

1. [50%] Toko furniture “ANTIQUE” menjual berbagai jenis furniture yang terbuat dari kayu jati. Toko ini memiliki 5 cabang, yakni di kota Jakarta, Bandung, Pekalongan, Yogyakarta, dan Lampung. Setiap minggu, pasokan kayu jati masuk untuk diolah menjadi berbagai furniture oleh para pekerjanya, di antaranya meja, kursi, lemari, rak, dipan, dan lain-lain. Selain pembuat furniture, karyawan toko juga terdiri dari staf kasir, staf keuangan, staf pemesanan, staf pengawas kualitas produk, dan kurir. Pemesanan dapat dilakukan dengan 2 cara, yakni membeli furniture yang sudah jadi atau pre-order (PO) di mana furniture dibuat sesuai spesifikasi dari pembeli. Di saat pembeli memesan secara PO, status pemesanannya dimulai dari pesanan dibuat, perhitungan biaya, pembayaran down payment (DP), pembuatan furniture, furnishing, pelunasan, siap dikirim, dikirim, dan terakhir diterima oleh pembeli. Metode pembayaran dapat dilakukan dengan metode cash, menggunakan kartu debit atau kartu kredit, dan fintech (OVO, Dana, dan Gopay). Selama 10 tahun berdiri, pemilik toko ingin mengetahui kemajuan bisnisnya tersebut, di antaranya:
- Berapa omzet dan keuntungan toko selama 5 tahun terakhir di setiap cabangnya?
 - Berapa lama waktu rata-rata yang dibutuhkan toko untuk menyelesaikan pesanan PO di setiap cabangnya?
 - Berapa persentase omzet toko yang diterima melalui fintech yang diperoleh setiap bulannya?

Berdasarkan deskripsi toko “ANTIQUE” di atas, lakukan analisis berikut ini serta tuliskan output-nya:

- a. [15%] Tentukan tipe proses bisnis dari pertanyaan a-c di atas! Selain itu, identifikasi pula fakta dan dimensi-dimensinya! Tentukan juga tipe faktanya!

Berapa omzet dan keuntungan toko selama 5 tahun terakhir di setiap cabangnya?

Quantity (Fact) Business Process (Sales) Duration of Time (Date Dimension) Store Location (Branch Dimension)

Berapa lama waktu rata-rata yang dibutuhkan toko untuk menyelesaikan pesanan PO di setiap cabangnya?

Quantity (Fact) Duration of Time (Date Dimension) Business Process (PurchaseOrder) Store Location (Branch Dimension)

Berapa persentase omzet toko yang diterima melalui fintech yang diperoleh setiap bulannya?

Quantity (Fact) Business Process (FintechPayment) Payment Option Duration of Time (Date Dimension)

Tipe proses bisnis saya tuangkan ke dalam Enterprise Bus Matrix sebagai berikut

| Business Process | Atomic Granularity | Metrics | Conformed Dimensions | | | | |
|------------------|-------------------------------|---|----------------------|---------|-----------|----------|---------|
| | | | Date | Product | Promotion | Salesrep | Fintech |
| Sales | 1 row per sales order line | Order quantity, amount (IDR), discount, profit, promo | x | x | x | x | |
| PurchaseOrder | 1 row per purchase order line | Order quantity, amount (IDR), discount, profit, promo | x | x | x | x | |

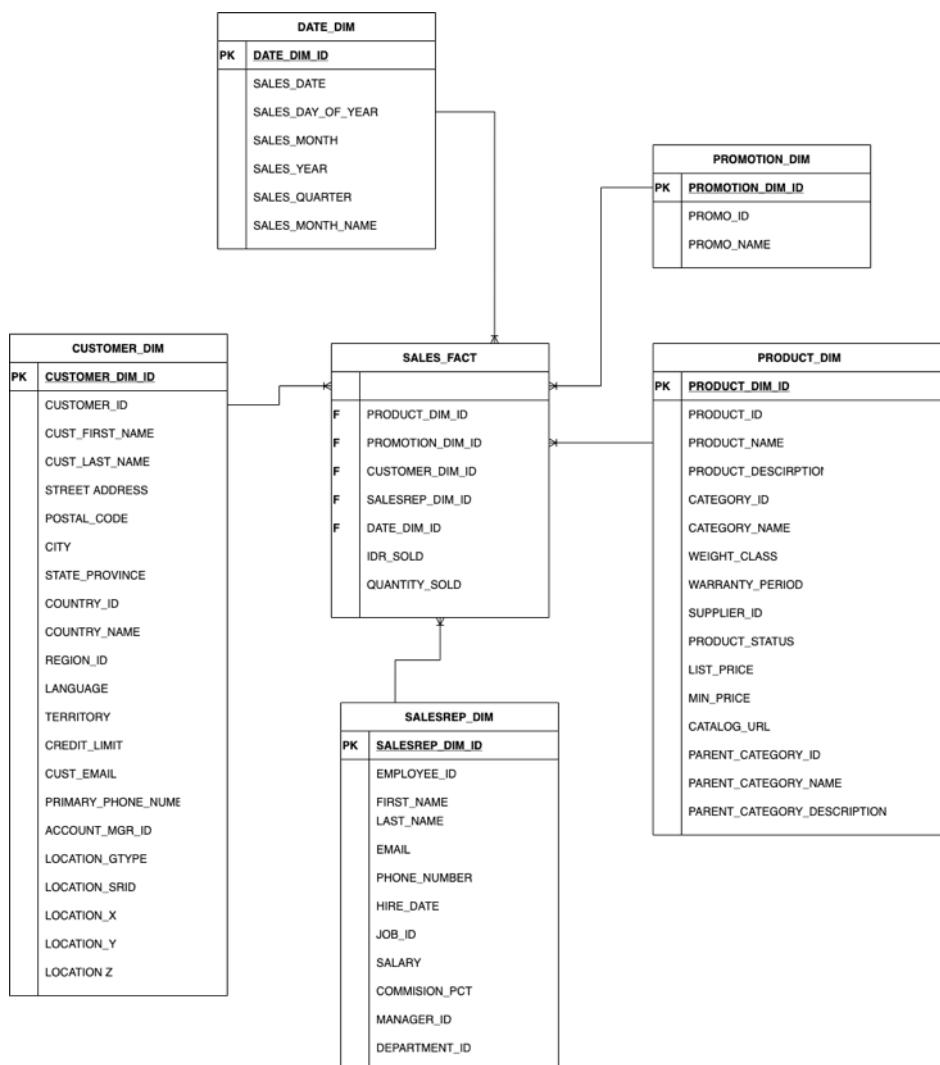
| | | | | | | | |
|----------------|---------------------------------|---|---|---|---|---|---|
| FintechPayment | 1 row per fintech payment order | Order quantity, amount (IDR), discount, profit, promo | x | x | x | x | x |
|----------------|---------------------------------|---|---|---|---|---|---|

b. [25%] Buatlah *star schema*-nya berdasarkan deskripsi studi kasus beserta pertanyaan a-c di atas!

- **SALES_FACT**

Berapa omzet dan keuntungan toko selama 5 tahun terakhir di setiap cabangnya?

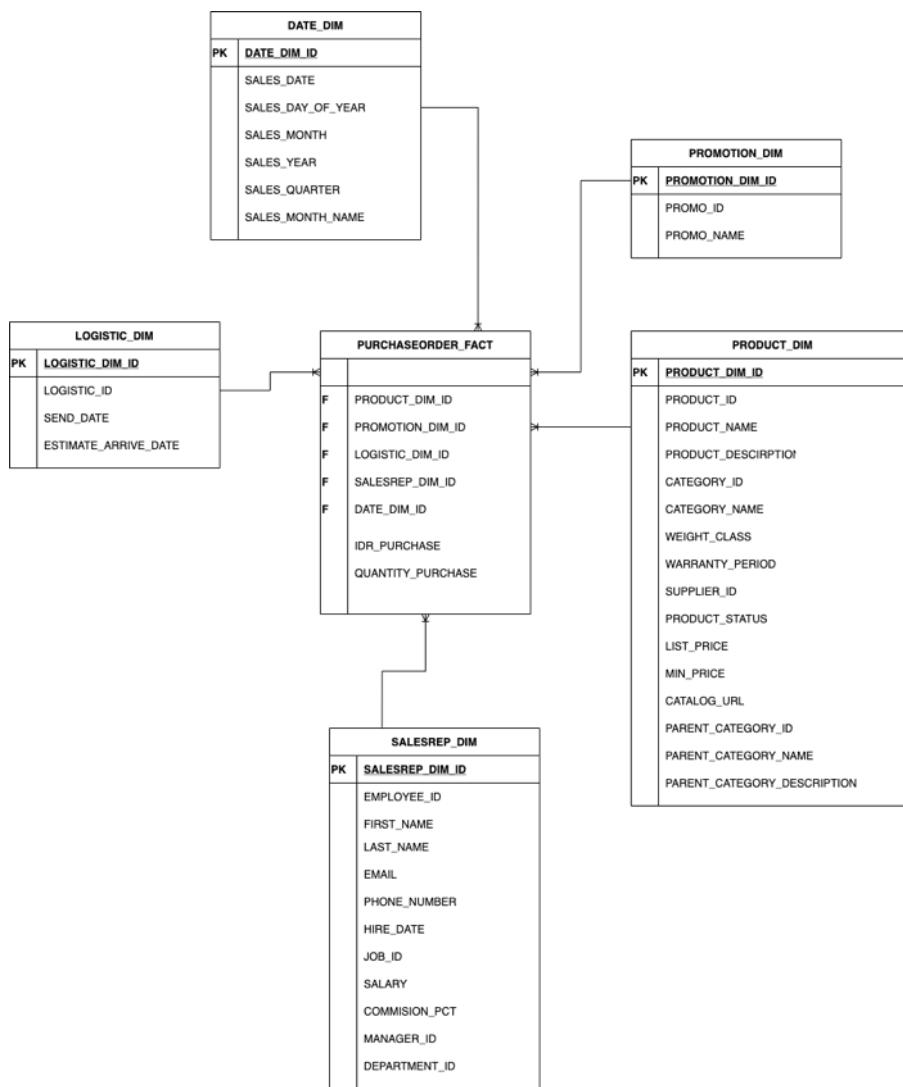
Quantity (Fact) Business Process (Sales) Duration of Time (Date Dimension) Store Location (Branch Dimension)



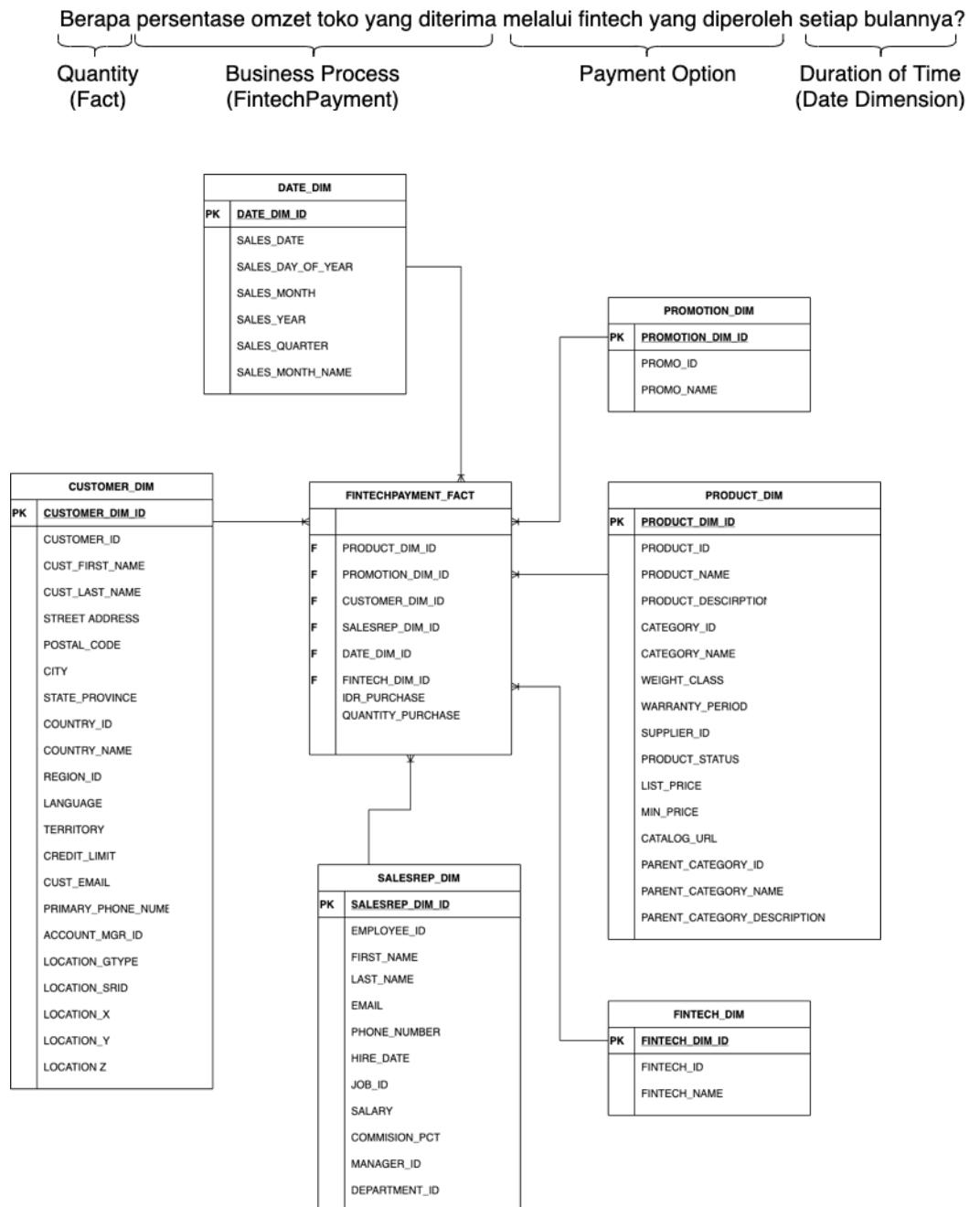
- PURCHASEORDER_FACT

Berapa lama waktu rata-rata yang dibutuhkan toko untuk menyelesaikan pesanan PO di setiap cabangnya?

Quantity (Fact) Duration of Time (Date Dimension) Business Process (PurchaseOrder) Store Location (Branch Dimension)



- FINTECHPAYMENT_FACT



- c. [10%] Proses *pre-processing* apa saja yang mungkin Anda terapkan sebagai bagian dari proses *Extract, Transform, Load* (ETL) terhadap data-data yang dimiliki Toko "ANTIQUE" tersebut?

Proses Extract dilakukan dengan membuat DataWarehouse dan menyesuaikan tipe data sehingga dapat diproses.

Proses Transform dilakukan untuk menyesuaikan data berdasarkan tanggal pada pertanyaan bisnis dengan generate dimensi baru yaitu Date.

Kemudian proses Load dilakukan untuk dapat menginput data baru ke dalam database di DataWarehouse dengan tipe SCD 2 terhadap keseluruhan dimensi tabel yang ada agar lebih mudah diproses.

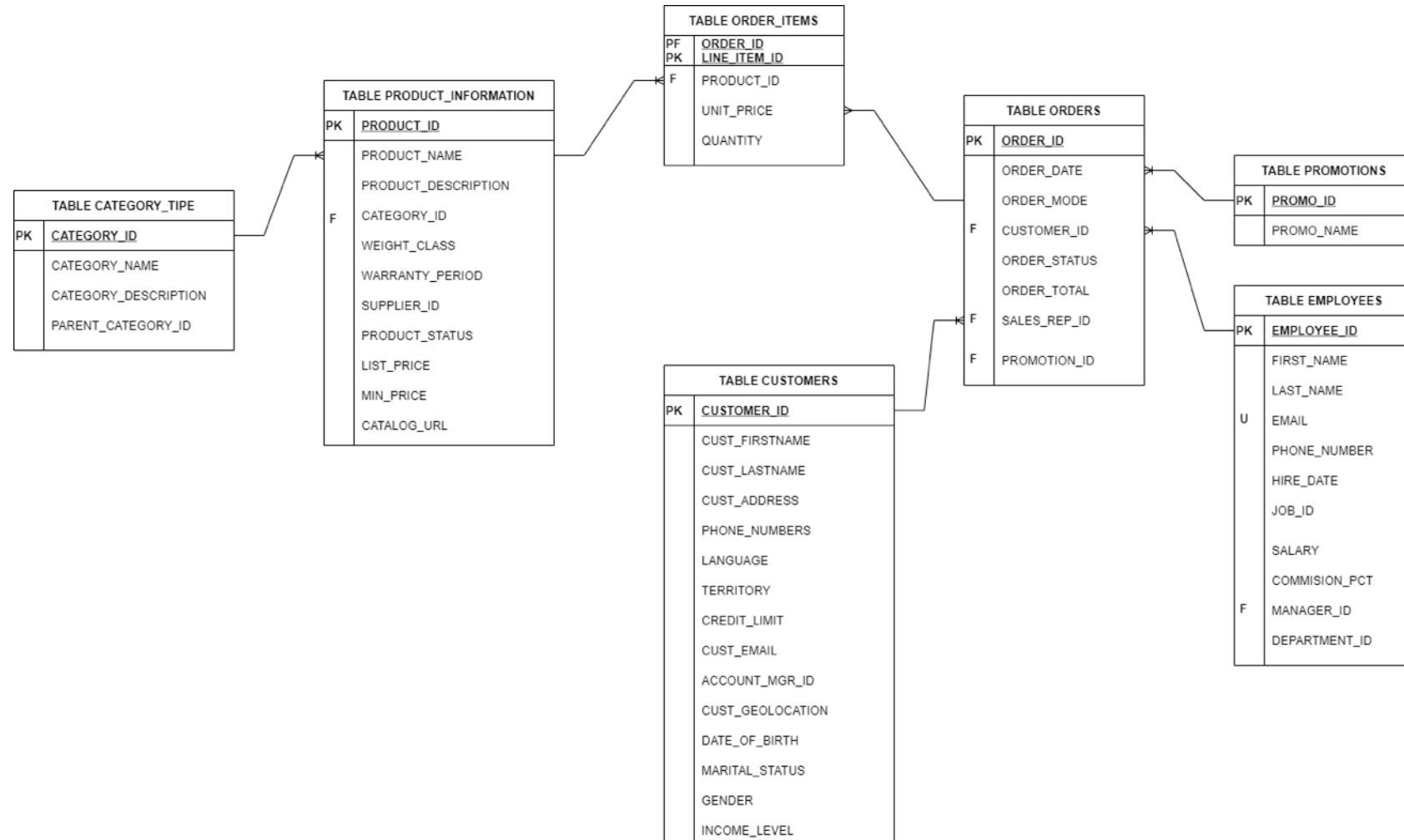
2. [50%] Buatlah video yang menjelaskan poin-poin berikut. Perhatikan bahwa wajah Anda juga harus masuk dalam video tersebut. Jawaban di-upload di Youtube kemudian *link*-nya diberikan dalam jawaban soal. Untuk memastikan keunikan video Anda dibandingkan dengan kawan Anda, maka sebaiknya tabel-tabel yang Anda gunakan sebagai contoh dalam video adalah tabel yang Anda buat sendiri atau dari kantor Anda sendiri.

Video telah diupload pada link

<https://youtu.be/En6vQvhUIDE>

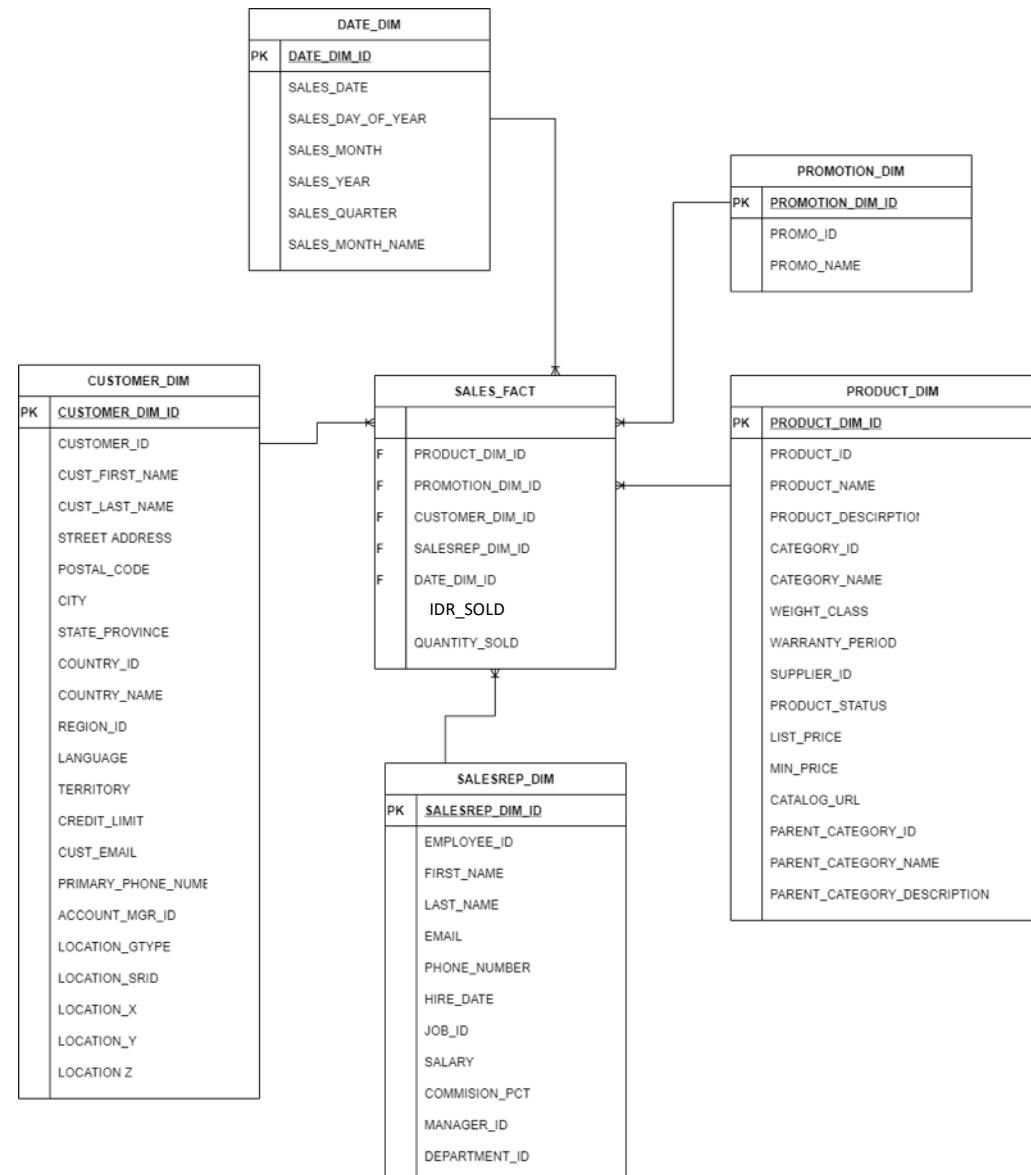
Namun suaranya kecil terkendala masalah pada mic

ERD ELECTRONIC STORE



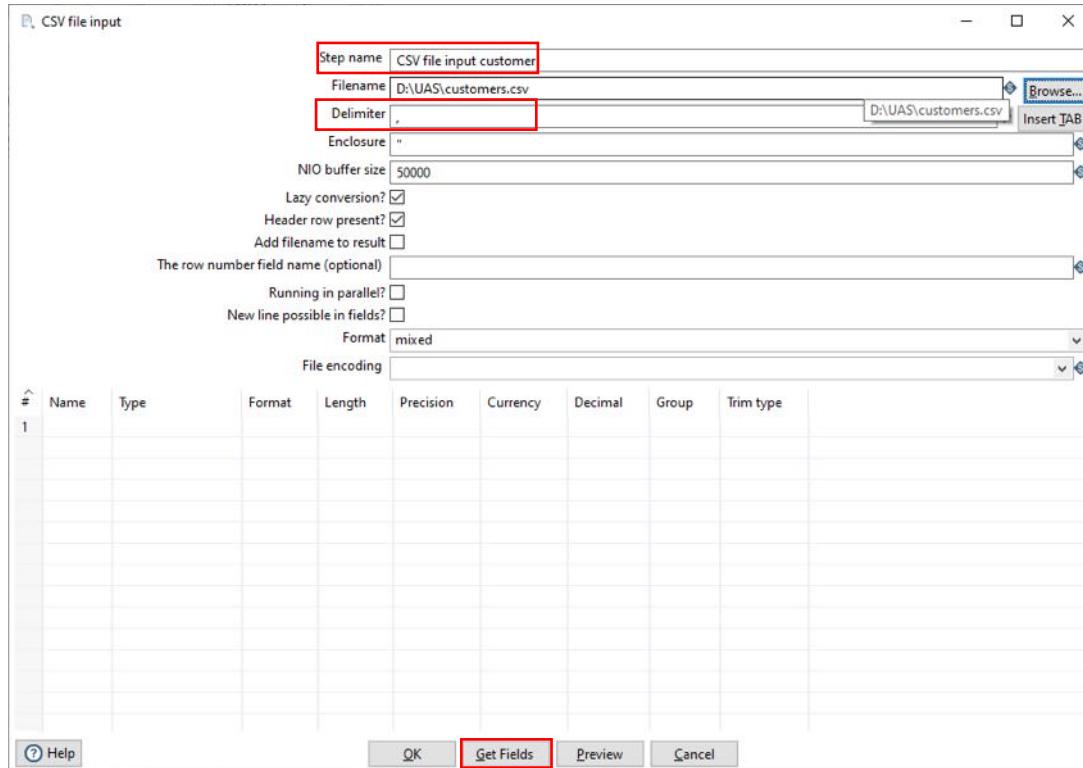
Untuk mencari proses bisnis dalam penjualan (SALES), maka perlu dibuatkan Star Schema dari ERD diatas sebagai berikut.

STAR SCHEMA SALES

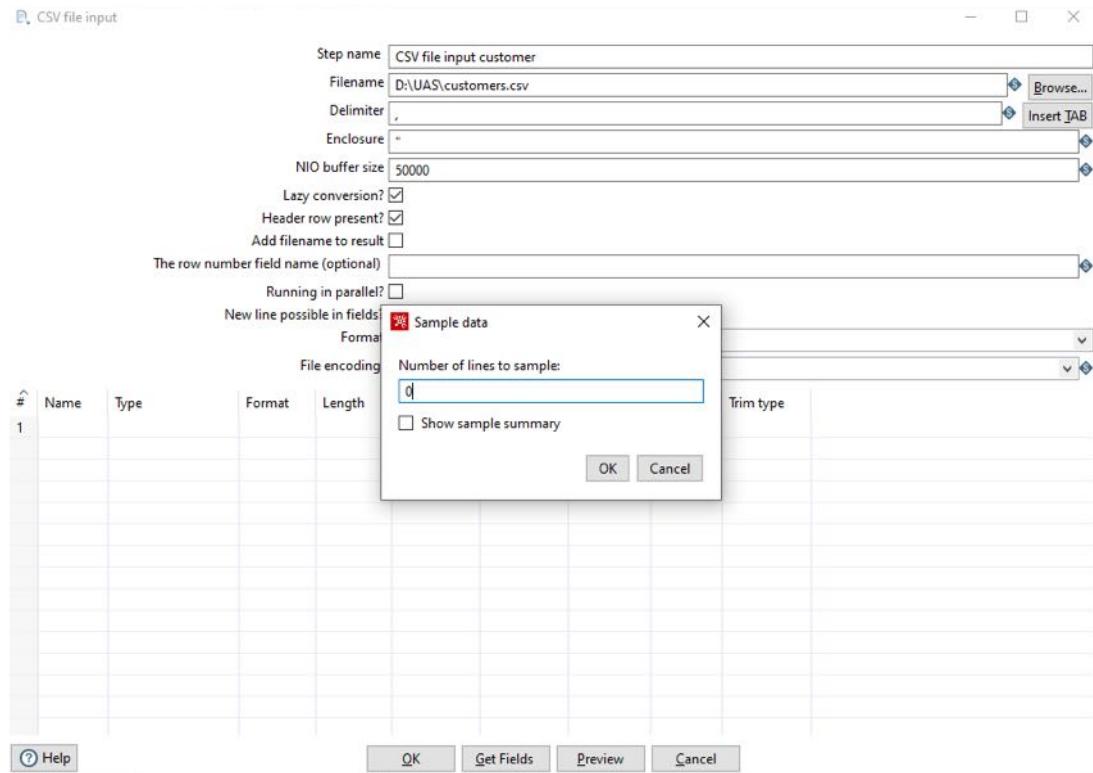


[15%] Jelaskan dengan detail bagaimana melakukan Type 2 SCD. Gunakan contoh *real*. *Step by step* dengan jelas bagaimana Anda mengimplementasikan Type 2 SCD tersebut, termasuk apa yang terjadi ketika data dalam tabel dimensi ada perubahan. Susunlah *scene* Anda sedemikian sehingga jelas perbedaan sebelum ada perubahan pada tabel dimensi, dan setelah ada perubahan.

