

## JOBSHEET IV

### MULTI RESOURCES

#### A. STUDI KASUS

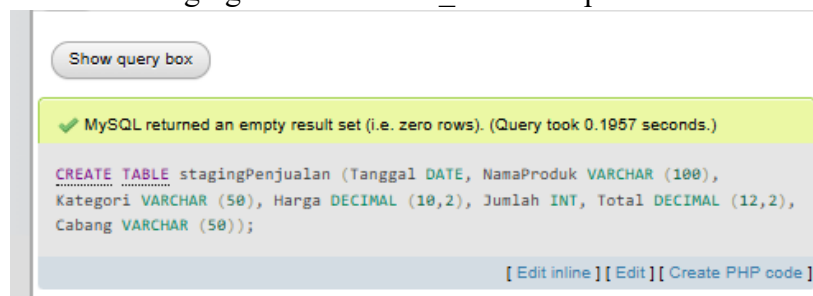
Pak Ozai merupakan staff di PT Indomarko Prismatama. PT Indomarko Prisma tama merupakan perusahaan retail yang memiliki usaha indomarti, superindi dan indigrosir. Pak ozai mempelajari proses bisnis dari perusahaan tersebut dan mencoba untuk membangun perusahaan sendiri dengan nama PT Ozai Enterprise dan membangun 3 cabang. Untuk dapat bersaing, pak ozai perlu melakukan analisa dari penjualan di ketiga cabang tersebut. 3 cabang tersebut mencatat penjualannya pada satu file excel.

Berikut ketiga file tersebut:

- Toko Azura** : <https://github.com/dik4rizky/datasources/blob/main/tokoazura.xls>
- Toko Zuko** : <https://github.com/dik4rizky/datasources/blob/main/tokozuko.xls>
- Toko Iroh** : <https://github.com/dik4rizky/datasources/blob/main/tokoiroh.xls>

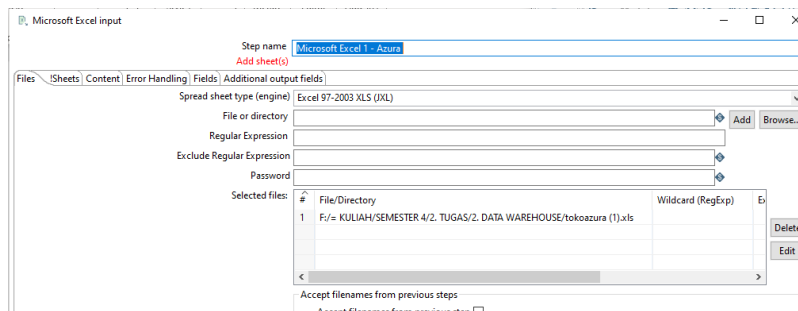
#### LANGKAH PRAKTIKUM

- Buat tabel staging di database dw\_OzaiEnterprise



- Drag and drop 3 **Excel Input** untuk data tokoazura, tokozuko, dan tokoiroh

- Toko azura**



Sesuaikan setiap field yang ada pada data excel toko azura :

Microsoft Excel input									
Step name: Microsoft Excel 1 - Azura									
Add sheet(s)									
Files	ISheets	Content	Error Handling	Fields	Additional output fields				
#	Name	Type	Length	Precision	Trim type	Repeat	Format	Currency	
1	kode	Number	-1	-1	none	N			
2	nomor_penjualan	String	-1	-1	none	N			
3	username	String	-1	-1	none	N			
4	qty	Number	-1	-1	none	N			
5	total_nilai	Number	-1	-1	none	N			
6	tanggal	Date	-1	-1	none	N			
7	Nama	String	-1	-1	none	N			
8	Harga	Number	-1	-1	none	N			
9	Satuan	String	-1	-1	none	N			

- **Toko zuko**

Step name: Microsoft Excel input 2 - Zuko

Spread sheet type (engine): Excel 97-2003 XLS (XLS)

File or directory:

Regular Expression:

Exclude Regular Expression:

Password:

Selected files:

#	File/Directory	Wildcard (RegExp)	Ex
1	F:/= KULIAH/SEMESTER 4/2. TUGAS/2. DATA WAREHOUSE/tokozuko (1).xls		
2			

Accept filenames from previous steps: ☐

Accept filenames from previous step: ☐

Step to read filenames from:

Field in the input to use as filename:

Show filename(s):

Sesuaikan setiap field yang ada pada data excel toko zuko :

Step name: Microsoft Excel input 2 - Zuko

Fields:

#	Name	Type	Length	Precision	Trim type	Repeat	Format
1	kode	Number	-1	-1	none	N	
2	nomor_penjualan	String	-1	-1	none	N	
3	username	String	-1	-1	none	N	
4	qty	Number	-1	-1	none	N	
5	total_nilai	Number	-1	-1	none	N	
6	tanggal	Date	-1	-1	none	N	
7	Nama	String	-1	-1	none	N	
8	Harga	Number	-1	-1	none	N	
9	Satuan	String	-1	-1	none	N	

- **Toko iroh**

Step name: Microsoft Excel input 3 - Iroh

Spread sheet type (engine): Excel 97-2003 XLS (XLS)

File or directory:

Regular Expression:

Exclude Regular Expression:

Password:

Selected files:

#	File/Directory	Wildcard (RegExp)	Ex
1	F:/= KULIAH/SEMESTER 4/2. TUGAS/2. DATA WAREHOUSE/tokoiroh (1).xls		

Accept filenames from previous steps: ☐

Accept filenames from previous step: ☐

Step to read filenames from:

Field in the input to use as filename:

Show filename(s):

Sesuaikan setiap field yang ada pada data excel toko iroh :

Step name: Microsoft Excel input 3 - Iroh

Fields:

#	Name	Type	Length	Precision	Trim type	Repeat	Format	Currency
1	kode	Number	-1	-1	none	N		
2	nomor_penjualan	String	-1	-1	none	N		
3	username	String	-1	-1	none	N		
4	qty	Number	-1	-1	none	N		
5	total_nilai	Number	-1	-1	none	N		
6	tanggal	Date	-1	-1	none	N		
7	Nama	String	-1	-1	none	N		
8	Harga	Number	-1	-1	none	N		
9	Satuan	String	-1	-1	none	N		

### 3. Tambahkan **add constants** (untuk mengisi field cabang dengan value constant)

- Constant1 Azura

The dialog box shows the step name 'Add constants 1 - Azura'. The fields table is as follows:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty string?
1	Cabang	String							Azura	N
2										

- Constant2 Zuko

The dialog box shows the step name 'Add constants 2 - Zuko'. The fields table is as follows:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty string?
1	Cabang	String							Zuko	N
2										

- Constant3 Iroh

The dialog box shows the step name 'Add constants 3 - Iroh'. The fields table is as follows:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty string?
1	Cabang	String							Iroh	
2										

### 4. Lakukan union dengan **Append Streams** atau **Merge Rows** untuk menggabungkan ketiga data menjadi satu

- Append streams 1 dan 2

The first screenshot shows the 'Append streams' dialog box with the step name 'Append streams'. The head hop is 'Add constants 1 - Azura' and the tail hop is 'Add constants 2 - Zuko'.

The second screenshot shows the 'Append streams' dialog box with the step name 'Append streams 2'. The head hop is 'Add constants 3 - Iroh' and the tail hop is 'Add constants 3 - Iroh'.

### 5. Konfigurasi untuk **tabel output**

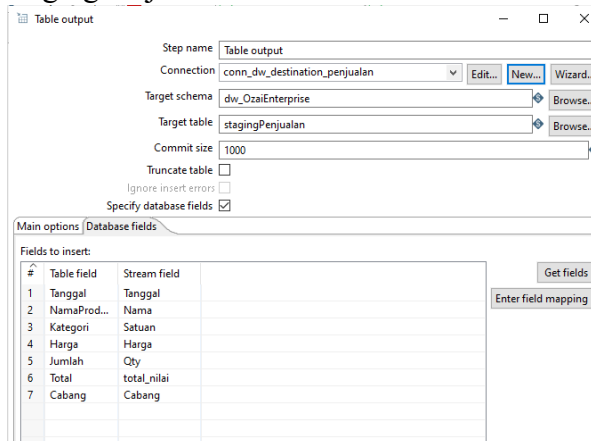
- Buat koneksi ke database dw\_OzaiEnterprise

