LAPORAN JOBSHEET 3 DATA WAREHOUSE DATABASE ANALYTICAL (A&B)



Disusun Oleh: Maharani Wirawan (2341760111) SIB 2B

SISTEM INFORMASI BISNIS JURUSAN TEKNOLOGI INFORMASI POLITEKNIK NEGERI MALANG 2025

A. Dimensi Waktu

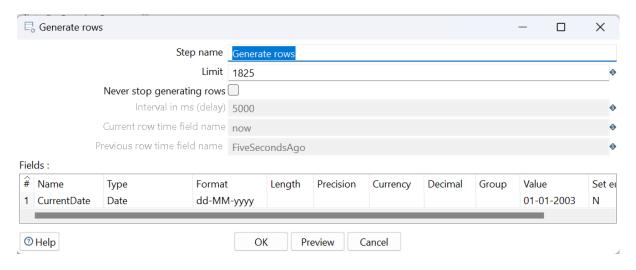
- 1. Buatlah sebuah database yang digunakan sebagai OLAP dengan nama dw LegendVehicle.
- 2. Buatlah table untuk menyimpan data master waktu atau yang disebut dengan tabel dimensi. Beri nama table tersebut dengan nama dimDate.



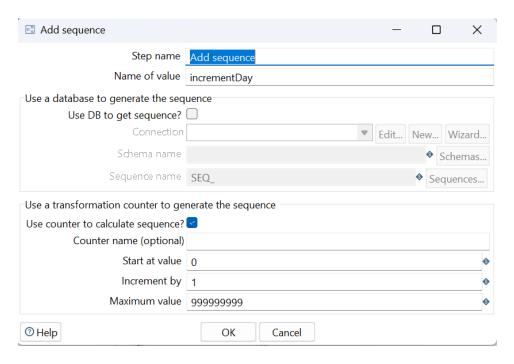
- 1. Buka PDI Spoon. Buat Transformation baru -> File New Transformation.
- 2. Drag and Drop beberapa objek yaitu:
- Generate Rows: digunakan untuk membuat baris data baru.
- Add Sequence: digunakan untuk membuat sequence, dalam hal ini membuat data di setiap harinya.
- Calculator: digunakan untuk menjumlahkan 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).



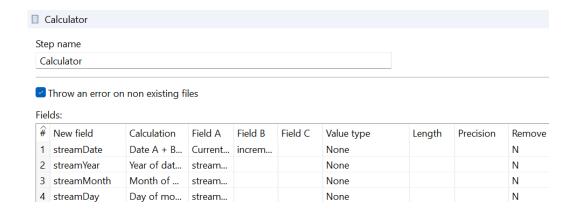
- 3. Konfigurasi pada Generate Rows adalah merubah limit menjadi 1825 dimana memiliki arti bahwa data yang akan dibuat sebanyak 1825 data. 1825 merupakan jumlah hari dalam 5 tahun (365 hari x 5 tahun).
- 4. Membuat fields baru bernama CurrentDate dengan type data Date dan format dd-MM-yyyy serta value awal 01-01-2003.



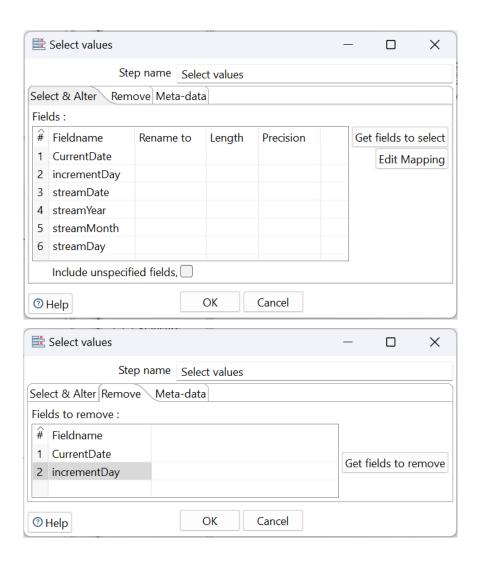
- 5. Hubungkan output dari Generate Rows menuju Add Sequence.
- 6. Konfigurasi pada Add Sequences adalah merubah Name of value menjadi incrementDay dengan start value bernilai 0 dan increment by bernilai 1



