

LAPORAN HASIL PRAKTIKUM

DATA WAREHOUSE

JOBSHEET 3



Oleh:

DZULFIKAR MUHAMMAD AL GHIFARI

NIM. 2341760071

SIB-2B / 08

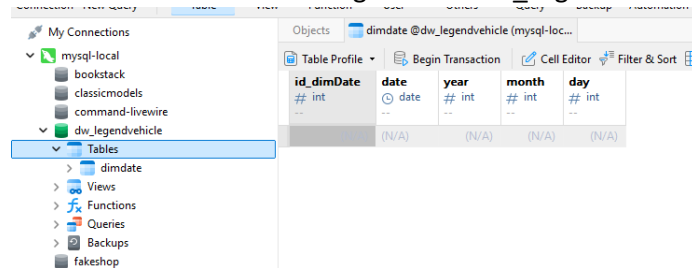
D-IV SISTEM INFORMASI BISNIS

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

A. DIMENSI WAKTU

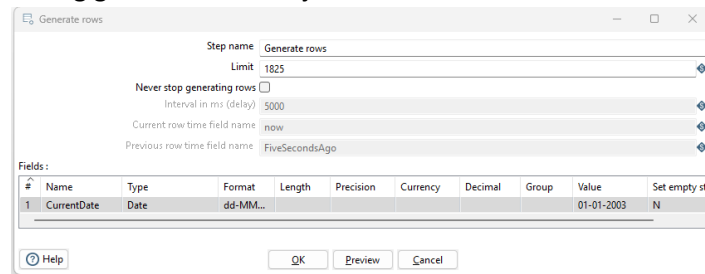
Membuat database OLAP dengan nama dw_LegendVehicle. Dan



Define objek



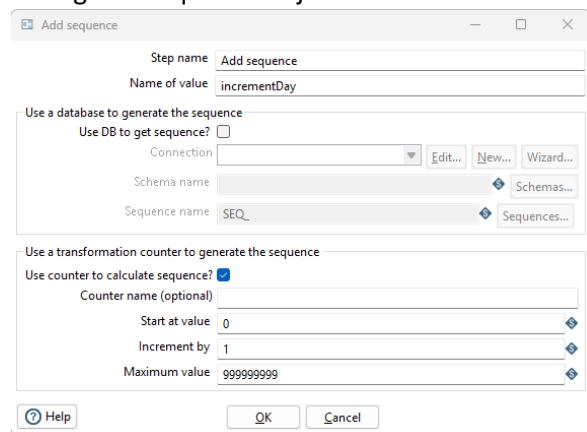
Config generate rows object



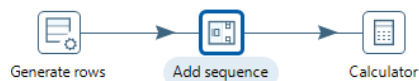
Hubungkan output generate rows -> add sequence



Config add sequence object



Hubungkan output dari add sequences menuju calculator.



Konfigurasi Calculator

Calculator

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 Days	CurrentDate	incrementDay		None			N
2	streamYear	Year of date A	streamDate			None			N
3	streamMonth	Month of date A	streamDate			None			N
4	streamDay	Day of month of d...	streamDate			None			N

Help

Hubungkan output dari calculator menuju Select values



Konfigurasi pada select values

Select values

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			

Get fields to select

Edit Mapping

Include unspecified fields, ordered by name ☐

OK Cancel

Select values

Step name: Select values

Select & Alter | Remove | Meta-data

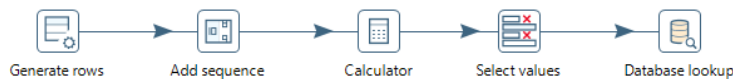
Fields to remove:

#	Fieldname
1	CurrentDate
2	incrementDay

Get fields to remove

OK Cancel

Hubungkan output select values menuju database lookup



Setting db connection & database lookup

Database Connection

General

Connection name: conn_dw_destination

Connection type: MySQL

Settings:

Host Name: localhost

Database Name: dw_legendevehicle

Port Number: 3306

Username: root

Password:

☒ Use Result Streaming Cursor

Test Feature List Explore

Connection tested successfully

Connection to conn_dw_destination was successful.

