**JOBSHEET 3**

**MATA KULIAH DATA WAREHOUSE**

**DIMENSI WAKTU DAN DIMENSI PEGAWAI**



**OLEH :**

**KARTIKA TRI JULIANA**

**SIB-2B / 10**

**2341760116**

**PROGRAM STUDI SISTEM INFORMASI BISNIS**

**JURUSAN TEKNOLOGI INFORMASI**

**POLITEKNIK NEGERI MALANG**

**TAHUN 2025**

**Studi Kasus**

Berdasarkan studi kasus sebelumnya, Pimpinan dari LegendVehicle mulai mengenal teknologi dan berpandangan visioner. LegendVehicle akan menerapkan data warehouse pada proses bisnisnya untuk menganalisa proses transaksi yang dilihat dari pembayaran yang masuk. Proses ini digunakan untuk menganalisa KPI "jumlah omset yang dimiliki ".

Untuk membentuk data warehouse tersebut, Data Engineer pada LegendVehicle perlu membuat sebuah database yang digunakan sebagai databasae OLAP.

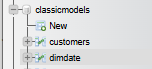
Data dari OLTP yang ada akan di "ETL" kan menuju database OLAP.

1. Dimensi Waktu

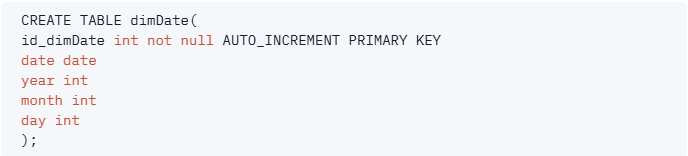
Buatlah sebuah database yang digunakan sebagai OLAP dengan nama dw\_LegendVehicle.

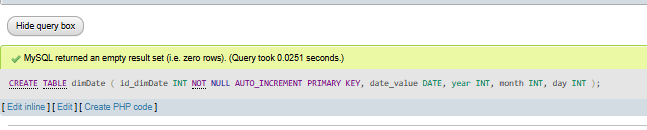


Karena perintah nya mengggunakan database sebelumnya maka saya memakai database classicmodels



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



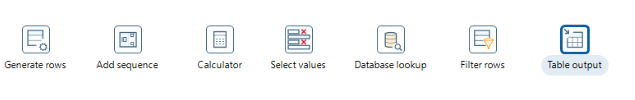


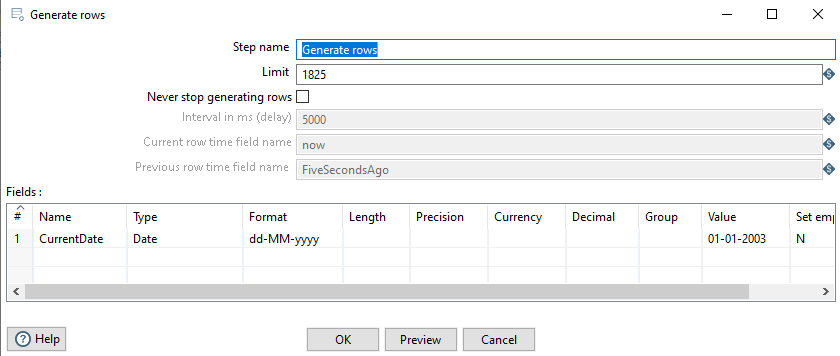
Pada tahapan selanjutnya, untuk membuat tabel dimensi dimDate , maka diperlukan generate data tanggal. Data tanggal yang disiapkan pada tabel dimDate menyesuaikan dengan proses bisnis yang berjalan.

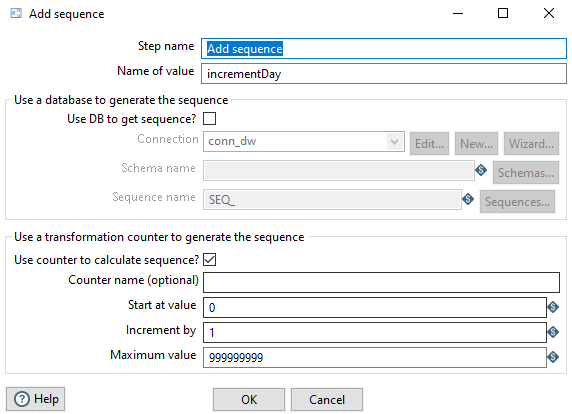
Proses bisnis pada LegendVehicle adalah 5 tahun. Sehingga data pada tabel dimdate yang harus tersedia adalah tanggal selama 5 tahun. Mulai dari 1 Januari 2023

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).

