

2241720145

Muhammad Baihaqi Aulia Asy’ari

Daftar Isi

[Jobsheet 1: Data Warehouse Sederhana dan Pengenalan PDI Spoon 3](#_Toc162436925)

# Jobsheet 1: Data Warehouse Sederhana dan Pengenalan PDI Spoon

**Tujuan Praktikum**

Setelah melakukan praktikum ini, mahasiswa diharapakan dapat lebih mengenal apa itu data warehouse berdasarkan uji coba membuat data warehouse sederhana

**Studi Kasus**

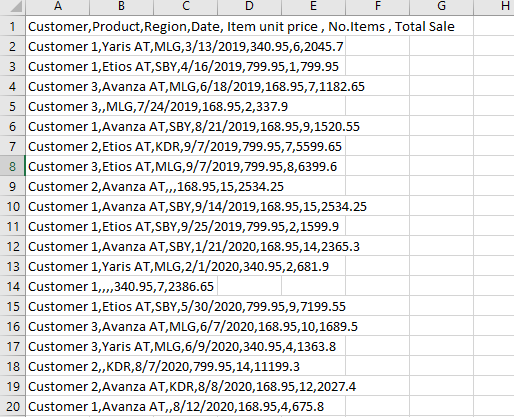
Pak Amir adalah Manager Sales Astro Mobil, yang bergerak di bidang distributor mobil wilayah Jawa Timur. Pak Amir meminta data penjualan dari beberapa cabang kepada supervisor. Data tersebut akan digunakan untuk membuat suatu Dashboard pengambilan keputusan. Dikarenakan proses permintaan ini dilakukan secara berulang setiap hari setelah jam kantor maka, Pak Amir akan membuat proses untuk menarik data yang ada pada file milik supervisor tersebut secara otomatis. Namun, kadang data tersebut tidak lengkap. Sehingga Pak Amir akan mengambil data yang lengkap saja dan mengembalikan data yang tidak lengkap.

Dari studi kasus tersebut maka, akan dilakukan:

1. Cek dan Analisa data penjualan.
2. Import data dari file (Extract)
3. Identifikasi data yang tidak lengkap (missing data) dan meletakkan pada file yang berbeda (Transform)
4. Memindahkan data yang sudah lengkap ke file dashboard. (Load)
5. Mengumpulkan data yang belum lengkap untuk dikembalikan.

\*Data penjualan tersebut dapat diakses dan di download melalui link berikut:

https://github.com/dik4rizky/datasources/blob/7e3f70f516a1b107635141d09862a8c30fbb6812/CarDistributionSales.csv



