

## **LAPORAN JOBSHEET 3**

### **DATABASE ANALYTICAL**

### **MATA KULIAH DATA WAREHOUSE**

Dosen Pengampu : Vipkas Al Hadid Firdaus, S.T., M.T.



**Disusun oleh :**

Dahnier Davina SIB-2A / 2341760023

**JURUSAN TEKNOLOGI INFORMASI**

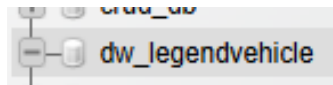
**PROGRAM STUDI D-IV SISTEM INFORMASI BISNIS**

**POLITEKNIK NEGERI MALANG**

**2025**

## A. DIMENSI WAKTU

1. Buatlah sebuah database yang digunakan sebagai OLAP dengan nama dw\_LegendVehicle

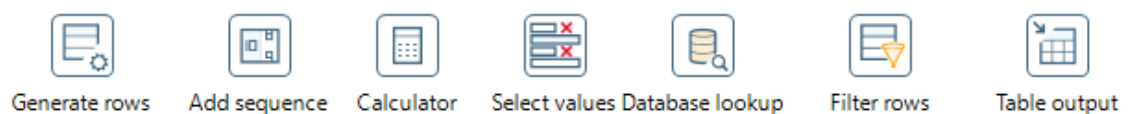


2. Buatlah table untuk menyimpan data master waktu atau yang disebut dengan dimensi. Beri nama tabel tersebut dengan nama dimDate.

```
✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0563 seconds.)  
  
CREATE TABLE dimDate( id_dimDate int not null AUTO_INCREMENT PRIMARY KEY, date date, year int,  
month int, day int );
```

3. Buatlah table dimensi dimDate, maka diperlukan generate data tanggal. 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 :

- a) Buka PDI Spoon . Buat Transformation baru -> File – new Transformation
- b) 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).



- c) 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 ).
- d) 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

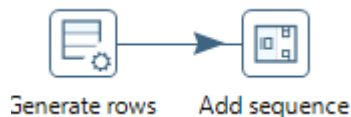
Current row time field name: now

Previous row time field name: FiveSecondsAgo

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty to null
1	CurrentDate	Date	dd-MM-...						01-01-2003	<input checked="" type="checkbox"/>

Buttons: ? Help, OK, Preview, Cancel

- e) Hubungkan output dari Generate Rows menuju Add Sequence.



- f) 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]

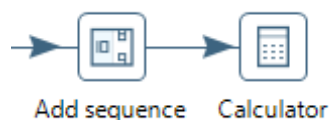
Start at value: 0

Increment by: 1

Maximum value: 99999999

Buttons: ? Help, OK, Cancel

- g) Hubungkan output dari add sequences menuju calculator.



h) Konfigurasi pada calculator dengan membuat fields baru sebagai Berikut :

- streamDate merupakan kalkulasi dari CurrentDate + incrementDay
- streamYear merupakan Year dari streamDate
- streamMonth merupakan Month dari streamDate
- streamDay merupakan Day of month dari streamDate

Step name  
Calculator

☒ Throw an error on non existing files

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove
1	streamDate	Date A + B D...	Current...	incrim...		None			N
2	streamYear	Year of date A	stream...			None			N
3	streamMonth	Month of da...	stream...			None			N
4	streamDay	Day of mont...	stream...			None			N

i) Hubungkan output dari calculator menuju Select values



j) 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.

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

Step name Select values

Select & Alter Remove Meta-data

Fields :

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

Konfig Select & alter

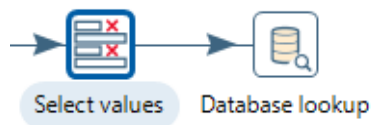
Select & Alter Remove Meta-data

Fields to remove :