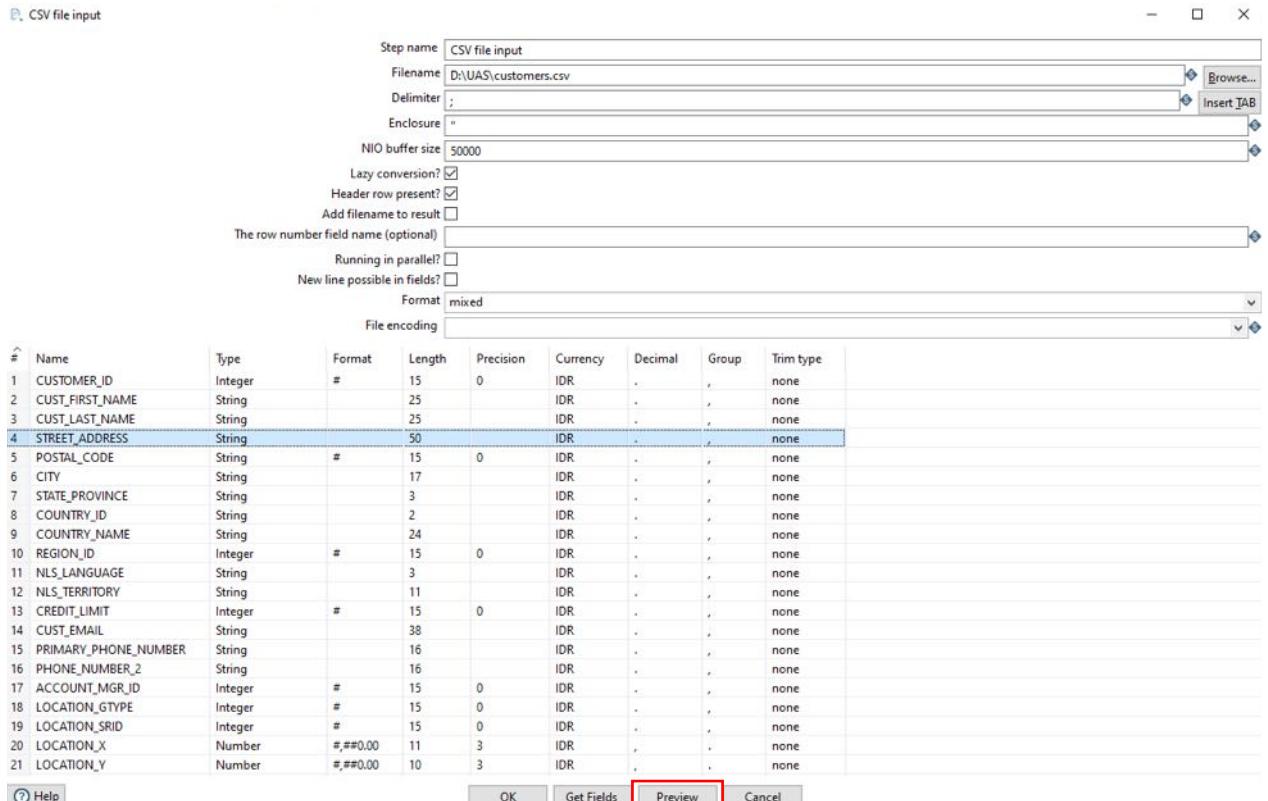
1. CUSTOMER_TRANSFORM (DIMENSION TABLE)



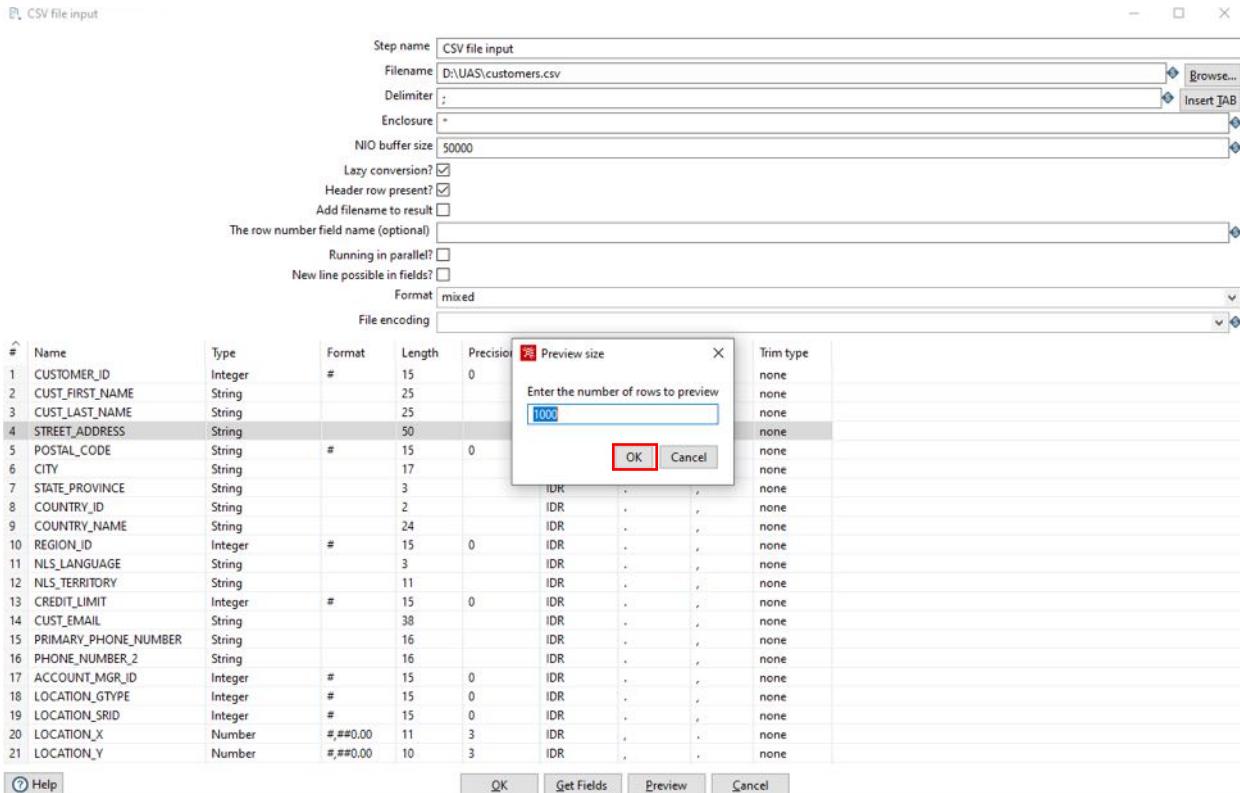
- Berikan nama pada Step name, dalam hal ini yang digunakan adalah customer
- Pilih filename
- Pastikan Delimeter telah sesuai
- Kemudian pilih GetFields



- Input 0 untuk menampilkan seluruh fields



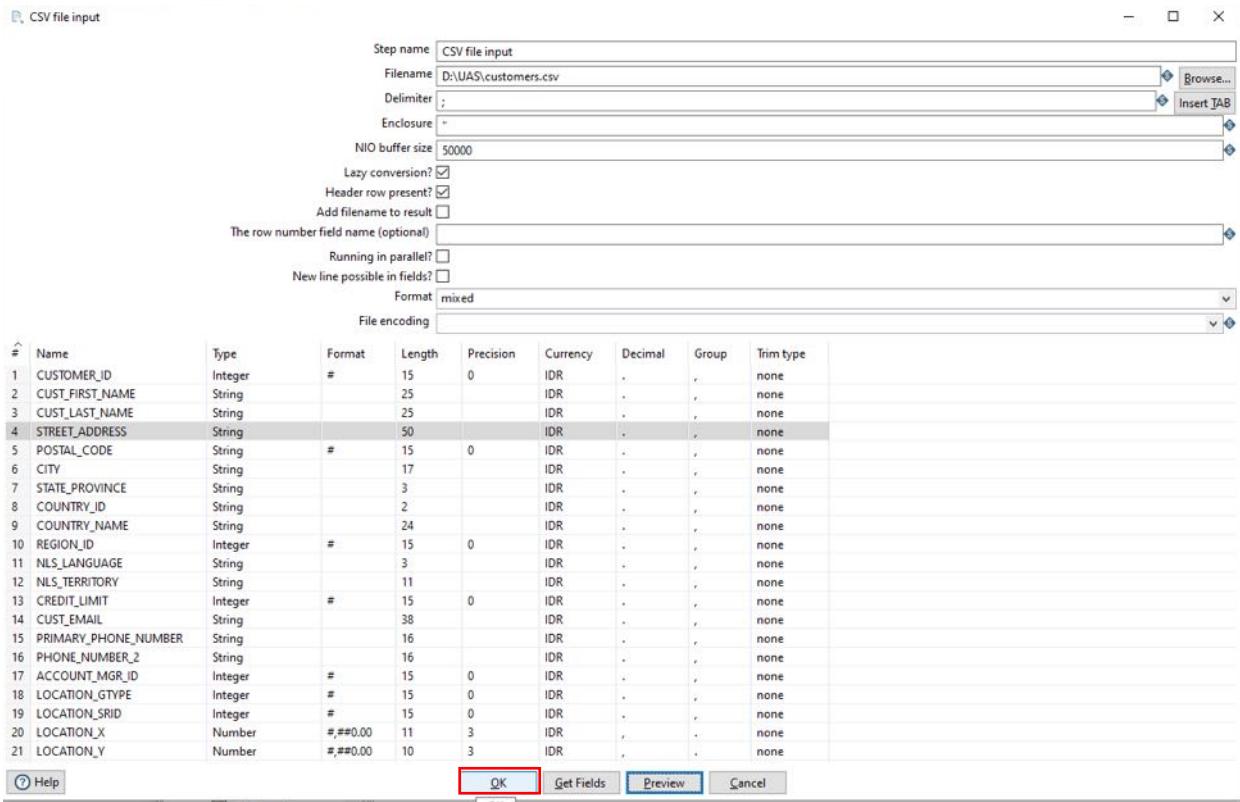
- Lakukan penyesuaian fields data
- Setelah sesuai maka dapat melihat data dengan pilih Preview



- Tentukan jumlah data yang ingin ditampilkan
- Pilih OK

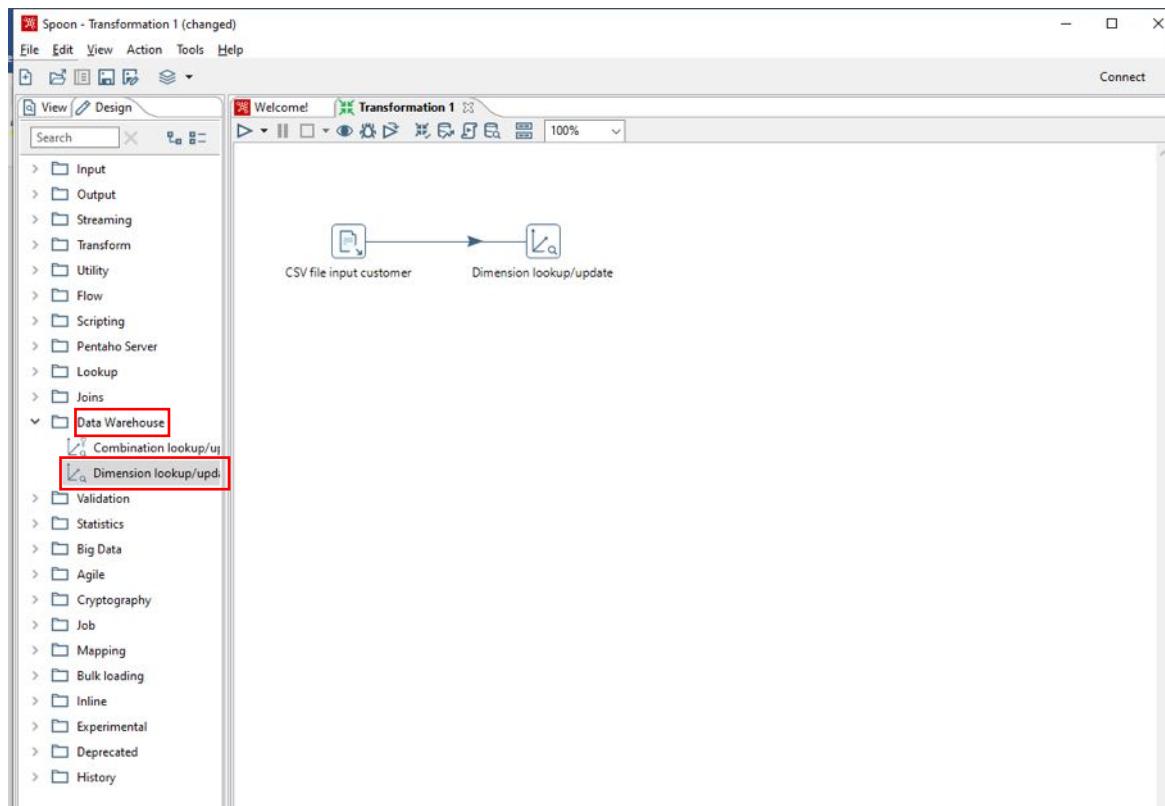
| Examine previous data | | | | | | | | | | | | | | | | |
|-----------------------|-------------|-----------------|----------------|----------------------------|-------------|---------------|----------------|------------|--------------------------|-----------|--------------|---------------|--------------|---------------------------------------|----------------------|----------------|
| # | CUSTOMER_ID | CUST_FIRST_NAME | CUST_LAST_NAME | STREET_ADDRESS | POSTAL_CODE | CITY | STATE_PROVINCE | COUNTRY_ID | COUNTRY_NAME | REGION_ID | NLS_LANGUAGE | NLS_TERRITORY | CREDIT_LIMIT | CUST_EMAIL | PRIMARY_PHONE_NUMBER | PHONE_NUMBER_2 |
| 1 | 145 | PATRICIA | JOHNSON | 1234 Bright St | 54701 | Eau Claire | WI | US | United States of America | 2 | us | AMERICA | 1000 | Maureen.Mindred@FREEEXAMPLE.COM | 555-1234-5678 | |
| 2 | 146 | ANITA | WILLIAMS | 1456 Rockwell Ct | 54717 | Madison | WI | US | United States of America | 2 | us | AMERICA | 200 | John.President@JACANA.EXAMPLE.COM | 555-1234-5678 | |
| 3 | 147 | BARBARA | JONES | 6555 W Good Hope Rd | 53228 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 600 | Janet.Warren@LAPWING.EXAMPLE.COM | 555-1234-5678 | |
| 4 | 148 | ELIZABETH | BROWN | 1234 N Shattock Rd | 53714 | Madison | WI | US | United States of America | 2 | us | AMERICA | 600 | Gustav.Steuberg@PNTIL.EEXAMPLE.COM | 555-1234-5678 | |
| 5 | 149 | JENNIFER | DAVIS | 4711 Sprecher Rd | 53704 | Madison | WI | US | United States of America | 2 | us | AMERICA | 600 | Markus.Rampling@PUFFNILEXAMPLE.COM | 555-1234-5678 | |
| 6 | 150 | MARIA | MILLER | 6161 N 54th St | 53218 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Godrie.State@PYRHULOXIA.EXAMPLE.COM | 555-1234-5678 | |
| 7 | 151 | SUSAN | WILSON | 11016 W Lincoln Ave | 53227 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Divine.Alykey@REDSTART.EXAMPLE.COM | 555-1234-5678 | |
| 8 | 152 | MARGARET | MOORE | 8600 W National Ave | 53227 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Dieter.Matthau@VERDORN.EXAMPLE.COM | 555-1234-5678 | |
| 9 | 153 | DOROTHY | TAYLOR | 615 N Sherman Ave | 53704 | Madison | WI | US | United States of America | 2 | us | AMERICA | 700 | Divine.Sheron@COVIBRD.EXAMPLE.COM | 555-1234-5678 | |
| 10 | 154 | KAREN | ANDERSON | 512 E Grand Ave | 53711 | Beloit | WI | US | United States of America | 2 | us | AMERICA | 700 | Fredricka.Romina@CURLYR.EXAMPLE.COM | 555-1234-5678 | |
| 11 | 155 | NANCY | THOMAS | 800 N Broadway Pl 1 | 53202 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Frederico.Romina@CURLYR.EXAMPLE.COM | 555-1234-5678 | |
| 12 | 156 | KAREN | JACKSON | 5235 N Ironwood Ln | 53217 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Godrie.Montana@DOPPERER.EXAMPLE.COM | 555-1234-5678 | |
| 13 | 157 | BETTY | WHITE | 322 E Michigan St | 53202 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Sidney.Caphane@DOUNIN.EXAMPLE.COM | 555-1234-5678 | |
| 14 | 158 | HELEN | HARRIS | 1402 Bellinger St Fl 4 | 54703 | Eau Claire | WI | US | United States of America | 2 | us | AMERICA | 700 | Frederic.Lyon@FLICKER.EXAMPLE.COM | 555-1234-5678 | |
| 15 | 159 | SANDRA | MARTIN | 411 E Wisconsin Ave # 2550 | 53205 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 700 | Eddie.Boyer@GALLINOLE.EXAMPLE.COM | 555-1234-5678 | |
| 16 | 160 | DONNA | THOMPSON | 808 3rd St # 100 | 54408 | Wisauau | WI | US | United States of America | 2 | us | AMERICA | 700 | Eddie.Stern@GOONTIE.EXAMPLE.COM | 555-1234-5678 | |
| 17 | 161 | CAROL | GARCIA | 300 Crooks St | 54301 | Green Bay | WI | US | United States of America | 2 | us | AMERICA | 900 | Ernest.Wesper@GROSEAK.EXAMPLE.COM | 555-1234-5678 | |
| 18 | 162 | PAUL | MARTINEZ | 123 E Dayton St | 53703 | Madison | WI | US | United States of America | 2 | us | AMERICA | 900 | Ernest.George@HORNING.EXAMPLE.COM | 555-1234-5678 | |
| 19 | 163 | SHARON | ROBISON | 2120 N 4th St | 54201 | La Crosse | WI | US | United States of America | 2 | us | AMERICA | 1200 | Frederic.Lyon@FLICKER.EXAMPLE.COM | 555-1234-5678 | |
| 20 | 164 | MICHELLE | CLARK | 3324 N Oakland Ave | 53211 | Milwaukee | WI | US | United States of America | 2 | us | AMERICA | 1200 | Charlette.Fonda@MOSOPHEN.EXAMPLE.COM | 555-1234-5678 | |
| 21 | 165 | Laura | RODRIGUEZ | 666 23rd Ave Ne | 55418 | Minneapolis | MN | US | United States of America | 2 | us | AMERICA | 1200 | Dheeriq.Alexander@THAZCHI.EXAMPLE.COM | 555-1234-5678 | |
| 22 | 166 | SARAH | LEWIS | 1501 Lovery Ave N | 55411 | Minneapolis | MN | US | United States of America | 2 | us | AMERICA | 1200 | Gerard.Hershey@PARULIA.EXAMPLE.COM | 555-1234-5678 | |
| 23 | 167 | KIMBERLY | LEE | 4500 Yosemite Ave # 100 | 55416 | Minneapolis | MN | US | United States of America | 2 | us | AMERICA | 1200 | Dheeriq.David@PUPPE.EXAMPLE.COM | 555-1234-5678 | |
| 24 | 168 | DEBORAH | WALKER | 2800 Chicago Ave # 100 | 55407 | Minneapolis | MN | US | United States of America | 2 | us | AMERICA | 1200 | Henna.Powell@SANDERLING.EXAMPLE.COM | 555-1234-5678 | |
| 25 | 169 | JESSICA | HALL | 200 1st St Sw | 55905 | Rochester | MN | US | United States of America | 2 | us | AMERICA | 1200 | Henna.Powell@SANDERLING.EXAMPLE.COM | 555-1234-5678 | |
| 26 | 170 | SHIRLEY | ALLEN | 314 W Superior St # 1015 | 55802 | Duluth | MN | US | United States of America | 2 | us | AMERICA | 1200 | Harry.McPackingah@DVRON.EXAMPLE.COM | 555-1234-5678 | |
| 27 | 171 | LYNN | YOUNG | 2725 N Franklin Rd | 55231 | Syracuse | NY | US | United States of America | 2 | us | AMERICA | 1200 | Janice.McPackingah@DVRON.EXAMPLE.COM | 555-1234-5678 | |
| 28 | 172 | ANGELA | HERNANDEZ | 2200 W Genesee St | 55219 | Syracuse | NY | US | United States of America | 2 | us | AMERICA | 1200 | Christie.Fonda@MOSOPHEN.EXAMPLE.COM | 555-1234-5678 | |
| 29 | 173 | MELISSA | KING | 1532 Main St | 14306 | Madison Falls | NY | US | United States of America | 2 | us | AMERICA | 1200 | Dheeriq.Oliver@EXILERI.EEXAMPLE.COM | 555-1234-5678 | |
| 30 | 174 | BRENDA | WRIGHT | 49 N Pine Ave | 12208 | Albany | NY | US | United States of America | 2 | us | AMERICA | 1200 | Gerhard.Siegner@JACANA.EXAMPLE.COM | 555-1234-5678 | |
| 31 | 175 | AMY | LOPEZ | 137 Lark St | 12210 | Albany | NY | US | United States of America | 2 | us | AMERICA | 1200 | HarryDean.Forest@KIXKADEE.EXAMPLE.COM | 555-1234-5678 | |
| 32 | 176 | ANNA | HILL | 33 Fulton St | 12601 | Poughkeepsie | NY | US | United States of America | 2 | us | AMERICA | 1200 | HarryDean.Cage@LAPWING.EXAMPLE.COM | 555-1234-5678 | |

- Pastikan kembali kesesuaian data
- Apabila sudah selesai maka dapat klik Close

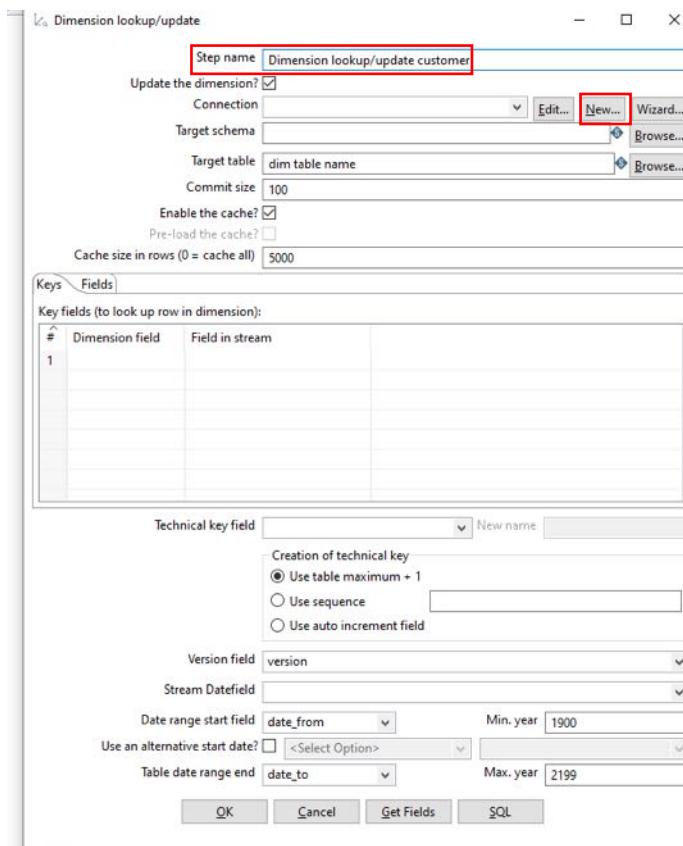


- Pilih OK

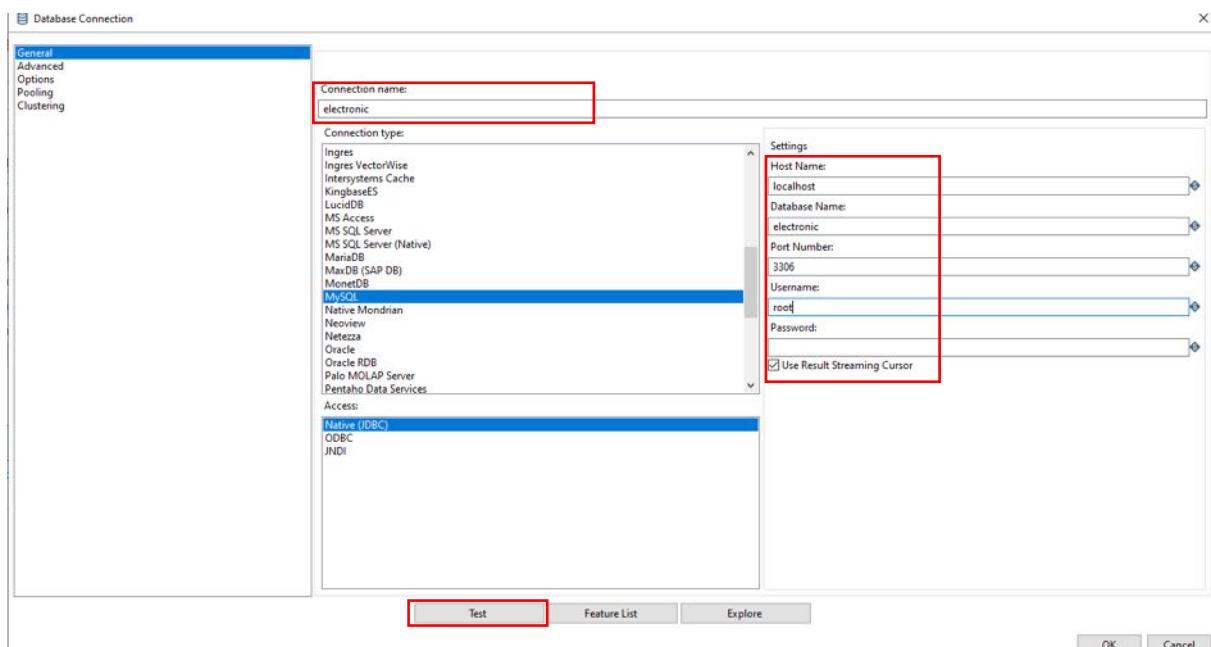
Setelah Langkah diatas sudah dilaksanakan, selanjutnya kita perlu membuat dimension customer (CUSTOMER_DIM) menggunakan Data Warehouse – Dimension lookup/update dan koneksi menggunakan konektor.



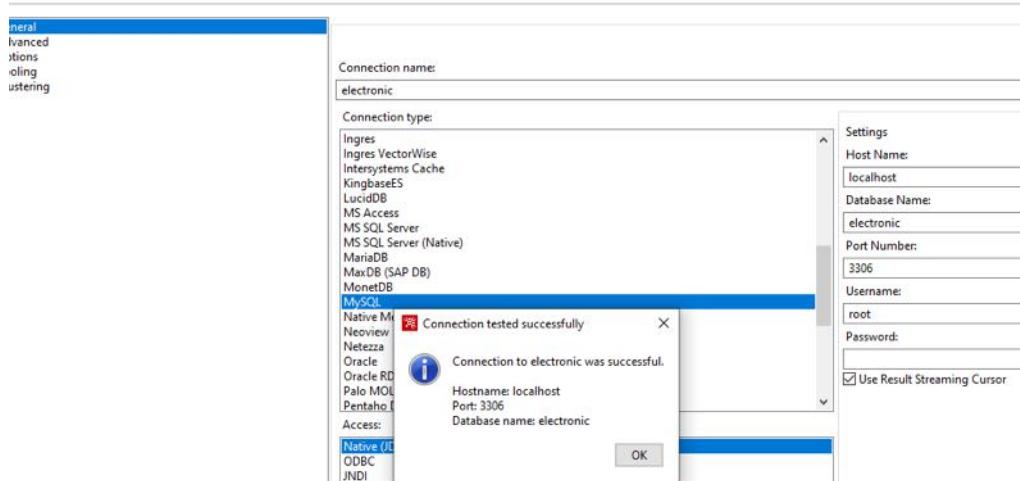
- Double klik pada Dimension lookup/update



- Berikan nama pada Step name
- Buat koneksi database, dalam hal ini saya akan gunakan MySQL dengan klik tombol New



- Berikan nama koneksi, untuk memudahkan disamakan saja dengan nama database pada MySQL yang sudah dibuat
- Sesuaikan jaringan host, port, username dan juga passwordnya
- Lakukan Test koneksi dengan klik pada tombol Test



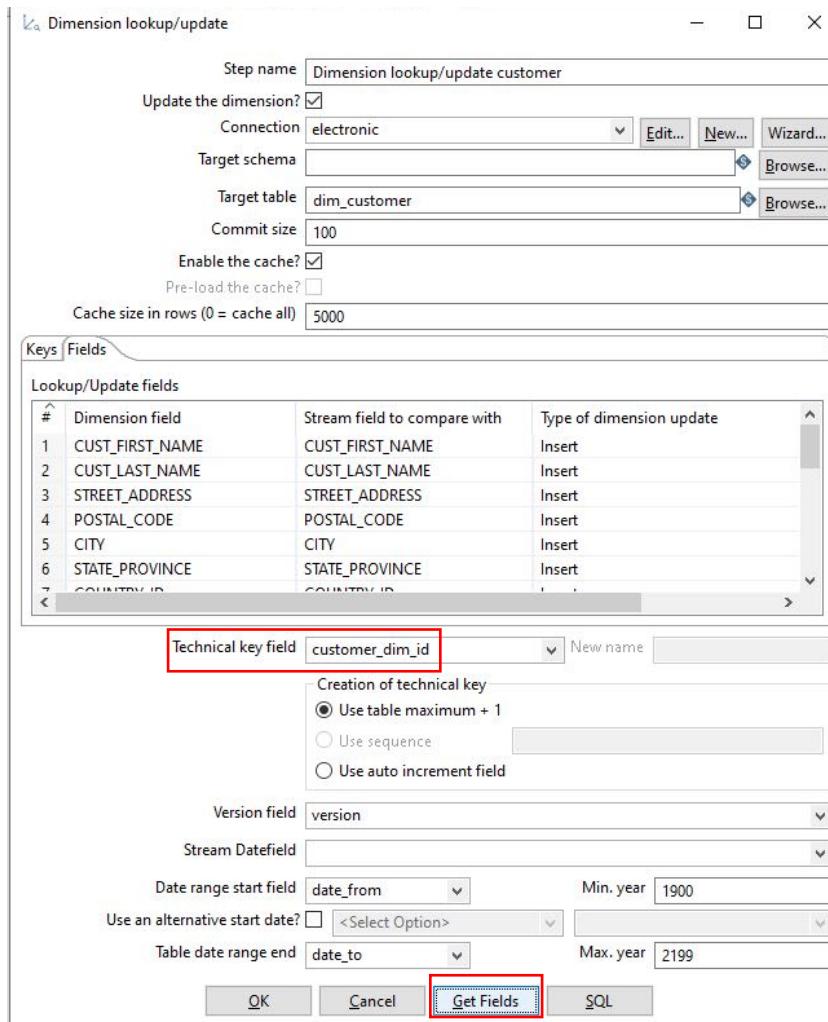
- Koneksi telah berhasil terhubung dan dapat melanjutkan ke tahapan selanjutnya
- Klik OK

This screenshot shows the 'Dimension lookup/update' configuration dialog box. At the top, the 'Step name' is set to 'Dimension lookup/update customer'. The 'Update the dimension?' checkbox is checked. The 'Connection' dropdown is set to 'electronic'. The 'Target schema' and 'Target table' (set to 'dim_customer') are also specified. Below these, 'Commit size' is set to '100', and 'Enable the cache?' and 'Pre-load the cache?' checkboxes are both unchecked. The 'Cache size in rows (0 = cache all)' is set to '5000'. The 'Fields' tab is selected under the 'Keys' tab. In the 'Key fields (to look up row in dimension)' section, there is a table:

| # | Dimension field | Field in stream |
|---|-----------------|-----------------|
| 1 | customer_id | customer_id |

Below this, the 'Technical key field' is set to 'customer_dim_id'. Under 'Creation of technical key', the radio button 'Use table maximum + 1' is selected. Other options include 'Use sequence' and 'Use auto increment field'. The 'Version field' is set to 'version'. The 'Stream Datefield' is empty. The 'Date range start field' is 'date_from' with a 'Min. year' of '1900'. The 'Table date range end' is 'date_to' with a 'Max. year' of '2199'. At the bottom are buttons for 'OK', 'Cancel', 'Get Fields', and 'SQL'.

- Checklist Update the dimension, untuk dilakukan Slowly Change Dimension (SCD) jika terdapat perubahan data akan tercatat.
- Untuk memetakannya kita perlu menentukan nama Dimension field dari Field in stream yang sudah ada pada CSV dan juga merupakan Primary Key aslinya.
- Pilih Tab Fields



- Pada Tab Fields, klik Get Fields.
- Berikan input nama pada Technical key field yang merupakan nama lain surrowgate key untuk memudahkan dalam pencatatan perubahan data.

Dimension lookup/update

Step name: Dimension lookup/update customer

Update the dimension?

Connection: electronic

Target schema:

Target table: dim_customer

Commit size: 100

Enable the cache?

Pre-load the cache?

