

C. Fakta Pembayaran

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 0.0014 seconds.)

CREATE TABLE FactOmset ( 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) );

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

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.



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.

```
select *
from employees e
left join employees r on e.reportsTo=r.employeeNumber
inner join customers c on e.employeeNumber=c.salesRepEmployeeNumber
inner join payments p on c.customerNumber=p.customerNumber;
```

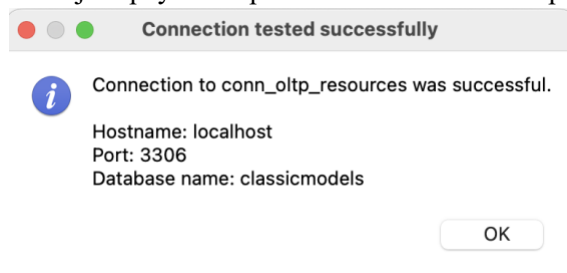


Table input

Step name: Table input

Connection: conn_oltp_resources Edit... New... Wizard...

SQL

Get SQL select statement...

```

select *
from employees e
left join employees r on e.reportsTo=r.employeeNumber
inner join customers c on e.employeeNumber=c.salesRepEmployeeNumber
inner join payments p on c.customerNumber=p.customerNumber;

```

Line 5 Column 59

Store column info in step meta data ☐

Enable lazy conversion ☐

Replace variables in script? ☐

Insert data from step

Execute for each row? ☐

Limit size 0

Help OK Preview Cancel

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.

Gambar kolom yang ditampilkan pada select & alter tab (1)

Select values

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			
32	paymentDate	streamDate		
33	amount	streamAmount		

Get fields to select Edit Mapping

Include unspecified fields, ordered by nar ☐

Help OK Cancel

Gambar kolom yang dihapus pada remove tab



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.

Connection tested successfully

Connection to conn_dw_destination was successful.

Hostname: localhost
Port: 3306
Database name: dw_LegendVehicle

OK

Database lookup

Step name

Connection Edit... New... Wizard...

Lookup schema Browse...

Lookup table Browse...

Enable cache? ☐

Cache size in rows (0=cache everything)

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_firstName	=	streamLastNameBoss	

Values to return from the lookup table :

#	Field	New name	Default	Type
1	id_dimEmployees			none
2				
3				

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by

OK
Cancel
Get Fields
Get lookup fields

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.

Database lookup

Step name: Database lookup 2

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

Lookup schema: dw_LegendVehicle [Browse...]

Lookup table: dimDate [Browse...]

Enable cache? ☐

Cache size in rows (0=cache everything): 0

Load all data from table ☐

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

#	Table field	Comparator	Field1	Field2
1	date	=	streamDate	
2				

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimDate			None
2				

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 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.

Database lookup

Step name: Database lookup 3

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

Lookup schema: dw_LegendVehicle [Browse...]

Lookup table: FactOmset [Browse...]

Enable cache? ☐

Cache size in rows (0=cache everything): 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	
4				

Values to return from the lookup table:

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

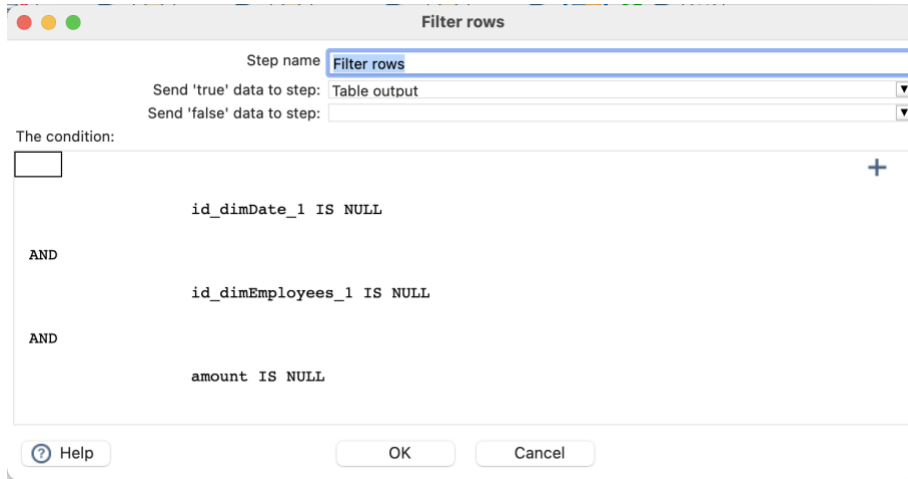
Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

[?] 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.



