

Nama : Arimbi Putri Hapsari

Kelas/No.Absen : SIB 2A/05

Laporan Praktikum Data Warehouse

Jobsheet 1: Data Warehouse Sederhana dan Pengenalan PDI Spoon

Tujuan Praktikum

Setelah melakukan praktikum ini, mahasiswa diharapkan 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.

	A	B	C	D	E	F	G	H
1	Customer,Product,Region,Date, Item unit price , No.Items , Total Sale							
2	Customer 1,	Yaris AT,	MLG,3/13/2019,	340.95,	6,	2045.7		
3	Customer 1,	Etiot AS,	SBY,4/16/2019,	799.95,	1,	799.95		
4	Customer 3,	Avanza AT,	MLG,6/18/2019,	168.95,	7,	1182.65		
5	Customer 3,,	MLG,7/24/2019,	168.95,	2,	337.9			
6	Customer 1,	Avanza AT,	SBY,8/21/2019,	168.95,	9,	1520.55		
7	Customer 2,	Etiot AT,	KDR,9/7/2019,	799.95,	7,	5599.65		
8	Customer 3,	Etiot AT,	MLG,9/7/2019,	799.95,	8,	6399.6		
9	Customer 2,	Avanza AT,,,	168.95,	15,	2534.25			
10	Customer 1,	Avanza AT,	SBY,9/14/2019,	168.95,	15,	2534.25		
11	Customer 1,	Etiot AT,	SBY,9/25/2019,	799.95,	2,	1599.9		
12	Customer 1,	Avanza AT,	SBY,1/21/2020,	168.95,	14,	2365.3		
13	Customer 1,	Yaris AT,	MLG,2/1/2020,	340.95,	2,	681.9		
14	Customer 1,,,,	340.95,	7,	2386.65				
15	Customer 1,	Etiot AS,	SBY,5/30/2020,	799.95,	9,	7199.55		
16	Customer 3,	Avanza AT,	MLG,6/7/2020,	168.95,	10,	1689.5		
17	Customer 3,	Yaris AT,	MLG,6/9/2020,	340.95,	4,	1363.8		
18	Customer 2,,	KDR,8/7/2020,	799.95,	14,	11199.3			
19	Customer 2,	Avanza AT,	KDR,8/8/2020,	168.95,	12,	2027.4		
20	Customer 1,	Avanza AT,,	8/12/2020,	168.95,	4,	675.8		

TUGAS 1

Analisa lah data tersebut!

1. Berapa jumlah kolom pada data tersebut?

Jawab: ada 7 kolom, yaitu:

- 1.Customer
- 2.Product
- 3.Region
- 4.Date
- 5.Item unit Price
- 6.No.items
- 7.Total Sale

2. Apa arti atau isi data dari setiap kolom yang ada?

Jawab:

- **Customer** : Nama atau ID pelanggan
- **Product** : Nama produk(mobil yang dijual)
- **Region** : Wilayah/cabang tempat transaksi dilakukan
- **Date** : Tanggal transaksi terjadi
- **Item unit price** : Harga satuan dari mobil yang dibeli
- **No. Items** : Jumlah unit mobil yang dibeli dalam transaksi tersebut
- **Total Sale** : Total penjualan dalam transaksi(Item unit price x No.items)

3. Adakah data yang memiliki nilai null / data yang tidak lengkap?

Jawab: Sepertinya ada beberapa data yang tidak lengkap(missing values), Contohnya

1. **Baris 8** (Customer 3 , Etios AT, MLG, 9/7/2019, 799.95, 8, 6399.6) → Kolom "Customer" terlihat tidak lengkap.
2. **Baris 11** (Customer 1, Avanza AT, SBY, 1/21/2020, 168.95, 14, 2365.3) → Customer terlihat tidak lengkap.
3. **Baris 13** (Customer 1, Yaris AT, MLG, 2/1/2020, 340.95, 2, 681.9) → Customer terlihat tidak lengkap.
4. **Baris 14** (Customer __, 340.95, 7, 2386.65) → Informasi Customer, Product, dan Region tidak ada.

Terdapat kemungkinan adanya kesalahan dalam pemisahan data karena format VSC, sehingga beberapa nilai tidak terdeteksi dengan benar, mungkin perlu dicek ulang dan diperbaiki.

4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?

Jawab: Ada kemungkinan beberapa data memiliki tipe yang berbeda dengan data lainnya dalam kolom yang sama. Beberapa perbedaan tipe data adalah:

1. Kolom "Customer"

Seharusnya berisi teks (Customer 1, Customer 2, dll.), tetapi pada baris 14 terlihat ada nilai yang hilang atau tidak sesuai format.

2. Kolom "Date"

Seharusnya semua berformat tanggal (MM/DD/YYYY), tetapi perlu dipastikan apakah ada yang tersimpan dalam format teks.