- 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

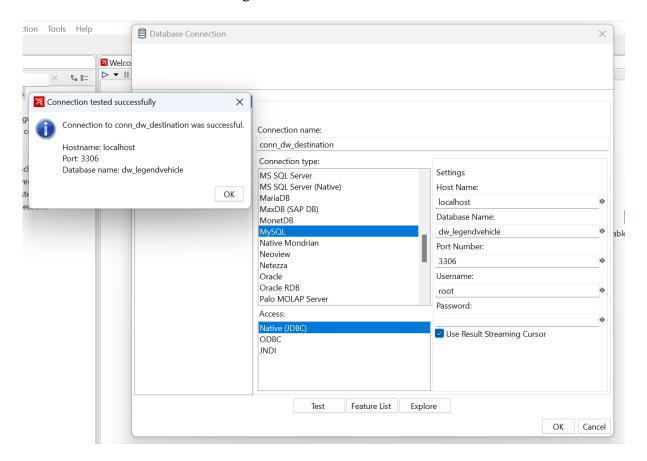


- 9. Hubungkan output dari calculator menuju Select values
- 10. Konfigurasi pada select values adalah dengan menekan tombol Get fields to select pada tab Select & Alter. Secara otomatis semua fields dari data input akan muncul.
- 11. Dikarenakan tidak semua fields digunakan, maka pada tab Remove diisikan fields CurrentDate dan incrementDay dikarenakan kedua fields tersebut tidak digunakan.

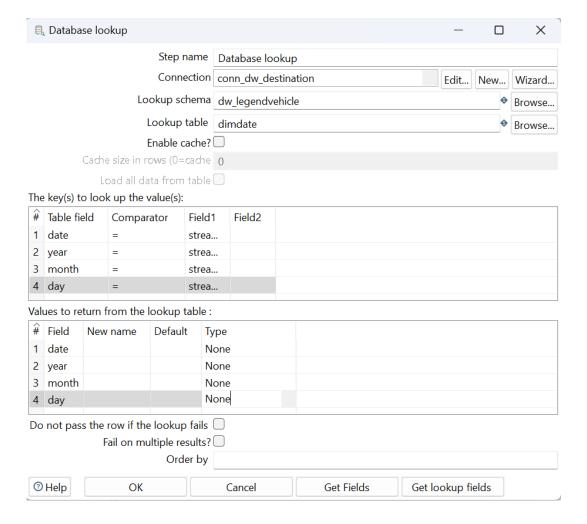


12. Hubungkan output select values menuju database lookup.

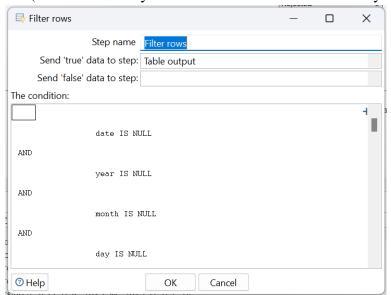
13. Sebelum melakukan konfigurasi pada database lookup, buatlah koneksi terlebih dahulu pada database melalui File - New - Database Connection. Gunakan Connection type MySQL dengan host name, database name, port number, username dan password sesuai konfigurasi MySQL pada device masing-masing. beri nama connection name tersebut dengan nama conn_dw_destination.



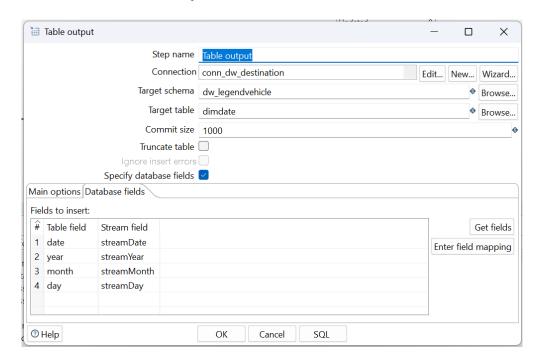
- 14. Konfigurasi pada database lookup adalah dengan memberikan connection dengan koneksi yang sudah dibuat pada step sebelumnya. dengan schema nama database yang digunakan dan tabel dimdate yang telah dibuat pada langkah pertama.
- 15. Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak 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 retrive adalah field yang ada pada table dimDate yaitu date, year, month, dan day.



