

<p>Nama: Rico Renaldy</p> <p>NIM: 065001800021</p>		<h2>MODUL 2</h2> <p>Nama Dosen: Ir. Teddy Siswanto, MMSi</p> <p>Hari/Tanggal: Kamis, 14 Maret 2024</p> <p>Praktikum Data Warehouse</p> <p>Nama Asisten Laboratorium: 1. Exchell S.G Ointu 065002100007 2. Nia Suhernawati – 065002100005</p>
--	---	---

Input Data pada Spoon

1. Teori Singkat

Data warehouse adalah jenis sistem manajemen data yang dirancang untuk memungkinkan dan mendukung kegiatan business intelligence (BI), terutama analitik. Gudang data semata-mata dimaksudkan untuk melakukan kueri dan analisis dan sering berisi sejumlah besar data historis. Data dalam gudang data biasanya berasal dari berbagai sumber seperti file log aplikasi dan aplikasi transaksi. Gudang data memusatkan dan mengkonsolidasikan sejumlah besar data dari berbagai sumber. Kemampuan analitisnya memungkinkan organisasi untuk memperoleh wawasan bisnis yang berharga dari data mereka untuk meningkatkan pengambilan keputusan. Seiring waktu, ia membangun catatan sejarah yang dapat sangat berharga bagi para ilmuwan data dan analis bisnis. Karena kemampuan ini, gudang data dapat dianggap sebagai "sumber kebenaran tunggal" organisasi.

2. Alat dan Bahan

Hardware : Laptop/PC

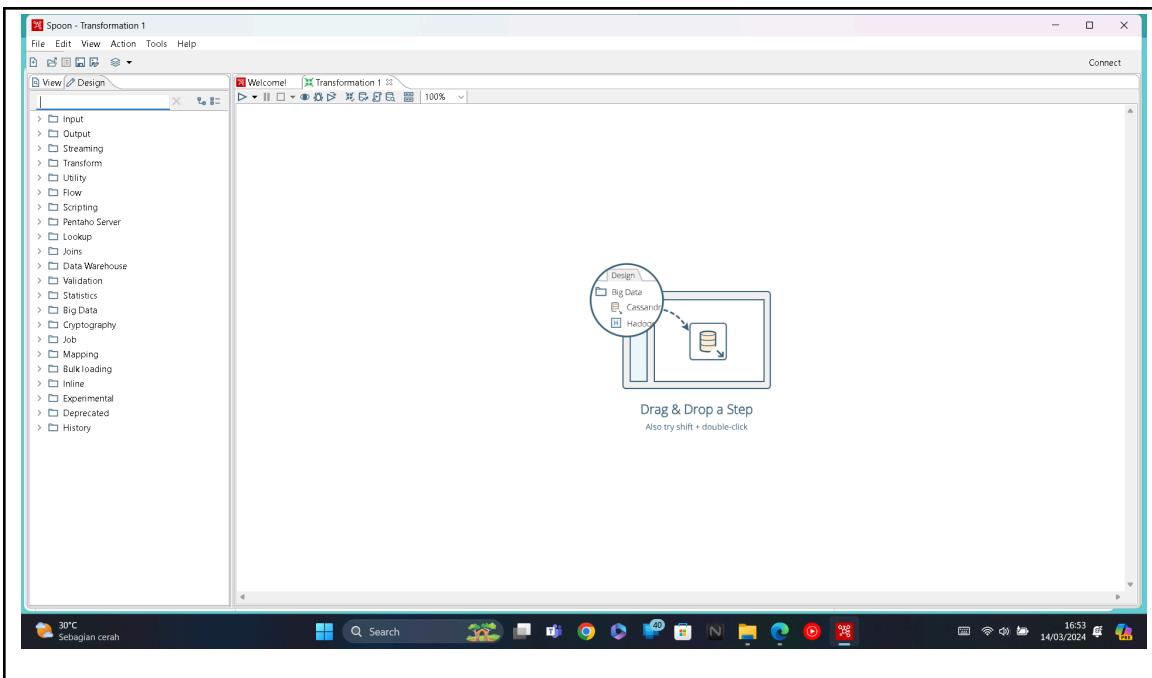
Software : Spoon Pentaho from Hitachi Vantara