The dialog box titled "Filter rows" is used to define a filter condition. It includes a "Step name" field set to "Filter rows", and dropdowns for "Send 'true' data to step:" (set to "Table output") and "Send 'false' data to step:". The "The condition:" section contains a list of conditions: "id_dimDate_1 IS NULL", "AND", "id_dimEmployees_1 IS NULL", "AND", and "amount IS NULL". At the bottom are "Help", "OK", and "Cancel" buttons.

Step name:

Send 'true' data to step:

Send 'false' data to step:

The condition:

☐

id_dimDate_1 IS NULL

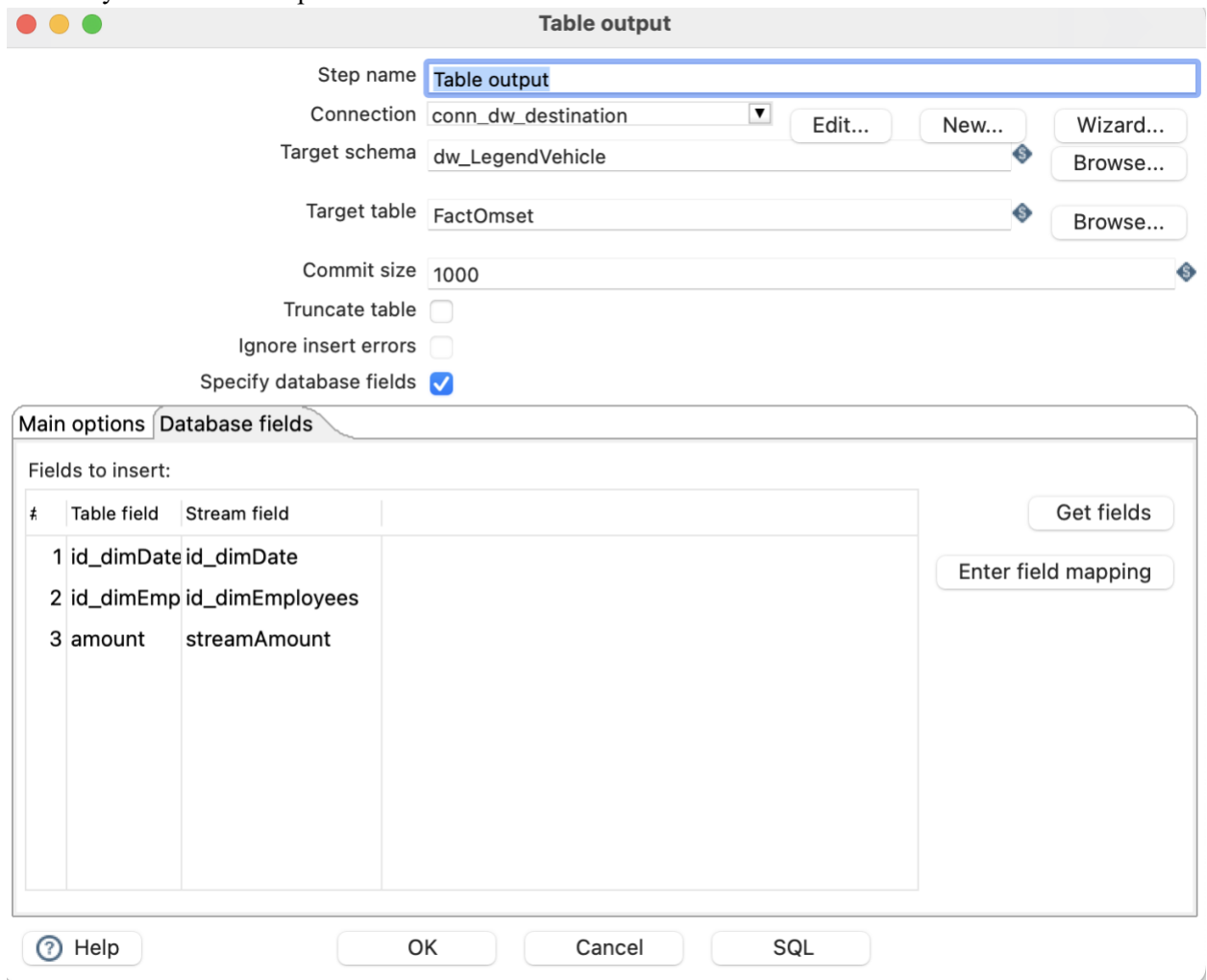
AND

id_dimEmployees_1 IS NULL

AND

amount IS NULL

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



The "Table output" dialog box is used to configure the output of a table. It includes fields for "Step name" (set to "Table output"), "Connection" (set to "conn_dw_destination"), "Target schema" (set to "dw_LegendVehicle"), and "Target table" (set to "FactOmset"). It also has a "Commit size" field (set to "1000") and checkboxes for "Truncate table", "Ignore insert errors", and "Specify database fields" (checked). Below these are tabs for "Main options" and "Database fields". The "Database fields" tab is active, showing a table with "Fields to insert:" and columns for "#", "Table field", and "Stream field". The table contains three rows: 1. id_dimDate, id_dimDate; 2. id_dimEmp, id_dimEmployees; 3. amount, streamAmount. To the right of the table are buttons for "Get fields" and "Enter field mapping". At the bottom are "Help", "OK", "Cancel", and "SQL" buttons.

Step name:

Connection:

Target schema:

Target table:

Commit size:

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_dimEmployees
3	amount	streamAmount

Spoon - Transformation 4

Connect

View