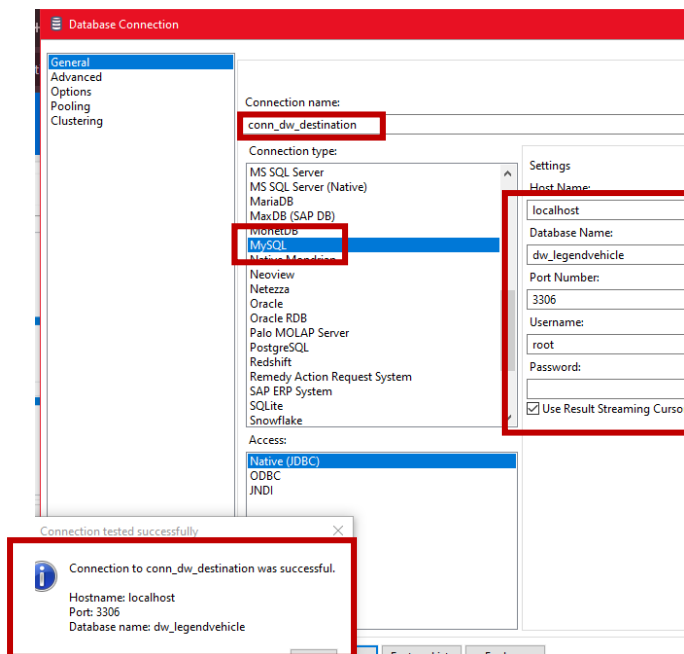
#	Fieldname
1	CurrentDate
2	incrementDay

Konfig remove

- l) Hubungkan output select values menuju database lookup.



- m) 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.



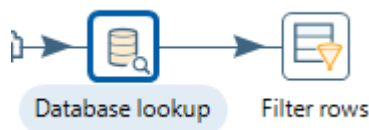
- n) Konfigurasi pada database lookup adalah dengan memberikan connection dengan koneksi yang sudah dibuat pada step sebelumnya. dengan schema nama database yang digunakan dan tabel dimdate yang telah dibuat pada langkah pertama.
- o) Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak 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

- p) Field yang akan di retrieve adalah field yang ada pada table dimDate yaitu date, year, month, dan day.

Table field	Comparison	Field1	Field2
date	=	stream...	
year	=	stream...	
month	=	stream...	
day	=	stream...	

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

- q) Hubungkan output dari database lookup dengan filter rows



- r) 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

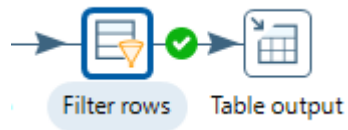
Send 'true' data to step: Table output

Send 'false' data to step:

The condition: Level 3, Select UP to go up one level

date IS NULL  
AND  
year IS NULL  
AND  
month IS NULL  
AND  
day IS NULL

- s) Hubungkan output dari filter rows menuju table output.



- t) Konfigurasi pada table output adalah memberikan koneksi pada conn\_dw\_destination dengan schema dw\_legendvehicle dan table dimdate.
- u) Aktifkan specify database fields.
- v) 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.

Connection	conn_dw_destination
Target schema	dw_legendvehicle
Target table	dimdate
Commit size	1000
Truncate table	<input type="checkbox"/>
Ignore insert errors	<input type="checkbox"/>
Specify database fields	<input checked="" type="checkbox"/>

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

- w) cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.

Showing rows 0 - 24 (1825 total, Query took 0.0004 seconds.)

```
SELECT * FROM `dimdate`
```

	id_dimDate	date	year	month	day
<input type="checkbox"/>	1	2003-01-01	2003	1	1
<input type="checkbox"/>	2	2003-01-02	2003	1	2
<input type="checkbox"/>	3	2003-01-03	2003	1	3
<input type="checkbox"/>	4	2003-01-04	2003	1	4
<input type="checkbox"/>	5	2003-01-05	2003	1	5
<input type="checkbox"/>	6	2003-01-06	2003	1	6

## 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 disetiap prosesnya.

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	Generate rows	0	0	1825	0	0	0	0	0	Finished	0.1s	34,434	-
2	Add sequence	0	1825	1825	0	0	0	0	0	Finished	0.1s	23,101	-
3	Calculator	0	1825	1825	0	0	0	0	0	Finished	0.2s	9,217	-
4	Select values	0	1825	1825	0	0	0	0	0	Finished	0.2s	9,080	-
5	Database lookup	0	1825	1825	0	0	0	0	0	Finished	2.1s	869	-
6	Filter rows	0	1825	1825	0	0	0	0	0	Finished	2.1s	867	-
7	Table output	0	1825	1825	0	1825	0	0	0	Finished	2.3s	781	-