3. Elemen Kompetensi

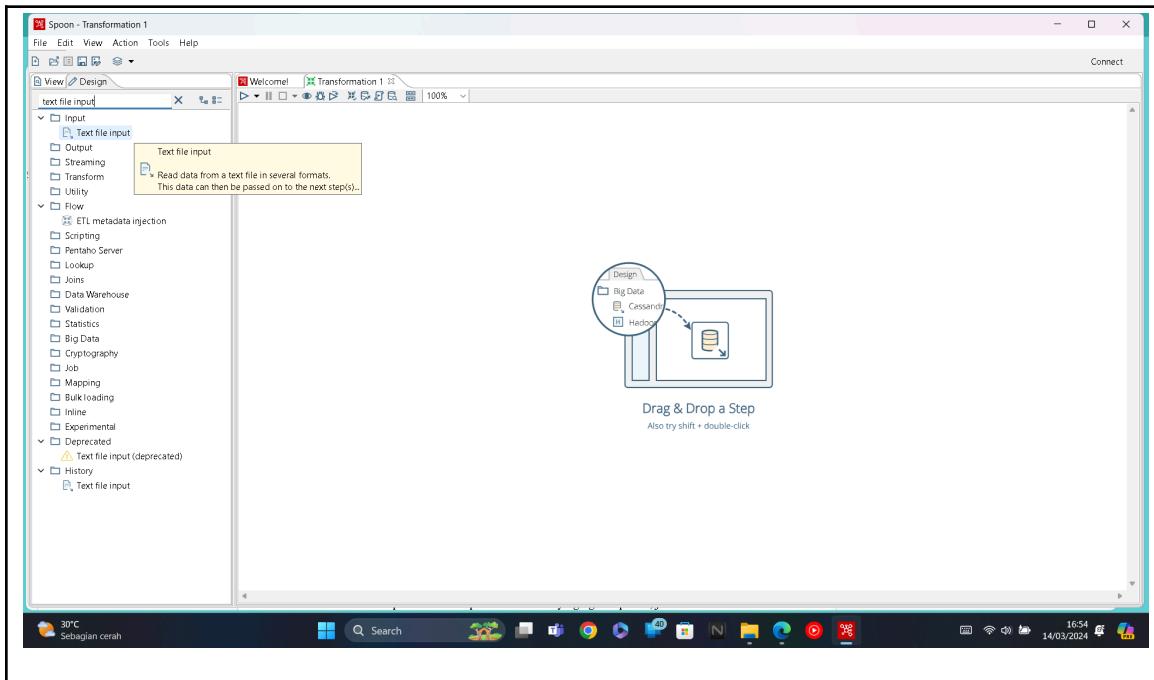
a. Latihan pertama – Input Data Text (TXT)

