

LAPORAN HASIL JOBSHEET 3 BAGIAN A DAN B



DISUSUN OLEH:

DIAJENG SEKAR ARUM

2341760070

PROGRAM STUDI D-IV SISTEM INFORMASI BISNIS

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

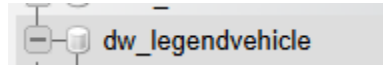
2024

A. DIMENSI WAKTU

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

```
CREATE DATABASES dw_LegendVehicle;
```

```
use dw_LegendVehicle;
```

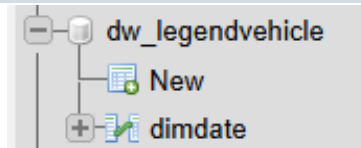


2. Buatlah table untuk menyimpan data master waktu atau yang disebut dengan tabel dimensi. Beri nama table tersebut dengan nama dimDate.

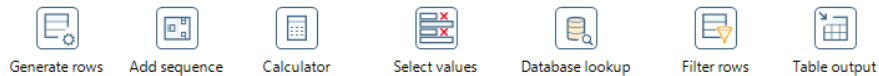
```
CREATE TABLE dimDate(  
id_dimDate int not null AUTO_INCREMENT PRIMARY KEY  
date date  
year int  
month int  
day int  
);
```

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.1359 seconds.)


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

[Edit inline](#) [\[Edit \]](#) [\[Create PHP code \]](#)

- Buka PDI Spoon. Buat Transformation baru -> File - New - Transformation.
- Drag and Drop beberapa objek yaitu:



- 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).

 Generate rows

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										

Help

OK

Preview

Cancel

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

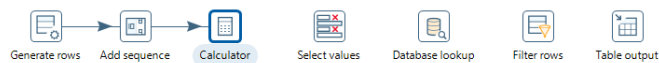
Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty
CurrentDate	Date	dd-MM...						01-01-2003	

- Hubungkan output dari Generate Rows menuju Add Sequence.



- Konfigurasi pada Add Sequences adalah merubah Name of value menjadi incrementDay dengan start value bernilai 0 dan increment by bernilai 1

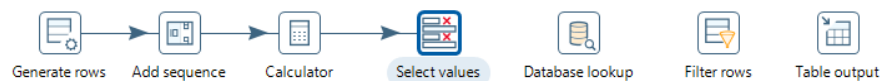
- Hubungkan output dari add sequences menuju calculator.



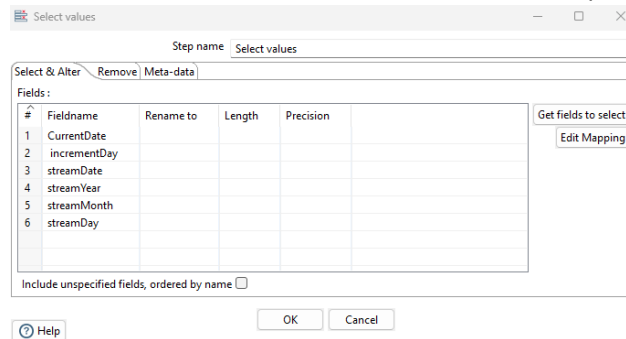
- Konfigurasi pada calculator dengan membuat fields baru sebagai berikut:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask
1	streamDate	Date A + B Days	CurrentDate	incrementDay						
2	streamYear	Year of date A	streamDate							
3	streamMonth	Month of date A	streamDate							
4	streamDay	Day of month of date A	streamDate							

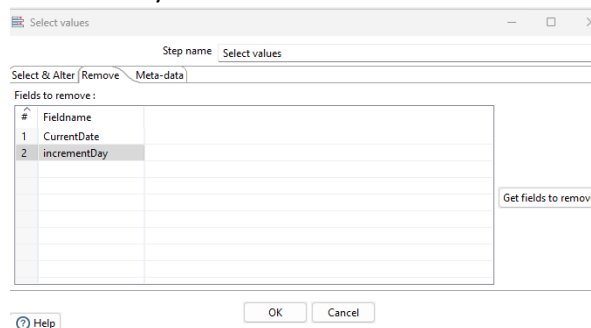
- Hubungkan output dari calculator menuju Select values



- Konfigurasi pada select values adalah dengan menekan tombol Get fields to select pada tab Select & Alter. Secara otomatis semua fields dari data input akan muncul.