The 'Database Connection' dialog box shows the connection name 'conn\_dw\_destination\_penjualan'. The connection type is 'H2'. The settings are as follows:

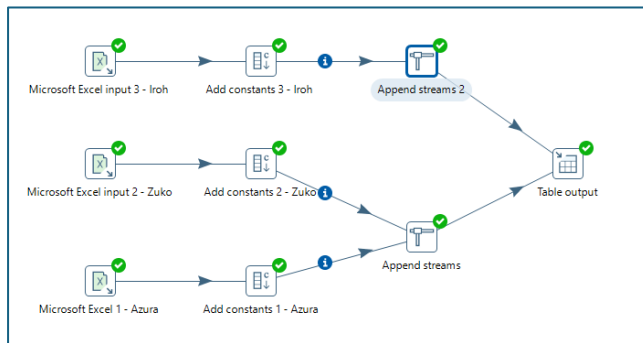
Setting	Value
Host Name	localhost
Database Name	dw_OzaiEnterprise
Port Number	3306
Username	root
Password	
Use Result Streaming Cursor	<input checked="" type="checkbox"/>

A success message dialog box is shown: 'Connection tested successfully. Connection to conn\_dw\_destination\_penjualan was successful. Hostname: localhost, Port: 3306, Database name: dw\_OzaiEnterprise'.

- Konfigurasi tabel output dan database fields nya sesuai dengan struktur tabel stagingPenjualan



## 6. Hasil akhir



### Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2025/04/19 00:45:37 - Spoon - Transformation opened.  
 2025/04/19 00:45:37 - Spoon - Launching transformation [Transformation JS4 Studi Kasus]...  
 2025/04/19 00:45:37 - Spoon - Started the transformation execution.  
 2025/04/19 00:45:37 - Transformation JS4 Studi Kasus - Dispatching started for transformation [Transformation JS4 Studi Kasus]  
 2025/04/19 00:45:37 - Table output.0 - Connected to database [conn\_dw\_destination\_penjualan] (commit=1000)  
 2025/04/19 00:45:38 - Microsoft Excel input 2 - Zuko.0 - Finished processing (I=1078, O=0, R=0, W=1078, U=0, E=0)  
 2025/04/19 00:45:38 - Add constants 2 - Zuko.0 - Finished processing (I=0, O=0, R=1078, W=1078, U=0, E=0)  
 2025/04/19 00:45:38 - Microsoft Excel 1 - Azura.0 - Finished processing (I=3862, O=0, R=0, W=3862, U=0, E=0)  
 2025/04/19 00:45:38 - Add constants 1 - Azura.0 - Finished processing (I=0, O=0, R=3862, W=3862, U=0, E=0)  
 2025/04/19 00:45:38 - Microsoft Excel input 3 - Iroh.0 - Finished processing (I=13069, O=0, R=0, W=13069, U=0, E=0)  
 2025/04/19 00:45:38 - Add constants 3 - Iroh.0 - Finished processing (I=0, O=0, R=13069, W=13069, U=0, E=0)  
 2025/04/19 00:45:38 - Append streams.0 - Finished processing (I=0, O=0, R=4940, W=4940, U=0, E=0)  
 2025/04/19 00:45:40 - Append streams.2.0 - Finished processing (I=0, O=0, R=13069, W=13069, U=0, E=0)  
 2025/04/19 00:45:42 - Table output.0 - Finished processing (I=0, O=18009, R=18009, W=18009, U=0, E=0)  
 2025/04/19 00:45:42 - Spoon - The transformation has finished!!

## Isi tabel pada database :

Showing rows 0 - 24 (18009 total, Query took 0.0005 seconds.)