1. Buat transformation sheer baru



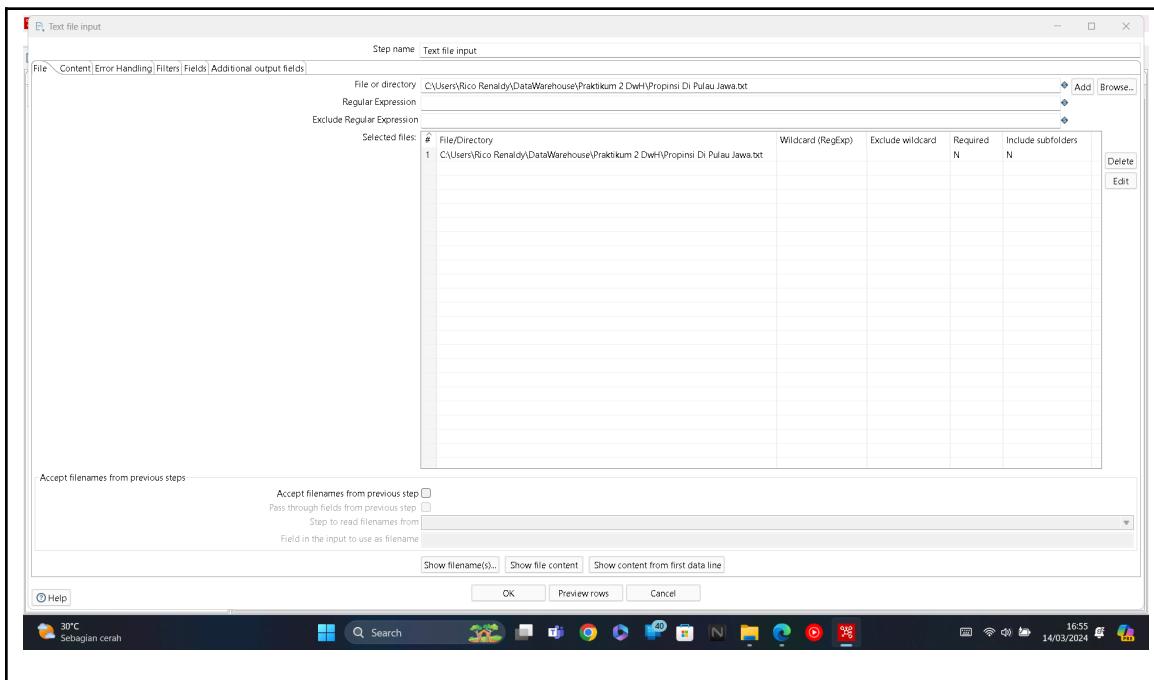
2. Cari dan pilih Text File Input.