Transformations

- Transformations
- Run
- Data
- Step
- Hop
- Parti
- Slav
- Kett
- VFS

1 input → Select values → Database lookup → Database lookup 2 → Database lookup 3 → Filter rows → Table output

Execution Results

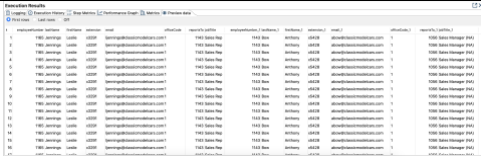
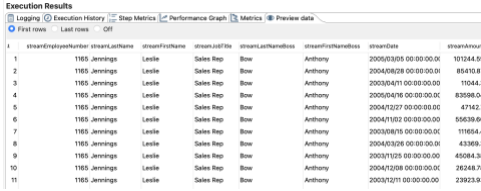
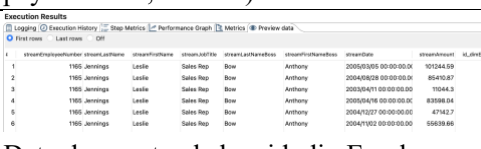
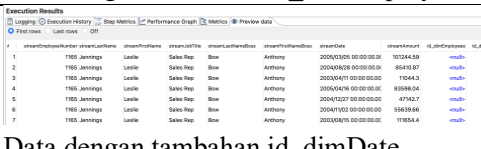
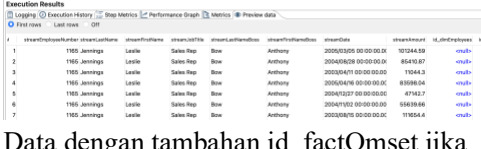
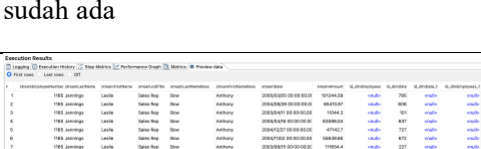

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2025-04-16 14:52:16.865 - Spoon - Transformation opened.
 2025-04-16 14:52:16.866 - Spoon - Launching transformation [Transformation 4]...
 2025-04-16 14:52:16.866 - Spoon - Started the transformation execution.
 2025-04-16 14:52:16.899 - Transformation 4 - Dispatching started for transformation [Transformation 4]
 2025-04-16 14:52:16.902 - Table output.0 - Connected to database [conn_dw_destination] (commit=1000)
 2025-04-16 14:52:16.940 - Table input.0 - Finished reading query, closing connection
 2025-04-16 14:52:16.941 - Table input.0 - Finished processing (I=273, O=0, R=0, W=273, U=0, E=0)
 2025-04-16 14:52:16.943 - Select values.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
 2025-04-16 14:52:16.986 - Database lookup.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
 2025-04-16 14:52:17.006 - Database lookup.2.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)
 2025-04-16 14:52:17.008 - Database lookup.3.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
 2025-04-16 14:52:17.011 - Filter rows.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
 2025-04-16 14:52:17.039 - Table output.0 - Finished processing (I=0, O=273, R=273, W=273, U=0, E=0)
 2025-04-16 14:52:17.040 - Spoon - The transformation has finished!!

