

LAPORAN PRAKTIKUM
DATA WAREHOUSE

JOBSHEET 3
DATABASE ANALYTICAL



Oleh:

Syifa Revalina Kamila

2341760065

SIB 2B

PROGRAM STUDI SISTEM INFORMASI BISNIS
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG 2025/2026

Jobsheet 3: Database Analytical

A. Dimensi Waktu

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

```
CREATE DATABASE dw_LegendVehicle;
use dw_LegendVehicle;
```



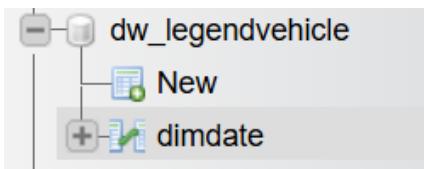
MySQL returned an empty result set (i.e. zero rows). (Query took 0.0731 seconds.)
CREATE DATABASE dw_LegendVehicle;
[Edit inline] [Edit] [Create PHP code]

MySQL returned an empty result set (i.e. zero rows). (Query took 0.0008 seconds.)
USE dw_LegendVehicle;
[Edit inline] [Edit] [Create PHP code]

MySQL returned an empty result set (i.e. zero rows). (Query took 0.0269 seconds.)

- 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
);
```



- Buka PDI Spoon. Buka Transformation baru -> File - New - Transformation.
- 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).



- 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**.

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value
1	CurrentDate	Date	dd-MM...	10					01-01-2003

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.

Name of value:	incrementDay
Use a database to generate the sequence:	<input checked="" type="checkbox"/>
Use DB to get sequence?	<input checked="" type="checkbox"/>
Connection:	[empty]
Schema name:	[empty]
Sequence name:	SEQ_
Use a transformation counter to generate the sequence:	<input checked="" type="checkbox"/>
Use counter to calculate sequence?	<input checked="" type="checkbox"/>
Counter name (optional):	[empty]
Start at value:	0
Increment by:	1
Maximum value:	999999999

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**

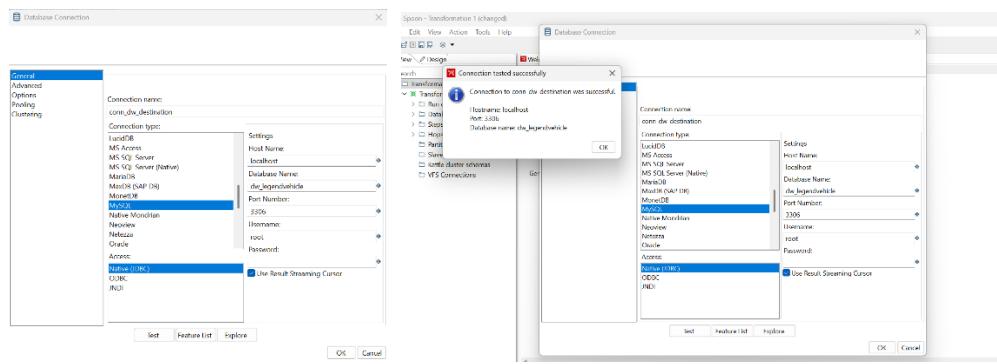
#	New field	Calculation	Field A	Field B	Field C	Value type	Length
1	streamDate	Date A + B...	Current...	increm...		None	
2	streamYear	Year of dat...	stream...			None	
3	streamMonth	Month of ...	stream...			None	
4	streamDay	Day of mo...	stream...			None	

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.

#	Fieldname	Rename to	Length	Precision
1	CurrentDate			
2	incrementDay			
3	streamDate			
4	streamYear			
5	streamMonth			
6	streamDay			

#	Fieldname
1	CurrentDate
2	incrementDay

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**.

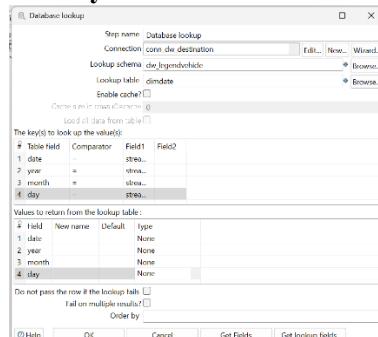


14. Konfigurasi pada **database lookup** adalah dengan memberikan **connection** dengan koneksi yang sudah dibuat pada step sebelumnya. dengan **schema** nama database yang digunakan dan **tabel dimdate** yang telah dibuat pada langkah pertama.

15. Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak kembar adalah:

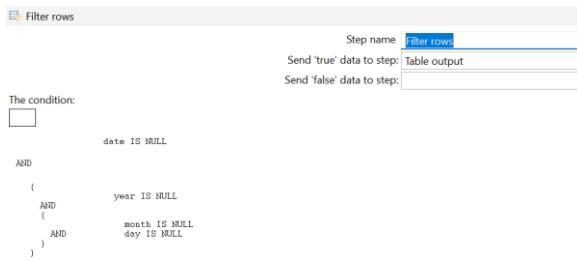
- field **date** pada table **dimdate** dengan field **streamDate**
- field **year** pada table **dimdate** dengan field **streamYear**
- field **month** pada table **dimdate** dengan field **streamMonth**
- field **day** pada table **dimdate** dengan field **streamDay**

16. Field yang akan di **retrieve** adalah field yang ada pada table **dimDate** yaitu **date**, **year**, **month**, dan **day**.

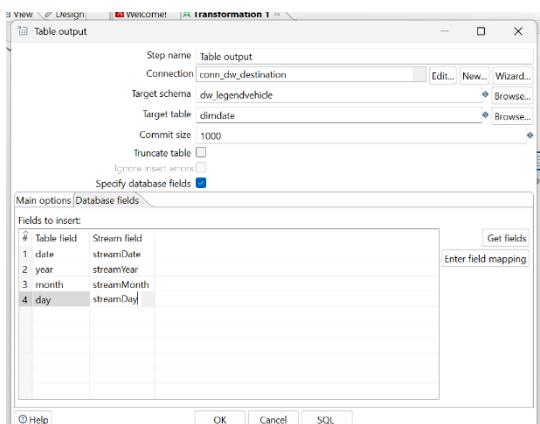


17. Hubungkan output dari **database lookup** dengan **filter rows**.

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.

		id_dimDate	date	year	month	day
□		1	2003-01-01	2003	1	1
□		2	2003-01-02	2003	1	2
□		3	2003-01-03	2003	1	3
□		4	2003-01-04	2003	1	4
□		5	2003-01-05	2003	1	5
□		6	2003-01-06	2003	1	6
□		7	2003-01-07	2003	1	7
□		8	2003-01-08	2003	1	8
□		9	2003-01-09	2003	1	9
□		10	2003-01-10	2003	1	10
■	Console	11	2003-01-11	2003	1	11

TUGAS 1

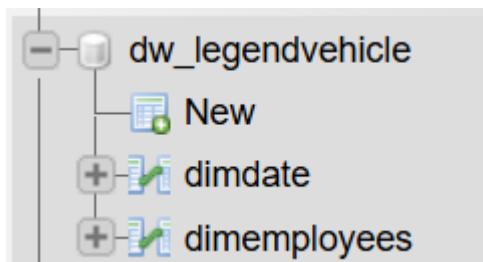
1. Buka preview tab pada execution result area di setiap proses object. amati input dan output data yang ada. bandingkan di setiap prosesnya. jelaskan perbedaan disetiap prosesnya.

