

**LAPORAN PRAKTIKUM
DATA WAREHOUSE**

**JOBSHEET 3
DATABASE ANALYTICAL**



Oleh:

DHEVINA AGUSTINA

2341760065

SIB 2B

PROGRAM STUDI SISTEM INFORMASI BISNIS

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

2025/2026

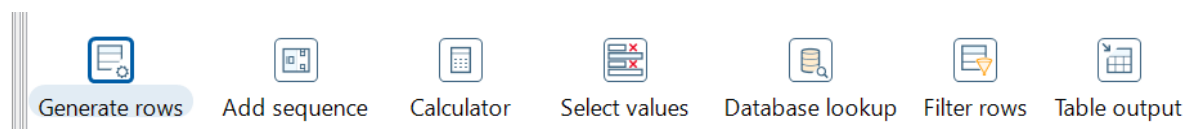
A. Dimensi Waktu

1. Buatlah sebuah database yang digunakan sebagai OLAP dengan nama dw_LegendVehicle.
CREATE DATABASES dw_LegendVehicle;
use dw_LegendVehicle;
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
);

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.

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

1. Buka PDI Spoon. Buat Transformation baru -> File - New - Transformation.
2. Drag and Drop beberapa objek yaitu:
 - Generate Rows: digunakan untuk membuat baris data baru.
 - Add Sequence: digunakan untuk membuat sequence, dalam hal ini membuat data di setiap harinya.
 - Calculator: digunakan untuk menjumlahkan hari dan mengambil data tahun, bulan dan hari.
 - Select Values: digunakan untuk memilih field yang digunakan.
 - 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.
 - Filter Rows: digunakan untuk mengambil data yang belum ada pada table dimDate setelah dicek sebelumnya.
 - Table Output: digunakan untuk menyimpan data pada tabel tujuan (dimDate).



3. 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).
4. Membuat fields baru bernama CurrentDate dengan type data Date dan format dd-MM-yyyy serta value awal 01-01-2003.

Step name: **Generate rows**

Limit: 1825

Never stop generating rows: ☐

Interval in ms (delay): 5000

Current row time field name: now

Previous row time field name: FiveSecondsAgo

Fields:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set
1	CurrentDate	Date	dd-MM-yyyy						01-01-2003	N

Buttons: Help, OK, Preview, Cancel

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

Step name: Add sequence

Name of value: incrementDay

Use a database to generate the sequence

Use DB to get sequence? ☐

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

Schema name: [dropdown] Schemas...

Sequence name: SEQ_ Sequences...

Use a transformation counter to generate the sequence

Use counter to calculate sequence? ☒

Counter name (optional): [text box]

Start at value: 0

Increment by: 1

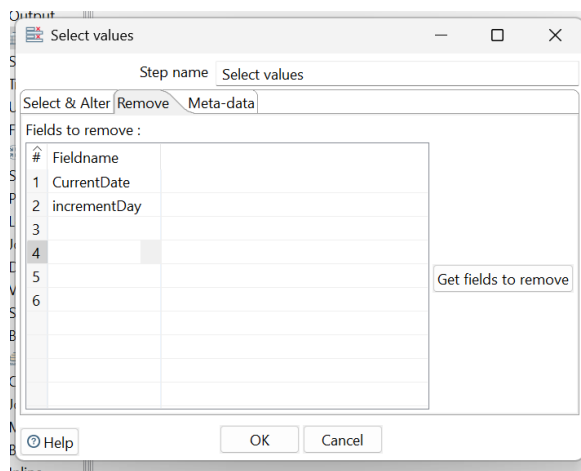
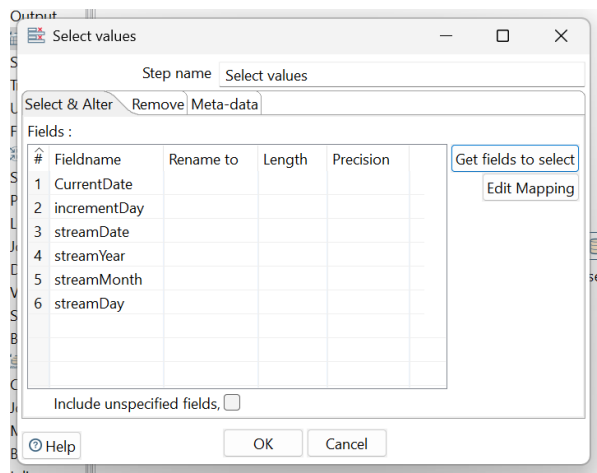
Maximum value: 999999999

