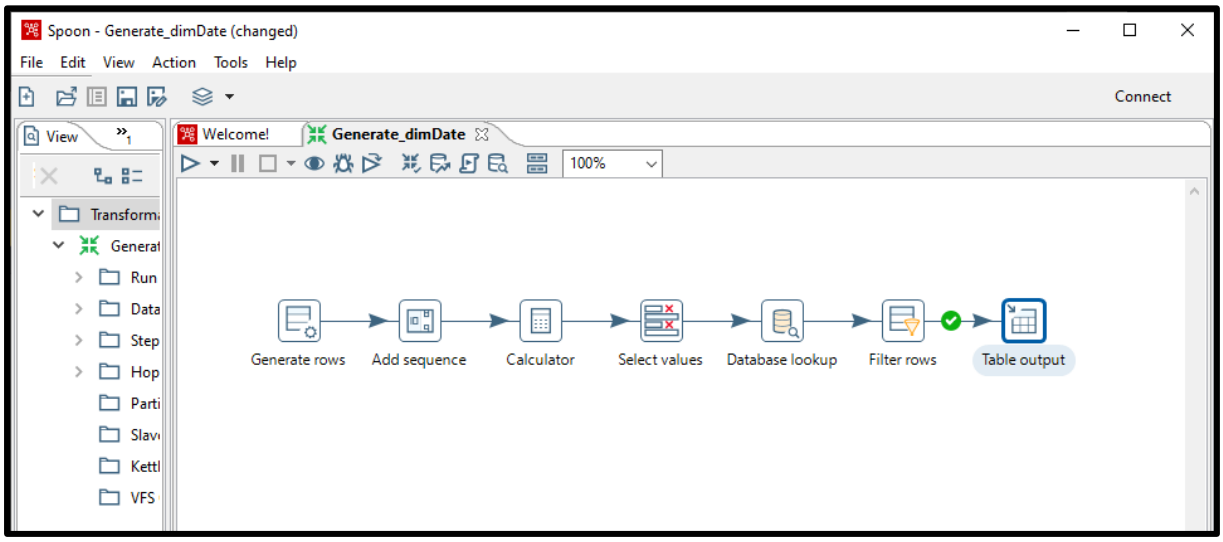
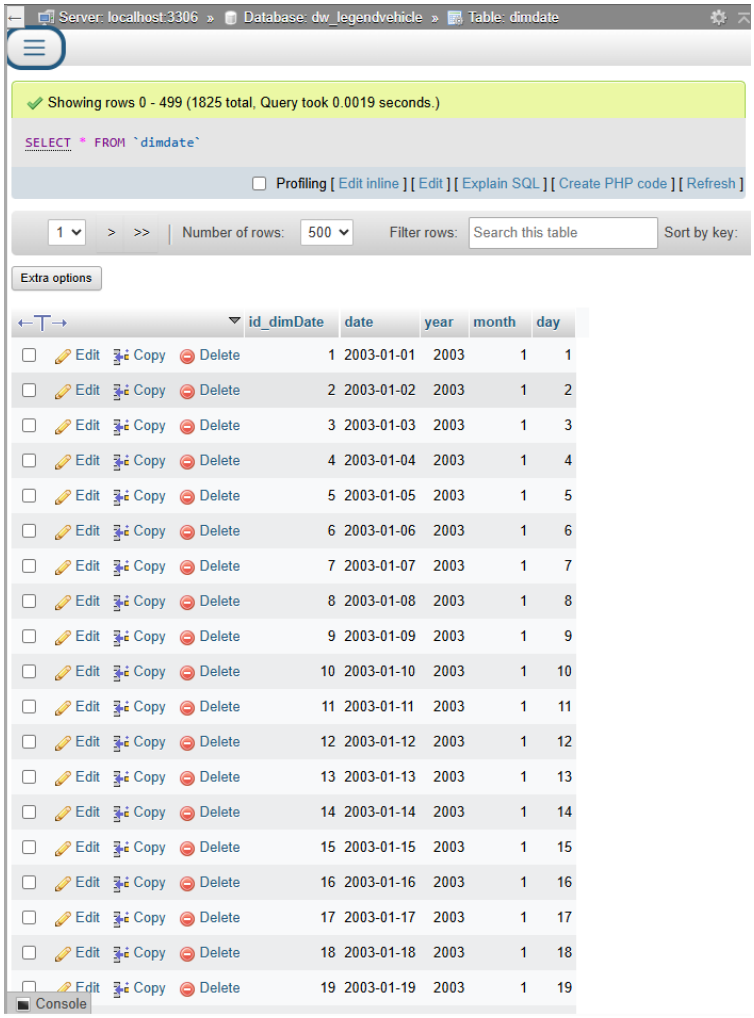
Pengerjaan Jobsheet 3