Hostname: localhost

Port: 3306

Database name: dw_legendevehicle

OK

Database lookup

Step name: Database lookup

Connection: conn_dw_destination

Lookup schema: dw_legendevehicle

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

OK Cancel Help Get Fields Get lookup fields

Hubungkan output dari database lookup dengan filter rows



Setting filter

Filter rows

Step name: **Filter rows**

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

Send 'false' data to step:

The condition:

☐ **date IS NULL**

AND

☐ **year IS NULL**

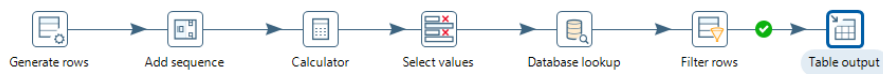
AND

☐ **month IS NULL**

AND

☐ **day IS NULL**

Hubungkan output dari filter rows menuju table output.



Setting table output

Table output

Step name: **Table output**

Connection: **conn_db_destination**

Target schema: **dim_warehouse**

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

Get fields

Enter field mapping

OK Cancel SQL

Result table dimdate

id_dimDate ## int	date date	year ## int	month ## int	day ## int
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
17	2003-01-17	2003	1	17

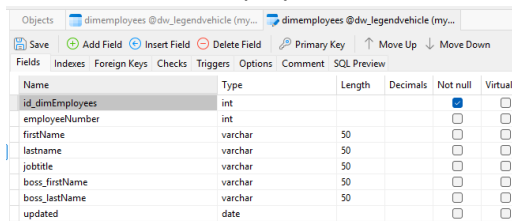
TUGAS 1

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	- (tidak ada input)	1825	Membuat 1825 baris data kosong atau default.
Add Sequences	1825	1825	Menambahkan nomor urut atau sequence ke setiap baris.
Calculator	1825	1825	Melakukan perhitungan berdasarkan nilai tertentu.
Select Values	1825	1825	Memilih kolom tertentu dari dataset.
Database Lookup	1825	1825	Mencari data dari database lain berdasarkan nilai kunci tertentu.
Filter Rows	1825	1825	Menyaring data berdasarkan kondisi tertentu.
Table Output	1825	1825 (ke database)	Menyimpan data hasil transformasi ke dalam tabel database.

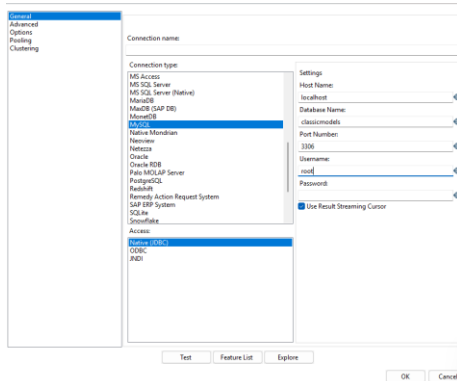
B. DIMENSI PEGAWAI

Create table dimemployees

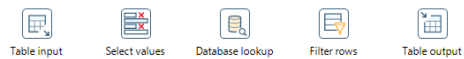


Name	Type	Length	Decimals	Not null	Virtual
id_dimEmployees	int			<input checked="" type="checkbox"/>	<input type="checkbox"/>
employeeNumber	int			<input type="checkbox"/>	<input type="checkbox"/>
firstName	varchar	50		<input type="checkbox"/>	<input type="checkbox"/>
lastName	varchar	50		<input type="checkbox"/>	<input type="checkbox"/>
jobtitle	varchar	50		<input type="checkbox"/>	<input type="checkbox"/>
boss_firstName	varchar	50		<input type="checkbox"/>	<input type="checkbox"/>
boss_lastName	varchar	50		<input type="checkbox"/>	<input type="checkbox"/>
updated	date			<input type="checkbox"/>	<input type="checkbox"/>