SELECT \* FROM 'stagingpenjualan'

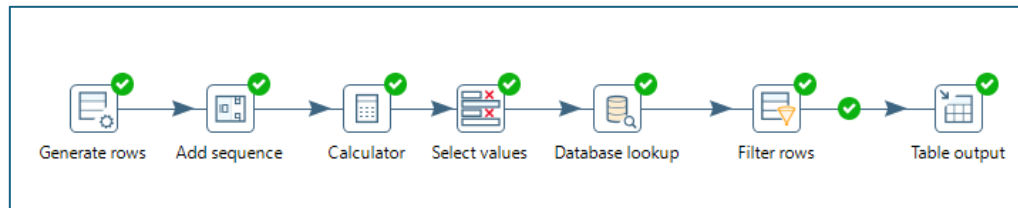
Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

1 > >> Number of rows: 25 Filter rows: Search this table

Extra options

Tanggal	NamaProduk	Kategori	Harga	Jumlah	Total	Cabang
2008-02-18	INDOCAFE CAPUCINO KTK 5X25G	PCS	6916.25	2	353700.00	Azura
2008-02-18	ALKALINE LR-03 (ISI 2)	LSN	7938.25	1	8250.00	Azura
2008-01-02	NUTRIJELL STRAW 15 GR	BOX	2867.50	1	33000.00	Azura
2008-01-02	NUTRIJELL ANGGUR 15 GR	BOX	2867.50	1	33000.00	Azura
2008-01-02	NUTRIJELL PLAIN 15 GR	BOX	2867.50	1	33000.00	Azura
2008-01-02	REJOICE FAMILY C 175	PCS	10890.00	1	187500.00	Azura
2008-01-02	WINGS BIRU 600 K 209G	DOS	2250.00	2	46050.00	Azura
2008-01-02	SOKLIN MATIC FRONT LOAD 1KG	DOS	20000.00	1	28850.00	Azura
2008-01-07	SOKLIN HIGINIS 900	DOS	15625.00	1	20250.00	Azura
2008-01-17	AJINOMOTO 50GR	PCS	1418.75	2	56500.00	Azura
2008-02-18	ABC SBLASU 135 ML	PCS	3537.50	1	154000.00	Azura
2008-02-18	KACANG SHANGHAI DK 450	PCS	12250.00	1	353700.00	Azura
2008-01-02	BATERAI ABC KECIL R-6	LSN	1450.00	4	18800.00	Azura
2008-01-17	ABC KCP ASIN 133 ML	PCS	2250.00	1	91325.00	Azura
2008-01-02	CARRERA POLES M 20GX8S	PCS	2800.00	2	70600.00	Azura

## B. DIMENSI WAKTU



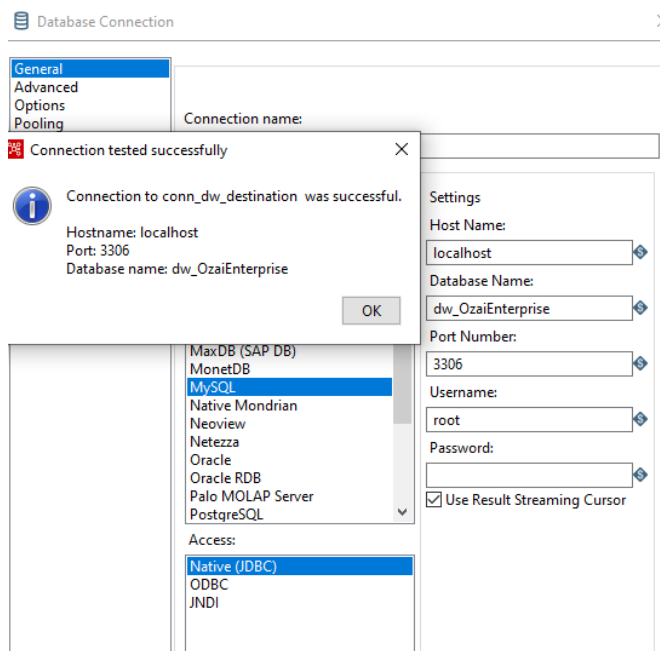
1. Membuat database yang digunakan sebagai OLAP dengan nama dw\_OzaiEnterprise

```
mysql> CREATE DATABASE dw_OzaiEnterprise;
Query OK, 1 row affected (0.10 sec)

mysql> use dw_OzaiEnterprise;
Database changed
```