Proses Objek	SS data input	SS data output	Keterangan
Generate rows			Step ini digunakan untuk mengasilkan

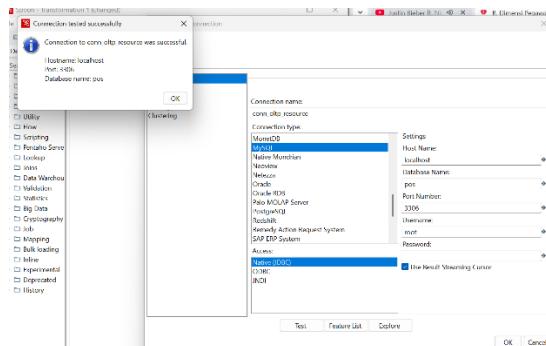
			data secara otomatis tanpa input dari yg lain.																																																																																																																					
Add Sequences	<p>Spoon name: Add sequences Name of value: incrementDay use a database to generate the sequence Use URL to get sequence Value or name: streamDate Sequence name: SGD to a transformation source to generate the sequence to a source to calculate sequences Create a new Spoolfile Start at value: 0 Increment by: 1 Maximum value: 99999999 Help OK Cancel</p>	<p>Rows of step: Add sequence (1000 rows)</p> <table border="1"> <thead> <tr> <th>#</th> <th>CurrentDate</th> <th>incrementDay</th> </tr> </thead> <tbody> <tr><td>1</td><td>01-01-2003</td><td>0</td></tr> <tr><td>2</td><td>01-01-2003</td><td>1</td></tr> <tr><td>3</td><td>01-01-2003</td><td>2</td></tr> <tr><td>4</td><td>01-01-2003</td><td>3</td></tr> </tbody> </table>	#	CurrentDate	incrementDay	1	01-01-2003	0	2	01-01-2003	1	3	01-01-2003	2	4	01-01-2003	3	Step ini digunakan untuk menambahkan nomor urut (sequence) ke setiap baris data yang dihasilkan sebelumnya.																																																																																																						
#	CurrentDate	incrementDay																																																																																																																						
1	01-01-2003	0																																																																																																																						
2	01-01-2003	1																																																																																																																						
3	01-01-2003	2																																																																																																																						
4	01-01-2003	3																																																																																																																						
Calculator	<p>Fields:</p> <table border="1"> <tr><td>1 streamDate</td><td>FieldA</td><td>FieldB</td><td>FieldC</td><td>Value type</td><td>Length</td><td>Precision</td><td>Remove</td></tr> <tr><td>1 streamYear</td><td>dateA + 2 days</td><td>CurrentDate</td><td>incrementDay</td><td>None</td><td>N</td><td></td><td></td></tr> <tr><td>2 streamYear</td><td>Year of dateA</td><td>streamDate</td><td></td><td>None</td><td>N</td><td></td><td></td></tr> <tr><td>3 streamMonth</td><td>Month of dateA</td><td>streamDate</td><td></td><td>None</td><td>N</td><td></td><td></td></tr> <tr><td>4 streamDay</td><td>Day of month of dateA</td><td>streamDate</td><td></td><td>None</td><td>N</td><td></td><td></td></tr> </table>	1 streamDate	FieldA	FieldB	FieldC	Value type	Length	Precision	Remove	1 streamYear	dateA + 2 days	CurrentDate	incrementDay	None	N			2 streamYear	Year of dateA	streamDate		None	N			3 streamMonth	Month of dateA	streamDate		None	N			4 streamDay	Day of month of dateA	streamDate		None	N			<p>Rows of step: Calculator (1000 rows)</p> <table border="1"> <thead> <tr> <th>#</th> <th>CurrentDate</th> <th>incrementDay</th> <th>streamDate</th> <th>streamYear</th> <th>streamMonth</th> <th>streamDay</th> </tr> </thead> <tbody> <tr><td>1</td><td>01-01-2003</td><td>0</td><td>2003-01-01 00:00:00</td><td>2003</td><td>1</td><td>1</td></tr> <tr><td>2</td><td>01-01-2003</td><td>1</td><td>2003-01-02 00:00:00</td><td>2003</td><td>1</td><td>2</td></tr> <tr><td>3</td><td>01-01-2003</td><td>2</td><td>2003-01-03 00:00:00</td><td>2003</td><td>1</td><td>3</td></tr> <tr><td>4</td><td>01-01-2003</td><td>3</td><td>2003-01-04 00:00:00</td><td>2003</td><td>1</td><td>4</td></tr> <tr><td>5</td><td>01-01-2003</td><td>4</td><td>2003-01-05 00:00:00</td><td>2003</td><td>1</td><td>5</td></tr> <tr><td>6</td><td>01-01-2003</td><td>5</td><td>2003-01-06 00:00:00</td><td>2003</td><td>1</td><td>6</td></tr> </tbody> </table>	#	CurrentDate	incrementDay	streamDate	streamYear	streamMonth	streamDay	1	01-01-2003	0	2003-01-01 00:00:00	2003	1	1	2	01-01-2003	1	2003-01-02 00:00:00	2003	1	2	3	01-01-2003	2	2003-01-03 00:00:00	2003	1	3	4	01-01-2003	3	2003-01-04 00:00:00	2003	1	4	5	01-01-2003	4	2003-01-05 00:00:00	2003	1	5	6	01-01-2003	5	2003-01-06 00:00:00	2003	1	6	Step ini digunakan untuk melakukan operasi matematika pada data.																												
