Nama: Adinda Ivanka Maysanda Putri

Kelas : SIB 2B NIM : 2341760058

JOBSHEET 1

DATA WAREHOUSE SEDERHANA DAN PENGENALAN PDI SPOON

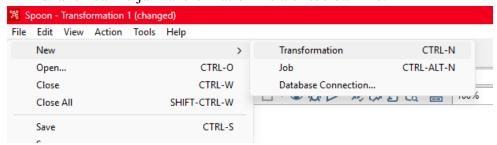
TUGAS 1

Analisa lah data tersebut!

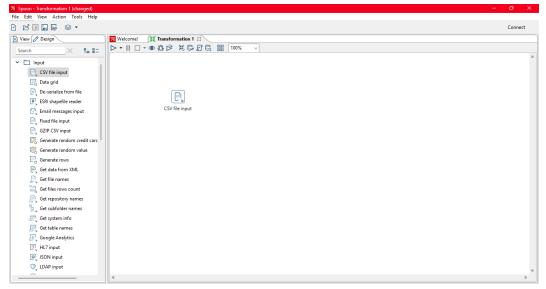
- 1. Berapa jumlah kolom pada data tersebut?
 - Pada data tersebut terdapat 7 kolom yaitu Customer, Product, Region, Date, Item unit price, No.Items, dan Total Sale
- 2. Apa arti atau isi data dari setiap kolom yang ada?
 - Customer: Berisi data setiap pelanggan yang melakukan pembelian.
 - Product: Berisi jenis produk yang dibeli oleh pelanggan.
 - Region: Menunjukkan lokasi atau cabang tempat transaksi/penjualan dilakukan.
 - Date: Menunjukkan tanggal terjadinya transaksi/penjualan.
 - Item Unit Price: Menunjukkan harga satuan dari produk yang dibeli.
 - No. Items: Menunjukkan jumlah produk yang dibeli oleh pelanggan.
 - Total Sale: Berisi total harga dari transaksi/penjualan tersebut.
- 3. Adakah data yang memiliki nilai null / data yang tidak lengkap?
 - Beberapa data memiliki nilai null atau tidak lengkap diantaranya:
 - Baris no 5, kolom product bernilai null
 - Baris no 9, kolom region dan juga date bernilai null
 - Baris no 14, kolom product, region, date bernilai null
 - Baris no 18, kolom product bernilai null
 - Baris no 10, kolom region bernilai null
- 4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?
 - Tidak ada tipe data yang berbeda dengan data lainnya, hanya ada beberapa data yang bernilai null

A. Pengambilan Data (Extract)

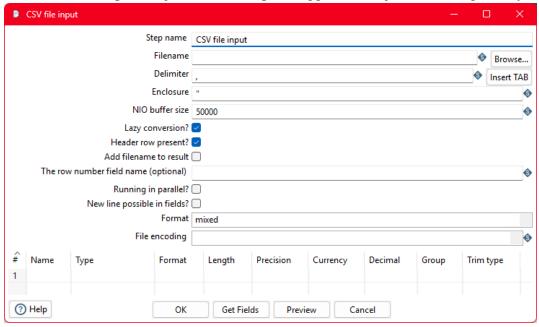
1. Buka lembar kerja Transformation melalui toolbar File.



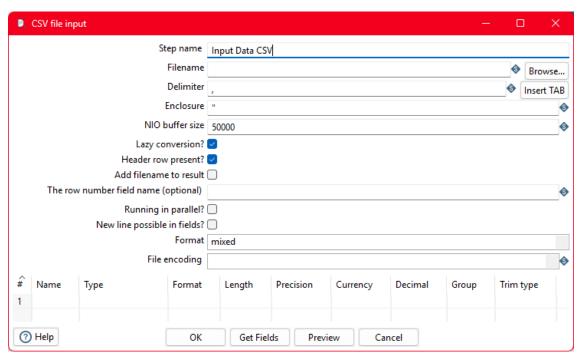
2. Cari lah objek CSV file input pada Design Area. Drag and drop objek tersebut menuju Work Area.