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

```
mysql> CREATE TABLE dimDate (
-> id_dimDate INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
-> date DATE,
-> year INT,
-> month INT,
-> day INT
-> );
Query OK, 0 rows affected (1.90 sec)
```

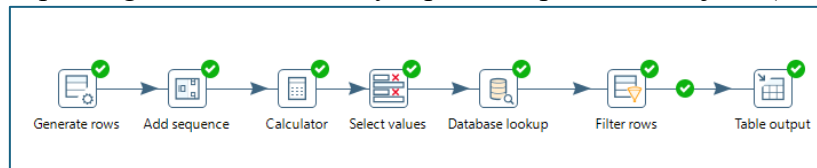


Pada tahapan selanjutnya, untuk membuat tabel dimensi dimDate, maka diperlukan generate data tanggal. Data tanggal yang disiapkan pada tabel dimDate menyesuaikan dengan proses bisnis yang berjalan.

Proses bisnis pada LegendVehicle adalah 20 tahun. Sehingga data pada tabel dimdate yang harus tersedia adalah tanggal selama 15 tahun. Mulai dari 1 Januari 2005

- 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).



3. Konfigurasi pada Generate Rows adalah merubah limit menjadi 1825 dimana memiliki arti bahwa data yang akan dibuat sebanyak 7300 data. 7300 merupakan jumlah hari dalam 20 tahun ( 365 hari x 20 tahun ).
4. Membuat fields baru bernama CurrentDate dengan type data Date dan format dd-MM-yyyy serta value awal 01-01-2005.

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

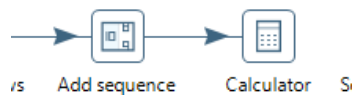
Fields:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty string?
1	currentDate	Date	dd-MM-yyyy						01-01-2005	N

Help OK Preview Cancel

5. Hubungkan output dari Generate Rows menuju Add Sequence.
6. Konfigurasi pada Add Sequences adalah merubah Name of value menjadi incrementDay dengan start value bernilai 0 dan increment by bernilai 1

7. Hubungkan output dari add sequences menuju calculator.



8. Konfigurasi pada calculator dengan membuat fields baru sebagai berikut:

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

Calculator

Step name

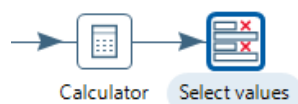
Calculator

☒ Throw an error on non existing files

Fields:

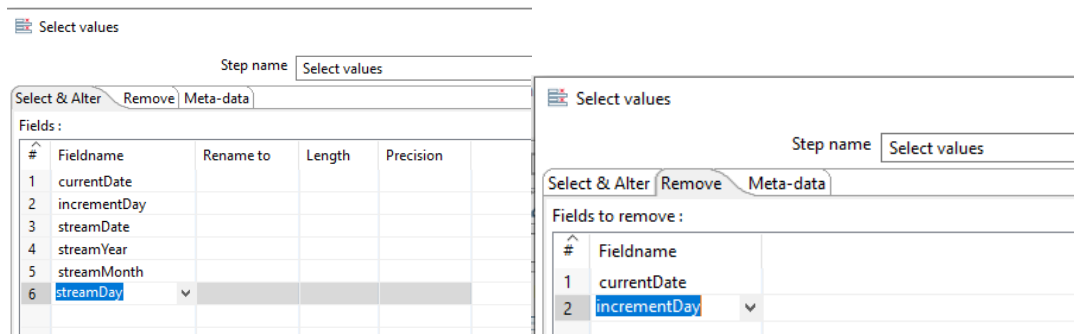
#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove
1	streamDate	Date A + B Days	currentDate	incrementDay		None			N
2	streamYear	Year of date A	streamDate			None			N
3	streamMonth	Month of date A	streamDate			None			N
4	streamDay	Day of month of date A	streamDate			None			N

