

Name : Azaria Cindy Sahasika

Number Id : 2341760169 / 06

Class : 2G – Business Information System

Lesson : Advanced Web Programing

Github Link : <https://github.com/azariacindy>

**Jobsheet 1**

**Data Warehouse Sederhana dan Pengenalan PDI Spoon**

Study Case:

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.

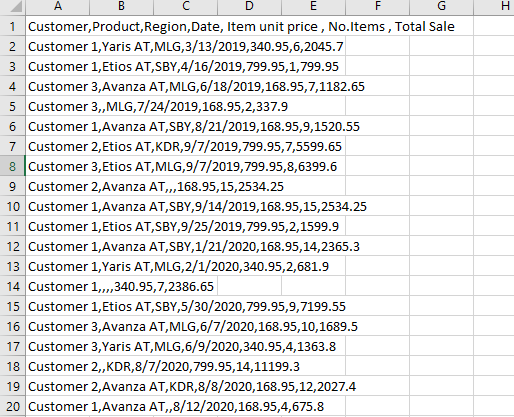
Dari studi kasus tersebut maka, akan dilakukan:

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?



* + Jumlah kolom ada 7 yaitu kolom ‘Customer’, ‘Product’, ‘Region’, ‘Date’, ‘Item unit price’, ‘No.Items’ dan ‘Total Sale’.

1. Apa arti atau isi data dari setiap kolom yang ada?
   * 1. Customer: nama pelanggan
     2. Product: nama product
     3. Region: wilayah tempat transaksi dilakukan
     4. Date: tanggal transaksi
     5. Item unit price: harga per unit product
     6. No. Item: jumlah unit produk yang dibeli dalam transaksi
     7. Total Sale: total transaksi dari perkalian ‘item unit price’ dan ‘no. item’
2. Adakah data yang memiliki nilai null/data yang tidak lengkap?
   * Ada beberapa baris yang isi kolomnya tidak lengkap, seperti pada kolom ‘product’, ‘region’ dan ‘date’.
3. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?
   * Pada beberapa kolom yang kosong atau ‘null’ otomatis akan memiliki tipe data yang berbeda. Tetapi pada dasarnya tidak ada tipe data yang berbeda dpada kolom yang sama.

**Tugas 2**

**Pengambilan Data (Extract)**

1. Apakah data hasil eksekusi sesuai dengan data aslinya?



* + Hasil eksekusi sama dengan data aslinya.

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?
   * + - Membaca sumber data
       - Membersihkan data
       - Transformasi data
       - Filtering dan sorting
       - Memvalidadi data sebelum diproses
       - Data dicatat pada file csv baru
2. Perhatikan gambar dibawah ini! Apa yang dimaksud dengan I, O, R, W, U, E?



* + - * I = input/jumlah data yang diambil
      * O = output/jumlah data yang dihasilkan
      * R = read/jumlah data yang dibaca
      * W = write/jumlah data yang berhasil dicatat
      * U = update/jumlah data yang diperbarui
      * E = error/jumlah error selama data diproses

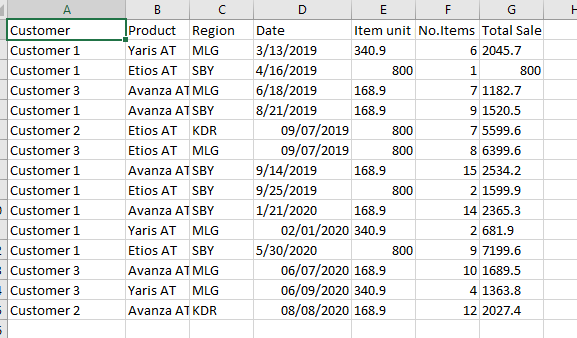
1. Berapa waktu yang dibutuhkan untuk melakukan eksekusi Extract data ini? Bandingkan dengan 3 teman yang lain, identifikasi pula spesifikasi perangkat yang digunakan (Processor, RAM, Storage, VGA, CPU). Bandingkan dalam bentuk tabel.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nama | Waktu | Processor | RAM | Storage | VGA | CPU |
| Cindy | 6s | Intel i5-10300H | 12 GB | HDD 1 TB | NVIDIA GTX 1050 | 60% |
| Evan | 5s | Intel i7-10750H | 16 GB | SSD NVMe 512 GB | NVIDIA GTX 1650 Ti | 70% |
| Kiya | 7s | Intel Celeron N4020 | 6 GB | HDD | Integrated | 50% |
| Afif | 6s | AMD Ryzen 5 4600H | 8 GB | SSD SATA 256 GB | Integrated Vega 6 | 70% |