Buttons: Help, OK, Cancel

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

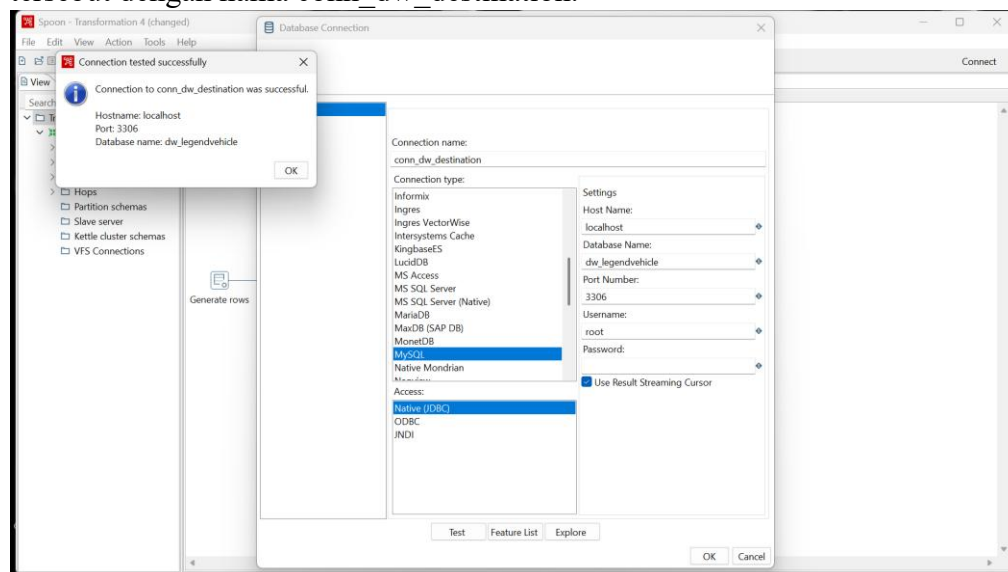
Step name												
Calculator												
<input checked="" type="checkbox"/> Throw an error on non existing files												
Fields:												
#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal symbol	Grouping symbol
1	streamDate	Date A + B D...	CurrentDate	IncrementDay		None			N			
2	streamYear	Year of date A	streamDate			None			N			
3	streamMonth	Month of dat...	streamDate			None			N			
4	streamDay	Day of mont...	streamDate			None			N			

9. Hubungkan output dari **calculator** menuju **Select values**
10. 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.
11. Dikarenakan tidak semua fields digunakan, maka pada tab **Remove** diisikan fields **CurrentDate** dan **incrementDay** dikarenakan kedua fields tersebut tidak digunakan.



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.



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.
15. Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak kembar adalah:
- field date pada table dimdate dengan field streamDate
 - field year pada table dimdate dengan field streamYear
 - field month pada table dimdate dengan field streamMonth
 - field day pada table dimdate dengan field streamDay
16. Field yang akan di **retrive** adalah field yang ada pada table **dimDate** yaitu **date, year, month, dan day**.

Database lookup

Step name: Database lookup

Connection: conn_dw_destination

Lookup schema: dw_legendvehicle

Lookup table: dimdate

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	=	streamDate	
2	year	=	streamYear	
3	month	=	streamMonth	
4	day	=	streamDay	

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

17. Hubungkan output dari **database lookup** dengan **filter rows**
18. 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