9. Hubungkan output dari calculator menuju Select values



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

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



12. Hubungkan output select values menuju database lookup.



13. Sebelum melakukan konfigurasi pada database lookup, buatlah koneksi terlebih dahulu pada database melalui File - New - Database Connection. Gunakan Connection type MySQL dengan host name , database name, port number, username dan password sesuai konfigurasi MySQL pada device masing-masing. beri nama connection name tersebut dengan nama conn\_dw\_destination.
14. Konfigurasi pada database lookup adalah dengan memberikan connection dengan koneksi yang sudah dibuat pada step sebelumnya. dengan schema nama database yang digunakan dan tabel dimdate yang telah dibuat pada langkah pertama.
15. Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak kembar adalah:
- field date pada table dimdate dengan field streamDate
  - field year pada table dimdate dengan field streamYear
  - field month pada table dimdate dengan field streamMonth
  - field day pada table dimdate dengan field streamDay
16. Field yang akan di retriive adalah field yang ada pada table dimDate yaitu date, year, month, dan day.

The screenshot shows the 'Database lookup' configuration dialog box. The 'Step name' is 'Database lookup'. The 'Connection' is 'conn\_dw\_destination'. The 'Lookup schema' is 'dw\_OzaiEnterprise'. The 'Lookup table' is 'dimdate'. The 'Enable cache?' checkbox is unchecked. The 'Cache size in rows (0=cache)' is 0. The 'Load all data from table' checkbox is unchecked. The 'The key(s) to look up the value(s):' table is as follows:

#	Table field	Comparator	Field1	Field2
1	date	=	stream...	
2	year	=	stream...	
3	month	=	stream...	
4	day	=	stream...	

The 'Values to return from the lookup table:' table is as follows:

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

Below the table, there is a text input field with the value 'of of these are' and a dropdown menu.



17. Hubungkan output dari database lookup dengan filter rows
18. Konfigurasi pada filter rows adalah dengan melakukan konfigurasi output true data pada table output. Pada bagian ini data yang tidak memiliki kesamaan pada tahapan sebelumnya akan dicek dimana jika fields Stream tidak memiliki kesamaan dengan field dimDate, maka field dimDate tersebut akan bernilai null. Pada pernyataan kondisi tuliskan ( date is null and year is null and month is null and day is null)

Filter rows

Step name: Filter rows    out put    outputs    outputting

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

19. Hubungkan output dari filter rows menuju table output.



20. Konfigurasi pada table output adalah memberikan koneksi pada conn\_dw\_destination dengan schema dw\_legendvehicle dan table dimdate.
21. Aktifkan specify database fields.
22. Pada tab Database fields, mapping data input streamDate, streamYear, streamMonth dan streamDay dengan fields yang ada pada dimDate. Pada tahapan ini akan dilakukan insert data menuju tabel dimDate.

Table output

Step name: Table output

Connection: conn\_dw\_destination    Edit...    New...    Wizard...

Target schema: dw\_OzaiEnterprise    Browse...

Target table: dimdate    Browse...

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

23. Cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.

**Execution Results**

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2025/03/20 09:50:31 - Spoon - Save file as...

2025/03/20 09:50:47 - Spoon - Transformation opened.

2025/03/20 09:50:47 - Spoon - Launching transformation [Transformation A - JS4]...

2025/03/20 09:50:47 - Spoon - Started the transformation execution.