id_dimEmployees	id_dimDate	amount
NULL	795	101244.59
NULL	606	85410.87
NULL	101	11044.30
NULL	837	83598.04
NULL	727	47142.70
NULL	672	55639.66
NULL	227	111654.40
NULL	451	43369.30
NULL	329	45084.38
NULL	708	26248.78
NULL	345	23923.93
NULL	99	16537.85
NULL	684	2434.25
NULL	322	50743.65
NULL	764	12692.19
NULL	217	38675.13
NULL	307	85559.12
NULL	805	46781.66
NULL	538	59551.38
NULL	271	29997.09
NULL	425	12573.28
NULL	717	14191.12
NULL	157	32641.98
NULL	598	33347.88
NULL	338	3879.96

Tugas 3

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.

No	Proses Objek	SS Data Input	SS Data Output	Keterangan
1	Table Input	Data Tansaksi dari OLTP (termasuk employee, manager, customer, dan payments)		Mengambil semua data transaksi dari database OLTP
2	Select Values	Semua field hasil query	 Hanya field yang dipilih (employeeNumber, lasname, firsnme, jobtitle, manager, paymentdate, amount)	Menghapus field yang tidak dibutuhkan
3	Database Lookup 1	Data dari select values	 Data dengan tambahan id_dimEmployees	Mencari id berdasarkan employeeNumber
4	Database Lookup 2	Mencari id_dimEmployees berdasarkan employeeNumber	 Data dengan tambahan id_dimDate	Mencari id_dimDate berdasarkan paymentDate
5	Database Lookup 3	Data dari Database lookup 2	 Data dengan tambahan id_factOmset jika sudah ada	Mengecek apakah kombinasi id_dimEmployees dan id_dimDate sudah ada di factOmset
6	Filter Rows	Data dari Database Lookup 3	 Data yang belum ada dalam factOmset	Menyaring data yang sudah ada agar tidak dimasukan lagi
7	Table Output	Data hasil filter	 Data masuk ke table factOmset	Memasukan data bar uke dalam table OLAP

Jika proses itu di ulangi (di run Kembali) apakah data akan redundant?

> Tidak karena,

- Database lookup 3 berfungsi untuk cek kombinasi id_dimEmployees dan id_dimDate sudah ada di factOmset.

- Filter rows memastikan hanya data baru yang akan masuk ke dalam table factOmset