Proses Objek	SS data Input	SS Data Output	Keterangan
<b>Generate Rows</b>	Tidak ada	1825 baris	Langkah ini membuat sejumlah baris data secara otomatis
<b>Add Sequences</b>	1825 baris	1825 baris dengan nomor urut	Langkah ini menambahkan nomor urut pada tiap baris data
<b>Calculator</b>	1825 baris + nomor urut	1825 baris dengan hasil perhitungan	Langkah ini melakukan perhitungan berdasar operasi matematika
<b>Select Values</b>	1825 baris + perhitungan	1825 baris dengan kolom yang dipilih	Langkah ini memilih kolom tertentu yang diperlukan & membuang kolom yang tidak dibutuhkan
<b>Database Lookup</b>	1825 baris + Kolom pilihan	1825 baris setelah lookup (tambahan data)	Langkah ini mengambil data tambahan dari database berdasar kunci tertentu
<b>Filter Rows</b>	1825 baris dari langkah sebelumnya	1825 baris data yang telah difilter	Langkah ini memfilter baris data berdasarkan kondisi tertentu
<b>Table Output</b>	1825 baris dari langkah sebelumnya	1825 baris yang dikirim ke database atau file tujuan	Langkah ini menyimpan data hasil transformasi kedalam tabel database atau file output lainnya.

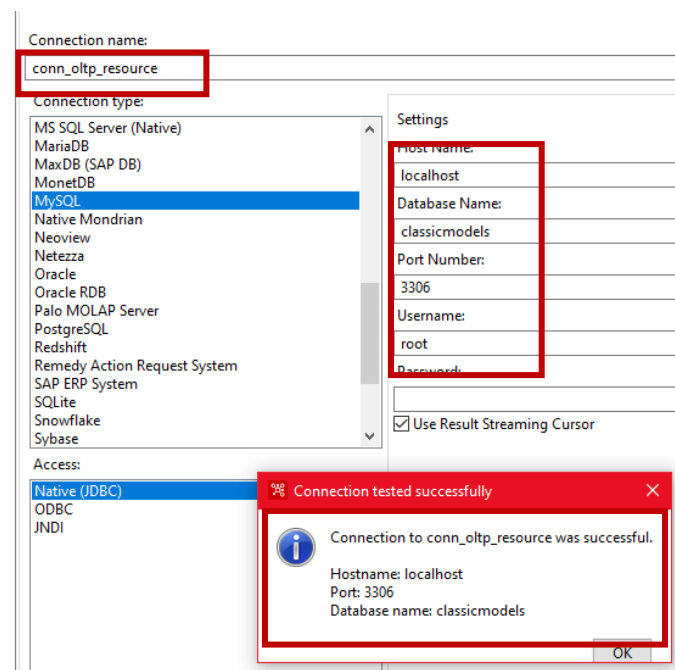


## B. Dimensi Pegawai

1. Buatlah table dimPegawai pada dw\_legendVehicle

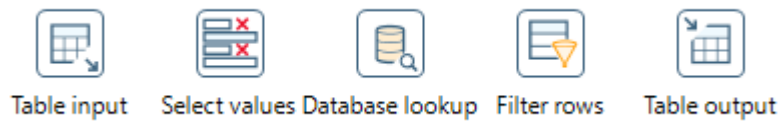
```
mysql> use classicmodels;
Database changed
mysql> CREATE TABLE dimEmployees (
  -> id_dimEmployees INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
  -> employeeNumber INT,
  -> firstName VARCHAR(50),
  -> lastName VARCHAR(50),
  -> jobTitle VARCHAR(50),
  -> boss_firstName VARCHAR(50),
  -> boss_lastName VARCHAR(50),
  -> updated_date TIMESTAMP DEFAULT CURRENT_TIMESTAMP
  -> );
Query OK, 0 rows affected (0.04 sec)
```

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



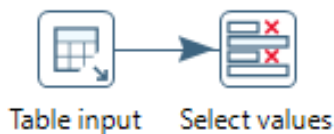
4. Konfigurasi pada table input dengan menghubungkan Connection pada koneksi conn\_oltp\_resources. Untuk mengambil data sumber menggunakan query dibawah ini.

Step name: Table input  
 Connection: conn\_oltp\_resource

SQL

```
SELECT *
FROM employees e
LEFT JOIN employees r
ON e.reportsTo = r.employeeNumber;
```

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.

Step name: Select values

Select & Alter    Remove    Meta-data

Fields:

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

Konfig Select & alter

Select & Alter Remove Meta-data

Fields to remove :

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

Konfig 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.

Step name Database lookup

Connection conn\_oltp\_resource Edit...

Lookup schema classicmodels

Lookup table dimemployees

Enable cache? ☐

Cache size in rows (0=cache) 0

Load all data from table ☐

Use 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

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.