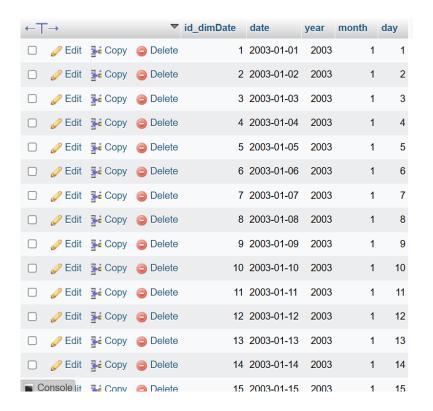
- 17. Hubungkan output dari database lookup dengan filter rows
- 18. Konfigurasi pada filter rows adalah dengan melakukan konfigurasi output true data pada table output. Pada bagian ini data yang tidak memiliki kesamaan pada tahapan sebelumnya akan dicek dimana jika fields Stream tidak memiliki kesamaan 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)



- 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.



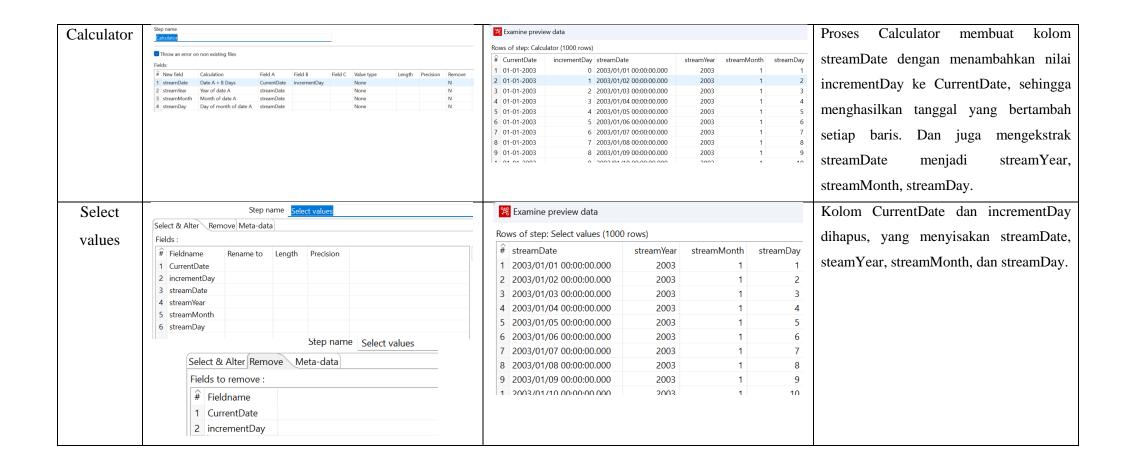
23. cek isi table dimdate pada database. Jika sukses maka pada table dimdate akan terisi 1825 data.