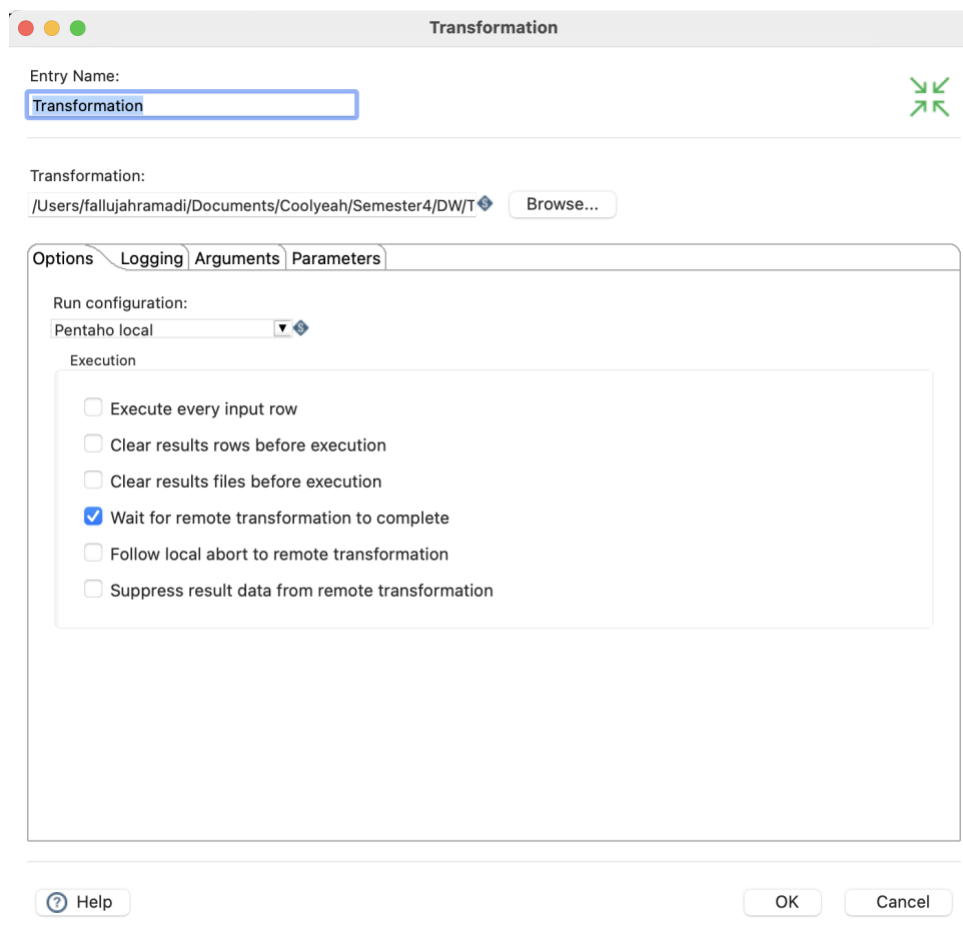
- Jika transformasi dijalankan ulang, data yang sudah ada tidak akan dimasukkan Kembali, sehingga tidak ada duplikasi data