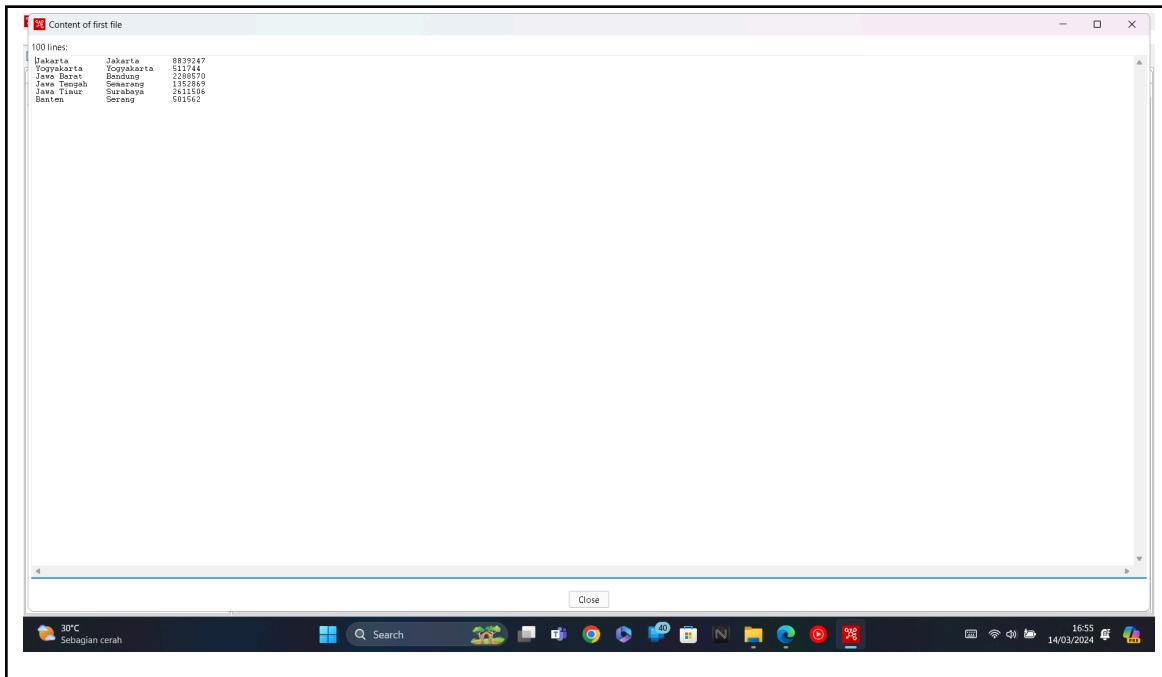


3. Klik 2x pada Text File Input lalu browse file yang ingin diinputkan, jika sudah setelah itu klik add. Maka tampilan sesudah di add akan seperti dibawah ini.





4. Jika ingin melihat preview data, maka pilih show file content, lalu akan ditampilkan preview datanya seperti gambar dibawah ini



5. Jika sudah melakukan pemeriksaan data yang diinputkan setelah itu klik OK.

The image displays two windows from the Talend Data Integration software, specifically the 'Text file input' step configuration.

Top Window (Step name: Text file input):

- File or directory:** C:\Users\Rico Renaldi\DataWarehouse\Praktikum 2 DwIT\Propinsi Di Pulau Jawa.txt
- Selected files:** 1 C:\Users\Rico Renaldi\DataWarehouse\Praktikum 2 DwIT\Propinsi Di Pulau Jawa.txt
- Buttons:** OK, Preview rows, Cancel

Bottom Window (Step name: Text file input):

- Fields Table:**

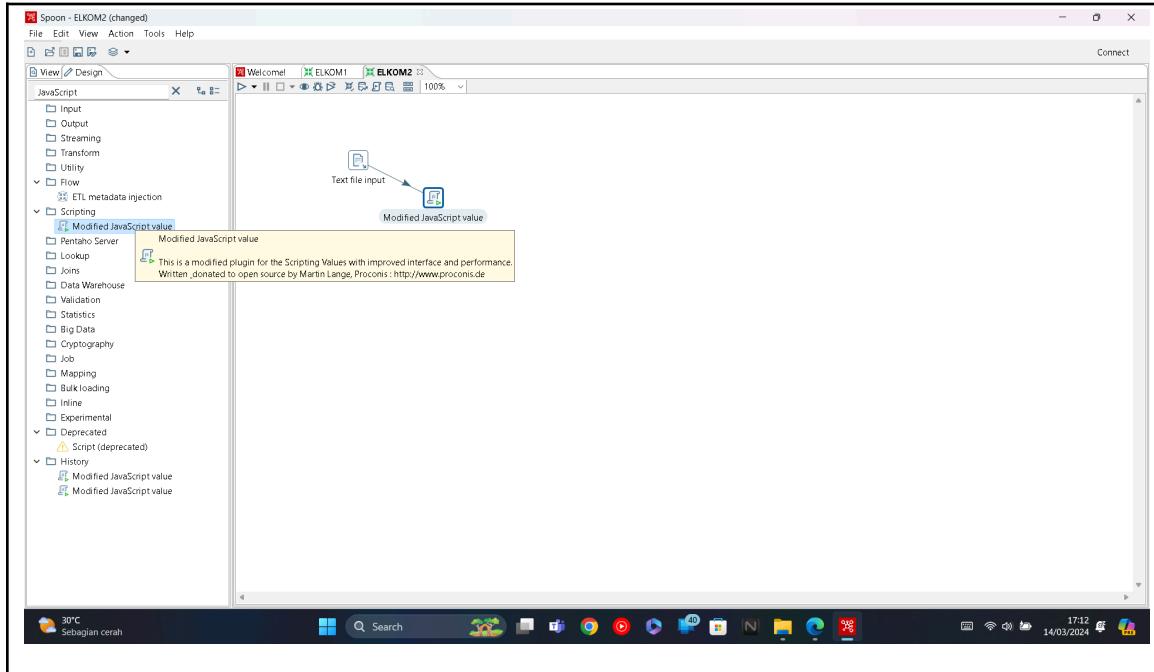
#	Name	Type	Format	Position	Length	Precision	Currency	Decimal	Group	Null if	Default	Trim type	Repeat
1	Field1	String		36		\$,	,	-			none	N