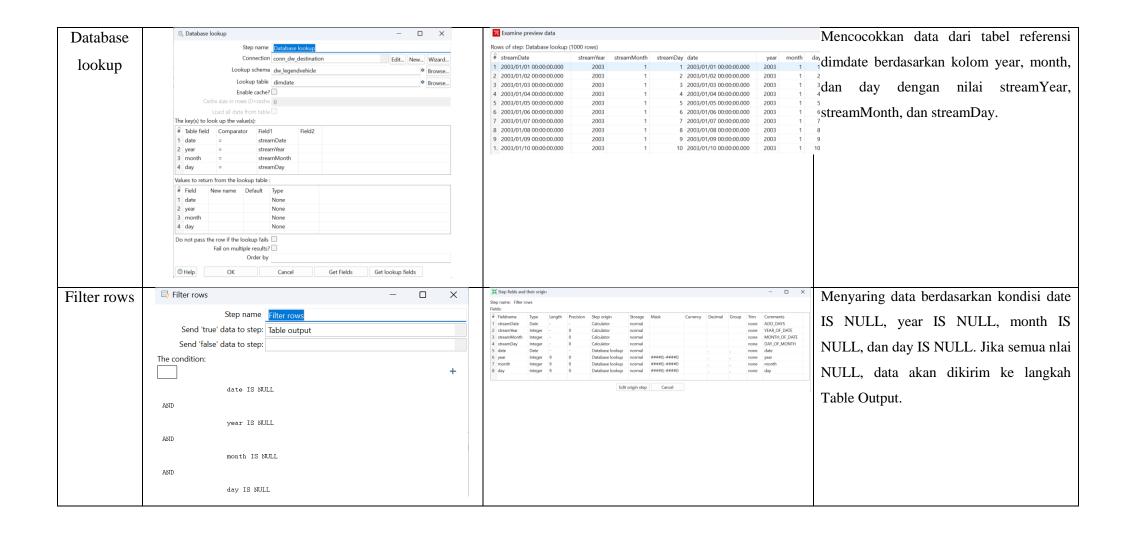


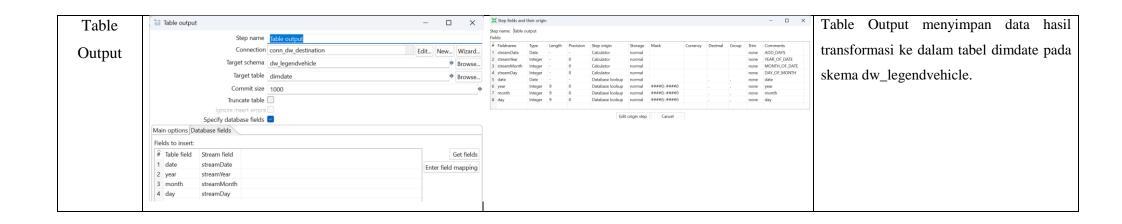
TUGAS 1

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	Step name Generate rows	Examine preview data	Data input menetapkan 01-01-2003 tanpa perubahan, dan data output menghasilkan 1000 baris dengan nilai yang sama
Add Sequences	Step name Add sequence Name of value incrementDay Use a database to generate the sequence Use DB to get sequence? Connection Schema name Sequence name Sequence name Sequence to calculate sequence? Counter to calculate sequence? Counter to alculate sequence? Counter name (optional) Start at value Increment by Maximum value 999999999	Rows of step: Add sequence (1000 rows) # CurrentDate incrementDay 1	Add Sequence menambahkan kolom incrementDay yang dimulai dari 0 dan bertambah 1 setiap baris, namu nilai CurrentDay masih 01-01-2003.

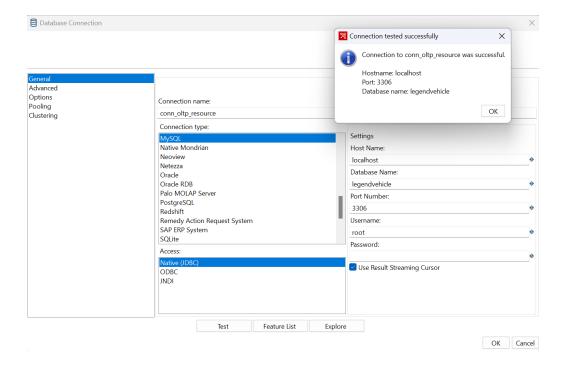




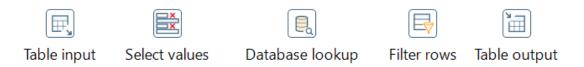


B. Dimensi Pegawai

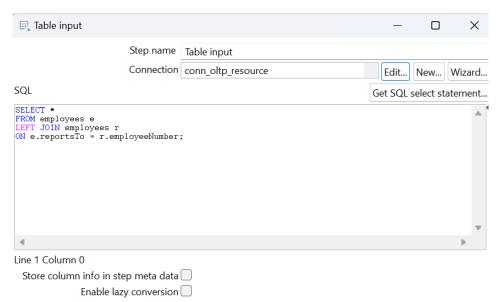
- 1. Buatlah tabel dimPegawai pada dw_legendVehicle.
- 2. Pada PDI Spoon buatlah koneksi baru dengan nama conn_oltp_resources yang menghubungkan dengan database oltp. sesuaikan hostname, database name, port number, username dan password dengan keadaan pada device masing-masing.