D. Jobs

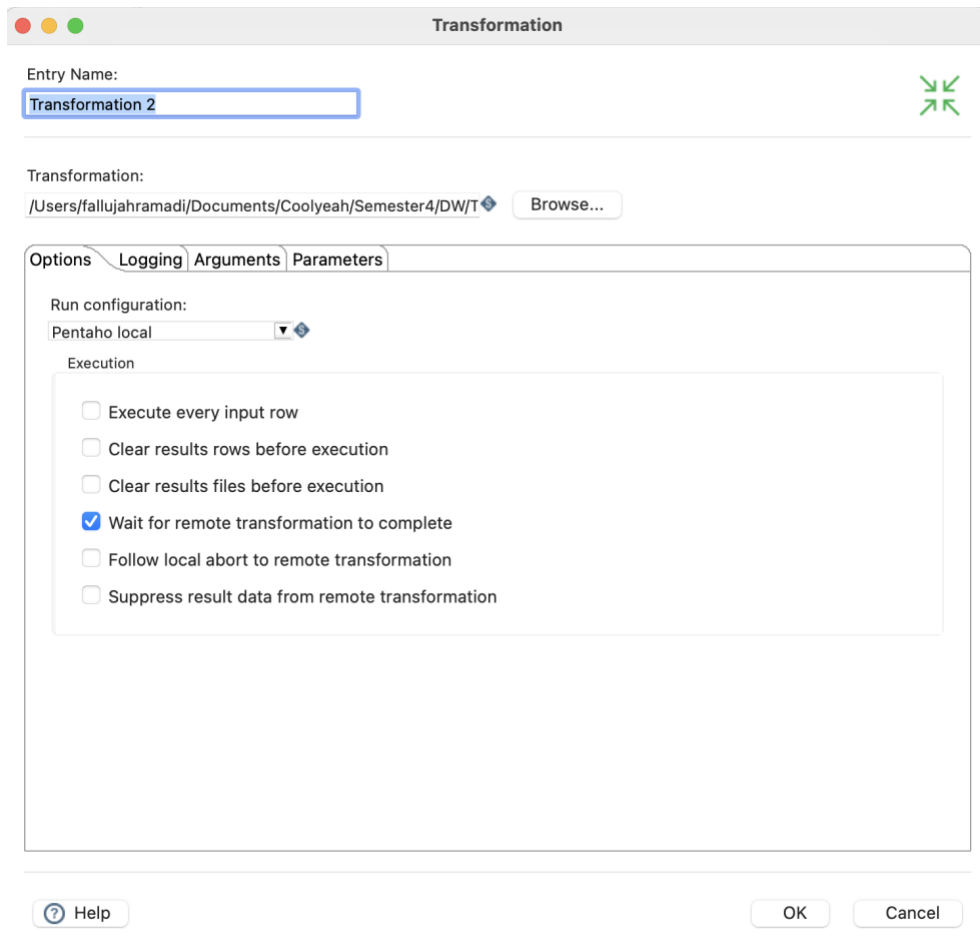
Buka Jobs pada File - New - Jobs

Gunakan 5 objects dan hububngkan sesuai urutan sebagai berikut:

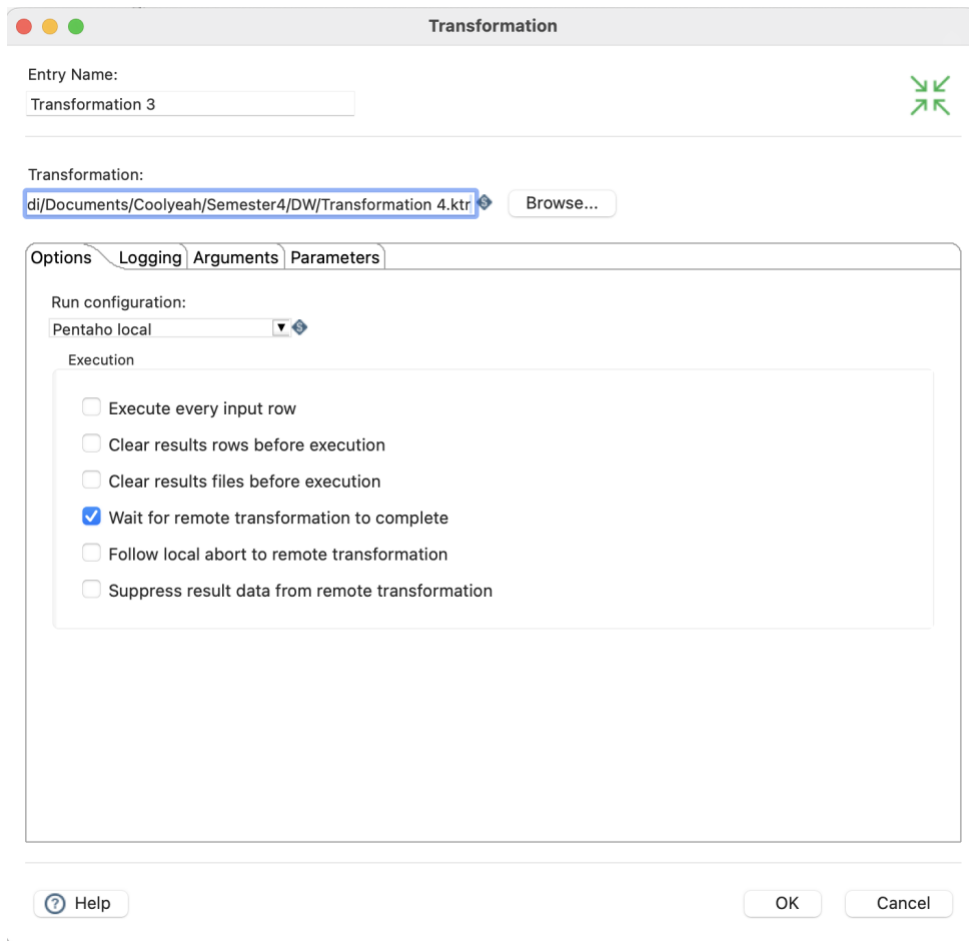
- Start: Objek untuk melakukan konfigurasi cron job dari proses ETL yag telah dibuat
- Transformation 1 : digunakan untuk menjalankan transformation pembuatan dimDate.
- Transformation 2 : digunakan untuk menjalankan transformation pembuatan dimEmployees.
- Transformation 3 : digunakan untuk menjalankan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