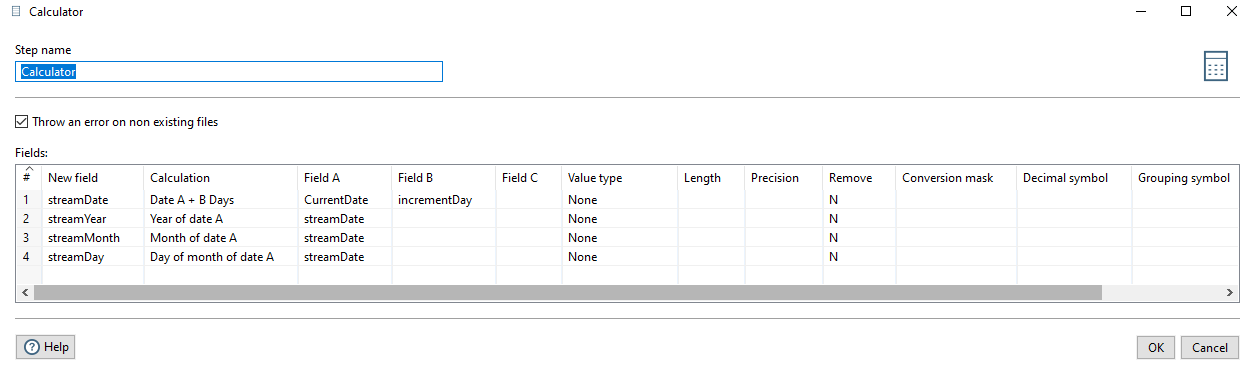


1. Konfigurasi pada Generate Rows adalah merubah limit menjadi 1825 dimana memiliki arti bahwa data yang akan dibuat sebanyak 1825 data. 1825 merupakan jumlah hari dalam 5 tahun ( 365 hari x 5 tahun ).
2. Membuat fields baru bernama CurrentDate dengan type data Date dan format dd-MM-yyyy serta value awal 01-01-2003.
3. Hubungkan output dari Generate Rows menuju Add Sequence.
4. Konfigurasi pada Add Sequences adalah merubah Name of value menjadi incrementDay dengan start value bernilai 0 dan increment by bernilai 1

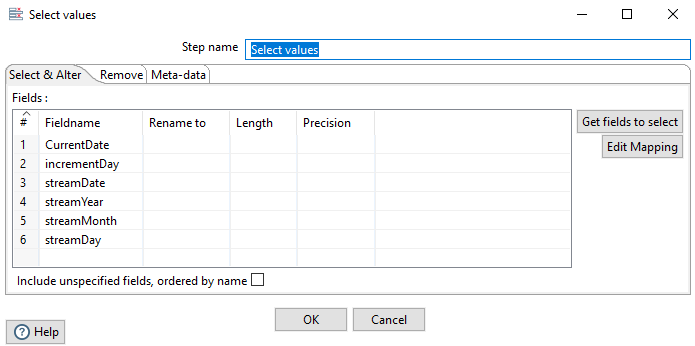
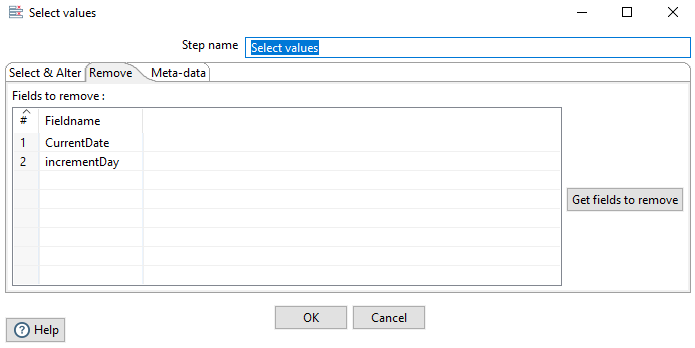


1. Hubungkan output dari add sequences menuju calculator.
2. Konfigurasi pada calculator dengan membuat fields baru sebagai berikut:

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

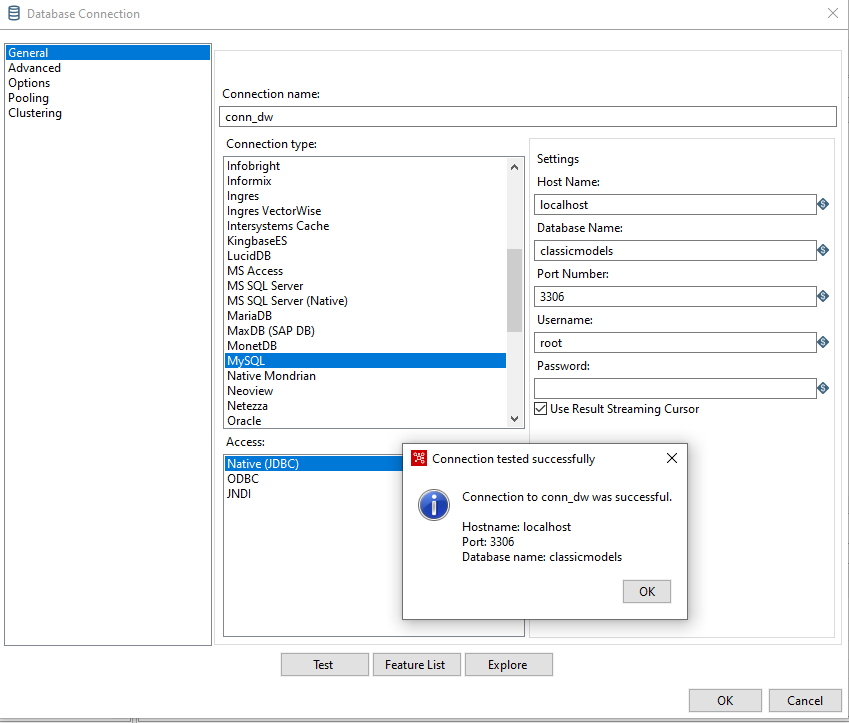


1. Hubungkan output dari calculator menuju Select values
2. Konfigurasi pada select values adalah dengan menekan tombol Get fields to select pada tab Select & Alter. Secara otomatis semua fields dari data input akan muncul.
3. Dikarenakan tidak semua fields digunakan, maka pada tab Remove diisikan fields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



1. Hubungkan output select values menuju database lookup.
2. 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.

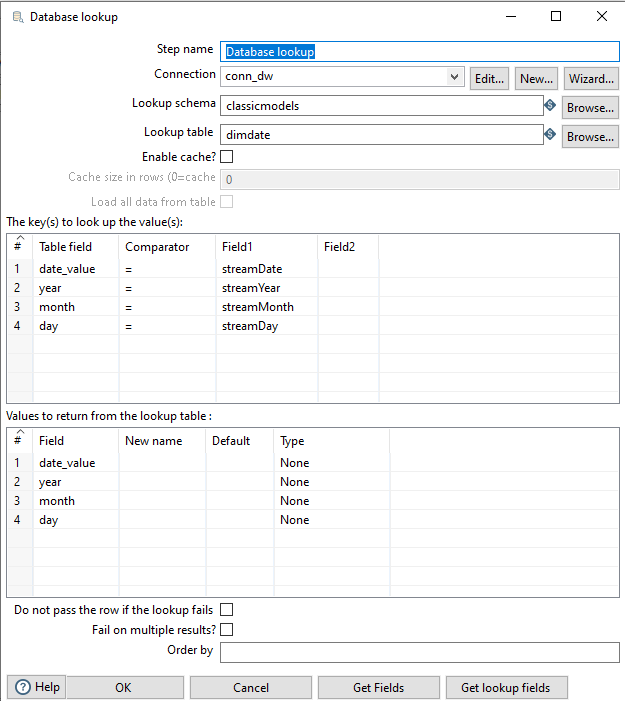
= Sebelumnya saya mengalami error pada hostname dan username dan ternyata hostname diisikan oleh localhost dan username dan password sesuaikan dengan yang ada di phpmyadmin masing-masing

S

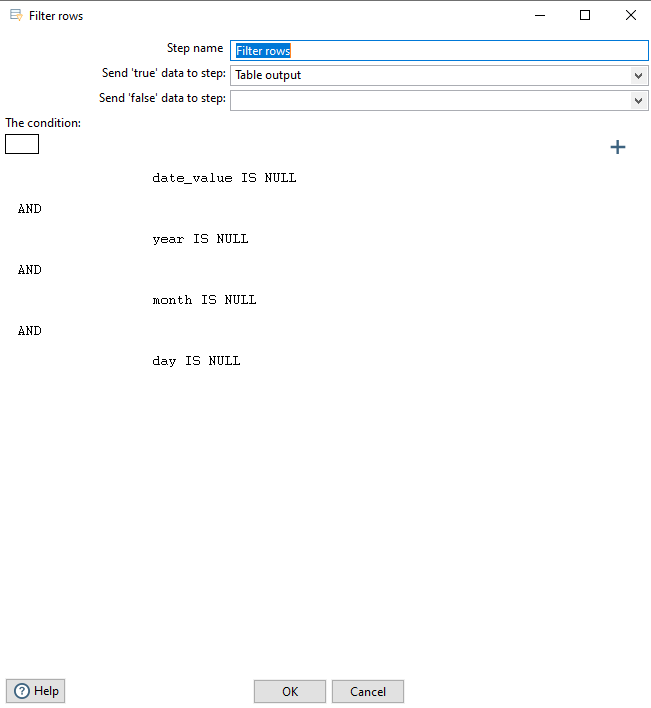
1. konfigurasi pada database lookup adalah dengan memberikan connection dengan koneksi yang sudah dibuat pada step sebelumnya. dengan schema nama database yang digunakan dan tabel dimdate yang telah dibuat pada langkah pertama.
2. Field yang akan dicek untuk melihat kesamaan isi datanya agar tidak kembar 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