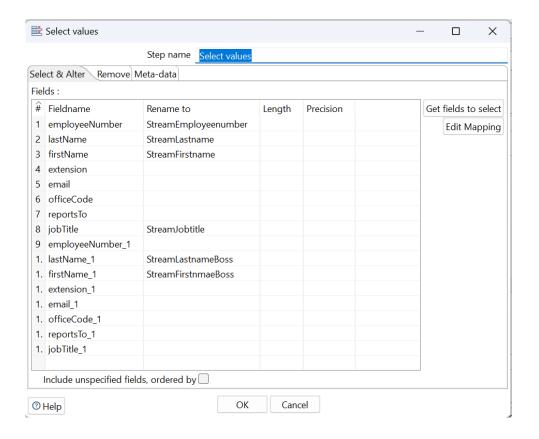
- 3. Drag and drop beberapa objek sebagai berikut:
 - Table input: digunakan mengambil data dari database OLTP.
 - Select values: memeilih field yang digunakan untuk proses Transform dan Load.
 - Database lookup: digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar
 - Filter rows: digunakan untuk memilih data stream yang masih belum ada apada tabel dimEmployees.
 - Table output: Memasukkan data ke dalam tavle dimEmployees

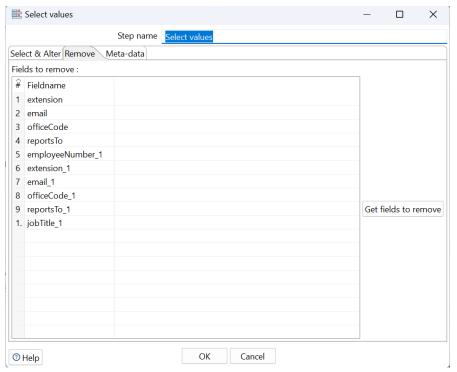


4. Konfigurasi pada table input dengan menghubungkan Connection pada konesi conn_oltp_resources. Untuk mengambil data sumber menggunakan query dibawah ini.



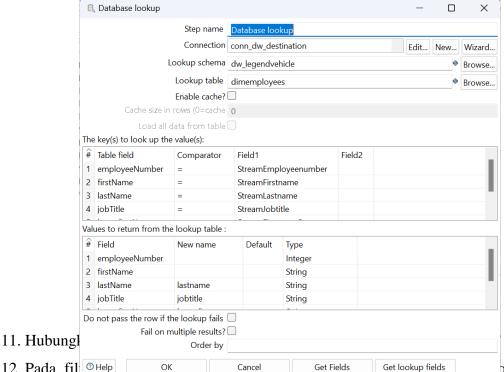
- 5. Hubungkan output table input pada select values.
- 6. Konfigurasi pada Select values yaitu mengambil data dari field employeenumber, lastname, firstname, jobtitle, lastname_1 dan firstname_1 sebagai data stream yang digunakan pada proses ETL pada tab select & alter.
- 7. Hilangkan field lain yang tidak digunakan pada tab remove.





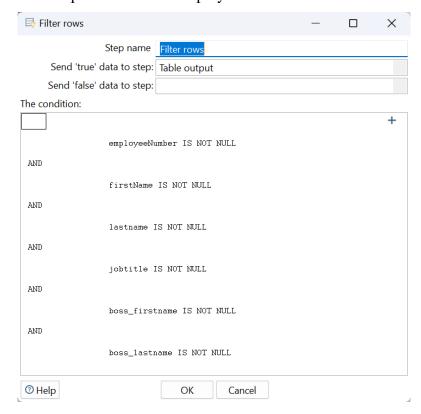
- 8. Hubungkan output select values pada database lookup
- Konfigurasi pada database lookup adalah dengan menghubungkan koneksi pada conn_dw_destination dengan table lookup dimEmployees yang telah dibuat pada tahap pertama.