3. Kolom "Item unit price", "No.Items", dan "Total Sale"

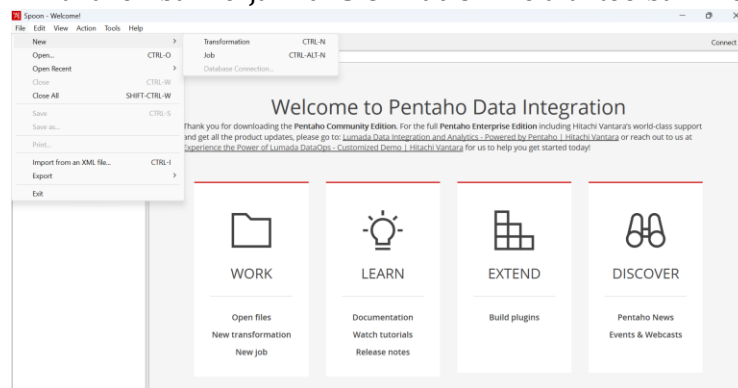
Seharusnya berisi angka (desimal atau integer), tetapi jika ada format teks atau simbol yang tidak sesuai, maka bisa menyebabkan perbedaan tipe data.

4. Kolom "Region"

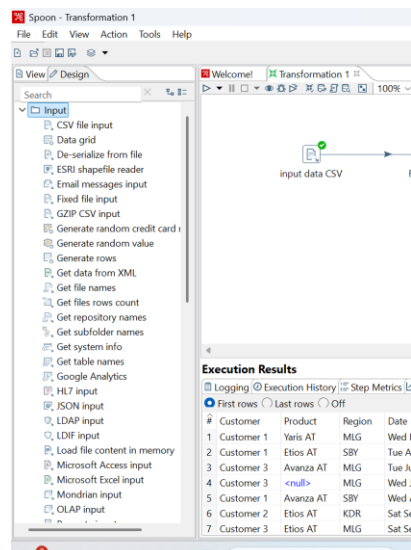
Biasanya berisi kode wilayah seperti MLG, SBY, atau KDR. Jika ada data kosong atau tidak sesuai pola, bisa dianggap berbeda tipe.

A. Pengambilan Data (Extract)

1. Buka lembar kerja **Transformation** melalui toolbar **File**.



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



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

CSV file input

Step name: input data CSV

Filename: D:\CarDistributionSales.csv [Browse...]

Delimiter: , [Insert TAB]

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

Format: mixed

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim
---	------	------	--------	--------	-----------	----------	---------	-------	------

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

CSV file input

Step name: CSV file input

Filename: D:\CarDistributionSales.csv [Browse...]

Delimiter: , [Insert TAB]

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

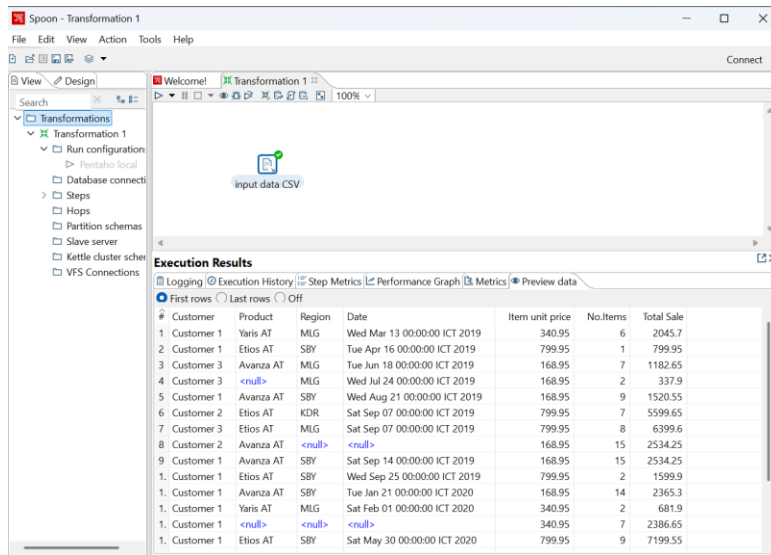
Format: mixed

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	Customer	String		10		\$.	,	none
2	Product	String		9		\$.	,	none
3	Region	String		3		\$.	,	none
4	Date	Date	MM/dd/yyyy			\$.	,	none
5	Item unit price	Number	##	6	2	\$.	,	none
6	No.Items	Integer	#	15	0	\$.	,	none
7	Total Sale	Number	##	7	2	\$.	,	none

[Help] [OK] [Get Fields] [Preview] [Cancel]

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



TUGAS 2

1. Apakah data hasil eksekusi sesuai dengan data aslinya?

Jawab: Dari hasil eksekusi sesuai dengan data aslinya jika tidak ada kesalahan dalam proses ekstraksi. Namun, perlu dicek apakah ada data yang hilang atau berubah format setelah proses extract.

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?

Jawab: Berdasarkan Modul jobsheet Langkah-langkahnya adalah:

1. Double-click pada objek **CSV file input** hingga muncul jendela konfigurasinya.
2. Mengubah **Step Name** menjadi **Input Data CSV**.
3. Pilih Lokasi file **CarDistributionSales.csv** melalui **Browse** pada **filename**.
4. Menyesuaikan delimiter (dalam kasus ini adalah koma ,)
5. Tekan **Get Fields** untuk mengambil judul pada setiap kolom
6. Tutup jendela konfigurasi tersebut.
7. Tekan tombol "**Run**" pada pojok kanan atas **Area Kerja**, maka akan muncul **Execution Result Area**. Pilih tab **Preview Data** pada **Execution Result Area**.