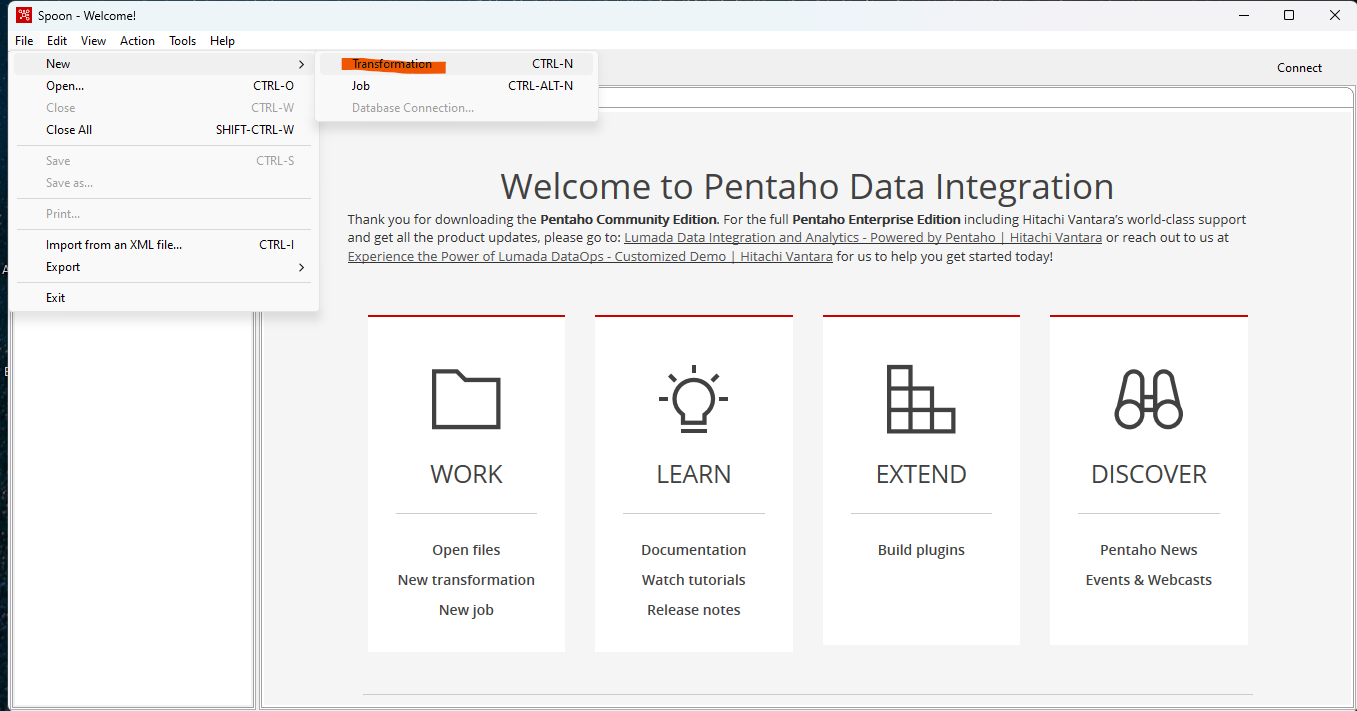
Gambar Data Penjualan Mobil

### TUGAS 1

1. Analisa lah data tersebut!
2. Berapa jumlah kolom pada data tersebut?
3. Apa arti atau isi data dari setiap kolom yang ada?
4. Adakah data yang memiliki nilai null / data yang tidak lengkap?
5. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?

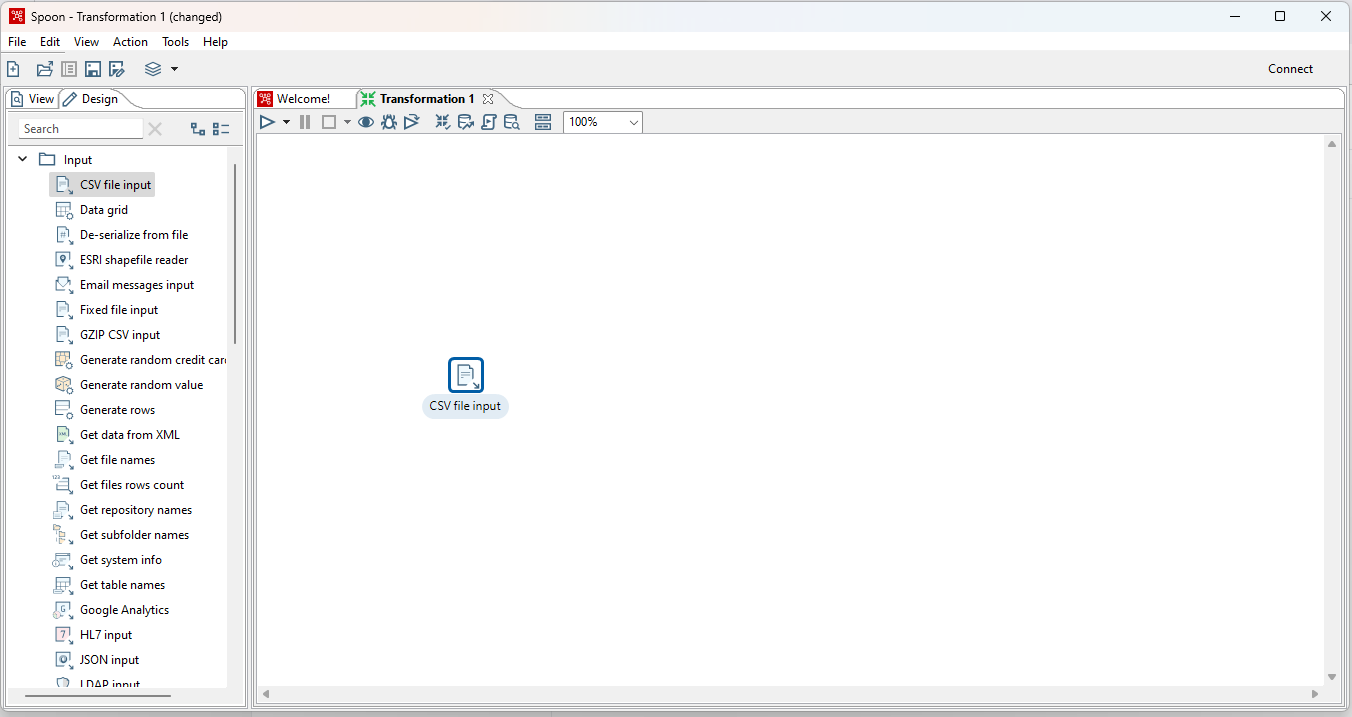
## A. Pengambilan Data (Extract)