Cache size in rows (0 = cache all): 5000

Keys Fields

Lookup/Update fields

| # | Dimension field | Stream field to compare with | Type of dimension update |
|---|-----------------|------------------------------|--------------------------|
| 1 | CUST_FIRST_NAME | CUST_FIRST_NAME | Insert |
| 2 | CUST_LAST_NAME | CUST_LAST_NAME | Insert |
| 3 | STREET_ADDRESS | STREET_ADDRESS | Insert |
| 4 | POSTAL_CODE | POSTAL_CODE | Insert |
| 5 | CITY | CITY | Insert |
| 6 | STATE_PROVINCE | STATE_PROVINCE | Insert |
| 7 | COUNTRY_ID | COUNTRY_ID | Insert |

Technical key field: customer_dim_id

Creation of technical key

Use table maximum + 1

Use sequence

Use auto increment field

Version field: version

Stream Datefield:

Date range start field: date_from Min. year: 1900

Use an alternative start date? <Select Option>

Table date range end: date_to Max. year: 2199

- Jika ada perubahan pada data fields aslinya, kita dapat memilih cara menanganinya pada kolom Type of dimension update.
Insert = SCD Tipe 2
- Dalam laporan ini saya akan membuat tahapan menggunakan SCD Tipe 2 pada Pentaho dengan 3 kolom tambahan (version, date_from, dan date_to).
- Pilih SQL

```

Simple SQL editor
SQL statements, separated by semicolon `;

CREATE TABLE dim_customer
(
    customer_dim_id BIGINT NOT NULL PRIMARY KEY
    ,version INT
    ,date_from DATETIME
    ,date_to DATETIME
    ,customer_id BIGINT
    ,CUST_FIRST_NAME VARCHAR(25)
    ,CUST_LAST_NAME VARCHAR(25)
    ,STREET_ADDRESS VARCHAR(50)
    ,POSTAL_CODE VARCHAR(15)
    ,CITY VARCHAR(17)
    ,STATE_PROVINCE VARCHAR(3)
    ,COUNTRY_ID VARCHAR(2)
    ,COUNTRY_NAME VARCHAR(24)
    ,REGION_ID BIGINT
    ,NLS_LANGUAGE VARCHAR(3)
    ,NLS_TERRITORY VARCHAR(11)
    ,CREDIT_LIMIT BIGINT
    ,CUST_EMAIL VARCHAR(38)
    ,PRIMARY_PHONE_NUMBER VARCHAR(16)
    ,PHONE_NUMBER_2 VARCHAR(16)
    ,ACCOUNT_MGR_ID BIGINT
    ,LOCATION_GTYPE BIGINT
    ,LOCATION_SRID BIGINT
    ,LOCATION_X DOUBLE
    ,LOCATION_Y DOUBLE
)
CREATE INDEX idx_dim_customer_lookup ON dim_customer(customer_id)
CREATE INDEX idx_dim_customer_tk ON dim_customer(customer_dim_id)
;

```

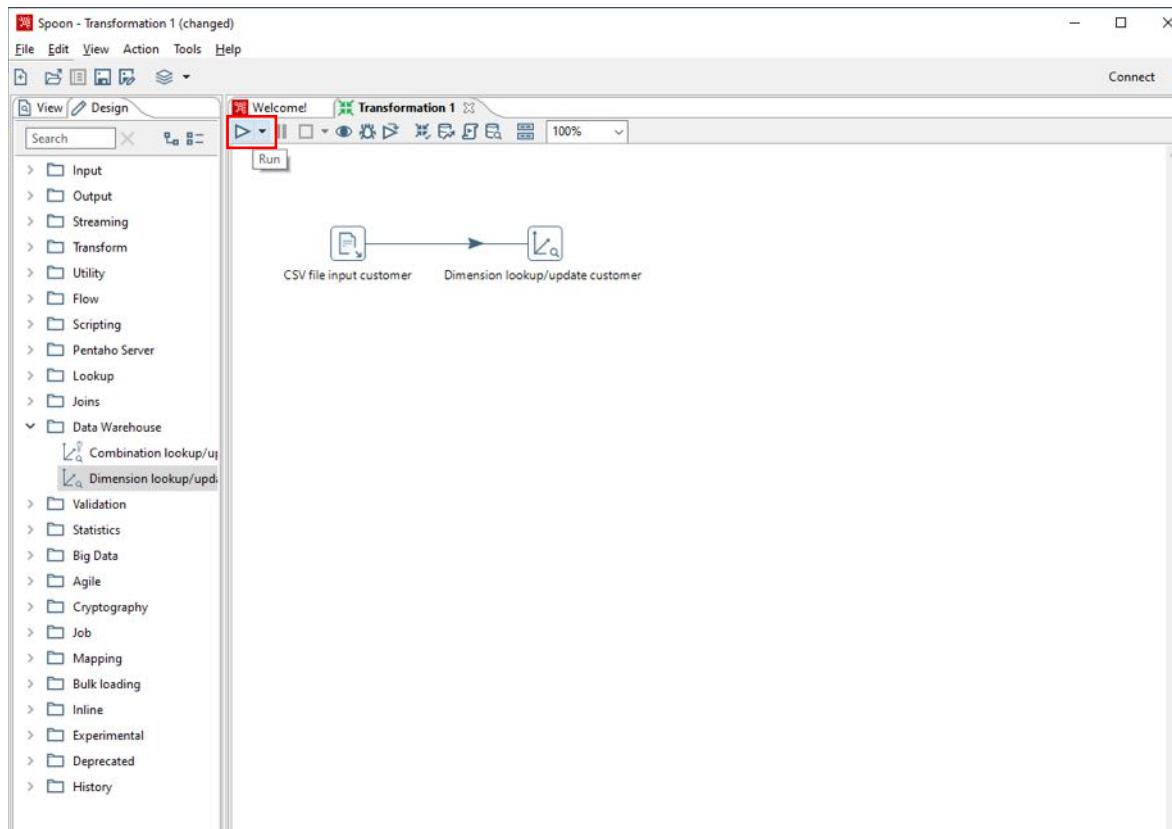
Line 1 column 0

Execute Clear cache Close

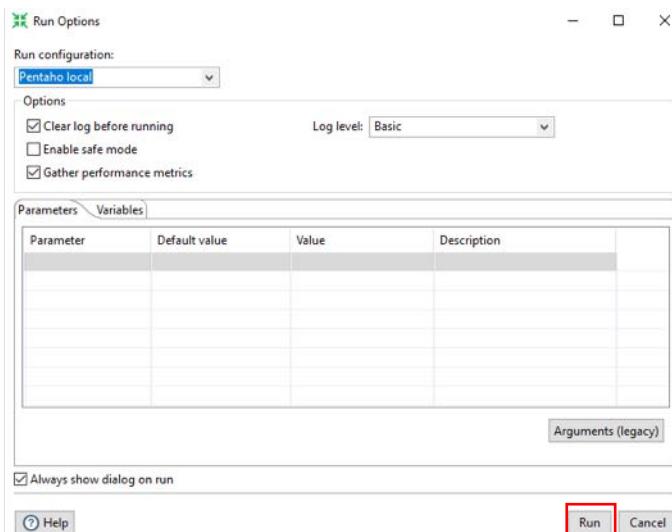
- customer_dim_id adalah tambahan yang merupakan surrowgate key
- version, date_from, dan date_to juga merupakan tambahan dari SCD Tipe 2 Pentaho
- sisanya merupakan data asli dari CSVnya
- Jika sesuai, pilih Execute

| # | Name | Type | Collation | Attributes | Null | Default | Comments | Extra | Action |
|----|----------------------|-------------|--------------------|------------|------|---------|----------|-------|--|
| 1 | customer_dim_id | bigint(20) | | | No | None | | | Change Drop More |
| 2 | version | int(11) | | | Yes | NULL | | | Change Drop More |
| 3 | date_from | datetime | | | Yes | NULL | | | Change Drop More |
| 4 | date_to | datetime | | | Yes | NULL | | | Change Drop More |
| 5 | customer_id | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 6 | CUST_FIRST_NAME | varchar(25) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 7 | CUST_LAST_NAME | varchar(25) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 8 | STREET_ADDRESS | varchar(50) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 9 | POSTAL_CODE | varchar(15) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 10 | CITY | varchar(17) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 11 | STATE_PROVINCE | varchar(3) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 12 | COUNTRY_ID | varchar(2) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 13 | COUNTRY_NAME | varchar(24) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 14 | REGION_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 15 | NLS_LANGUAGE | varchar(3) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 16 | NLS_TERRITORY | varchar(11) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 17 | CREDIT_LIMIT | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 18 | CUST_EMAIL | varchar(38) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 19 | PRIMARY_PHONE_NUMBER | varchar(16) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 20 | PHONE_NUMBER_2 | varchar(16) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 21 | ACCOUNT_MGR_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 22 | LOCATION_GTYPE | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 23 | LOCATION_SRID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 24 | LOCATION_X | double | | | Yes | NULL | | | Change Drop More |
| 25 | LOCATION_Y | double | | | Yes | NULL | | | Change Drop More |

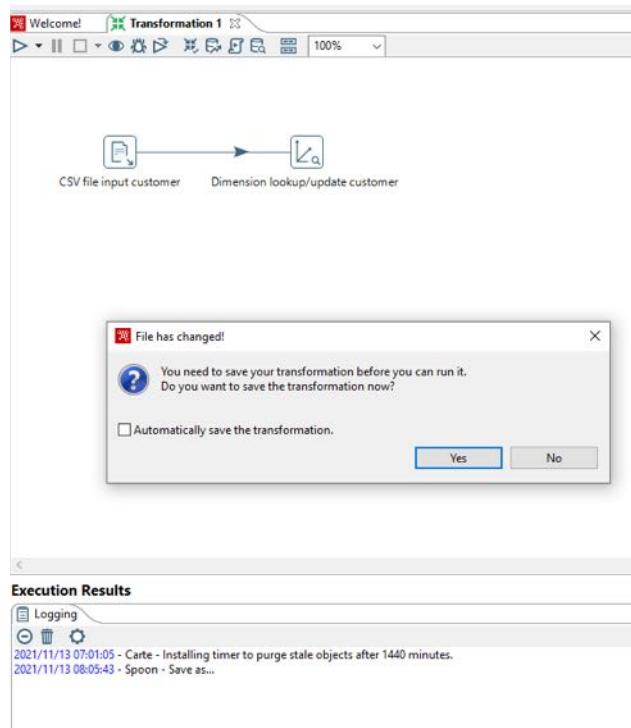
- Pentaho secara langsung menggenerate data ke MySQL dengan nama table dim_customer



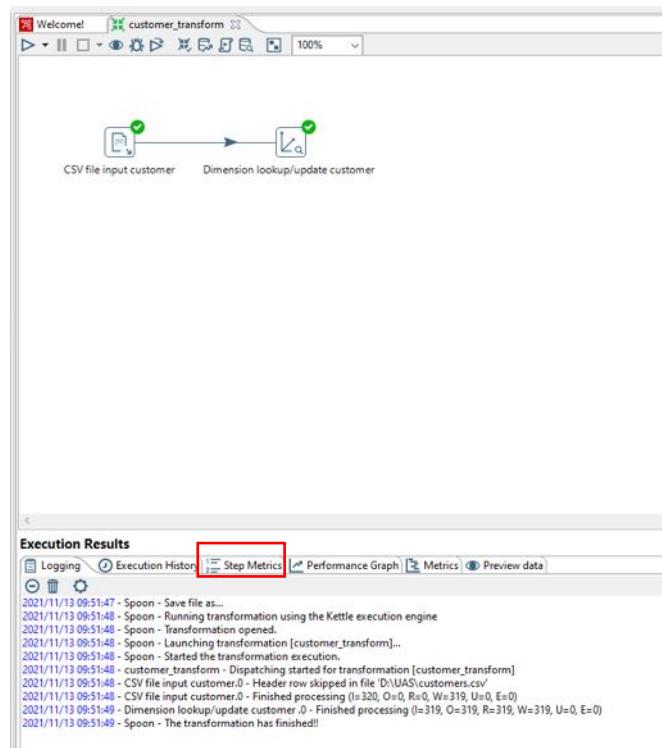
- Jalankan customer_transform dengan klik Run



- Klik Run



- Lakukan save dengan klik Yes
- Berikan nama customer_transform



- Data telah berhasil dieksekusi dan dapat diperiksa pada Tab Step Metrics

Execution Results

| # | Stepname | Copynr | Read | Written | Input | Output | Updated | Rejected | Errors | Active | Time |
|---|----------------------------------|--------|------|---------|-------|--------|---------|----------|--------|----------|------|
| 1 | CSV file input customer | 0 | 0 | 319 | 320 | 0 | 0 | 0 | 0 | Finished | 0.0s |
| 2 | Dimension lookup/update customer | 0 | 319 | 319 | 319 | 0 | 0 | 0 | 0 | Finished | 0.4s |

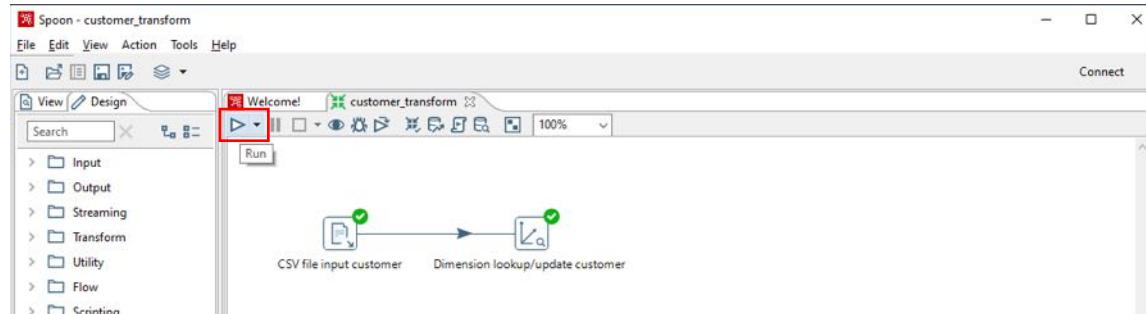
- Terdapat total data 320, yang dimana 1 data merupakan data template dari Pentaho sehingga data valid berjumlah 319 yang telah terinput ke database electronic pada table dim_customer dengan SCD Tipe 2

| version | date_from | date_to | customer_id | CUST_FIRST_NAME | CUST_LAST_NAME | STREET_ADDRESS | POSTAL_CODE | CITY | STATE_PROVINCE | COUNTRY_ID | COUNT |
|---------|---------------------|---------------------|-------------|-----------------|----------------|---------------------|-------------|------------|----------------|------------|---------------|
| 1 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 145 | PATRICIA | JOHNSON | 2120 Heights Dr | 54701 | Eau Claire | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 146 | LINDA | WILLIAMS | 14 Robbinsville Ct | 53217 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 147 | BARBARA | JONES | 6555 W Good Hope Rd | 53223 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 148 | ELIZABETH | BROWN | 1314 N Stoughton Rd | 53714 | Madison | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 149 | JENNIFER | DAVIS | 4715 Sprecher Rd | 53704 | Madison | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 150 | MARIA | MILLER | 6161 N 64Th St | 53218 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 151 | SUSAN | WILSON | 11016 W Lincoln Ave | 53227 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 152 | MARGARET | MOORE | 8600 W National Ave | 53227 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 153 | DOROTHY | TAYLOR | 615 N Sherman Ave | 53704 | Madison | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 154 | LISA | ANDERSON | 512 E Grand Ave | 53511 | Beloit | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 155 | NANCY | THOMAS | 600 N Broadway Fl 1 | 53202 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 156 | KAREN | JACKSON | 5235 N Ironwood Ln | 53217 | Milwaukee | WI | US | United States |
| 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 157 | BETTY | WHITE | 322 E Michigan St | 53202 | Milwaukee | WI | US | United States |

- Karena kita menggunakan SCD Tipe 2 maka akan terdapat kolom version, date_from, dan date_to yang merupakan template dari Pentaho
- Kali ini kita akan melakukan perubahan data dan melihat perbedaannya Ketika sesudah data berubah.

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|----|-------------|-----------------|----------------|---------------------|-------------|------------|----------------|------------|---------------|---|----|---------|-------------------------------|-----|------|------|---|---|
| 1 | CUSTOMER_ID | CUST_FIRST_NAME | CUST_LAST_NAME | STREET_ADDRESS | POSTAL_CODE | CITY | STATE_PROVINCE | COUNTRY_ID | COUNT | | | | | | | | | |
| 2 | 145 | PATRICIA | JOHNSON | 2120 Heights Dr | 54701 | Eau Claire | WI | US | United States | 2 | us | AMERICA | 500 Mammutt+1 745 123 4306 | 145 | 2001 | 8307 | # | |
| 3 | 146 | LINDA | WILLIAMS | 14 Robbinsville Ct | 53217 | Milwaukee | WI | US | United States | 2 | us | AMERICA | 500 Elia.Fawc+1 414 123 4307 | 145 | 2001 | 8307 | # | |
| 4 | 147 | BARBARA | JONES | 6555 W Good Hope Rd | 53223 | Milwaukee | WI | US | United States | 2 | us | AMERICA | 600 Ishwarya.+1 414 123 4308 | 145 | 2001 | 8307 | # | |
| 5 | 148 | ELIZABETH | BROWN | 1314 N Stoughton Rd | 53714 | Madison | WI | US | United States | 2 | us | AMERICA | 600 Gustav.Str+1 608 123 4309 | 145 | 2001 | 8307 | # | |
| 6 | 149 | JENNIFER | DAVIS | 4715 Sprecher Rd | 53704 | Madison | WI | US | United States | 2 | us | AMERICA | 600 Markus.Ri+1 608 123 4318 | 145 | 2001 | 8307 | # | |
| 7 | 150 | MARIA | MILLER | 6161 N 64Th St | 53227 | Milwaukee | WI | US | United States | 2 | us | AMERICA | 700 Goldie.Sli+1 414 123 4323 | 145 | 2001 | 8307 | # | |
| 8 | 151 | SUSAN | WILSON | 11016 W Lincoln Ave | 53202 | Milwaukee | WI | US | United States | 2 | us | AMERICA | 700 Divine.Ay+1 414 123 4324 | 145 | 2001 | 8307 | # | |
| 9 | 145 | ABDUL AZZAM | AJHARI | 2120 Heights Dr | 54701 | Eau Claire | WI | US | United States | 2 | us | AMERICA | 500 Mammutt+1 745 123 4306 | 145 | 2001 | 8307 | # | |
| 10 | 146 | LINDA | WILLIAMS | 14 Robbinsville Ct | 53217 | Milwaukee | WI | US | United States | 2 | us | AMERICA | 500 Elia.Fawc+1 414 123 4307 | 145 | 2001 | 8307 | # | |
| 11 | 147 | BARBARA | JONES | 6555 W Good Hope Rd | 53223 | Milwaukee | WI | US | United States | 2 | us | AMERICA | 600 Ishwarya.+1 414 123 4308 | 145 | 2001 | 8307 | # | |
| 12 | 148 | ELIZABETH | BROWN | 1314 N Stoughton Rd | 53714 | Madison | WI | US | United States | 2 | us | AMERICA | 600 Gustav.Str+1 608 123 4309 | 145 | 2001 | 8307 | # | |
| 13 | 149 | JENNIFER | DAVIS | 4715 Sprecher Rd | 53704 | Madison | WI | US | United States | 2 | us | AMERICA | 600 Markus.Ri+1 608 123 4318 | 145 | 2001 | 8307 | # | |

- Untuk mengetahui perubahannya, saya akan merubah fields CUST_FIRST_NAME dan CUST_LAST_NAME yang sebelumnya PATRICIA dan JOHNSON menjadi ABDUL AZZAM dan AJHARI pada CSV
- CUSTOMER_ID yang diubah adalah 145
- Save CSV



- Jalankan kembali dengan klik Run pada Pentaho

| Execution Results | | | | | | | | | | | |
|-------------------|----------------------------------|--------|------|---------|-------|--------|---------|----------|--------|----------|------|
| | Stepname | Copynr | Read | Written | Input | Output | Updated | Rejected | Errors | Active | Time |
| 1 | CSV file input customer | 0 | 0 | 319 | 320 | 0 | 0 | 0 | 0 | Finished | 0.0s |
| 2 | Dimension lookup/update customer | 0 | 319 | 319 | 319 | 1 | 0 | 0 | 0 | Finished | 0.2s |

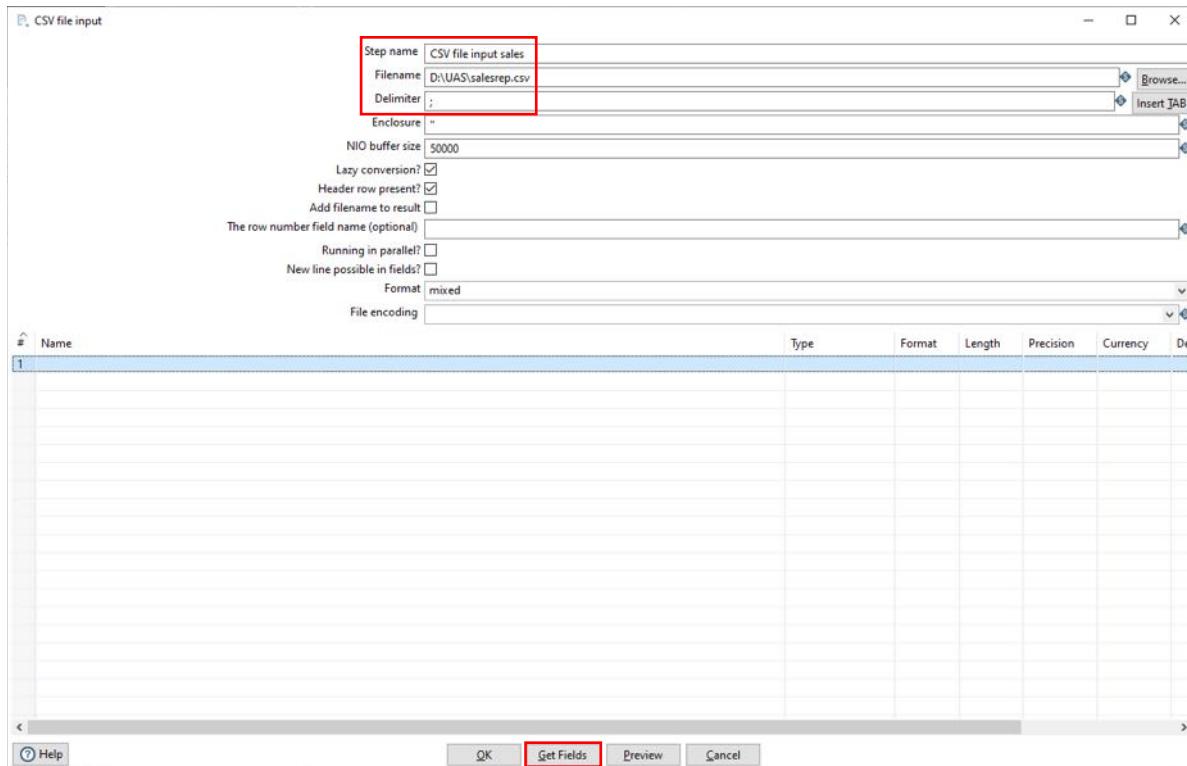
- Dapat dilihat terjadi perubahan pada Output data yang kita ubah adalah menjadi 1
- Kemudian kita periksa pada MySQL berdasarkan data CUSTOMER_ID = 145 di table dim_customer

```
SELECT * FROM `dim_customer` WHERE `customer_id` = 145;
```

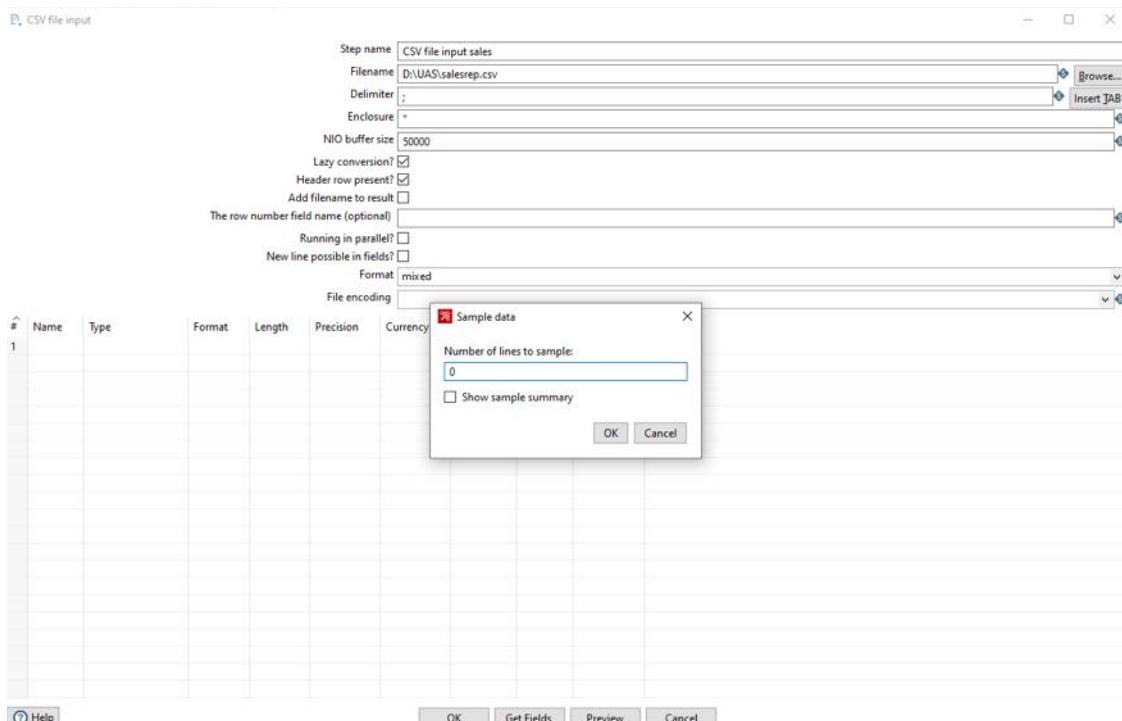
| customer_dim_id | version | date_from | date_to | customer_id | CUST_FIRST_NAME | CUST_LAST_NAME | STREET_ADDRESS | POSTAL_CODE | CITY | STATE_PROVINCE | COUNTRY_ID | COUNTRY_NAME |
|-----------------|---------|---------------------|---------------------|-------------|-----------------|----------------|-----------------|-------------|------------|----------------|------------|--------------------------|
| 1 | 1 | 1900-01-01 00:00:00 | 2021-11-13 10:02:42 | 145 | PATRICIA | JOHNSON | 2120 Heights Dr | 54701 | Eau Claire | WI | US | United States of America |
| 320 | 2 | 2021-11-13 10:02:42 | 2199-12-31 23:59:59 | 145 | ABDULAZZAM | AJHARI | 2120 Heights Dr | 54701 | Eau Claire | WI | US | United States of America |

- Terdapat perubahan version, date_from, dan date_to
- SCD Tipe 2 tidak menghapus data sebelumnya, namun menambah data baru berdasarkan data yang diubah

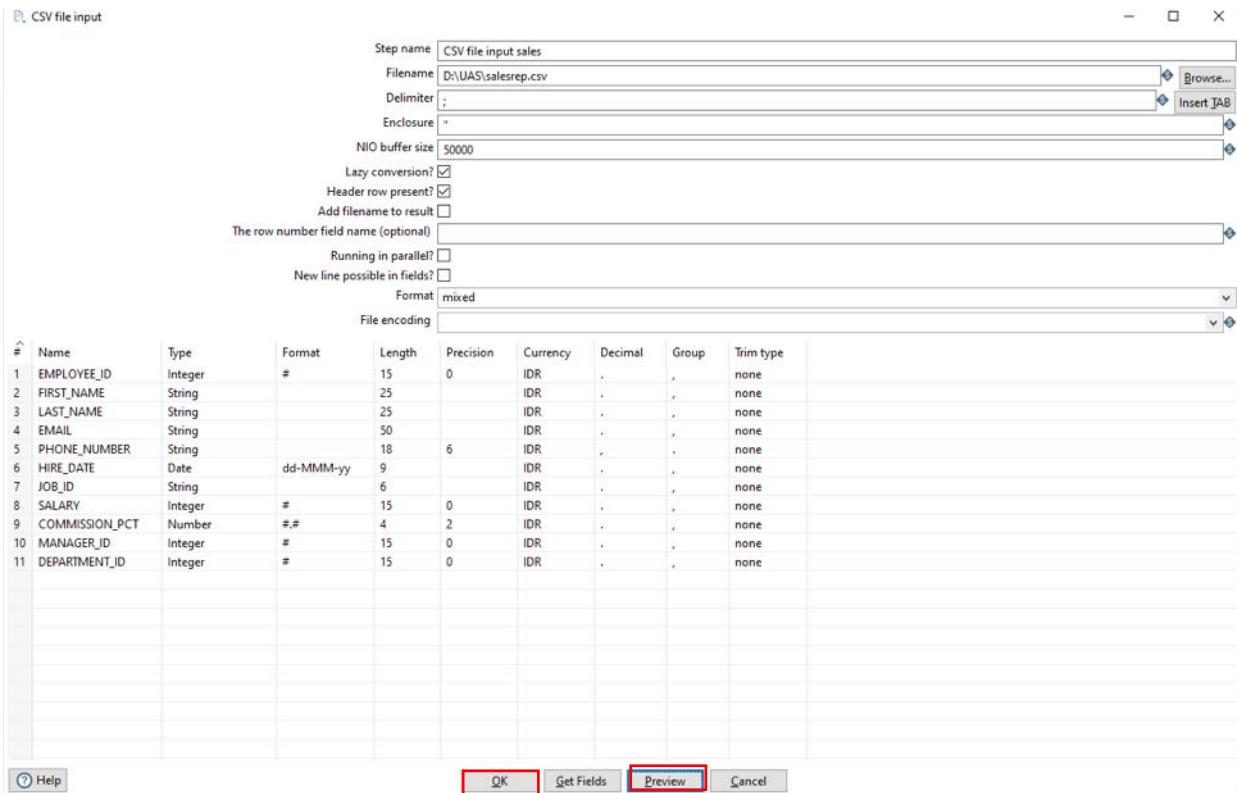
2. SALESREP_TRANSFORM (DIMENSION TABLE)



- Berikan nama pada Step name, dalam hal ini yang digunakan adalah salesrep
- Pilih filename
- Pastikan Delimeter telah sesuai
- Kemudian pilih GetFields

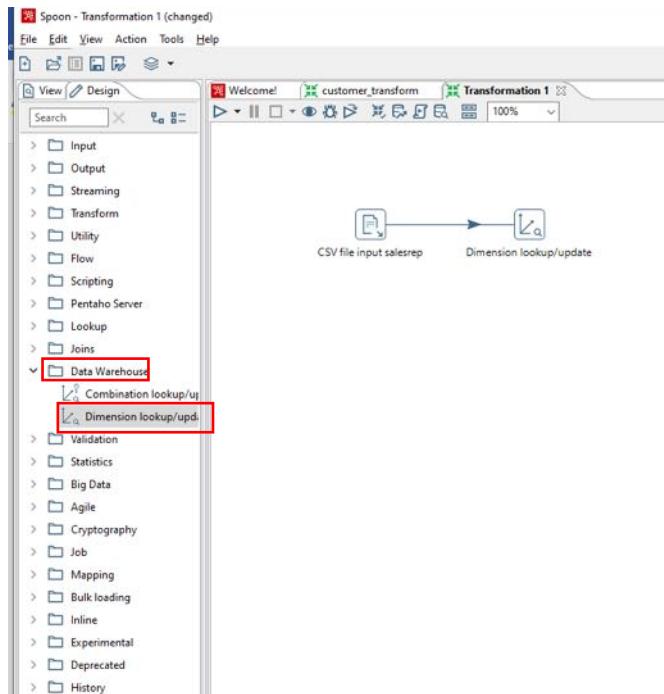


- Input 0 untuk menampilkan seluruh fields

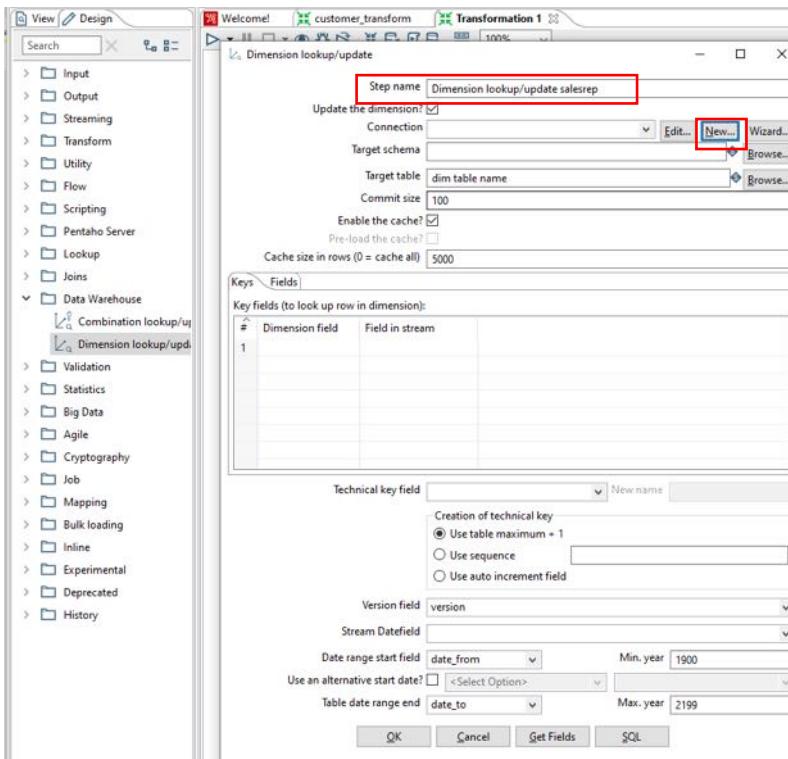


- Periksa data dengan klik tombol Preview
- Jika sudah sesuai boleh lanjutkan klik tombol OK

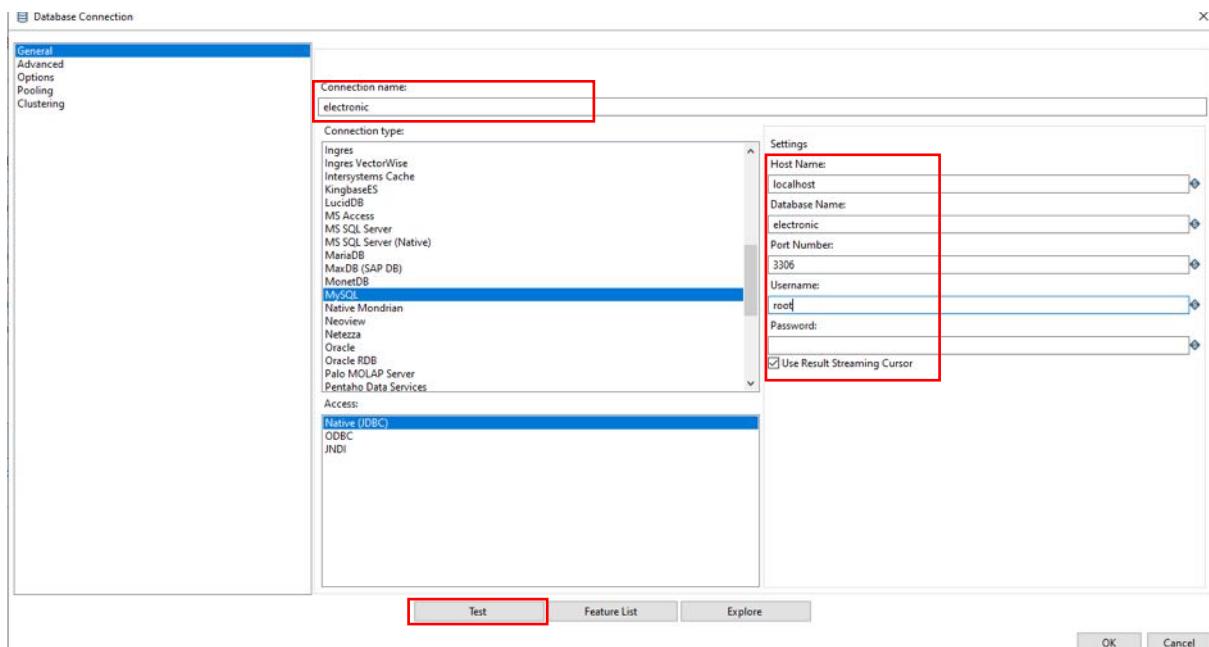
Setelah Langkah diatas sudah dilaksanakan, selanjutnya kita perlu membuat dimension salesrep (SALESREP_DIM) menggunakan Data Warehouse – Dimension lookup/update dan koneksi menggunakan konektor.



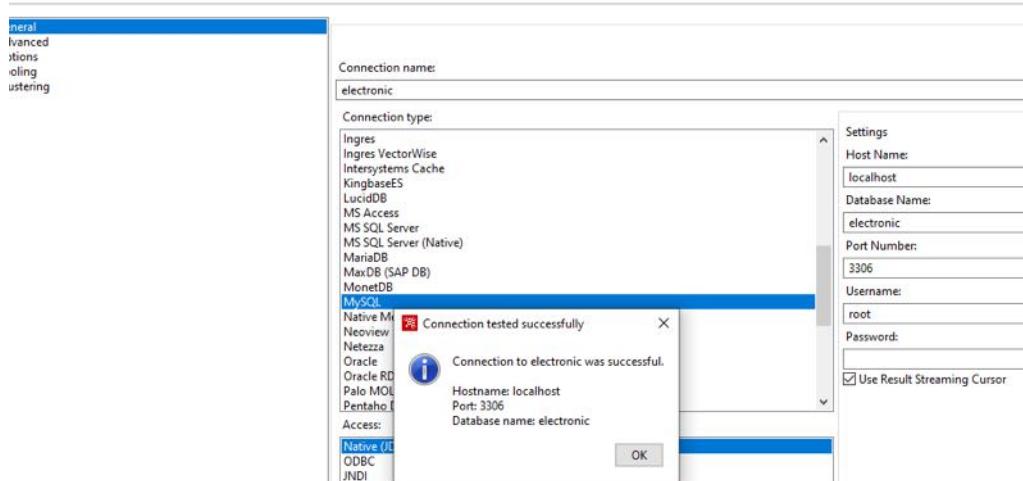
- Double klik pada Dimension lookup/update



- Berikan nama pada Step name
- Buat koneksi database, dalam hal ini saya akan gunakan MySQL dengan klik tombol New



- Berikan nama koneksi, untuk memudahkan disamakan saja dengan nama database pada MySQL yang sudah dibuat
- Sesuaikan jaringan host, port, username dan juga passwordnya
- Lakukan Test koneksi dengan klik pada tombol Test



- Koneksi telah berhasil terhubung dan dapat melanjutkan ke tahapan selanjutnya
- Klik OK

Dimension lookup/update

Step name: Dimension lookup/update salesrep

Update the dimension?