3. Perhatikan gambar dibawah ini! Apa yang dimaksud dengan I,O,R,W,U,E ?

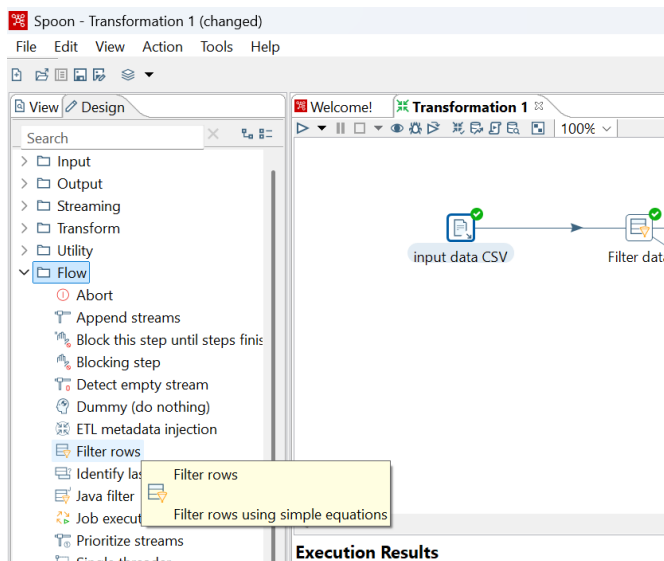
2024/02/17 06:51:48 - Input Data CSV.0 - Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)

Jawab:

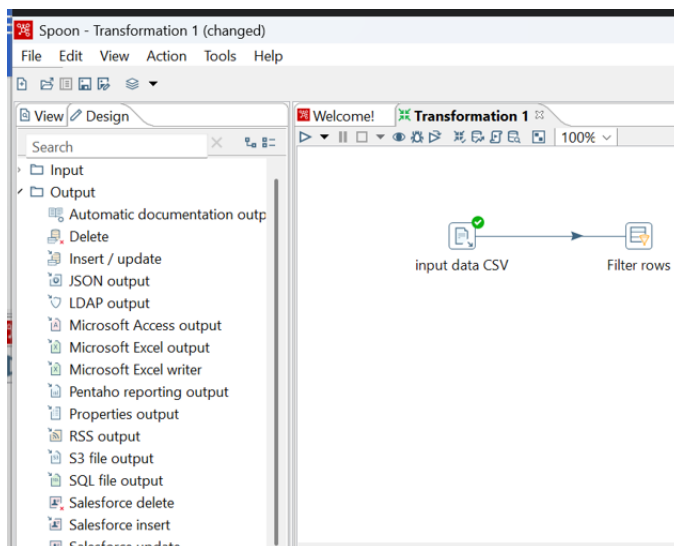
- **I(Input):** Jumlah data yang masuk ke dalam proses
 - **O(Output) :** Jumlah data yang keluar setelah proses
 - **R(Read) :** Jumlah data yang telah dibaca dari sumber
 - **W(Write) :** Jumlah data yang ditulis ke output
 - **U(Update) :** Jumlah data yang diperbarui
 - **E(Error) :** Jumlah data yang mengalami kesalahan dalam proses
4. 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)

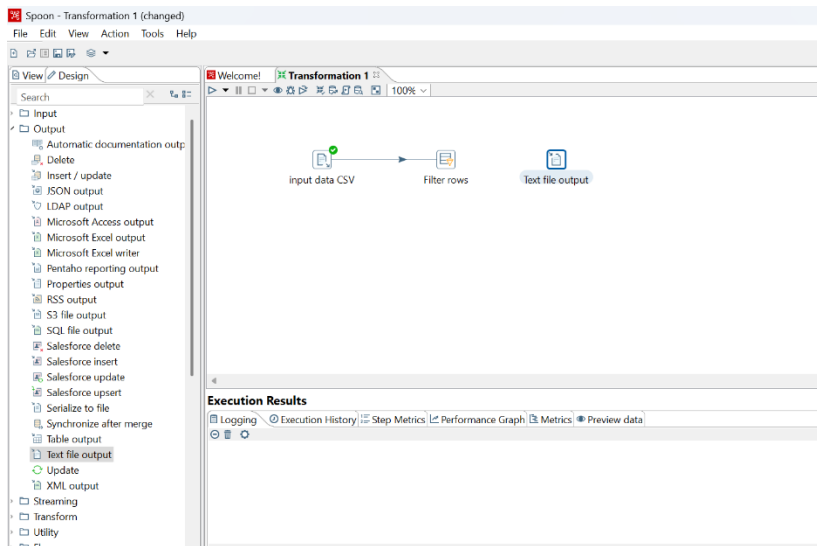
1. Carilah objek **Filter rows** pada **Design Area**. Drag and drop objek tersebut pada **Work Area**.



