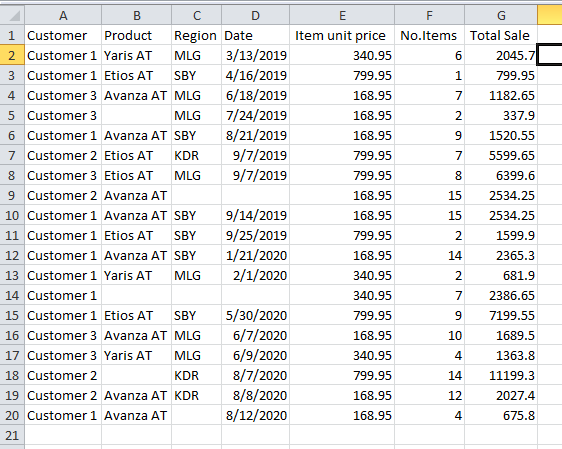
**TUGAS 1**

Analisa lah data tersebut!

1. Berapa jumlah kolom pada data tersebut?

= terdapat 7 kolom yakni customer, product, region, date, item unit price, no.items dan total sale

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

= menunjukkan data penjualan product oleh beberapa customer di daerah Malang, Surabaya dan Kediri

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

= ada yakni ada pada kolom product ada 3, kolom region ada 3, date ada 2 yang dimana data yang memiliki nilai null tersebut berbeda-beda baris dan penempatannya

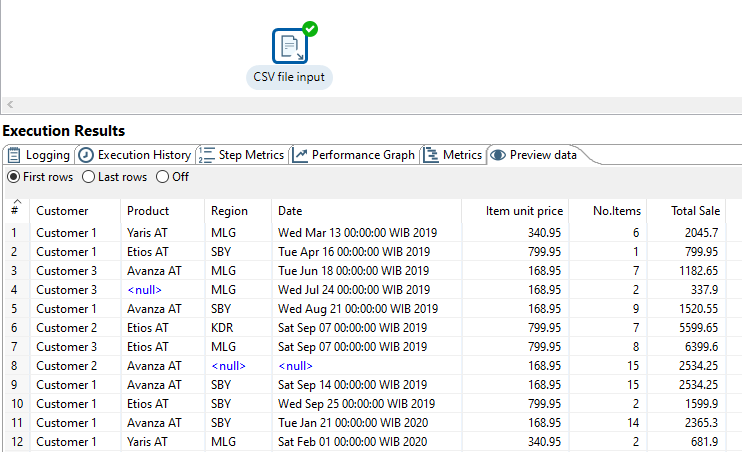
1. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?

= tidak ada

**TUGAS 2**

1. Apakah data hasil eksekusi sesuai dengan data aslinya?

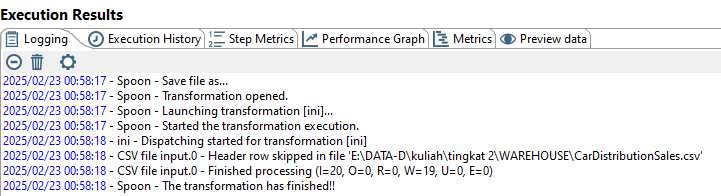
= ya, data sudah sesuai semua



1. 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?

= berikut prosesnya

1. Dilakukan save file pada folder
2. Pembukaan transformasi
3. Menjalankan transformasi
4. Ekseskusi transformasi
5. Pengiriman proses
6. Membaca file csv
7. Transformasi selesai



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



= berikut penjelasannya

* **I (Input Count) → 20 =** Jumlah total baris yang dibaca dari sumber data (dalam hal ini, file CSV
* **O (Output Count) → 0 =** Jumlah total baris yang dikirim ke langkah berikutnya dalam transformasi (jika ada). Dalam log ini, nilainya 0, mungkin karena tidak ada langkah output setelah membaca file.
* **R (Read Count) → 0 =** Jumlah baris yang dibaca dari sumber data lain yang bukan bagian dari langkah ini. Nilai 0 berarti langkah ini hanya membaca data dari file CSV tanpa membaca tambahan dari sumber lain.
* **W (Write Count) → 19 =** Jumlah baris yang berhasil ditulis setelah transformasi. Di sini, 19 baris berhasil diproses dan disimpan, kemungkinan setelah mengabaikan baris header.
* **U (Update Count) → 0 =** Jumlah baris yang diperbarui dalam suatu proses. Nilai 0 menunjukkan tidak ada pembaruan data yang dilakukan.
* **E (Error Count) → 0 =** Jumlah baris yang mengalami kesalahan dalam proses. Nilai 0 berarti tidak ada kesalahan yang terjadi selama transformasi.

