**DATA WAREHOUSE  
Week 1**



**Nama:** Paudra Akbar Buana

**NIM:** 2341760063

**JURUSAN TEKNOLGI INFORMASI**

**PRODI D-IV SISTEM INFORMASI BISNIS**

**POLITEKNIK NEGERI MALANG**

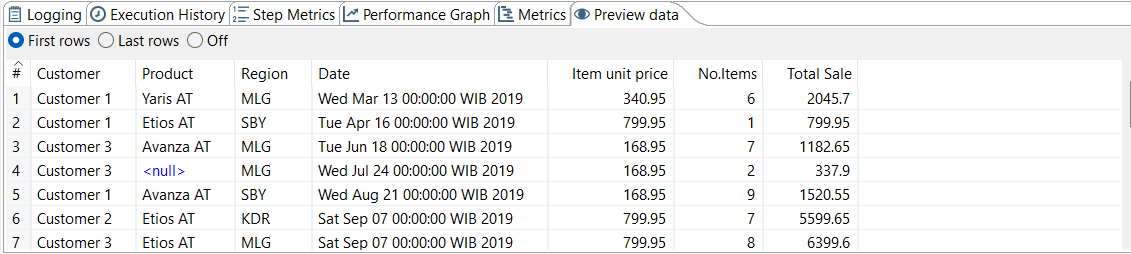
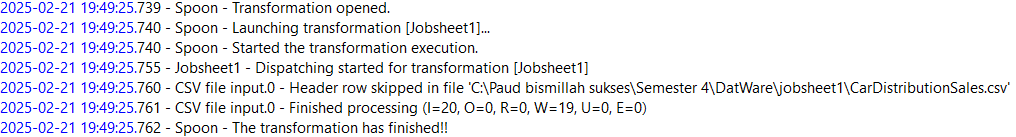
**2023/2024**

**TUGAS 1:**

Analisa lah data tersebut!

1. Berapa jumlah kolom pada data tersebut?  
   - 7
2. Apa arti atau isi data dari setiap kolom yang ada?  
   -Customer = Adalah id customer/pembeli   
   -Product = Adalah nama produk yang dibeli  
   -Region = Daerah/ Alamat dari pembeli  
   -Date = Tanggal pembelian produk  
   -Item unit price = Harga produk  
   -No.item = Jumlah produk yang dibeli  
   -Total Sales = Total harga produk yang dibeli
3. Adakah data yang memiliki nilai null / data yang tidak lengkap?  
   - Ada, pada baris 5,9,14,18,20
4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?  
   -Tidak ada

**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 ?

  
  
- **I (Input)**: Jumlah record (baris) yang dibaca dari sumber data.

- **O (Output)**: Jumlah record yang dikirim ke langkah berikutnya dalam transformasi.

- **R (Read)**: Jumlah record yang dibaca oleh langkah tertentu.

- **W (Written)**: Jumlah record yang ditulis ke target.

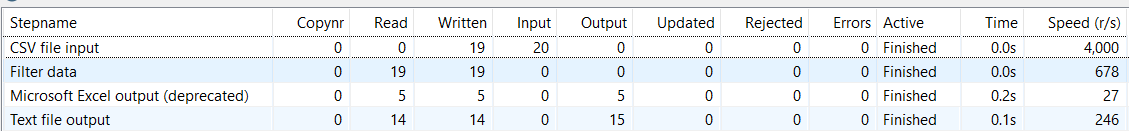
- **U (Updated)**: Jumlah record yang diperbarui.

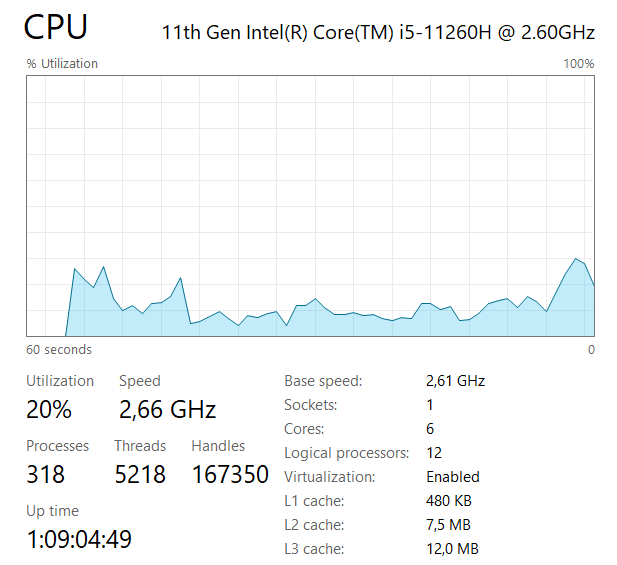
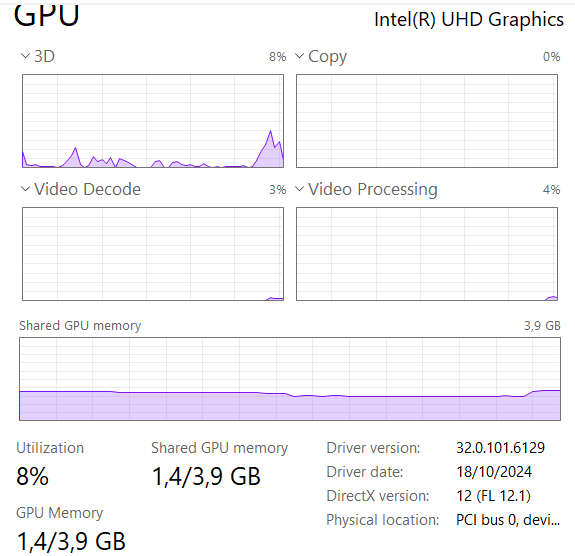
- **E (Error)**: Jumlah record yang mengalami kesalahan dalam pemrosesan.

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.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nama | Processor | RAM(GB) | Storage(GB) | VGA | CPU | Runtime(s) |
| Paudra | Intel | 8 | 500 | Intel(R) UHD Graphics | 11th Gen Intel(R) Core(TM) i5-11260H @2.60GHz | 0,00 |
| Filla | Amd | 6 | 1000 | Radeon Vega 8 Mobile GFX | AMD Ryzen 5 3500 u Processor with AMD Rageon Vega 8 Graphics | 0.4 |
| Devin | Intel | 4 | 128 | Intel HD Graphics 4000 | Intel Core i5 3437u | 1 |

**TUGAS 3:**

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?  
   -CSV: Pemisahan data pada tiap kolom dipisahkan dengan tanda koma(,)  
   -XLS: Pemisahan data pada tiap kolom dipisahkan dengan garis table(|)
2. Jelaskan apa yang terjadi pada proses **Filter rows**!  
   -Pada proses *Filter Rows (Filter Data)* dalam proses ini berperan untuk penyaringan data dari file *Input Data CSV* sebelum data tersebut diteruskan pada output yang berbeda yaitu *Text file output* (Result is true) jika memenuhi syarat dan *Microsoft Excel Output* (Result is false) jika tidak memenuhi
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  
     
   Processor:  
     
   RAM:  
   

CPU:  
  
VGA/GPU:  
  
Storage:  