1. Buka lembar kerja Transformation melalui toolbar File.



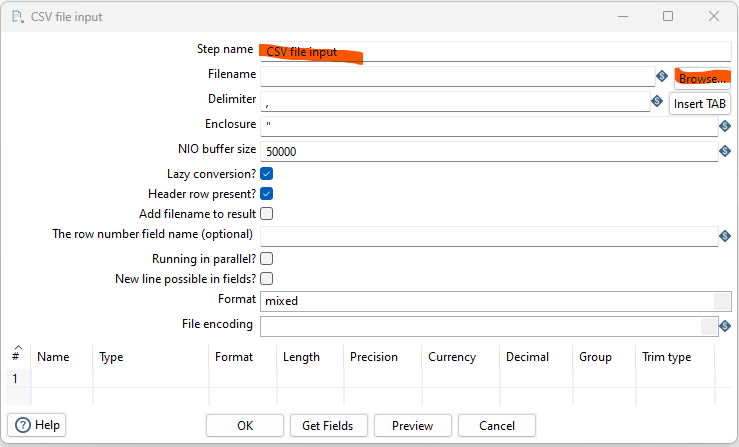
Gambar memulai membuka lembar kerja Transformation

1. Cari lah objek CSV file input pada Design Area. Drag and drop objek tersebut menuju Work Area.

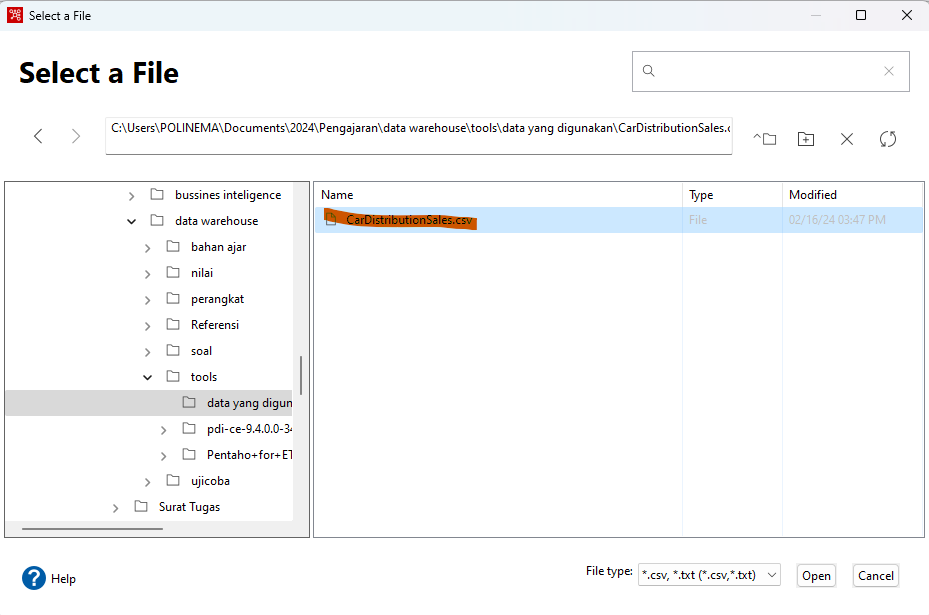


Gambar objek CSV file input pada Work Area

1. Double-click pada objek CSV file input hingga muncul jendela konfigurasinya.
2. Ubah nama Step name menjadi Input Data CSV, (penamaan digunakan untuk identifikasi saja, selanjutnya penamaan boleh menyesuaikan).
3. Pilih lokasi file CarDistributionSales.csv melalui Browse pada filename.

