

# **DATA WAREHOUSE**

**JS01**



**Oleh:**

**WAHYU TRISNANTOADI PRAKOSO**

**NIM. 2341760153**

**SIB-2B / 25**

**D-IV SISTEM INFORMASI BISNIS  
JURUSAN TEKNOLOGI INFORMASI  
POLITEKNIK NEGERI MALANG**

- CSV file input

Step name: Input data CSV

Filename: D:\APLIKASI\polinema\DW\JST\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	Trailing
1	Customer	String		10		IDR	.	,	no
2	Product	String		9		IDR	.	,	no
3	Region	String		3		IDR	.	,	no
4	Date	Date	MM/dd/yyyy			IDR	.	,	no
5	Item unit price	BigNumber	#,##	6	2	IDR	.	,	no
6	No.Items	Integer	#	15	0	IDR	.	,	no
7	Total Sale	BigNumber	#,##	7	2	IDR	.	,	no
8									

Help OK Get Fields Preview Cancel

### 3. Run csv

Run Options

Run configuration: Pentaho local

Options

☒ Clear log before running Log level: Basic

☐ Enable safe mode

☒ Gather performance metrics

Parameters Variables

Parameter	Default value	Value	Description

Arguments (legacy)

☒ Always show dialog on run

Help Run Cancel

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
1	Customer 1	Yaris AT	MLG	Wed Mar 13 00:00:00 WIB 2019	340.95	6	2045.7
2	Customer 1	Etios AT	SBY	Tue Apr 16 00:00:00 WIB 2019	799.95	1	799.95
3	Customer 3	Avanza AT	MLG	Tue Jun 18 00:00:00 WIB 2019	168.95	7	1182.65
4	Customer 3	<null>	MLG	Wed Jul 24 00:00:00 WIB 2019	168.95	2	337.9
5	Customer 1	Avanza AT	SBY	Wed Aug 21 00:00:00 WIB 2019	168.95	9	1520.55
6	Customer 2	Etios AT	KDR	Sat Sep 07 00:00:00 WIB 2019	799.95	7	5599.65
7	Customer 3	Etios AT	MLG	Sat Sep 07 00:00:00 WIB 2019	799.95	8	6399.6
8	Customer 2	Avanza AT	<null>	<null>	168.95	15	2534.25
9	Customer 1	Avanza AT	SBY	Sat Sep 14 00:00:00 WIB 2019	168.95	15	2534.25
10	Customer 1	Etios AT	SRV	Wed Sep 25 00:00:00 WIB 2019	799.95	2	1599.9

### TUGAS 2

1. Apakah data hasil eksekusi sesuai dengan data aslinya?

- Iya sama

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
1	Customer 1	Yaris AT	MLG	Wed Mar 13 00:00:00 WIB 2019	340.95	6	2045.7
2	Customer 1	Etios AT	SBY	Tue Apr 16 00:00:00 WIB 2019	799.95	1	799.95
3	Customer 3	Avanza AT	MLG	Tue Jun 18 00:00:00 WIB 2019	168.95	7	1182.65
4	Customer 3	<null>	MLG	Wed Jul 24 00:00:00 WIB 2019	168.95	2	337.9
5	Customer 1	Avanza AT	SBY	Wed Aug 21 00:00:00 WIB 2019	168.95	9	1520.55
6	Customer 2	Etios AT	KDR	Sat Sep 07 00:00:00 WIB 2019	799.95	7	5599.65
7	Customer 3	Etios AT	MLG	Sat Sep 07 00:00:00 WIB 2019	799.95	8	6399.6
8	Customer 2	Avanza AT	<null>	<null>	168.95	15	2534.25
9	Customer 1	Avanza AT	SBY	Sat Sep 14 00:00:00 WIB 2019	168.95	15	2534.25
10	Customer 1	Etios AT	SRV	Wed Sep 25 00:00:00 WIB 2019	799.95	2	1599.9

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?

- Membuka transformasi
- Meluncurkan transformasi
- Memulai eksekusi transformasi
- Mengirim tugas ke transformasi
- Memproses data input CSV
- Menyelesaikan transformasi

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)

1. **I (Input)**: Jumlah baris yang dibaca/dimasukkan ke dalam transformasi = 20

- **O (Output)**: Jumlah baris yang dikeluarkan/ditulis dari transformasi = 0
- **R (Read)**: Jumlah baris yang dibaca = 0
- **W (Written)**: Jumlah baris yang ditulis = 19
- **U (Updated)**: Jumlah baris yang diperbarui = 0
- **E (Error)**: Jumlah error yang terjadi selama pemrosesan = 0

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.