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



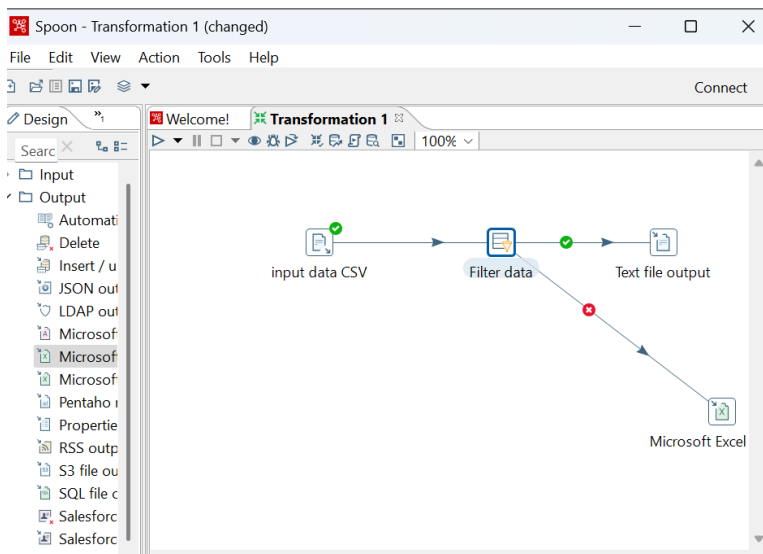
3. Carilah objek **Text file output** pada **Design Area**. Drag and drop objek tersebut pada **Work Area**.



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

5. Hubungkan **filter rows** dengan **text file output** menggunakan **konektor output** dari **filter rows**. Pilih **Result is true** sebagai jenis konektornya.

6. Hubungkan **filter rows** dengan **microsoft excel output** menggunakan **konektor output** dari **filter rows**. Pilih **Result is false** sebagai jenis konektornya.

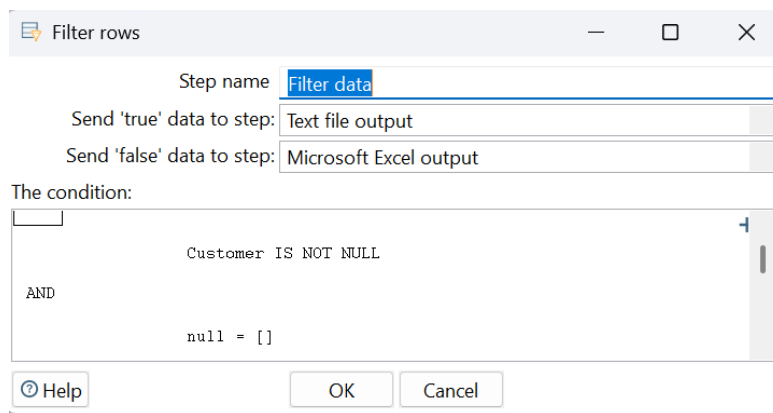


7. **Double click** pada **filter rows** hingga muncul jendela konfigurasi. ubah **step name** menjadi **Filter data**.

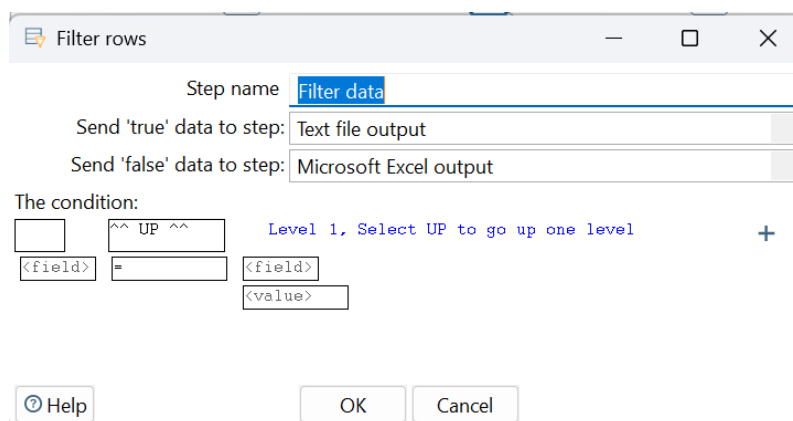
8. Cek bagian **send 'true' data to step** menjadi **Text file output** dan cek bagian **Send 'false' data to step** menjadi **Microsoft Excel output**.

9. Klik **<Field>** pada bagian **condition**

10. Pilih field yang diinginkan untuk digunakan sebagai kondisi. Pada kesempatan ini pilih **Customer**, kemudian klik **OK**.
11. Pada bagian **value** pilih **IS NOT NULL** kemudian tekan **OK** , hal ini dilakukan karena data yang dipilih adalah data yang tidak Null
12. Tekan tombol + untuk menambah kondisi.
13. Pilih statement yang masih null untuk menambah kondisi.
14. Ulangi proses nomor 9,10,11 dengan mengganti semua field yang ada. hingga statement menjadi seperti pada gambar dibawah ini.
15. Gunakan operator AND untuk menghubungkan