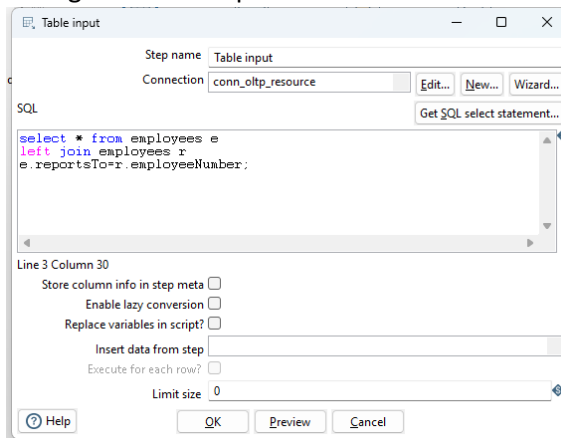
Create new connection for dimemployees



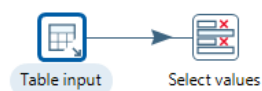
Define object



Konfigurasi table input



Hubungkan output table input pada select values



Konfigurasi pada Select values

Select values - Fields

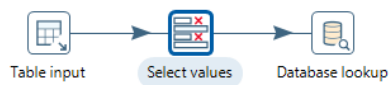
#	Fieldname	Rename to	Length	Precision
1	employeeNumber	StreamEmployeeNumber		
2	lastName	StreamLastname		
3	firstName	StreamFirstname		
4	extension			
5	email			
6	officeCode			
7	reportsTo			
8	jobTitle	StreamJobtitle		
9	employeeNumber_1			
10	lastName_1	StreamLastnameBoss		
11	firstName_1	StreamFirstnameBoss		
12	extension_1			
13	email_1			
14	officeCode_1			
15	reportsTo_1			
16	jobTitle_1			

Include unspecified fields, ordered by name ☐

Select values - 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

Hubungkan output select values pada database lookup



Konfigurasi pada database lookup

Database lookup

Step name: Database lookup

Connection: conn_dw_destination

Look up schema: dw_legendvehicle

Look up table: dimemployees

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the values:

#	Table field	Co...	Field1	Field2
1	employeeNumber	=	StreamEmployeeNumber	
2	firstName	=	StreamFirstname	
3	lastName	=	StreamLastname	
4	jobtitle	=	StreamJobtitle	
5	boss_firstname	=	StreamFirstnameBoss	
6	boss_lastname	=	StreamLastnameBoss	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	employeeNumber	employeeNumber		Integer
2	firstName	firstName		String
3	lastName	lastName		String
4	jobtitle	jobtitle		String
5	boss_firstname	boss_firstname		String
6	boss_lastname	boss_lastname		String

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

Hubungkan output database lookup dengan filter rows.

Filter rows

Step name: Filter rows

Send 'true' data to step:

Send 'false' data to step:

The condition:

☐ employeeNumber IS NULL

AND

☐ firstName IS NULL

AND

☐ lastName IS NULL

AND

☐ jobTitle IS NULL

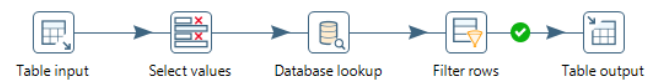
AND

☐ boss_firstname IS NULL

AND

☐ boss_lastname IS NULL

Hubungkan output dari filter rows dengan table output.



Konfigurasi table output

Table output

Step name: **Table output**

Connection: **conn_dw_destination** [Edit... New... Wizard...]

Target schema: **dw_legendvehicle** [Browse...]

Target table: **dimemployees** [Browse...]

Commit size: **1000**

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☒

Main options - Database fields

Partition data over tables: ☐

Partitioning field: []

Partition data per month: ☐

Partition data per day: ☐

Use batch update for inserts: ☒

Is the name of the table defined in a field?: ☐

Field that contains name of table: []

Store the tablename field: ☐

Return auto-generated key: ☐

Name of auto-generated key field: []

[Help] [OK] [Cancel] [SQL]