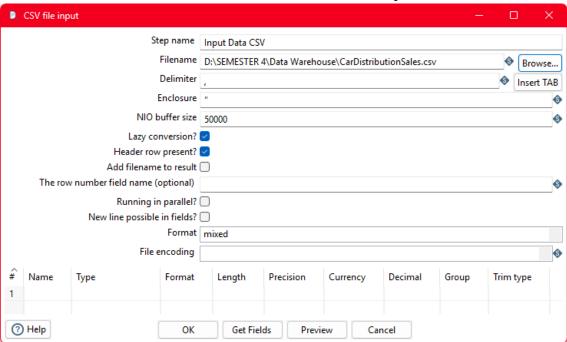
3. Double-click pada objek CSV file input hingga muncul jendela konfigurasinya.



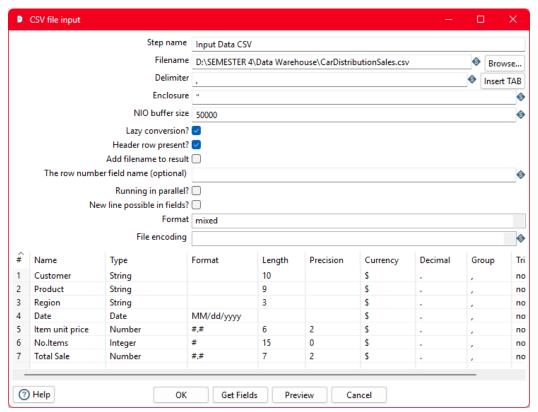
4. Ubah nama Step name menjadi Input Data CSV, (penamaan digunakan untuk identifikasi saja, selanjutnya penamaan boleh menyesuaikan).



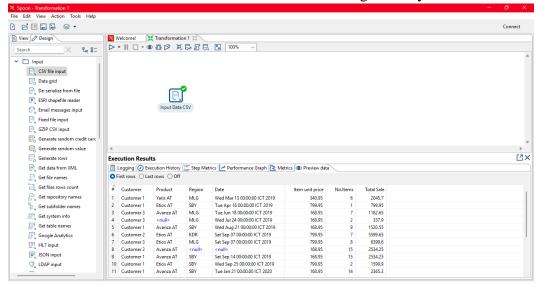
5. Pilih lokasi file CarDistributionSales.csv melalui Browse pada filename.



- 6. Jika data telah dipilih pastikan pada bagian Delimiter sesuai dengan kondisi pemisah data pada file yang dipilih. (Pada kesempatan ini adalah comma ", ").
 - Sudah sesuai
- 7. Tekan Get Fields untuk mengambil judul pada setiap kolom, pastikan nama-nama kolom telah sesuai
- 8. Tutup jendela konfigurasi tersebut.



9. Tekan tombol "Run" pada pojok kanan atas Area Kerja, maka akan muncul Execution Result Area. Pilih tab Prieview Data pada Execution Result Area. Jika proses yang dilakukan benar maka data akan muncul sesuai dengan aslinya



TUGAS 2

- 1. Apakah data hasil eksekusi sesuai dengan data aslinya?
- Iya, hasil eksekusi sudah sesuai dengan data aslinya