Connection: electronic

Target schema:

Target table: dim_salesrep

Commit size: 100

Enable the cache?

Pre-load the cache?

Cache size in rows (0 = cache all): 5000

Keys Fields

Key fields (to look up row in dimension):

| # | Dimension field | Field in stream |
|---|-----------------|-----------------|
| 1 | salesrep_id | EMPLOYEE_ID |

Technical key field: salesrep_dim_id

Creation of technical key:

- Use table maximum + 1
- Use sequence
- Use auto increment field

Version field: version

Stream Datefield:

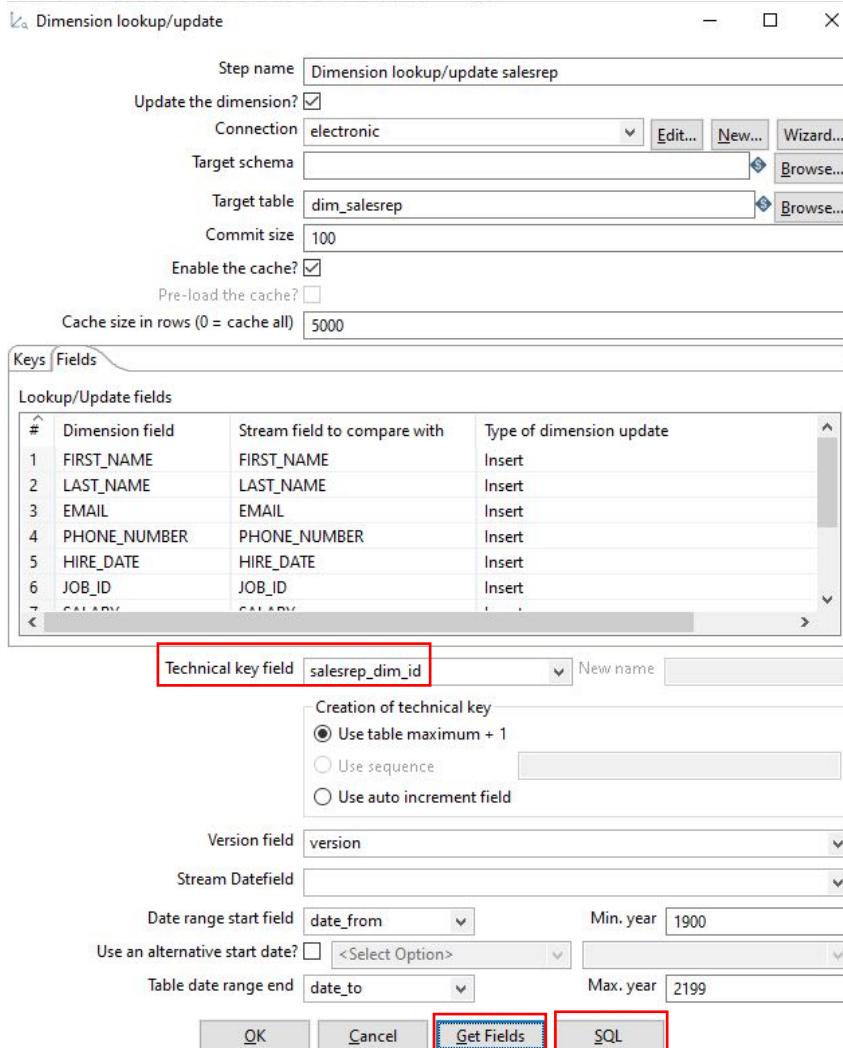
Date range start field: date_from Min. year: 1900

Use an alternative start date? <Select Option>

Table date range end: date_to Max. year: 2199

- Checklist Update the dimension, untuk dilakukan Slowly Change Dimension (SCD) jika terdapat perubahan data akan tercatat.
- Untuk memetakannya kita perlu menentukan nama Dimension field dari Field in stream yang sudah ada pada CSV dan juga merupakan Primary Key aslinya.

- Pilih Tab Fields



- Pada Tab Fields, klik Get Fields.
- Berikan input nama pada Technical key field yang merupakan nama lain surrowgate key untuk memudahkan dalam pencatatan perubahan data.
- Pilih SQL

```

Simple SQL editor
SQL statements, separated by semicolon ;
CREATE TABLE dim_salesrep
(
    salesrep_dim_id BIGINT NOT NULL PRIMARY KEY
    ,version INT
    ,date_from DATETIME
    ,date_to DATETIME
    ,salesrep_id BIGINT
    ,FIRST_NAME VARCHAR(25)
    ,LAST_NAME VARCHAR(25)
    ,EMAIL VARCHAR(50)
    ,PHONE_NUMBER VARCHAR(18)
    ,HIRE_DATE DATETIME
    ,JOB_ID VARCHAR(6)
    ,SALARY BIGINT
    ,COMMISSION_PCT DOUBLE
    ,MANAGER_ID BIGINT
    ,DEPARTMENT_ID BIGINT
)
CREATE INDEX idx_dim_salesrep_lookup ON dim_salesrep(salesrep_id)
CREATE INDEX idx_dim_salesrep_pk ON dim_salesrep(dim_salesrep_id)
;

```

Line 1 column 0

Execute Clear cache Close

- salesrep_dim_id adalah tambahan yang merupakan surrounge key
- version, date_from, dan date_to juga merupakan tambahan dari SCD Tipe 2 Pentaho
- sisanya merupakan data asli dari CSVnya
- Jika sesuai, pilih Execute

phpMyAdmin

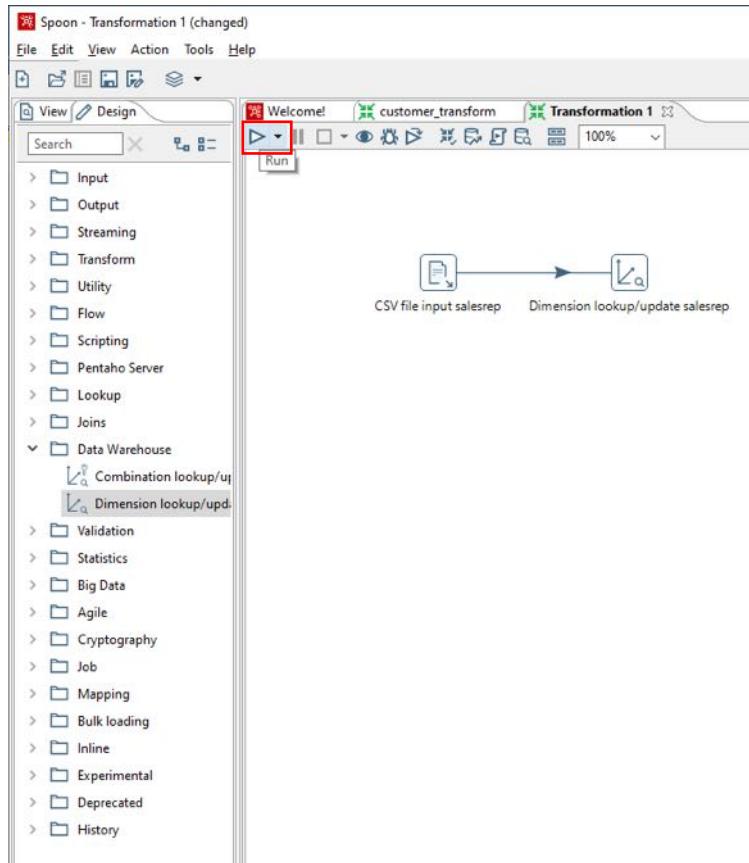
Structure of dim_salesrep table:

| # | Name | Type | Collation | Attributes | Null | Default | Comments | Extra | Action |
|----|-----------------|-------------|--------------------|------------|------|---------|----------|-------|--|
| 1 | salesrep_dim_id | bigint(20) | | | No | None | | | Change Drop More |
| 2 | version | int(11) | | | Yes | NULL | | | Change Drop More |
| 3 | date_from | datetime | | | Yes | NULL | | | Change Drop More |
| 4 | date_to | datetime | | | Yes | NULL | | | Change Drop More |
| 5 | salesrep_id | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 6 | FIRST_NAME | varchar(25) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 7 | LAST_NAME | varchar(25) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 8 | EMAIL | varchar(50) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 9 | PHONE_NUMBER | varchar(18) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 10 | HIRE_DATE | datetime | | | Yes | NULL | | | Change Drop More |
| 11 | JOB_ID | varchar(6) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 12 | SALARY | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 13 | COMMISSION_PCT | double | | | Yes | NULL | | | Change Drop More |
| 14 | MANAGER_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 15 | DEPARTMENT_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |

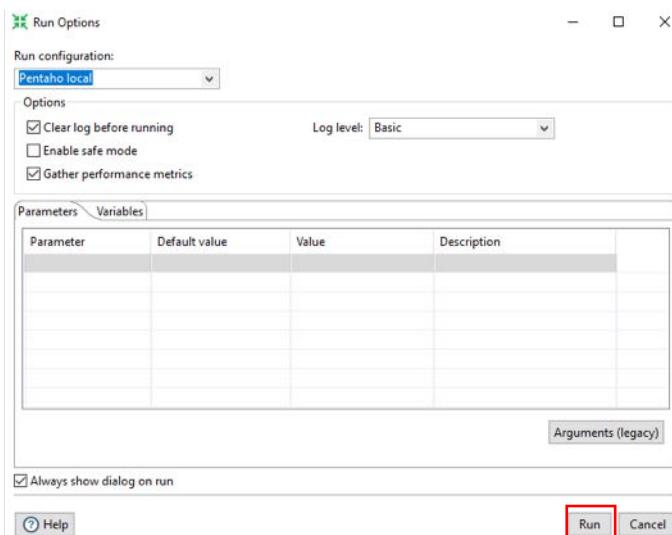
Indexes:

| Action | Keyname | Type | Unique | Packed | Column | Cardinality | Collation | Null | Comment |
|--|-------------------------|-------|--------|--------|-----------------|-------------|-----------|------|---------|
| Edit Rename Drop | PRIMARY | BTREE | Yes | No | salesrep_dim_id | 0 | A | No | |
| Edit Rename Drop | idx_dim_salesrep_lookup | BTREE | No | No | salesrep_id | 0 | A | Yes | |
| Edit Rename Drop | idx_dim_salesrep_pk | BTREE | No | No | salesrep_dim_id | 0 | A | No | |

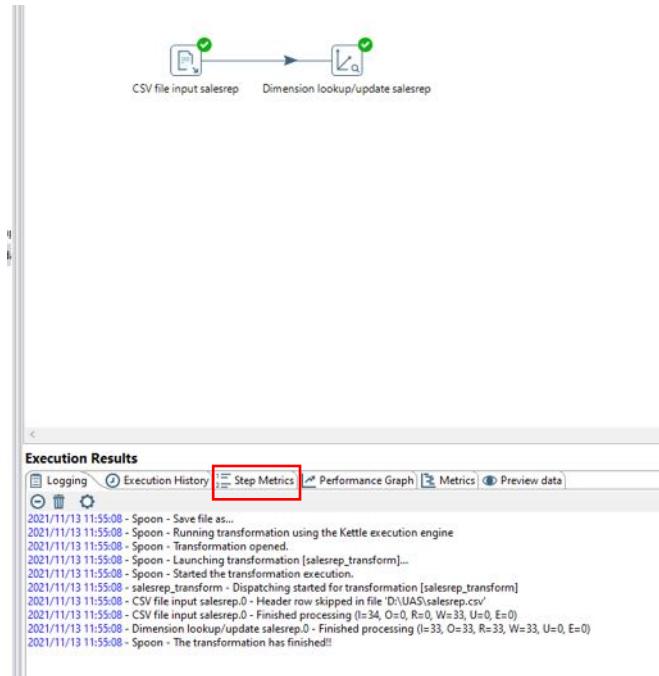
- Pentaho secara langsung menggenerate data ke MySQL dengan nama table dim_salesrep



- Jalankan salesrep_transform dengan klik Run



- Klik Run



- Data telah berhasil dieksekusi dan dapat diperiksa pada Tab Step Metrics

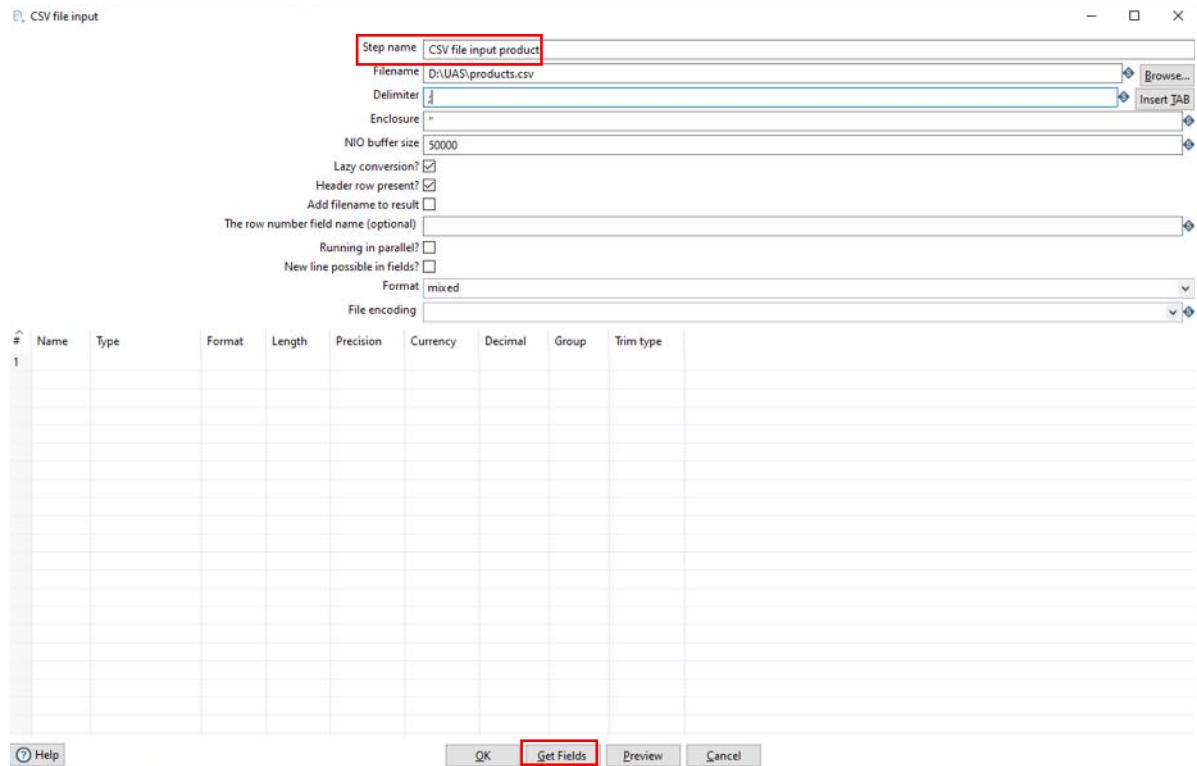
| # | Stepname | Copynr | Read | Written | Input | Output | Updated | Rejected | Errors | Active | Time |
|---|----------------------------------|--------|------|---------|-------|--------|---------|----------|--------|----------|------|
| 1 | CSV file input salesrep | 0 | 0 | 33 | 34 | 0 | 0 | 0 | 0 | Finished | 0.0s |
| 2 | Dimension lookup/update salesrep | 0 | 33 | 33 | 33 | 33 | 0 | 0 | 0 | Finished | 0.1s |

- Terdapat total data 34, yang dimana 1 data merupakan data template dari Pentaho sehingga data valid berjumlah 33 yang telah terinput ke database electronic pada table dim_salesrep dengan SCD Tipe 2

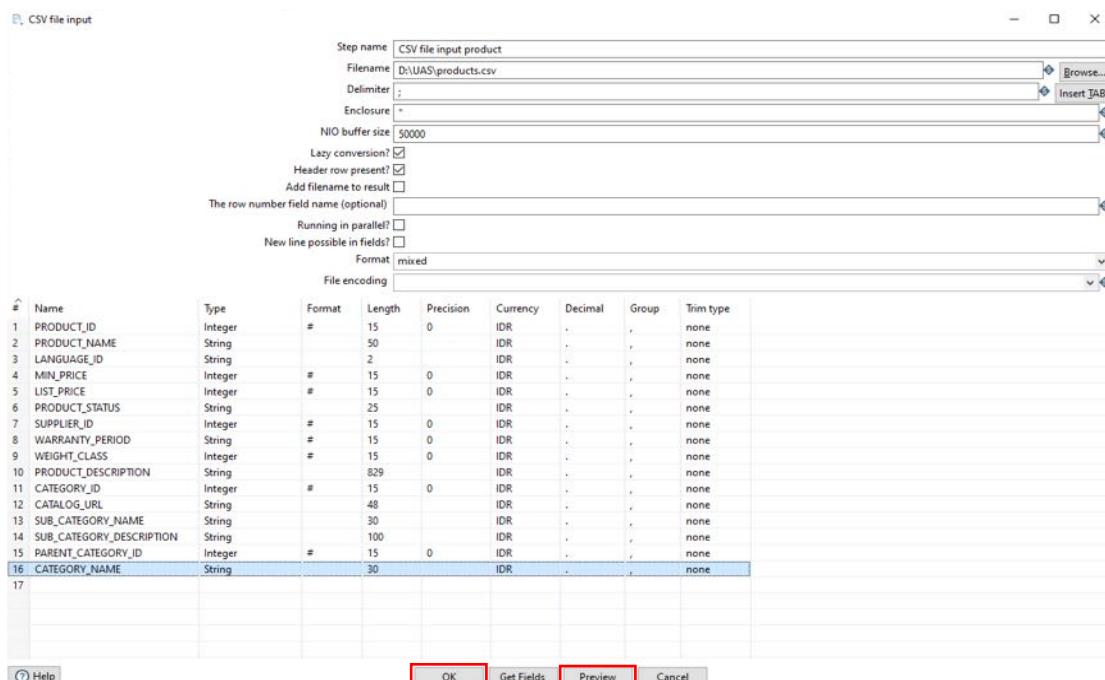
```
SELECT * FROM `dim_salesrep`
```

| Showing rows 0 - 24 (34 total, Query took 0.0000 seconds.) | | | | | | | | | | | | | | | | | | |
|--|-----------------|---------|-----------|---------|-------------|---------------------|---------------------|-------|--------------|---------------------|----------|--------------------|---------------------|------------|---------|------|-----|----|
| SELECT * FROM `dim_salesrep` | | | | | | | | | | | | | | | | | | |
| Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh] | | | | | | | | | | | | | | | | | | |
| Options | salesrep.dim_id | version | date_from | date_to | salesrep_id | FIRST_NAME | LAST_NAME | EMAIL | PHONE_NUMBER | HIRE_DATE | JOB_ID | SALARY | COMMISSION_PCT | MANAGER_ID | DEPT_ID | | | |
| <input type="checkbox"/> | Edit | Copy | Delete | 0 | 1 | NULL | NULL | NULL | NULL | 2005-03-10 00:00:00 | SA_MAN | 12000 | 0.3 | 100 | 80 | | | |
| <input type="checkbox"/> | Edit | Copy | Delete | 1 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 147 | PENELOPE | GUINNESS | AERAZU | 011.44.1344.429278 | 2005-10-15 00:00:00 | SA_MAN | 11000 | 0.3 | 100 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 2 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 148 | NICK | WAHLBERG | GCAMBRAU | 011.44.1344.619268 | 2007-01-29 00:00:00 | SA_MAN | 10500 | 0.2 | 100 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 3 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 149 | ED | CHASE | EZLOTKEY | 011.44.1344.429018 | 2006-03-10 00:00:00 | SA_MAN | 10500 | 0.2 | 100 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 4 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 150 | JENNIFER | DAVIS | PTUCKER | 011.44.1344.129268 | 2005-01-30 00:00:00 | SA REP | 10000 | 0.3 | 145 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 5 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 151 | JOHNNY | LOLOBRIGIDA | DOERNSTE | 011.44.1344.345268 | 2005-03-24 00:00:00 | SA REP | 9500 | 0.25 | 145 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 6 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 152 | BETTE | NICHOLSON | PHALL | 011.44.1344.478968 | 2005-08-20 00:00:00 | SA REP | 9000 | 0.25 | 145 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 7 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 153 | GRACE | MOSTEL | COLESEN | 011.44.1344.498718 | 2006-03-30 00:00:00 | SA REP | 8000 | 0.2 | 145 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 8 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 154 | MATTHEW | JOHANSSON | NCAMBRAU | 011.44.1344.987668 | 2006-12-09 00:00:00 | SA REP | 7500 | 0.2 | 145 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 9 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 155 | JOE | SWANK | OTUVAULI | 011.44.1344.486500 | 2007-11-23 00:00:00 | SA REP | 7000 | 0.15 | 145 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 10 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 156 | CHRISTIAN | GABLE | JKING | 011.44.1345.428268 | 2004-01-30 00:00:00 | SA REP | 10000 | 0.35 | 146 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 11 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 157 | ZERO | CAGE | PSULLY | 011.44.1345.929268 | 2004-03-04 00:00:00 | SA REP | 9500 | 0.35 | 146 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 12 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 158 | KARL | BERRY | AMCEWEV | 011.44.1345.829268 | 2004-08-01 00:00:00 | SA REP | 9000 | 0.35 | 146 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 13 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 159 | UMA | WOOD | LSMITH | 011.44.1345.729268 | 2005-03-10 00:00:00 | SA REP | 8000 | 0.3 | 146 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 14 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 160 | VIVIEN | BERGEN | LDORAN | 011.44.1345.629268 | 2005-12-15 00:00:00 | SA REP | 7500 | 0.3 | 146 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 15 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 161 | CUBA | OLIVIER | SSEWALL | 011.44.1345.529268 | 2006-11-03 00:00:00 | SA REP | 7000 | 0.25 | 146 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 16 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 162 | FRED | COSTNER | CVISHNEY | 011.44.1346.129268 | 2005-11-11 00:00:00 | SA REP | 10500 | 0.25 | 147 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 17 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 163 | HELEN | VOIGHT | DGREENE | 011.44.1346.229268 | 2007-03-19 00:00:00 | SA REP | 9500 | 0.15 | 147 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 18 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 164 | DAN | TORN | MMARVIN | 011.44.1346.329268 | 2008-01-24 00:00:00 | SA REP | 7200 | 0.1 | 147 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 19 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 165 | BOB | FAWCETT | DLEE | 011.44.1346.529268 | 2008-02-23 00:00:00 | SA REP | 6800 | 0.1 | 147 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 20 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 166 | LUCILLE | TRACY | SANDE | 011.44.1346.629268 | 2008-03-24 00:00:00 | SA REP | 6400 | 0.1 | 147 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 21 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 167 | KIRSTEN | PALTROW | ABANDA | 011.44.1346.729268 | 2008-04-21 00:00:00 | SA REP | 6200 | 0.1 | 147 | 80 |
| <input type="checkbox"/> | Edit | Copy | Delete | 22 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 168 | ELVIS | MARX | LOZER | 011.44.1343.929268 | 2005-03-11 00:00:00 | SA REP | 11500 | 0.25 | 148 | 80 |

3. PRODUCT_TRANSFORM (DIMENSION TABLE)

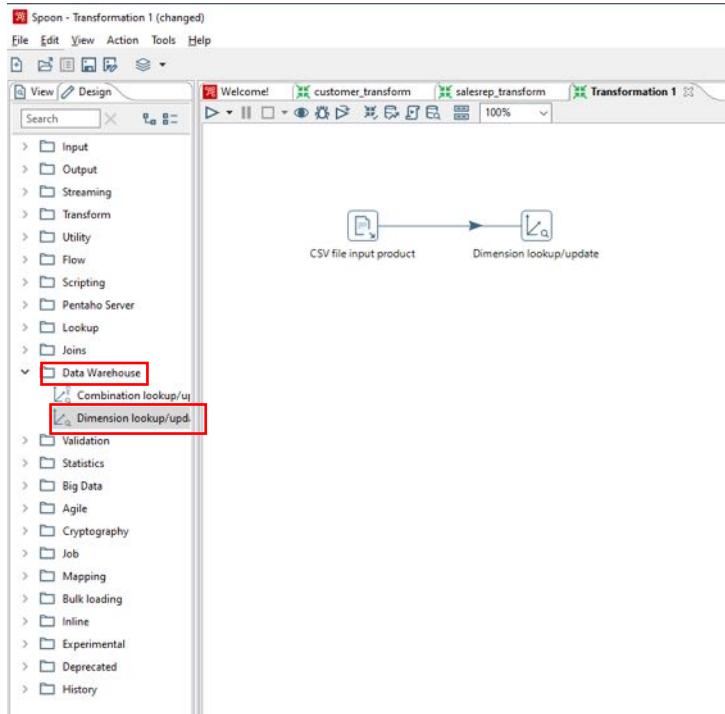


- Berikan nama pada Step name, dalam hal ini yang digunakan adalah product
- Pilih filename
- Pastikan Delimeter telah sesuai
- Kemudian pilih GetFields

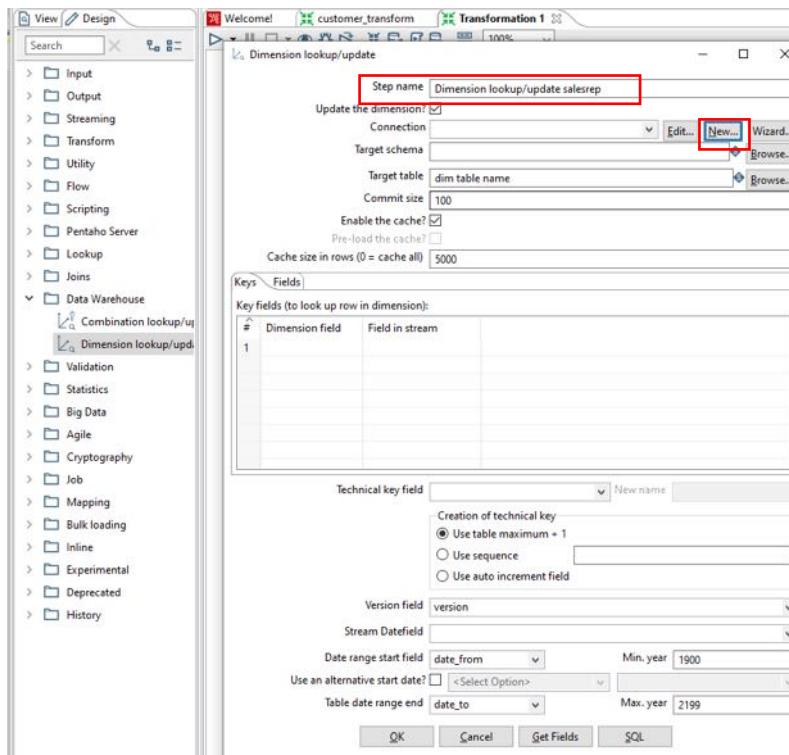


- Periksa data dengan klik tombol Preview
- Jika sudah sesuai boleh lanjutkan klik tombol OK

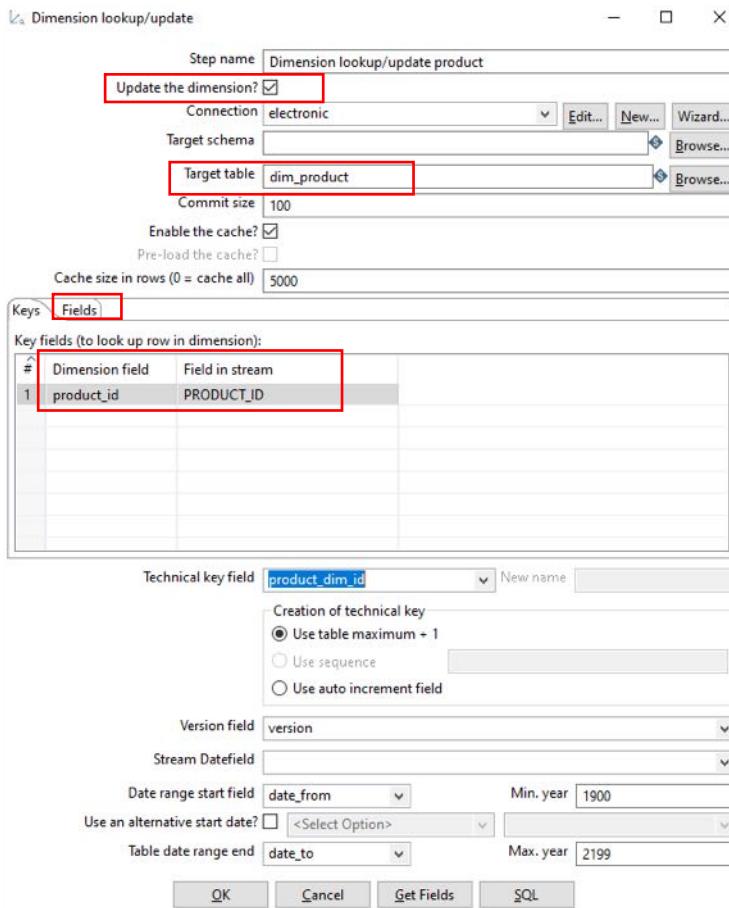
Setelah Langkah diatas sudah dilaksanakan, selanjutnya kita perlu membuat dimension product (PRODUCT_DIM) menggunakan Data Warehouse – Dimension lookup/update dan koneksi menggunakan konektor.



- Double klik pada Dimension lookup/update



- Berikan nama pada Step name
- Pilih koneksi yang sudah dibuat dan terhubung ke MySQL



- Checklist Update the dimension, untuk dilakukan Slowly Change Dimension (SCD) jika terdapat perubahan data akan tercatat.
- Untuk memetakannya kita perlu menentukan nama Dimension field dari Field in stream yang sudah ada pada CSV dan juga merupakan Primary Key aslinya.
- Pilih Tab Fields

Dimension lookup/update

Step name: Dimension lookup/update product

Update the dimension?

Connection: electronic

Target schema:

Target table: dim_product

Commit size: 100

Enable the cache?

Pre-load the cache?

