

Laporan Praktikum Data Warehouse

Jobsheet 3 : Database Analytical

Dosen Pengampu : Vipkas Al Hadid Firdaus, ST., MT



Disusun Oleh:

Queenadhynar Azarine Dwipa A.

2341760109

SIB 2B

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

2023/2024

A. Dimensi Waktu

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



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

Table	Action	Rows	Type	Collation	Size	Overhead
dimdate	Browse Structure Search Insert Empty Drop	0	InnoDB	utf8mb4_0900_ai_ci	16.0 KiB	-
1 table	Sum	0	InnoDB	utf8mb4_0900_ai_ci	16.0 KiB	0 B

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 empty string
1	CurrentDate	Date	dd:MM:...						01-01-2003	<input checked="" type="checkbox"/>

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]

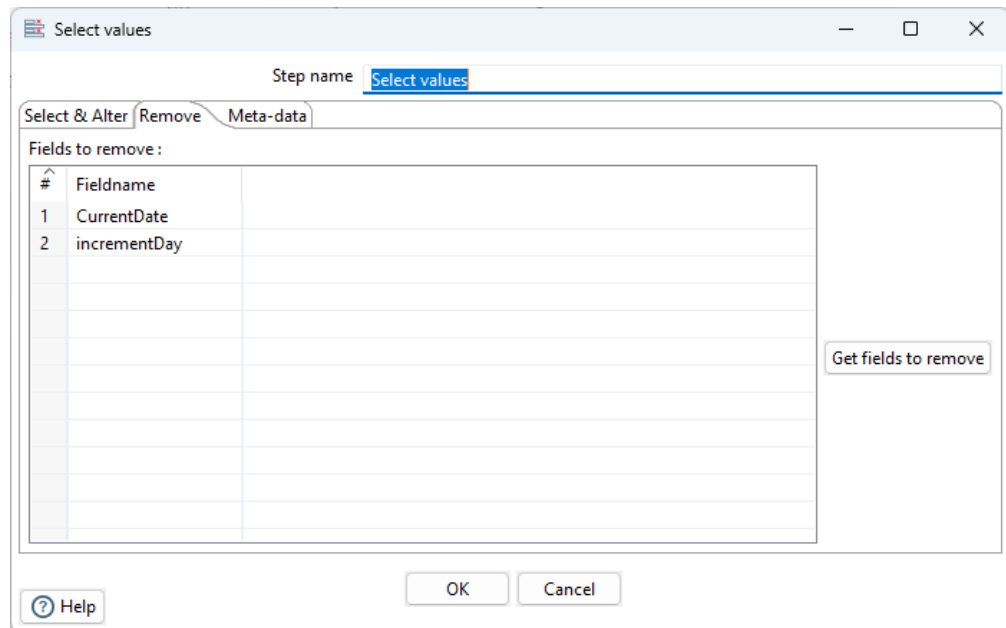
Start at value: 0

Increment by: 1

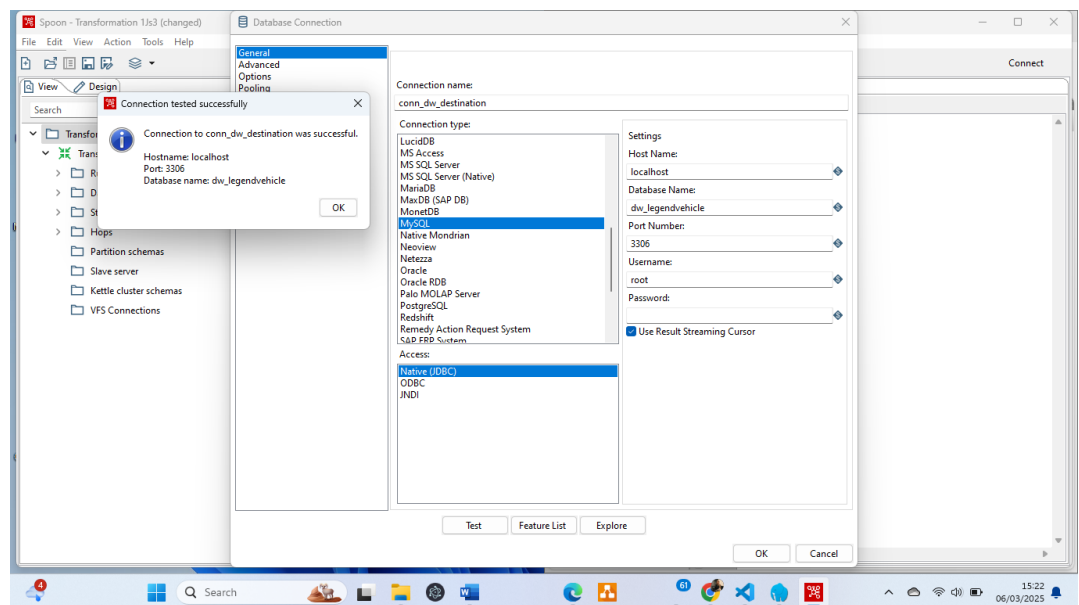
Maximum value: 999999999

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



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 [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	=	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 ☐