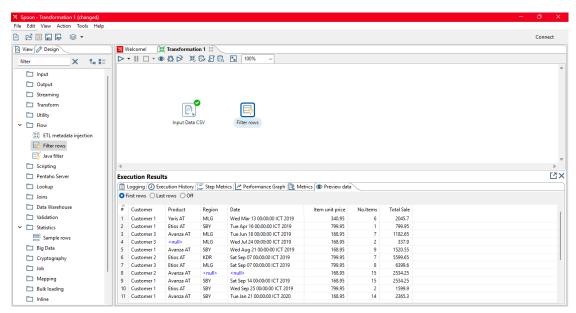
- 2. PDI Spoon melakukan proses extract, perhatikan pada Tab Logging di Execution Results Area, langkah langkah apa sajakah yang dilakukan PDI Spoon untuk melakukan extract data?
- Langkah langkah yang dilakukan untuk extract file data tersebut diantaranya adalah:
 - 1. Simpan file terlebih dahulu
 - 2. Kemudian membuka file Transformation
 - 3. Launching file Transformation
 - 4. Memulai eksekusi file Transformation
 - 5. Mengirimkan data untuk file Transformation
 - 6. Proses input data (header skipped)
 - 7. Proses input data (finishing)
- 3. Perhatikan gambar dibawah ini, apakah yang dimaksud dengan I,O,R,W,U,E? 2024/02/17 06:51:48 Input Data CSV.0 Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)
 - ⇒ Huruf huruf tersebut merepresentasikan ;
 - I (Input) adalah jumlah data yang diinput
 - O (Output) adalah jumlah data yang berhasil dikeluarkan
 - R (Rejected) adalah jumlah data yang tertolak atau gagal diproses
 - W (Written) adalah jumlah data yang berhasil tertulis
 - U (Updated) adalah jumlah data yang diperbarui
 - E (Errors) adalah jumlah kesalahan yang terjadi pada saat proses input data
- 4. 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
Adinda	Intel® Celeron® N4020 (2C / 2T, 1.1 / 2.8GHz,	8 GB	256GB	Intel® UHD Graphics 600	Intel® Celeron ® N4020	0.4s
	4MB)					
Dini	AMD Ryzen 7 5800HS with Radeon Graphics 3.20 GHz	16 GB	475GB	AMD Radeon Graphics	AMD Ryzen 7 5800HS with Radeon Graphics	0.1s
Dhevina	12th Gen Intel(R) Core(TM) i5	8 GB	152 GB	Inter(R) UHD Graphics	12th Gen Intel(R) Core(T	0.6s

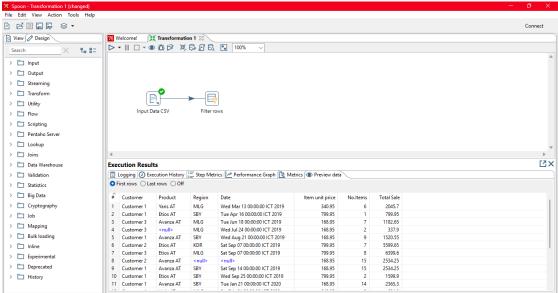
- 1235U 1.30		M) i5 -	
GHz		1235U	

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

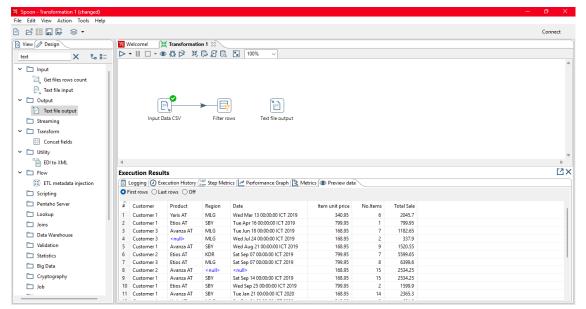
1. Carilah objek Filter rows pada Design Area. Drag and drop objek tersebut pada Work Area



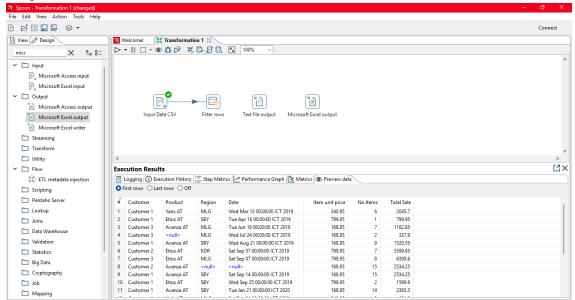
2. Hubungkan output pada Input Data CSV menuju Filter rows. Pilih koneksi Main output of step saat menghubungkan kedua objek tersebut.