- Buttons:** Get Fields, Minimal width, OK, Preview rows, Cancel

b. Latihan Kedua – Menginput Data Text yang Tidak Terstruktur

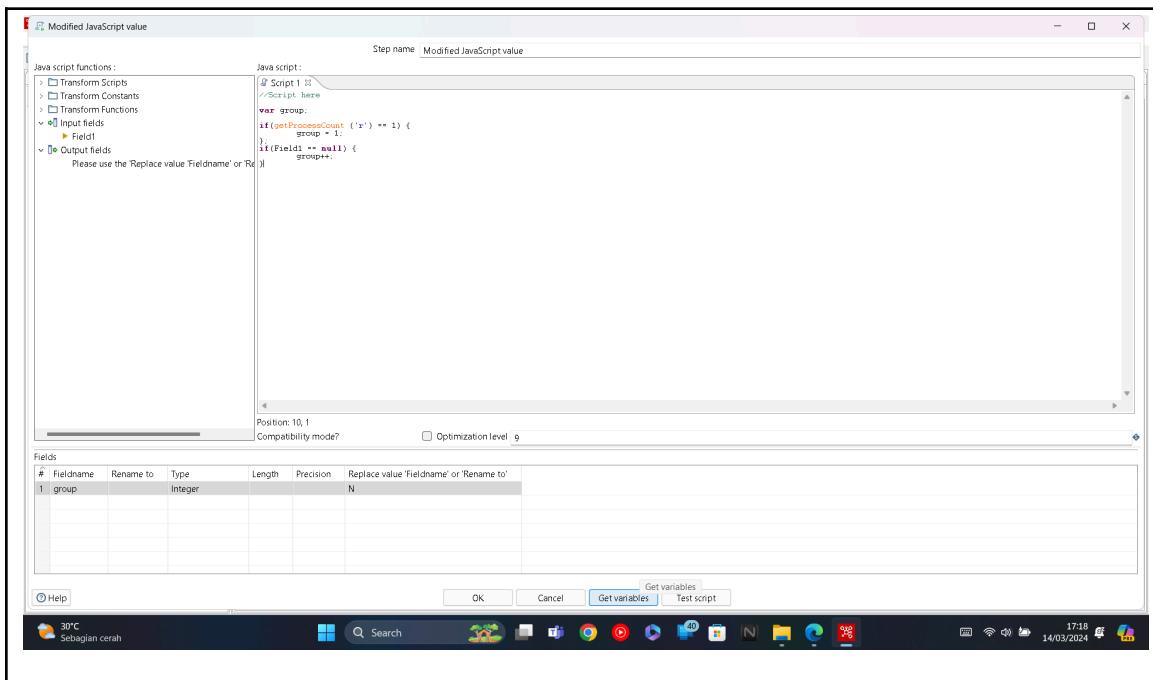


1. Ulangi semua langkah di **Latihan Pertama** dengan menginputkan data yang berbeda. dan setelah itu cari dan tambahkan Modified JavaScript value lalu hubungkan dengan Hop.