2025/03/20 09:50:48 - Transformation A - JS4 - Dispatching started for transformation [Transformation A - JS4]

2025/03/20 09:50:48 - Table output.0 - Connected to database [conn\_dw\_destination] (commit=1000)

2025/03/20 09:50:49 - Generate rows.0 - Finished processing (I=0, O=0, R=0, W=1825, U=0, E=0)

2025/03/20 09:50:49 - Add sequence.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)

2025/03/20 09:50:49 - Calculator.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)

2025/03/20 09:50:49 - Select values.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)

2025/03/20 09:50:51 - Database lookup.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)

2025/03/20 09:50:51 - Filter rows.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)

2025/03/20 09:50:52 - Table output.0 - Finished processing (I=0, O=1825, R=1825, W=1825, U=0, E=0)

2025/03/20 09:50:52 - Spoon - The transformation has finished!!

	id_dimDate	date	year	month	day
<input type="checkbox"/> Edit Copy Delete	1	2005-01-01	2005	1	1
<input type="checkbox"/> Edit Copy Delete	2	2005-01-02	2005	1	2
<input type="checkbox"/> Edit Copy Delete	3	2005-01-03	2005	1	3
<input type="checkbox"/> Edit Copy Delete	4	2005-01-04	2005	1	4
<input type="checkbox"/> Edit Copy Delete	5	2005-01-05	2005	1	5
<input type="checkbox"/> Edit Copy Delete	6	2005-01-06	2005	1	6
<input type="checkbox"/> Edit Copy Delete	7	2005-01-07	2005	1	7
<input type="checkbox"/> Edit Copy Delete	8	2005-01-08	2005	1	8
<input type="checkbox"/> Edit Copy Delete	9	2005-01-09	2005	1	9
<input type="checkbox"/> Edit Copy Delete	10	2005-01-10	2005	1	10
<input type="checkbox"/> Edit Copy Delete	11	2005-01-11	2005	1	11
<input type="checkbox"/> Edit Copy Delete	12	2005-01-12	2005	1	12
<input type="checkbox"/> Edit Copy Delete	13	2005-01-13	2005	1	13
<input type="checkbox"/> Edit Copy Delete	14	2005-01-14	2005	1	14
<input type="checkbox"/> Edit Copy Delete	15	2005-01-15	2005	1	15
<input type="checkbox"/> Edit Copy Delete	16	2005-01-16	2005	1	16
<input type="checkbox"/> Edit Copy Delete	17	2005-01-17	2005	1	17
<input type="checkbox"/> Edit Copy Delete	18	2005-01-18	2005	1	18
<input type="checkbox"/> Edit Copy Delete	19	2005-01-19	2005	1	19
<input type="checkbox"/> Edit Copy Delete	20	2005-01-20	2005	1	20

Console

## 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 di setiap prosesnya.

Execution Results													
<div> <a href="#">Logging</a> <a href="#">Execution History</a> <a href="#">Step Metrics</a> <a href="#">Performance Graph</a> <a href="#">Metrics</a> <a href="#">Preview data</a> </div>													
#	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	13,931	-
2	Add sequence	0	1825	1825	0	0	0	0	0	Finished	0.2s	12,006	-
3	Calculator	0	1825	1825	0	0	0	0	0	Finished	0.3s	7,300	-
4	Select values	0	1825	1825	0	0	0	0	0	Finished	0.3s	6,636	-
5	Database lookup	0	1825	1825	0	0	0	0	0	Finished	2.5s	719	-
6	Filter rows	0	1825	1825	0	0	0	0	0	Finished	2.6s	714	-
7	Table output	0	1825	1825	0	1825	0	0	0	Finished	3.1s	589	-

Proses Objek	SS Data input	SS Data Output	Keterangan
Generate Rows	0	1825	Data bertambah menjadi 1825. Generate rows berfungsi untuk membuat 1825 baris data
Add Sequences	1825	1825	Menambahkan nilai sequence/nomor urut pada data
Calculator	1825	1825	Melakukan perhitungan atau manipulasi pada data
Select Values	1825	1825	Memilih kolom yang akan digunakan
Database Lookup	1825	1825	Mencari data yang cocok berdasarkan kondisi/nilai dalam dataset
Filter rows	1825	1825	Filtering data berdasarkan kondisi jika tanggal, bulan, tahun serta hari null.
Table output	1825	1825	Menghasilkan transformasi dan menyimpan ke dalam tabel database.  Langkah ini hanya memproses data dengan nilai <b>NULL</b> pada kolom <b>date</b> , <b>year</b> , <b>month</b> , dan <b>day</b> , sedangkan data yang valid akan diabaikan.