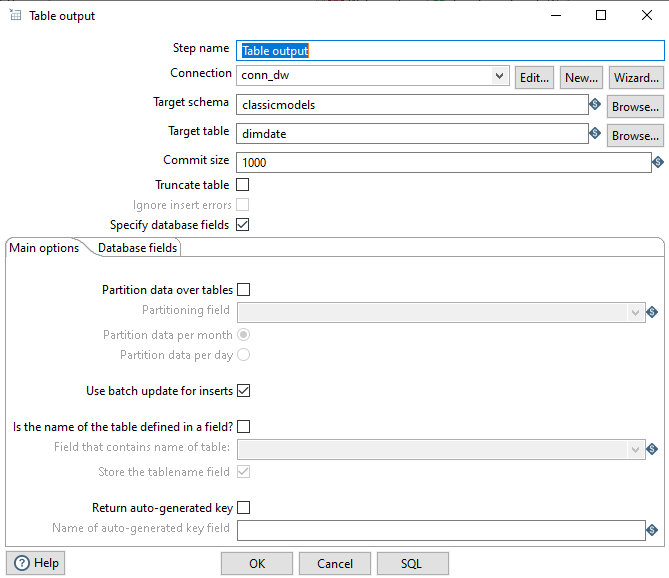
1. Field yang akan di retrive adalah field yang ada pada table dimDate yaitu date, year, month, dan day.

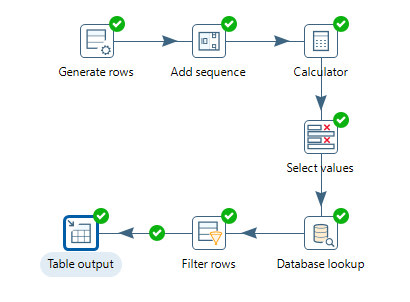


1. Hubungkan output dari database lookup dengan filter rows
2. Konfigurasi pada filter rows adalah dengan melakukan konfigurasi output true data pada table output. Pada bagian ini data yang tidak memiliki kesamaan pada tahapan sebelumnya akan dicek dimana jika fields Stream tidak memiliki kesamaan dengan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



1. Hubungkan output dari filter rows menuju table output.
2. Konfigurasi pada table output adalah memberikan koneksi pada conn\_dw\_destination dengan schema dw\_legendvehicle dan table dimdate.
3. Aktifkan specify database fields.
4. 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.



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

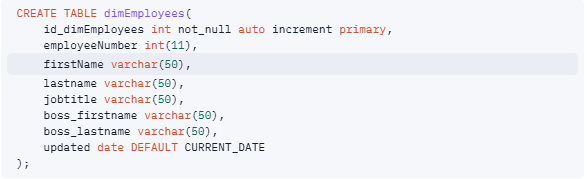
**TUGAS 1**

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.

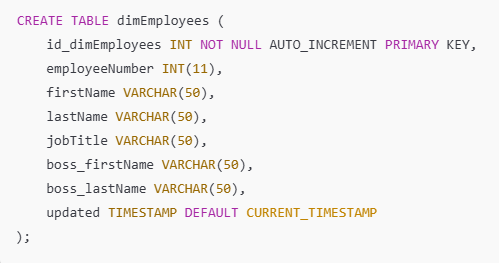
|  |  |  |  |
| --- | --- | --- | --- |
| **Proses Objek** | **Input** | **Output** | **keterangan** |
| Generate Rows |  |  | Output pada generate rows berisikan currentdata dengan format dd-mm-yyy yakni 01-01-01 sebanyak inputan selama 1825 hari atau 5 tahun lamanya |
| Add Sequence |  |  | Dengan memasukkan start at value 0 yang artinya value terisi dimulai dengan 0 dan increment by yakni setiap baris menambah 1 |
| Calculator |  |  | Menampilkan field baru yakni streamDate, streamYear, streamMonth, streamDay. Dengan adanya incrementDay pada field 2 yang menandakan penambahan 1 pada hari nya yang pengaruh ke streamDay |
| Select Values |  |  | Select dan alter menampilkan currentDate, incrementDay, streamDate, streamYear, streamMonth, streamDay dan pada remove memasukkan currentDate, incrementDay sehingga yang ditampilkan hanya row streamDate, streamYear, streamMonth, streamDay |
| Database Lookup |  |  | Koneksi nyambung ke conn\_dw dengan lookup table dimdat di lookup skema classicmodels. Key yang ditampilkan adalah date\_value sebagai streamDate, year sebagai streamYear, month sebagai streamMonth, day sebagai streamDay. Kemudian adanya return lookup dengan type None sehingga ketika di view hasil menjadi null |
| Filter Rows |  |  | Inputa yang diberikan menunjukkan jika filter rows berguna untuk memisahkan data data dengan nilai null yang akan diteruskan pada table output |
| Table Output |  |  | Dengan konekso conn\_dw dan target skema classicmodels dan target table dimdae dan centang pada field database yang spesifik untuk dimunculkan pada table output |

1. **DIMENSI PEGAWAI**
2. Buatlah tabel dimPegawai pada dw\_legendVehicle.