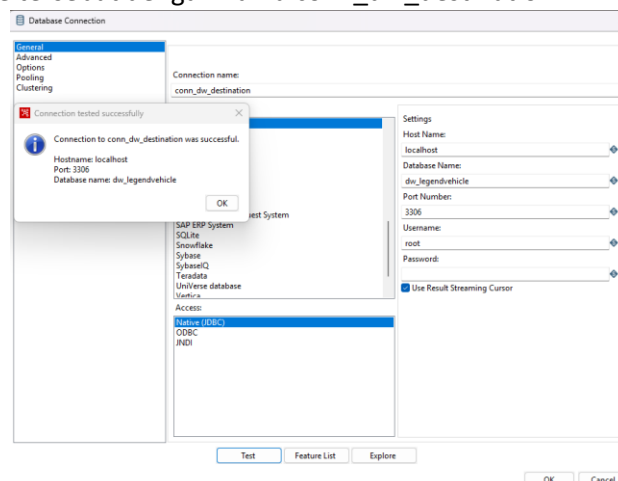
- Dikarenakan tidak semua fields digunakan, maka pada tab Remove diisikan fields CurrentDate dan incrementDay dikarenakan kedua fields tersebut tidak digunakan.



- Hubungkan output select values menuju database lookup.



- 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.



- 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.
- Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak kembar adalah:
- Field yang akan di retriive adalah field yang ada pada table dimDate yaitu date, year, month, dan day.

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:

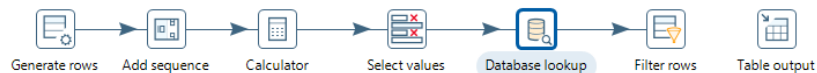
#	Field	New name	Default	Type
1	date			None
2	year			None
3	month			None
4	day			None

Do not pass the row if the lookup fails: ☐

Fail on multiple results?: ☐

Order by:

- Hubungkan output dari database lookup dengan filter rows



- 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)

Step name: Filter rows

Send 'true' data to step: Table output

Send 'false' data to step:

The condition:

```

AND
  date IS NULL
AND
  year IS NULL
AND
  (
    month IS NULL
    AND
    day IS NULL
  )
  
```

Help OK Cancel

- Hubungkan output dari filter rows menuju table output.



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

Step name: Table output

Connection: conn_dw_destination

Target schema: dw_legendvehicle

Target table: dimdate

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☒

Main options / Database fields

Fields to insert:

#	Table field	Stream field
1	date	streamDate
2	year	streamYear
3	month	streamMonth
4	day	streamDay

Buttons: Get fields, Enter field mapping, Help, OK, Cancel, SQL

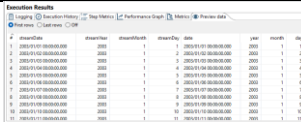
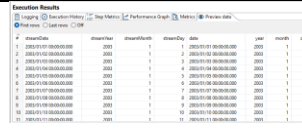
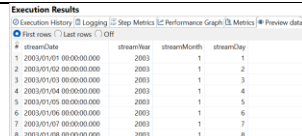
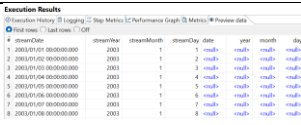
- Pada tab Database fields, mapping data input streamDate, streamYear, streamMonth dan streamDay dengan fields yang ada pada dimDate. Pada tahapan ini akan dilakukan insert data menuju tabel dimDate.
- cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.