(

AND

(

AND

AND

year IS NULL

month IS NULL

day IS NULL

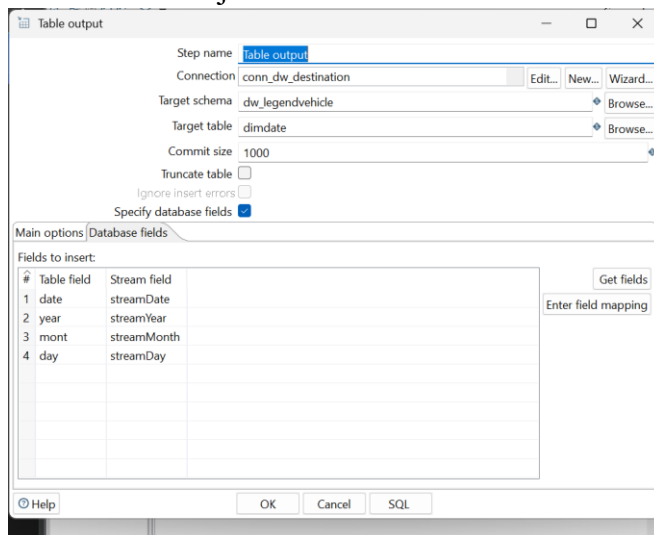
)

)

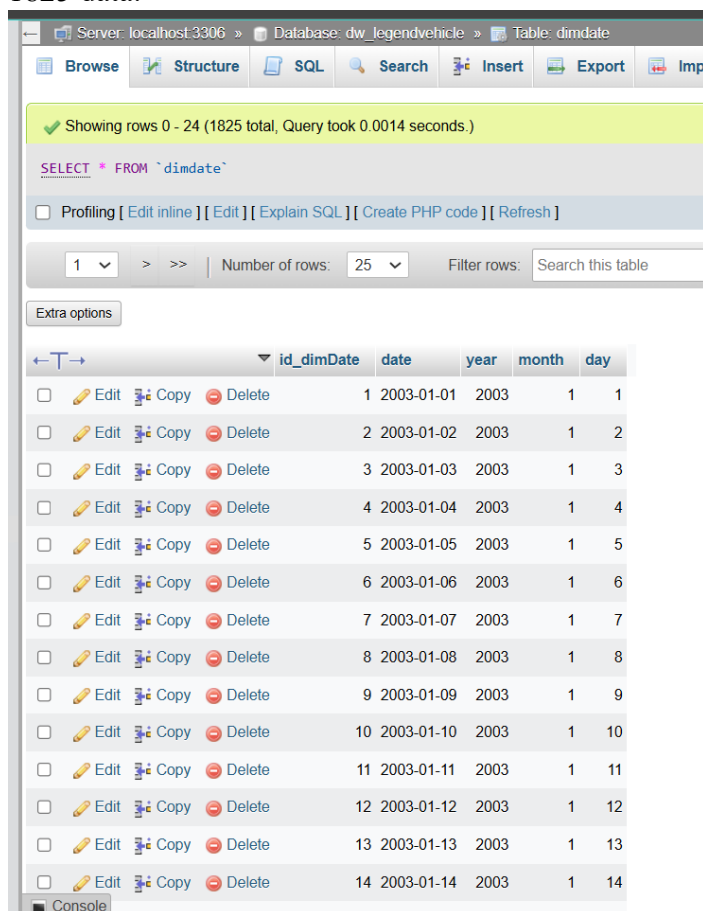
Help OK Cancel

19. Hubungkan output dari filter rows menuju table output.

20. Konfigurasi pada table output adalah memberikan koneksi pada conn_dw_destination dengan schema dw_legendvehicle dan table dimdate.
21. Aktifkan specify database fields.
22. 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.

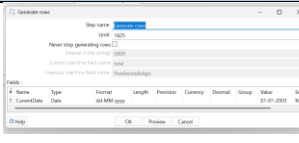
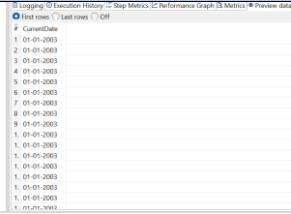
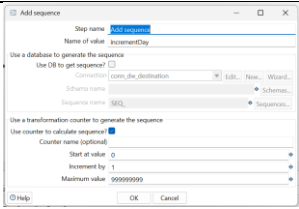
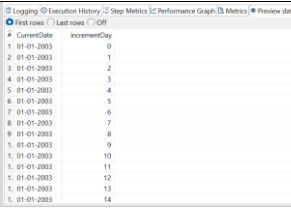
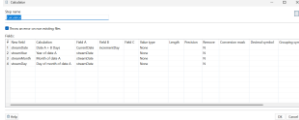
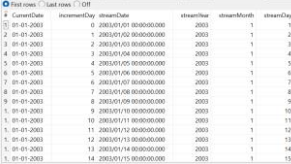


23. cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.