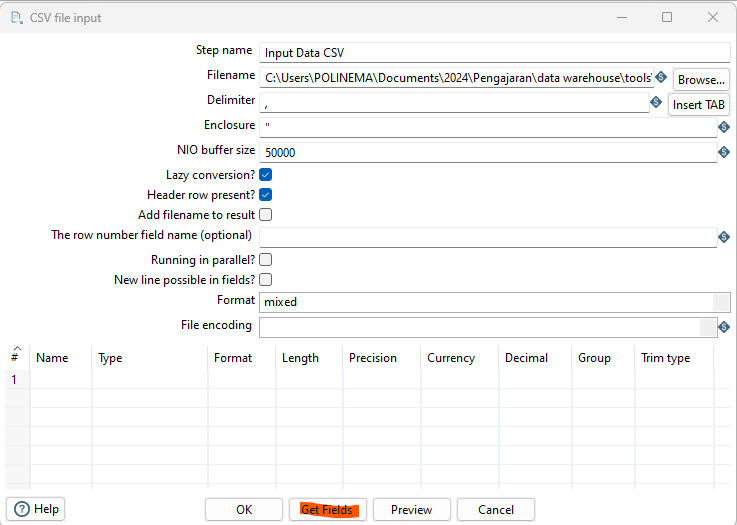


Jendela konfigurasi CSV file input

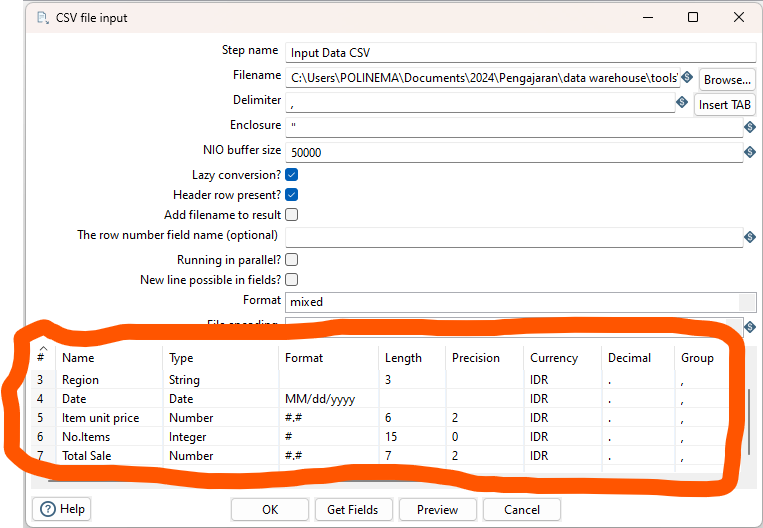


Gambar lokasi file CarDistributionSales.csv

1. Jika data telah dipilih pastikan pada bagian Delimiter sesuai dengan kondisi pemisah data pada file yang dipilih. (Pada kesempatan ini adalah comma " , " ).
2. Tekan Get Fields untuk mengambil judul pada setiap kolom, pastikan nama-nama kolom telah sesuai

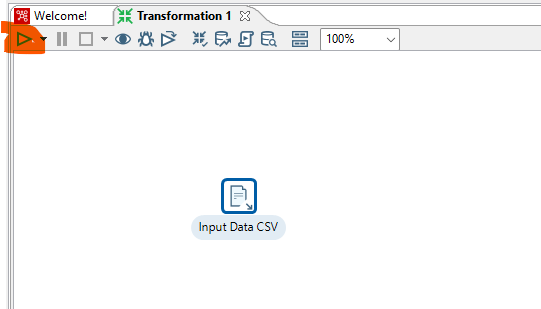


Gambar jendela konfigurasi setelah file dipilih

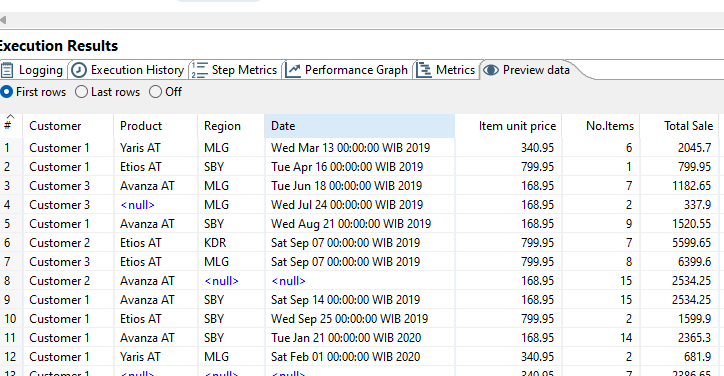


Gambar jendela konfigurasi setelah Get Fields ditekan

1. Tutup jendela konfigurasi tersebut.
2. Tekan tombol "Run" pada pojok kanan atas Area Kerja, makan akan muncul Execution Result Area. Pilih tab Prieview Data pada Execution Result Area. Jika proses yang dilakukan benar maka data akan muncul sesuai dengan aslinya.



Gambar jendela transformation area



Gambar hasil dari proses

Jika langkah-langkah diatas telah berhasil, maka proses extract telah berhasil dilakukan.

