	William Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				21	1621 Mami Nishi	Sales Rep	Mary Patterson	2025-03-08 13:45:14
<input type="checkbox"/>				22	1625 Yoshimi Kato	Sales Rep	Mami Nishi	2025-03-08 13:45:14
<input type="checkbox"/>				23	1702 Martin Gerard	Sales Rep	Gerard Bondur	2025-03-08 13:45:14

TUGAS 2

- Buka preview tab pada execution result area di setiap proses object. amati input dan output data yang ada. bandingkan di setiap prosesnya. jelaskan perbedaan disetiap prosesnya.

No.	Proses Objek	SS Data Input	SS Data Output	Keterangan
1.	Table Input	Data mentah dari sumber OLTP (MySQL)	Data dari tabel sumber (belum diubah)	Mengambil data langsung dari database sumber
2.	Select Values	Data dari Table Input 	Data dengan perubahan nama kolom atau pemilihan kolom tertentu 	Memilih kolom yang dibutuhkan dan mengganti nama field
3.	Database Lookup	Data hasil dari Select Values	Data dengan tambahan informasi dari tabel lain	Melakukan lookup data dari tabel referensi (misalnya, mencari nama atasan dari reportsTo).

		 		
4.	Filter Rows	Data dari Database Lookup	Data yang sudah difilter (misalnya, hanya karyawan dengan job tertentu)	Memfilter data berdasarkan kondisi tertentu

[illegible]

Data hasil dari Filter Rows

Filter rows

Step name: Filter rows

Send 'true' data to step: Table output

Send 'false' data to step:

The condition:

☐ To edit a subcondition, simply click on it +

- employeeNumber IS NULL
- AND
- firstName IS NULL
- AND
- lastName IS NULL
- AND
- jobtitle IS NULL
- AND
- boss_firstname IS NULL
- AND
- boss_lastname IS NULL

[illegible][illegible][illegible]

Menyimpan hasil transformasi ke database dw_legendvehicle (tujuan).

2. Jika proses itu di ulangi (di run kembali) apakah data akan redudant?

Jawaban :

Ya, bisa redundant jika tidak ada mekanisme untuk menghindari duplikasi data.

Jika Table Output langsung menambahkan data tanpa validasi, maka setiap kali transformasi dijalankan, data yang sama akan tersimpan berulang kali.

Ini ada solusi untuk menghindari redundansi :

- Gunakan truncate sebelum insert jika hanya ingin data terbaru.
 - Terapkan primary key atau unique constraint di tabel tujuan. (biar tidak ada yang sama)
 - Gunakan "Insert/Update" daripada langsung "Table Output" untuk mencegah duplikasi.
3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?

Jawaban :

Pada database classicmodels saya menambahkan data diri saya pada tabel employees

```

1 row inserted. (Query took 0.3936 seconds.)

INSERT INTO employees (employeeNumber, lastName, firstName, extension, email, officeCode, reportsTo, jobTitle) VALUES (1202, 'Regita', 'Aqueena', 'x9001', 'aqueena@classicmodelcars.com', 1, 1002, 'Data Scientist');
[ Edit inline ] [ Edit ] [ Create PHP code ]

```

Hasil di database classicmodels:

Showing rows 0 - 0 (1 total, Query took 0.0003 seconds.)

```
SELECT * FROM employees WHERE lastName = 'Regita';
```

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

☐ Show all | Number of rows: 25 | Filter rows: Search this table

Extra options

	employeeNumber	lastName	firstName	extension	email	officeCode	reportsTo	jobTitle
<input type="checkbox"/> Edit Copy Delete	1202	Regita	Aqueena	x9001	aqueena@classicmodelcars.com	1	1002	Data Scientist

Setelah itu, lakukan migrasi dan check di database dw_legendvehicle.

Hasil di database dw_legendvehicle:

Server: localhost3306 > Database: dw_legendvehicle > Table: dimemployees

Showing rows 0 - 0 (1 total, Query took 0.0002 seconds.)

```
SELECT * FROM `dimemployees` WHERE lastName = 'Regita';
```

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

☐ Show all | Number of rows: 25 | Filter rows: Search this table

Extra options

	id_dimEmployees	employeeNumber	firstName	lastName	jobTitle	boss_firstName	boss_lastName	updated
<input type="checkbox"/> Edit Copy Delete	25	1202	Aqueena	Regita	Data Scientist	Diane	Murphy	2025-03-08 15:19:31

☐ Check all | With selected: Edit Copy Delete Export

- Table Input akan membaca data terbaru termasuk entri baru.
- Select Values dan Database Lookup akan memproses data termasuk nama baru (Aqueena).
- Jika Filter Rows tidak memfilter data bernama Aqueena, maka nama Aqueena akan ikut masuk ke Table Output.
- Jika Table Output hanya menambah data tanpa validasi, maka nama Aqueena akan tersimpan di data warehouse.