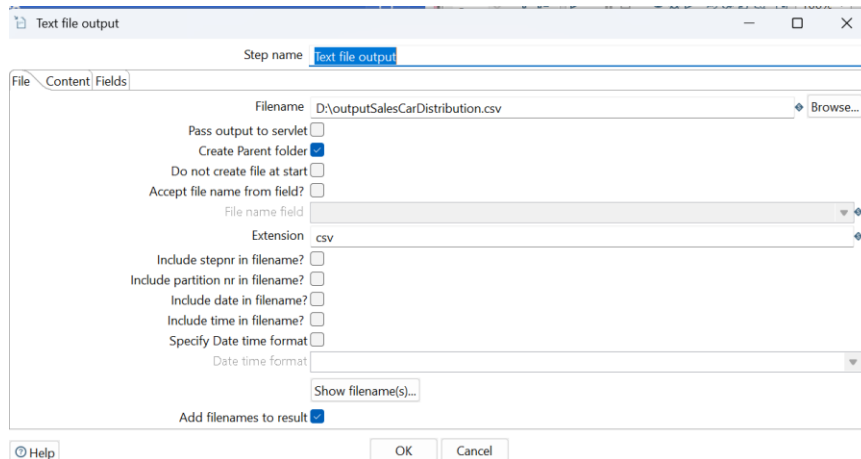


Gambar mengubah statement yang null

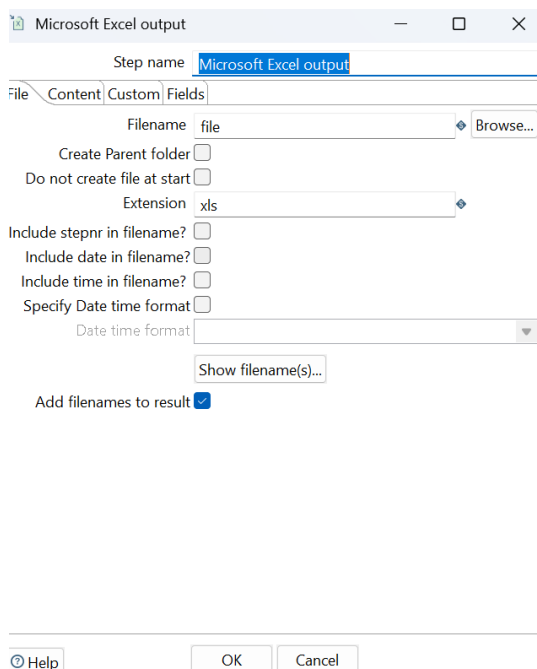


16. Jika semua field sudah masuk dalam kondisi maka tekan **OK**.

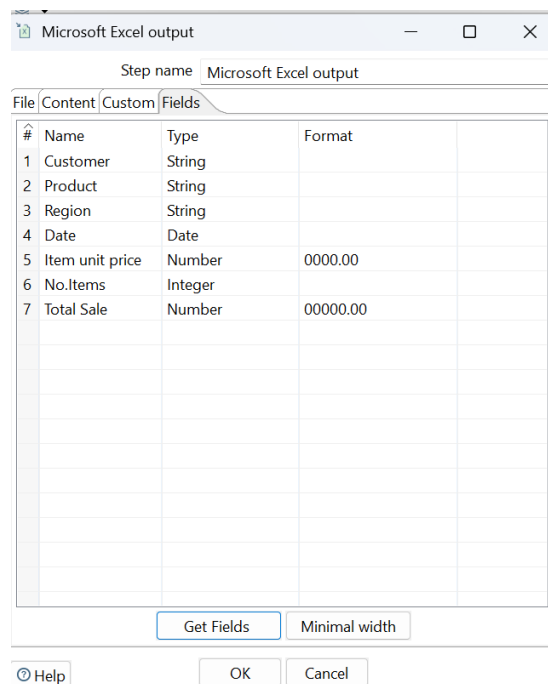
17. Double-click pada **text file output**, pilih lokasi file untuk menyimpan hasil output file dengan menekan tombol **browse**. beri nama **outputSalesCarDistribution**.
18. Ubah extension menjadi **csv**.
19. Pada tab **Fields**, tekan tombol **Get fields** untuk mengambil field data.
20. Tekan **ok** untuk kembali ke **work area**



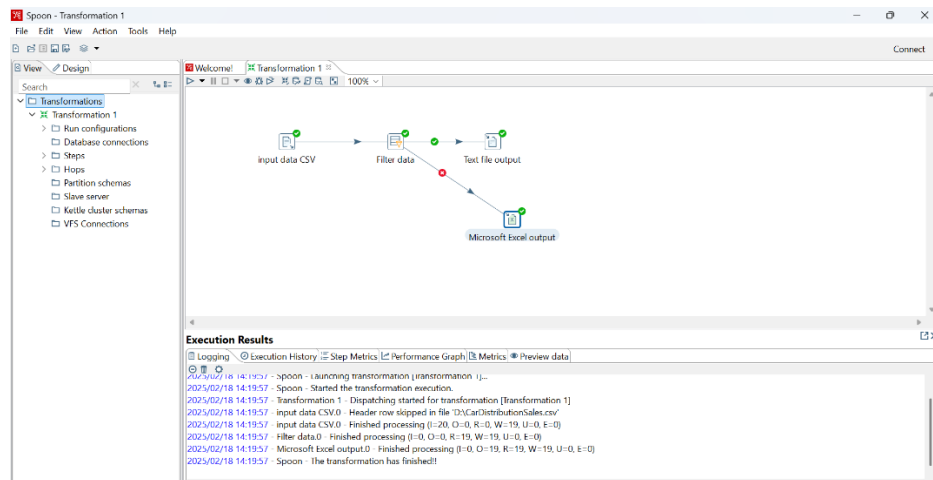
21. Hal yang sama dilakukan pada konfigurasi **Microsoft Excel output**, Double-click pada objek **Microsoft Excel output** hingga muncul jendela konfigurasi.
22. Pilih lokasi file untuk menyimpan file output dengan menekan tombol **browse**. Beri nama **failedSalesCarDistribution**. untuk extension tetap dengan extension **xls**.
23. Pada tab **fields** tekan tombol **Get fields** untuk mengambil field data.
24. Tekan **ok** untuk kembali ke **work area**.