### TUGAS 2

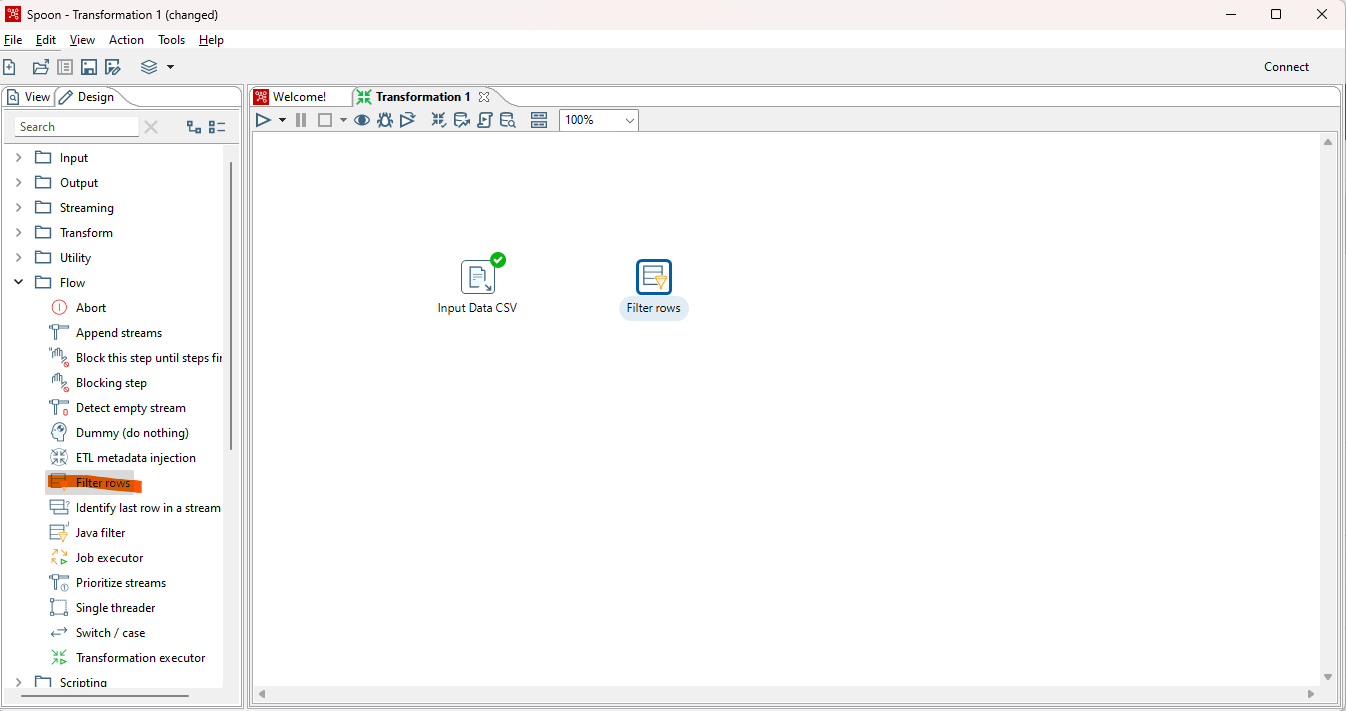
1. Apakah data hasil eksekusi sesuai dengan data aslinya?
2. PDI Spoon melakukan proses extract, perhatikan pada Tab Logging di Execution Results Area, langkah-langkah apa sajakah yang dilakukan PDI Spoon untuk melakukan extract data?
3. Perhatikan gambar dibawah ini! Apa yang dimaksud dengan I,O,R,W,U,E ?



1. Berapa waktu yang dibutuhkan untuk melakukan eksekusi Extract data ini? Bandingkan dengan 3 teman yang lain, identifikasi pula spesifikasi perangkat yang digunakan (Processor, RAM, Storage, VGA, CPU ). Bandingkan dalam bentuk tabel.

## B. Filter Data (Transform) dan Pengemasan data (Load)

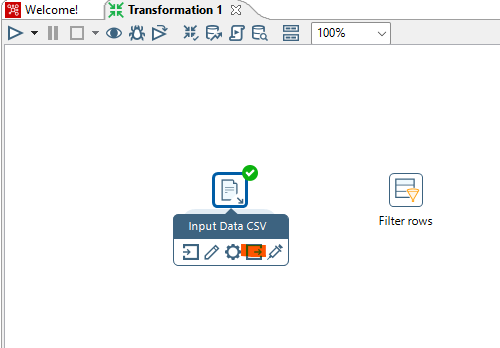
1. Carilah objek Filter rows pada Design Area. Drag and drop objek tersebut pada Work Area.



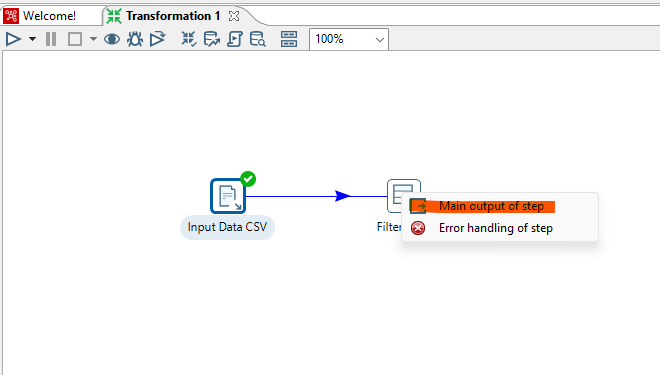
Gambar objek filter rows pada area kerja

1. Hubungkan output pada Input Data CSV menuju Filter rows. Pilih koneksi Main output of step saat menghubungkan kedua objek tersebut.

Pada tahap ini Input Data CSV dihubungkan oleh konektor pada Filter rows memiliki makna bahwa setelah proses Input Data CSV dilakukan maka proses selanjutnya adalah Filter rows.