1 streamDate	FieldA	FieldB	FieldC	Value type	Length	Precision	Remove																																																																																																																	
1 streamYear	dateA + 2 days	CurrentDate	incrementDay	None	N																																																																																																																			
2 streamYear	Year of dateA	streamDate		None	N																																																																																																																			
3 streamMonth	Month of dateA	streamDate		None	N																																																																																																																			
4 streamDay	Day of month of dateA	streamDate		None	N																																																																																																																			
#	CurrentDate	incrementDay	streamDate	streamYear	streamMonth	streamDay																																																																																																																		
1	01-01-2003	0	2003-01-01 00:00:00	2003	1	1																																																																																																																		
2	01-01-2003	1	2003-01-02 00:00:00	2003	1	2																																																																																																																		
3	01-01-2003	2	2003-01-03 00:00:00	2003	1	3																																																																																																																		
4	01-01-2003	3	2003-01-04 00:00:00	2003	1	4																																																																																																																		
5	01-01-2003	4	2003-01-05 00:00:00	2003	1	5																																																																																																																		
6	01-01-2003	5	2003-01-06 00:00:00	2003	1	6																																																																																																																		
Select Values	<p>Fields :</p> <table border="1"> <tr><td>#</td><td>Fieldname</td><td>Rename to</td></tr> <tr><td>1</td><td>CurrentDate</td><td></td></tr> <tr><td>2</td><td>incrementDay</td><td></td></tr> <tr><td>3</td><td>streamDate</td><td></td></tr> <tr><td>4</td><td>streamYear</td><td></td></tr> <tr><td>5</td><td>streamMonth</td><td></td></tr> <tr><td>6</td><td>streamDay</td><td></td></tr> </table>	#	Fieldname	Rename to	1	CurrentDate		2	incrementDay		3	streamDate		4	streamYear		5	streamMonth		6	streamDay		<p>Rows of step: Select values (1000 rows)</p> <table border="1"> <thead> <tr> <th>#</th> <th>streamDate</th> <th>streamYear</th> <th>streamMonth</th> <th>streamDay</th> </tr> </thead> <tbody> <tr><td>1</td><td>2003-01-01 00:00:00</td><td>2003</td><td>1</td><td>1</td></tr> <tr><td>2</td><td>2003-01-02 00:00:00</td><td>2003</td><td>1</td><td>2</td></tr> <tr><td>3</td><td>2003-01-03 00:00:00</td><td>2003</td><td>1</td><td>3</td></tr> <tr><td>4</td><td>2003-01-04 00:00:00</td><td>2003</td><td>1</td><td>4</td></tr> <tr><td>5</td><td>2003-01-05 00:00:00</td><td>2003</td><td>1</td><td>5</td></tr> <tr><td>6</td><td>2003-01-06 00:00:00</td><td>2003</td><td>1</td><td>6</td></tr> </tbody> </table>	#	streamDate	streamYear	streamMonth	streamDay	1	2003-01-01 00:00:00	2003	1	1	2	2003-01-02 00:00:00	2003	1	2	3	2003-01-03 00:00:00	2003	1	3	4	2003-01-04 00:00:00	2003	1	4	5	2003-01-05 00:00:00	2003	1	5	6	2003-01-06 00:00:00	2003	1	6	Step ini digunakan untuk memilih kolom yang akan dipertahankan dalam proses selanjutnya.																																																													