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.

PROSES OBJEK	SS DATA INPUT	SS DATA OUTPUT	KET.
Generate rows			<ul style="list-style-type: none"> - Step ini digunakan untuk menghasilkan data secara otomatis tanpa input dari sumber lain. - Pada contoh ini, 1000 baris data dibuat dengan nilai tetap "01-01-2003" pada kolom CurrentDate. - Hasil output menunjukkan 1000 baris dengan tanggal yang sama.
Add Sequences			<ul style="list-style-type: none"> - Step ini digunakan untuk menambahkan nomor urut (sequence) ke setiap baris data yang dihasilkan sebelumnya. - Field baru IncrementDay ditambahkan dengan nilai yang bertambah mulai dari 0, 1, hingga jumlah baris data yang ada. - Hasil output menunjukkan bahwa setiap baris memiliki tanggal CurrentDate yang sama, tetapi memiliki sequence yang bertambah dari 1, 2, 3, ... hingga baris terakhir.
Calculator			<ul style="list-style-type: none"> - Step ini digunakan untuk melakukan operasi matematika pada data. - Pada contoh ini,

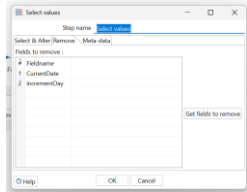
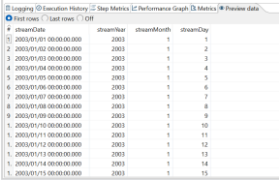
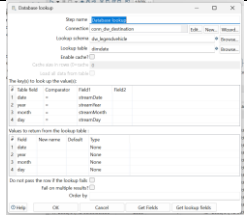

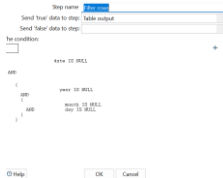
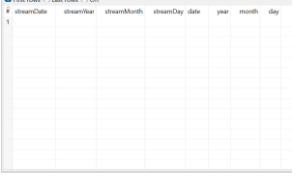
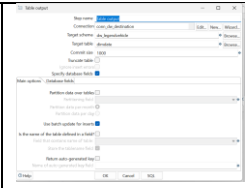
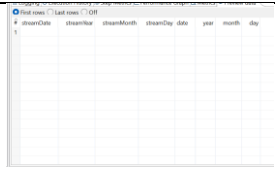
			<p>dilakukan perhitungan $\text{streamDate} = \text{CurrentDate} + \text{incrementDay}$, sehingga menghasilkan tanggal yang bertambah setiap barisnya.</p> <ul style="list-style-type: none"> - Selain itu, data dipecah menjadi beberapa bagian: streamYear, streamMonth, streamDay.
Select values			<ul style="list-style-type: none"> - Step ini digunakan untuk memilih kolom yang akan dipertahankan dalam proses selanjutnya. - Kolom yang tidak dibutuhkan akan dibuang agar proses lebih efisien.
Database lookup			<ul style="list-style-type: none"> - Step ini digunakan untuk mencocokkan data dengan tabel referensi dari database. - Data yang telah dihitung (misalnya, streamDate) dibandingkan dengan data dari database untuk mendapatkan informasi tambahan. - Contohnya, hasil output menunjukkan data yang telah diperiksa terhadap tabel referensi dan menambahkan informasi baru seperti date, year, month, day.
Filter rows			<p>Step ini digunakan untuk menyaring data berdasarkan kondisi tertentu.</p>