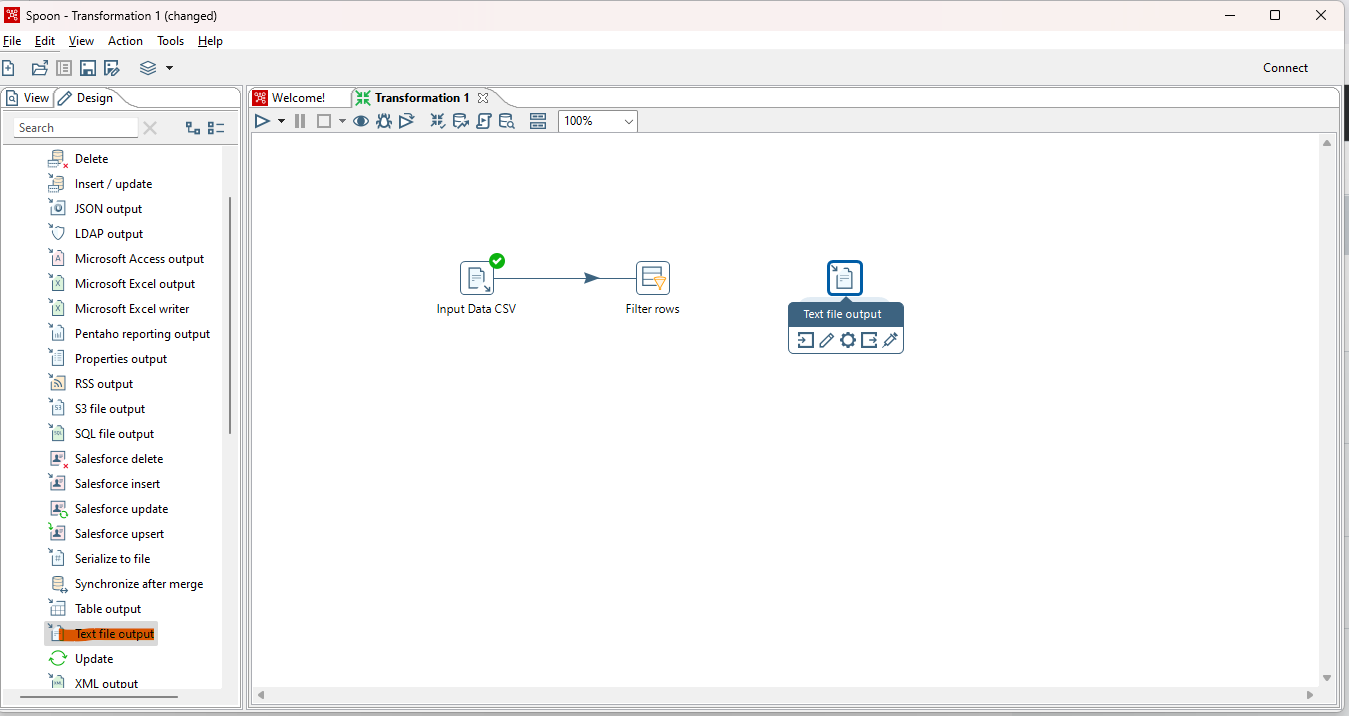


Gambar output konektor pada Input data CSV



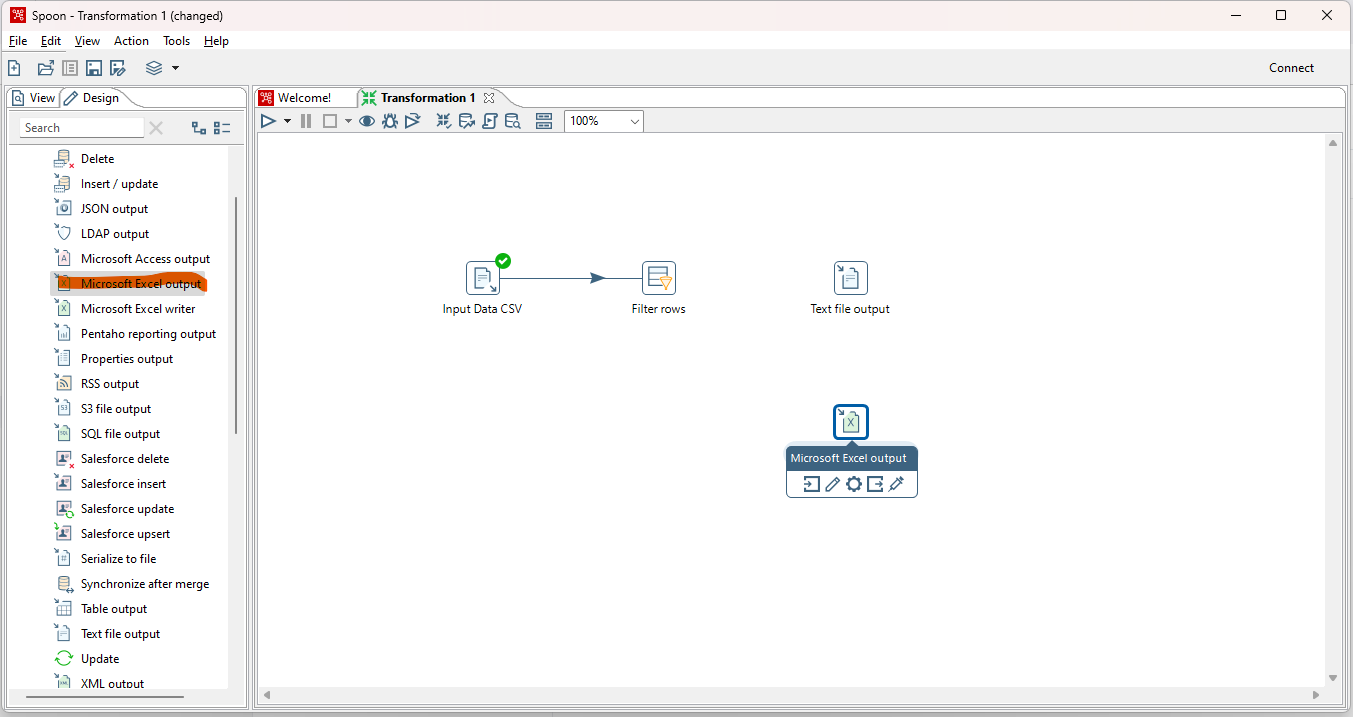
Gambar Input Data CSV terhubung dengan Filter rows