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
11	2003-01-11	2003	1	11
12	2003-01-12	2003	1	12
13	2003-01-13	2003	1	13
14	2003-01-14	2003	1	14
15	2003-01-15	2003	1	15
16	2003-01-16	2003	1	16

#	streamDate	streamYear	streamMonth	streamDay	date	year	month	day
1	2003-06-21 00:00:00.000	2003	6	21	<null>	<null>	<null>	<null>
2	2003-06-22 00:00:00.000	2003	6	22	<null>	<null>	<null>	<null>
3	2003-06-23 00:00:00.000	2003	6	23	<null>	<null>	<null>	<null>
4	2003-06-24 00:00:00.000	2003	6	24	<null>	<null>	<null>	<null>
5	2003-06-25 00:00:00.000	2003	6	25	<null>	<null>	<null>	<null>
6	2003-06-26 00:00:00.000	2003	6	26	<null>	<null>	<null>	<null>
7	2003-06-27 00:00:00.000	2003	6	27	<null>	<null>	<null>	<null>
8	2003-06-28 00:00:00.000	2003	6	28	<null>	<null>	<null>	<null>
9	2003-06-29 00:00:00.000	2003	6	29	<null>	<null>	<null>	<null>
10	2003-06-30 00:00:00.000	2003	6	30	<null>	<null>	<null>	<null>
11	2003-07-01 00:00:00.000	2003	7	1	<null>	<null>	<null>	<null>

TUGAS:

Proses objek Generate rows	Ss input tidak ada	Ss output	Keterangan Langkah ini digunakan untuk secara otomatis menghasilkan data tanggal awal. Karena merupakan sumber data utama, langkah ini tidak memerlukan input.
Add sequence			Langkah ini untuk menambahkan deretan angka yang dapat dimanfaatkan untuk berbagai perhitungan, seperti membuat ID unik atau menentukan urutan.
Calculator			Tahapan ini digunakan untuk melakukan manipulasi data, misalnya menambahkan jumlah hari atau memecah tanggal menjadi bagian tahun, bulan, dan hari. Ini bermanfaat untuk analisis lanjutan atau pembuatan dimensi waktu.
Select values			data disaring dengan memilih hanya kolom yang relevan. Pada tahap ini juga bisa dilakukan penggantian nama kolom atau penghapusan kolom yang tidak diperlukan.

Database lookup			sistem melakukan pencarian informasi tambahan dari database berdasarkan nilai <i>streamDate</i> . Bila tidak ditemukan kecocokan di database, nilai tersebut akan dikembalikan sebagai <i>null</i> .
Filter rows			baris data akan difilter berdasarkan kondisi di mana tanggal, bulan, tahun, dan hari bernilai <i>null</i> .
Table output			<i>Table Output</i> hanya akan menampung data dengan nilai <i>null</i> pada kolom tanggal, tahun, bulan, dan hari. Sementara data dengan nilai valid akan diabaikan dan tidak disertakan dalam <i>Table Output</i> .

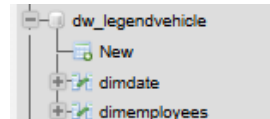
B. DIMENSI PEGAWAI

- Buatlah tabel dimPegawai pada dw_legendVehicle.

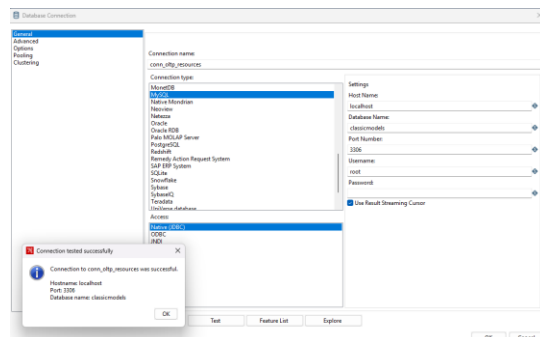
```

1 CREATE TABLE dimEmployees(
2   id_dimEmployees int NOT NULL auto_increment PRIMARY KEY,
3   employeeNumber int(11),
4   firstname varchar(50),
5   lastname varchar(50),
6   jobtitle varchar(50),
7   boss_firstname varchar(50),
8   boss_lastname varchar(50),
9   updated date DEFAULT (CURRENT_DATE)
10 );

```

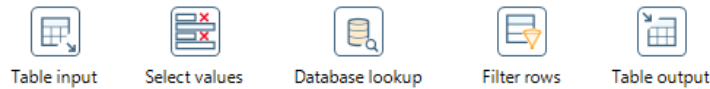


- Pada PDI Spoon buatlah koneksi baru dengan nama conn_oltp_resources yang menghubungkan dengan database oltp. sesuaikan hostname, database name, port number, username dan password dengan keadaan pada device masing-masing.