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.

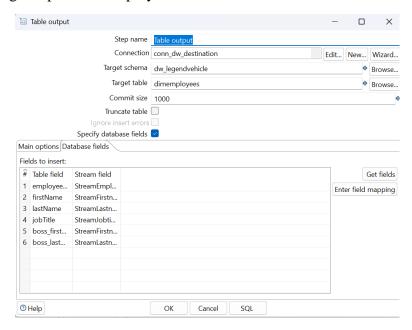


12. Pada fil Help Cancel Get Fields Get lookup fields oyees untuk

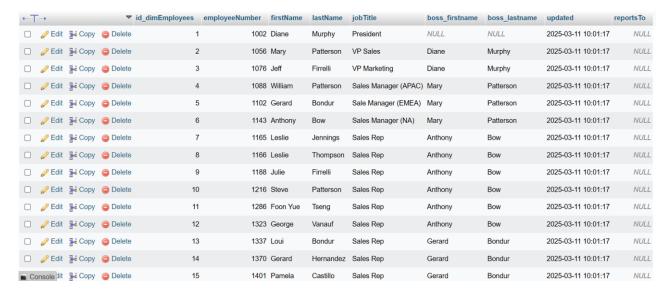
dimasukkan pada proses selanjutnya. Hal itu menandakan bahwa data stream belum memiliki kesamaan pada data di dimemployees.



- 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 databse fields, dan mapping data stream input dari oltp terhadap field yang ada pada dimemployees.

