

NIM : 2241720255

Nama Lengkap : Davis Maulana Hermanto

Daftar Isi

[Use this style for TOC Level 1 / Chapter (Bab) 3](#_Toc98155632)

[Use this style for TOC Level 2 / Sub Chapter (Sub Bab) 3](#_Toc98155633)

[Use this style for TOC Level 3 / Sub Sub Chapter (Sub Sub Bab) 3](#_Toc98155634)

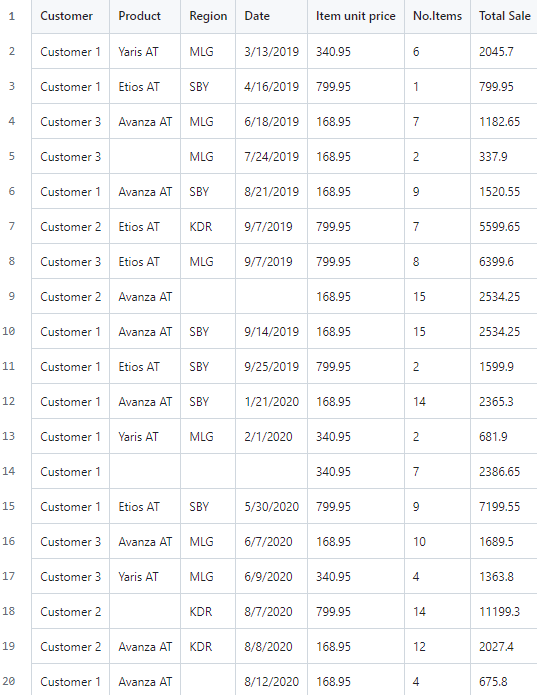
# Chapter 1(Arsitektur Data Warehouse)

**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



Tugas 1

Analisa lah data tersebut

1. Berapa jumlah kolom pada data tersebut?

* Ada 7 kolom yang terdiri dari kolom customer, product, region, date, item unit price, no. Items, dan Total Sale.

1. Apa arti atau isi data dari setiap kolom yang ada?
2. Customer : Berisi nama lengkap pelanggan
3. Product : Berisi nama model mobil yang terjual
4. Region : Berisi wilayah penjualan mobil, seperti MLG (Malang), SBY (Surabaya), dan KDR (Kediri)
5. Date : Berisi tanggal penjualan mobil dalam format dd/mm/yyyy
6. Item Unit Price : Berisi harga satuan produk mobil
7. No. Items : Berisi jumlah produk mobil yang terjual pada tanggal tersebut
8. Total Sale : Berisi total penjualan produk mobil pada tanggal tersebut
9. Adakah data yang memiliki nilai null / data yang tidak lengkap?

* Ada, pada baris 5 (kolom Product), baris 9 (kolom Region dan Date), baris 14(kolom Product, Region, dan Date), baris 18 (kolom Product), dan baris 20 (kolom Region)

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

* Tidak ada

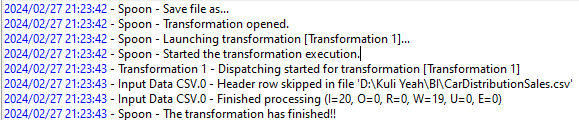
## Sub Chapter 1 (Pengambilan Data (Extract))

Tugas 2

1. Apakah data hasil eksekusi sesuai dengan data aslinya?

* Ya, datanya sama

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?
   1. PDI Spoon akan menyimpan data CarDistributionSales tersebut terlebih dahulu
   2. Setelah itu, data tersebut akan ditransform
   3. Setelah itu, data yng sudah ditransform tersebut akan dieksekusi



1. Perhatikan gambar dibawah ini! Apa yang dimaksud dengan I,O,R,W,U,E?
   1. I : Input Data CSV.0, Menunjukkan bahwa data yang diproses berasal dari file CSV dengan nama "Data CSV.0"
   2. O : Finished processing, Menunjukkan bahwa proses pengolahan data telah selesai.
   3. R : 0, Menunjukkan jumlah baris data yang ditolak karena memiliki kesalahan
   4. W: 19, Menunjukkan jumlah baris data yang diterima dan disimpan
   5. U : 0, Menunjukkan jumlah baris data yang diperbarui
   6. E : 0, Menunjukkan jumlah baris data yang dihapus



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.

* Saya hanya memerlukan waktu 1 detik saja untuk melakukan semua proses tersebut

## Sub Chapter 1 (Filter Data (Transform) dan Pengemasan data (Load))

Tugas 3

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?
   1. Pada file csv, data yang ditampilkan telah tersaring sehingga tidak ada data null pada suatu kolom setiap barisnya. Sedangkan pada file excel, data yang ditampilkan merupakan baris baris yang memiliki data null pada suatu kolomnya.
2. Jelaskan apa yang terjadi pada proses **Filter rows**!
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

### Use this style for TOC Level 3 / Sub Sub Chapter (Sub Sub Bab)