#	Fieldname	Rename to																																																																																																																						
1	CurrentDate																																																																																																																							
2	incrementDay																																																																																																																							
3	streamDate																																																																																																																							
4	streamYear																																																																																																																							
5	streamMonth																																																																																																																							
6	streamDay																																																																																																																							
#	streamDate	streamYear	streamMonth	streamDay																																																																																																																				
1	2003-01-01 00:00:00	2003	1	1																																																																																																																				
2	2003-01-02 00:00:00	2003	1	2																																																																																																																				
3	2003-01-03 00:00:00	2003	1	3																																																																																																																				
4	2003-01-04 00:00:00	2003	1	4																																																																																																																				
5	2003-01-05 00:00:00	2003	1	5																																																																																																																				
6	2003-01-06 00:00:00	2003	1	6																																																																																																																				
Databse lookup	<p>The logic to look up the values:</p> <table border="1"> <tr><td>#</td><td>Value Field</td><td>Comparator</td><td>Field</td></tr> <tr><td>1</td><td>date</td><td>=</td><td>streamDate</td></tr> <tr><td>2</td><td>year</td><td>=</td><td>streamYear</td></tr> <tr><td>3</td><td>month</td><td>=</td><td>streamMonth</td></tr> <tr><td>4</td><td>day</td><td>=</td><td>streamDay</td></tr> </table> <p>Values to return from the lookup table :</p> <table border="1"> <tr><td>#</td><td>Field</td><td>New name</td><td>Default</td><td>Type</td></tr> <tr><td>1</td><td>date</td><td>None</td><td></td><td>Date</td></tr> <tr><td>2</td><td>year</td><td>None</td><td></td><td>Date</td></tr> <tr><td>3</td><td>month</td><td>None</td><td></td><td>Date</td></tr> <tr><td>4</td><td>day</td><td>None</td><td></td><td>Date</td></tr> </table>	#	Value Field	Comparator	Field	1	date	=	streamDate	2	year	=	streamYear	3	month	=	streamMonth	4	day	=	streamDay	#	Field	New name	Default	Type	1	date	None		Date	2	year	None		Date	3	month	None		Date	4	day	None		Date	<p>Rows of step: Database lookup (1000 rows)</p> <table border="1"> <thead> <tr> <th>#</th> <th>streamDate</th> <th>streamYear</th> <th>streamMonth</th> <th>streamDay</th> <th>year</th> <th>month</th> <th>day</th> </tr> </thead> <tbody> <tr><td>1</td><td>2003-01-01 00:00:00</td><td>2003</td><td>1</td><td>1</td><td>2003-01-01 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>2</td><td>2003-01-02 00:00:00</td><td>2003</td><td>1</td><td>2</td><td>2003-01-02 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>3</td><td>2003-01-03 00:00:00</td><td>2003</td><td>1</td><td>3</td><td>2003-01-03 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>4</td><td>2003-01-04 