Result table dimemployees

id_dimEmployees	employeeNumber	firstName	lastName	jobtitle	boss_firstName	boss_lastName	update
# int	# int	# varchar(50)	# varchar(50)	# varchar(50)	# varchar(50)	# varchar(50)	date
1	1002	Diane	Murphy	President	(Null)	(Null)	(Null)
2	1056	Mary	Patterson	VP Sales	Diane	Murphy	(Null)
3	1076	Jeff	Firrelli	VP Marketing	Diane	Murphy	(Null)
4	1088	William	Patterson	Sales Manager (A	Mary	Patterson	(Null)
5	1102	Gerard	Bondur	Sale Manager (E)	Mary	Patterson	(Null)
6	1143	Anthony	Bow	Sales Manager (h	Mary	Patterson	(Null)
7	1165	Leslie	Jennings	Sales Rep	Anthony	Bow	(Null)
8	1166	Leslie	Thompson	Sales Rep	Anthony	Bow	(Null)
9	1188	Julie	Firrelli	Sales Rep	Anthony	Bow	(Null)
10	1216	Steve	Patterson	Sales Rep	Anthony	Bow	(Null)
11	1286	Foon Yue	Tseng	Sales Rep	Anthony	Bow	(Null)
12	1323	George	Vanauf	Sales Rep	Anthony	Bow	(Null)
13	1337	Loui	Bondur	Sales Rep	Gerard	Bondur	(Null)
14	1370	Gerard	Bondur	Sales Rep	Gerard	Bondur	(Null)

TUGAS 2

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

Proses Objek	SS Data Input	SS Data Output	Keterangan
Table Input	23	0	Membaca 23 data dari database, lalu mengirimkannya ke proses selanjutnya.
Select Values	23	23	Memilih atau mengubah nama kolom tertentu. Tidak ada perubahan jumlah data.
Database Lookup	23	23	Mencari referensi data dari database lain, tetapi tidak ada perubahan jumlah record.
Filter Rows	23	23	Memfilter data berdasarkan kondisi tertentu, tetapi di sini tidak ada data yang difilter.
Table Output	23	23	Menulis hasil transformasi ke database tujuan. Semua 23 record berhasil diproses.

2. Jika proses itu di ulangi (di run kembali) apakah data akan redudant?
Tidak, tidak ada data redundant Ketika di run kembali
3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini.
Amati hasilnya, apa yang terjadi?

1621	Nishi	Mami	x101	mnishi@classicmodelcars.	5	1056	Sales Rep
1625	Kato	Yoshimi	x102	ykato@classicmodelcars.c	5	1621	Sales Rep
1702	Gerard	Martin	x2312	mgerard@classicmodelcar	4	1102	Sales Rep
1788	Muhammad	Dzulfikar	x999	dzul@gmail.com	1	1002	VP Backend Dev

1	1002	Diane	Murphy	President	(Null)	(Null)	(Null)
2	1056	Mary	Patterson	VP Sales	Diane	Murphy	(Null)
3	1076	Jeff	Firrelli	VP Marketing	Diane	Murphy	(Null)
4	1088	William	Patterson	Sales Manager (A	Mary	Patterson	(Null)
5	1102	Gerard	Bondur	Sale Manager (E)	Mary	Patterson	(Null)
6	1143	Anthony	Bow	Sales Manager (B	Mary	Patterson	(Null)
7	1165	Leslie	Jennings	Sales Rep	Anthony	Bow	(Null)
8	1166	Leslie	Thompson	Sales Rep	Anthony	Bow	(Null)
9	1188	Julie	Firrelli	Sales Rep	Anthony	Bow	(Null)
10	1216	Steve	Patterson	Sales Rep	Anthony	Bow	(Null)
11	1286	Foon Yue	Tseng	Sales Rep	Anthony	Bow	(Null)
12	1323	George	Vanauf	Sales Rep	Anthony	Bow	(Null)
13	1337	Loui	Bondur	Sales Rep	Gerard	Bondur	(Null)
14	1370	Gerard	Hernandez	Sales Rep	Gerard	Bondur	(Null)
15	1401	Pamela	Castillo	Sales Rep	Gerard	Bondur	(Null)
16	1501	Larry	Bott	Sales Rep	Gerard	Bondur	(Null)
17	1504	Barry	Jones	Sales Rep	Gerard	Bondur	(Null)
18	1611	Andy	Fuiter	Sales Rep	William	Patterson	(Null)
19	1612	Peter	Marsh	Sales Rep	William	Patterson	(Null)
20	1619	Tom	King	Sales Rep	William	Patterson	(Null)
21	1621	Mami	Nishi	Sales Rep	Mary	Patterson	(Null)
22	1625	Yoshimi	Kato	Sales Rep	Mami	Nishi	(Null)
23	1702	Martin	Gerard	Sales Rep	Gerard	Bondur	(Null)
24	1002	Diane	Murphy	President	(Null)	(Null)	(Null)
25	1002	Diane	Murphy	President	(Null)	(Null)	(Null)
26	1788	Dzulfikar	Muhammad	VP Backend Dev	Diane	Murphy	(Null)