Cache size in rows (0 = cache all): 5000

Fields

Lookup/Update fields

| # | Dimension field | Stream field to compare with | Type of dimension update |
|---|-----------------|------------------------------|--------------------------|
| 1 | PRODUCT_NAME | PRODUCT_NAME | Insert |
| 2 | LANGUAGE_ID | LANGUAGE_ID | Insert |
| 3 | MIN_PRICE | MIN_PRICE | Insert |
| 4 | LIST_PRICE | LIST_PRICE | Insert |
| 5 | PRODUCT_STATUS | PRODUCT_STATUS | Insert |
| 6 | SUPPLIER_ID | SUPPLIER_ID | Insert |
| 7 | WARRANTY_PERIOD | WARRANTY_PERIOD | Insert |

Technical key field: product_dim_id

Creation of technical key:

Use table maximum + 1

Use sequence

Use auto increment field

Version field: version

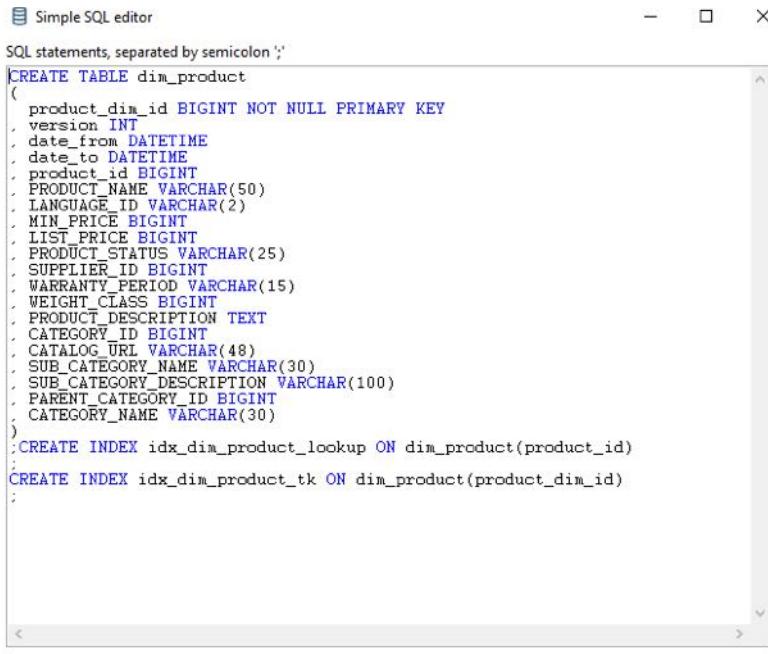
Stream Datefield:

Date range start field: date_from Min. year: 1900

Use an alternative start date? <Select Option>

Table date range end: date_to Max. year: 2199

- Pada Tab Fields, klik Get Fields.
- Berikan input nama pada Technical key field yang merupakan nama lain surrowgate key untuk memudahkan dalam pencatatan perubahan data.
- Pilih SQL



```

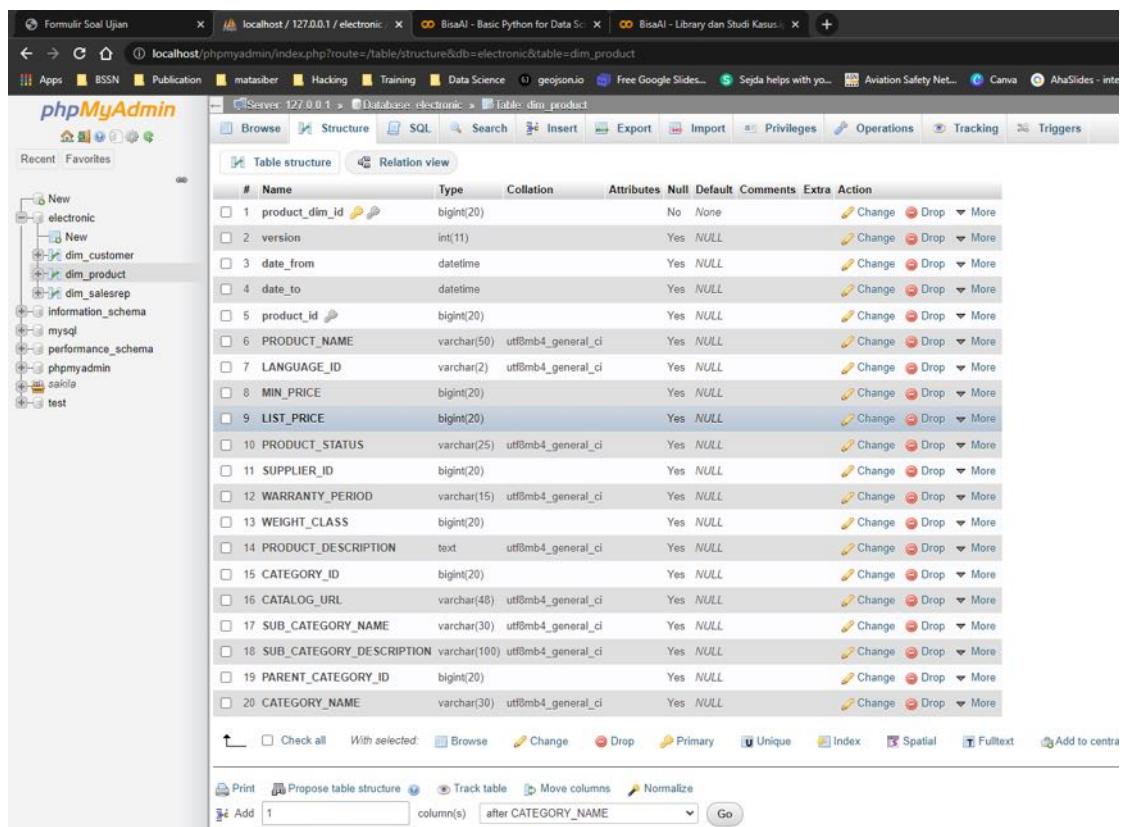
Simple SQL editor

SQL statements, separated by semicolon `;`:
CREATE TABLE dim_product
(
    product_dim_id BIGINT NOT NULL PRIMARY KEY
    , version INT
    , date_from DATETIME
    , date_to DATETIME
    , product_id BIGINT
    , PRODUCT_NAME VARCHAR(50)
    , LANGUAGE_ID VARCHAR(2)
    , MIN_PRICE BIGINT
    , LIST_PRICE BIGINT
    , PRODUCT_STATUS VARCHAR(25)
    , SUPPLIER_ID BIGINT
    , WARRANTY_PERIOD VARCHAR(15)
    , WEIGHT_CLASS BIGINT
    , PRODUCT_DESCRIPTION TEXT
    , CATEGORY_ID BIGINT
    , CATALOG_URL VARCHAR(48)
    , SUB_CATEGORY_NAME VARCHAR(30)
    , SUB_CATEGORY_DESCRIPTION VARCHAR(100)
    , PARENT_CATEGORY_ID BIGINT
    , CATEGORY_NAME VARCHAR(30)
)
CREATE INDEX idx_dim_product_lookup ON dim_product(product_id)
CREATE INDEX idx_dim_product_tk ON dim_product(product_dim_id)
;
```

Line 1 column 0

[Execute](#) [Clear cache](#) [Close](#)

- product_dim_id adalah tambahan yang merupakan surrouge key
- version, date_from, dan date_to juga merupakan tambahan dari SCD Tipe 2 Pentaho
- sisanya merupakan data asli dari CSVnya
- Jika sesuai, pilih Execute

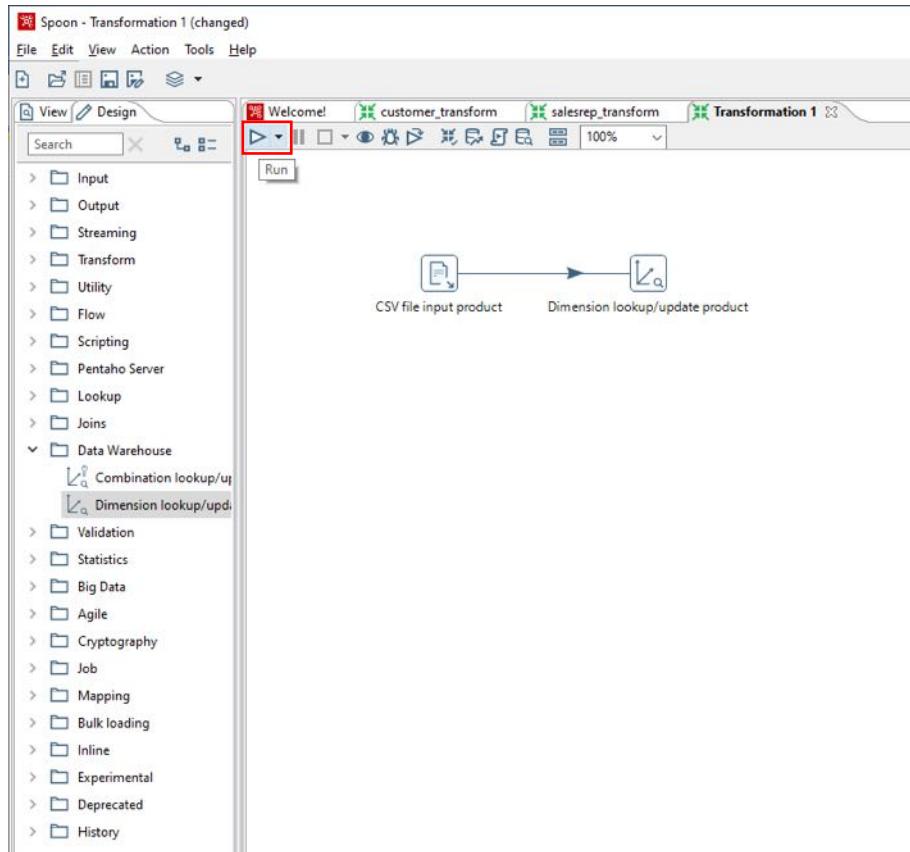


The screenshot shows the phpMyAdmin interface for the 'electronic' database. The 'dim_product' table is selected. The table structure is displayed with 20 columns:

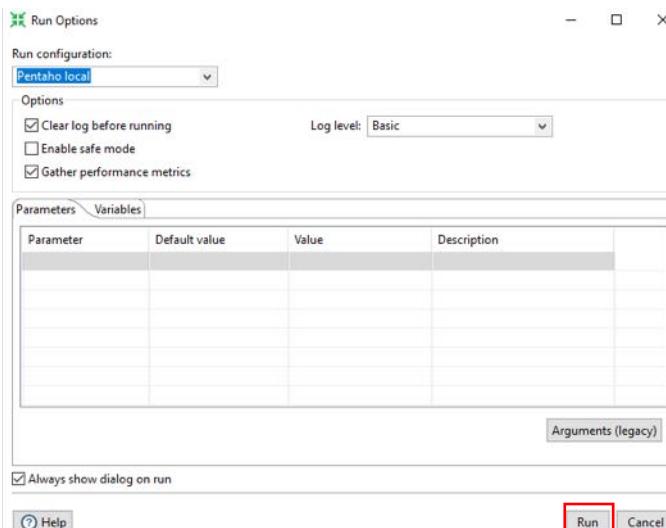
| # | Name | Type | Collation | Attributes | Null | Default | Comments | Extra | Action |
|----|--------------------------|--------------|--------------------|------------|------|---------|----------|-------|--|
| 1 | product_dim_id | bigint(20) | | | No | None | | | Change Drop More |
| 2 | version | int(11) | | | Yes | NULL | | | Change Drop More |
| 3 | date_from | datetime | | | Yes | NULL | | | Change Drop More |
| 4 | date_to | datetime | | | Yes | NULL | | | Change Drop More |
| 5 | product_id | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 6 | PRODUCT_NAME | varchar(50) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 7 | LANGUAGE_ID | varchar(2) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 8 | MIN_PRICE | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 9 | LIST_PRICE | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 10 | PRODUCT_STATUS | varchar(25) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 11 | SUPPLIER_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 12 | WARRANTY_PERIOD | varchar(15) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 13 | WEIGHT_CLASS | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 14 | PRODUCT_DESCRIPTION | text | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 15 | CATEGORY_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 16 | CATALOG_URL | varchar(48) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 17 | SUB_CATEGORY_NAME | varchar(30) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 18 | SUB_CATEGORY_DESCRIPTION | varchar(100) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
| 19 | PARENT_CATEGORY_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 20 | CATEGORY_NAME | varchar(30) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |

At the bottom, there are buttons for Print, Propose table structure, Track table, Move columns, Normalize, Add, and Go.

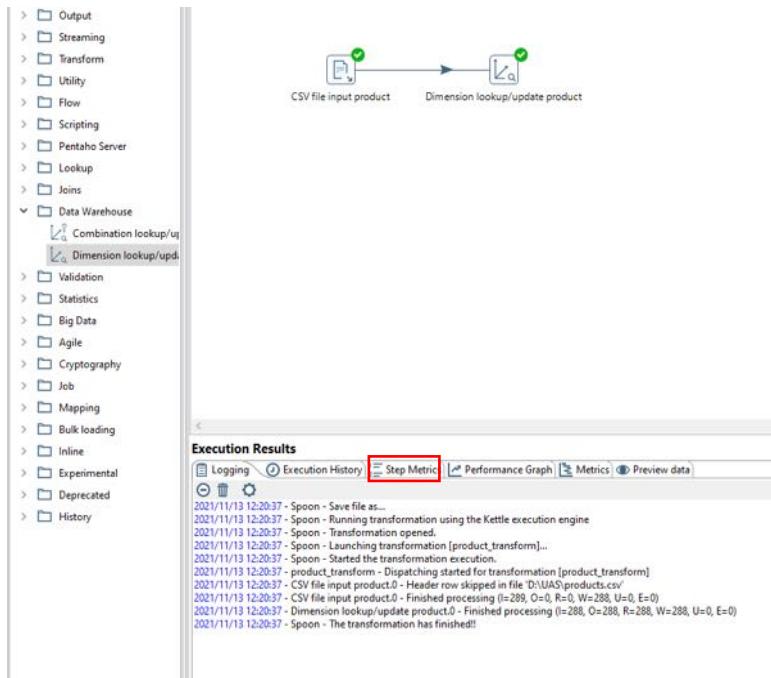
- Pentaho secara langsung menggenerate data ke MySQL dengan nama table dim_product



- Jalankan product_transform dengan klik Run



- Klik Run



- Data telah berhasil dieksekusi dan dapat diperiksa pada Tab Step Metrics

| Execution Results | | | | | | | | | | | | |
|-------------------|---------------------------------|-------|-------------------|---------|--------------|--------|-------------------|----------|---------|----------|--------------|--|
| | Logging | | Execution History | | Step Metrics | | Performance Graph | | Metrics | | Preview data | |
| # | Stepname | Copyn | Read | Written | Input | Output | Updated | Rejected | Errors | Active | Time | |
| 1 | CSV file input product | 0 | 0 | 288 | 289 | 0 | 0 | 0 | 0 | Finished | 0.0s | |
| 2 | Dimension lookup/update product | 0 | 288 | 288 | 288 | 288 | 0 | 0 | 0 | Finished | 0.3s | |

- Terdapat total data 289, yang dimana 1 data merupakan data template dari Pentaho sehingga data valid berjumlah 288 yang telah terinput ke database electronic pada table dim_salesrep dengan SCD Tipe 2

```
SELECT * FROM `dim product`
```

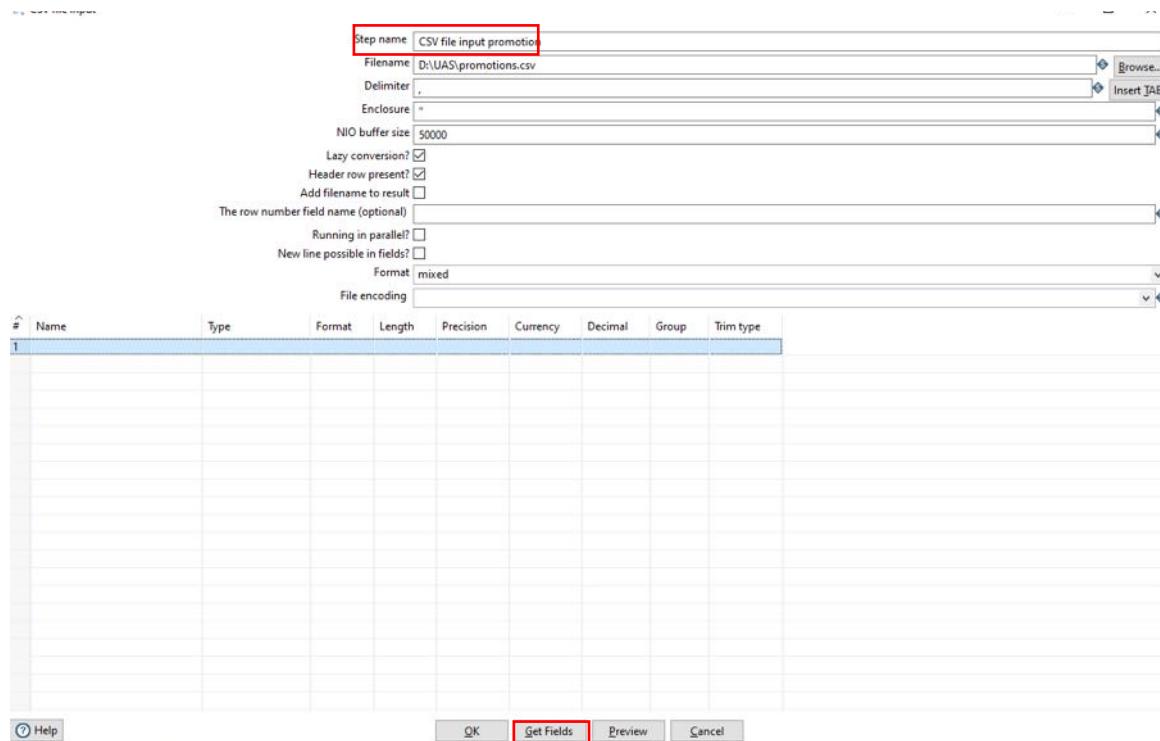
Showing rows 0 - 24 (289 total. Query took 0.0009 seconds.)

[SELECT](#) | [FROM](#) `dim_product`

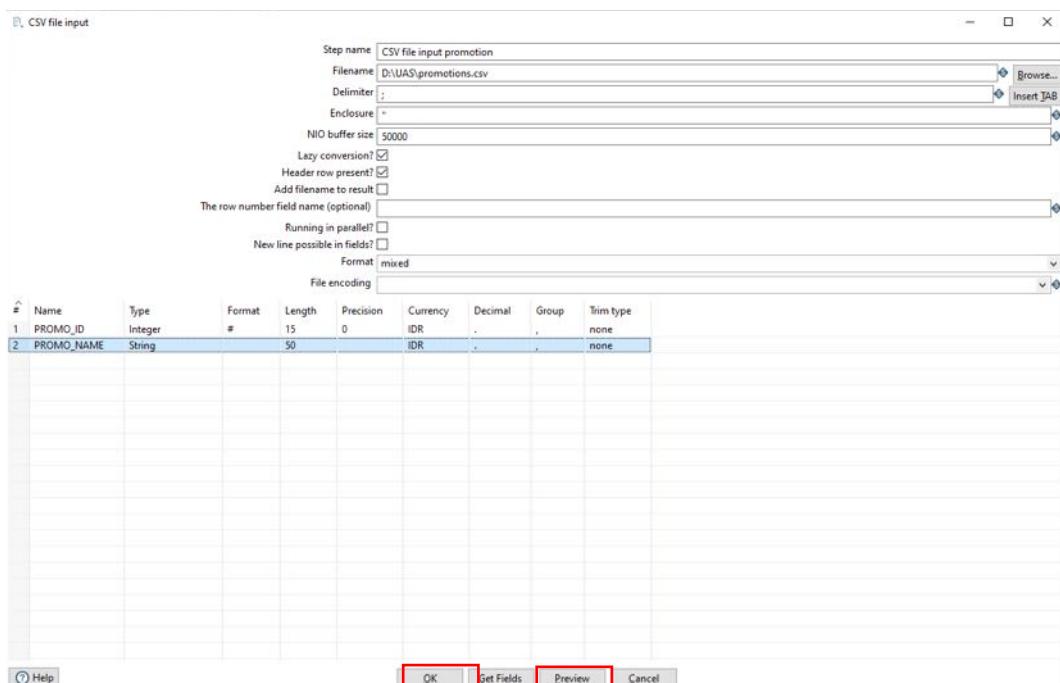
Profiling | [Edit inline](#) | [Edit](#) | [Explain SQL](#) | [Create PHP code](#) | [Refresh](#)

| | product_dim_id | version | date_from | date_to | product_id | PRODUCT_NAME | LANGUAGE_ID | MIN_PRICE | LIST_PRICE | PRODUCT_STATUS | SUPPLIER_ID | WARRANTY_PERIOD | WEIGHT_CLASS | PR | |
|-------------------------------------|-------------------------|----------------------|------------------------|---------|------------|---------------------|---------------------|-----------|--------------------|----------------|-------------|-----------------|-------------------|--------|----|
| <input type="checkbox"/> | Edit | Copy | Delete | 0 | 1 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | |
| <input type="checkbox"/> | Edit | Copy | Delete | 1 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 3054 | 519 | 600 | obsolete | 102060 | 1 | 3 | |
| <input type="checkbox"/> | Edit | Copy | Delete | 2 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 1782 | Compact 400/DQ | US | 108 | 125 | obsolete | 102088 | -5 |
| <input type="checkbox"/> | Edit | Copy | Delete | 3 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2430 | Compact 400/LQ | US | 143 | 175 | orderable | 102087 | 2 |
| <input type="checkbox"/> | Edit | Copy | Delete | 4 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 1792 | Industrial 600/DQ | US | 180 | 225 | orderable | 102088 | 5 |
| <input type="checkbox"/> | Edit | Copy | Delete | 5 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 1791 | Industrial 700/HQ | US | 239 | 275 | orderable | 102086 | 5 |
| <input type="checkbox"/> | Edit | Copy | Delete | 6 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2302 | Injet B/6 | US | 121 | 150 | obsolete | 102096 | 2 |
| <input type="checkbox"/> | Edit | Copy | Delete | 7 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 4453 | Injet C/4 | US | 174 | 195 | orderable | 102090 | 2 |
| <input type="checkbox"/> | Edit | Copy | Delete | 8 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 1797 | Injet C/BHQ | US | 288 | 349 | orderable | 102094 | 2 |
| <input type="checkbox"/> | Edit | Copy | Delete | 9 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2459 | LaserPro 1200/B/BW | US | 568 | 699 | under development | 102099 | 3 |
| <input type="checkbox"/> | Edit | Copy | Delete | 10 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 3127 | LaserPro 600/B/BW | US | 444 | 498 | orderable | 102087 | 3 |
| <input type="checkbox"/> | Edit | Copy | Delete | 11 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2254 | HD 10GB/L | US | 371 | 453 | obsolete | 102071 | 2 |
| <input type="checkbox"/> | Edit | Copy | Delete | 12 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 3353 | HD 10GB/R | US | 413 | 489 | obsolete | 102071 | 3 |
| <input type="checkbox"/> | Edit | Copy | Delete | 13 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 3069 | HD 10GB/S | US | 350 | 436 | obsolete | 102051 | 2 |
| <input type="checkbox"/> | Edit | Copy | Delete | 14 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2253 | HD 10GB/@5400 | US | 322 | 399 | obsolete | 102069 | 3 |
| <input checked="" type="checkbox"/> | Console | Copy | Delete | 15 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 3088 | SDRAM - 32 MB | US | 220 | 258 | orderable | 102057 | -9 |

4. PROMOTION_TRANSFORM (DIMENSION TABLE)

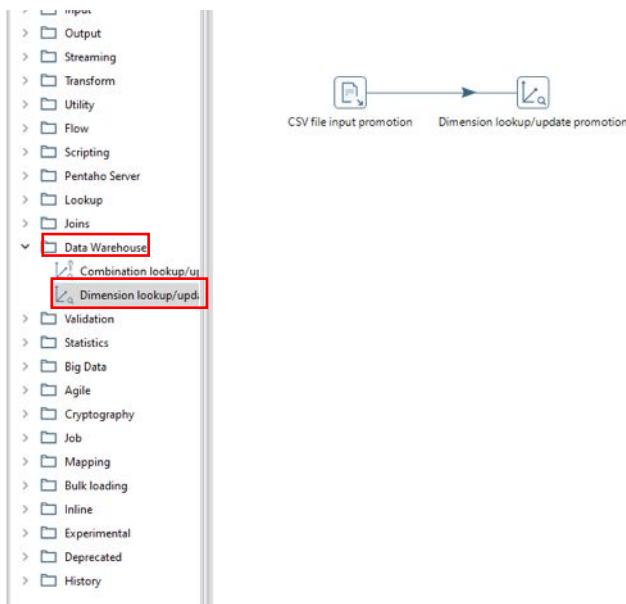


- Berikan nama pada Step name, dalam hal ini yang digunakan adalah promotion
- Pilih filename
- Pastikan Delimeter telah sesuai
- Kemudian pilih GetFields

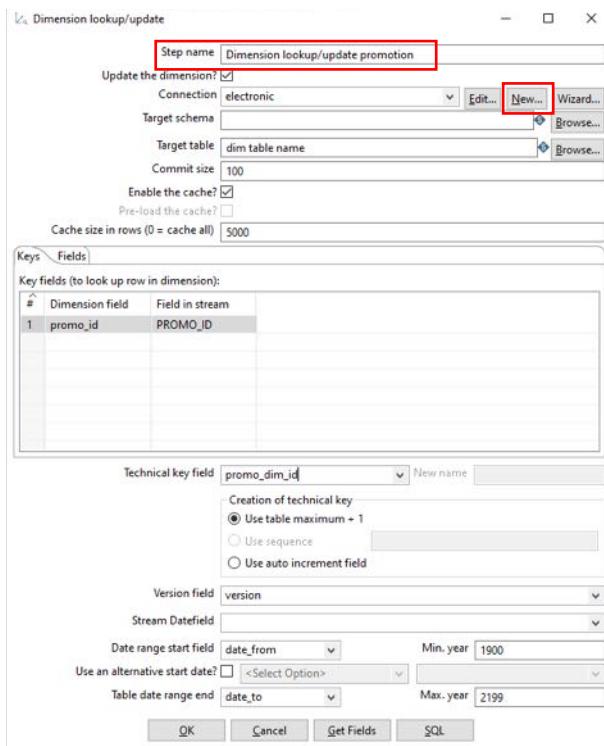


- Periksa data dengan klik tombol Preview
- Jika sudah sesuai boleh lanjutkan klik tombol OK

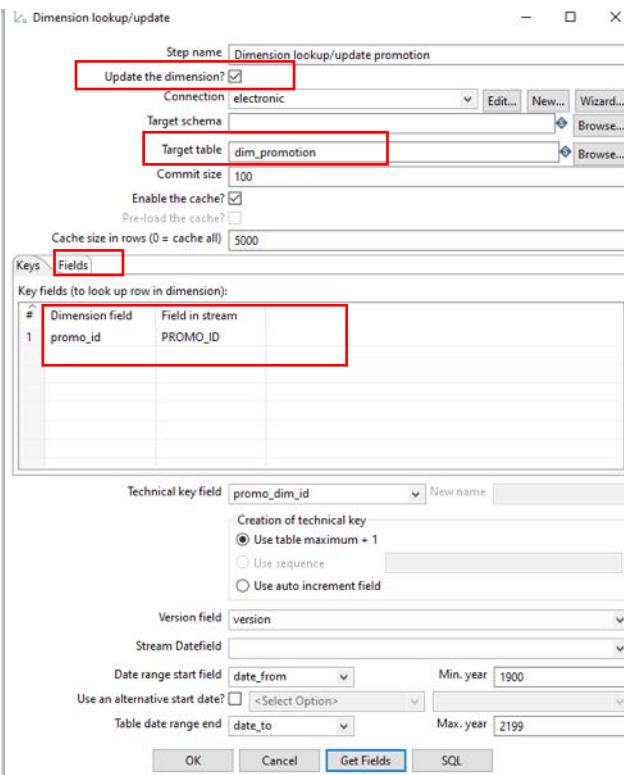
Setelah Langkah diatas sudah dilaksanakan, selanjutnya kita perlu membuat dimension promotion (PROMOTION_DIM) menggunakan Data Warehouse – Dimension lookup/update dan koneksi menggunakan konektor.



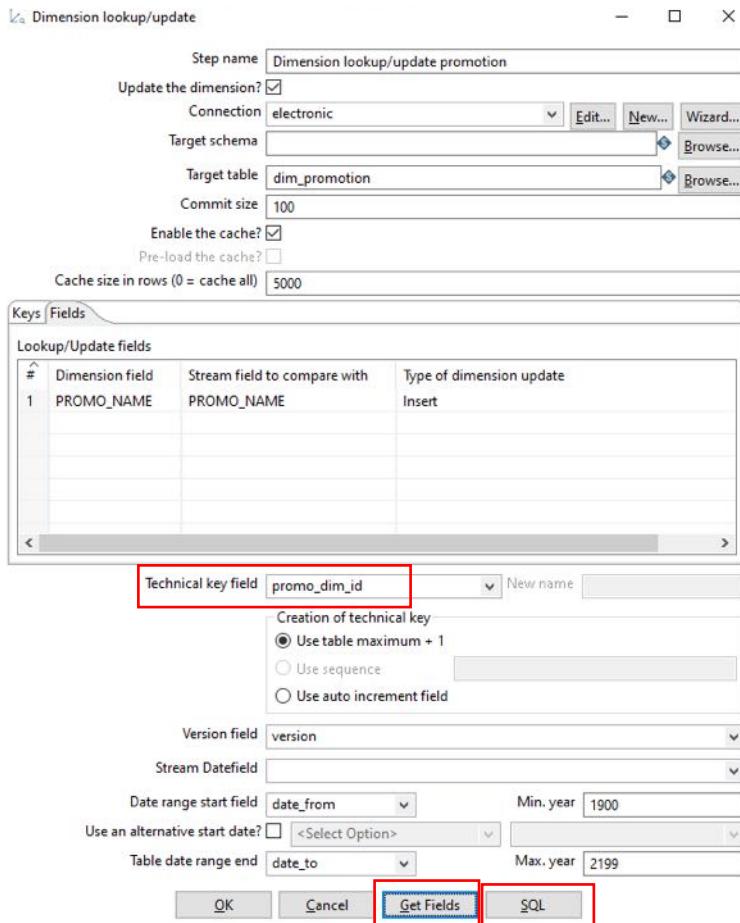
- Double klik pada Dimension lookup/update



- Berikan nama pada Step name
- Pilih koneksi yang sudah dibuat dan terhubung ke MySQL



- Checklist Update the dimension, untuk dilakukan Slowly Change Dimension (SCD) jika terdapat perubahan data akan tercatat.
- Untuk memetakannya kita perlu menentukan nama Dimension field dari Field in stream yang sudah ada pada CSV dan juga merupakan Primary Key aslinya.
- Pilih Tab Fields



- Pada Tab Fields, klik Get Fields.
- Berikan input nama pada Technical key field yang merupakan nama lain surrowgate key untuk memudahkan dalam pencatatan perubahan data.
- Pilih SQL

```

Simple SQL editor

SQL statements, separated by semicolon ';

CREATE TABLE dim_promotion
(
    promo_dim_id BIGINT NOT NULL PRIMARY KEY
    ,version INT
    ,date_from DATETIME
    ,date_to DATETIME
    ,promo_id BIGINT
    PROMO_NAME VARCHAR(50)
)
CREATE INDEX idx_dim_promotion_lookup ON dim_promotion(promo_id)
CREATE INDEX idx_dim_promotion_tk ON dim_promotion(promo_dim_id)
;

```

Line 1 column 0

Execute Clear cache Close

- promo_dim_id adalah tambahan yang merupakan surrounge key
- version, date_from, dan date_to juga merupakan tambahan dari SCD Tipe 2 Pentaho
- sisanya merupakan data asli dari CSVnya
- Jika sesuai, pilih Execute

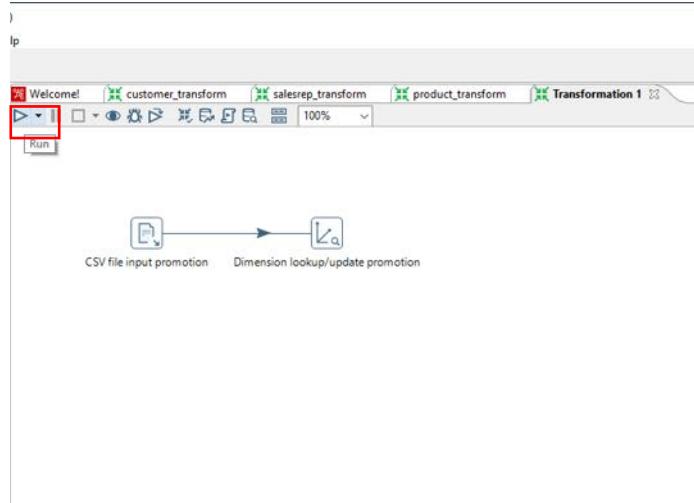
The screenshot shows the phpMyAdmin interface with the following details:

- Database:** electronic
- Table:** dim_promotion
- Structure View:** Shows columns:

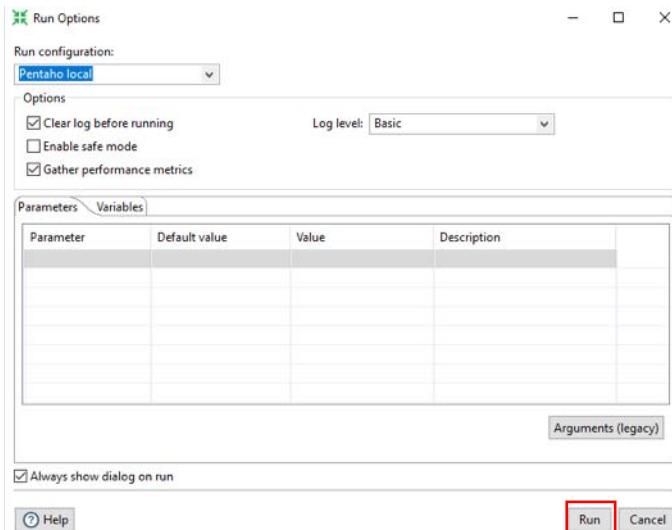
| # | Name | Type | Collation | Attributes | Null | Default | Comments | Extra | Action |
|---|--------------|-------------|--------------------|------------|------|---------|----------|-------|--|
| 1 | promo_dim_id | bigint(20) | | | No | None | | | Change Drop More |
| 2 | version | int(11) | | | Yes | NULL | | | Change Drop More |
| 3 | date_from | datetime | | | Yes | NULL | | | Change Drop More |
| 4 | date_to | datetime | | | Yes | NULL | | | Change Drop More |
| 5 | promo_id | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 6 | PROMO_NAME | varchar(50) | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |
- Indexes View:** Shows three indexes:

| Action | Keyname | Type | Unique | Packed | Column | Cardinality | Collation | Null | Comment |
|--|--------------------------|-------|--------|--------|--------------|-------------|-----------|------|---------|
| Edit Rename Drop | PRIMARY | BTREE | Yes | No | promo_dim_id | 0 | A | No | |
| Edit Rename Drop | idx_dim_promotion_lookup | BTREE | No | No | promo_id | 0 | A | Yes | |
| Edit Rename Drop | idx_dim_promotion_tk | BTREE | No | No | promo_dim_id | 0 | A | No | |
- Partitions View:** Shows a message: "No partitioning defined!"