1. Carilah objek Text file output pada Design Area. Drag and drop objek tersebut pada Work Area.



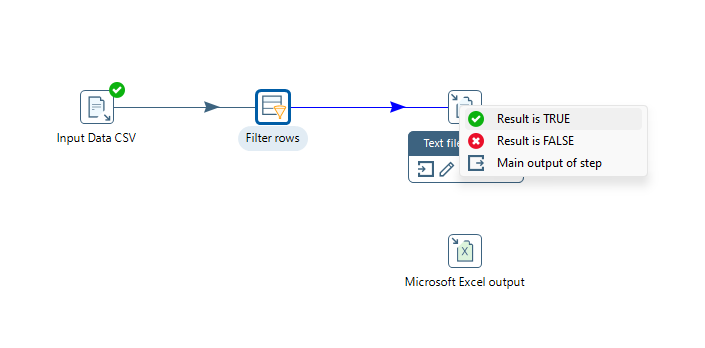
Gambar text file output pada work area

1. Carilah objek Microsoft excel output pada Design Area. Drag and drop objek tersebut pada Work Area.



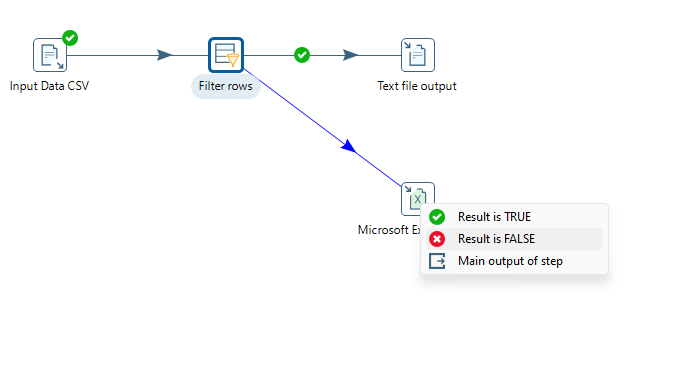
Gambar Microsofft excel output pada work area

1. Hubungkan filter rows dengan text file output menggunakan konektor output dari filter rows. Pilih Result is true sebagi jenis konektornya.



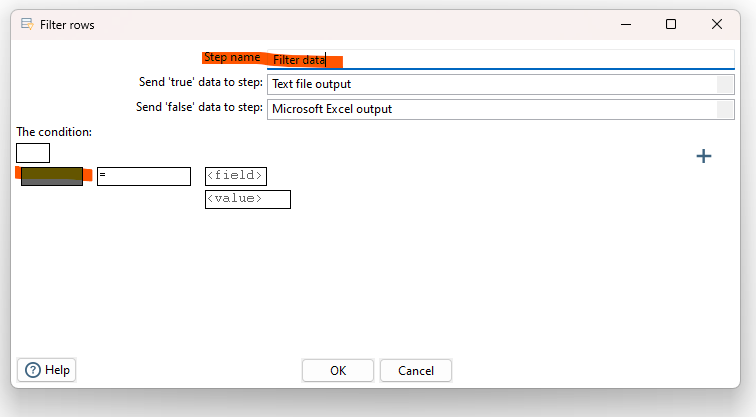
Gambar konktor text file output

1. Hubungkan filter rows dengan microsoft excel output menggunakan konektor output dari filter rows. Pilih Result is false sebagi jenis konektornya.