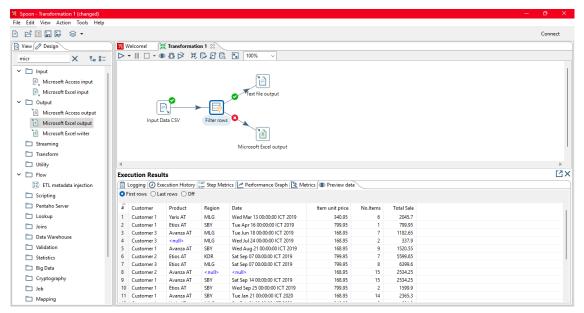
3. Carilah objek Text file output pada Design Area. Drag and drop objek tersebut pada Work Area.



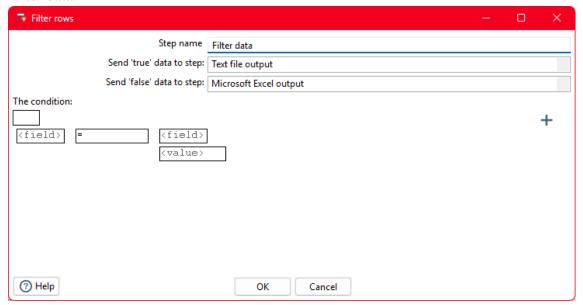
4. Carilah objek Microsoft excel output pada Design Area. Drag and drop objek tersebut pada Work Area.



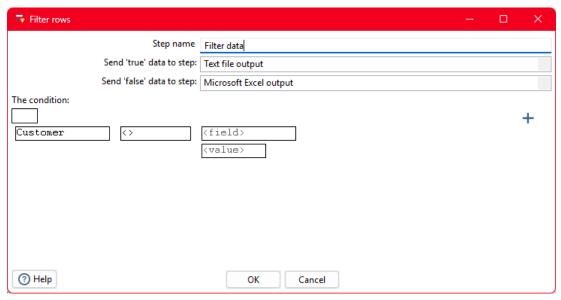
- 5. Hubungkan filter rows dengan text file output menggunakan konektor output dari filter rows. Pilih Result is true sebagi jenis konektornya.
- 6. Hubungkan filter rows dengan microsoft excel output menggunakan konektor output dari filter rows. Pilih Result is false sebagi jenis konektornya.



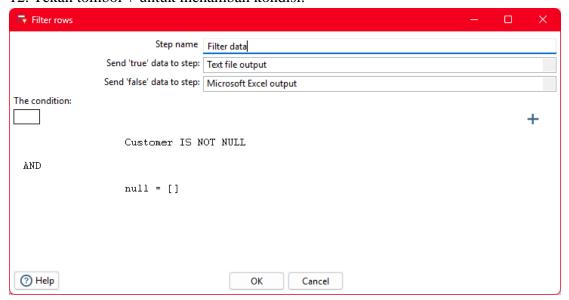
7. Double click pada filter rows hingga muncul jendela konfigurasi. ubah step name menjadi Filter data.



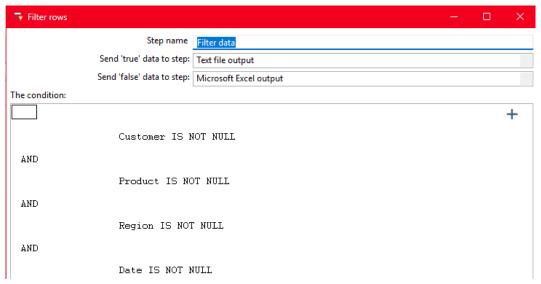
- 8. Cek bagian send 'true' data to step menjadi Text file output dan cek bagian Send 'false' data to step menjadi Microsoft Excel output.
- 9. Klik pada bagian condition
- 10. Pilih field yang diinginkan untuk digunakan sebagai kondisi. Pada kesempatan ini pilih Customer, kemudian klik OK.