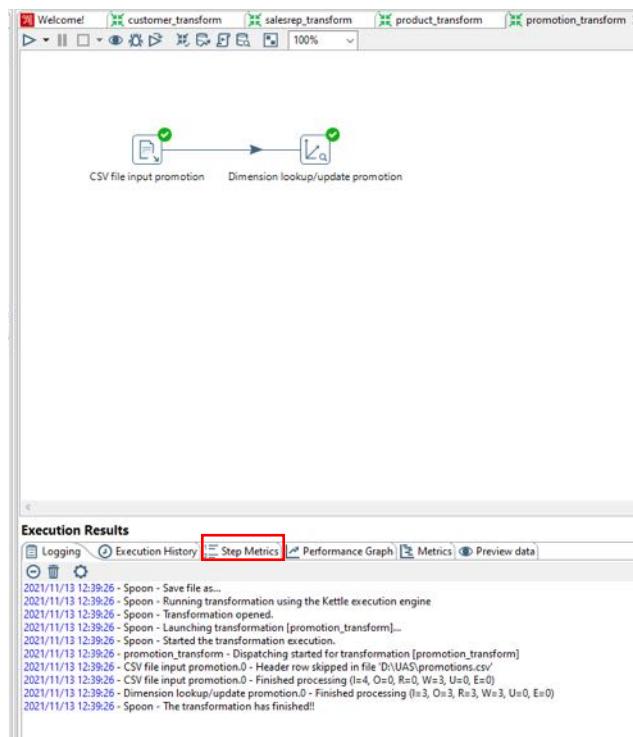
- Pentaho secara langsung menggenerate data ke MySQL dengan nama table dim_promotion



- Jalankan promotion_transform dengan klik Run



- Klik Run



- Data telah berhasil dieksekusi dan dapat diperiksa pada Tab Step Metrics

Execution Results

| # | Stepname | Copynr | Read | Written | Input | Output | Updated | Rejected | Errors | Active | Time |
|---|-----------------------------------|--------|------|---------|-------|--------|---------|----------|--------|----------|------|
| 1 | CSV file input promotion | 0 | 0 | 3 | 4 | 0 | 0 | 0 | 0 | Finished | 0.0s |
| 2 | Dimension lookup/update promotion | 0 | 3 | 3 | 3 | 3 | 0 | 0 | 0 | Finished | 0.1s |

- Terdapat total data 4, yang dimana 1 data merupakan data template dari Pentaho sehingga data valid berjumlah 3 yang telah terinput ke database electronic pada table dim_promotion dengan SCD Tipe 2

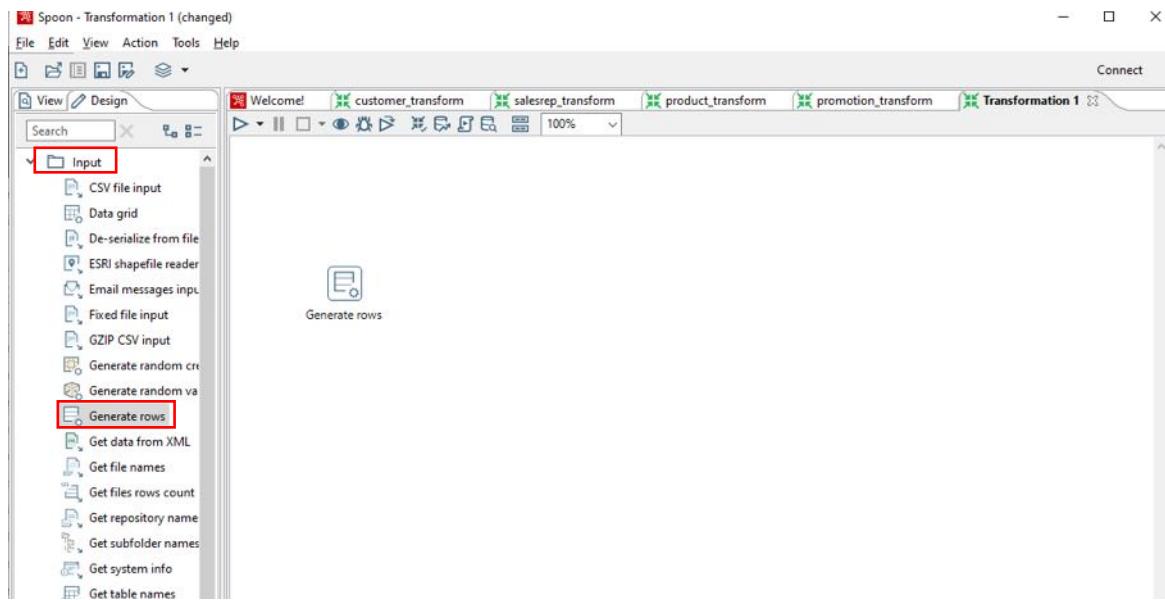
```
SELECT * FROM `dim_promotion`
```

Showing rows 0 - 3 (4 total). Query took 0.0006 seconds.

| + Options | | | | | |
|--------------------------|----------------------|---------|---------------------|-----------------------|---------------------|
| | promo_dim_id | version | date_from | date_to | promo_id PROMO_NAME |
| <input type="checkbox"/> | Edit Copy Delete 0 | 1 | NULL | NULL | NULL NULL |
| <input type="checkbox"/> | Edit Copy Delete 1 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 1 | summer sale |
| <input type="checkbox"/> | Edit Copy Delete 2 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 2 | christmast |
| <input type="checkbox"/> | Edit Copy Delete 3 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 3 | new year |

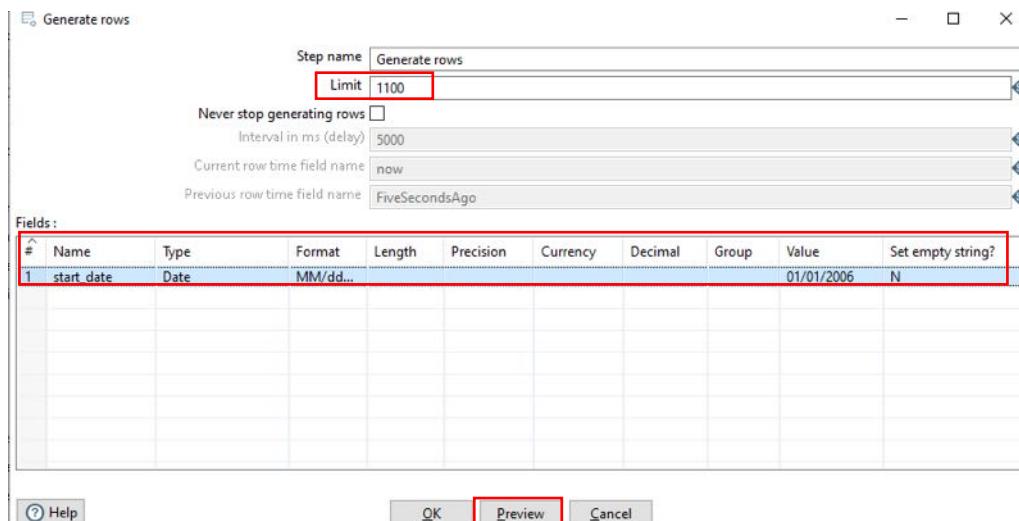
5. DATE_TRANSFORM (DIMENSION TABLE)

Merupakan transformasi yang akan dibuat/generate untuk membentuk sebuah dimensi tanggal secara manual.



- Pilih Generate rows pada Input
- Double klik pada Generate rows

Data yang dimiliki merupakan data penjualan barang elektronik selama 3 tahun sejak Januari tahun 2006 – Desember tahun 2008 yang dapat dilihat pada CSV orders.



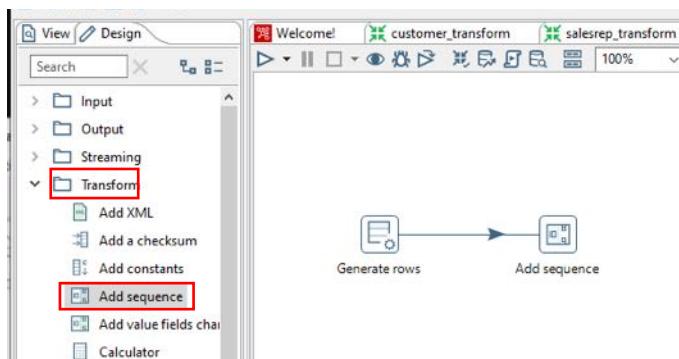
- 1 tahun berjumlah 365 hari, karena 3 tahun membutuhkan $365 \times 3 = 1095$, disini saya membulatkan menjadi 1100 yang input pada Limit.
- Dalam membuat tanggal dasar selama 3 tahun diperlukan field baru yang diberi nama start_date, dengan type Date, format MM/dd/yy, value start 01/01/2006, dan tanpa set empty string menjadi N.
- Lalu pilih Preview untuk memeriksa.

| # | start_date |
|----|------------|
| 1 | 01/01/06 |
| 2 | 01/01/06 |
| 3 | 01/01/06 |
| 4 | 01/01/06 |
| 5 | 01/01/06 |
| 6 | 01/01/06 |
| 7 | 01/01/06 |
| 8 | 01/01/06 |
| 9 | 01/01/06 |
| 10 | 01/01/06 |
| 11 | 01/01/06 |
| 12 | 01/01/06 |
| 13 | 01/01/06 |
| 14 | 01/01/06 |
| 15 | 01/01/06 |
| 16 | 01/01/06 |
| 17 | 01/01/06 |
| 18 | 01/01/06 |
| 19 | 01/01/06 |
| 20 | 01/01/06 |
| 21 | 01/01/06 |
| 22 | 01/01/06 |
| 23 | 01/01/06 |
| 24 | 01/01/06 |
| 25 | 01/01/06 |
| 26 | 01/01/06 |
| 27 | 01/01/06 |
| 28 | 01/01/06 |
| 29 | 01/01/06 |
| 30 | 01/01/06 |
| 31 | 01/01/06 |
| 32 | 01/01/06 |
| 33 | 01/01/06 |

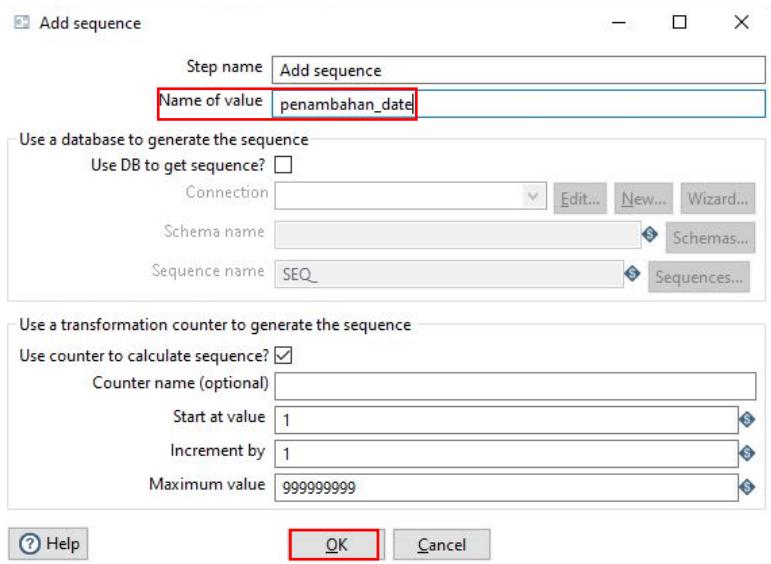
Close **Show Log**

- Ini merupakan tanggal dasar yang nanti menjadi dasar pada Langkah selanjutnya
- Pilih Close

Perlu ditambahkan setiap baris dengan 1/2/3 nilai/value dengan transformasi tabelnya menggunakan AddSequence.

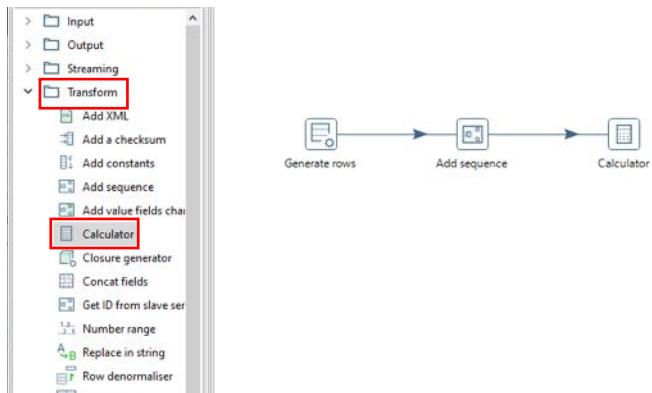


- Pilih Transform
- Pilih Add sequence
- Koneksikan Generate rows menuju Add sequence
- Double klik pada Add sequence

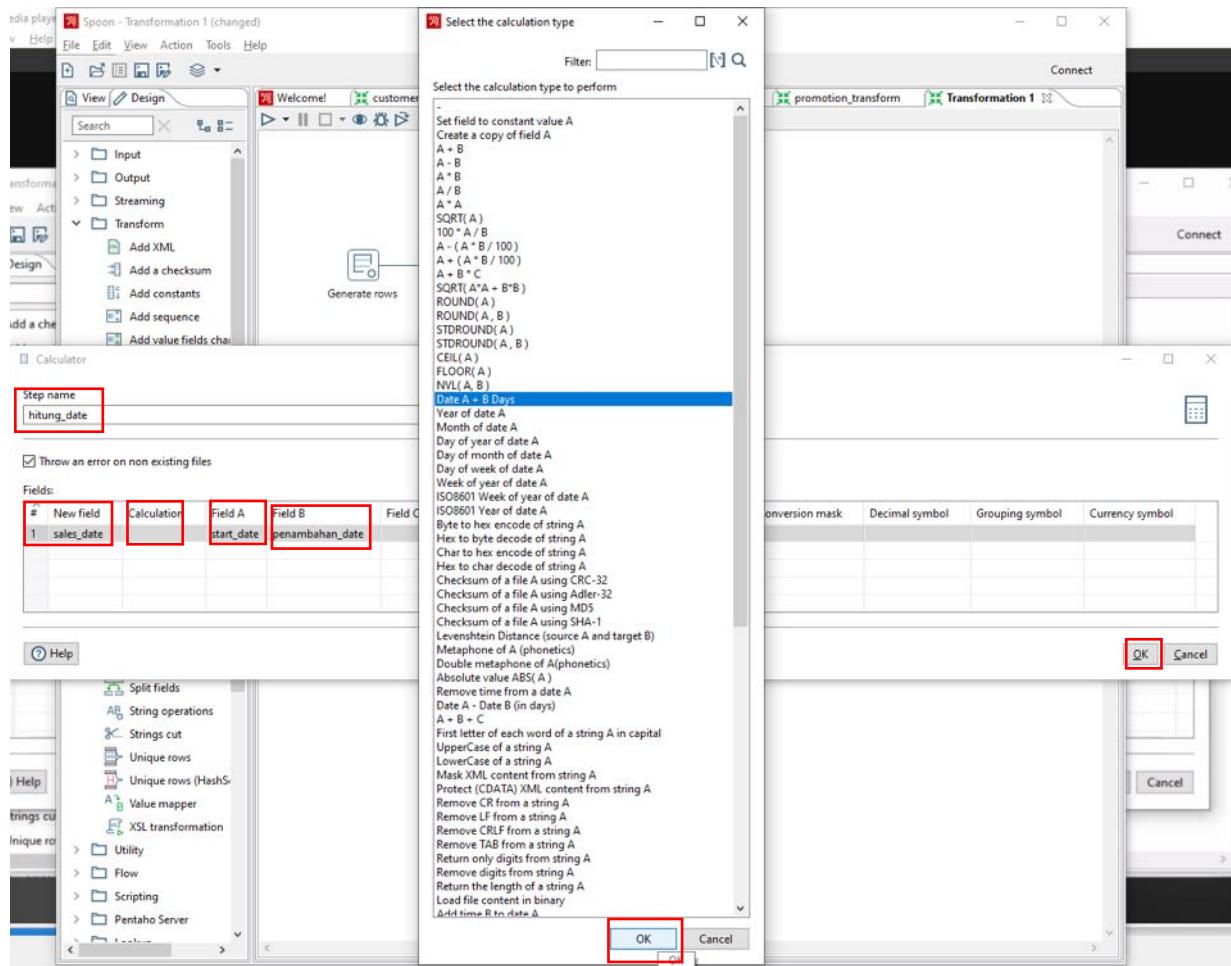


- Input nama pada Name of value, karena fungsi ini untuk penambahan tanggal secara berurutan maka diberi nama penambahan_date
- Klik OK

Karena diperlukan fungsi penambahan dari Generate rows dan Add sequence, kita membutuhkan fungsi pada Transform Calculator

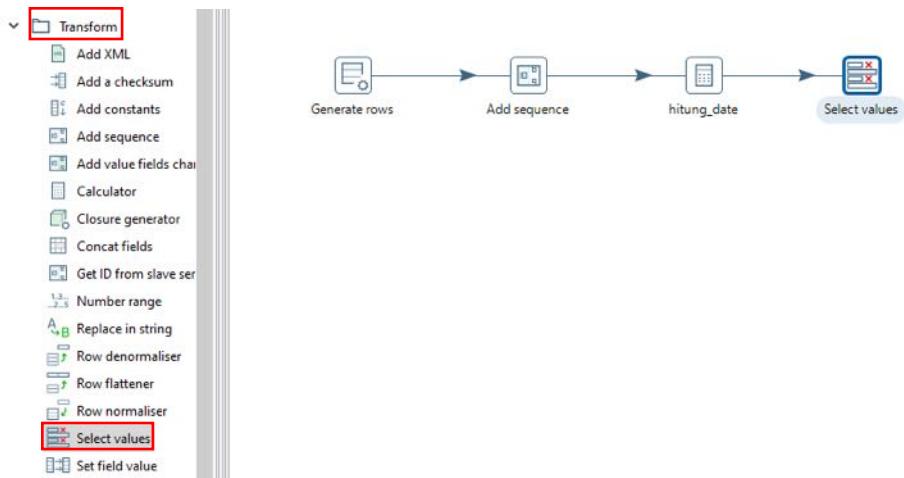


- Pilih Transform
- Pilih Calculator
- Koneksikan Add sequence ke Calculator
- Double klik pada Calculator

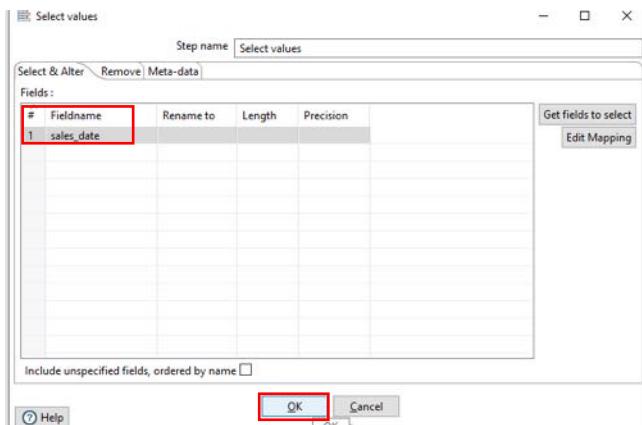


- Input step name dengan nama hitung_date
Karena ini akan menjadi bagian dari sales, maka kita perlu membuat field baru untuk agregasi dari start_date dan penambahan_date.
- Beri nama New field adalah sales_date
- Field A pilih start_date
- Field B pilih penambahan_date
- Pada bagian Calculation sales_date karena kita akan menambah antara start_date (A) dan penambahan_date (B) maka rumus yang dipilih adalah Date A + B Days
- Klik OK

Karena yang dibutuhkan hanya hasil generate saja, perlu ditampilkan sales_date dengan memanfaatkan fungsi Select values

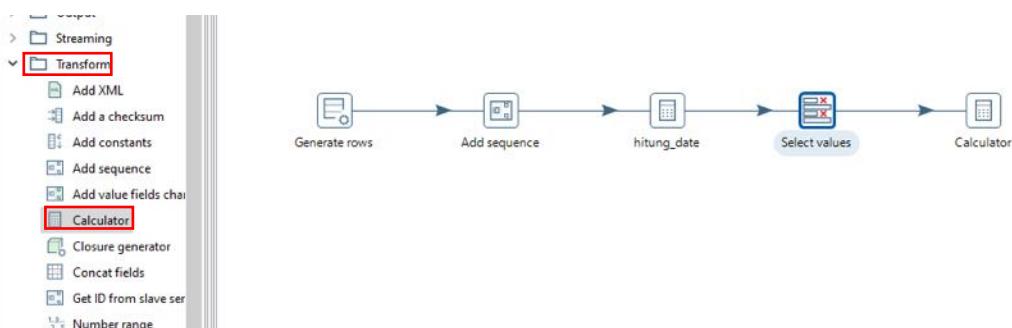


- Pilih Transform
- Pilih Select values
- Koneksikan Calculator hitung_date dengan fungsi Select values
- Double klik pada Select values

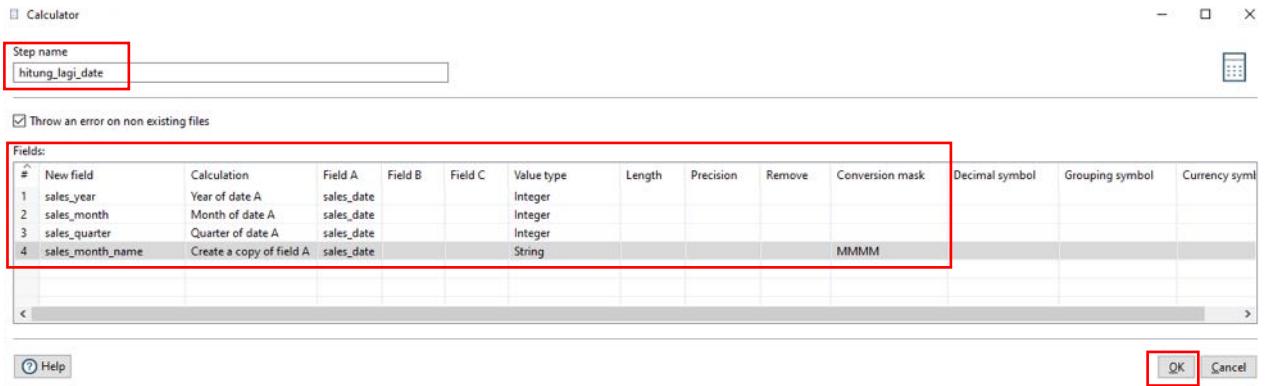


- Pada Fieldname pilih sales_date
- Klik OK

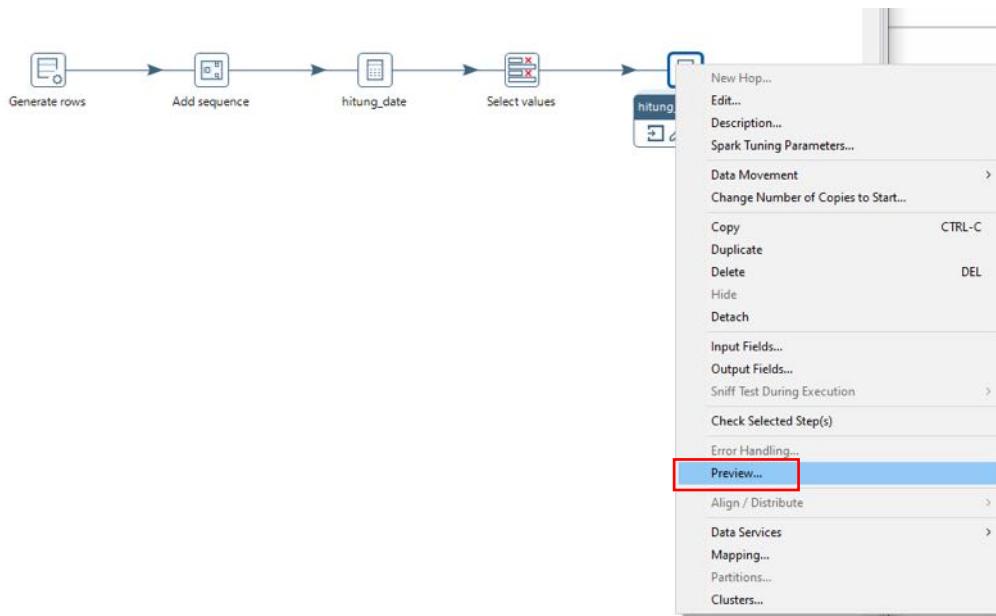
Kita akan lakukan pemisahan berdasarkan kebutuhan Tahun, Bulan, dan Hari maka kita perlu fungsi Calculator lagi



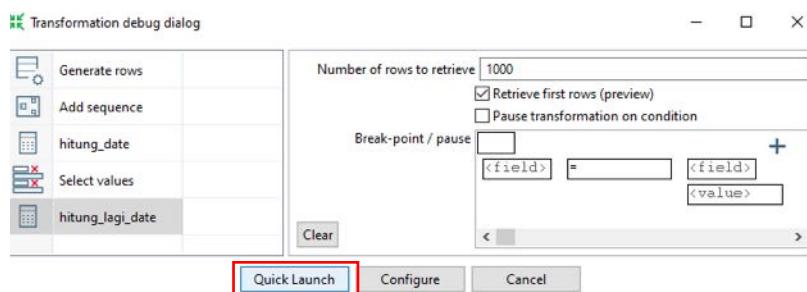
- Pilih Transform
- Pilih Calculator
- Koneksikan fungsi Select values dengan Calculator
- Double klik pada Calculator



- Input Step name menjadi hitung_lagi_date dalam satuan integer
- New field sales_year untuk mengambil tahun dalam satuan integer
- New field sales_month untuk mengambil bulan dalam satuan integer
- New field sales_quarter untuk mengambil periode/semester dalam satuan integer
- New field sales_month_name untuk mengambil nama Bulan dalam satuan String
- Klik OK



- Periksa data dengan klik kanan pada calculator hitung_lagi date
- Pilih Preview



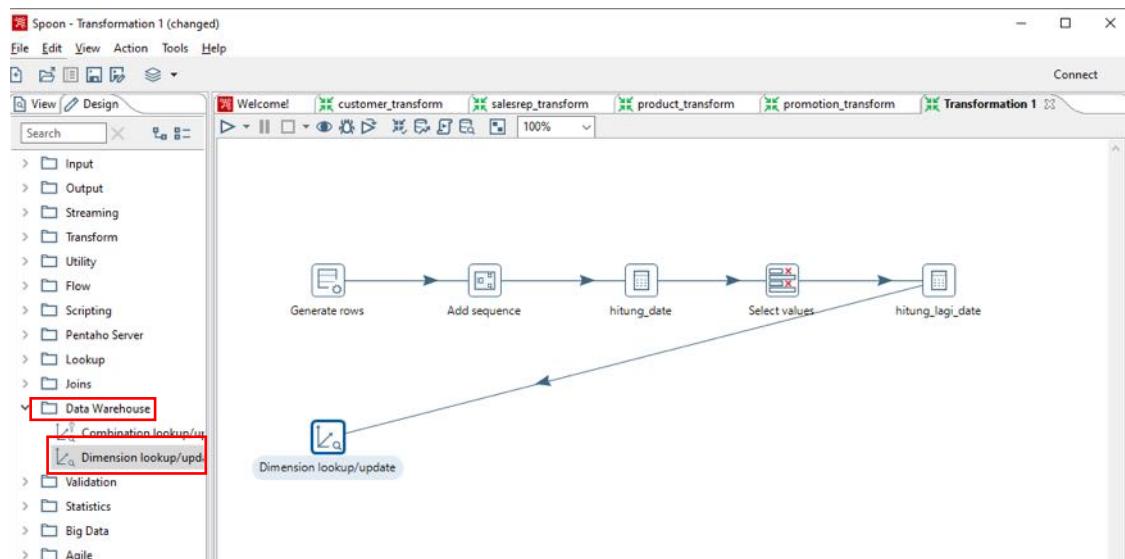
- Pilih Quick Launch

Examine preview data

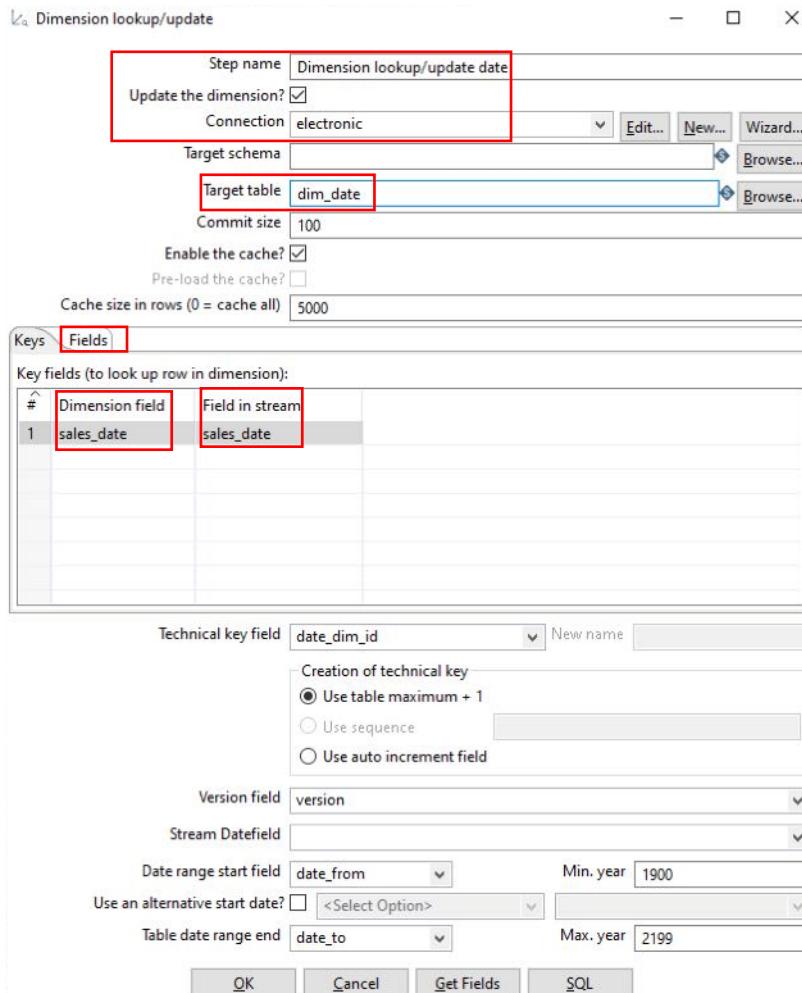
Rows of step: hitung_lagi_date (1000 rows)

| # | sales_date | sales_year | sales_month | sales_quarter | sales_month_name |
|----|-------------------------|------------|-------------|---------------|------------------|
| 1 | 2006/01/02 00:00:00.000 | 2006 | 1 | 1 | January |
| 2 | 2006/01/03 00:00:00.000 | 2006 | 1 | 1 | January |
| 3 | 2006/01/04 00:00:00.000 | 2006 | 1 | 1 | January |
| 4 | 2006/01/05 00:00:00.000 | 2006 | 1 | 1 | January |
| 5 | 2006/01/06 00:00:00.000 | 2006 | 1 | 1 | January |
| 6 | 2006/01/07 00:00:00.000 | 2006 | 1 | 1 | January |
| 7 | 2006/01/08 00:00:00.000 | 2006 | 1 | 1 | January |
| 8 | 2006/01/09 00:00:00.000 | 2006 | 1 | 1 | January |
| 9 | 2006/01/10 00:00:00.000 | 2006 | 1 | 1 | January |
| 10 | 2006/01/11 00:00:00.000 | 2006 | 1 | 1 | January |
| 11 | 2006/01/12 00:00:00.000 | 2006 | 1 | 1 | January |
| 12 | 2006/01/13 00:00:00.000 | 2006 | 1 | 1 | January |
| 13 | 2006/01/14 00:00:00.000 | 2006 | 1 | 1 | January |
| 14 | 2006/01/15 00:00:00.000 | 2006 | 1 | 1 | January |
| 15 | 2006/01/16 00:00:00.000 | 2006 | 1 | 1 | January |
| 16 | 2006/01/17 00:00:00.000 | 2006 | 1 | 1 | January |
| 17 | 2006/01/18 00:00:00.000 | 2006 | 1 | 1 | January |
| 18 | 2006/01/19 00:00:00.000 | 2006 | 1 | 1 | January |
| 19 | 2006/01/20 00:00:00.000 | 2006 | 1 | 1 | January |
| 20 | 2006/01/21 00:00:00.000 | 2006 | 1 | 1 | January |
| 21 | 2006/01/22 00:00:00.000 | 2006 | 1 | 1 | January |
| 22 | 2006/01/23 00:00:00.000 | 2006 | 1 | 1 | January |
| 23 | 2006/01/24 00:00:00.000 | 2006 | 1 | 1 | January |
| 24 | 2006/01/25 00:00:00.000 | 2006 | 1 | 1 | January |
| 25 | 2006/01/26 00:00:00.000 | 2006 | 1 | 1 | January |
| 26 | 2006/01/27 00:00:00.000 | 2006 | 1 | 1 | January |
| 27 | 2006/01/28 00:00:00.000 | 2006 | 1 | 1 | January |
| 28 | 2006/01/29 00:00:00.000 | 2006 | 1 | 1 | January |
| 29 | 2006/01/30 00:00:00.000 | 2006 | 1 | 1 | January |
| 30 | 2006/01/31 00:00:00.000 | 2006 | 1 | 1 | January |
| 31 | 2006/02/01 00:00:00.000 | 2006 | 2 | 1 | February |
| 32 | 2006/02/02 00:00:00.000 | 2006 | 2 | 1 | February |
| 33 | 2006/02/03 00:00:00.000 | 2006 | 2 | 1 | February |