Data dengan nama saya akan tertambahkan pada table dimemployees

C. FAKTA PEMBAYARAN

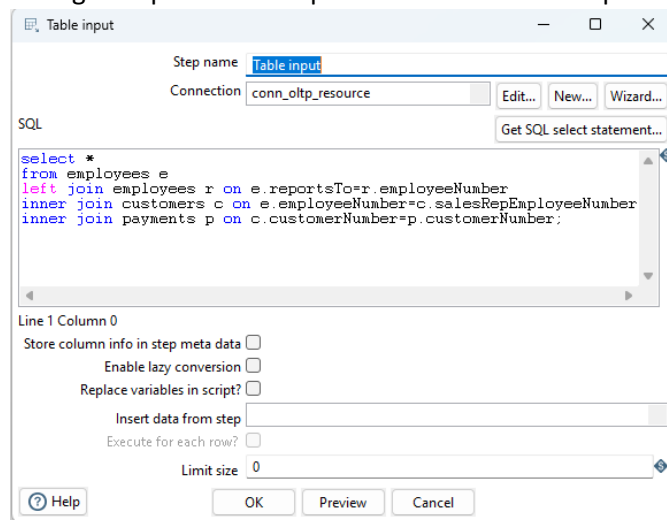
Buat table factOmset

```
mysql-local dw_legendvehicle Run Stop Explain
1 CREATE TABLE factomset (
2   id_factomset INT AUTO_INCREMENT PRIMARY KEY,
3   id_dimEmployees INT,
4   id_dimDate INT,
5   amount DECIMAL(10,2),
6   FOREIGN KEY (id_dimEmployees) REFERENCES dimemployees(id_dimEmployees),
7   FOREIGN KEY (id_dimDate) REFERENCES dimdate(id_dimDate)
8 );
```

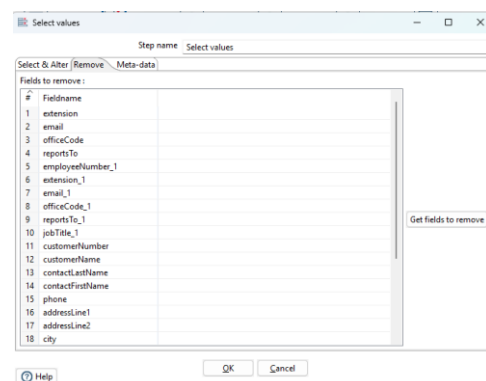
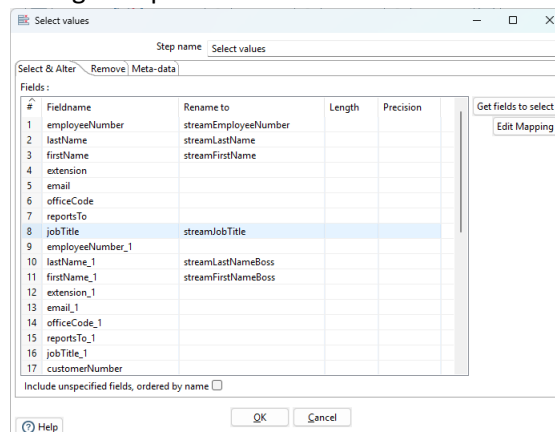
Buat Transformation baru pada PDI Spoon.



Konfigurasi pada Table input adalah untuk mendapatkan data dari OLTP



Konfigurasi pada Select Values



Konfigurasi pada tabel database lookup

Database lookup

Step name: Database lookup

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

Lookup schema: dw_legendvehicle [Browse...]

Lookup table: dimemployees [Browse...]

