

### LAPORAN JOBSHEET 3

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

Copy

```
CREATE DATABASES dw_LegendVehicle;
```

```
use dw_LegendVehicle;
```

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

Copy

```
CREATE TABLE dimDate(
```

```
id_dimDate int not null AUTO_INCREMENT PRIMARY KEY
```

```
date date
```

```
year int
```

```
month int
```

```
day int
```

```
);
```

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 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.

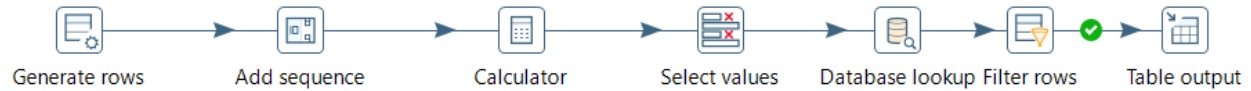
Calculator: digunakan untuk menjumlahkan hari dan mengambil data tahun, bulan dan hari.

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).



3. 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 ).
4. 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

Fields:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set em
1	CurrentDate	Date	dd-MM-yyyy						01-01-2003	N

Buttons: 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

Step name: Add sequence

Name of value: incrementDay

Use a database to generate the sequence: ☐

Use DB to get sequence? ☐

Connection: conn\_dw\_destination

Schema name:

Sequence name: SEQ\_

Use a transformation counter to generate the sequence: ☒

Use counter to calculate sequence? ☒

Counter name (optional):

Start at value: 0

Increment by: 1

Maximum value: 99999999

Buttons: Help, OK, Cancel

7. Hubungkan output dari add sequences menuju calculator.

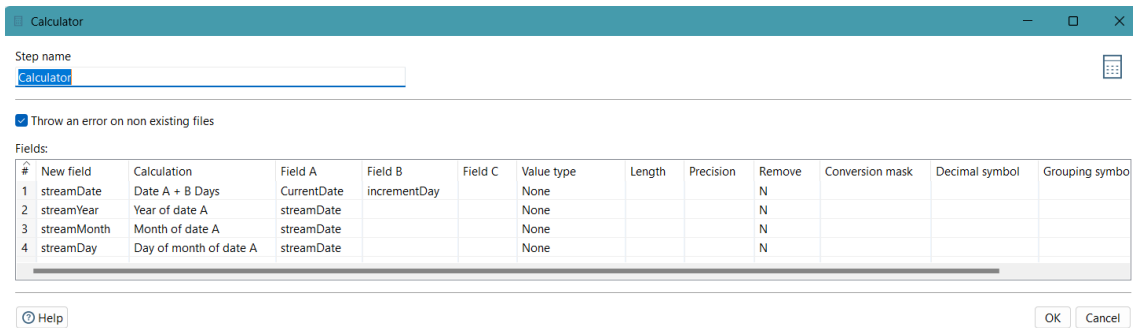
8. Konfigurasi pada calculator dengan membuat fields baru sebagai berikut:

streamDate merupakan kalkulasi dari CurrentDate + incrementDay

streamYear merupakan Year dari streamDate

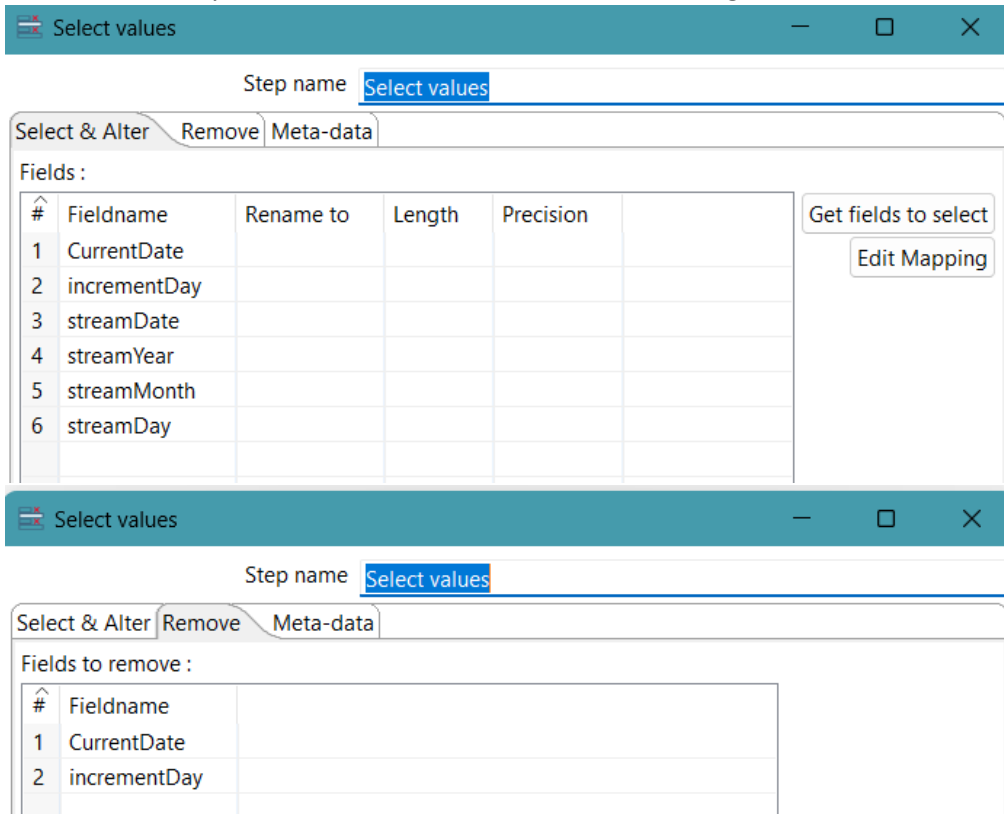
streamMonth merupakan Month dari streamDate