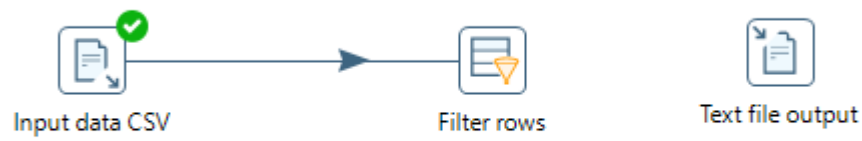
- 43 ms
- AMD ryzen 5000 series 7
- 8GB
- SSD 160 GB

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

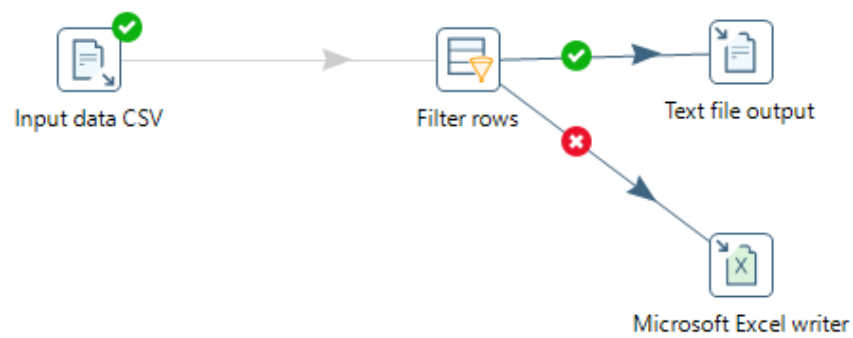
1. Menambahkan filter rows



2. Menambahkan text file output



- Menambahkan Microsoft excel output dan setting filter rows ke text file input true dan excel output false



#### 4. Set filter

Step name	Filter rows
Send 'true' data to step:	Text file output
Send 'false' data to step:	Microsoft Excel writer

The condition:

Customer IS NOT NULL

AND

Product IS NOT NULL

AND

Region IS NOT NULL

AND

Date IS NOT NULL

AND

Item unit price IS NOT NULL

AND

No. Items IS NOT NULL

AND

Total Sale IS NOT NULL

## 5. Set file text output

Step name

Fields

Filename

Pass output to servlet ☐

Create Parent folder ☒

Do not create file at start ☐

Accept file name from field? ☐

File name field

Extension

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 ☒

#	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	BigNumber	#,##	6	2	IDR	.	,	none
6	No.Items	Integer	#	15	0	IDR	.	,	none
7	Total Sale	BigNumber	#,##	7	2	IDR	.	,	none

## 6. Set excel output

Step name

Filename  [Browse...](#)

Extension

Create parent folder ☐

Stream XLSX data ☐

Split every ... data rows

Include stepnr in filename? ☐

Include date in filename? ☐

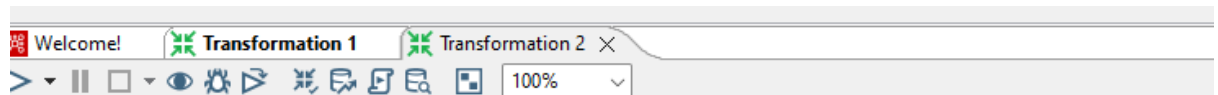
Include time in filename? ☐

Specify Date time format ☐

Date time format

Fields						
#	Name	Type	Format	Style from cell	Field title	Header/Footer sty
1	Customer	String			Customer	
2	Product	String			Product	
3	Region	String			Region	
4	Date	Date			Date	
5	Item unit price	Number			Item unit price	
6	No.Items	Integer			No.Items	
7	Total Sale	Number			Total Sale	