Send 'true' data to step: Table output  
Send 'false' data to step:

The condition:

☐ To edit a subcondition, s

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

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\_oltp\_resource

Target schema classicmodels

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	employeeNumber	StreamEmployeeenumber
2	firstName	StreamFirstName
3	lastName	StreamLastName
4	jobTitle	StreamJobTitle
5	boss_firstName	StreamFirstnameBoss
6	boss_lastName	StreamLastnameBoss

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

	id_dimEmployees	employeeNumber	firstName	lastName	jobTitle	boss_firstName	boss_lastName	updated_date
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	1	1002	Diane	Murphy	President	NULL	NULL	2025-03-11 11:17:05
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	2	1056	Mary	Patterson	VP Sales	Diane	Murphy	2025-03-11 11:17:05
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	3	1076	Jeff	Firrelli	VP Marketing	Diane	Murphy	2025-03-11 11:17:05
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	4	1088	William	Patterson	Sales Manager (APAC)	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	5	1102	Gerard	Bondur	Sale Manager (EMEA)	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	6	1143	Anthony	Bow	Sales Manager (NA)	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	7	1165	Leslie	Jennings	Sales Rep	Anthony	Bow	2025-03-11 11:17:05

## 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.

Execution Results													
Logging Execution History Step Metrics Performance Graph Metrics Preview data													
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	Table input	0	0	23	23	0	0	0	0	0	Finished	0.0s	852
2	Select values	0	23	23	0	0	0	0	0	0	Finished	0.0s	523
3	Database lookup	0	23	23	0	0	0	0	0	0	Finished	0.1s	303
4	Filter rows	0	23	23	0	0	0	0	0	0	Finished	0.1s	295
5	Table output	0	23	23	0	23	0	0	0	0	Finished	0.2s	146

Proses Objek	SS data Input	SS Data Output	Keterangan
Table Input	23 baris	23 baris yang diambil dari database	Langkah ini membaca data dari tabel dalam database sebagai sumber utama utk transform
Select Values	23 baris	23 baris data setelah pemilihan kolom	Langkah ini digunakan untuk kolom yang diperlukan dan membuang kolom yg tdk perlu

<b>Database Lookup</b>	23 baris	23 baris + tambahan info dari database	Langkah ini melakukan pencarian data tambahan dari database
<b>Filter Rows</b>	23 baris	23 baris data yang telah difilter	Langkah ini menyaring data berdasarkan kondisi tertentu
<b>Table Output</b>	23 baris + sudah difilter	23 baris data yang disimpan kedalam tabel database	Langkah ini digunakan untuk menulis data hasil transform dari database

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

➤ Ya, terlihat pada database, dengan firstName Diane muncul 2 kali.

		id_dimEmployees	employeeNumber	firstName	lastName	jobTitle	boss_firstName	boss_lastName	updated_date
<input type="checkbox"/>	Edit Copy Delete		1	1002 Diane	Murphy	President	NULL	NULL	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		2	1056 Mary	Patterson	VP Sales	Diane	Murphy	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		3	1076 Jeff	Firrelli	VP Marketing	Diane	Murphy	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		4	1088 William	Patterson	Sales Manager (APAC)	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		5	1102 Gerard	Bondur	Sale Manager (EMEA)	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		6	1143 Anthony	Bow	Sales Manager (NA)	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		7	1165 Leslie	Jennings	Sales Rep	Anthony	Bow	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		8	1166 Leslie	Thompson	Sales Rep	Anthony	Bow	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		9	1188 Julie	Firrelli	Sales Rep	Anthony	Bow	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		10	1216 Steve	Patterson	Sales Rep	Anthony	Bow	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		11	1286 Foon Yue	Tseng	Sales Rep	Anthony	Bow	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		12	1323 George	Vanauf	Sales Rep	Anthony	Bow	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		13	1337 Loui	Bondur	Sales Rep	Gerard	Bondur	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		14	1370 Gerard	Hernandez	Sales Rep	Gerard	Bondur	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		15	1401 Pamela	Castillo	Sales Rep	Gerard	Bondur	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		16	1501 Larry	Bott	Sales Rep	Gerard	Bondur	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		17	1504 Barry	Jones	Sales Rep	Gerard	Bondur	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		18	1611 Andy	Foxter	Sales Rep	William	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		19	1612 Peter	Marsh	Sales Rep	William	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		20	1619 Tom	King	Sales Rep	William	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		21	1621 Mami	Nishi	Sales Rep	Mary	Patterson	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		22	1625 Yoshimi	Kato	Sales Rep	Mami	Nishi	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		23	1702 Martin	Gerard	Sales Rep	Gerard	Bondur	2025-03-11 11:17:05
<input type="checkbox"/>	Edit Copy Delete		24	1002 Diane	Murphy	President	NULL	NULL	2025-03-11 11:47:29