streamDay merupakan Day of month dari streamDate



#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal symbol	Grouping symbol
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 diisikan fields CurrentDate dan incrementDay dikarenakan kedua fields tersebut tidak digunakan.



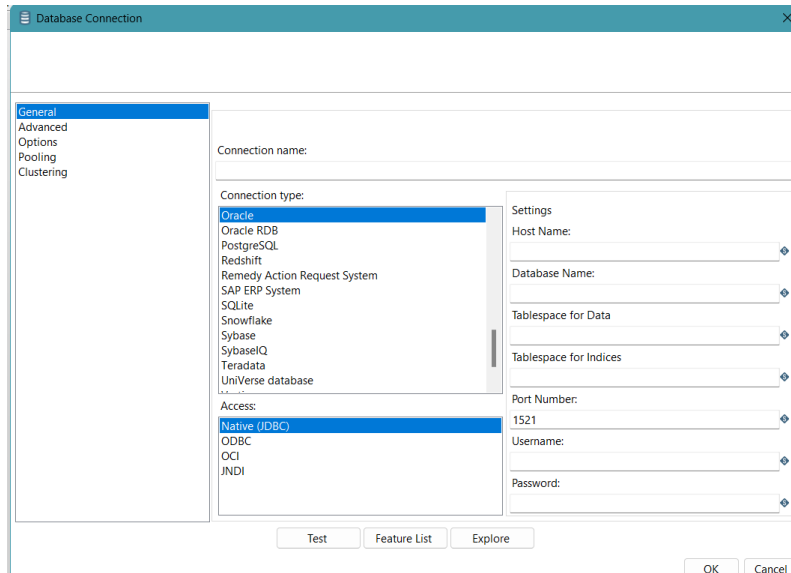
#	Fieldname	Rename to	Length	Precision
1	CurrentDate			
2	incrementDay			
3	streamDate			
4	streamYear			
5	streamMonth			
6	streamDay			

#	Fieldname
1	CurrentDate
2	incrementDay

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 retrieve adalah field yang ada pada table dimDate yaitu date, year, month, dan day.

Database lookup

Step name: Database lookup

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

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

Lookup table: dimdet [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	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:

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

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

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.

Transformation properties

Transformation name: dimensi waktu

Transformation filename: C:\Renald Tugas\Semester 4\Data Warehouse\dimensi waktu.ktr

Description:

Extended description:

Status:

Version:

Directory: /

Created by: -

Created at: Thu Mar 06 14:23:33 ICT 2025

Last modified by: -

Last modified at: Thu Mar 06 14:23:33 ICT 2025

OK SQL Cancel

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

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

## 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.

Proses Objek	SS data input	SS data output	Keterangan
Generate rows	Tidak ada input	Data dengan sejumlah baris yang dibuat secara otomatis	Membuat data dummy atau awal dengan jumlah baris tertentu.
Add Sequences	Data dari "Generate rows"	Data dengan kolom tambahan yang berisi nomor urut	Menambahkan nomor urut ke setiap baris data.
Calculator	Data dari "Add Sequences"	Data dengan kolom hasil perhitungan baru	Melakukan operasi matematika berdasarkan aturan tertentu.
Select values	Data dari "Calculator"	Data dengan kolom tertentu yang dipilih	Memilih kolom tertentu yang akan digunakan lebih lanjut.
Database lookup	Data dari "Select values"	Data dengan informasi tambahan dari database	Mengambil informasi tambahan berdasarkan pencocokan data dengan database.
Filter rows	Data dari "Database lookup"	Data yang sudah disaring sesuai kriteria	Menyaring data berdasarkan kondisi tertentu.
Table Output	Data dari "Filter rows"	Data yang siap disimpan dalam tabel output	Menyimpan hasil akhir ke dalam tabel tujuan.

