

2241720145

Muhammad Baihaqi Aulia Asy’ari

Daftar Isi

[Jobsheet 1: Data Warehouse Sederhana dan Pengenalan PDI Spoon 3](#_Toc169367702)

[TUGAS 1 3](#_Toc169367703)

[A. Pengambilan Data (Extract) 4](#_Toc169367704)

[TUGAS 2 8](#_Toc169367705)

[B. Filter Data (Transform) dan Pengemasan data (Load) 9](#_Toc169367706)

[TUGAS 3 15](#_Toc169367707)

# Jobsheet 1: Data Warehouse Sederhana dan Pengenalan PDI Spoon

### TUGAS 1

Analisa lah data tersebut!

1. Berapa jumlah kolom pada data tersebut?
2. Apa arti atau isi data dari setiap kolom yang ada?
3. Adakah data yang memiliki nilai null / data yang tidak lengkap?
4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?

### JAWABAN

1. Terdapat 7 kolom di table tersebut.
2. Customer berisi index customer yang membeli, Product berisi nama produk yang dibeli atau jenis produk yang terjual, Region berisi kode daerah dari penjualan, date berisi tanggal dari penjualan, item unit price berisi harga dari produk, no item berisi jumlah barang yang terjual dan total sale berisi jumlah penjualan di transaksi tersebut.
3. Ya, terdapat beberapa data yang tidak lengkap.
4. Tidak ada data di kolom yang sama dengan tipe data yang berbeda.

## A. Pengambilan Data (Extract)

### 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 ?



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.

### ANSWER

1. Yes, the executed data is the same as the original data
2. The Pentaho DI opened, launched, started, and dispatched the transformation. then it skipped the header row of the CSV file. After that, it processes every row and finishes the transformation.
3. The I, O, R, W, U, E stands for input, output, read, written, updated, and errors.
4. –

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CPU | Ryzen 3 3200 | Ryzen 5 3500X | Intel i5-11300H | Intel i3-1115g4 |
| RAM | 8 GB | 16 GB | 8 GB | 12 GB |
| Storage | HDD | HDD | SSD | SSD |
| GPU | Radeon Vega | RTX 3060ti | GTX 1650 | Intel HD Graphic |
| Time | 102 ms | 63 ms | 13 ms | 18 ms |

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

### TUGAS 3

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

### ANSWER

1. The csv file has complete data in every column, the excel file has the incomplete data in some column.
2. The filter row separates the data that fits with the condition and data that doesn’t into the true or false connector.
3. –

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CPU | Ryzen 3 3200 | Ryzen 5 3500X | Intel i5-11300H | Intel i3-1115g4 |
| RAM | 8 GB | 16 GB | 8 GB | 12 GB |
| Storage | HDD | HDD | SSD | SSD |
| GPU | Radeon Vega | RTX 3060ti | GTX 1650 | Intel HD Graphic |
| Time | 514 ms |  |  |  |

## Studi Kasus

KASUS:

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>

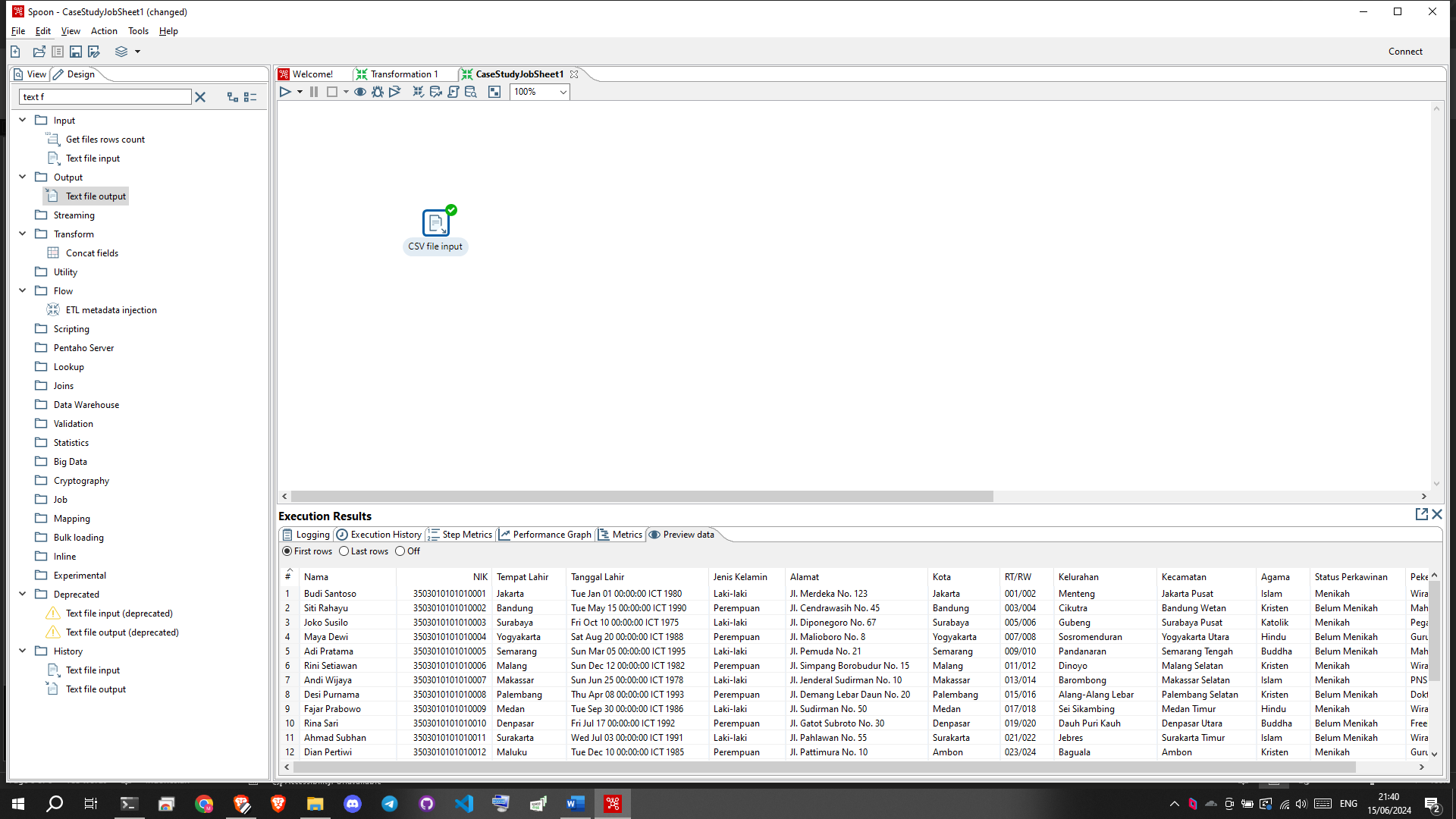
Bantulah Pak Pascanowo untuk memishkan data tersebut menjadi 4 file:

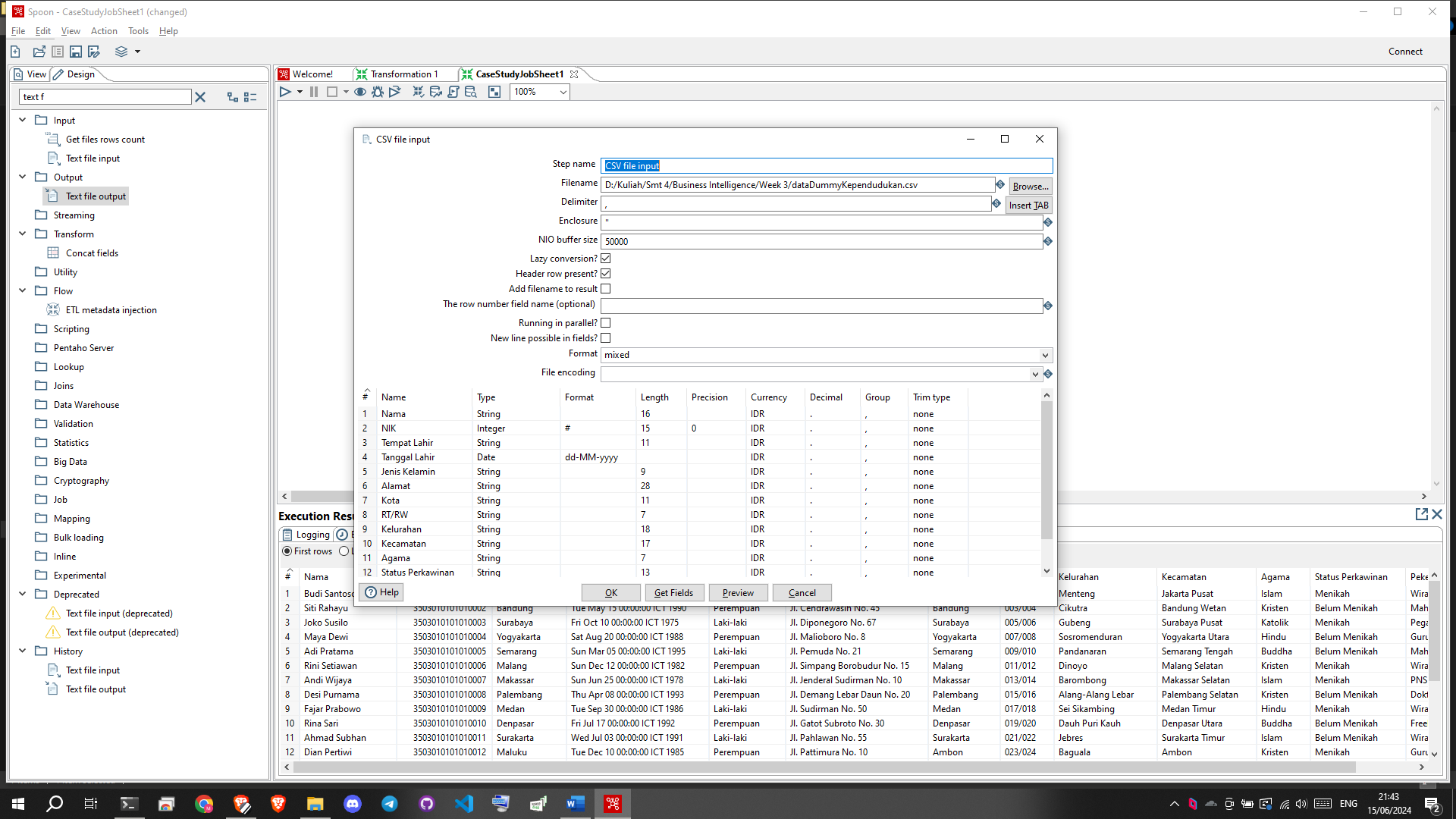
MasyarakatMahasiswa, MasyarakatPNS, MasyarakatDokter, MasyarakatLainnya.

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

### Case Study Steps

1. Drag the ‘CSV file input’ object to the ‘Design Area’ and configure to open the ‘dataDummyKependudukan’ file.





1. Drag the ‘Filter Row’ object to the ‘Design Area’ and configure to open the ‘dataDummyKependudukan’ file.