Gambar jendela konfigurasi Microsoft Excel output



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



TUGAS 3

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?

Jawab:

- ⇒ **File CSV:** Berisi data yang sudah difilter, yaitu hanya data yang **lengkap**(tanpa nilai NULL)
- ⇒ **File Excel:** Berisi data yang **tidak lengkap**(mengandung nilai NULL) untuk dikembalikan ke supervisor

2. Jelaskan apa yang terjadi pada proses **Filter rows**!

Jawab:

- ⇒ **Filter Rows** digunakan untuk menyaring data berdasarkan kondisi tertentu
- ⇒ Dalam tugas ini, filter digunakan untuk memisahkan data yang memiliki **nilai NULL** dan yang **lengkap**
- ⇒ Data yang **tidak memiliki nilai NULL** dikirim ke **Text File Output(CSV)**
- ⇒ Data yang **Memiliki nilai NULL** dikirim ke **Microsoft Excel Output(XLS)** untuk dianalisis lebih lanjut

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

Studi Kasus

Disclaimer: Data yang tercantum di studi kasus ini adalah fiktif dan hanya digunakan untuk tujuan pembelajaran saja. Nama tempat, tanggal lahir, alamat, dan informasi lainnya yang tertera hanya tipuan belaka dan tidak memiliki kaitan dengan individu yang sebenarnya.

KASUS:

Pak Pascanowo seorang ketua RW di Ibu Kota Nusabangsa (IKN). Terdapat 20 orang dari negara Konoha yang akan pindah dan berdomisili di IKN. Tata letak pemukiman pada IKN telah dikondisikan agar berkelompok disesuaikan dengan lokasi pekerjaannya agar akses terjangkau.

Sebagai contoh, mahasiswa akan dikumpulkan dan berdomisili di wilayah edukasi yang akses dekat dengan kampus-kampus perguruan tinggi. Sedangkan dokter akan berdomisili dekat dengan rumah sakit atau layanan kesehatan dan PNS akan berdomisili dekat dengan perkantoran dan layanan publik. Selain itu, orang yang bukan dari ketiga kategori tersebut bebas memilih domisili atau tempat tinggalnya di IKN.

Pak Pascanowo akan mendapatkan data dari pemerintah pusat tentang Masyarakat Konoha yang akan pindah ke IKN. dari data tersebut Pak Pascanowo akan membagi menjadi 4 kelompok,

Data masyarakat berprofesi mahasiswa akan diberikan ke Pak Emir Makarena selaku ketua RT wilayah pendidikan.

Data masyarakat berprofesi PNS akan diberikan ke Pak Yasana Laili selaku ketua RT wilayah pemerintahan

Data masyarakat berprofesi Dokter akan diberikan ke Pak Budi Pekerti selaku ketua RT wilayah kesehatan

data-data tersebut akan digunakan untuk analisa di masing-masing wilayah dan juga Dashboard untuk membantu mengambil keputusan.

Pak Pascanowo mendapatkan data penduduk yang akan pindah ke wilayahnya melalui link: <https://github.com/dik4rizky/datasources/blob/82821b8ec341176f979956c84b79720d4aead012/dummyKependudukan.csv>

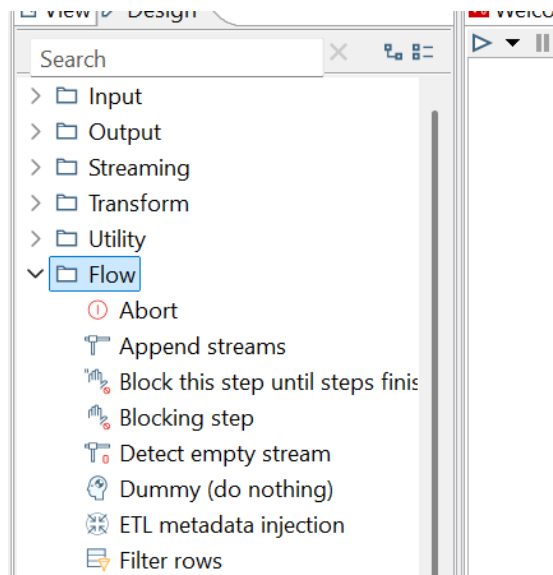
Bantulah Pak Pascanowo untuk memisahkan data tersebut menjadi 4 file: MasyarakatMahasiswa, MasyarakatPNS, MasyarakatDokter, MasyarakatLainnya.

Dikarenakan proses tersebut akan repetisi setiap ada masyarakat Konoha yang berpindah ke IKN maka buatlah sistem tersebut menggunakan PDI Spoon.