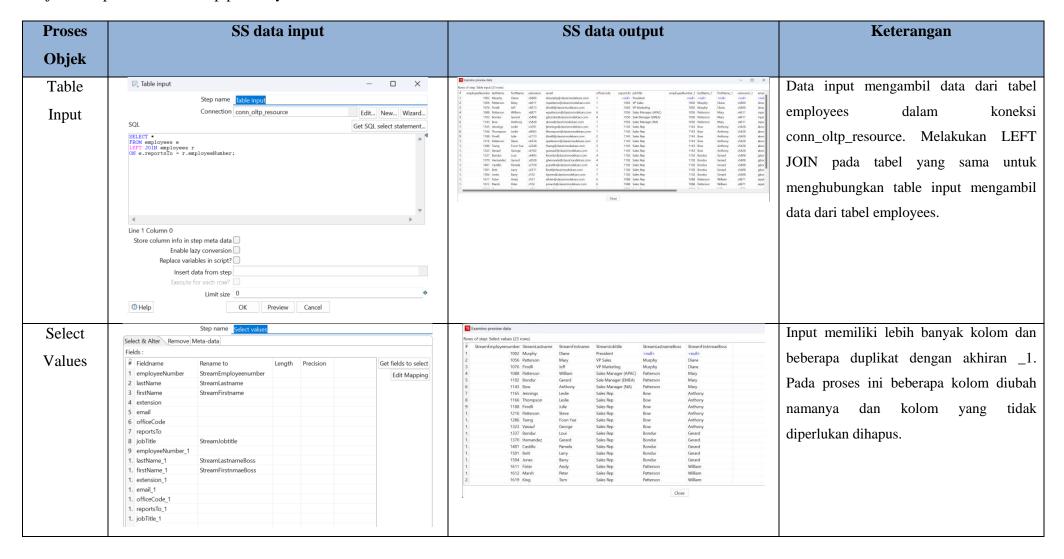


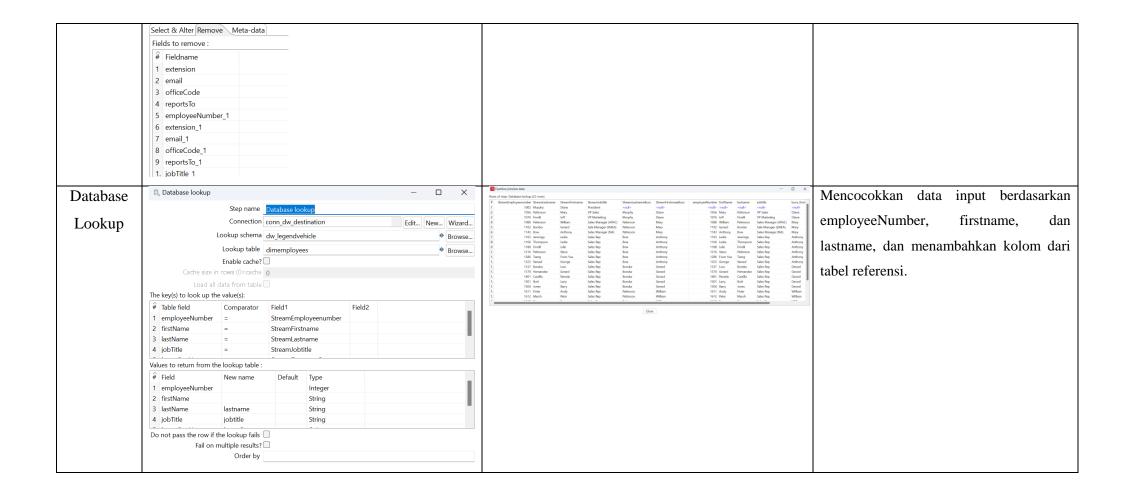
16. jika proses keseluruhan berhasil maka tabel dimemployees akan terisi data pegawai dari database OLTP.

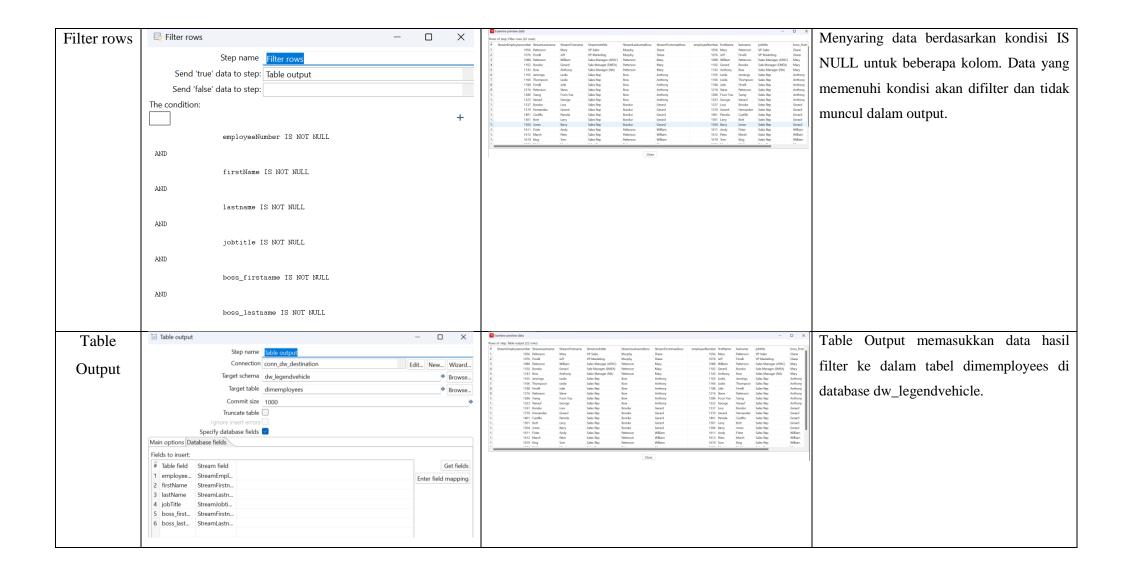


TUGAS 2

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.







- 2. Jika proses itu di ulangi (di run kembali) apakah data akan redudant?
 - = Jika proses di ulangi, tidak aka nada redudansi data karena terdapar proses yang mencegah terjadinya duplikasi yaitu Database Lookup yang mengecek data yang sudah ada sebelum dimasukkan dan filter rows yang menyaring data supaya tidak terjadi duplikasi.
- 3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?