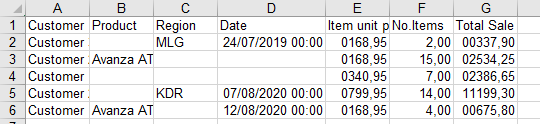
**Tugas 3**

**Filter Data (Transform) dan Pengemasan Data (Load)**

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



* + File csv berisi data yang tidak ada nilai null



* + File xls berisi data yang ada nilai null atau tidak lengkap

1. Jelaskan apa yang terjadi pada proses Filter rows!
   * Langkah transformasi yang bertujuna untuk menyaring data berdasrkan kondisi tertentu. Jika true akan diteruskan ke langkah selanjutnya, jika false maka dapat dihapus atau diteruskan ke langkah yang lainnya.
2. 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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nama | Waktu | Processor | RAM | Storage | VGA | CPU |
| Cindy | 6s | Intel i5-10300H | 12 GB | HDD 1 TB | NVIDIA GTX 1050 | 60% |
| Evan | 5s | Intel i7-10750H | 16 GB | SSD NVMe 512 GB | NVIDIA GTX 1650 Ti | 70% |
| Kiya | 7s | Intel Celeron N4020 | 6 GB | HDD | Integrated | 50% |
| Afif | 6s | AMD Ryzen 5 4600H | 8 GB | SSD SATA 256 GB | Integrated Vega 6 | 70% |

**Studi 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. Dari 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