1. Tambahkan CSV File Input, Pilih file **dataDummyKependudukan.csv**, Klik **Get Fields** untuk membaca struktur kolom.

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group
1	Nama	String		16		\$	-	,
2	NIK	Integer	#	15	0	\$	-	,
3	Tempat Lahir	String		11		\$	-	,
4	Tanggal Lahir	Date	dd-MM-yyyy			\$	-	,
5	Jenis Kelamin	String		9		\$	-	,
6	Alamat	String		28		\$	-	,
7	Kota	String		11		\$	-	,
8	RT/RW	String		7		\$	-	,
9	Kelurahan	String		18		\$	-	,
1.	Kecamatan	String		17		\$	-	,
1.	Agama	String		7		\$	-	,
1.	Status Perkawinan	String		13		\$	-	,
1.	Pekerjaan	String		14		\$	-	,

2. Tambahkan Step "Filter Rows", Dari panel Design, pilih **Flow** -> **Filter Rows**, Seret ke area kerja dan hubungkan dengan CSV File Input.



3. Kita akan membuat **4 Filter Rows** untuk masing-masing kategori:
- **Filter Mahasiswa** (Profesi = 'Mahasiswa')

The screenshot shows a 'Filter rows' dialog box with the following fields and options:

- Step name:** Filter rows
- Send 'true' data to step:** MasyarakatMahasiswa.output
- Send 'false' data to step:** (empty)
- The condition:** A dropdown menu is open, showing 'Pekerjaan' selected, followed by an equals sign, another dropdown menu showing 'Mahasiswa' selected, and '(String)' in parentheses.
- Buttons:** Help, OK, and Cancel.

- **Filter PNS** (Profesi = 'PNS')

The screenshot shows a 'Filter rows' dialog box with the following fields and options:

- Step name:** Filter rows 2
- Send 'true' data to step:** MasyarakatPNS.csv
- Send 'false' data to step:** (empty)
- The condition:** A dropdown menu is open, showing 'Pekerjaan' selected, followed by an equals sign, another dropdown menu showing 'PNS' selected, and '(String)' in parentheses.
- Buttons:** Help, OK, and Cancel.

- **Filter Dokter** (Profesi = 'Dokter')

The screenshot shows a 'Filter rows' dialog box with the following fields and options:

- Step name:** Filter rows 3
- Send 'true' data to step:** MasyarakatDokter.csv
- Send 'false' data to step:** (empty)
- The condition:** A dropdown menu is open, showing 'Pekerjaan' selected, followed by an equals sign, another dropdown menu showing 'Dokter' selected, and '(String)' in parentheses.
- Buttons:** Help, OK, and Cancel.

- **Filter Lainnya (jika Profesi bukan Mahasiswa, PNS, atau Dokter)**

Filter rows

Step name: Filter rows 4

Send 'true' data to step: MasyarakatLainnya.csv

Send 'false' data to step:

The condition:

<field> = <field> +

<value>

Help OK Cancel

4. Kita akan menambahkan **Text File Output** untuk masing-masing Filter Rows:
- Hubungkan ke **"Text File Output"** bernama MasyarakatMahasiswa.csv.

Text file output

Step name: MasyarakatMahasiswa.output

File Content Fields

Filename: D:\dataDummyKependudukan.csv Browse...

Pass output to servlet: ☐

Create Parent folder: ☒

Do not create file at start: ☐

Accept file name from field: ☐

File name field:

Extension: csv

Include stepnr in filename: ☐

Include partition nr in filename: ☐

Include date in filename: ☐

Include time in filename: ☐

Specify Date time format: ☐

Date time format:

Show filename(s)...

Add filenames to result: ☒

Help OK Cancel

- Hubungkan ke **"Text File Output"** bernama MasyarakatPNS.csv.

The screenshot shows the 'Text file output' configuration window. The 'Step name' is 'MasyarakatPNS.csv'. The 'File' tab is selected. The 'Filename' is 'D:\dataDummyKependudukan.csv'. The 'Extension' is 'csv'. The 'Create Parent folder' checkbox is checked. The 'Add filenames to result' checkbox is checked. The 'Show filename(s)...' button is visible.

Step name: MasyarakatPNS.csv

File Content Fields

Filename: D:\dataDummyKependudukan.csv [Browse...]

Pass output to servlet: ☐

Create Parent folder: ☒

Do not create file at start: ☐

Accept file name from field: ☐

File name field: []

Extension: csv

Include stepnr in filename: ☐

Include partition nr in filename: ☐

Include date in filename: ☐

Include time in filename: ☐

Specify Date time format: []

Date time format: []

Show filename(s)...

Add filenames to result: ☒

Help OK Cancel

- Hubungkan ke **"Text File Output"** bernama MasyarakatDokter.csv.

The screenshot shows the 'Text file output' configuration window. The 'Step name' is 'MasyarakatDokter.csv'. The 'File' tab is selected. The 'Filename' is 'D:\dataDummyKependudukan.csv'. The 'Extension' is 'csv'. The 'Create Parent folder' checkbox is checked. The 'Add filenames to result' checkbox is checked. The 'Show filename(s)...' button is visible.

Step name: MasyarakatDokter.csv

File Content Fields

Filename: D:\dataDummyKependudukan.csv [Browse...]