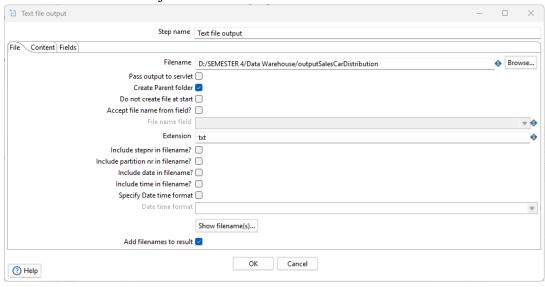
- 11. Pada bagian value pilih IS NOT NULL kemudian tekan OK , hal ini dilakukan karena data yang dipilih adalah data yang tidak Null
- 12. Tekan tombol + untuk menambah kondisi.



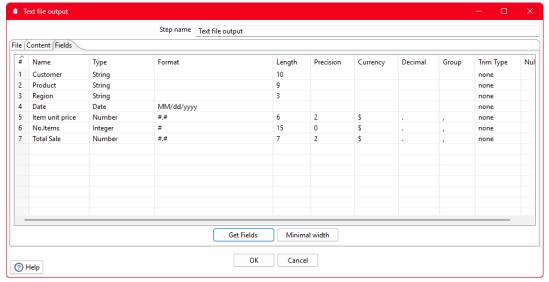
- 13. Pilih statement yang masih null untuk menambah kondisi.
- 14. Ulangi proses nomor 9,10,11 dengan mengganti semua field yang ada. hingga statement menjadi seperti pada gambar dibawah ini.
- 15. Gunakan operator AND untuk menghubung
- 16. Jika semua field sudah masuk dalam kondisi maka tekan OK.



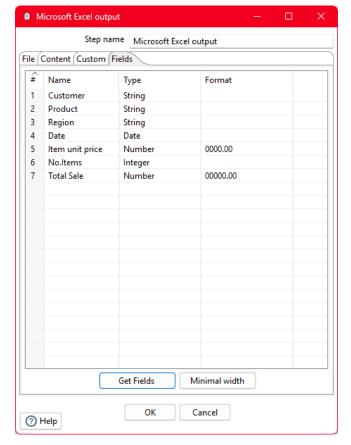
- 17. Double-click pada text file output, pilih lokasi file untuk menyimpan hasil output file dengan menekan tombol browse. beri nama outputSalesCarDistribution.
- 18. Ubah extension menjadi csv.



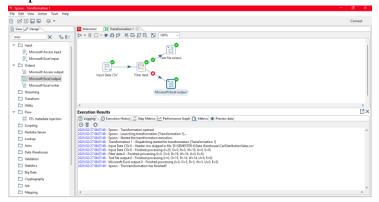
19. Pada tab Fields, tekan tombol Get fields untuk mengambil field data.



- 20. Tekan ok untuk kembali ke work area
- 21. Hal yang sama dilakukan pada konfigurasi Microsoft Excel output, Double click pada objek Microsoft Excel output hingga muncul jendela konfigurasi.
- 22. Pilih lokasi file untuk menyimpan file output dengan menekan tombol browse. Beri nama failedSalesCarDistribution. untuk extension tetap dengan extension xls.
- 23. Pada tab fields tekan tombol Get fields untuk mengambil field data.
- 24. Tekoan ok untuk kembali ke work area.



25. Tekan tombul Run pada pojok kiri atas Work Area. jika berhasil maka akan muncul tanda centang hijau di setiap proses. Proses berhasil dijalankan failedSalesCarDistribution outputSalesCarDistribution



TUGAS 3

- 1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?
 - File CSV adalah hasil extract awal sebelum dilakukan transformasi data, sedangkan file Excel sudah melewati proses pembersihan data.
- 2. Jelaskan apa yang terjadi pada proses Filter rows!
 - Filter rows memisahkan data sebelumnya menjadi dua bagian, yaitu data yang lengkap disimpan pada file csv, sedangkan data yang masih kosong atau memiliki nilai null disimpan pada file excel.
- 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