00:00:00</td><td>2003</td><td>1</td><td>4</td><td>2003-01-04 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>5</td><td>2003-01-05 00:00:00</td><td>2003</td><td>1</td><td>5</td><td>2003-01-05 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>6</td><td>2003-01-06 00:00:00</td><td>2003</td><td>1</td><td>6</td><td>2003-01-06 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>7</td><td>2003-01-07 00:00:00</td><td>2003</td><td>1</td><td>7</td><td>2003-01-07 00:00:00</td><td>2003</td><td>1</td></tr> <tr><td>8</td><td>2003-01-08 00:00:00</td><td>2003</td><td>1</td><td>8</td><td>2003-01-08 00:00:00</td><td>2003</td><td>1</td></tr> </tbody> </table>	#	streamDate	streamYear	streamMonth	streamDay	year	month	day	1	2003-01-01 00:00:00	2003	1	1	2003-01-01 00:00:00	2003	1	2	2003-01-02 00:00:00	2003	1	2	2003-01-02 00:00:00	2003	1	3	2003-01-03 00:00:00	2003	1	3	2003-01-03 00:00:00	2003	1	4	2003-01-04 00:00:00	2003	1	4	2003-01-04 00:00:00	2003	1	5	2003-01-05 00:00:00	2003	1	5	2003-01-05 00:00:00	2003	1	6	2003-01-06 00:00:00	2003	1	6	2003-01-06 00:00:00	2003	1	7	2003-01-07 00:00:00	2003	1	7	2003-01-07 00:00:00	2003	1	8	2003-01-08 00:00:00	2003	1	8	2003-01-08 00:00:00	2003	1	Step ini digunakan untuk mencocokkan data dengan tabel referensi dari database.
#	Value Field	Comparator	Field																																																																																																																					
1	date	=	streamDate																																																																																																																					
2	year	=	streamYear																																																																																																																					
3	month	=	streamMonth																																																																																																																					
4	day	=	streamDay																																																																																																																					
#	Field	New name	Default	Type																																																																																																																				
1	date	None		Date																																																																																																																				
2	year	None		Date																																																																																																																				
3	month	None		Date																																																																																																																				
4	day	None		Date																																																																																																																				
#	streamDate	streamYear	streamMonth	streamDay	year	month	day																																																																																																																	
1	2003-01-01 00:00:00	2003	1	1	2003-01-01 00:00:00	2003	1																																																																																																																	
2	2003-01-02 00:00:00	2003	1	2	2003-01-02 00:00:00	2003	1																																																																																																																	
3	2003-01-03 00:00:00	2003	1	3	2003-01-03 00:00:00	2003	1																																																																																																																	
4	2003-01-04 00:00:00	2003	1	4	2003-01-04 00:00:00	2003	1																																																																																																																	
5	2003-01-05 00:00:00	2003	1	5	2003-01-05 00:00:00	2003	1																																																																																																																	
6	2003-01-06 00:00:00	2003	1	6	2003-01-06 00:00:00	2003	1																																																																																																																	