Mata Kuliah: Data Warehouse

Database Analytical

Nama : Aditya Yuhanda Putra
Kelas : SIB 2A





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.

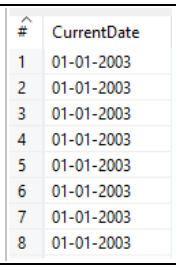
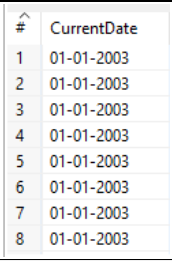
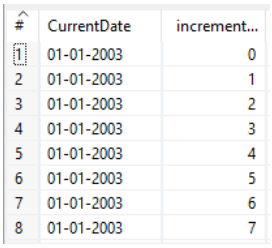
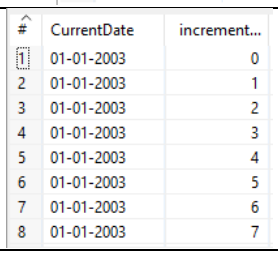
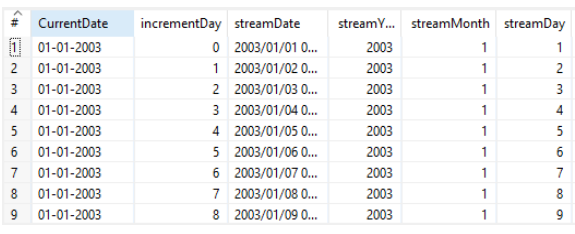
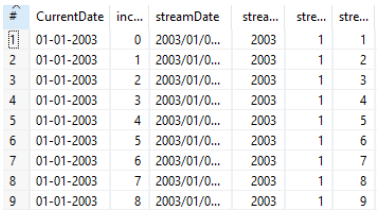
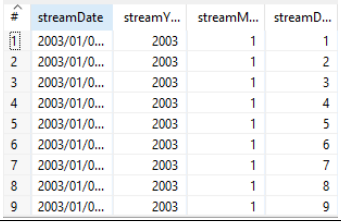
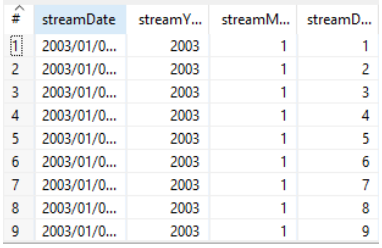
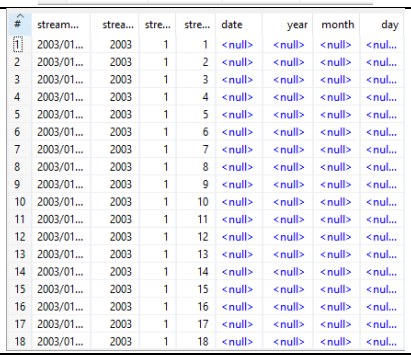
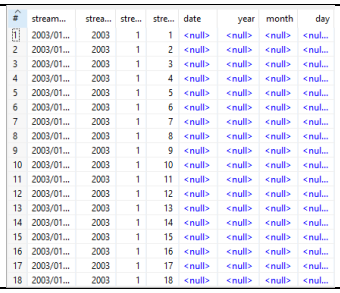
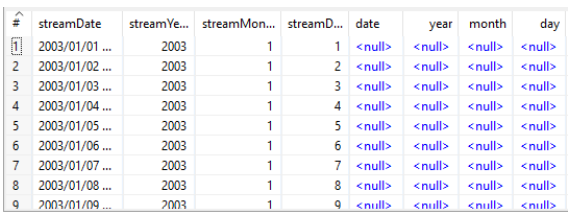
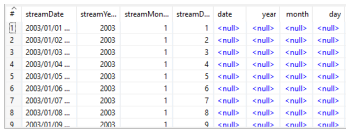
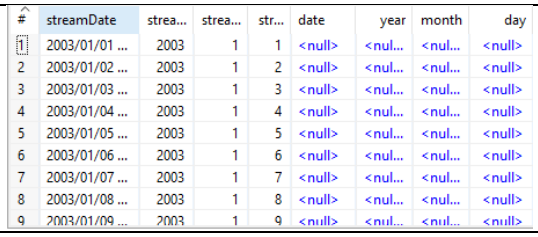
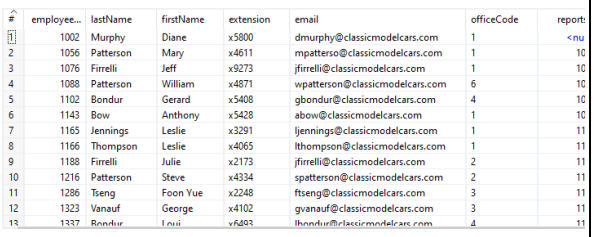
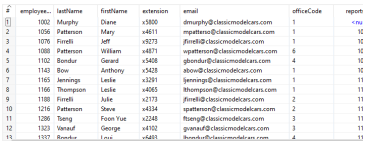
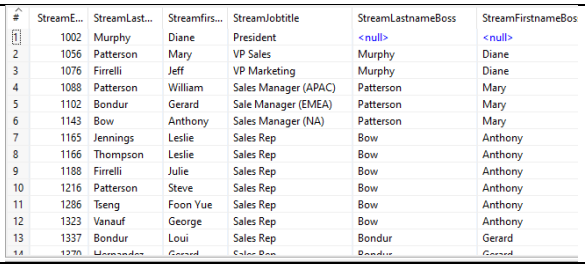
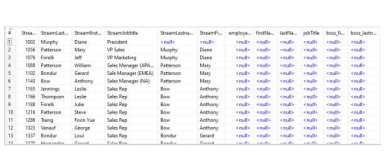
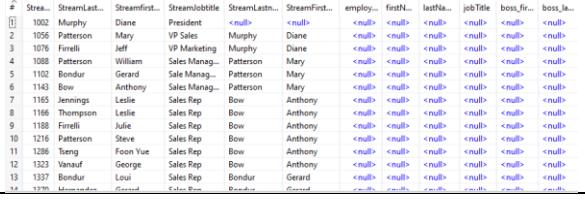
Proses Objek	SS data input	SS data output	Keterangan
Generate rows			Proses ini menghasilkan sejumlah baris data sesuai konfigurasi. Biasanya digunakan untuk membuat data dummy atau tanggal secara otomatis.
Add Sequences			Menambahkan kolom baru yang berisi nomor urut (sequence number). Biasanya digunakan untuk memberikan *ID unik* pada setiap baris data.
Calculator			Proses ini melakukan operasi kalkulasi (misalnya perhitungan tanggal, konversi nilai, atau manipulasi angka).
Select values			Memilih hanya kolom yang diperlukan dan menghilangkan kolom yang tidak diperlukan sebelum diteruskan ke proses berikutnya.
Database lookup			Proses ini mengambil referensi data dari database berdasarkan nilai tertentu, biasanya untuk mendapatkan informasi tambahan seperti nama bulan berdasarkan nomor bulan.
Filter rows			Memfilter data berdasarkan kondisi tertentu, misalnya hanya mengambil data dengan tahun >= 2000.

Table Output			Menyimpan hasil akhir ke dalam database sesuai dengan struktur tabel yang telah ditentukan.

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.