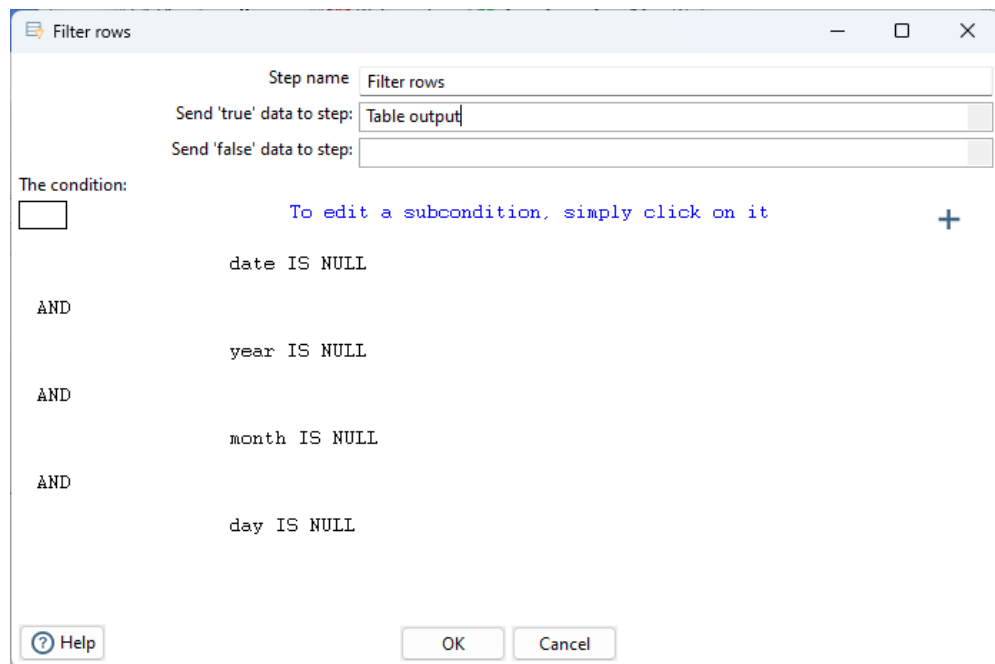
Fail on multiple results? ☐

Order by:

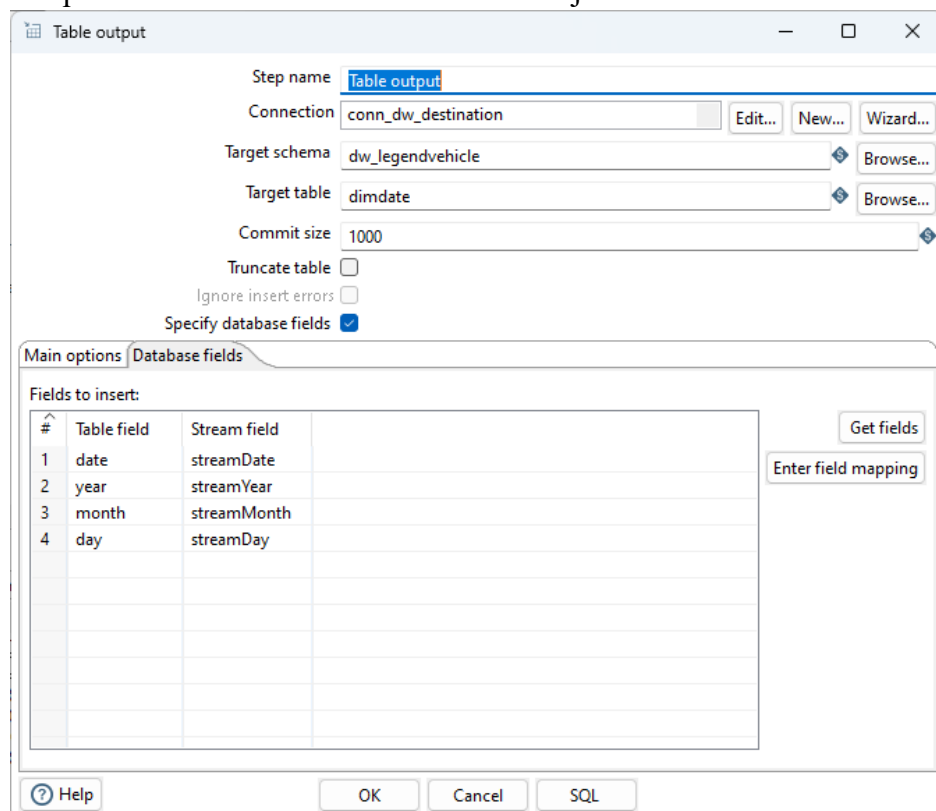
[?] Help [OK] [Cancel] [Get Fields] [Get lookup fields]

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)