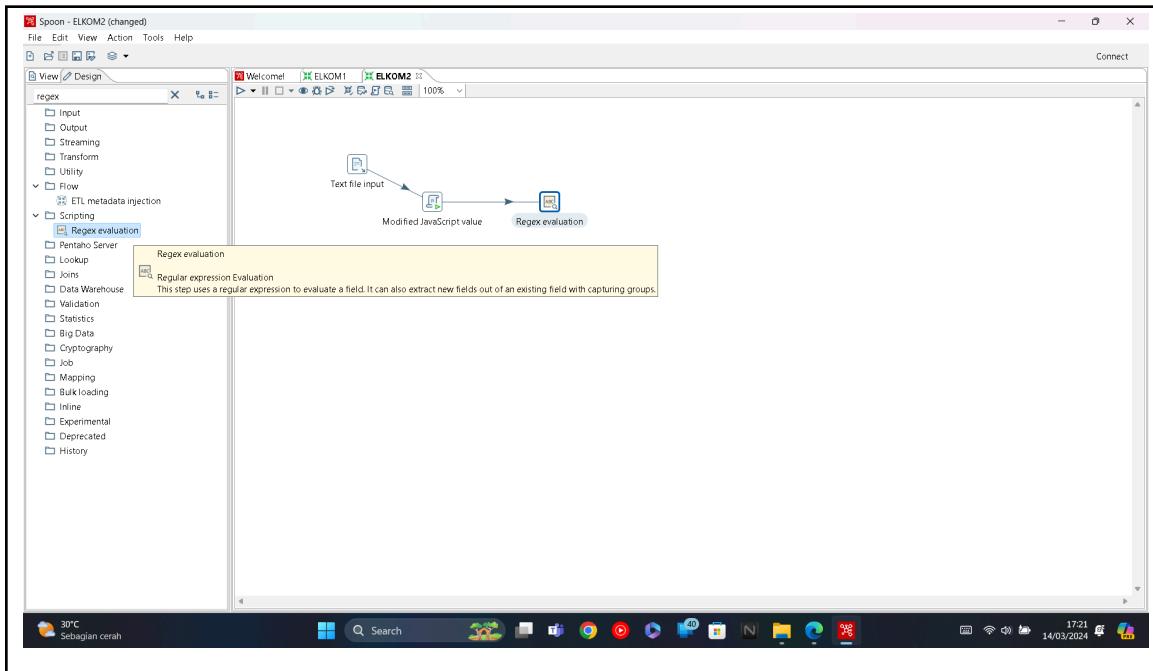


2. Selanjutnya tulis script seperti pada gambar dibawah dan klik Get variables setelah itu ubah type data pada field group menjadi Integer. Setelah itu klik OK.