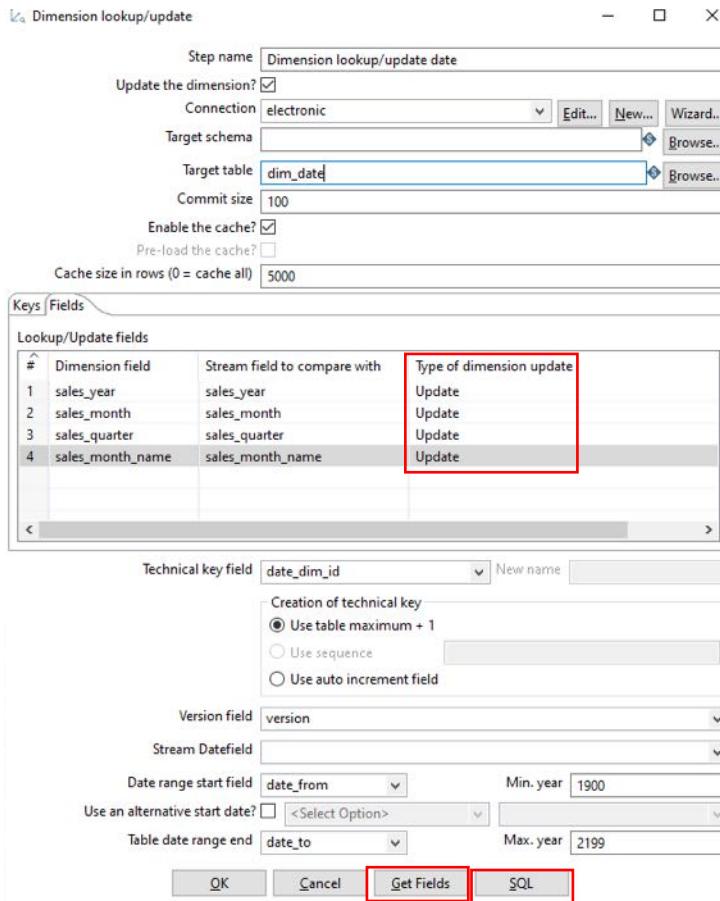
Setelah preview hitung_lagi_date telah sesuai, kemudian akan dijadikan sebuah dimensi tanggal pada Langkah selanjutnya sebagai berikut



- Pilih Data Warehouse
- Pilih Dimension lookup/update
- Koneksikan jaringan dari fungsi calculator hitung_lagi_date ke Dimension lookup/update
- Double klik pada Dimension lookup/update



- Input step name
- Input connection
- Input target table menjadi dim_date
- Input Field in stream menjadi sales_date
- Input Dimension field menjadi sales_date
- Input Technical key field untuk SCD/Surrow gate key menjadi date_dim_id
- Klik pada tab Fields



- Klik Get Fields
- Karena dimensi ini digenerate dan tidak mengganggu table/database operasional, maka Type of dimension update dapat menggunakan Update saja
- Klik SQL
- Pilih Execute

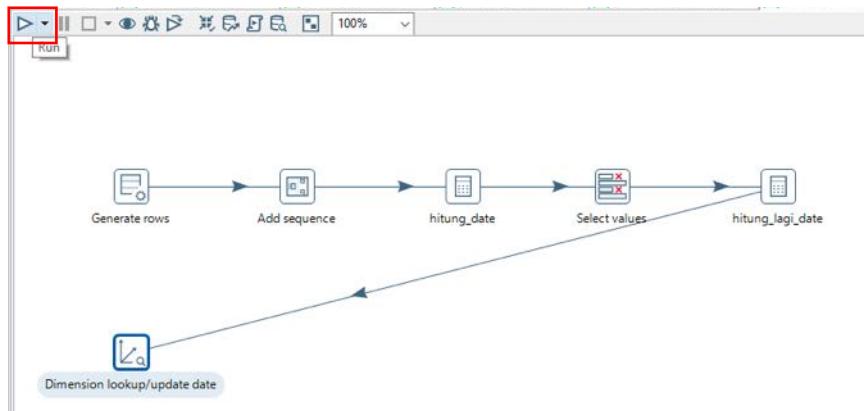
The screenshot shows the 'Table structure' for the 'dim_date' table in the 'electronic' database. The table has nine columns:

| # | Name | Type | Collation | Attributes | Null | Default | Comments | Extra | Action |
|---|------------------|------------|--------------------|------------|------|---------|----------|-------|--|
| 1 | date_dim_id | bigint(20) | | | No | None | | | Change Drop More |
| 2 | version | int(11) | | | Yes | NULL | | | Change Drop More |
| 3 | date_from | datetime | | | Yes | NULL | | | Change Drop More |
| 4 | date_to | datetime | | | Yes | NULL | | | Change Drop More |
| 5 | sales_date | datetime | | | Yes | NULL | | | Change Drop More |
| 6 | sales_year | int(11) | | | Yes | NULL | | | Change Drop More |
| 7 | sales_month | int(11) | | | Yes | NULL | | | Change Drop More |
| 8 | sales_quarter | int(11) | | | Yes | NULL | | | Change Drop More |
| 9 | sales_month_name | tinytext | utf8mb4_general_ci | | Yes | NULL | | | Change Drop More |

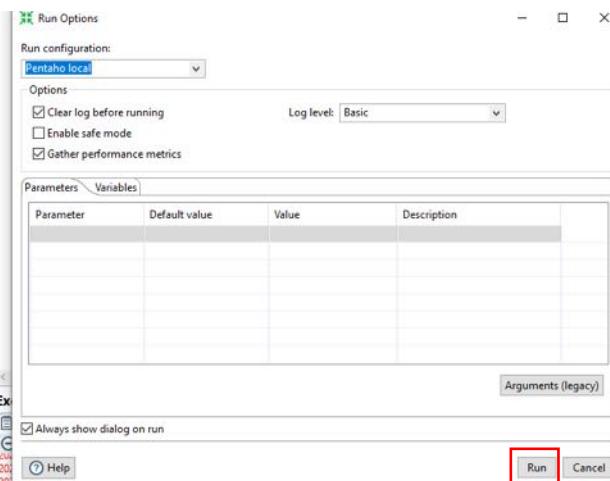
Below the table structure, there is an 'Indexes' section with three entries:

| Action | Keyname | Type | Unique | Packed | Column | Cardinality | Collation | Null | Comment |
|--|---------------------|-------|--------|--------|-------------|-------------|-----------|------|---------|
| Edit Rename Drop | PRIMARY | BTREE | Yes | No | date_dim_id | 0 | A | No | |
| Edit Rename Drop | idx_dim_date_lookup | BTREE | No | No | sales_date | 0 | A | Yes | |
| Edit Rename Drop | idx_dim_date_pk | BTREE | No | No | date_dim_id | 0 | A | No | |

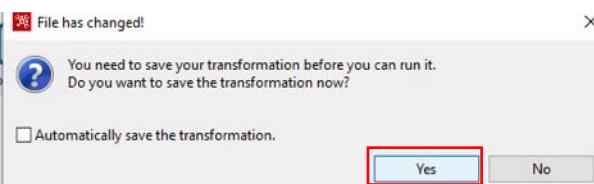
- Table dim_date secara otomatis dibuat pada database electronic di MySQL



- Jalankan proses untuk mengisi table dim_date menggunakan perintah Run pada Pentaho



- Pilih Run



- Pilih Yes
- Simpan dengan nama date_transform

```
SELECT * FROM `dim_date`
```

Showing rows 0 - 24 (1101 total, Query took 0.0006 seconds.)

SELECT * FROM `dim_date`

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

Number of rows: 25 Filter rows Search this table Sort by key: None

+ Options

| | | date_dim_id | version | date_from | date_to | sales_date | sales_year | sales_month | sales_quarter | sales_month_name |
|--------------------------|---------------------|-------------|---------------------|---------------------|---------------------|------------|------------|-------------|---------------|------------------|
| <input type="checkbox"/> | Edit Copy Delete 0 | 1 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| <input type="checkbox"/> | Edit Copy Delete 1 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-02 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 2 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-03 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 3 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-04 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 4 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-05 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 5 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-06 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 6 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-07 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 7 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-08 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 8 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-09 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 9 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-10 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 10 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-11 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 11 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-12 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 12 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-13 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 13 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-14 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 14 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-15 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 15 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-16 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 16 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-17 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 17 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-18 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 18 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-19 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 19 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-20 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 20 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-21 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 21 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-22 00:00:00 | 2006 | 1 | 1 | 1 | January |
| <input type="checkbox"/> | Edit Copy Delete 22 | 1 | 1900-01-01 00:00:00 | 2199-12-31 23:59:59 | 2006-01-23 00:00:00 | 2006 | 1 | 1 | 1 | January |

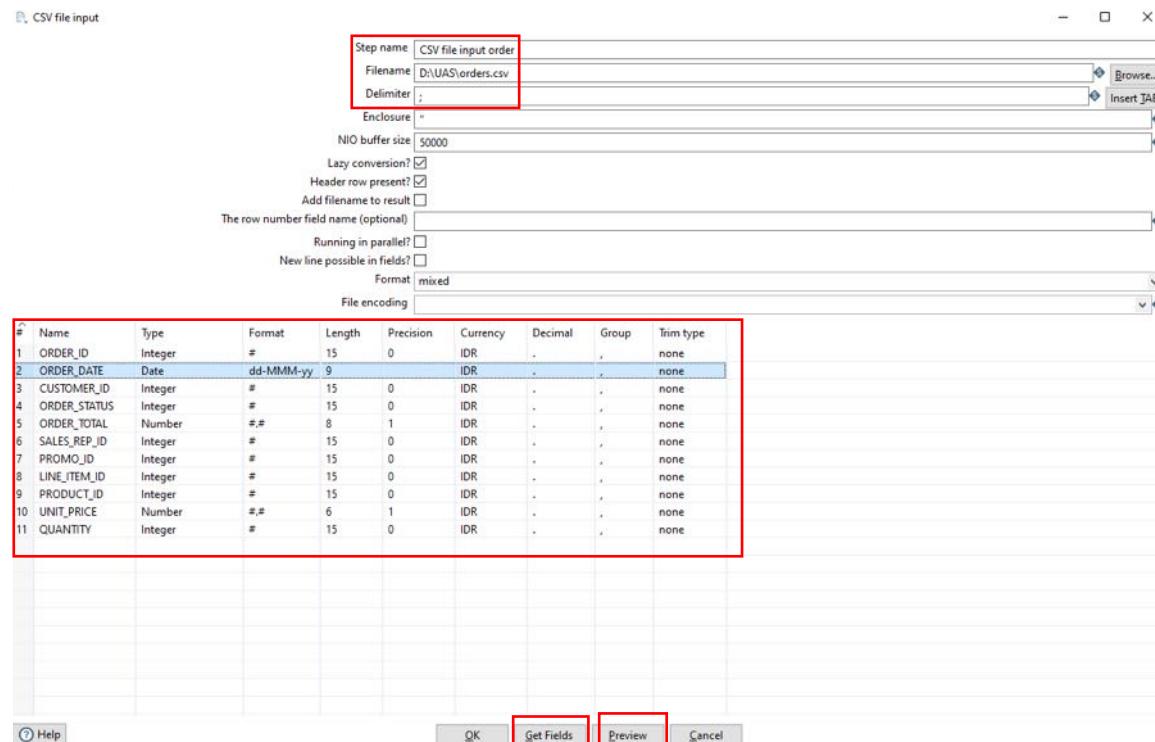
Data Date telah terisi pada table dim_date.

Seluruh table dimensi sudah dibuat, Langkah terakhir adalah membuat table Fact yang akan dibahas pada poin selanjutnya dengan transform baru.

[25%] Pilih satu ETL tools, jelaskan dengan detil setiap langkah yang dilakukan pada proses ETL nya. Harus ada contoh nyata. Jadi, dari tabel-tabel operasionalnya menjadi tabel dimensi dan tabel faktanya bagaimana. Jelaskan bagaimana memilih atribut- atribut dari *source table* lalu memasukkannya ke tabel dimensi. Jelaskan juga bagaimana agregasi dilakukan untuk mengisi fakta pada tabel fakta.

6. ORDER_TRANSFORM (FACT TABLE)

Base fact tablenya adalah table order, perlu dibuat transform baru dengan input CSV sebagai berikut



- Input step name menjadi order
- Pilih Filename
- Periksa Delimeter
- Pilih Get Fields
- Lakukan penyesuaian Type data pada Fields
- Pilih Preview untuk memvalidasi kembali kesesuaian data
- Close dan OK

| SALES_FACT | |
|------------|------------------|
| F | PRODUCT_DIM_ID |
| F | PROMOTION_DIM_ID |
| F | CUSTOMER_DIM_ID |
| F | SALESREP_DIM_ID |
| F | DATE_DIM_ID |
| | IDR_SOLD |
| | QUANTITY SOLD |

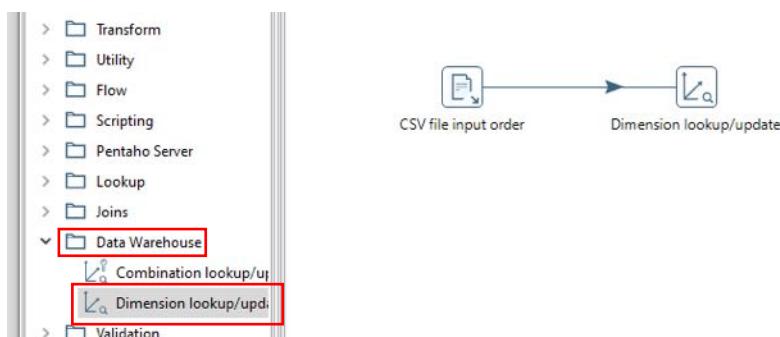
- PRODUCT_DIM_ID diambil dari table dim_product
- PROMOTION_DIM_ID diambil dari table dim_promotion
- CUSTOMER_DIM_ID diambil dari table dim_salesrep
- DATE_DIM_ID diambil dari table dim_date

Perlu dihitung total terjual di group berdasarkan table dimension dengan memeriksa

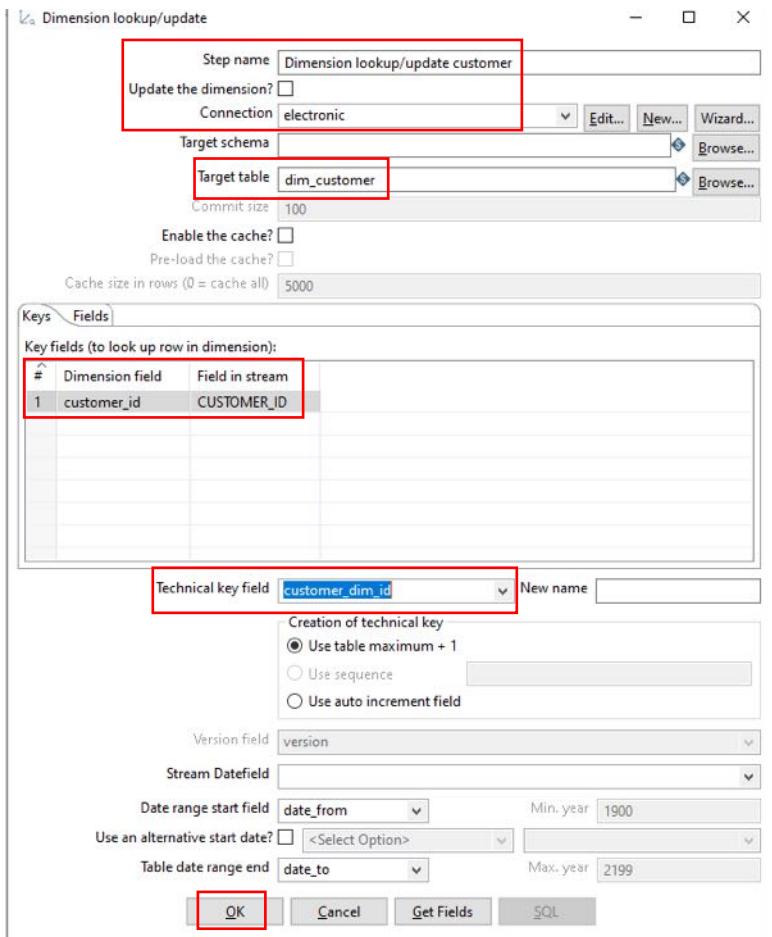
- CUSTOMER_ID
- SALES REP_ID

Karena perlu Lookup, dibutuhkan fungsi Dimension Lookup (bukan untuk update tapi lebih hanya untuk lookup saja)

a) Lakukan pencocokan dari sisi customer (GROUP berdasarkan customer)

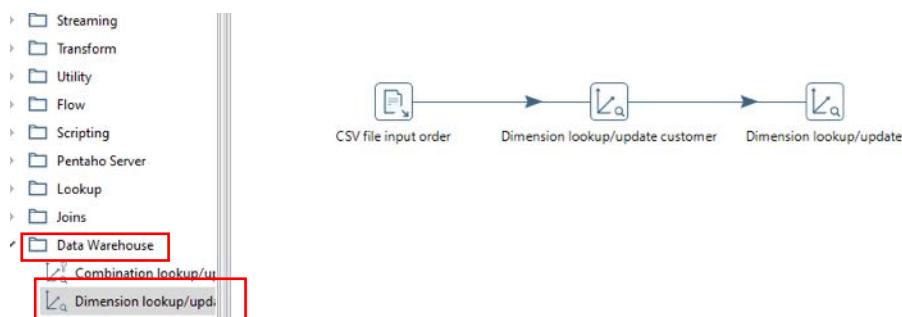


- Pilih Data Warehouse
- Pilih Dimension lookup/update
- Koneksikan CSV file input oder ke fungsi Dimension lookup/update untuk customer
- Double klik pada Dimension lookup/update



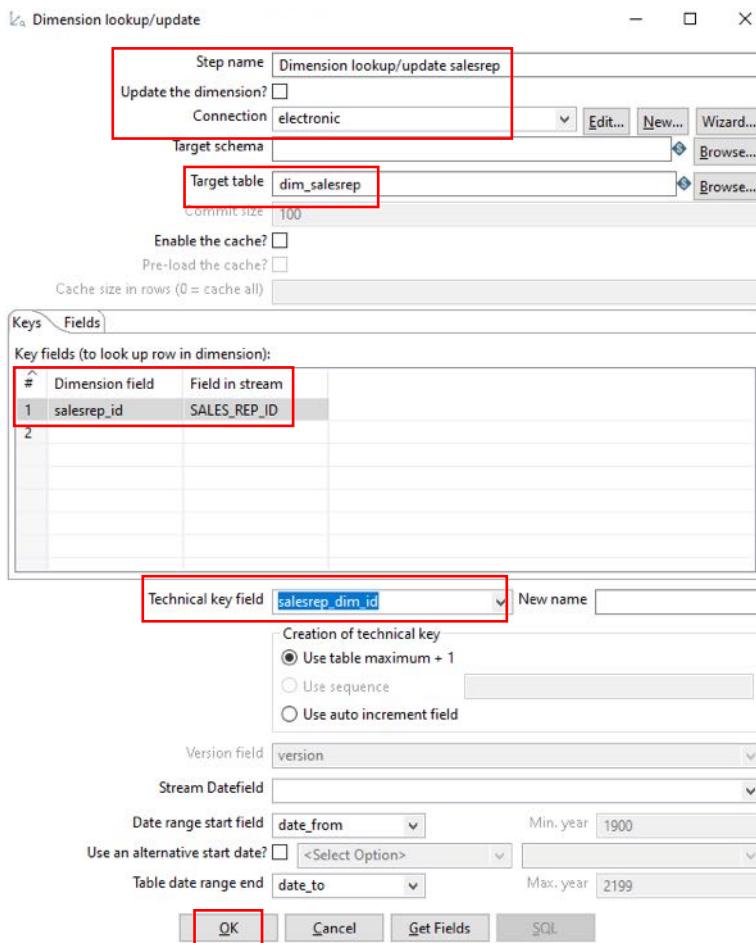
- Input Step name
- Input Connection
- Uncheck pada Update the dimension
- Uncheck pada Enable the cache
- Cocokkan Dimension Field dan Field in stream
- Pilih Technical key field dengan customer_dim_id
- Klik OK

b) Lakukan pencocokan dari sisi salesrep (GROUP berdasarkan salesrep)



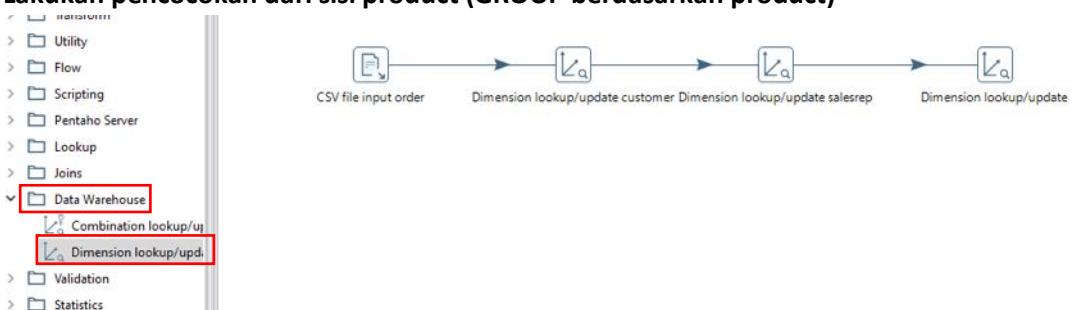
- Pilih Data Warehouse
- Pilih Dimension lookup/update
- Koneksikan CSV file input oder ke fungsi Dimension lookup/update untuk salesrep

- Double klik pada Dimension lookup/update



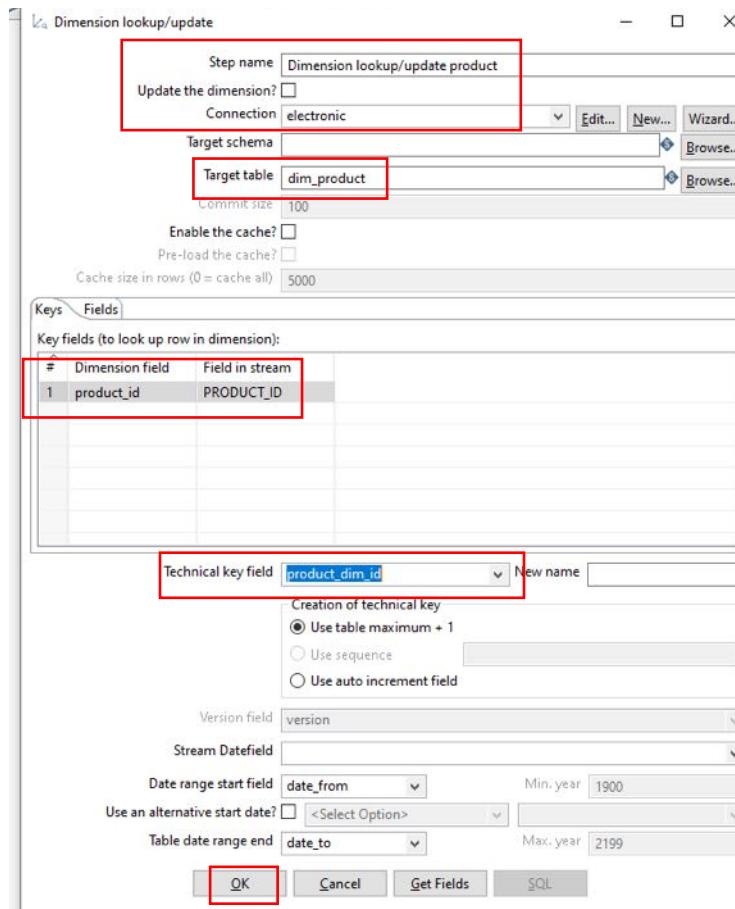
- Input Step name
- Input Connection
- Uncheck pada Update the dimension
- Uncheck pada Enable the cache
- Cocokkan Dimension Field dan Field in stream
- Pilih Technical key field dengan salesrep_dim_id
- Klik OK

c) Lakukan pencocokan dari sisi product (GROUP berdasarkan product)



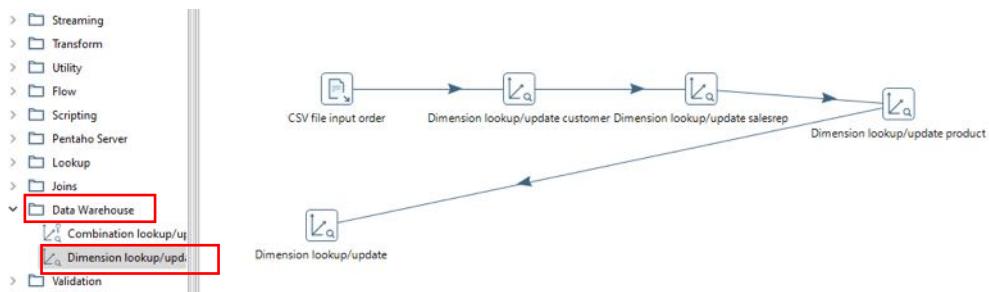
- Pilih Data Warehouse
- Pilih Dimension lookup/update
- Koneksikan CSV file input oder ke fungsi Dimension lookup/update untuk product

- Double klik pada Dimension lookup/update



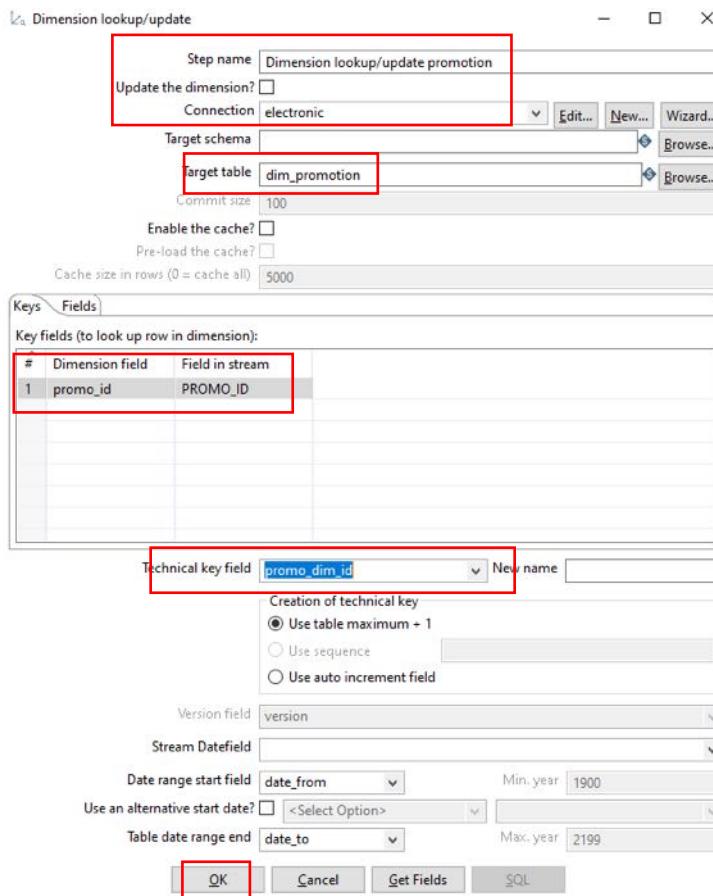
- Input Step name
- Input Connection
- Uncheck pada Update the dimension
- Uncheck pada Enable the cache
- Cocokkan Dimension Field dan Field in stream
- Pilih Technical key field dengan product_dim_id
- Klik OK

d) Lakukan pencocokan dari sisi promotion (GROUP berdasarkan promotion)



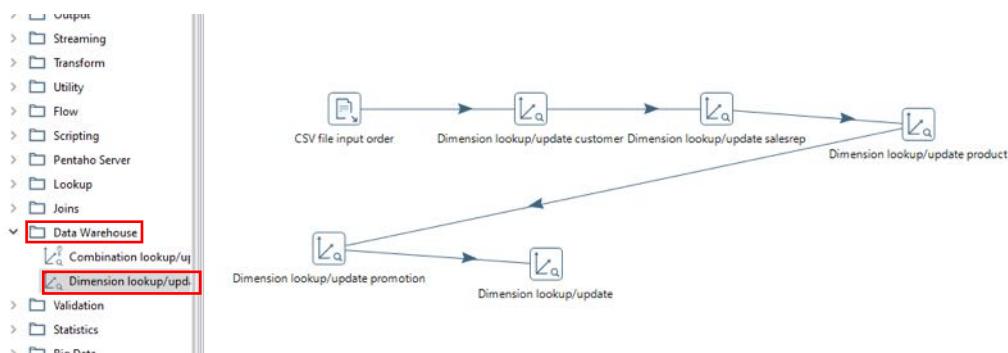
- Pilih Data Warehouse
- Pilih Dimension lookup/update
- Koneksikan CSV file input oder ke fungsi Dimension lookup/update untuk promotion

- Double klik pada Dimension lookup/update



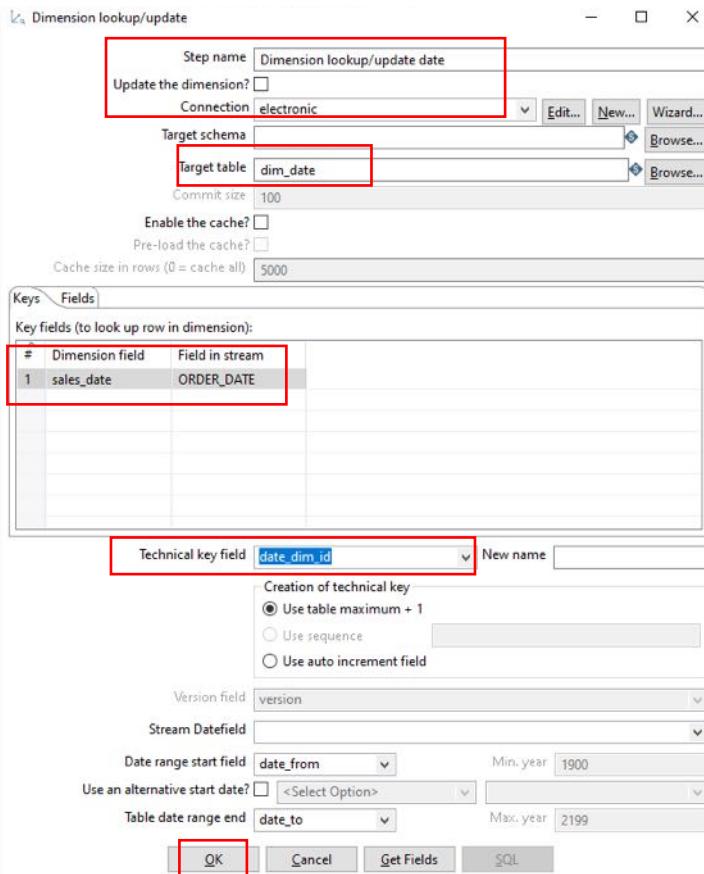
- Input Step name
- Input Connection
- Uncheck pada Update the dimension
- Uncheck pada Enable the cache
- Cocokkan Dimension Field dan Field in stream
- Pilih Technical key field dengan promo_dim_id
- Klik OK

e) Lakukan pencocokan dari sisi date (GROUP berdasarkan date)



- Pilih Data Warehouse
- Pilih Dimension lookup/update

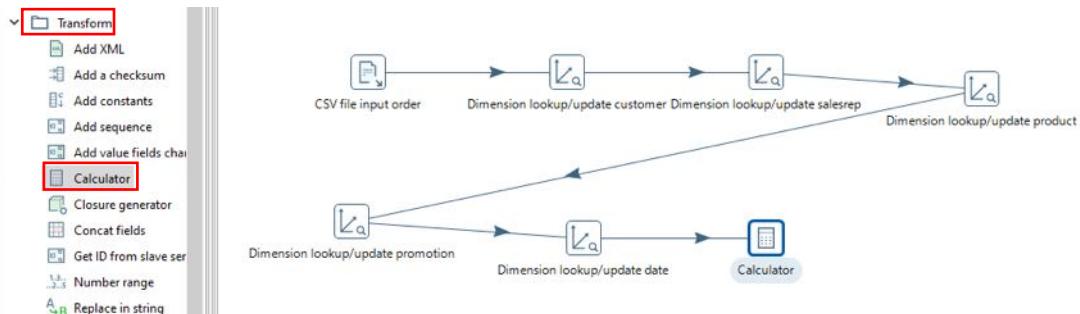
- Koneksikan CSV file input oder ke fungsi Dimension lookup/update untuk date
- Double klik pada Dimension lookup/update



- Input Step name
- Input Connection
- Uncheck pada Update the dimension
- Uncheck pada Enable the cache
- Cocokkan Dimension Field dan Field in stream
- Pilih Technical key field dengan date_dim_id
- Klik OK

f) Lakukan pencocokan dari sisi table fakta

Setelah semua data sudah dilakukan pencocokan, kita perlu lakukan perhitungan dengan membuat table fakta menggunakan fungsi calculator



- Pilih Transform
- Pilih Calculator
- Koneksikan Dimension lookup/update Date ke fungsi Calculator
- Double klik pada Calculator

Calculator

| | |
|-----------|----------|
| Step name | faktanya |
|-----------|----------|

Throw an error on non existing files

- Input Step name dengan faktanya

Dalam mencari perhitungan harga dari satuan barang perlu dibuat fungsi pada New field

Fields:

| # | New field | Calculation | Field A | Field B | Field C | Value type |
|---|-----------|-------------|------------|----------|---------|------------|
| 1 | idr_sold | A * B | UNIT_PRICE | QUANTITY | | |

- Input New field dengan idr_sold
- Input field A dengan UNIT_PRICE
- Input field B dengan QUANTITY
- Pilih Calculation dengan A*B