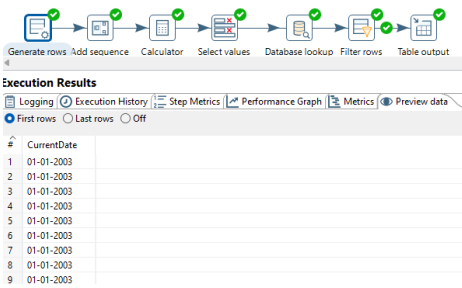
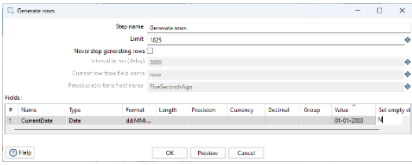
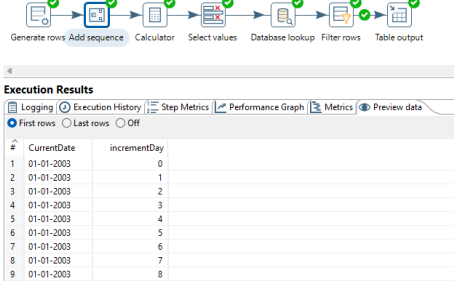
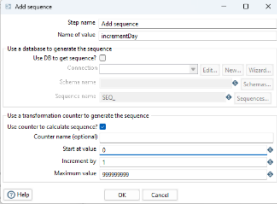
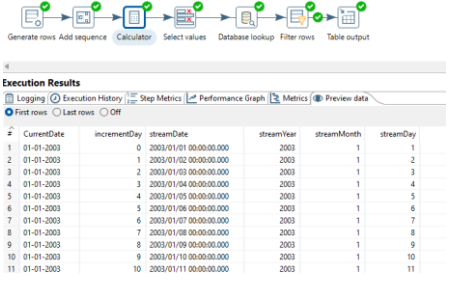
19. Hubungkan output dari filter rows menuju table output.
20. Konfigurasi pada table output adalah memberikan koneksi pada conn_dw_destination dengan schema dw_legendvehicle dan table dimdate.
21. Aktifkan specify database fields.
22. Pada tab Database fields, mapping data input streamDate, streamYear,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

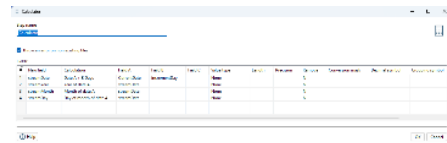
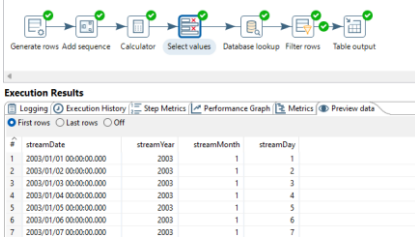


23. cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.

<div> <div>← T →</div> <div>id_dimDate</div> <div>date</div> <div>year</div> <div>month</div> <div>day</div> </div>									
<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	

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		Keterangan
Generate rows	Tidak ada input, data dibuat dari 0		<p>Output : beberapa baris dengan kolom CurrentDate, semua berisi 01-01-2003</p> <p>Proses ini menghasilkan sejumlah baris data dengan nilai tetap pada kolom CurrentDate</p>
Add Sequences			<p>Input : data dari Generate Rows</p> <p>Output : Data yang sama dengan tambahan kolom incrementDay, yang berisi angka urut mulai dari 0, 1, 2, ..</p> <p>Menambahkan kolom incrementDay sebagai penanda urutan setiap baris data.</p>
Calculator			<p>Input : Data dari Add Sequence:</p> <p>Output : Data yang sama dengan tambahan beberapa kolom:</p> <ul style="list-style-type: none"> - streamDate: Format timestamp dari CurrentDate ditambah incrementDay

			<div>- streamYear: Tahun dari streamDate</div> <div>- streamMonth: Bulan dari streamDate</div> <div>- streamDay: Hari dari streamDate</div> <div>Proses ini melakukan perhitungan tanggal berdasarkan incrementDay, kemudian memecahnya menjadi tahun, bulan, dan hari.</div>																																								
<div>Select values</div>		 <table><tr><th>#</th><th>streamDate</th><th>streamYear</th><th>streamMonth</th><th>streamDay</th></tr><tr><td>1</td><td>2003/01/01 00:00:00.000</td><td>2003</td><td>1</td><td>1</td></tr><tr><td>2</td><td>2003/01/02 00:00:00.000</td><td>2003</td><td>1</td><td>2</td></tr><tr><td>3</td><td>2003/01/03 00:00:00.000</td><td>2003</td><td>1</td><td>3</td></tr><tr><td>4</td><td>2003/01/04 00:00:00.000</td><td>2003</td><td>1</td><td>4</td></tr><tr><td>5</td><td>2003/01/05 00:00:00.000</td><td>2003</td><td>1</td><td>5</td></tr><tr><td>6</td><td>2003/01/06 00:00:00.000</td><td>2003</td><td>1</td><td>6</td></tr><tr><td>7</td><td>2003/01/07 00:00:00.000</td><td>2003</td><td>1</td><td>7</td></tr></table>	#	streamDate	streamYear	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	<div>Input : Data dari Calculator</div> <div>Output : Data yang sudah diseleksi hanya menyisakan: - streamDate - streamYear - streamMonth - streamDay</div> <div>Proses ini menyaring kolom yang tidak diperlukan (CurrentDate, incrementDay).</div>
#	streamDate	streamYear	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																																							