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

```
INSERT INTO employees (
    employeeNumber, lastName, firstName, extension, email, officeCode, reportsTo, jobTitle
) VALUES (
    8888, 'Dahniar', 'Davina', 'x8888', 'dahniardavina123@gmail.com', 1, 1002, 'Data Scientist');
```

Edit

Copy

Delete

8888

Dahnir

Davina

x8888

dahnirdavina123@gmail.com

1

1002

Data Scientist

#

StreamEmployeeNumber

StreamLastName

StreamFirstName

StreamJobTitle

StreamLastnameBoss

StreamFirstnameBoss

employeeNumber

firstName

lastName

1002

Murphy

Diane

President

<null>

<null>

<null>

<null>

<null>

8888

Dahnir

Davina

Data Scientist

Murphy

Diane

<null>

<null>

<null>

Edit

Copy

Delete

26

8888

Davina

Dahnir

Data Scientist

Diane

Murphy

2025-03-11

15:40:46

### C. Fakta Pembayaran

1. Buatlah sebuah tabel pada database OLAP yang telah dibuat (database dw\_legendVehicle).
  - nama tabel: FactOmset
  - Field: id\_dimEmployees int FK tabel dimEmployee
  - Field: id\_dimDate int FK tabel dimDate
  - Field: amount decimal(10.2)

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0859 seconds.)

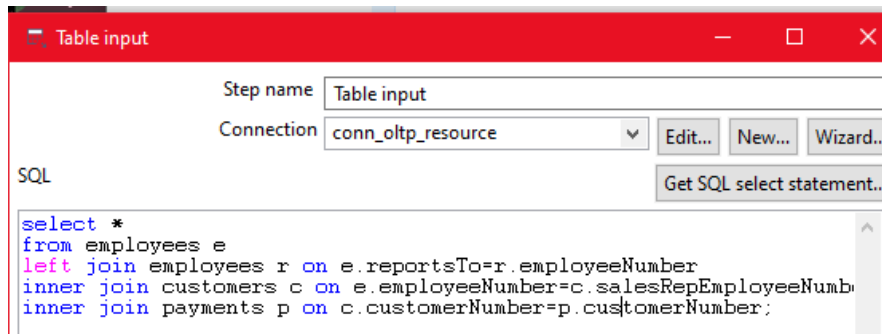
```
CREATE TABLE FactOmsset ( id_dimEmployees INT, id_dimDate INT, amount DECIMAL(10,2),
FOREIGN KEY (id_dimEmployees) REFERENCES dimEmployees(id_dimEmployees), FOREIGN KEY
(id_dimDate) REFERENCES dimDate(id_dimDate) );
```

2. Buat Transformation baru pada PDI Spoon. Gunakan object-object yang dihubungkan sebagai berikut:
- Table Input: Digunakan untuk mengambil data transaksi dari OLTP
  - Select values: digunakan untuk memilih field yang akan digunakan untuk OLAP
  - Database lookup (1) : digunakan untuk mencocokkan data pada tabel dimEmployee untuk mengambil id\_dimEmployee.
  - Database lookup (2) : digunakan untuk mencocokkan data pada tabel dimDate untuk mengambil id\_dimDate.
  - Database lookup (3): digunakan untuk mencocokkan data pada tabel factomset untuk melihat data yang sama atau tidak.
  - Filter rows: digunakan untuk memilih data yang sudah ada pada tabel factomset tidak dimasukkan lagi.

- Table output: digunakan untuk memasukkan data pada tabel factOmset.



3. Konfigurasi pada Table input adalah untuk mendapatkan data dari OLTP sesuai dengan hasil dari query berikut. query tersebut akan menampilkan seluruh isi field dari employee hingga transaksi yang didapat dilihat dari payments yang didapat.



4. Konfigurasi pada Select Values adalah untuk menghapus semua field kecuali employeeNumber, lastname dari employee, firstname dari employee, jobTitle dari employee, lastname dari manager, firstname dari manager, payment date dan amount.