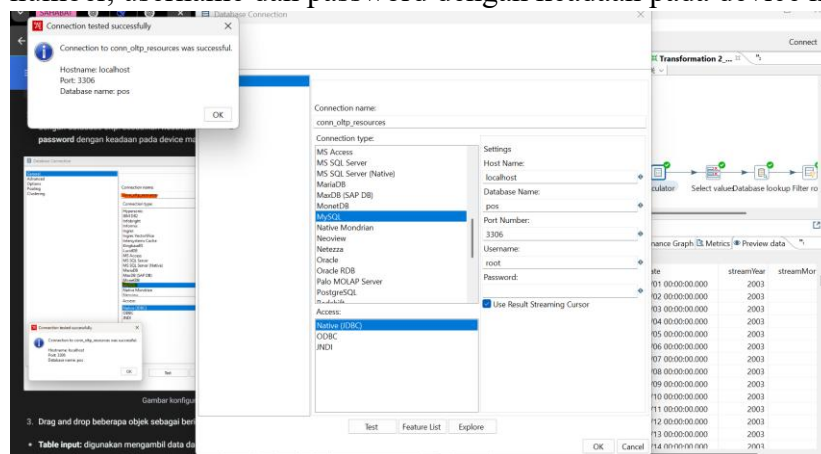
Table Output			Step ini digunakan untuk menyimpan data hasil transformasi ke dalam database atau file output.
--------------	---	--	--

B. Dimensi Pegawai

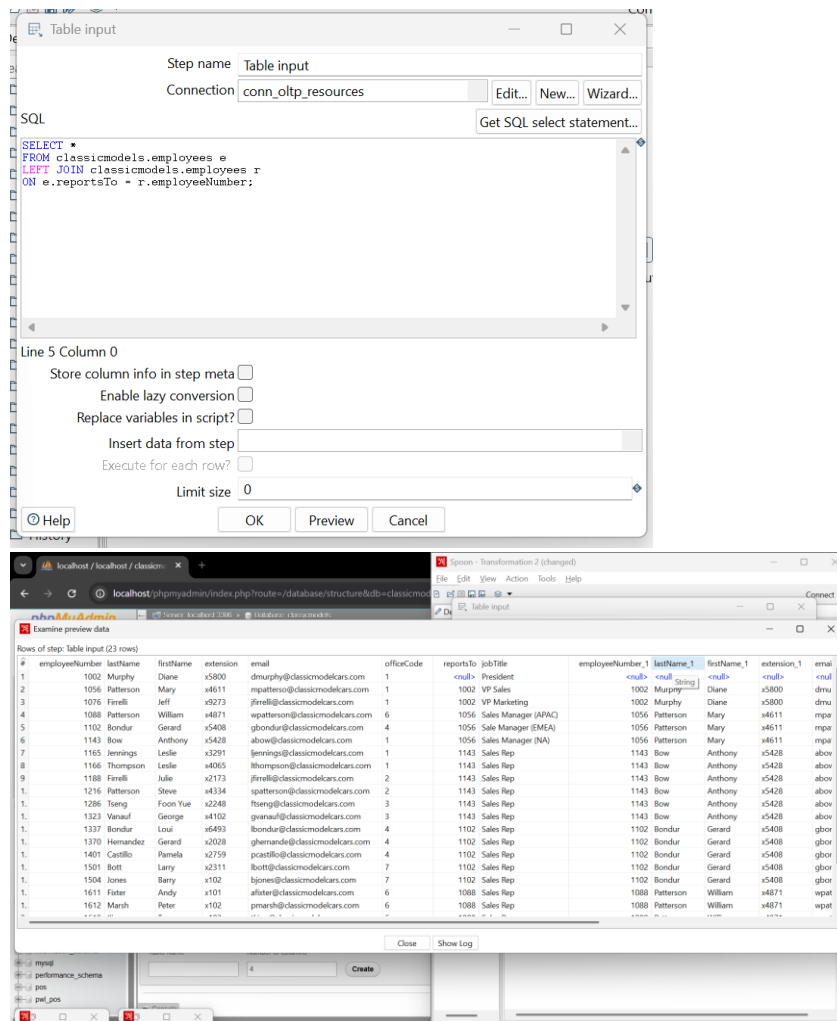
1. Buatlah tabel dimPegawai pada dw_legendVehicle.

```
CREATE TABLE dimEmployees (
  id_dimEmployees INT NOT NULL AUTO_INCREMENT,
  employeeNumber INT,
  firstName VARCHAR(50),
  lastName VARCHAR(50),
  jobTitle VARCHAR(50),
  boss_firstName VARCHAR(50),
  boss_lastName VARCHAR(50),
  updated DATETIME DEFAULT CURRENT_TIMESTAMP,
  PRIMARY KEY (id_dimEmployees)
);
```

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.