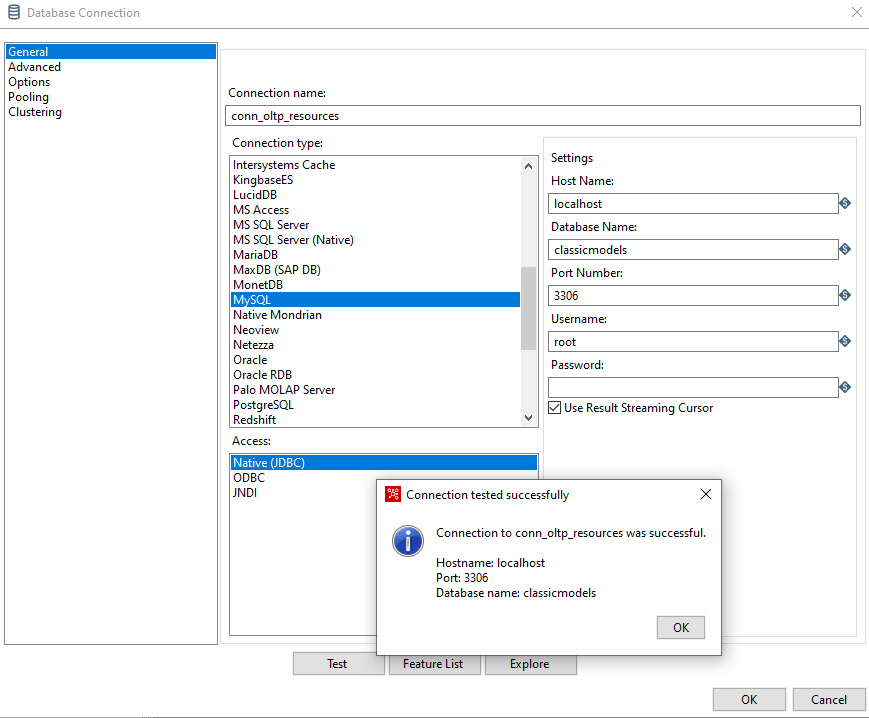
= karena saya memakai classicmodels maka saya membuatnya di classicmodels



Saya mengalami error sehingga memperbaiki kode nya menjadi seperti ini

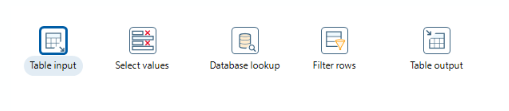


1. Pada PDI Spoon buatlah koneksi baru dengan nama conn\_oltp\_resources yang menghubungkan dengan database oltp. sesuaikan hostname, database name, port number, username dan password dengan keadaan pada device masing-masing.

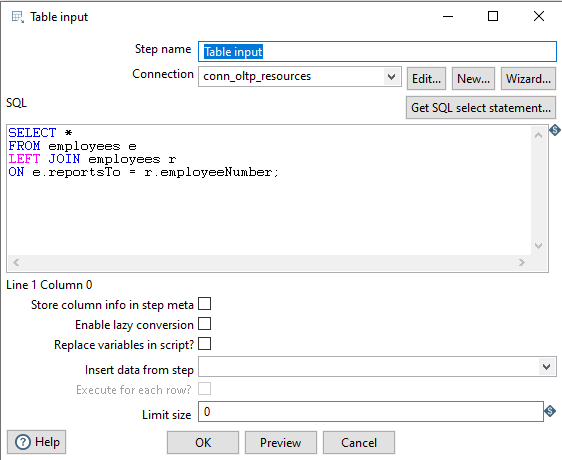


1. Drag and drop beberapa objek sebagai berikut:Table input:

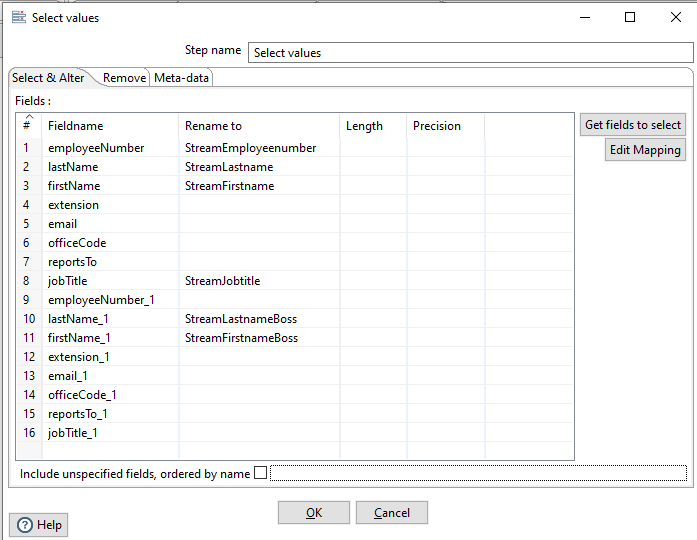
* digunakan mengambil data dari database OLTP.
* Select values: memeilih field yang digunakan untuk proses Transform dan Load.
* Database lookup: digunakan untuk melihat data pada tabel dimEmployees untuk memastikan data tidak kembar
* Filter rows: digunakan untuk memilih data stream yang masih belum ada apada tabel dimEmployees.
* Table output: Memasukkan data ke dalam tavle dimEmployees

