

## Jobsheet 1: Data Warehouse Sederhana dan Pengenalan PDI Spoon

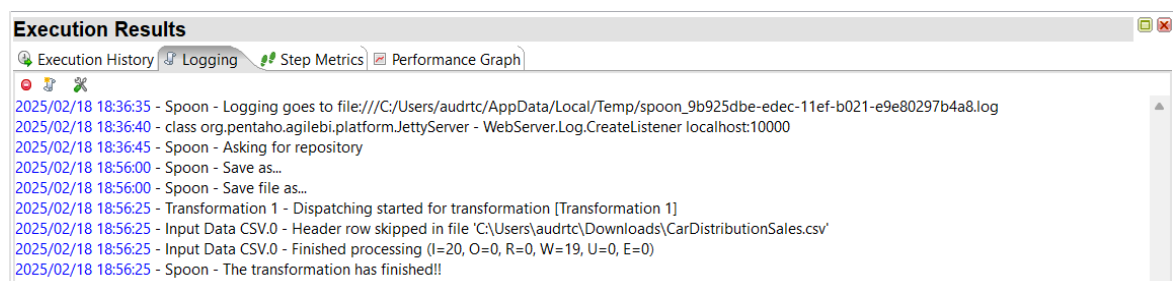
### TUGAS 1

Analisa lah data tersebut!

1. Berapa jumlah kolom pada data tersebut?  
Data tersebut memiliki 7 kolom
2. Apa arti atau isi data dari setiap kolom yang ada?
  - Customer: Nama pelanggan (String)
  - Product: Jenis mobil yang dibeli (String)
  - Region: Wilayah pembelian (String)
  - Date: Tanggal transaksi (Date)
  - Item unit price: Harga satuan mobil (Float)
  - No.Items: Jumlah unit yang dibeli (Integer)
  - Total Sale: Total penjualan untuk transaksi tersebut (Float)
3. Adakah data yang memiliki nilai null / data yang tidak lengkap?
  - Product: 3 nilai kosong
  - Region: 3 nilai kosong
  - Date: 2 nilai kosong
4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?  
Tidak ada data yang memiliki tipe yang berbeda dalam kolom yang sama

### TUGAS 2

1. Apakah data hasil eksekusi sesuai dengan data aslinya?  
Data hasil eksekusi sesuai dengan data asli, hanya ada perbedaan format angka karena pembulatan
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?



- 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)

- I (Input) → Jumlah record (baris data) yang dibaca dari sumber data
- O (Output) → Jumlah record yang ditulis ke target output
- R (Read) → Jumlah record yang dibaca oleh langkah tertentu dalam transformasi
- W (Written) → Jumlah record yang berhasil ditulis ke target (misalnya database, file, atau sistem lain)
- U (Updated) → Jumlah record yang diperbarui selama proses transformasi
- E (Error) → Jumlah record yang mengalami error atau gagal diproses

- 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. Membutuhkan waktu 0 sec untuk melakukan eksekusi extract data

Execution Results													
Execution History   Logging   Step Metrics   Performance Graph													
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	Input Data CSV	0	0	19	20	0	0	0	0	Finished	0.0s	952	

No	Nama	Waktu Eksekusi (detik)	Processor	RAM	Storage (SSD/HDD)	VGA	CPU Usage (%)
1	Satria	0.1 detik	AMD Ryzen 7 5800H	8GB	SSD 512GB	RTX 3050	40%
2	Reishi	1 detik	Inter(R) Core™ i7-7820HQ	16GB	SSD 256GB	NVIDIA GeForce 940MX	25%

### TUGAS 3

- Apa perbedaan isi data output dilihat dari isi file csv dan file excel? Pada file csv data yang ditampilkan adalah data yang valuenya not null, sedangkan file excel yang ditampilkan data yang memiliki value null
- Jelaskan apa yang terjadi pada proses **Filter rows**! Setiap baris data dicek berdasarkan kondisi yang telah ditentukan. Memisahkan data berdasarkan hasil filter, jika kondisi True data akan diteruskan ke Langkah berikutnya, sedangkan jika kondisi False data akan dialihkan ke jalur lain
- 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 Untuk melakukan eksekusi Extract Transfer Load membutuhkan waktu 0.1 sec

Execution Results

Execution History

Logging

Step Metrics

Performance Graph

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input
1	Input Data CSV	0	0	19	20	0	0	0	0	Finished	0.0s	2,000	
2	Filter data	0	19	19	0	0	0	0	0	Finished	0.0s	1,357	
3	Text file output	0	14	14	0	15	0	0	0	Finished	0.1s	134	
4	Microsoft Excel Output	0	5	5	0	5	0	0	0	Finished	0.1s	45	

No	Nama	Waktu Eksekusi (detik)	Processor	RAM	Storage (SSD/HDD)	VGA	CPU Usage (%)
1	Satria	0.1 detik	AMD Ryzen 7 5800H	8GB	SSD 512GB	RTX 3050	40%
2	Reishi	1 detik	Inter(R) CoreTM i7-7820HQ	16GB	SSD 256GB	NVIDIA FeForce 940MX	25%