3. Drag and drop beberapa objek sebagai berikut:
 - Table input: digunakan mengambil data dari database OLTP.
 - Select values: memilih field yang digunakan untuk proses Transform dan Load.
 - Database lookup: digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar
 - Filter rows: digunakan untuk memilih data stream yang masih belum ada apada tabel dimEmployees.
 - Table output: Memasukkan data ke dalam tavle dimEmployees
4. Konfigurasi pada table input dengan menghubungkan **Connection** pada koneksi **conn_oltp_resources**. Untuk mengambil data sumber menggunakan query dibawah ini.



- 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.

Select & Alter Remove Meta-data

#	Fieldname	Rename to	Length	Precision
1	employeeNumber	StreamEmployeeNumber		
2	lastName	StreamLastName		
3	firstName	StreamFirstName		
4	extension			
5	email			
6	officeCode			
7	reportsTo			
8	jobTitle	StreamJobTitle		
9	employeeNumber_1			
10	lastName_1	StreamLastNameBoss		
11	firstName_1	StreamFirstNameBoss		
12	extension_1			
13	email_1			
14	officeCode_1			
15	reportsTo_1			
16	jobTitle_1			

Get fields to select
Edit Mapping

- Hilangkan field lain yang tidak digunakan pada tab remove.

Select & Alter Remove Meta-data

Fields to remove :

#	Fieldname
1	extension
2	email
3	officeCode
4	reportsTo
5	employeeNumber_1
6	extension_1
7	email_1
8	officeCode_1
9	reportsTo_1
10	jobTitle_1

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

Lookup schema: dw_legendvehicle

Lookup table: dimemployees

Enable cache? ☐

Cache size in rows (0=cache every row): 0

Load all data from table ☐

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

#	Table field	Comparator	Field1	Field2
3	lastName	=	Stream/lastName	
4	jobTitle	=	Stream/jobTitle	
5	boss_firstName	=	Stream/firstNameBoss	
6	boss_lastName	=	Stream/lastNameBoss	

Values to return from the lookup table:

#	Field	New name	Default	Type
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:

OK Cancel Get Fields Get lookup fields

11. Hubungkan output databse 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:

☐ employeeNumber IS NOT NULL

AND

☐ firstName IS NOT NULL

AND

☐ lastName IS NOT NULL

AND

☐ jobTitle IS NOT NULL

AND

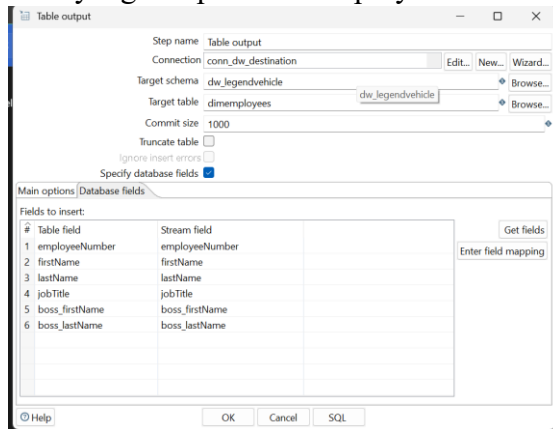
☐ boss_firstName IS NOT NULL

AND

☐ boss_lastName IS NOT NULL

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.



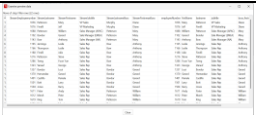
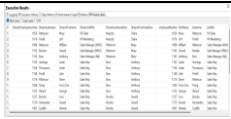
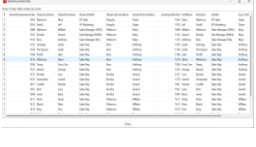

16. jika proses keseluruhan berhasil maka tabel **dimemployees** akan terisi data pegawai dari database OLTP.