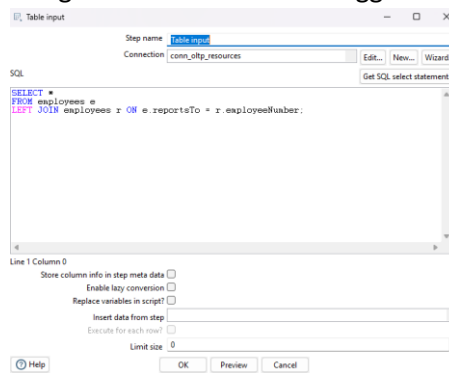


3. Drag and drop beberapa objek sebagai berikut:

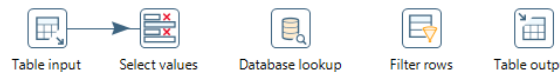
- Table input: digunakan mengambil data dari database OLTP.
- Select values: memilih field yang digunakan untuk proses Transform dan Load.
- Database lookup: digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar
- Filter rows: digunakan untuk memilih data stream yang 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.

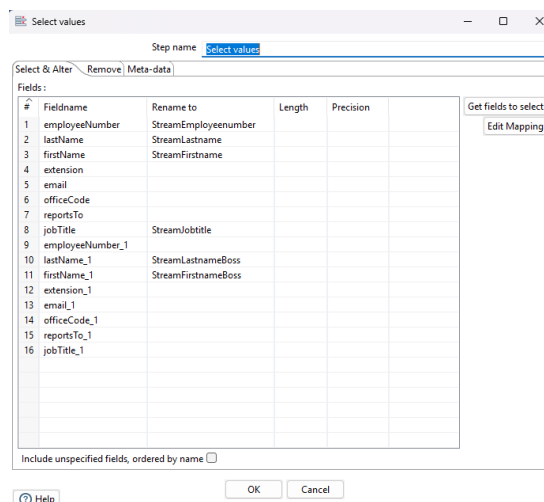


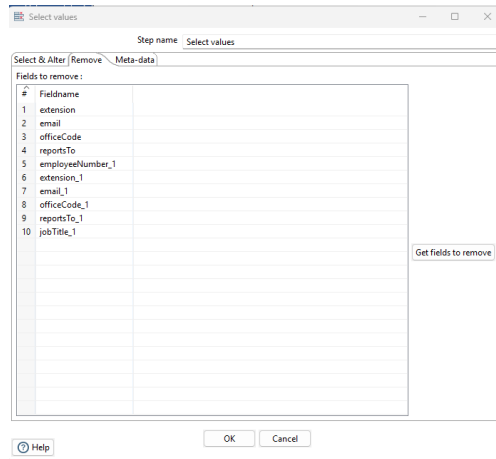
5. Hubungkan output table input pada select values.



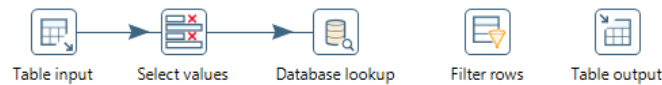
6. Konfigurasi pada Select values yaitu mengambil data dari field employeeNumber, lastname, firstname, jobtitle , lastname_1 dan firstname_1 sebagai data stream yang digunakan pada proses ETL pada tab select & alter.