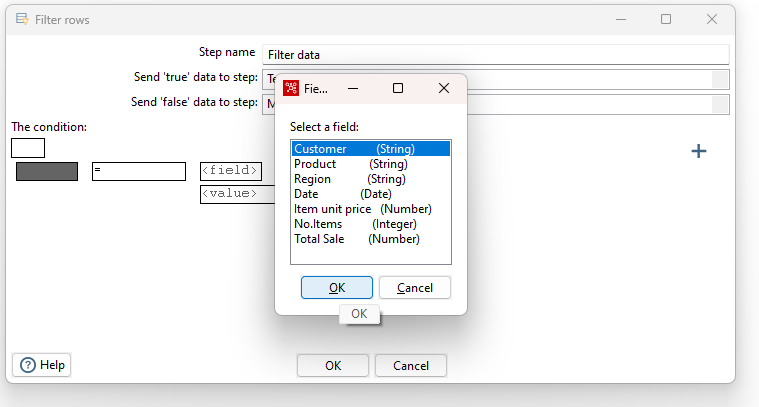
Gambar konektor microsoft excel output

1. Double click pada filter rows hingga muncul jendela konfigurasi. ubah step name menjadi Filter data.
2. Cek bagian send 'true' data to step menjadi Text file output dan cek bagian Send 'false' data to step menjadi Microsoft Excel output.
3. Klik <Field> pada bagian condition



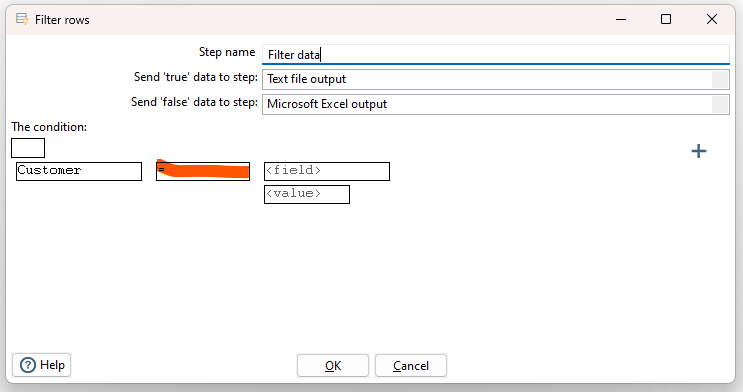
Gambar jendela konfigurasi filter rows

1. Pilih field yang diinginkan untuk digunakan sebagai kondisi. Pada kesempatan ini pilih Customer, kemudian klik OK.



Gambara jendela pilihan field

1. Pada bagian value pilih IS NOT NULL kemudian tekan OK , hal ini dilakukan karena data yang dipilih adalah data yang tidak Null



Gambar jendela filter rows bagian value

Gambar jendela filter rows pemilihan value

Tekan tombol + untuk menambah kondisi.

Pilih statement yang masih null untuk menambah kondisi.

Ulangi proses nomor 9,10,11 dengan mengganti semua field yang ada. hingga statement menjadi seperti pada gambar dibawah ini.

Gunakan operator AND untuk menghubung

Gambar hasil statement

Gambar untuk menambah statement

Gambar memilih operator menjadi AND

Gambar mengubah statement yang null

Gambar pemilihan field untuk statement

Jika semua field sudah masuk dalam kondisi maka tekan OK.

Double-click pada text file output, pilih lokasi file untuk menyimpan hasil output file dengan menekan tombol browse. beri nama outputSalesCarDistribution.

Ubah extension menjadi csv.

Pada tab Fields, tekan tombol Get fields untuk mengambil field data.

Tekan ok untuk kembali ke work area

Jendela konfigurasi text file output

Jendela konigurasi text file input ekstensi csv

Tab get fields

Hal yang sama dilakukan pada konfigurasi Microsoft Excel output, Double-click pada objek Microsoft Excel output hingga muncul jendela konfigurasi.

Pilih lokasi file untuk menyimpan file output dengan menekan tombol browse. Beri nama failedSalesCarDistribution. untuk extension tetap dengan extension xls.

Pada tab fields tekan tombol Get fields untuk mengambil field data.

Tekoan ok untuk kembali ke work area.

Gambar jendela konfigurasi Microsoft Excel output

Gambar tab field microsoft excel output

Tekan tombul Run pada pojok kiri atas Work Area. jika berhasil maka akan muncul tanda centang hijau di setiap proses.

Gambar proses berhasil dijalankan

Gambar failedSalesCarDistribution

Gambar outputSalesCarDistribution

Langkah-langkah diatas merupakan proses melakukan filter data, dimana data yang lengkap akan diletakkan pada text file csv dan data yang belum lengkap akan diletakkan pada ms excel.

### TUGAS 3

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?
2. Jelaskan apa yang terjadi pada proses Filter rows!
3. Berapa waktu yang dibutuhkan untuk melakukan eksekusi Extract Transfer Load data ini? Bandingkan dengan 3 teman yang lain, identifikasi pula spesifikasi perangkat yang digunakan (Processor, RAM, Storage, VGA, CPU ). bandingkan dalam bentuk tabel