Database lookup

	streamDate	streamYear	streamMonth	streamDay	date	year	month	day
1	2003/01/01 00:00:00.000	2003	1	1	2003/01/01 00:00:00.000	2003	1	1
2	2003/01/02 00:00:00.000	2003	1	2	2003/01/02 00:00:00.000	2003	1	2
3	2003/01/03 00:00:00.000	2003	1	3	2003/01/03 00:00:00.000	2003	1	3
4	2003/01/04 00:00:00.000	2003	1	4	2003/01/04 00:00:00.000	2003	1	4
5	2003/01/05 00:00:00.000	2003	1	5	2003/01/05 00:00:00.000	2003	1	5
6	2003/01/06 00:00:00.000	2003	1	6	2003/01/06 00:00:00.000	2003	1	6
7	2003/01/07 00:00:00.000	2003	1	7	2003/01/07 00:00:00.000	2003	1	7
8	2003/01/08 00:00:00.000	2003	1	8	2003/01/08 00:00:00.000	2003	1	8
9	2003/01/09 00:00:00.000	2003	1	9	2003/01/09 00:00:00.000	2003	1	9
10	2003/01/10 00:00:00.000	2003	1	10	2003/01/10 00:00:00.000	2003	1	10

Input :
Data dari **Select Values**

Output :
Data yang sama dengan tambahan kolom hasil lookup:
- date
- year
- month
- day

Proses ini mengambil data tambahan dari database berdasarkan streamDate, menghasilkan informasi yang lebih lengkap.

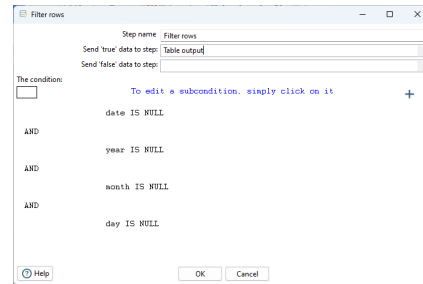
Filter rows

Input :
Data dari **Database Lookup**

Output :
Data setelah difilter berdasarkan kondisi tertentu.

Proses ini digunakan untuk menyaring baris data yang memenuhi kriteria tertentu sebelum disimpan ke tabel output.

Table Output



Input :
Data dari **Filter Rows**

Output :
Data yang disimpan dalam
database atau tabel output

Proses ini menyimpan hasil
akhir ke dalam database atau
file output untuk digunakan
lebih lanjut.