	id_dimEmployees	employeeNumber	firstName	lastName	jobTitle	boss_firstName	boss_lastName	updated
<input type="checkbox"/>		1	1002 Diane	Murphy	President	NULL	NULL	2025-03-09 01:25:35
<input type="checkbox"/>		2	1056 Mary	Patterson	VP Sales	Diane	Murphy	2025-03-09 01:25:35
<input type="checkbox"/>		3	1076 Jeff	Firrelli	VP Marketing	Diane	Murphy	2025-03-09 01:25:35
<input type="checkbox"/>		4	1088 William	Patterson	Sales Manager (APAC)	Mary	Patterson	2025-03-09 01:25:35
<input type="checkbox"/>		5	1102 Gerard	Bondur	Sale Manager (EMEA)	Mary	Patterson	2025-03-09 01:25:35
<input type="checkbox"/>		6	1143 Anthony	Bow	Sales Manager (NA)	Mary	Patterson	2025-03-09 01:25:35
<input type="checkbox"/>		7	1165 Leslie	Jennings	Sales Rep	Anthony	Bow	2025-03-09 01:25:35
<input type="checkbox"/>		8	1166 Leslie	Thompson	Sales Rep	Anthony	Bow	2025-03-09 01:25:35
<input type="checkbox"/>		9	1188 Julie	Firrelli	Sales Rep	Anthony	Bow	2025-03-09 01:25:35
<input type="checkbox"/>		10	1216 Steve	Patterson	Sales Rep	Anthony	Bow	2025-03-09 01:25:35
<input type="checkbox"/>		11	1286 Foon Yue	Tsang	Sales Rep	Anthony	Bow	2025-03-09 01:25:35
<input type="checkbox"/>		12	1323 George	Vanauf	Sales Rep	Anthony	Bow	2025-03-09 01:25:35
<input type="checkbox"/>		13	1337 Loui	Bondur	Sales Rep	Gerard	Bondur	2025-03-09 01:25:35
<input type="checkbox"/>		14	1370 Gerard	Hernandez	Sales Rep	Gerard	Bondur	2025-03-09 01:25:35

TUGAS 2

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.

PROSES OBJEK	SS DATA INPUT	SS DATA OUTPUT	KET.
Table Input			Hasil output adalah seluruh dataset awal sebelum transformasi.
Select values			Input berupa semua kolom dari Table Input, sedangkan output hanya kolom yang dipilih
Database lookup			Menambahkan data dari tabel lain berdasarkan key,

			output akan memiliki kolom tambahan dari tabel lain
Filter Rows			Input memiliki semua data sebelum filtering, sedangkan output hanya menyisakan data yang memenuhi kondisi.
Table output			Data yang masuk ke sini adalah hasil akhir setelah semua proses sebelumnya.

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

Jawaban: Jika proses ini diulang (di-run kembali), data tidak akan redundant. Hal ini karena ada mekanisme yang mencegah duplikasi

- Database Lookup: Mengecek apakah data sudah ada sebelum dimasukkan.
- Filter Rows: Menyaring data yang akan diproses agar tidak terjadi duplikasi.

3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?

Jawaban: Data karyawan baru dengan nama saya berhasil ditambahkan pada table employees di OLTP dengan memastikan employeeNumber unik untuk menghindari duplikasi.

Run SQL query/queries on table classicmodels.employees:

```

1 INSERT INTO employees (employeeNumber, lastName, firstName, extension, email, officeCode, reportsTo, jobTitle)
2 VALUES (2508, 'Agustina', 'Dhevina', 'x8004', 'dhevina@gmail.com', 1, NULL, 'Business Analyst');
3

```

✓ 1 row inserted. (Query took 0.0227 seconds)

```

INSERT INTO employees (employeeNumber, lastName, firstName, extension, email, officeCode, reportsTo, jobTitle) VALUES (2508, 'Agustina', 'Dhevina', 'x8004', 'dhevina@gmail.com', 1, NULL, 'Business Analyst');

```

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

employeeNumber	lastName	firstName	extension	email	officeCode	reportsTo	jobTitle
2508	Agustina	Dhevina	x8004	dhevina@gmail.com	1	NULL	Business Analyst

✓ 1 row inserted. (Query took 0.0045 seconds)

```

INSERT INTO employees (employeeNumber, lastName, firstName, extension, email, officeCode, reportsTo, jobTitle) VALUES (1400, 'Agustina', 'Dhevina', 'x7001', 'dhevinagustina@classicmodelcars.com', 1, 1002, 'Sales Rep');

```

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