Select & Alter	Remove	Meta-data
Fields:		
#	Fieldname	Rename to
0	jobTitle	streamJobTitle
1	employeeNumber_1	
2	lastName_1	streamLastNameBoss
3	firstName_1	streamFirstNameBoss
4	extension_1	
5	email_1	
6	officeCode_1	
7	reportsTo_1	
8	jobTitle_1	
9	customerNumber	
10	customerName	
11	contactLastName	
12	contactFirstName	
13	phone	
14	addressLine1	
15	addressLine2	
16	city	
17	state	
18	postalCode	
19	country	
20	salesRepEmployeeNumber	
21	creditLimit	
22	customerNumber_1	
23	checkNumber	
24	paymentDate	streamDate
25	amount	streamAmount

Select & Alter	Remove	Meta-data
Fields to remove:		
#	Fieldname	
1	extension	
2	email	
3	officeCode	
4	reportsTo	
5	employeeNumber_1	
6	extension_1	
7	email_1	
8	officeCode_1	
9	reportsTo_1	
10	jobTitle_1	
11	customerNumber	
12	customerName	
13	contactLastName	
14	contactFirstName	
15	phone	
16	addressLine1	
17	addressLine2	
18	city	
19	state	
20	postalCode	
21	country	
22	salesRepEmployee...	
23	creditLimit	
24	customerNumber_1	
25	checkNumber	



5. Konfigurasi pada tabel database lookup adalah dengan melakukan komparasi field stream (output dari proses sebelumnya) dengan field isi data pada tabel dimEmployees. Jika data tersebut cocok maka akan diambil id\_dimEmployees nya.

Step name: Database lookup

Connection: conn\_dw\_destination

Lookup schema: dw\_legendvehicle

Lookup table: dimemployees

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	=	streamFirstName	
6	boss_lastName	=	streamLastName	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimEmployees			

6. Konfigurasi pada tabel database lookup yang kedua adalah dengan melakukan komparasi field stream (output dari proses sebelumnya untuk field date) dengan field isi data pada tabel dimDate. Jika data tersebut cocok maka akan diambil id\_dimDate nya.

Step name: Database lookup 2

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	=	stream...	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimDate			None

7. Konfigurasi pada tabel database lookup yang ketiga adalah dengan melakukan komparasi id\_dimDate dan id\_dimEmployees yang diambil dari proses lookup sebelumnya dengan field isi data pada tabel factOmset. Jika data tersebut cocok data tidak akan dimasukkan dalam tabel factOmset.

Step name: Database lookup 3  
 Connection: conn\_dw\_destination  
 Lookup schema: dw\_legendvehicle  
 Lookup table: factomset  
 Enable cache: ☐  
 Cache size in rows (0=cache all): 0  
 Load all data from table: ☐

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

#	Table field	Comparator	Field1	Field2
1	id_dimDate	=	id_dimDate	
2	id_dimEmployees	=	id_dimEmployees	
3	amount	=	streamAmount	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimDate			None
2	id_dimEmployees			None
3	amount			None

8. Pada bagian filter rows untuk melihat data belum ada pada tabel factOmset dengan melihat bahwa output dari lookup sebelumnya bernilai NULL.

Step name: Filter rows  
 Send 'true' data to step: Table output  
 Send 'false' data to step:

The condition:

☐ ☐ UP ☐ Level 1. Select UP to ,

```

id_dimDate_1 IS NULL
AND
id_dimEmployees_1 IS NULL
AND
amount IS NULL
  
```

9. Konfigurasi terakhir pada table output adalah dengan melakukan mapping data output dari proses sebelumnya ke dalam field pada tabel factOmset.

**Table output**

Step name: Table output

Connection: conn\_dw\_dest [Edit...] [New...] [Wizard...]

Target schema: dw\_legendvehicle [Browse...]

Target table: factomset [Browse...]

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☒

Main options | Database fields

Fields to insert:

#	Table field	Stream field
1	id_dimDate	id_dimDate
2	id_dimEmployees	id_dimEmployees
3	amount	streamAmount

[Get fields] [Enter field mapping]

10. Jalankan proses transformation tersebut.

✓ Showing rows 0 - 272 (273 total, Query took 0.0011 seconds.)

`SELECT * FROM `factomset``

☐ Profiling [Edit inline] [Edit] [Ex]

☐ Show all | Number of rows: 500 ▼

Extra options

id_dimEmployees	id_dimDate	amount
NULL	658	6066.78
NULL	156	14571.44
NULL	718	1676.14
NULL	717	14191.12
NULL	157	32641.98
NULL	598	33347.88
NULL	140	45864.03
NULL	715	82261.22
NULL	151	7565.08
NULL	435	44894.74
NULL	684	19501.82
NULL	586	47924.19