Proses Objek	SS data input	SS data output	Keterangan
Table Input			Proses **Extract**, mengambil data dari database OLTP dengan query SELECT * FROM employees e LEFT JOIN employees r ON e.reportsTo=r.employeeNumber;
Select Values			Proses Transform, memilih hanya kolom yang akan digunakan dan membuang kolom tidak relevan
Database Lookup			Proses Lookup, mengecek apakah data employeeNumber sudah ada di dimEmployees untuk menghindari duplikasi
Filter Rows			Proses Filter, hanya menyisakan data baru yang akan dimasukkan ke dimEmployees
Table Output			Proses Load, memasukkan data baru ke dalam tabel dimEmployees

2. Jika proses itu di ulangi (di run kembali) apakah data akan redudant?
   - **Ya,** Datanya redundan dan ketika saya menambahkan data saya sendiri juga redundan

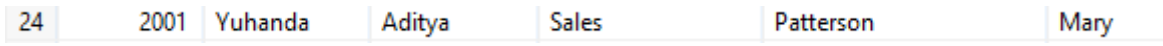
3. Tambahkan nama anda pada table employee di OLTP. jalankan kembali transformasi ini. Amati hasilnya, apa yang terjadi?
   - Tambahkan Nama Anda ke Tabel `employees` di OLTP

     Gunakan query berikut untuk menyesuaikan data dengan format tabel `classicmodels.employees`:

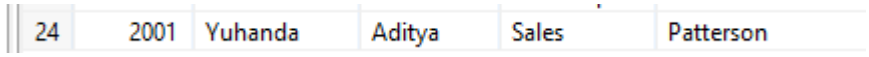
```
INSERT INTO employees (employeeNumber, lastName, firstName, extension,
email, officeCode, reportsTo, jobTitle)
VALUES (2001, 'Yuhanda', 'Aditya', 'x9999',
'ayuhanda@classicmodelcars.com', 1, 1056, 'Sales');
```

   - Jalankan Kembali Transformasi di PDI Spoon
     - **Table Input:** Data baru Aditya Yuhanda muncul.

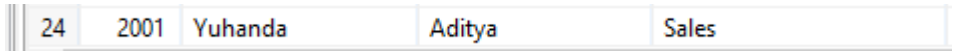
     - **Select Values:** Data tetap ada setelah seleksi kolom.



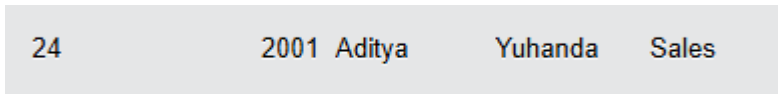
     - **Database Lookup:** Mendeteksi bahwa employeeNumber = 2001 **belum ada di dimEmployees**.



     - **Filter Rows:** Data **tidak difilter keluar**, karena merupakan pegawai baru.

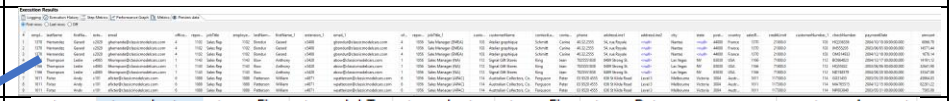
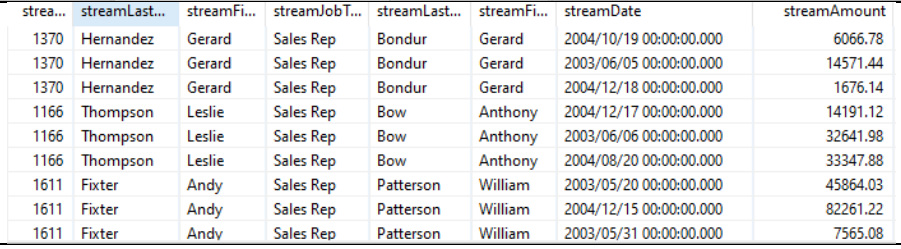