B. Dimesi Pegawai

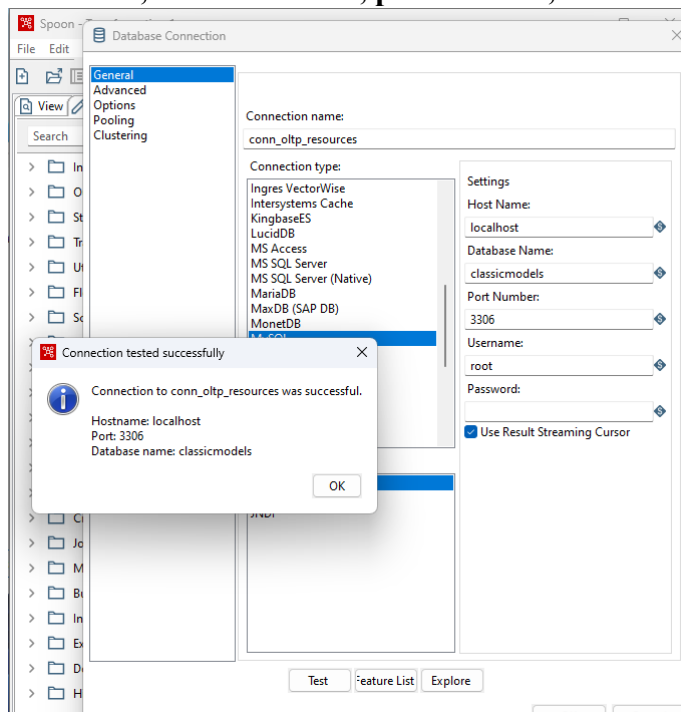
1. Buatlah tabel **dimPegawai** pada **dw_legendVehicle**.

```
MySQL returned an empty result set (i.e. zero rows). (Query took 0.1718 seconds.)

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 TIMESTAMP DEFAULT CURRENT_TIMESTAMP, PRIMARY KEY (id_dimEmployees) );

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

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.



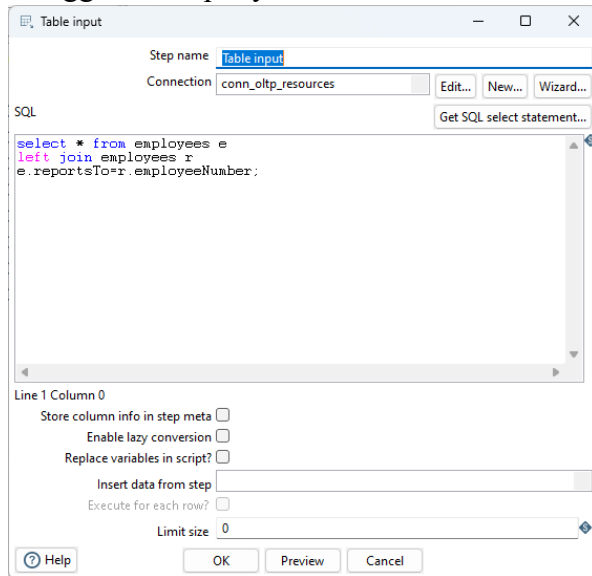
3. Drag and drop beberapa objek sebagai berikut:

- 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 table dimEmployees



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



5. Hubungkan output table input pada select values.
6. Konfigurasi pada Select values yaitu mengambil data dari field employeeenumber, lastname, firstname, jobtitle , lastname_1 dan firstname_1 sebagai data stream yang digunakan pada proses ETL pada tab select & alter.
7. Hilangkan field lain yang tidak digunakan pada tab remove.

Select values

Step nameSelect values

Select & AlterRemoveMeta-data

Fields :

#	Fieldname	Rename to	Length	Precision
1	employeeNumber	StreamEmployeeenumber		
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 selectEdit Mapping

Include unspecified fields, ordered by name☐

Help

OKCancel

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
1	employeenumber	=	StreamEmployeenumber	
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.

Table output

Step name

Table output

Connection

conn_dw_destination

Edit...

New...

Wizard...

Target schema

dw_legendvehicle

Browse...

Target table

dimemployees

Browse...

Commit size

1000

Truncate table

☐

Ignore insert errors

☐

Specify database fields

☒

Main options

Database fields

Fields to insert:

#	Table field	Stream field
1	employeeNumber	StreamEmployeeNumber
2	firstName	StreamFirstname
3	lastName	StreamLastname
4	jobtitle	StreamJobtitle
5	boss_firstname	StreamFirstnameBoss
6	boss_lastname	StreamLastnameBoss

Get fields

Enter field mapping

Help

OK

Cancel

SQL

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

✓ Showing rows 0 - 22 (23 total, Query took 0.0012 seconds.)

`SELECT * FROM `employees``

☐ Profiling [\[Edit inline \]](#) [\[Edit \]](#) [\[Explain SQL \]](#) [\[Create PHP code \]](#) [\[Refresh \]](#)

☐ Show all | Number of rows: 25 ▼ | Filter rows: | Sort by key: None ▼

Extra options

		employeeNumber	lastName	firstName	extension	email	officeCode	reportsTo	jobTitle
<input type="checkbox"/>	Edit Copy Delete	1002	Murphy	Diane	x5800	dmurphy@classicmodelcars.com	1	NULL	President
<input type="checkbox"/>	Edit Copy Delete	1056	Patterson	Mary	x4611	mpatterso@classicmodelcars.com	1	1002	VP Sales
<input type="checkbox"/>	Edit Copy Delete	1076	Firrelli	Jeff	x9273	jfirrelli@classicmodelcars.com	1	1002	VP Marketing
<input type="checkbox"/>	Edit Copy Delete	1088	Patterson	William	x4871	wpatterson@classicmodelcars.com	6	1056	Sales Manager (APAC)
<input type="checkbox"/>	Edit Copy Delete	1102	Bondur	Gerard	x5408	gbondur@classicmodelcars.com	4	1056	Sale Manager (EMEA)
<input type="checkbox"/>	Edit Copy Delete	1143	Bow	Anthony	x5428	abow@classicmodelcars.com	1	1056	Sales Manager (NA)
<input type="checkbox"/>	Edit Copy Delete	1165	Jennings	Leslie	x3291	ljennings@classicmodelcars.com	1	1143	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1166	Thompson	Leslie	x4065	lthompson@classicmodelcars.com	1	1143	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1188	Firrelli	Julie	x2173	jfirrelli@classicmodelcars.com	2	1143	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1216	Patterson	Steve	x4334	spatterson@classicmodelcars.com	2	1143	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1286	Tseng	Foon Yue	x2248	ftseng@classicmodelcars.com	3	1143	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1323	Vanauf	George	x4102	gvanauf@classicmodelcars.com	3	1143	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1337	Bondur	Loui	x6493	lbondur@classicmodelcars.com	4	1102	Sales Rep
<input type="checkbox"/>	Edit Copy Delete	1370	Hernandez	Gerard	x2028	ghernande@classicmodelcars.com	4	1102	Sales Rep

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	Keterangan																																																																																																				
Table Input	Data mentah dari sumber OLTP	<div>Execution Results</div> <div><div>Logging</div><div>Execution History</div><div>Step Metrics</div><div>Performance Graph</div><div>Metrics</div><div>Preview data</div></div> <div><div>First rows</div><div>Last rows</div><div>Off</div></div> <table><thead><tr><th>#</th><th>employeeNumber</th><th>lastName</th><th>firstName</th><th>extension</th><th>email</th><th>officeCode</th><th>reportsTo</th><th>jobTitle</th><th>employeeNumber</th></tr></thead><tbody><tr><td>1</td><td>1002</td><td>Murphy</td><td>Diane</td><td>x5800</td><td>dmurphy@classicmodelcars.com</td><td>1</td><td><null></td><td>President</td><td></td></tr><tr><td>2</td><td>1056</td><td>Patterson</td><td>Mary</td><td>x4811</td><td>mpatterson@classicmodelcars.com</td><td>1</td><td>1002</td><td>VP Sales</td><td></td></tr><tr><td>3</td><td>1076</td><td>Finelli</td><td>Jeff</td><td>x9273</td><td>jfinelli@classicmodelcars.com</td><td>1</td><td>1002</td><td>VP Marketing</td><td></td></tr><tr><td>4</td><td>1088</td><td>Patterson</td><td>William</td><td>x4871</td><td>wpatterson@classicmodelcars.com</td><td>6</td><td>1056</td><td>Sales Manager (APAC)</td><td></td></tr><tr><td>5</td><td>1102</td><td>Bondur</td><td>Gerard</td><td>x5408</td><td>gbondur@classicmodelcars.com</td><td>4</td><td>1056</td><td>Sale Manager (EMEA)</td><td></td></tr><tr><td>6</td><td>1143</td><td>Bow</td><td>Anthony</td><td>x5438</td><td>abow@classicmodelcars.com</td><td>1</td><td>1056</td><td>Sales Manager (NA)</td><td></td></tr><tr><td>7</td><td>1165</td><td>Jennings</td><td>Leslie</td><td>x3291</td><td>ljennings@classicmodelcars.com</td><td>1</td><td>1143</td><td>Sales Rep</td><td></td></tr><tr><td>8</td><td>1166</td><td>Thompson</td><td>Leslie</td><td>x4865</td><td>lthompson@classicmodelcars.com</td><td>1</td><td>1143</td><td>Sales Rep</td><td></td></tr><tr><td>9</td><td>1188</td><td>Finelli</td><td>Julie</td><td>x2173</td><td>jfinelli@classicmodelcars.com</td><td>2</td><td>1143</td><td>Sales Rep</td><td></td></tr></tbody></table>	#	employeeNumber	lastName	firstName	extension	email	officeCode	reportsTo	jobTitle	employeeNumber	1	1002	Murphy	Diane	x5800	dmurphy@classicmodelcars.com	1	<null>	President		2	1056	Patterson	Mary	x4811	mpatterson@classicmodelcars.com	1	1002	VP Sales		3	1076	Finelli	Jeff	x9273	jfinelli@classicmodelcars.com	1	1002	VP Marketing		4	1088	Patterson	William	x4871	wpatterson@classicmodelcars.com	6	1056	Sales Manager (APAC)		5	1102	Bondur	Gerard	x5408	gbondur@classicmodelcars.com	4	1056	Sale Manager (EMEA)		6	1143	Bow	Anthony	x5438	abow@classicmodelcars.com	1	1056	Sales Manager (NA)		7	1165	Jennings	Leslie	x3291	ljennings@classicmodelcars.com	1	1143	Sales Rep		8	1166	Thompson	Leslie	x4865	lthompson@classicmodelcars.com	1	1143	Sales Rep		9	1188	Finelli	Julie	x2173	jfinelli@classicmodelcars.com	2	1143	Sales Rep		Mengambil data langsung dari database sumber
#	employeeNumber	lastName	firstName	extension	email	officeCode	reportsTo	jobTitle	employeeNumber																																																																																														
1	1002	Murphy	Diane	x5800	dmurphy@classicmodelcars.com	1	<null>	President																																																																																															
2	1056	Patterson	Mary	x4811	mpatterson@classicmodelcars.com	1	1002	VP Sales																																																																																															
3	1076	Finelli	Jeff	x9273	jfinelli@classicmodelcars.com	1	1002	VP Marketing																																																																																															
4	1088	Patterson	William	x4871	wpatterson@classicmodelcars.com	6	1056	Sales Manager (APAC)																																																																																															
5	1102	Bondur	Gerard	x5408	gbondur@classicmodelcars.com	4	1056	Sale Manager (EMEA)																																																																																															
6	1143	Bow	Anthony	x5438	abow@classicmodelcars.com	1	1056	Sales Manager (NA)																																																																																															
7	1165	Jennings	Leslie	x3291	ljennings@classicmodelcars.com	1	1143	Sales Rep																																																																																															
8	1166	Thompson	Leslie	x4865	lthompson@classicmodelcars.com	1	1143	Sales Rep																																																																																															
9	1188	Finelli	Julie	x2173	jfinelli@classicmodelcars.com	2	1143	Sales Rep																																																																																															
Select values	<div>Table input</div> <div><div>Step name</div><div>Table input</div></div> <div><div>Connection</div><div>conn_oltp_resources</div><div>Edit...</div><div>New...</div><div>Wizard...</div></div> <div><div>SQL</div><div><div>Get SQL select statement...</div></div><div><pre>select * from employees e left join employees r e.reportsTo=r.employeeNumber;</pre></div></div> <div><div>Line 1 Column 0</div><div><div>Store column info in step meta</div><div>Enable lazy conversion</div><div>Replace variables in script?</div><div>Insert data from step</div><div>Execute for each row?</div><div>Limit size</div><div>0</div></div></div> <div><div>Help</div><div>OK</div><div>Preview</div><div>Cancel</div></div>	<div>Execution Results</div> <div><div>Logging</div><div>Execution History</div><div>Step Metrics</div><div>Performance Graph</div><div>Metrics</div><div>Preview data</div></div> <div><div>First rows</div><div>Last rows</div><div>Off</div></div> <table><thead><tr><th>#</th><th>StreamEmployeeNumber</th><th>StreamLastname</th><th>StreamFirstname</th><th>StreamJobtitle</th><th>StreamLastnameBoss</th><th>StreamFirstnameBoss</th></tr></thead><tbody><tr><td>1</td><td>1002</td><td>Murphy</td><td>Diane</td><td>President</td><td><null></td><td><null></td></tr><tr><td>2</td><td>1056</td><td>Patterson</td><td>Mary</td><td>VP Sales</td><td>Murphy</td><td>Diane</td></tr><tr><td>3</td><td>1076</td><td>Finelli</td><td>Jeff</td><td>VP Marketing</td><td>Murphy</td><td>Diane</td></tr><tr><td>4</td><td>1088</td><td>Patterson</td><td>William</td><td>Sales Manager (APAC)</td><td>Patterson</td><td>Mary</td></tr><tr><td>5</td><td>1102</td><td>Bondur</td><td>Gerard</td><td>Sale Manager (EMEA)</td><td>Patterson</td><td>Mary</td></tr></tbody></table>	#	StreamEmployeeNumber	StreamLastname	StreamFirstname	StreamJobtitle	StreamLastnameBoss	StreamFirstnameBoss	1	1002	Murphy	Diane	President	<null>	<null>	2	1056	Patterson	Mary	VP Sales	Murphy	Diane	3	1076	Finelli	Jeff	VP Marketing	Murphy	Diane	4	1088	Patterson	William	Sales Manager (APAC)	Patterson	Mary	5	1102	Bondur	Gerard	Sale Manager (EMEA)	Patterson	Mary	Memilih kolom yang dibutuhkan dan mengganti nama field																																																										