## 7. Run pada work area



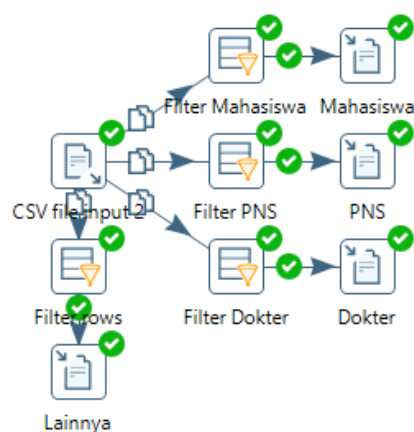
Customer;Product;Region;Date;Item unit price;No.Items;Total Sale		
Customer 1;Yaris AT ;MLG;3/13/2019;340.95;6;2045.7		
Customer 1;Etios AT ;SBY;4/16/2019;799.95;1;799.95		
Customer 3;Avanza AT;MLG;6/18/2019;168.95;7;1182.65		
Customer 1;Avanza AT;SBY;8/21/2019;168.95;9;1520.55		
Customer 2;Etios AT ;KDR;9/7/2019;799.95;7;5599.65		
Customer 3;Etios AT ;MLG;9/7/2019;799.95;8;6399.6		
Customer 1;Avanza AT;SBY;9/14/2019;168.95;15;2534.25		
Customer 1;Etios AT ;SBY;9/25/2019;799.95;2;1599.9		
Customer 1;Avanza AT;SBY;1/21/2020;168.95;14;2365.3		
Customer 1;Yaris AT ;MLG;2/1/2020;340.95;2;681.9		
Customer 1;Etios AT ;SBY;5/30/2020;799.95;9;7199.55		
Customer 3;Avanza AT;MLG;6/7/2020;168.95;10;1689.5		
Customer 3;Yaris AT ;MLG;6/9/2020;340.95;4;1363.8		
Customer 2;Avanza AT;KDR;8/8/2020;168.95;12;2027.4		

	A	B	C	D	E	F	G	H
1	Customer	Product	Region	Date	Item unit p	No.Items	Total Sale	
2	Customer 3		MLG	43670	168.95	2	337.9	
3	Customer	Avanza AT			168.95	15	2534.25	
4	Customer 1				340.95	7	2386.65	
5	Customer 2		KDR	44050	799.95	14	11199.3	
6	Customer	Avanza AT		44055	168.95	4	675.8	
7								
8								

### TUGAS 3

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?
  - **File csv data nya lengkap, file excel terdapat data kosong**
2. Jelaskan apa yang terjadi pada proses **Filter rows**!
  - **Memisahkan data:** baris yang memenuhi kondisi
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 table
  - 43 ms
  - AMD ryzen 5000 series 7
  - 8GB
  - SSD 160 GB

### Studi Kasus





MAHASISWA

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Nama;NIK	Tempat Lahir;Tanggal Lahir;Jenis Kelamin;Alamat;Kota;RT/RW;Kelurahan;Kecamatan;Agama;Status Perkawinan;Pekerjaan																
2	Siti Rahayu	;3503010101010002	;Bandung	;15-05-1990	;Perempuan	;Jl. Cendrawasih No. 45	;Bandung	;003/004	;Cikutra	;Bandung Wetan	;Kristen	;Belum Menikah	;Mahasiswa					
3	Adi Pratama	;3503010101010005	;Semarang	;05-03-1995	;Laki-laki	;Jl. Pemuda No. 21	;Semarang	;009/010	;Pandanaran	;Semarang Tengah	;Buddha	;Belum Menikah	;Mahasiswa					
4	Maya Wulandari	;3503010101010020	;Jayapura	;23-10-1993	;Perempuan	;Jl. Proklamasi No. 80	;Jayapura	;039/040	;Kota Jayapura	;Jayapura	;Kristen	;Belum Menikah	;Mahasiswa					
5																		

PNS

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Nama;NIK	Tempat Lahir;Tanggal Lahir;Jenis Kelamin;Alamat;Kota;RT/RW;Kelurahan;Kecamatan;Agama;Status Perkawinan;Pekerjaan															
2	Andi Wijaya	;3503010101010007	;Makassar	;25-06-1978	;Laki-laki	;Jl. Jenderal Sudirman No. 10	;Makassar	;013/014	;Barombong	;Makassar Selatan	;Islam	;Menikah	;PNS				
3	Agus Santoso	;3503010101010015	;Padang	;02-04-1982	;Laki-laki	;Jl. Sudirman No. 20	;Padang	;029/030	;Padang Barat	;Padang	;Islam	;Menikah	;PNS				
4																	