### C. Dimensi Pegawai

1. Buatlah tabel dimPegawai pada dw\_legendVehicle.

Copy

```
CREATE TABLE dimEmployees(  
    id_dimEmployees int not_null auto increment primary,  
    employeeNumber int(11),  
    firstName varchar(50),  
    lastname varchar(50),  
    jobtitle varchar(50),  
    boss_firstname varchar(50),  
    boss_lastname varchar(50),  
    updated date DEFAULT CURRENT_DATE  
);
```

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.

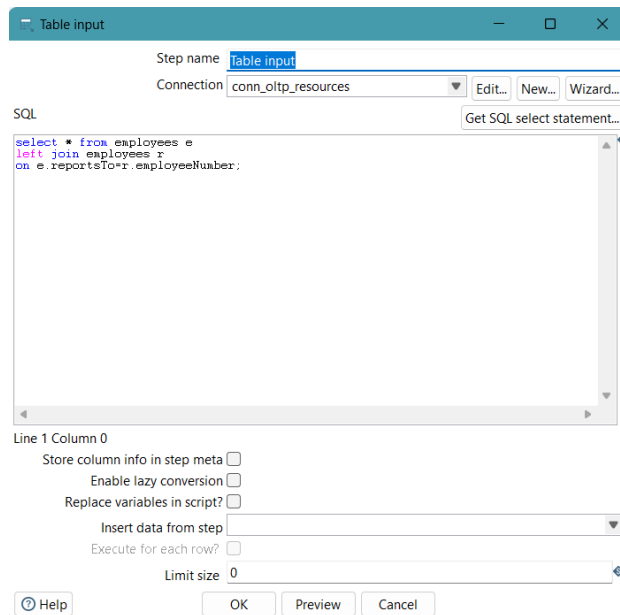
Database lookup: digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar

Filter rows: digunakan untuk memilih data stream yang masih belum ada apada tabel dimEmployees.

Table output: Memasukkan data ke dalam tavle dimEmployees



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



## Copy

```
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 employeeenumber, 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.



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			
1..	lastName_1	StreamLastnameBoss		
1..	firstName_1	StreamFirstnameBoss		
1..	extension_1			
1..	email_1			
1..	officeCode_1			
1..	reportsTo_1			
1..	jobTitle_1			

Get fields to select

Edit Mapping

☐ Include unspecified fields, ordered

Help

OK

Cancel

Select values

Step name

Select values

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
1..	jobTitle_1

Get fields to remove

Help

OK

Cancel

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.

Database lookup

Step name: Database lookup

Connection: conn\_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	employeeNumber			Integer
2	firstName			String
3	lastName			String
4	jobTitle			String
5	boss_firstname			String
6	boss_lastname			String

Do not pass the row if the lookup fails ☐

Fail on multiple results? ☐

Order by:

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

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.

Filter rows

Step name: Filter rows

Send 'true' data to step: Table output

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

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\_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: [Dropdown]

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: [Dropdown]

Store the tablename field: ☒

Return auto-generated key: ☐

Name of auto-generated key field: [Dropdown]

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

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
<input type="checkbox"/>	Edit Copy Delete		117	1002 Diane	Murphy	President	NULL	NULL	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		118	1056 Mary	Patterson	VP Sales	Diane	Murphy	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		119	1076 Jeff	Firrelli	VP Marketing	Diane	Murphy	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		120	1088 William	Patterson	Sales Manager (APAC)	Mary	Patterson	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		121	1102 Gerard	Bondur	Sale Manager (EMEA)	Mary	Patterson	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		122	1143 Anthony	Bow	Sales Manager (NA)	Mary	Patterson	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		123	1165 Leslie	Jennings	Sales Rep	Anthony	Bow	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		124	1166 Leslie	Thompson	Sales Rep	Anthony	Bow	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		125	1188 Julie	Firrelli	Sales Rep	Anthony	Bow	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		126	1216 Steve	Patterson	Sales Rep	Anthony	Bow	2025-03-13
<input type="checkbox"/>	Edit Copy Delete		127	1286 Foon Yue	Tseng	Sales Rep	Anthony	Bow	2025-03-13

## 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.

StreamEmployeeNumber	StreamLastname	StreamFirstname	StreamJobtitle	StreamLastnameBoss	StreamFirstnameBoss	employeeNumber	firstName	lastName	jobTitle	boss_firstName
1002	Murphy	Diane	President	<null>	<null>	<null>	<null>	<null>	<null>	<null>

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

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

<input type="checkbox"/>	Edit Copy Delete	140	2001 Renald	Putra	Software Engineer	Diane	Murphy	2025-03-13
--------------------------	------------------	-----	-------------	-------	-------------------	-------	--------	------------