#	StreamEmployeeNumber	StreamLastname	StreamFirstname	StreamJobtitle	StreamLastnameBoss	StreamFirstnameBoss																																																																																																	
1	1002	Murphy	Diane	President	<null>	<null>																																																																																																	
2	1056	Patterson	Mary	VP Sales	Murphy	Diane																																																																																																	
3	1076	Finelli	Jeff	VP Marketing	Murphy	Diane																																																																																																	
4	1088	Patterson	William	Sales Manager (APAC)	Patterson	Mary																																																																																																	
5	1102	Bondur	Gerard	Sale Manager (EMEA)	Patterson	Mary																																																																																																	

Database
lookup

Select values

Step nameSelect values

Select & AlterRemoveMeta-data

Fields:

#	Fieldname	Rename to	Length	Precision
1	employeeNumber	StreamEmployeeenumber		
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

Include unspecified fields, ordered by name☐

Help

OK

Cancel

#	StreamEmployeeenumber	StreamLastname	StreamFirstname	StreamJobtitle	StreamLastnameBoss	StreamFirstnameBoss	employeeNumber	firstNam
1	1002	Murphy	Diane	President	<null>	<null>	<null>	<null>
2	1056	Patterson	Mary	VP Sales	Murphy	Diane	<null>	<null>
3	1076	Finnell	Jeff	VP Marketing	Murphy	Diane	<null>	<null>
4	1088	Patterson	William	Sales Manager (APAC)	Patterson	Mary	<null>	<null>
5	1102	Bondur	Gerard	Sale Manager (EMEA)	Patterson	Mary	<null>	<null>
6	1143	Bow	Anthony	Sales Manager (NA)	Patterson	Mary	<null>	<null>
7	1165	Jennings	Leslie	Sales Rep	Bow	Anthony	<null>	<null>
8	1166	Thompson	Leslie	Sales Rep	Bow	Anthony	<null>	<null>
9	1188	Finnell	Julie	Sales Rep	Bow	Anthony	<null>	<null>

Melakukan
lookup data dari
tabel referensi
(misalnya,
mencari nama
atasan dari
reportsTo)

Filter rows

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
1	employeenumber	=	StreamEmployeenumber	
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

Execution Results

Logging

Execution History

Step Metrics

Performance Graph

Metrics

Preview data

First rows

Last rows

0/0

#	StreamEmployeenumber	StreamLastname	StreamFirstname	StreamJobtitle	StreamLastnameBoss	StreamFirstnameBoss	employeenumber	firstName
1	1052	Murphy	Diane	President	<null>	<null>	<null>	<null>
2	1056	Patterson	Mary	VP Sales	Murphy	Diane	<null>	<null>
3	1076	Finelli	Jeff	VP Marketing	Murphy	Diane	<null>	<null>
4	1088	Patterson	William	Sales Manager (APAC)	Patterson	Mary	<null>	<null>
5	1102	Bondur	Gerard	Sale Manager (EMEA)	Patterson	Mary	<null>	<null>
6	1143	Bow	Anthony	Sales Manager (NA)	Patterson	Mary	<null>	<null>

Memfilter data berdasarkan kondisi tertentu

Table Output

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

Execution Results

Logging

Execution History

Step Metrics

Performance Graph

Metrics

Preview data

First rows

Last rows

OH

#	StreamEmployeeNumber	StreamLastname	StreamFirstname	StreamJobtitle	StreamLastnameBoss	StreamFirstnameBoss	employeeNumber	firstName
1	1002	Murphy	Diane	President	<null>	<null>	<null>	<null>
2	1056	Patterson	Mary	VP Sales	Murphy	Diane	<null>	<null>
3	1076	Firelli	Jeff	VP Marketing	Murphy	Diane	<null>	<null>
4	1088	Patterson	William	Sales Manager (APAC)	Patterson	Mary	<null>	<null>
5	1102	Bondur	Gerard	Sale Manager (EMEA)	Patterson	Mary	<null>	<null>
6	1140	Bow	Anthony	Sales Manager (NA)	Patterson	Mary	<null>	<null>
7	1165	Jennings	Leslie	Sales Rep	Bow	Anthony	<null>	<null>
8	1166	Thompson	Leslie	Sales Rep	Bow	Anthony	<null>	<null>

Menyimpan hasil transformasi ke database dw legendvehicle

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

Welcome!

Transformation 2153

75%

Table input: select values Database lookup

Filter rows

Table output

Execution Results

Logging

Execution History

Step Metrics

Performance Graph

Metrics

Preview data

2025/03/11 15:08:43 - Spoon - Transformation opened.

2025/03/11 15:08:43 - Spoon - Launching transformation [Transformation 2153]...

2025/03/11 15:08:43 - Spoon - Started the transformation execution.

2025/03/11 15:08:43 - Transformation 2153 - Dispatching started for transformation [Transformation 2153]

2025/03/11 15:08:44 - Table output.0 - Connected to database [conn_dw_destination] (commit=1000)

2025/03/11 15:08:44 - Table input.0 - Finished reading query, closing connection

2025/03/11 15:08:44 - Table input.0 - Finished processing (I=23, O=0, R=0, W=23, U=0, E=0)

2025/03/11 15:08:44 - Select values.0 - Finished processing (I=0, O=0, R=23, W=23, U=0, E=0)

2025/03/11 15:08:44 - Database lookup.0 - Finished processing (I=22, O=0, R=23, W=23, U=0, E=0)

2025/03/11 15:08:44 - Filter rows.0 - Finished processing (I=0, O=0, R=23, W=1, U=0, E=0)

2025/03/11 15:08:44 - Table output.0 - Finished processing (I=0, O=1, R=1, W=1, U=0, E=0)

2025/03/11 15:08:44 - Spoon - The transformation has finished!!

- Iya, bisa redundant tidak memiliki 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
3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?