Pass output to servlet: ☐

Create Parent folder: ☒

Do not create file at start: ☐

Accept file name from field: ☐

File name field: []

Extension: csv

Include stepnr in filename: ☐

Include partition nr in filename: ☐

Include date in filename: ☐

Include time in filename: ☐

Specify Date time format: []

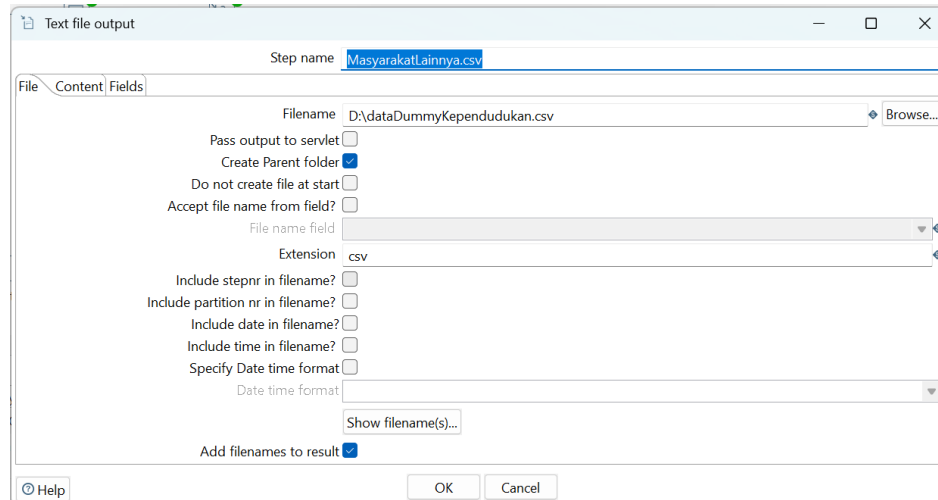
Date time format: []

Show filename(s)...

Add filenames to result: ☒

Help OK Cancel

- Hubungkan ke "Text File Output" bernama MasyarakatLainnya.csv.



5. Klik **Run (F9)** untuk menjalankan proses.

Execution Results

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan
1	Budi Santoso	3503010101010001	Jakarta	Tue Jan 01 00:00:00 ICT 1980	Laki-laki	Jl. Merdeka No. 123	Jakarta	001/002	Menteng
2	Siti Rahayu	3503010101010002	Bandung	Tue May 15 00:00:00 ICT 1990	Perempuan	Jl. Cendrawasih No. 45	Bandung	003/004	Cikutra
3	Joko Susilo	3503010101010003	Surabaya	Fri Oct 10 00:00:00 ICT 1975	Laki-laki	Jl. Diponegoro No. 67	Surabaya	005/006	Gubeng
4	Maya Dewi	3503010101010004	Yogyakarta	Sat Aug 20 00:00:00 ICT 1988	Perempuan	Jl. Malioboro No. 8	Yogyakarta	007/008	Sosromeni
5	Adi Pratama	3503010101010005	Semarang	Sun Mar 05 00:00:00 ICT 1995	Laki-laki	Jl. Pemuda No. 21	Semarang	009/010	Pandanarai
6	Rini Setiawan	3503010101010006	Malang	Sun Dec 12 00:00:00 ICT 1982	Perempuan	Jl. Simpang Borobudur No. 15	Malang	011/012	Dinoyo
7	Andi Wijaya	3503010101010007	Makassar	Sun Jun 25 00:00:00 ICT 1978	Laki-laki	Jl. Jenderal Sudirman No. 10	Makassar	013/014	Barombon
8	Desi Purnama	3503010101010008	Palembang	Thu Apr 08 00:00:00 ICT 1993	Perempuan	Jl. Demang Lebar Daun No. 20	Palembang	015/016	Alang-Alar
9	Fajar Prabowo	3503010101010009	Medan	Tue Sep 30 00:00:00 ICT 1986	Laki-laki	Jl. Sudirman No. 50	Medan	017/018	Sei Sikamb

Spoon - Transformation 2

File Edit View Action Tools Help

View Design

Search

- Input
- Output
- Streaming
- Transform
- Utility
- Flow
 - Abort
 - Append streams
 - Block this step until steps finish
 - Blocking step
 - Detect empty stream
 - Dummy (do nothing)
 - ETL metadata injection
 - Filter rows**
 - Identify last row in a stream
 - Java filter
 - Job executor
 - Prioritize streams
 - Single threader
 - Switch / case
 - Transformation executor
- Scripting
- Pentaho Server
- Lookup
- Joins
- Data Warehouse
- Validation
- Statistics

Transformation 2

CSV file input → Filter rows → Select values → Mahasiswa Output

Filter rows 2 → Filter rows 3 → Filter rows 4

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan
1	Siti Rahayu	3503010101010002	Bandung	Tue May 15 00:00:00 ICT 1990	Perempuan	Jl. Cendrawasih No. 45	Bandung	003/004	Cikutra	Bandung
2	Adi Pratama	3503010101010005	Semarang	Sun Mar 05 00:00:00 ICT 1995	Laki-laki	Jl. Pemuda No. 21	Semarang	009/010	Pandanaran	Semarang
3	Maya Wulandari	3503010101010020	Jayapura	Sat Oct 23 00:00:00 ICT 1993	Perempuan	Jl. Proklamasi No. 80	Jayapura	039/040	Kota Jayapura	Jayapura

14.43
25/02/2025