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.

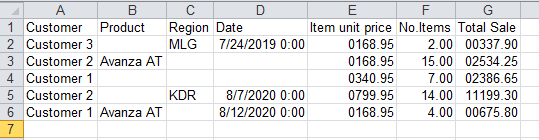
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | Processor | RAM | Storage | VGA | CPU | Time |
| Kartika | Intel Core i3-6006U 2.0 GHz | 8 GB | 535 GB | Intel HD dan NVDIA GeForce 920MX 2GB | Intel Core i3-6006U | 1 s |
| Dini | AMD Ryzzen 7 5800HS with Radeon Graphics 3.20 GHz | 16 GB | 475 GB | AMD Radeon Graphics | AMD Ryzzen 7 5800HS with Radeon Graphics | 1 s |
| Nimas | AMD Ryzen 3 5300U | 16 GB | 475 GB | AMD Radeon Graphics | AMD Ryzen 3 5300U  with Radeon Graphics | 1 s |
| Sufyan | I7 gen 9 | 16 GB | 256 GB | GTX 1650 | I7 gen 9 | 4 s |

**TUGAS 3**

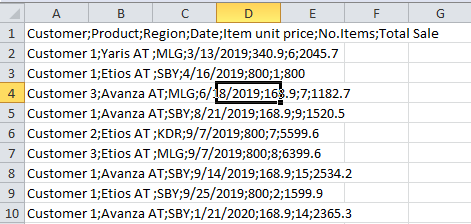
1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?

= perbedaan di kedua file tersebut adalah jika excel lebih terstruktur dan sudah pisah sesuai kolom-kolomnya sedangkan file csv masih berantaan dan perlu di pisah lagi sesuai kolomnya

1. File excel



1. File csv



1. Jelaskan apa yang terjadi pada proses Filter rows!

= Langkah **"Filter rows"** menyaring data dengan memastikan semua kolom **tidak kosong (NULL)**.

* **Data lengkap** (tidak ada NULL) = dikirim ke **Text file output**
* **Data tidak lengkap** (ada NULL) = dikirim ke **Microsoft Excel output**

Tujuannya untuk **memisahkan data valid dan tidak valid** sebelum diproses lebih lanjut.

1. 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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | Processor | RAM | Storage | VGA | CPU | Time |
| Kartika | Intel Core i3-6006U 2.0 GHz | 8 GB | 535 GB | Intel HD dan NVDIA GeForce 920MX 2GB | Intel Core i3-6006U | 0 s |
| Dini | AMD Ryzzen 7 5800HS with Radeon Graphics 3.20 GHz | 16 GB | 475 GB | AMD Radeon Graphics | AMD Ryzzen 7 5800HS with Radeon Graphics | 1 s |
| Nimas | AMD Ryzen 3 5300U | 16 GB | 475 GB | AMD Radeon Graphics | AMD Ryzen 3 5300U  with Radeon Graphics | 4 s |

**STUDY CASE**

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. darai 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

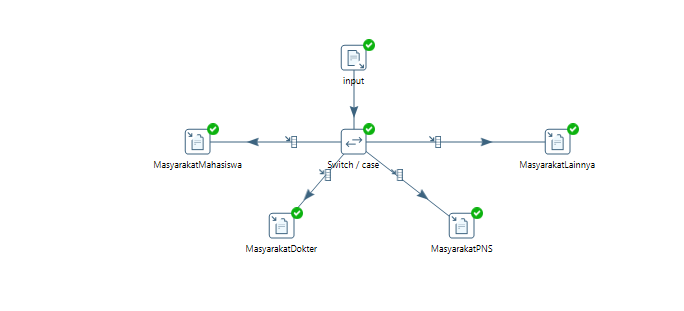
**Data masyarakat berprofesi Dokter** akan diberikan ke Pak Budi Pekerti selaku ketua RT wilayah kesehatan

data-data tersbut akan digunakan untuk analisa di masing-masing wilayah dan juga Dashboard untuk emmbantu mengambil keputusan.

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

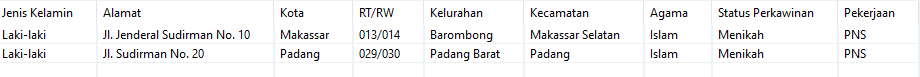
Bantulah Pak Pascanowo untuk memishkan data tersebut menjadi 4 file: MasyarakatMahasiswa, MasyarakatPNS, MasyarakatDokter, MasyarakatLainnya.

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

****

1. **Mahasiswa**

****

1. **Dokter**
2. **PNS**
3. **Lainnya**