7. Hilangkan field lain yang tidak digunakan pada tab remove.



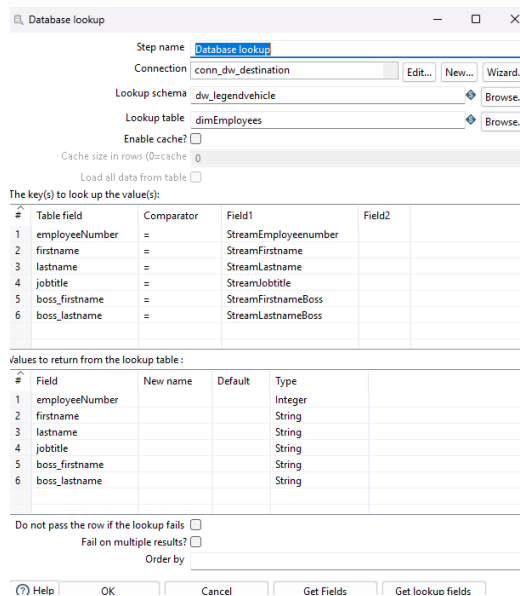


8. Hubungkan output select values pada database lookup

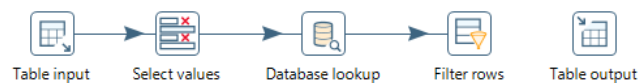


9. Konfigurasi pada database lookup adalah dengan menghubungkan koneksi pada conn_dw_destination dengan table lookup dimEmployees yang telah dibuat pada tahap pertama.

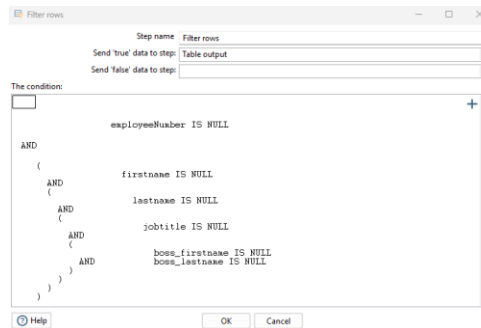
10. Field yang di lookup adalah field pada tabel dimEmployees dengan field stream input dari OLTP. sedangkan field yang di retrieve adalah field dari dimEmployees itu sendiri. Jika tidak ada data yang sama maka akan muncul null.



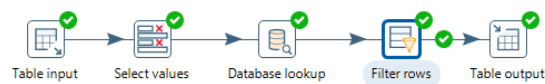
11. Hubungkan output database lookup dengan filter rows.



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.

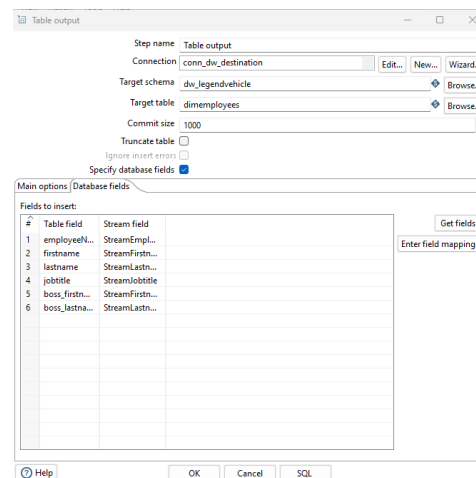


13. Hubungkan output dari filter rows dengan table output



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.


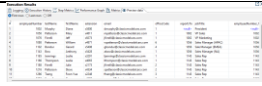
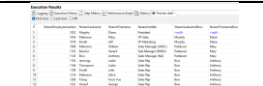
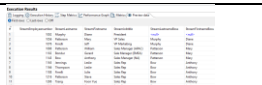
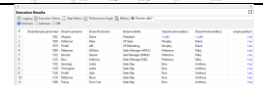
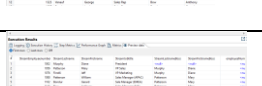
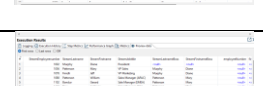
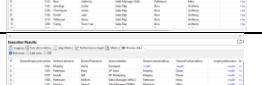
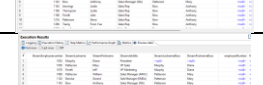