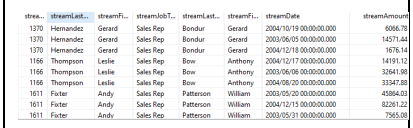
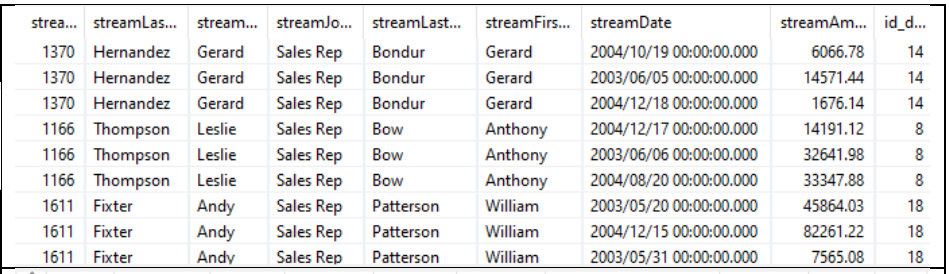
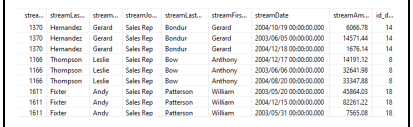
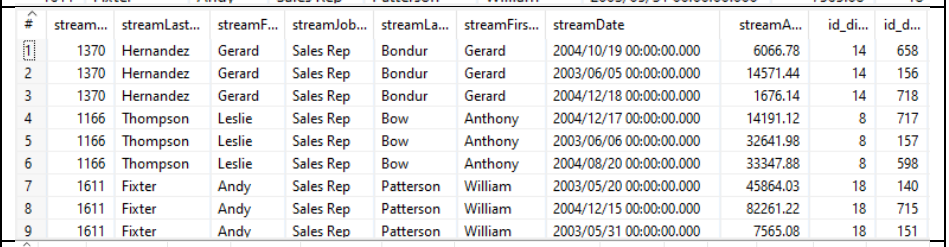
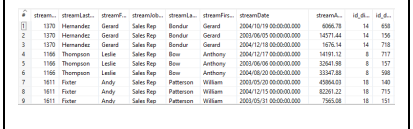
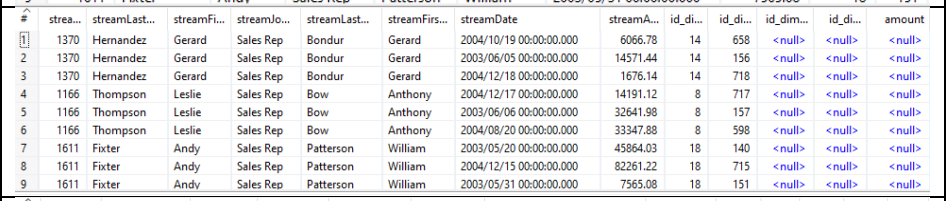
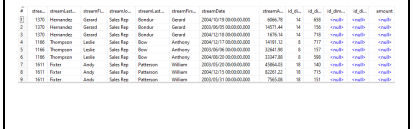
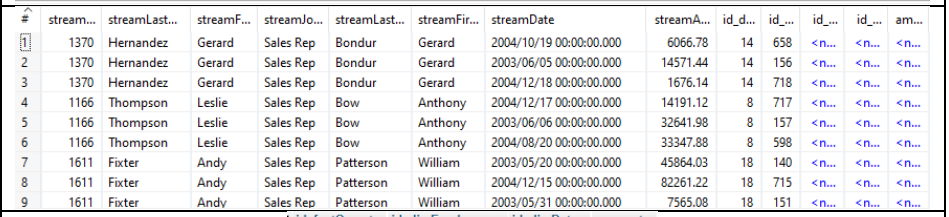
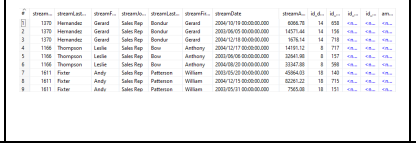
     - **Table Output:** Data Aditya Yuhanda berhasil masuk ke dimEmployees.



TUGAS 3

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
Table Input		
Select Values		

Database Lookup 1		
Database Lookup 2		
Database Lookup 3		
Filter Rows		
Table Output		