```

1 row inserted.
Inserted row id: 25 (Query took 0.0184 seconds.)

INSERT INTO dimEmployees (employeeNumber, firstName, lastName, jobTitle, boss_firstName, boss_lastName, updated) VALUES (1203, 'Azarine', 'Dhynar', 'Data Analyst', 'Mary', 'Patterson', NOW());
[ Edit inline ] [ Edit ] [ Create PHP code ]

```

id	employeeNumber	firstName	lastName	jobTitle	boss_firstName	boss_lastName	updated		
<input type="checkbox"/>	Edit	Copy	Delete	22	1625	Yoshimi Kato	Sales Rep	Mami Nishi	2025-03-11 14:04:22
<input type="checkbox"/>	Edit	Copy	Delete	23	1702	Martin Gerard	Sales Rep	Gerard Bondur	2025-03-11 14:04:22
<input type="checkbox"/>	Edit	Copy	Delete	24	1002	Diane Murphy	President	NULL	2025-03-11 15:08:44
<input type="checkbox"/>	Edit	Copy	Delete	25	1203	Azarine Dhynar	Data Analyst	Mary Patterson	2025-03-11 15:27:05

- Table input akan membaca data terbaru termasuk entri baru
- Select values dan database lookup akan memproses data termasuk nama baru (Dhynar)
- Jika filter rows tidak memfilter data bernama Dhynar, maka nama Dhynar akan ikut masuk ke table output