- Hasil pada phpMyAdmin

#	id_dimemployees	employeeNumber	firstname	lastname	jobtitle	boss_firstname	boss_lastname	updated
1	1	1000	Steve	Murphy	President	NULL	NULL	2025-04-26
2	2	1005	Mary	Paterson	VP Sales	Steve	Murphy	2025-04-26
3	3	1078	Jeff	Frawell	VP Marketing	Steve	Murphy	2025-04-26
4	4	1088	William	Paterson	Sales Manager (SME)	Mary	Paterson	2025-04-26
5	5	1102	Gerard	Bondur	Sales Manager (SME)	Mary	Paterson	2025-04-26
6	6	1143	Anthony	Bow	Sales Manager (SME)	Mary	Paterson	2025-04-26
7	7	1195	Leslie	Jermolov	Sales Rep	Anthony	Bow	2025-04-26
8	8	1198	Leslie	Thompson	Sales Rep	Anthony	Bow	2025-04-26
9	9	1188	Julie	Finelli	Sales Rep	Anthony	Bow	2025-04-26
10	10	1218	Steve	Paterson	Sales Rep	Anthony	Bow	2025-04-26
11	11	1285	Fran Yue	Tsang	Sales Rep	Anthony	Bow	2025-04-26
12	12	1323	Georgie	Veraut	Sales Rep	Anthony	Bow	2025-04-26
13	13	1337	Lou	Bondur	Sales Rep	Gerard	Bondur	2025-04-26
14	14	1370	Gerard	Hernandez	Sales Rep	Gerard	Bondur	2025-04-26
15	15	1401	Pamela	Cassillo	Sales Rep	Gerard	Bondur	2025-04-26
16	16	1501	Larry	Bott	Sales Rep	Gerard	Bondur	2025-04-26

TUGAS:

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	-		Mengambil data dari tabel sumber
Select Values			Memilih kolom tertentu yang diperlukan
Database Lookup			Mengambil data tambahan dari database lain berdasarkan key tertentu
Filter Rows			Menyaring data berdasarkan kondisi tertentu
Table Output			Menyimpan hasil akhir ke tabel tujuan

2. Jika proses itu di ulangi (di run kembali) apakah data akan redudant?
 - Apabila data yang digunakan tetap sama dan telah melalui proses penyaringan yang benar, maka saat transformasi dijalankan kembali, tidak akan terjadi redundansi data. Hal ini karena data yang masuk ke tahap output sudah difilter sesuai kriteria tertentu, sehingga hanya data yang diperlukan saja yang diproses dan disimpan. Dengan kata lain, proses filter berfungsi untuk memastikan bahwa tidak ada data yang sama dimasukkan lebih dari satu kali ke dalam tabel tujuan. Oleh karena itu, meskipun transformasi dilakukan berulang kali, selama data input tidak berubah dan penyaringan berjalan sesuai aturan, maka tidak akan ada duplikasi atau penggandaan data di database
3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?

```

1 INSERT INTO employees (
2   employeeNumber, lastName, firstName, extension, email, officeCode, reportsTo, jobTitle)
3 VALUES (
4   9999, 'Diajeng Sekar', 'Arum', 'x9999', 'aqilanurazzal23@gmail.com', 1, 1002, 'Data Scientist'
5 );

```

- Hasil saat di pentaho ketika dijaankan lagi

Execution Results									
	StreamEmployeeNumber	StreamLastname	StreamFirstname	StreamJobTitle	StreamLastnameBoss	StreamFirstnameBoss	employeeNumber	firstname	
1	1002	Murphy	Diane	President	<null>	<null>	<null>	<null>	
2	9999	Diajeng Sekar	Arum	Data Scientist	Murphy	Diane	<null>	<null>	

- Hasil pada phpMyAdmin

<input type="checkbox"/>				23	1702	Martin	Gerard	Sales Rep	Gerard	Bondur	2025-04-26
<input type="checkbox"/>				24	1002	Diane	Murphy	President	NULL	NULL	2025-04-26
<input type="checkbox"/>				25	9999	Arum	Diajeng Sekar	Data Scientist	Diane	Murphy	2025-04-26

- Yang terjadi adalah datanya akan redudant