Gambar konfigurasi pada bagian Transformation A

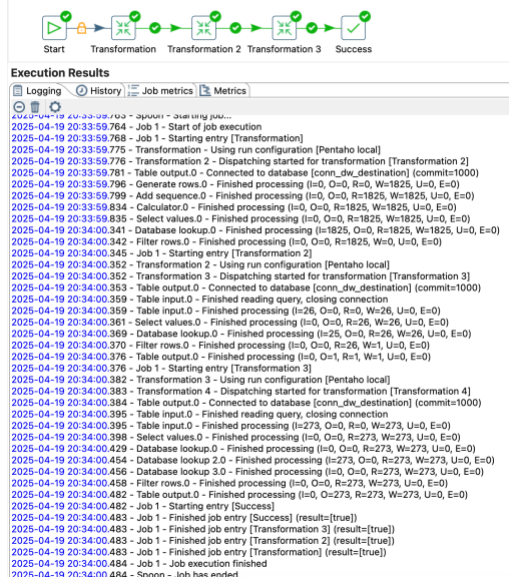


Gambar konfigurasi pada bagian Transformation B



Gambar konfigurasi pada bagian Transformation C

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



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.

No	Aspek	OLTP (Operational Database)	OLAP (Data Warehouse)
1	Tujuan	Digunakan untuk transaksi harian seperti penjualan, update data, dll	Digunakan untuk analisis data dalam jangka Panjang
2	Struktur Data	Normalisasi tinggi (banyak table dengan relasi) untuk efisiensi penyimpanan	Denormalisasi (mengurangi jumlah join) agar query lebih cepat.
3	Contoh Tabel	Customers, orders, vehicles	dimDate, dimEmployees, factOmset
4	Performansi	Dioptimalkan untuk transaksi cepat	Dioptimalkan untuk query analitik dan laporan

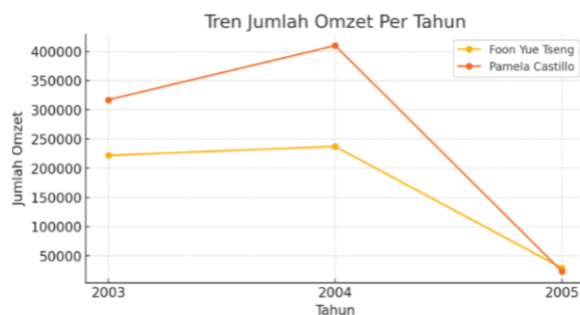
Buatlah report pertahun untuk KPI "Jumlah omset yang didapat" pada Foon Yue Tseng dan Pamela Castillo. Serta gambarkan grafiknya (grafik garis).

MySQL returned an empty result set (i.e. zero rows). (Query took 0.0032 seconds.)

```
SELECT e.firstName AS Nama, SUM(CASE WHEN d.year = 2003 THEN f.amount ELSE 0 END) AS '2003', SUM(CASE WHEN d.year = 2004 THEN f.amount ELSE 0 END) AS '2004', SUM(CASE WHEN d.year = 2005 THEN f.amount ELSE 0 END) AS '2005' FROM factOmset f JOIN dimEmployees e ON f.id_dimEmployees = e.id_dimEmployees JOIN dimDate d ON f.id_dimDate = d.id_dimDate WHERE (e.firstName = 'Foon Yue' AND e.lastName = 'Tseng') OR (e.firstName = 'Pamela' AND e.lastName = 'Castillo') GROUP BY e.firstName ORDER BY e.firstName;
```

☐ Profiling ☐ Edit inline ☐ Edit ☐ Explain SQL ☐ Create PHP code ☐ Refresh

Nama	2003	2004	2005
Foon Yue	221887.03	237255.26	29070.38
Pamela	317104.78	409910.07	23187.02



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

- query di jobsheet 2 (OLTP) ; menggunakan table transaksi langsung, misalnya orders, payments, dll
- query di tugas (OLAP) ; menggunakan dimensional modelling, yaitu factOmset, dimEmployees, dan dimDate, sehingga lebih fokus ke agresi. Data.

Simpulkan dengan bahasa sendiri, apa perbedaan OLTP dan OLAP?

- OLTP -> database untuk transaksi sehari hari, cepat dalam pengoperasian CRUD
- OLAP -> database untuk analisis dan pelaporan, berfokus pada agresi dan performa query.