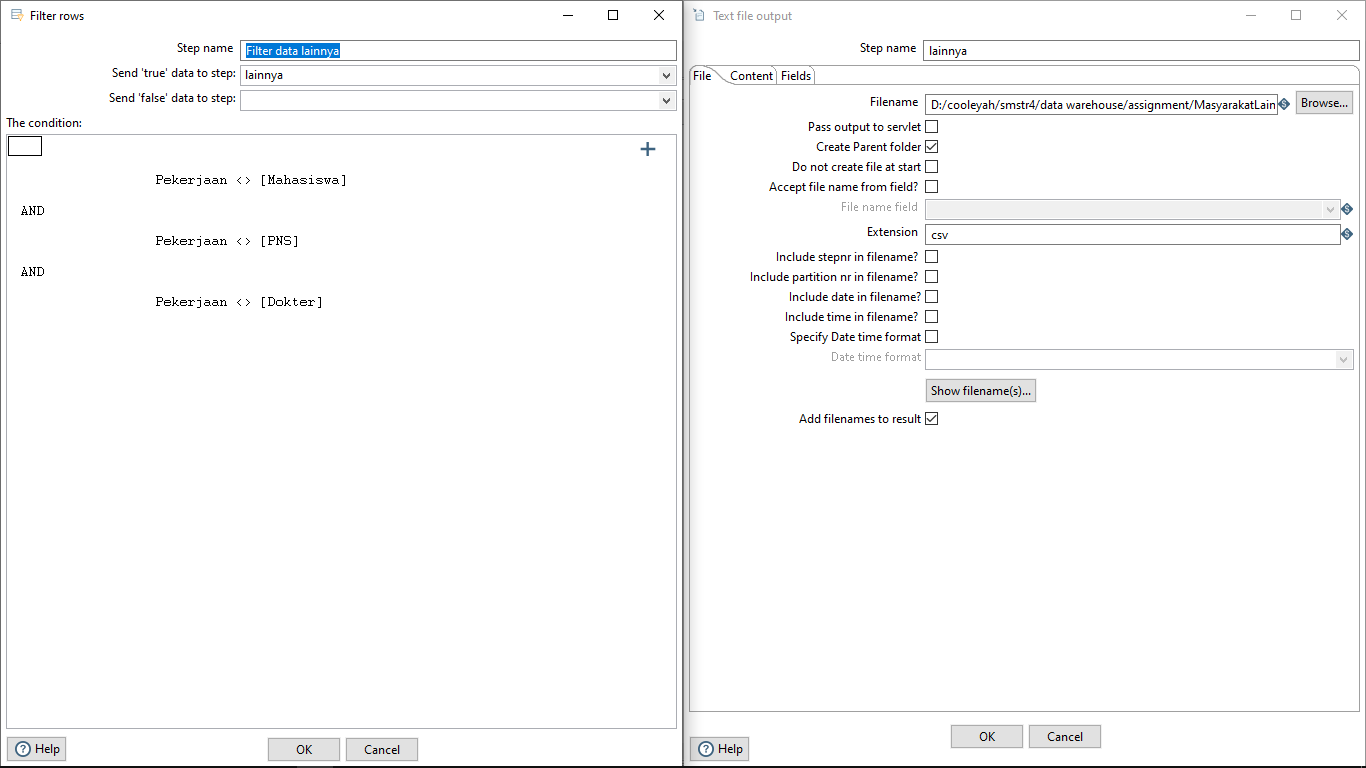
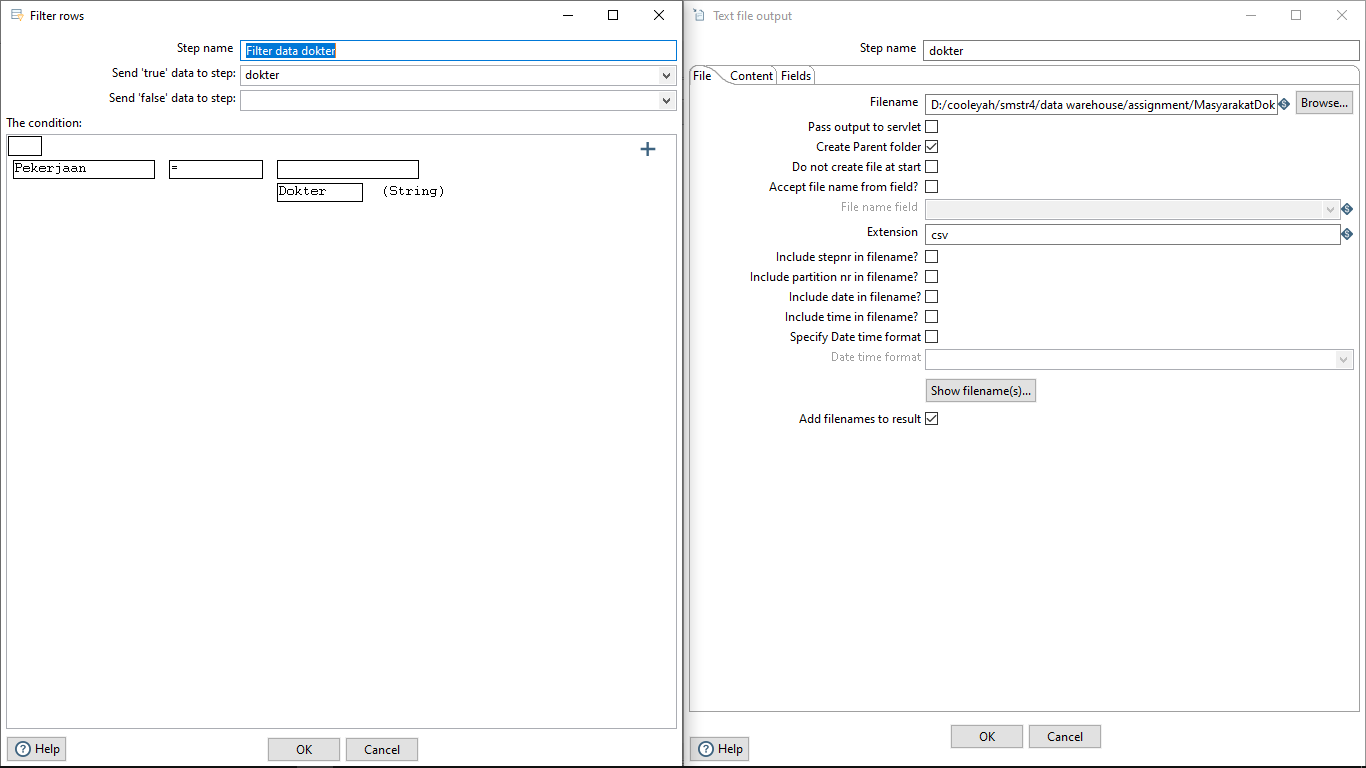
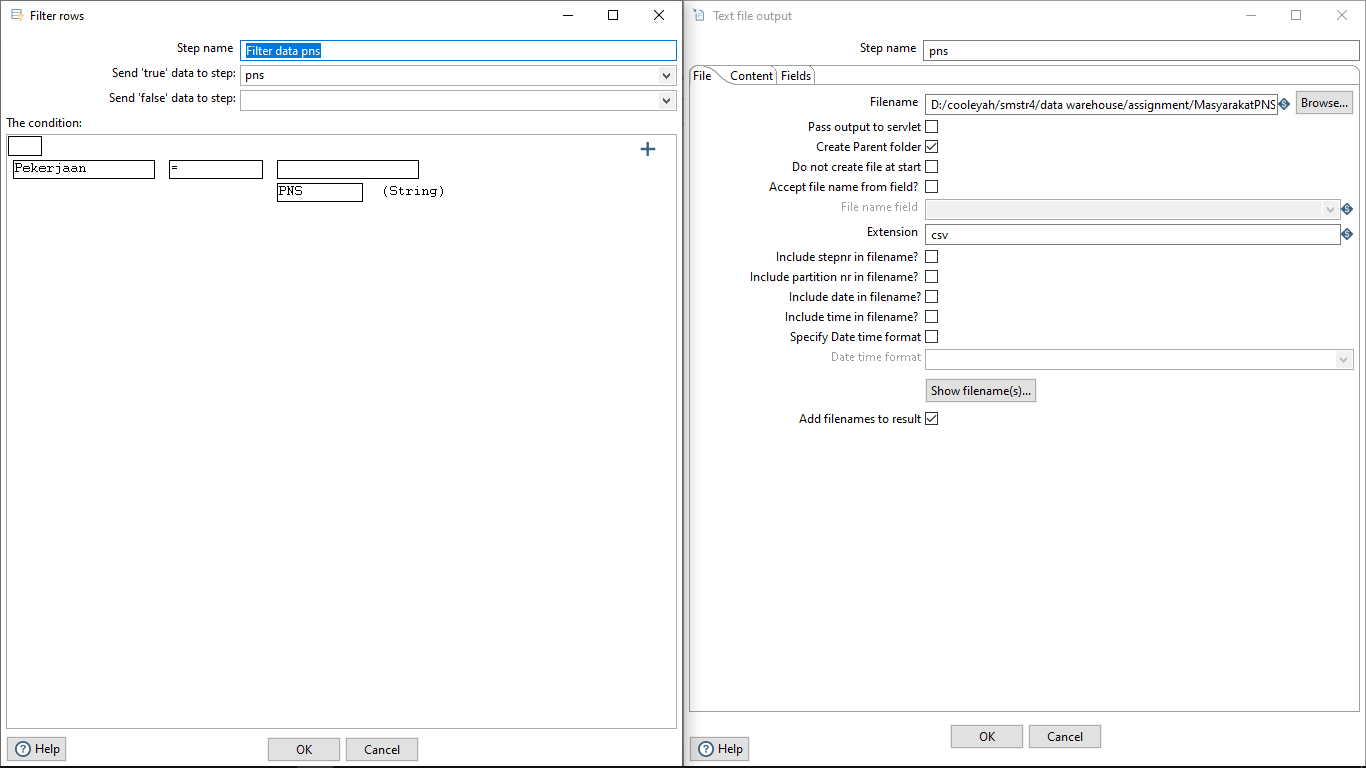
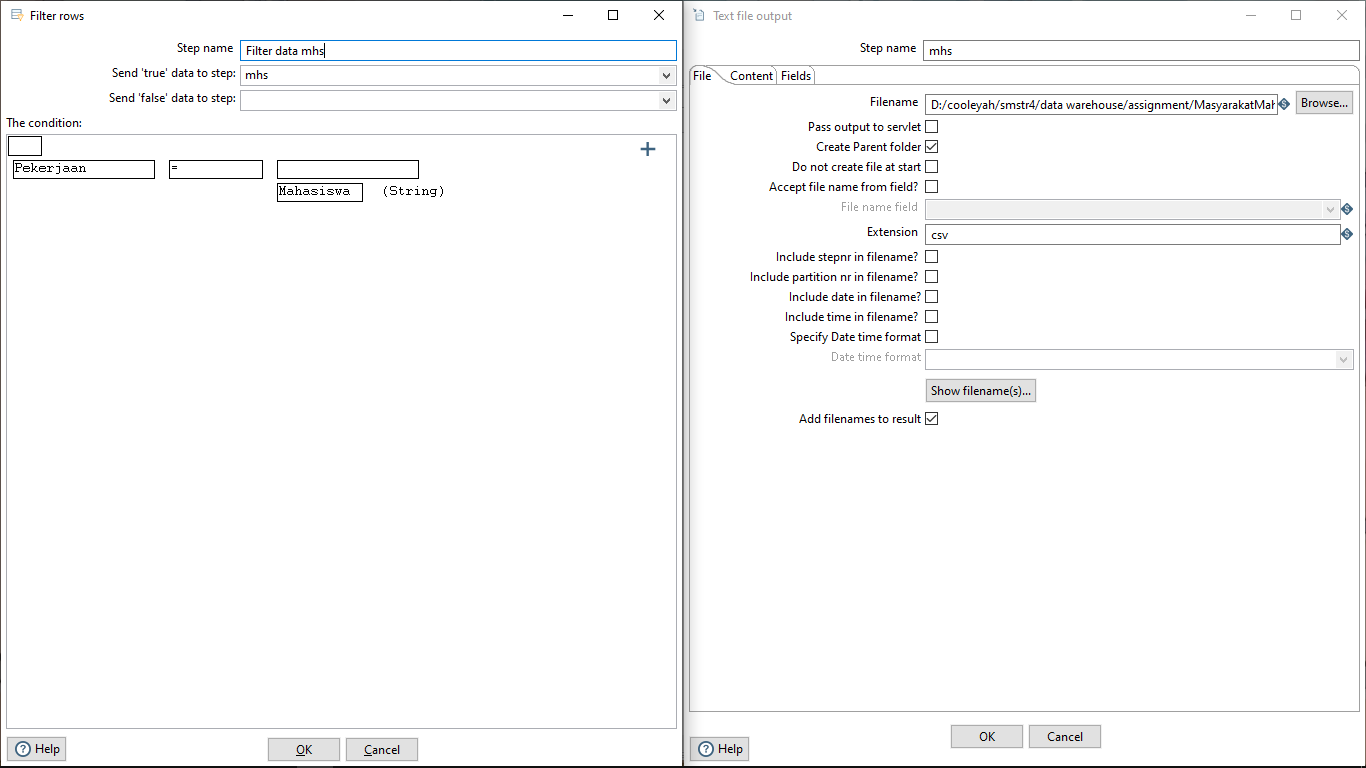
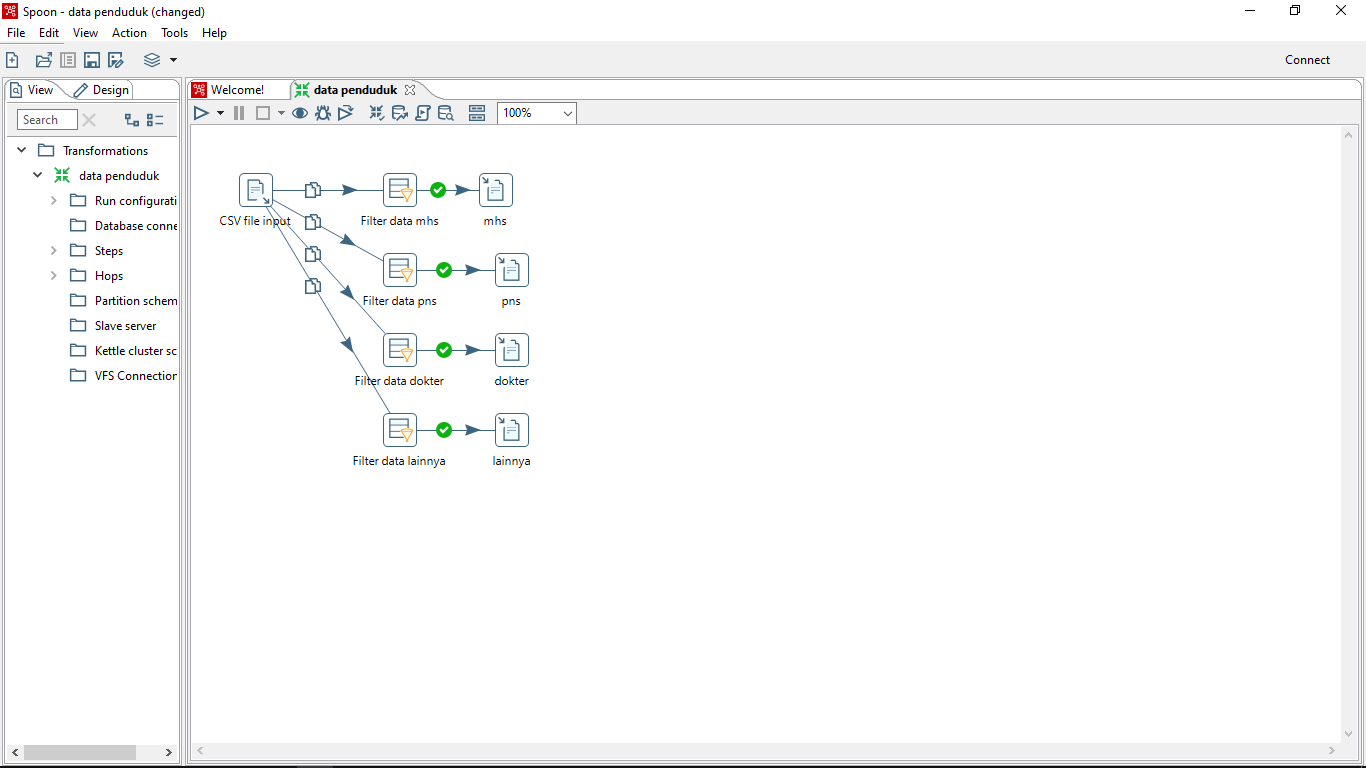
Data-data tersbut akan digunakan untuk analisa di masing-masing wilayah dan juga Dashboard untuk emmbantu mengambil keputusan. Pak Pascanowo mendapatakan data penduduk yang akan pindah ke wilayahnya melalui link: <https://github.com/dik4rizky/datasources/blob/82821b8ec341176f979956c84b79720d4aead012/dataDummyKependudukan.csv>

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.

**Jawaban Studi Kasus**



Output hasil data:

* + - * MasyarakatMahasiswa



* + - * MasyarakatPNS



* + - * MasyarakatDokter



* + - * MasyarakatLainnya