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

➢ Run Pertama :

```
2025/03/17 22:54:41 - Spoon - Transformation opened.
2025/03/17 22:54:41 - Spoon - Launching transformation [Fakta_Pembayaran]...
2025/03/17 22:54:41 - Spoon - Started the transformation execution.
2025/03/17 22:54:42 - Fakta_Pembayaran - Dispatching started for transformation [Fakta_Pembayaran]
2025/03/17 22:54:42 - Table output.0 - Connected to database [conn_dw_destination] (commit=1000)
2025/03/17 22:54:43 - Select values.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
2025/03/17 22:54:43 - Table input.0 - Finished reading query, closing connection
2025/03/17 22:54:43 - Table input.0 - Finished processing (I=273, O=0, R=0, W=273, U=0, E=0)
2025/03/17 22:54:43 - Database lookup.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)
2025/03/17 22:54:43 - Database lookup 2.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)
2025/03/17 22:54:43 - Database lookup 3.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
2025/03/17 22:54:43 - Filter rows.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
2025/03/17 22:54:44 - Table output.0 - Finished processing (I=0, O=273, R=273, W=273, U=0, E=0)
2025/03/17 22:54:44 - Spoon - The transformation has finished!!
```

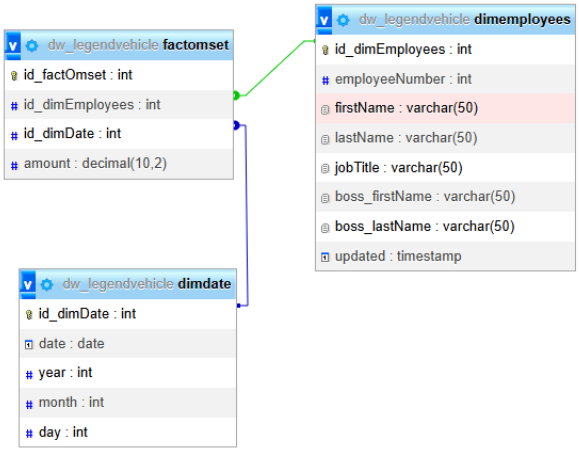
➢ Run kedua

```
2025/03/17 23:14:43 - Spoon - Transformation opened.
2025/03/17 23:14:43 - Spoon - Launching transformation [Fakta_Pembayaran]...
2025/03/17 23:14:43 - Spoon - Started the transformation execution.
2025/03/17 23:14:43 - Fakta_Pembayaran - Dispatching started for transformation [Fakta_Pembayaran]
2025/03/17 23:14:43 - Table output.0 - Connected to database [conn_dw_destination] (commit=1000)
2025/03/17 23:14:43 - Table input.0 - Finished reading query, closing connection
2025/03/17 23:14:43 - Table input.0 - Finished processing (I=273, O=0, R=0, W=273, U=0, E=0)
2025/03/17 23:14:43 - Select values.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)
2025/03/17 23:14:43 - Database lookup.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)
2025/03/17 23:14:43 - Database lookup 2.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)
2025/03/17 23:14:44 - Database lookup 3.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)
2025/03/17 23:14:44 - Filter rows.0 - Finished processing (I=0, O=0, R=273, W=0, U=0, E=0)
2025/03/17 23:14:44 - Spoon - The transformation has finished!!
```

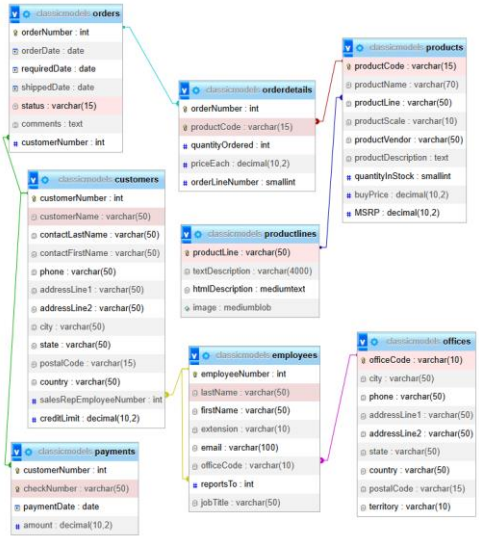
- Terlihat ketika filter rows yang di run kedua itu tidak melakukan Write lagi, yang berarti tidak ada penambahan data di tabel factomset selama tidak ada data baru yang ditambahkan

TUGAS 4

1. Buka desain database dari dw_legendvehicle pada DBMS, bandingkan design tersebut dengan desain db OLTP legendVehicle pada jobsheet 2. analisalah dan ceritakan perbedaannya.
   - OLAP dw_legendvehicle DBMS



   - OLTP classicmodel



   - Perbedaan:

Aspek	OLTP (legendVehicle - Jobsheet 2)	OLAP (dw_legendVehicle - Jobsheet 4)
Tujuan	Menyimpan dan memproses transaksi operasional harian (CRUD)	Menganalisis data transaksi dalam bentuk ringkasan dan histori
Struktur Data	Normalisasi tinggi untuk efisiensi penyimpanan dan menghindari redundansi	Denormalisasi untuk mempercepat query analisis
Kompleksitas Query	Query pendek dan sering dilakukan	Query kompleks untuk agregasi dan analisis
Frekuensi Perubahan Data	Sering diperbarui dengan transaksi baru	Data statis dan hanya diperbarui secara periodik

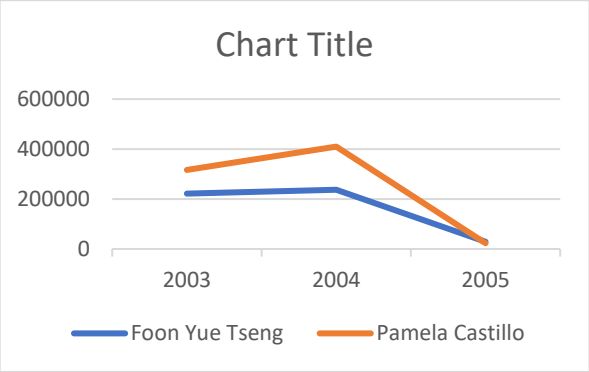
2. Buatlah report pertahun untuk KPI **"Jumlah omset yang didapat"** pada **Foon Yue Tseng** dan **Pamela Castillo**. Serta gambarkan grafiknya (grafik garis).
   - Query :

```
SELECT
    e.firstName, e.lastName,
    d.year,
    SUM(f.amount) AS total_omset
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 IN ('Foon Yue', 'Pamela') AND e.lastName IN ('Tseng', 'Castillo')
GROUP BY e.firstName, e.lastName, d.year
ORDER BY e.firstName, e.lastName, d.year;
```

   - Data:

firstName ▲ 1	lastName	year	total_omset
Foon Yue	Tseng	2003	221887.03
Foon Yue	Tseng	2004	237255.26
Foon Yue	Tseng	2005	29070.38
Pamela	Castillo	2003	317104.78
Pamela	Castillo	2004	409910.07
Pamela	Castillo	2005	23187.02

   - Grafik :



3. Jelaskan perbedaan query saat mendapatkan data pada nomor 2 dengan query pada saat Jobsheet 2!
   - Sumber Data
     - Jobsheet 2: Mengambil data langsung dari OLTP (employees, customers, payments).
     - Nomor 2: Mengambil data dari OLAP (factOmset, dimEmployees, dimDate).
   - Struktur Tabel
     - Jobsheet 2 masih menggunakan tabel transaksi asli, jadi harus melakukan banyak JOIN untuk mendapatkan nama karyawan dan tahun pembayaran.
     - Nomor 2 sudah menggunakan dimensi (dimEmployees, dimDate), jadi query lebih simpel karena tabel sudah diringkas dalam data warehouse.
   - Efisiensi Query
     - Jobsheet 2 memproses data langsung dari transaksi, jadi lebih berat.
     - Nomor 2 lebih cepat karena data sudah diproses dan disimpan di OLAP.
   - Penggunaan Tahun
     - Jobsheet 2: Harus pakai YEAR(p.paymentDate) untuk ambil tahun dari transaksi.
     - Nomor 2: Tinggal ambil d.year karena sudah ada di dimDate.

4. Simpulkan dengan bahasa sendiri, apa perbedaan OLTP dan OLAP?
   - OLTP itu buat nyimpen transaksi harian, kayak data penjualan atau order pelanggan. Struktur datanya rapi banget (ternormalisasi) biar update cepat.
   - OLAP lebih ke analisis. Datanya udah diringkas, jadi lebih gampang buat laporan atau cari tren bisnis. Ini yang dipakai kalau bos mau lihat omset tahunan tanpa nunggu lama.
   - Singkatnya, OLTP = buat kerjaan harian, OLAP = buat laporan dan analisis.