DOKTER

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Nama;NIK	Tempat Lahir;Tanggal Lahir;Jenis Kelamin;Alamat;Kota;RT/RW;Kelurahan;Kecamatan;Agama;Status Perkawinan;Pekerjaan																	
2	Desi Purnama	;3503010101010008	;Palembang	;08-04-1993	;Perempuan	;Jl. Demang Lebar Daun No. 20	;Palembang	;015/016	;Alang-Alang Lebar	;Palembang Selatan	;Kristen	;Belum Menikah	;Dokter						
3	Bambang Kusumo	;3503010101010017	;Banjarماسin	;25-09-1980	;Laki-laki	;Jl. A. Yani No. 60	;Banjarماسin	;033/034	;Banjarماسin Tengah	;Banjarماسin	;Hindu	;Menikah	;Dokter						

LAINNYA

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Nama;NIK	Tempat Lahir;Tanggal Lahir;Jenis Kelamin;Alamat;Kota;RT/RW;Kelurahan;Kecamatan;Agama;Status Perkawinan;Pekerjaan																
2	Budi Santoso	;3503010101010001	;Jakarta	;01-01-1980	;Laki-laki	;Jl. Merdeka No. 123	;Jakarta	;001/002	;Menteng	;Jakarta Pusat	;Islam	;Menikah	;Wiraswasta					
3	Joko Susilo	;3503010101010003	;Surabaya	;10-10-1975	;Laki-laki	;Jl. Diponegoro No. 67	;Surabaya	;005/006	;Gubeng	;Surabaya Pusat	;Katolik	;Menikah	;Pegawai Swasta					
4	Maya Dewi	;3503010101010004	;Yogyakarta	;20-08-1988	;Perempuan	;Jl. Malioboro No. 8	;Yogyakarta	;007/008	;Sosromenduran	;Yogyakarta Utara	;Hindu	;Belum Menikah	;Guru					
5	Rini Setiawan	;3503010101010006	;Malang	;12-12-1982	;Perempuan	;Jl. Simpang Borobudur No. 15	;Malang	;011/012	;Dinoyo	;Malang Selatan	;Kristen	;Menikah	;Wiraswasta					
6	Fajar Prabowo	;3503010101010009	;Medan	;30-09-1986	;Laki-laki	;Jl. Sudirman No. 50	;Medan	;017/018	;Sei Sikambing	;Medan Timur	;Hindu	;Menikah	;Wiraswasta					
7	Rina Sari	;3503010101010010	;Denpasar	;17-07-1992	;Perempuan	;Jl. Gatot Subroto No. 30	;Denpasar	;019/020	;Dauh Puri Kauh	;Denpasar Utara	;Buddha	;Belum Menikah	;Freelancer					
8	Ahmad Subhan	;3503010101010011	;Surakarta	;03-07-1991	;Laki-laki	;Jl. Pahlawan No. 55	;Surakarta	;021/022	;Jebres	;Surakarta Timur	;Islam	;Belum Menikah	;Wiraswasta					
9	Dian Pertiwi	;3503010101010012	;Maluku	;10-12-1985	;Perempuan	;Jl. Pattimura No. 10	;Ambon	;023/024	;Baguala	;Ambon	;Kristen	;Menikah	;Guru					
10	Rudi Hartono	;3503010101010013	;Pontianak	;28-05-1973	;Laki-laki	;Jl. Ahmad Yani No. 30	;Pontianak	;025/026	;Pontianak Barat	;Pontianak Kota	;Hindu	;Menikah	;Wiraswasta					
11	Lina Fitriani	;3503010101010014	;Manado	;15-08-1994	;Perempuan	;Jl. Wolter Monginsidi No. 17	;Manado	;027/028	;Mapanget	;Manado Selatan	;Kristen	;Belum Menikah	;Akuntan					
12	Rina Agustina	;3503010101010016	;Bengkulu	;19-11-1990	;Perempuan	;Jl. Imam Bonjol No. 45	;Bengkulu	;031/032	;Ratu Samban	;Bengkulu Tengah	;Kristen	;Menikah	;Wiraswasta					
13	Nia Dewi Lestari	;3503010101010018	;Manokwari	;14-06-1987	;Perempuan	;Jl. Diponegoro No. 40	;Manokwari	;035/036	;Manokwari Barat	;Manokwari	;Kristen	;Belum Menikah	;Penulis					
14	Doni Setiawan	;3503010101010019	;Palu	;07-03-1977	;Laki-laki	;Jl. Soekarno-Hatta No. 70	;Palu	;037/038	;Palu Barat	;Palu	;Islam	;Menikah	;Wiraswasta					