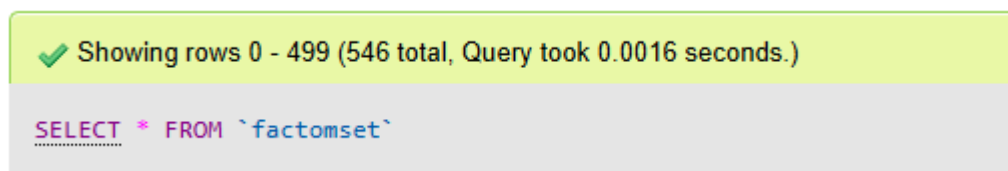
1. Tabel factomset saat ini berisi 273 baris data (seperti terlihat pada "Showing rows 0 - 272 (273 total)")
2. Struktur tabel terdiri dari 3 kolom: id\_dimEmployees, id\_dimDate, dan amount
  - a) Kolom id\_dimEmployees berisi nilai NULL untuk semua baris yang terlihat
  - b) Kolom id\_dimDate berisi nilai numerik (seperti 658, 156, 718, dll)
  - c) Kolom amount berisi nilai numerik dengan desimal (seperti 6066.78, 14571.44, dll)

### ➤ TUGAS 3

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
<b>Table Input</b>	273 baris data dari sumber	273 baris data
<b>Select Values</b>	273 baris data	0 (tidak menghasilkan output langsung)
<b>Database Lookup (1)</b>	273 baris data	0 (tidak menghasilkan output langsung)
<b>Database Lookup (2)</b>	273 baris data	273 baris data
<b>Database Lookup (3)</b>	273 baris data	0 (tidak menghasilkan output langsung)
<b>Filter Rows</b>	273 baris data	0 (tidak menghasilkan output langsung)
<b>Table Output</b>	273 baris data	273 baris data

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



- Ya, akan redundant karena tidak melakukan DELETE atau TRUNCATE pada tabel factomset sebelum dijalankan kembali, kemudian karena tidak adanya constraint unique atau primary key yang diterapkan tabel.

### D. Jobs

1. Buka Jobs pada File - New - Jobs
2. Gunakan 5 objects dan hububngkan sesuai urutan sebagai berikut:
  - Start : Objek untuk melakukan konfigurasi cron job dari proses ETL yag telah dibuat
  - Transformation 1 : digunakan untuk menjalankan transformation pembuatan dimDate.

- Transformation 2 : digunakan untuk menjalankan transformation pembuatan dimEmployees.
- Transformation 3 : digunakan untuk menjalankan transformation pembuatan factOmset.
- Success : Objek untuk menandakan bahwa proses telah selesai.



3. Konfigurasi dilakukan pada setiap transformation untuk mengambil file tranformation yang telah dibuat pada bagian A, B dan C.

a) Transformation 1

The screenshot shows the 'Transformation' configuration window. The 'Entry Name' field is set to 'Transformation 1'. The 'Transformation' field is highlighted with a red box and contains the path 'D:\pentaho\Transformation.js3A.ktr'. Below this, the 'Options' tab is selected. Under 'Run configuration', 'Pentaho local' is chosen. In the 'Execution' section, the checkbox 'Wait for remote transformation to complete' is checked, while the others are unchecked.

b) Transformation 2

The screenshot shows the 'Transformation' configuration window. The 'Entry Name' field is set to 'Transformation 2'. The 'Transformation' field is highlighted with a red box and contains the path 'D:\pentaho\Transformation.js3B.ktr'. The 'Browse...' button is highlighted with a blue box.

c) Transformation 3

**Transformation**

Entry Name:

Transformation 3

---

Transformation:

D:\pentaho\Transformation.js3C.ktr

Browse...

4. Jalankan jobs tersebut ketika semua transformation telah dihubungkan dengan file tranformation yang telah dibuat sebelumnya.

**Execution Results**

Logging History Job metrics Metrics

2025/03/12 15:17:00 - Table input.0 - Finished reading query, closing connection

2025/03/12 15:17:00 - Table input.0 - Finished processing (I=273, O=0, R=0, W=273, U=0, E=0)

2025/03/12 15:17:00 - Select values.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/03/12 15:17:00 - Database lookup.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/03/12 15:17:00 - Database lookup 2.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)

2025/03/12 15:17:00 - Database lookup 3.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/03/12 15:17:00 - Filter rows.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/03/12 15:17:00 - Table output.0 - Finished processing (I=0, O=273, R=273, W=273, U=0, E=0)

2025/03/12 15:17:00 - Job js3 - Starting entry [Success]

2025/03/12 15:17:00 - Job js3 - Finished job entry [Success] (result=[true])

2025/03/12 15:17:00 - Job js3 - Finished job entry [Transformation 3] (result=[true])

2025/03/12 15:17:00 - Job js3 - Finished job entry [Transformation 2] (result=[true])

2025/03/12 15:17:00 - Job js3 - Finished job entry [Transformation 1] (result=[true])

2025/03/12 15:17:00 - Job js3 - Job execution finished

2025/03/12 15:17:00 - Spoon - Job has ended.