Name	Processor	RAM	Storage	VGA	CPU	Time
Adinda	Intel® Celeron®	8 GB	256GB	Intel® UHD	Intel®	0.4s
	N4020 (2C / 2T,			Graphics 600	Celeron®	
	1.1 / 2.8GHz,				N4020	
	4MB)					
Dini	AMD Ryzen 7	16 GB	475GB	AMD	AMD Ryzen	0.1s
	5800HS with			Radeon	7 5800HS	
	Radeon Graphics			Graphics	with Radeon	
	3.20 GHz				Graphics	
Dhevina	12th Gen	8 GB	152 GB	Inter(R)	12th Gen	0.6s
	Intel(R)			UHD	Intel(R)	
	Core(TM) i5 -			Graphics	Core(TM) i5 -	
	1235U 1.30 GHz				1235U	

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. 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.

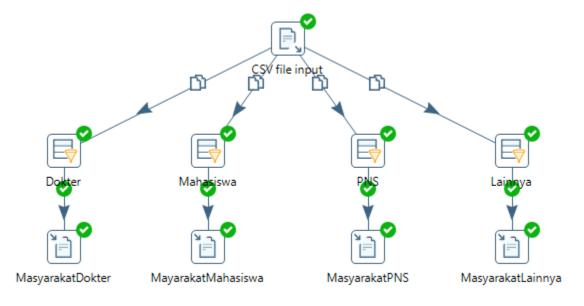
Data masyarakat berprofesi PNS akan diberikan ke Pak Yasana Laili selaku ketua RT wilayah pemerintahan

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

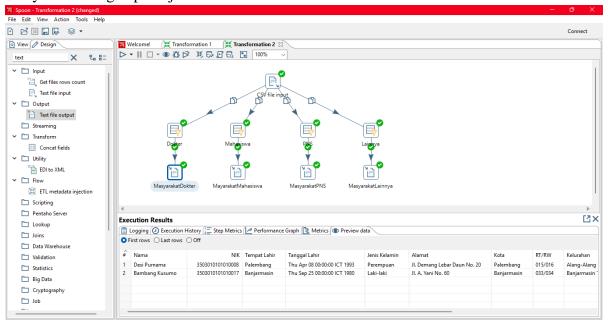
data-data tersbut akan digunakan untuk analisa di masing-masing wilayah dan juga Dashboard untuk me=mbantu mengambil keputusan.

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

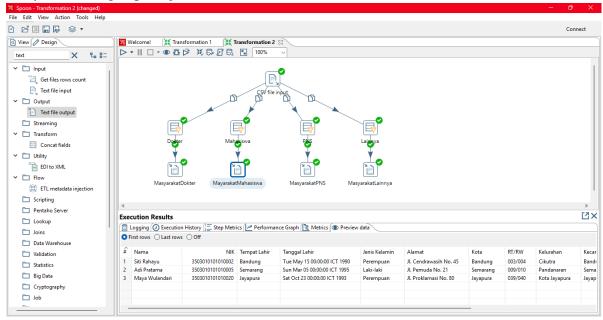
Bantulah Pak Pascanowo untuk memishkan data tersebut menjadi 4 file: MasyarakatMahasiswa, MasyarakatPNS, MasyarakatDokter, MasyarakatLainnya



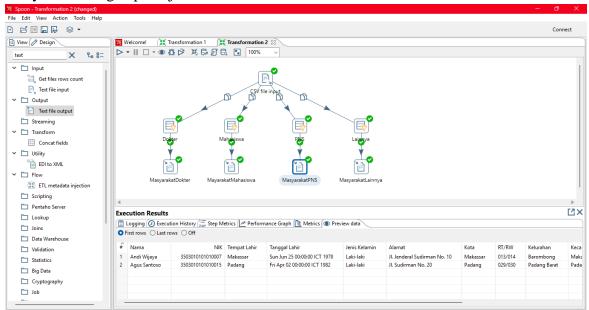
- Masyarakat dengan pekerjaan Dokter



- Masyarakat dengan pekerjaan Mahasiswa



- Masyarakat dengan pekerjaan PNS



Masyarakat dengan pekerjaan lainnya