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

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

#	Table field	Comparator	Field1	Field2
1	employeeNumber	=	streamEmployeeNumber	
2	firstName	=	streamFirstName	
3	lastName	=	streamLastName	
4	jobtitle	=	streamJobTitle	
5	boss_firstname	=	streamFirstNameBoss	
6	boss_lastname	=	streamLastNameBoss	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimEmployees			None

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

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

Konfigurasi pada tabel database lookup yang kedua

Database lookup

Step name: Database lookup 2

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

Lookup schema: dw_legendvehicle [Browse...]

Lookup table: dimdate [Browse...]

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

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

#	Table field	Com...	Field1	Field2
1	date	=	streamDate	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimDate			None

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

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

Konfigurasi pada tabel database lookup yang ketiga

Step name: Database lookup 3

Connection: com_dw_destination

Lookup schema: dw_legendvehicle

Lookup table: factomset

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	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

Do not pass the row if the lookup fails: ☐

Fail on multiple results?: ☐

Order by:

Buttons: Help, OK, Cancel, Get Fields, Get lookup fields

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:

☐ id_DimDate_1 IS NULL

AND

☐ id_DimEmployees_1 IS NULL

AND

☐ amount IS NULL

Buttons: Help, OK, Cancel, SQL

Konfigurasi tabel output

Step name: Table output

Connection: com_dw_destination

Target schema: dw_legendvehicle

Target table: factomset

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☒

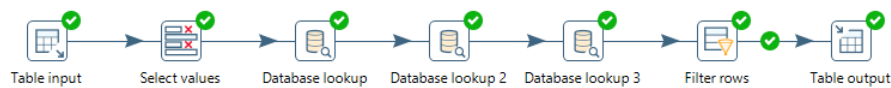
Main options | Database fields

Fields to insert:

#	Table field	Stream field
1	id_DimDate	id_dimDate
2	id_DimEmp...	id_dimEmplo...
3	amount	streamAmou...

Buttons: Help, OK, Cancel, SQL

Run transformation



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 (I)	SS Data Output (O)	Keterangan
Table Input	273	0	Membaca data dari database sumber dan meneruskannya ke proses berikutnya.
Select values	0	0	Memilih kolom yang diperlukan tanpa mengubah jumlah data.
Database lookup (1)	273	0	Lookup pertama untuk mengambil data tambahan dari tabel referensi.
Database lookup (2)	273	0	Lookup kedua untuk mendapatkan informasi tambahan.
Database lookup (3)	0	0	Tidak ada input langsung, tetapi tetap meneruskan 273 baris data. Ini bisa terjadi jika lookup hanya digunakan sebagai referensi tanpa input eksplisit.
Filter rows	0	0	Menyaring data sesuai kondisi tertentu tanpa mengurangi jumlah data.
Table Output	0	273	Menyimpan data hasil transformasi ke database tujuan.

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

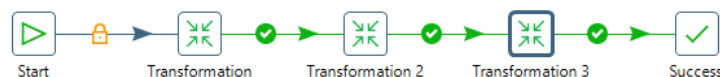
Tidak, tidak ada redundansi data

D. FAKTA PEMBAYARAN

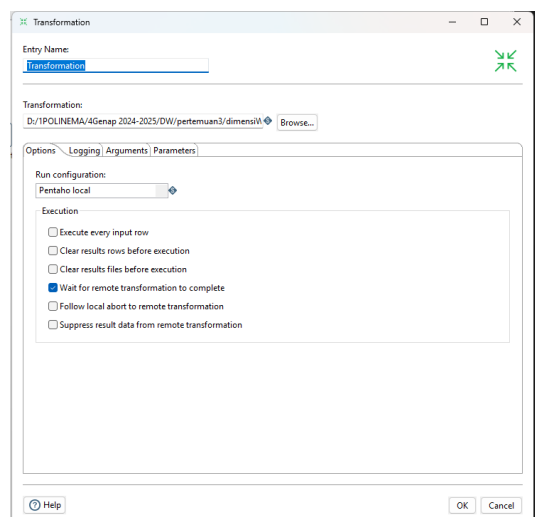
Buka Jobs pada File - New - Jobs

Gunakan 5 objects dan hububngkan sesuai urutan sebagai berikut:

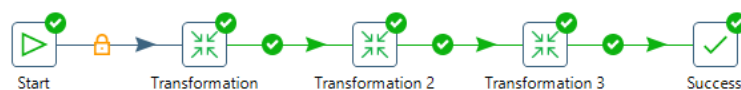
- Start: Objek untuk melakukan konfigurasi cron job dari proses ETL yang telah dibuat
- Transformation 1 : digunakan untuk menjalankan transformation pembuatan dimDate.
- Transformation 2 : digunakan untuk menjalankan transformation pembuatan dimEmployees.
- Transformation 3 : digunakan untuk menjalankan transformation pembuatan factOmset.
- Success: Objek untuk menandakan bahwa proses telah selesai.