➤ **TUGAS 4**

- Buka desain database dari dw\_legendvehicle pada DBMS, bandingkan design tersebut dengan desain db OLTP legendVehicle pada jobsheet 2. analisislah dan ceritakan perbedaannya.

OLAP dw_legendvehicle	OLTP legendvehicle
Digunakan untuk analisis laporan dan analisis tren	Digunakan untuk operasional dan transaksi sehari-hari dan update data
Tabel sederhana menggunakan dimensi dan fakta	Tabel rinci dan kompleks
Ada redudansi untuk mempercepat	Minimal terhadap redudansi
Berbasis multidimensi	Berbasis relasional

2. Buatlah report pertahun untuk KPI "Jumlah omset yang didapat" pada Foon Yue Tseng dan Pamela Castillo. Serta gambarkan grafiknya (grafik garis).

```
1 SELECT
2     CONCAT(e.firstname, ' ', e.lastname) AS name,
3     d.year AS Tahun,
4     SUM(t.amount) AS Total_Omset
5 FROM factomset t
6 JOIN dimemployees e ON t.id_dimEmployees = e.id_dimEmployees
7 JOIN dimdate d ON t.id_dimDate = d.id_dimDate
8 WHERE CONCAT(e.firstname, ' ', e.lastname) IN ('Foon Yue Tseng', 'Pamela
9 Castillo')
10 GROUP BY name, d.year
11 ORDER BY name, Tahun;
```

name	Tahun	Total_Omset
------	-------	-------------

Tidak ada data yang muncul, sehingga tidak ada data yang bisa dibandingkan dan grafik yang muncul

3. Jelaskan perbedaan query saat mendapatkan data pada nomor 2 dengan query pada saat Jobsheet 2!

➤ Query nomor 2

```
1 SELECT
2     CONCAT(e.firstname, ' ', e.lastname) AS name,
3     d.year AS Tahun,
4     SUM(t.amount) AS Total_Omset
5 FROM factomset t
6 JOIN dimemployees e ON t.id_dimEmployees = e.id_dimEmployees
7 JOIN dimdate d ON t.id_dimDate = d.id_dimDate
8 WHERE CONCAT(e.firstname, ' ', e.lastname) IN ('Foon Yue Tseng', 'Pamela
9 Castillo')
10 GROUP BY name, d.year
11 ORDER BY name, Tahun;
```

➤ Query jobsheet 2

```
mysql> SELECT
-> e.employeeNumber AS id_pegawai,
-> CONCAT(e.firstName, ' ', e.lastName) AS Nama_Pegawai,
-> YEAR(p.paymentDate) AS Tahun,
-> SUM(p.amount) AS Total_Omset
-> FROM employees e
-> JOIN customers c ON e.employeeNumber = c.salesRepEmployeeNumber
-> JOIN payments p ON c.customerNumber = p.customerNumber
-> WHERE e.firstName IN ('Foon Yue', 'Pamela')
-> AND e.lastName IN ('Tseng', 'Castillo')
-> GROUP BY e.employeeNumber, YEAR(p.paymentDate)
-> ORDER BY Tahun, Nama_Pegawai;
```

➤ Perbandingan Query no 2 dengan JS 2

Query No 2	Query Jobsheet 2
Mengambil data dari factOmset (data warehouse)	Mengambil data dari employees, customers, orders, orderDetails (database transaksi)
Menganalisis total omzet berdasarkan tahun dan nama karyawan	Menghitung total penjualan berdasarkan jumlah barang yang dipesan dan harga satuan
Bergabung dengan dimEmployees dan dimDate untuk mendapatkan informasi karyawan dan tanggal	Bergabung dengan employees, customers, orders, orderDetails untuk mendapatkan detail transaksi
Menggunakan skema dimensi dan fakta (denormalisasi)	Menggunakan tabel transaksi dengan normalisasi tinggi
Lebih cepat karena data sudah disiapkan dalam bentuk agregat	Lebih lambat karena harus menghitung total dari detail transaksi

4. Simpulkan dengan bahasa sendiri, apa perbedaan OLTP dan OLAP?

- OLTP (Online Transaction Processing) adalah sistem yang dirancang untuk menangani transaksi operasional sehari-hari yang artinya lebih sederhana, cepat dan efisien. Data yang digunakan harus realtime. Sedangkan OLAP (Online Analytical Processing) adalah sistem yang dirancang untuk menganalisis data yang lebih kompleks dan pembuatan laporan.