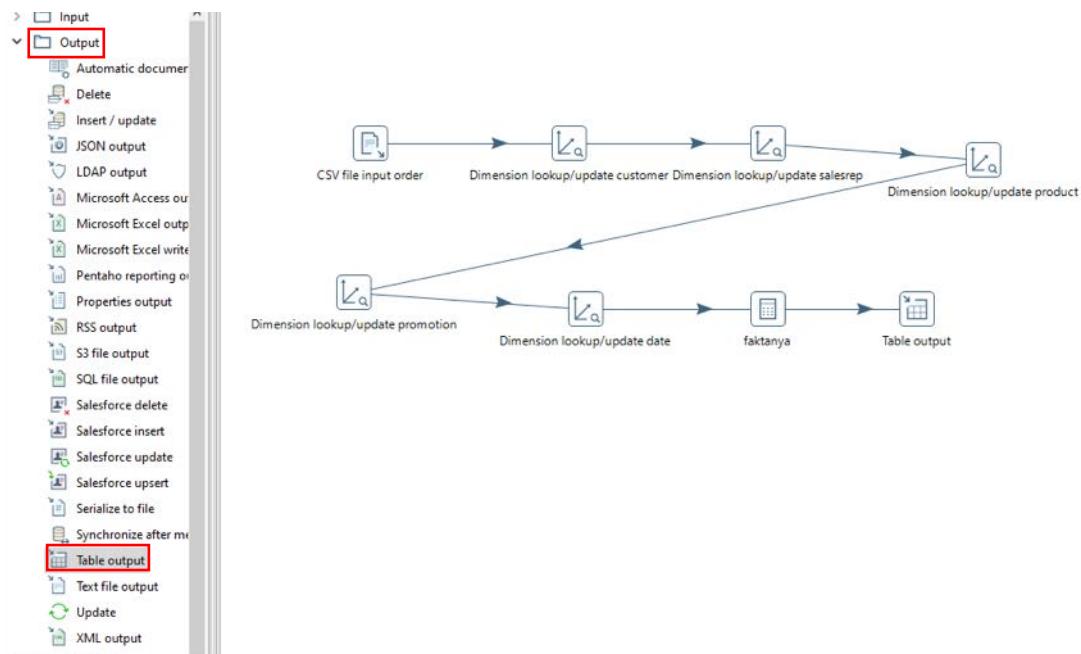
Dalam mencari perhitungan barang yang terjual perlu dibuat fungsi pada New field

Fields:

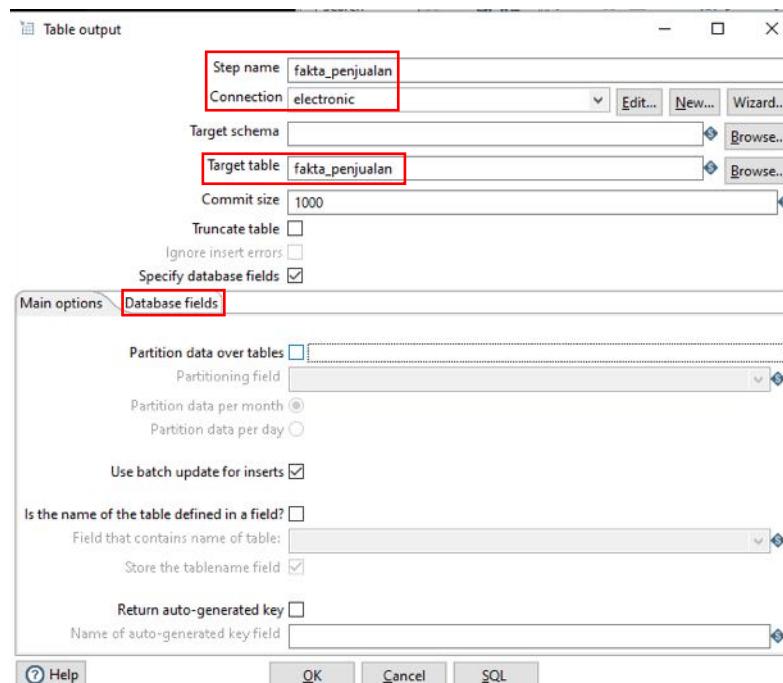
| # | New field | Calculation | Field A | Field B | Field C | Value type | Length | Precision | Remove | Conversion mask | Decimal symbol | Grouping symbol | Currency symb |
|---|-----------|--------------------------|------------|----------|---------|------------|--------|-----------|--------|-----------------|----------------|-----------------|---------------|
| 1 | idr_sold | A * B | UNIT_PRICE | QUANTITY | | Number | N | N | N | | | | |
| 2 | banyaknya | Create a copy of field A | QUANTITY | | | | | | | | | | |

-
- Input New field dengan banyaknya
 - Input Field A dengan QUANTITY
 - Pilih Calculation dengan Create a copy of field A
 - Pilih Value type dengan Number
 - Pilih OK

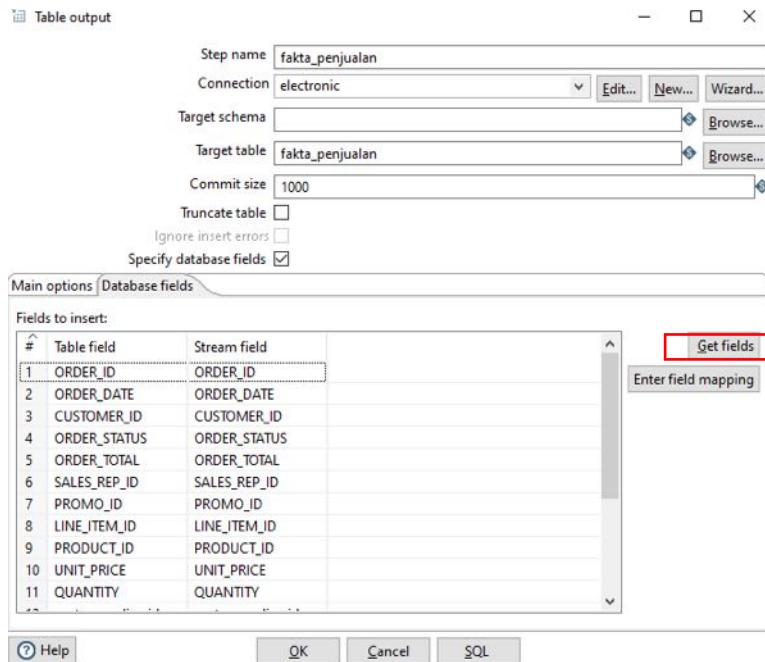
Setelah semua table dimensi dan fakta di group, keluaran data perlu ditampilkan pada table yang berbeda. Dengan demikian dibutuhkan sebuah fungsi Table output sebagai berikut.



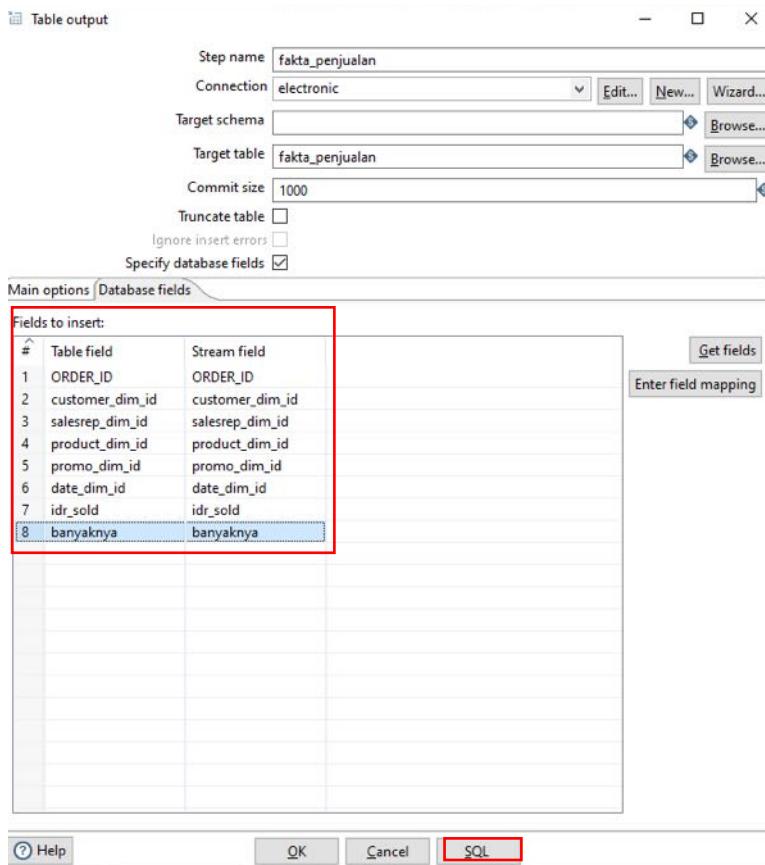
- Pilih Output
- Pilih Table output
- Koneksikan fungsi calculator faktanya ke fungsi Table output
- Double klik pada Table output



- Input Step name dengan faktा_penjualan
- Input Connection
- Input Target table dengan faktа_penjualan
- Pilih Tab Database fields



- Klik Get fields

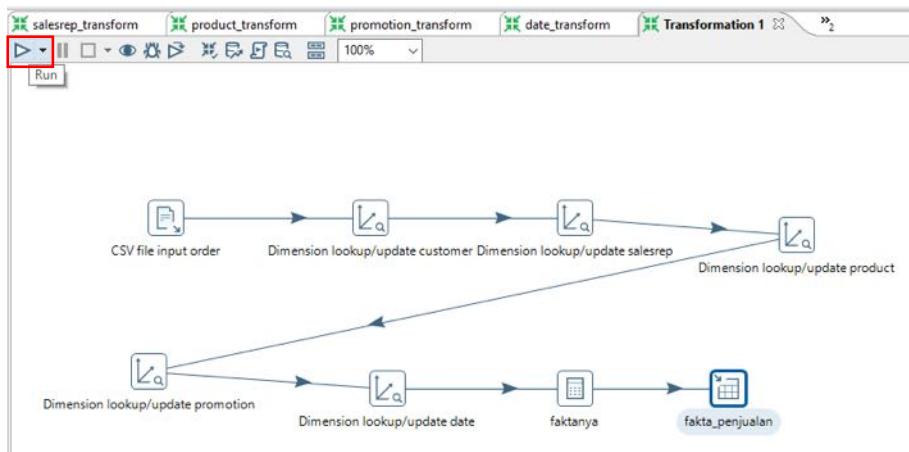


- Sesuaikan Database fields dengan data yang ingin dikeluarkan pada table faktanya
- Pilih SQL, lalu Execute

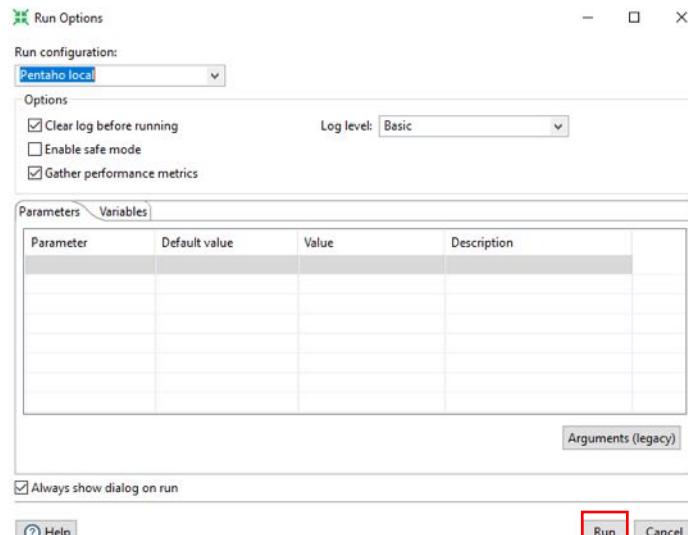
The screenshot shows the phpMyAdmin interface for the 'electronic' database. The left sidebar lists tables: 'electronic', 'dim_customer', 'dim_date', 'dim_product', 'dim_promotion', 'dim_salesrep', 'fakta_penjualan', 'information_schema', 'mysql', and 'performance_schema'. The 'fakta_penjualan' table is selected. The main area displays the table structure with 8 columns: #, Name, Type, Collation, Attributes, Null, Default, Comments, Extra, and Action. The columns are: 1 ORDER_ID (bigint(20)), 2 customer_dim_id (int(11)), 3 salesrep_dim_id (int(11)), 4 product_dim_id (int(11)), 5 promo_dim_id (int(11)), 6 date_dim_id (int(11)), 7 idr_sold (double), and 8 banyaknya (double). Each column has a 'Change' button under 'Action'.

| # | Name | Type | Collation | Attributes | Null | Default | Comments | Extra | Action |
|---|-----------------|------------|-----------|------------|------|---------|----------|-------|--|
| 1 | ORDER_ID | bigint(20) | | | Yes | NULL | | | Change Drop More |
| 2 | customer_dim_id | int(11) | | | Yes | NULL | | | Change Drop More |
| 3 | salesrep_dim_id | int(11) | | | Yes | NULL | | | Change Drop More |
| 4 | product_dim_id | int(11) | | | Yes | NULL | | | Change Drop More |
| 5 | promo_dim_id | int(11) | | | Yes | NULL | | | Change Drop More |
| 6 | date_dim_id | int(11) | | | Yes | NULL | | | Change Drop More |
| 7 | idr_sold | double | | | Yes | NULL | | | Change Drop More |
| 8 | banyaknya | double | | | Yes | NULL | | | Change Drop More |

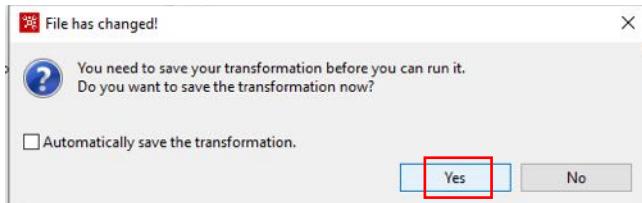
Table fakta_penjualan telah tergenerate dengan sesuai pada MySQL, kita perlu menjalankan transformasi untuk menginput data yang telah diatur dengan tahapan sebagai berikut.



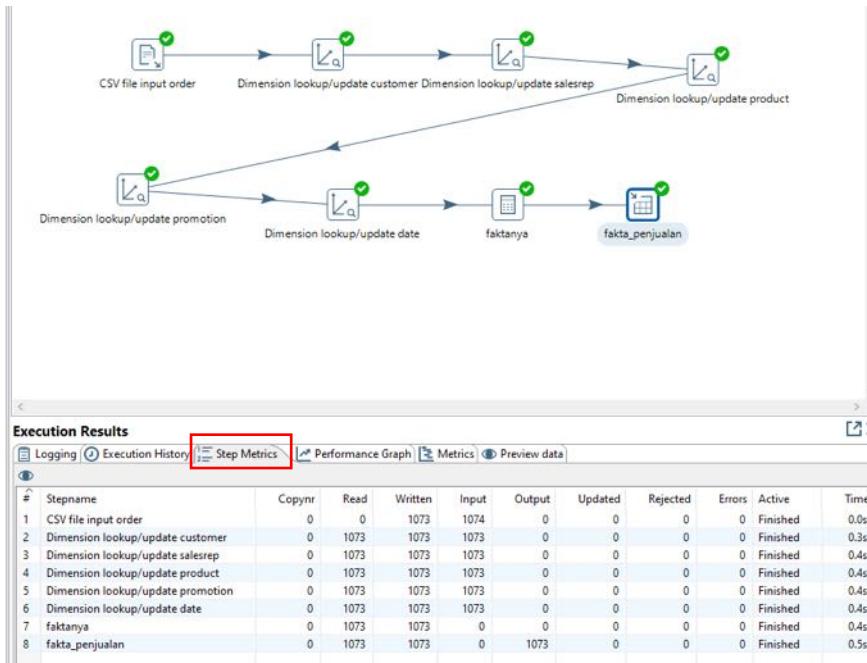
- Pilih Run



- Pilih Run



- Pilih Yes
- Simpan dengan nama fakta_transform



- Pilih Step Metrics
- Data sudah tersimpan sempurna

```
SELECT * FROM `fakta_penjualan`
```

✓ Showing rows 0 - 24 (1073 total, Query took 0.0007 seconds.)

SELECT * FROM `fakta_penjualan`

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

Number of rows: 25 Filter rows: Search this table

Options

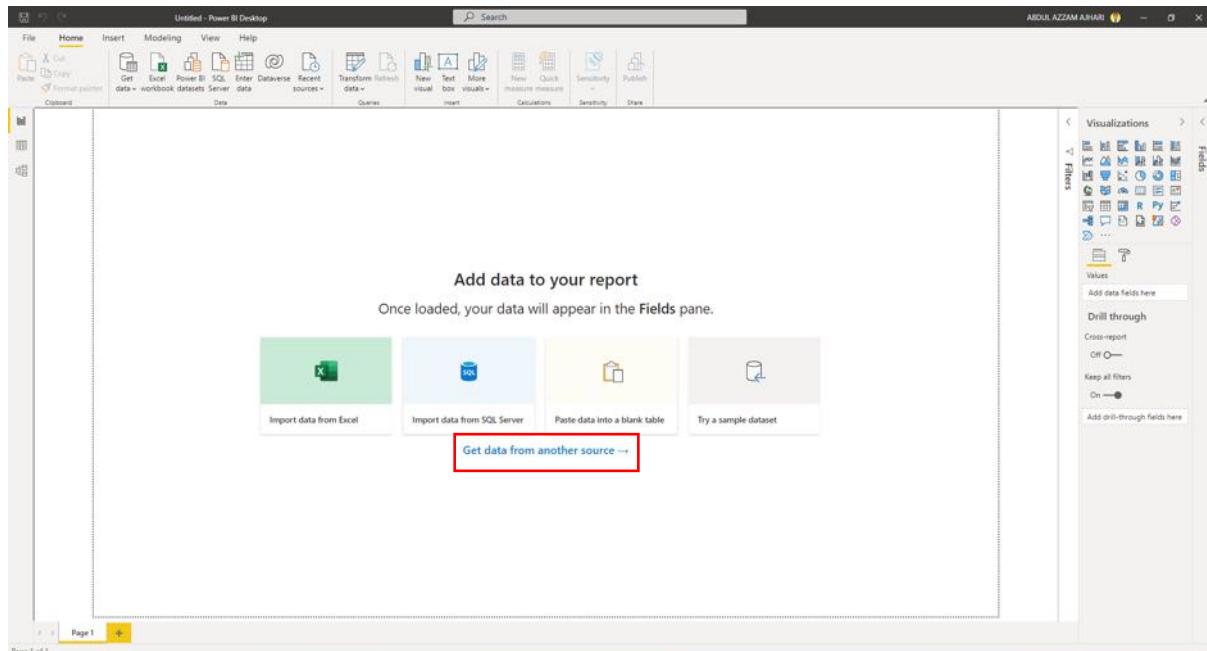
| ORDER_ID | customer_dim_id | salesrep_dim_id | product_dim_id | promo_dim_id | date_dim_id | idr_sold | banyaknya |
|----------|-----------------|-----------------|----------------|--------------|-------------|-------------------|-----------|
| 2354 | 160 | 9 | 250 | 0 | 925 | 848 | 53 |
| 2354 | 160 | 9 | 240 | 0 | 925 | 3713 | 47 |
| 2354 | 160 | 9 | 174 | 0 | 925 | 1927 | 47 |
| 2354 | 160 | 9 | 253 | 1 | 925 | 1008 | 48 |
| 2354 | 160 | 9 | 179 | 0 | 925 | 2928 | 61 |
| 2354 | 160 | 9 | 173 | 0 | 925 | 4162.4 | 43 |
| 2354 | 160 | 9 | 163 | 0 | 925 | 2368 | 64 |
| 2354 | 160 | 9 | 261 | 0 | 925 | 4697 | 77 |
| 2354 | 160 | 9 | 164 | 1 | 925 | 3468 | 68 |
| 2354 | 160 | 9 | 100 | 0 | 925 | 10164 | 70 |
| 2354 | 160 | 9 | 154 | 0 | 925 | 986 | 58 |
| 2354 | 160 | 9 | 162 | 0 | 925 | 1830 | 61 |
| 2354 | 160 | 9 | 105 | 0 | 925 | 8157.599999999999 | 72 |
| 2355 | 160 | 0 | 244 | 3 | 25 | 16337.2 | 188 |
| 2355 | 160 | 0 | 176 | 3 | 25 | 10545 | 185 |
| 2355 | 160 | 0 | 66 | 1 | 25 | 46226.4 | 204 |
| 2355 | 160 | 0 | 180 | 3 | 25 | 9200 | 200 |
| 2355 | 160 | 0 | 113 | 3 | 25 | 211.2000000000002 | 192 |
| 2355 | 160 | 0 | 146 | 3 | 25 | 4975 | 199 |
| 2355 | 160 | 0 | 258 | 3 | 25 | 3572 | 188 |
| 2355 | 160 | 0 | 251 | 3 | 25 | 3230 | 190 |
| 2355 | 160 | 0 | 114 | 3 | 25 | 216.7000000000002 | 197 |
| 2356 | 161 | 0 | 239 | 1 | 755 | 3168 | 44 |
| 2356 | 161 | 0 | 172 | 3 | 755 | 3920 | 40 |
| 2356 | 161 | 0 | 61 | 3 | 755 | 5049 | 34 |

Number of rows: 25 Filter rows: Search this table

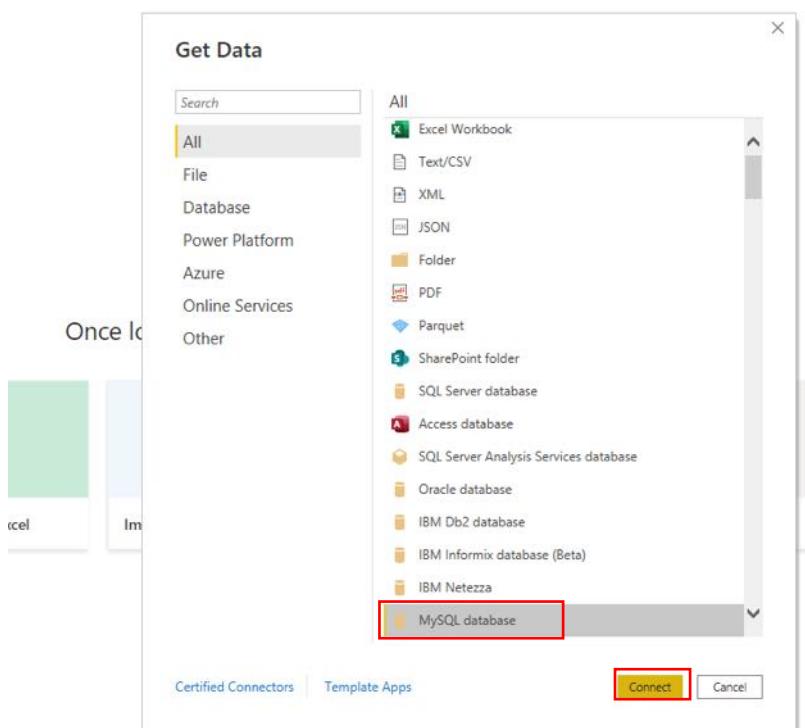
[10%] Jelaskan bagaimana membuat visualisasi dari tabel fakta yang dihasilkan dari soal No.b. Jelaskan dalam video Anda mengapa bentuk visualisasi tertentu anda pilih, dan apa makna dari visualisasi yang muncul.

VISUALISASI

Visualisasi menggunakan Microsoft Power BI Desktop yang dihubungkan pada Database MySQL



Pilih Get data from another source



Pilih MySQL database, lalu klik Connect



Input server dengan nama localhost dan database adalah electronic, lalu klik OK

The screenshot shows the Power BI Navigator interface. On the left, under 'Display Options', there's a tree view showing 'localhost: electronic [6]' with several tables selected. On the right, a preview of the 'electronic.fakta_penjualan' table is displayed with columns: ORDER_ID, customer_dim_id, salesrep_dim_id, product_dim_id, and promo_dim_id. The table contains numerous rows of data. At the bottom, there are buttons for 'Select Related Tables', 'Load' (highlighted with a red box), 'Transform Data', and 'Cancel'.

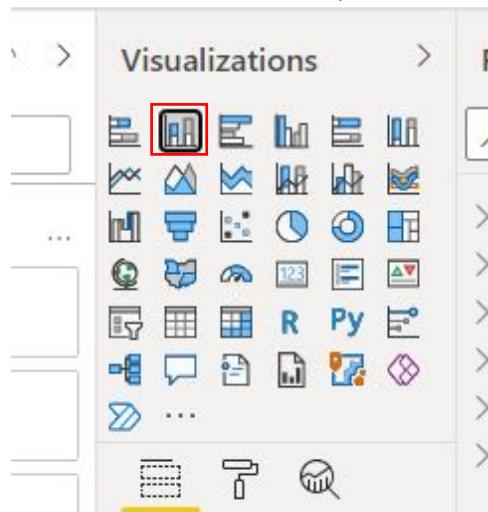
Pilih seluruh table dimensi dan fakta, kemudian klik Load

The screenshot shows the Power BI Fields pane. On the left, there are sections for 'Visualizations' and 'Filters'. On the right, the 'Fields' section is highlighted with a red box, showing a list of tables: 'electronic.dim_customer', 'electronic.dim_date', 'electronic.dim_product', 'electronic.dim_promotion...', 'electronic.dim_salesrep', and 'electronic.fakta_penjualan...'. There is also a 'Search' bar at the top of the Fields section.

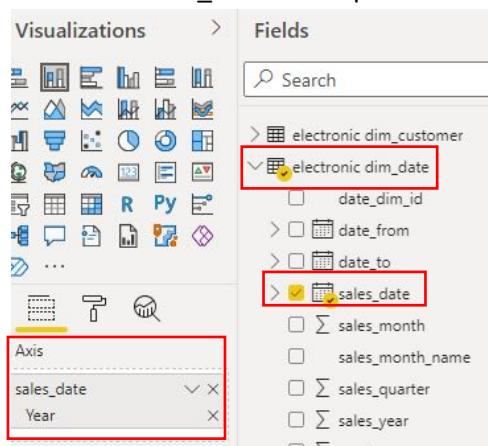
Apabila table berhasil di Load akan muncul di pojok kanan pada menu Fields

- Bisnis Questions pertama: Berapa Pendapatan Toko per Tahun dari 2006 – 2008

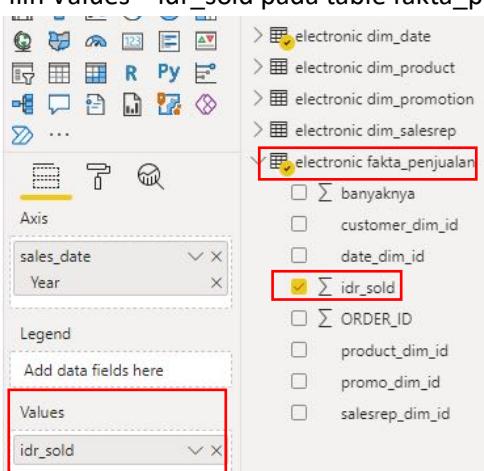
1. Pilih Stacked Column Chart pada Visualizations



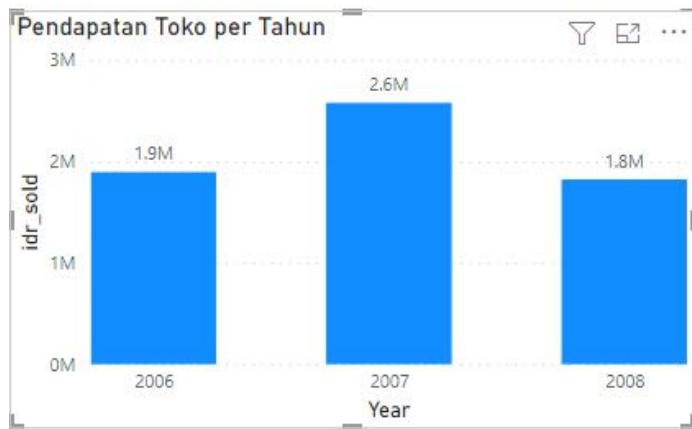
2. Atur Axis - sales_date – Year pada table dim_date



3. Pilih Values – idr_sold pada table fakta_penjualan



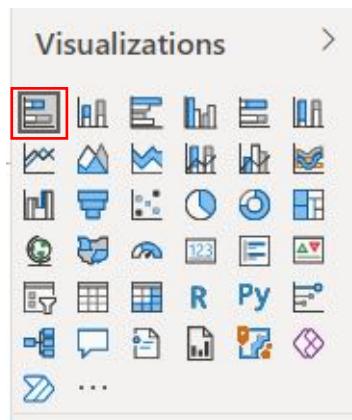
4. Visualisasi yang terjadi



Pendapatan toko pada tahun 2006 berjumlah 1,9 juta dollar atau sekitar Rp. 1,9 Miliar. Pendapatan toko pada tahun 2007 berjumlah 2,6 juta dollar atau sekitar Rp. 2,6 Miliar. Pendapatan toko pada tahun 2008 berjumlah 1,8 juta dollar atau sekitar Rp. 1,8 Miliar. Pendapatan tertinggi berada pada tahun 2007 dan mengalami penurunan pada tahun 2008. Column Chart dipilih karena kita akan membandingkan nilai pendapatan toko setiap tahunnya.

- **Bisnis Questions kedua: Berapa Jumlah barang yang terjual dengan harga diatas 20K dollar atau sekitar Rp. 20 juta dengan total penjualan lebih dari 20 Pcs pada setiap Promo yang ada.**

1. Pilih Stacked Bar Chart



2. Atur Axis dengan PRODUCT_NAME dari dim_product

| Axis | Values | Legend | Small multiples |
|--------------|----------|------------|-----------------|
| PRODUCT_NAME | idr_sold | PROMO_NAME | PRODUCT_NAME |

Legend: PROMO_NAME

Values: idr_sold

Small multiples: PRODUCT_NAME

3. Atur Values dengan idr_sold dari table fakta_penjualan

The screenshot shows the Tableau Data Source pane. In the 'Values' section, the field 'idr_sold' is selected and highlighted with a red box.

4. Atur Tooltips dengan banyaknya dari table fakta_penjualan

The screenshot shows the Tableau Data Source pane. In the 'ToolTips' section, the field 'banyaknya' is selected and highlighted with a red box.

5. Atur Legend dengan PROMO_NAME dari dim_promotion

The screenshot shows the Tableau Data Source pane. In the 'Legend' section, the field 'PROMO_NAME' is selected and highlighted with a red box. The legend items listed on the right include 'electronic dim_customer', 'electronic dim_date', 'electronic dim_product', 'electronic dim_prom...', 'date_from', 'date_to', 'promo_dim_id', 'sum(promo_id)', 'PROMO_NAME', 'sum(version)', 'electronic dim_salesrep', and 'electronic fakta_penjual...'.

6. Visualisasi yang terjadi



Terjadi penjualan terbesar pada promo new year yaitu produk Desk-W/48 dengan penjualan lebih dari 20 buah dengan total harga adalah 255.267 dollar atau sekitar Rp. 2 miliar 552 juta-an.

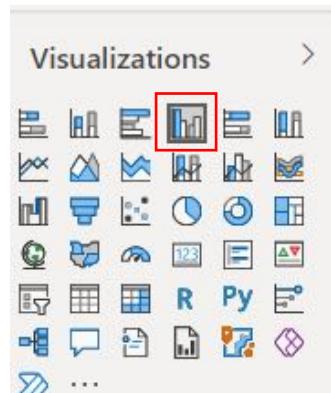
Penjualan pada promo summer sale terbesar yaitu produk LaserPro 600/6/8W dengan total harga adalah 111.355 dollar atau sekitar Rp. 1 miliar 355 juta-an.

Terakhir, penjualan pada promo christmast terbesar yaitu produk KB 101/ES dengan total harga adalah 24.840 dollar atau sekitar Rp. 248 juta.

Stacked bar chart dipilih untuk membandingkan nilai dimensi penjualan dari dimensi produk dan promosi.

- Bisnis Questions ketiga: Siapakah Sales yang paling produktif dengan penjualan terbanyak.**

1. Pilih Clustered Column Chart



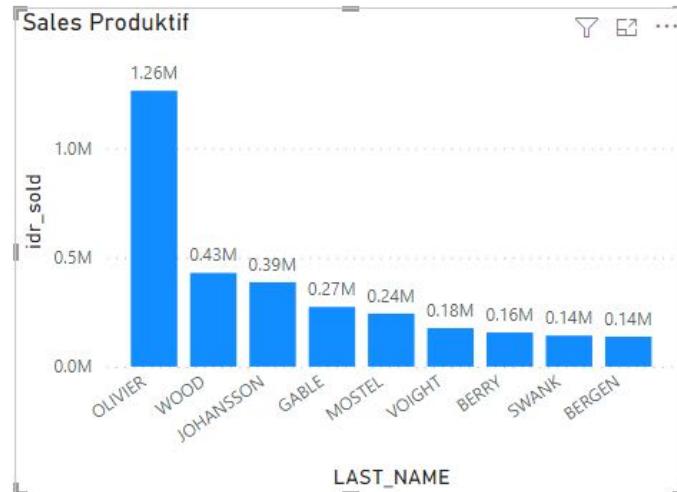
2. Atur Axis dengan LAST_NAME pada dim_salesrep

| Axis | Values |
|-----------|----------|
| LAST_NAME | idr_sold |

3. Atur Values dengan idr_sold pada tabel fakta_penjualan

The screenshot shows the Power BI Data View interface. In the 'Values' section, the field 'idr_sold' is highlighted with a red box. To its right, under the 'electronic fakta_penju...' node, the field 'sum(idr_sold)' is also highlighted with a red box.

4. Visualisasi yang terjadi



Olivier merupakan Sales produktif dengan total penjualan yang dilakukan adalah 1,26 juta dollar atau sekitar Rp. 1 miliar 260 juta.

Clustered column chart dipilih digunakan untuk membandingkan nilai dari grup kategori salesrep dengan total penjualannya.



Total idr_sold was higher for new year (300,452.90) than summer sale (249,833.20).

Desk - W/48 in PROMO_NAME made up 44.38% of idr_sold.

Average idr_sold was higher for new year (100,150.97) than summer sale (62,458.30).