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



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



TUGAS 4

1. Buka desain database dari dw_legendvehicle pada DBMS, bandingkan design tersebut dengan desain db OLTP legendVehicle pada jobsheet 2. analisislah dan ceritakan perbedaannya.

Database OLTP legendVehicle dirancang untuk menangani transaksi harian dengan struktur tabel yang sangat ternormalisasi, memastikan kecepatan dan integritas dalam operasi CRUD,

sementara data warehouse dw_legendvehicle menggunakan denormalisasi dengan skema seperti star schema untuk analisis historis dan pembuatan laporan strategis.

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

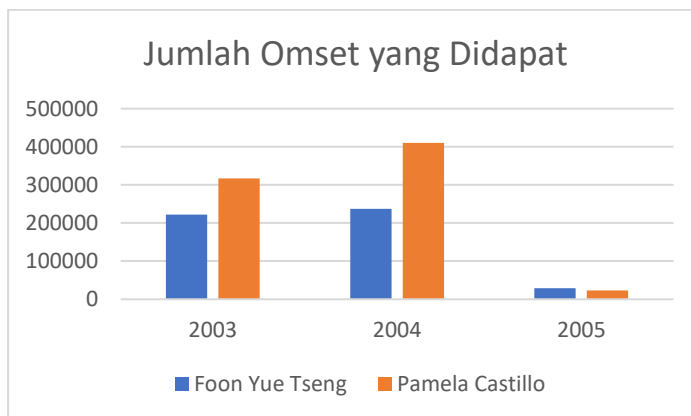
```

1 SELECT
2   CONCAT(e.firstName, ' ', e.lastName) AS Nama,
3   d.year AS Tahun,
4   SUM(f.amount) AS Total_Omset
5 FROM factomset f
6 JOIN dimemployees e ON f.id_dimEmployees = e.id_dimEmployees
7 JOIN dimdate d ON f.id_dimDate = d.id_dimDate
8 WHERE CONCAT(e.firstName, ' ', e.lastName) IN ('Foon Yue Tseng', 'Pamela Castillo')
9 GROUP BY Nama, Tahun;
10 ORDER BY Nama, Tahun;
11

```

Nama	Tahun	Total_Omset
Foon Yue Tseng	2003	221887,03
Foon Yue Tseng	2004	237255,26
Foon Yue Tseng	2005	29070,38
Pamela Castillo	2003	317104,78
Pamela Castillo	2004	409910,07
Pamela Castillo	2005	23187,02

Nama	2003	2004	2005
Foon Yue Tseng	221887,03	237255,3	29070,38
Pamela Castillo	317104,78	409910,1	23187,02



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

Jobsheet 2 terletak pada tujuan dan struktur datanya. Kueri di Jobsheet 2 bekerja dalam OLTP, yang fokus pada transaksi harian dengan tabel sangat ternormalisasi untuk operasi cepat dan akurat.

kueri pada nomor 2 menggunakan OLAP, yang dirancang untuk analisis data historis dengan tabel yang didenormalisasi, serta menggunakan agregasi seperti *SUM* dan *GROUP BY* untuk mendukung pengambilan keputusan bisnis.

OLTP mengutamakan efisiensi transaksi, sementara OLAP dioptimalkan untuk analisis skala besar.

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

OLTP berfungsi untuk memproses transaksi harian secara cepat dan akurat, seperti pencatatan penjualan atau transaksi perbankan.

sedangkan, OLAP berfungsi untuk menganalisis data historis dengan agregasi dan pelaporan, membantu dalam pengambilan keputusan bisnis.