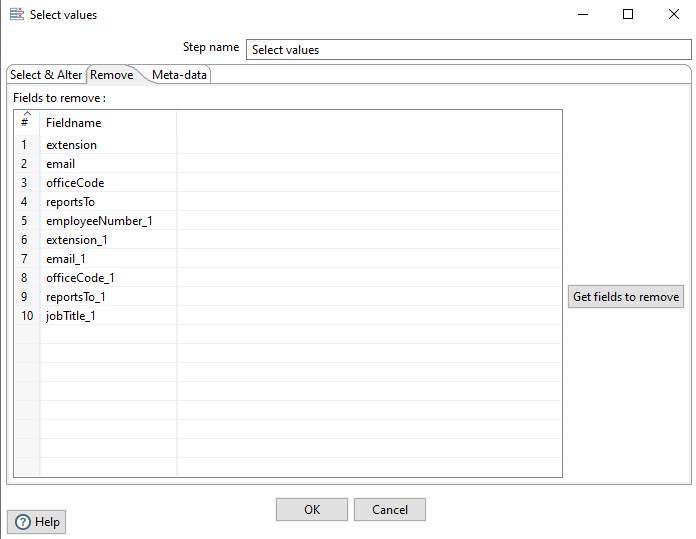


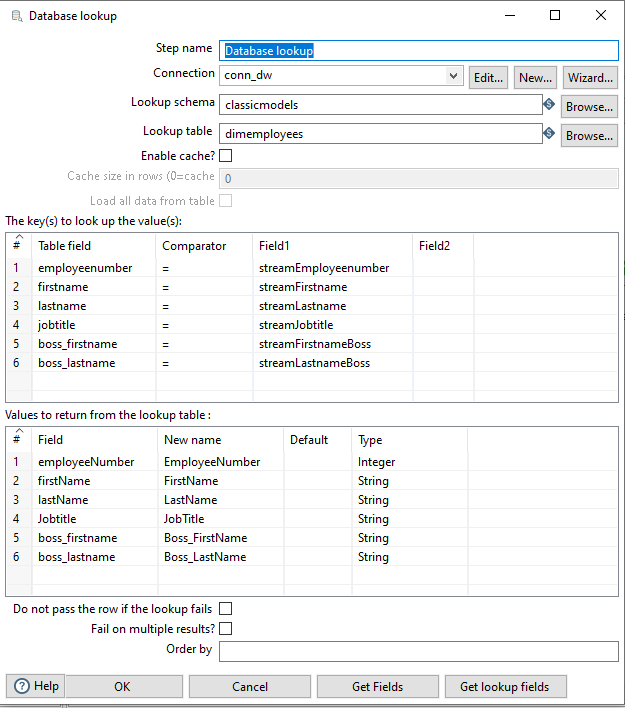
1. Konfigurasi pada table input dengan menghubungkan Connection pada konesi conn\_oltp\_resources. Untuk mengambil data sumber menggunakan query dibawah ini.



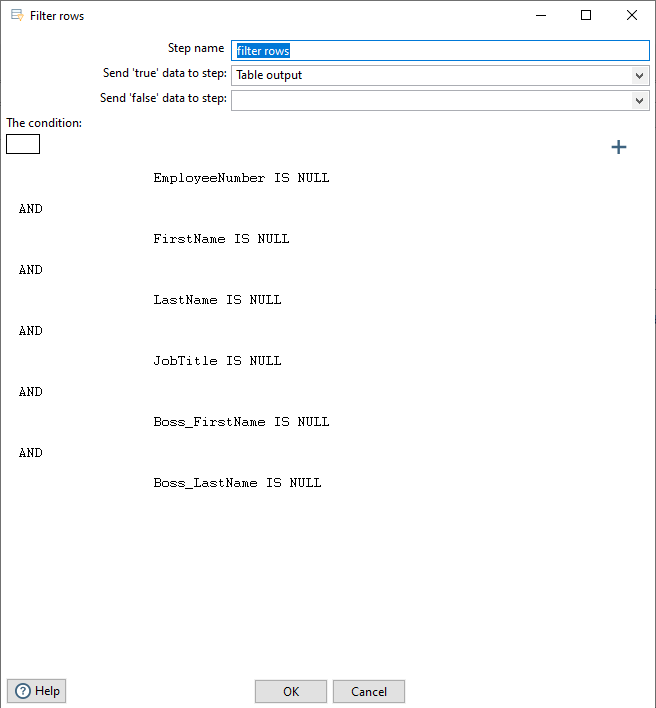
1. Hubungkan output table input pada select values.
2. Konfigurasi pada Select values yaitu mengambil data dari field employeenumber, lastname, firstname, jobtitle , lastname\_1 dan firstname\_1 sebagai data stream yang digunakan pada proses ETL pada tab select & alter.