7	2003-01-07 00:00:00	2003	1	7	2003-01-07 00:00:00	2003	1																																																																																																																	
8	2003-01-08 00:00:00	2003	1	8	2003-01-08 00:00:00	2003	1																																																																																																																	
Filter rows	<p>DATA IS NULL</p> <p>AND</p> <p>year IS NULL</p> <p>and</p> <p>month IS NULL</p> <p>and</p> <p>day IS NULL</p>	<p>Rows of step: Filter rows (1000 rows)</p> <table border="1"> <thead> <tr> <th>#</th> <th>insertion</th> <th>streamDate</th> <th>streamYear</th> <th>streamMonth</th> <th>streamDay</th> <th>year</th> <th>month</th> <th>day</th> </tr> </thead> <tbody> <tr><td>1</td><td>insertion</td><td>streamDate</td><td>streamYear</td><td>streamMonth</td><td>streamDay</td><td>year</td><td>month</td><td>day</td></tr> </tbody> </table>	#	insertion	streamDate	streamYear	streamMonth	streamDay	year	month	day	1	insertion	streamDate	streamYear	streamMonth	streamDay	year	month	day	Digunakan untuk menyaring data berdasarkan kondisi tertentu																																																																																																			
#	insertion	streamDate	streamYear	streamMonth	streamDay	year	month	day																																																																																																																
1	insertion	streamDate	streamYear	streamMonth	streamDay	year	month	day																																																																																																																
Table Output	<p>Rows of step: Table output (1000 rows)</p> <table border="1"> <tr><td>#</td><td>streamDate</td><td>streamYear</td><td>streamMonth</td><td>streamDay</td><td>year</td><td>month</td><td>day</td></tr> </table>	#	streamDate	streamYear	streamMonth	streamDay	year	month	day	<p>Rows of step: Table output (1000 rows)</p> <table border="1"> <tr><td>#</td><td>streamDate</td><td>streamYear</td><td>streamMonth</td><td>streamDay</td><td>year</td><td>month</td><td>day</td></tr> </table>	#	streamDate	streamYear	streamMonth	streamDay	year	month	day	Digunakan untuk menyimpan data hasil transformasi ke dalam database atau file output.																																																																																																					
#	streamDate	streamYear	streamMonth	streamDay	year	month	day																																																																																																																	
#	streamDate	streamYear	streamMonth	streamDay	year	month	day																																																																																																																	

B. Dimensi Pegawai

- Buatlah tabel dimPegawai pada dw_legendVehicle.



- 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.

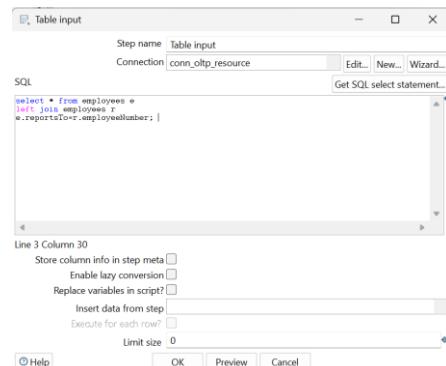


3. Drag and drop beberapa objek sebagai berikut:

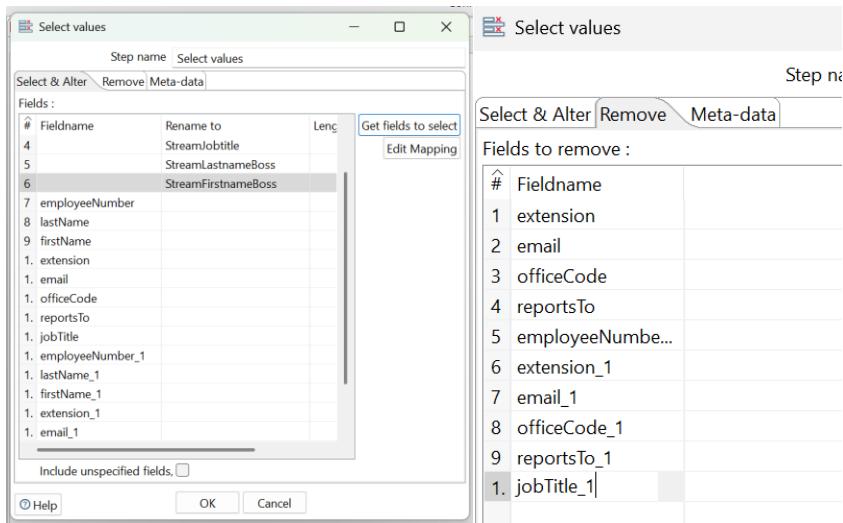
- **Table input:** digunakan mengambil data dari database OLTP.
- **Select values:** memilih field yang digunakan untuk proses Transform dan Load.
- **Database lookup:** digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar
- **Filter rows:** digunakan untuk memilih data stream yang masih belum ada apada tabel dimEmployees.
- **Table output:** Memasukkan data ke dalam tavle dimEmployees



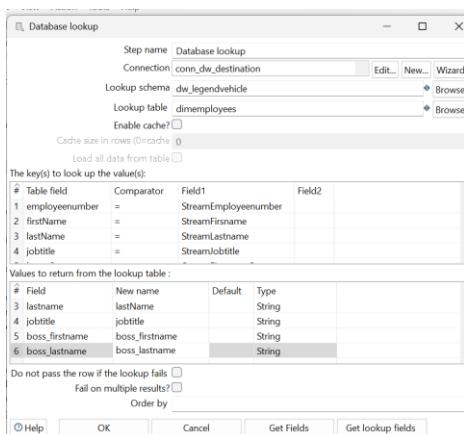
4. Konfigurasi pada table input dengan menghubungkan **Connection** pada koneksi **conn_oltp_resources**. Untuk mengambil data sumber menggunakan query dibawah ini.



5. Hubungkan output **table input** pada **select values**.
6. 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**.
7. Hilangkan field lain yang tidak digunakan pada tab **remove**.



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.



Filter rows

Step name: Filter rows

Send 'true' data to step: Table output

Send 'false' data to step:

The condition:

```

    [ ] ^ UP ^
    Level 3, Select UP to go up one level

    JobTitle IS NULL
    AND
    (
        boss_firstname IS NULL
    )

```

Table output

Step name: Table output

Connection: conn_dw_destination

Target schema: dw_legendvehicle

Target table: dimemployees

Commit size: 1000

Truncate table:

Ignore insert errors:

Specify database fields:

Main options Database fields

Fields to insert:

#	Table field	Stream field
1	employee...	StreamEmplo...
2	firstName	StreamFirstn...
3	lastName	StreamLastn...
4	jobtitle	StreamJobt...
5	boss_firstname	StreamFirstn...
6	boss_lastname	StreamLastn...

Get fields Enter field mapping

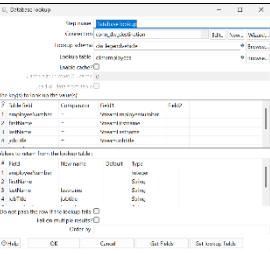
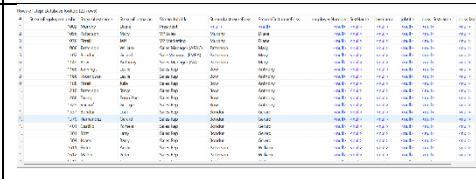
OK Cancel SQL

		id_dimEmployees	employeeNumber	firstName	lastName	JobTitle	boss_firstname	boss_lastname	updated
□	Edit	1	1001	John	Doe	Manager	Jane	Smith	2020-03-08 09:09:44
□	Edit	2	1002	Diane	Murphy	President	NULL	NULL	2020-03-11 22:06:56
□	Edit	3	1006	Mary	Patterson	VI Sales	Diane	Murphy	2020-03-11 22:06:06
□	Edit	4	1078	Jeff	Ferrini	VP Marketing	Diane	Murphy	2020-03-11 22:06:56
□	Edit	5	1080	William	Patterson	Sales Manager (APAC)	Mary	Patterson	2020-03-11 22:06:56
□	Edit	6	1092	Genard	Bondur	Sales Manager (EMEA)	Mary	Patterson	2020-03-11 22:06:56
□	Edit	7	1143	Anthony	Tow	Sales Manager (NA)	Mary	Patterson	2020-03-11 22:06:56
□	Edit	8	1160	Leslie	Jennings	Sales Rep	Anthony	Bow	2020-03-11 22:06:56
□	Edit	9	1161	Leslie	Thompson	Sales Rep	Anthony	Bow	2020-03-11 22:06:56
□	Edit	10	1188	Juli	Ferrini	Sales Rep	Anthony	Bow	2020-03-11 22:06:06
□	Edit	11	1216	Steve	Pollermann	Sales Rep	Anthony	Bow	2020-03-11 22:06:56
□	Edit	12	1280	Foon Yip	Tseng	Sales Rep	Anthony	Bow	2020-03-11 22:06:56
□	Edit	13	1321	George	Weasif	Sales Rep	Anthony	Bow	2020-03-11 22:06:56
□	Edit	14	1337	Touli	Bondur	Sales Rep	Genard	Bondur	2020-03-11 22:06:56
□	Edit	15	1370	Genard	Hernandez	Sales Rep	Genard	Bondur	2020-03-11 22:06:56
□	Edit	16	1401	Pamela	Castillo	Sales Rep	Genard	Bondur	2020-03-11 22:06:56
□	Edit	17	1501	Lenny	Bott	Sales Rep	Genard	Bondur	2020-03-11 22:06:56

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.

Proses Objek	SS data input	SS data output	Keterangan
Table Input	<pre> SELECT * FROM classicmodels.employees e LEFT JOIN classicmodels.employees r ON e.reportsTo = r.employeeNumber; </pre>		Hasil output adalah seluruh dataset awal sebelum transformasi.
Select values	<p>Select & Alter Remove Meta data </p> <p>Fields to remove :</p> <ul style="list-style-type: none"> 1 Fieldname 2 extension 3 officeCode 4 reportsTo 5 employeeNumber 6 extension_1 7 email_1 8 officeCode_1 9 reportsTo_1 10 jobTitle 		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	<p>The condition:</p> <pre> 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 </pre>	<p>The condition:</p> <pre> 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 </pre>	Input memiliki semua data sebelum filtering, sedangkan output hanya menyisakan data yang memenuhi kondisi.																																																																																																																									
Table output	<p>Step name: <u>table output</u></p> <p>Connection: conn_dw_destination</p> <p>Target schema: dw_legendvehicle</p> <p>Target table: dimentemployees</p> <p>Commit size: 1000</p> <p>Truncate table: <input type="checkbox"/></p> <p>Ignore insert errors: <input type="checkbox"/></p>	<table border="1"> <tbody> <tr><td></td><td></td><td></td><td>2</td><td>1002</td><td>Dene</td><td>Misty</td><td>President</td><td>NULL</td><td>NULL</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>3</td><td>1005</td><td>May</td><td>Patricia</td><td>Vice Pres</td><td>None</td><td>Maggy</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>4</td><td>1009</td><td>Jeff</td><td>Rich</td><td>VP Marketing</td><td>Done</td><td>Maggy</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>5</td><td>1003</td><td>Wilson</td><td>Peter</td><td>Sales Manager (NA)</td><td>Maggy</td><td>Paterson</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>6</td><td>1012</td><td>Gerald</td><td>Brent</td><td>Sales Manager (EU)</td><td>Maggy</td><td>Paterson</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>7</td><td>1010</td><td>Armeny</td><td>Isaac</td><td>Sales Manager (AP)</td><td>Maggy</td><td>Paterson</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>8</td><td>1005</td><td>Leanne</td><td>Jenings</td><td>Sales Rep</td><td>Anthony</td><td>Eve</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>9</td><td>1005</td><td>Leslie</td><td>Thompson</td><td>Sales Rep</td><td>Anthony</td><td>Eve</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>10</td><td>1005</td><td>Jesse</td><td>Fried</td><td>Sales Rep</td><td>Anthony</td><td>Eve</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>11</td><td>1015</td><td>Shawn</td><td>Paterson</td><td>Sales Rep</td><td>Anthony</td><td>Eve</td><td>2023-01-11 22:05:59</td></tr> <tr><td></td><td></td><td></td><td>12</td><td>1005</td><td>Tonya</td><td>Reng</td><td>Sales Rep</td><td>Anthony</td><td>Eve</td><td>2023-01-11 22:05:59</td></tr> </tbody> </table>				2	1002	Dene	Misty	President	NULL	NULL	2023-01-11 22:05:59				3	1005	May	Patricia	Vice Pres	None	Maggy	2023-01-11 22:05:59				4	1009	Jeff	Rich	VP Marketing	Done	Maggy	2023-01-11 22:05:59				5	1003	Wilson	Peter	Sales Manager (NA)	Maggy	Paterson	2023-01-11 22:05:59				6	1012	Gerald	Brent	Sales Manager (EU)	Maggy	Paterson	2023-01-11 22:05:59				7	1010	Armeny	Isaac	Sales Manager (AP)	Maggy	Paterson	2023-01-11 22:05:59				8	1005	Leanne	Jenings	Sales Rep	Anthony	Eve	2023-01-11 22:05:59				9	1005	Leslie	Thompson	Sales Rep	Anthony	Eve	2023-01-11 22:05:59				10	1005	Jesse	Fried	Sales Rep	Anthony	Eve	2023-01-11 22:05:59				11	1015	Shawn	Paterson	Sales Rep	Anthony	Eve	2023-01-11 22:05:59				12	1005	Tonya	Reng	Sales Rep	Anthony	Eve	2023-01-11 22:05:59	Data yang masuk ke sini adalah hasil akhir setelah semua proses sebelumnya.
			2	1002	Dene	Misty	President	NULL	NULL	2023-01-11 22:05:59																																																																																																																		
			3	1005	May	Patricia	Vice Pres	None	Maggy	2023-01-11 22:05:59																																																																																																																		
			4	1009	Jeff	Rich	VP Marketing	Done	Maggy	2023-01-11 22:05:59																																																																																																																		
			5	1003	Wilson	Peter	Sales Manager (NA)	Maggy	Paterson	2023-01-11 22:05:59																																																																																																																		
			6	1012	Gerald	Brent	Sales Manager (EU)	Maggy	Paterson	2023-01-11 22:05:59																																																																																																																		
			7	1010	Armeny	Isaac	Sales Manager (AP)	Maggy	Paterson	2023-01-11 22:05:59																																																																																																																		
			8	1005	Leanne	Jenings	Sales Rep	Anthony	Eve	2023-01-11 22:05:59																																																																																																																		
			9	1005	Leslie	Thompson	Sales Rep	Anthony	Eve	2023-01-11 22:05:59																																																																																																																		
			10	1005	Jesse	Fried	Sales Rep	Anthony	Eve	2023-01-11 22:05:59																																																																																																																		
			11	1015	Shawn	Paterson	Sales Rep	Anthony	Eve	2023-01-11 22:05:59																																																																																																																		
			12	1005	Tonya	Reng	Sales Rep	Anthony	Eve	2023-01-11 22:05:59																																																																																																																		

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

Jawaban: jika proses ini diulangi (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 tabel *employee* di OLTP. Jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?

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