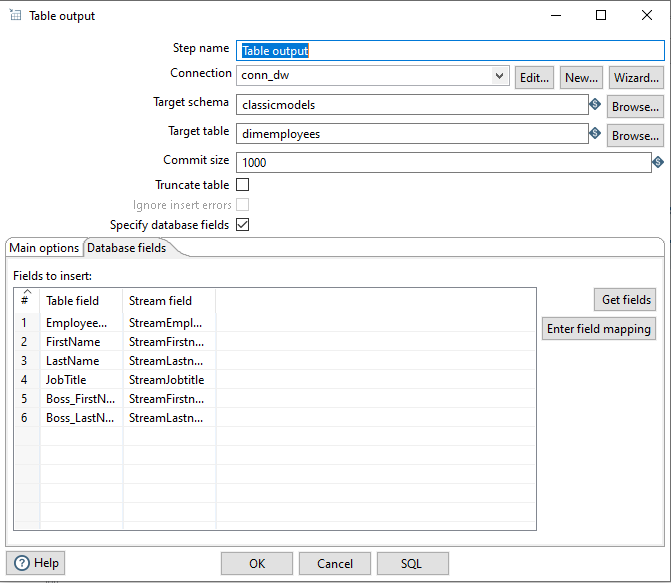
1. Hilangkan field lain yang tidak digunakan pada tab remove.
2. Hubungkan output select values pada database lookup
3. Konfigurasi pada database lookup adalah dengan menghubungkan koneksi pada conn\_dw\_destination dengan table lookup dimEmployees yang telah dibuat pada tahap pertama.
4. Field yang 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.



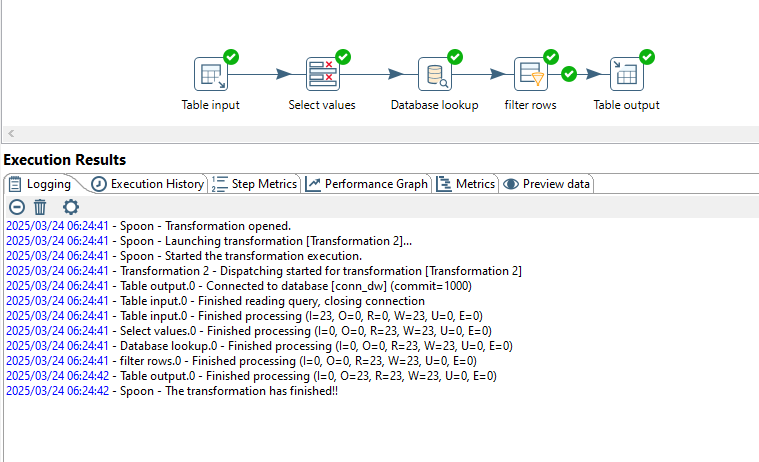
1. Hubungkan output databse lookup dengan filter rows.
2. Pada filter rows berikan kondisi field yang null pada field dimemployees untuk dimasukkan pada proses selanjutnya. Hal itu menandakan bahwa data stream belum memiliki kesamaan pada data di dimemployees.

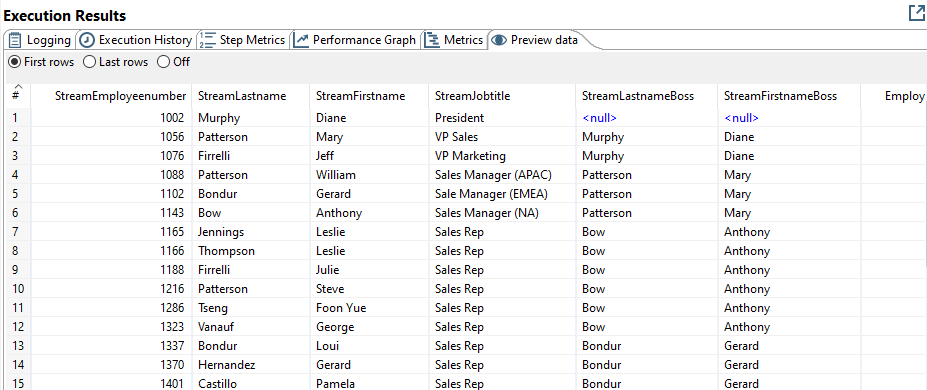


1. Hubungkan output dari filter rows dengan table output.
2. Pada table output, gunakan connection conn\_dw\_destination untuk memasukkan data pada tabel dimemployees.
3. Aktifkan specify databse fields, dan mapping data stream input dari oltp terhadap field yang ada pada dimemployees.



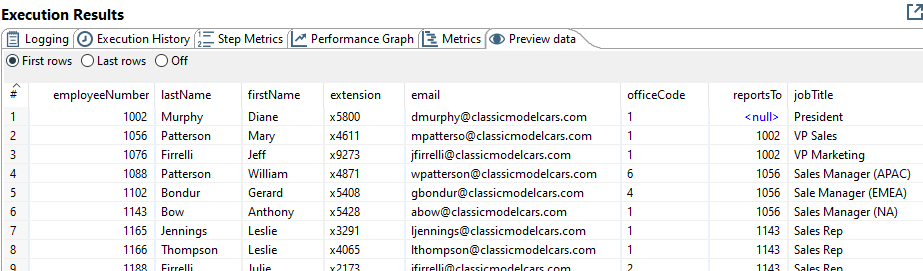
1. jika proses keseluruhan berhasil maka tabel dimemployees akan terisi data pegawai dari database OLTP.



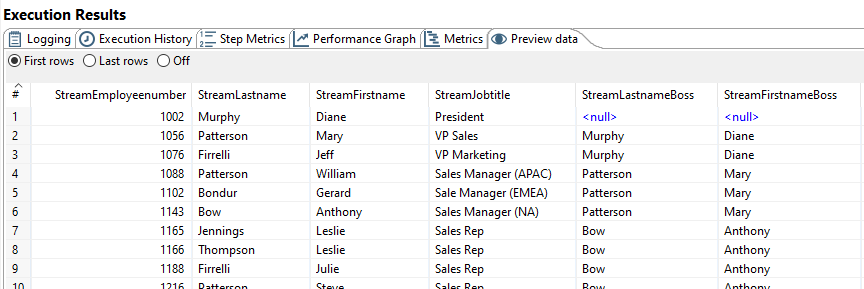


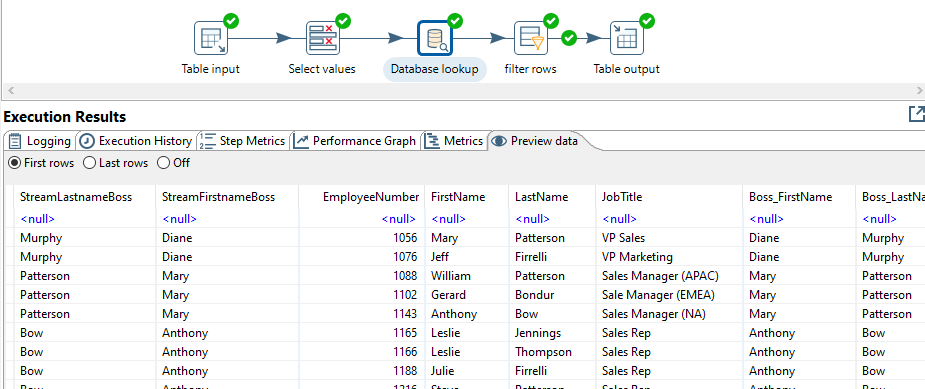
**TUGAS 2**

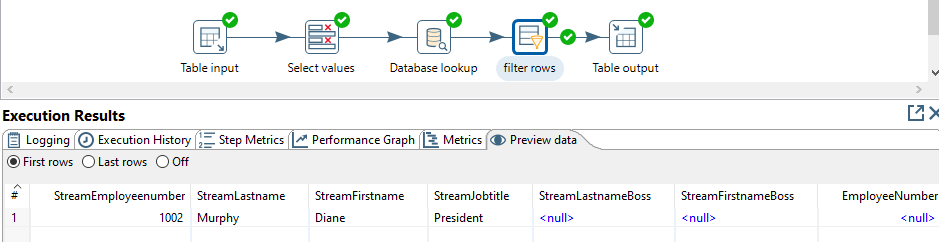
1. Buka preview tab pada execution result area di setiap proses object. amati input dan output data yang ada. bandingkan di setiap prosesnya. jelaskan perbedaan disetiap prosesnya.
2. Tabel input : pada tabel input memberikan data yang lengkap pada suatu perusahaan



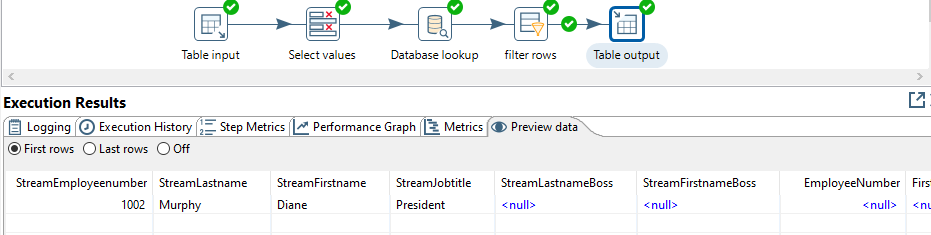
1. Select value : pada tahap ini di filter lagi yang akan di munculkan atau preview, terdapat 6 field yang dimunculkan dan di rename



1. Database lookup : data yang ada pada select value akan di teruskan ke database dan diberikan type data nya sehingga bisa diisi sesuai dengan tipe data nya 
2. Filter rows : disetting IS NULL karena data yang dimunculkan di filter rows adalah StreamLastName, StreamFirstName, EmployeeNumber, FirstName, LastName, JobTitle, Boss\_FirstName, Boss\_LastName yang berrnilai NULL



1. Table output : memberikan hasil dari filter rows yakni data yang bernilai null pada database



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

= tidak terjadi redudant dan untuk hasil tetap yakni 1 data dengan StreamEmployeeNumber 1002 (Murphy)

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

= data berhadil ditambahkan