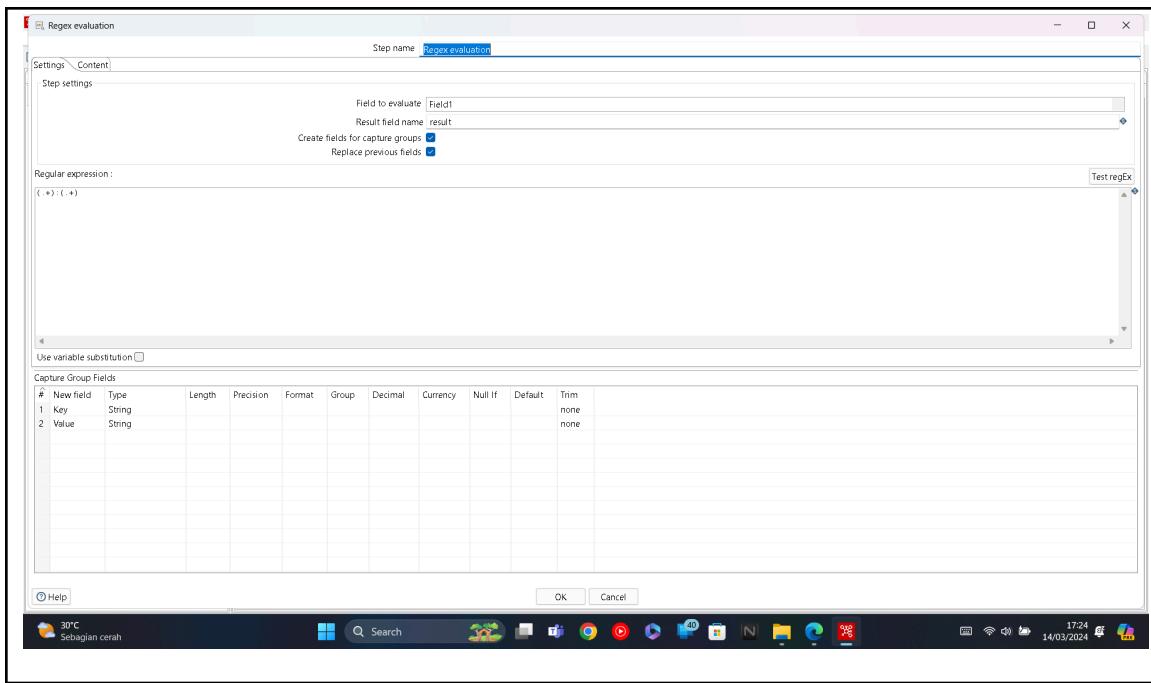




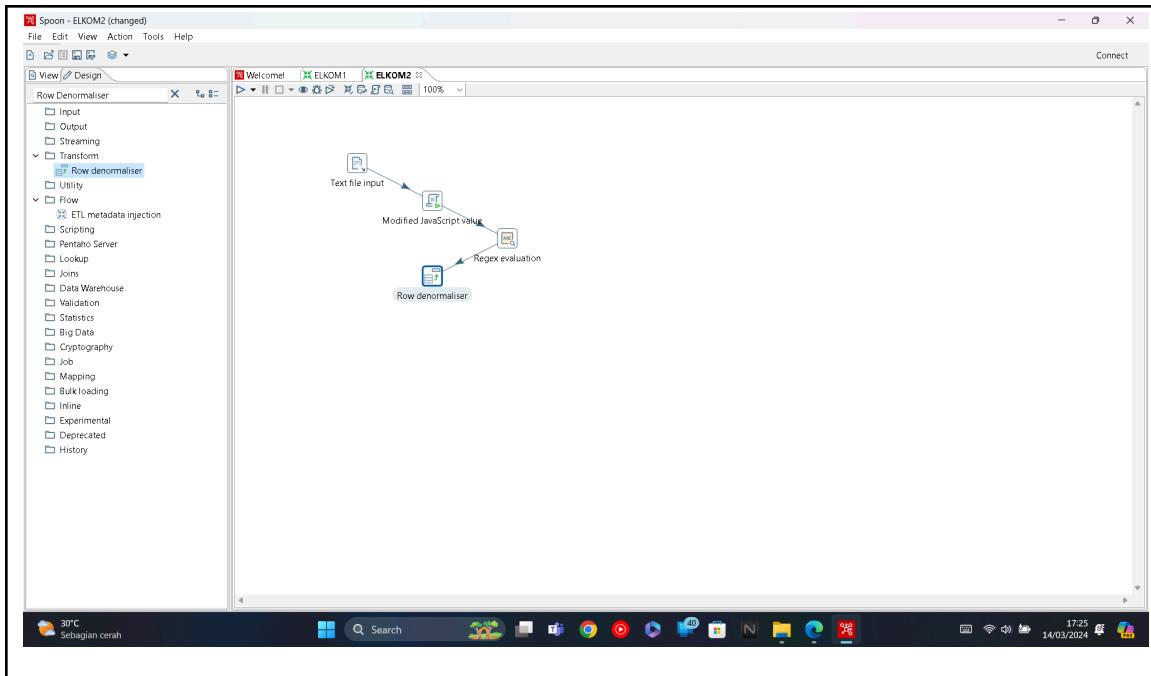
3. Selanjutnya cari dan tambahkan Regex evaluation. Setelah itu klik 2x pada Regex Evaluation



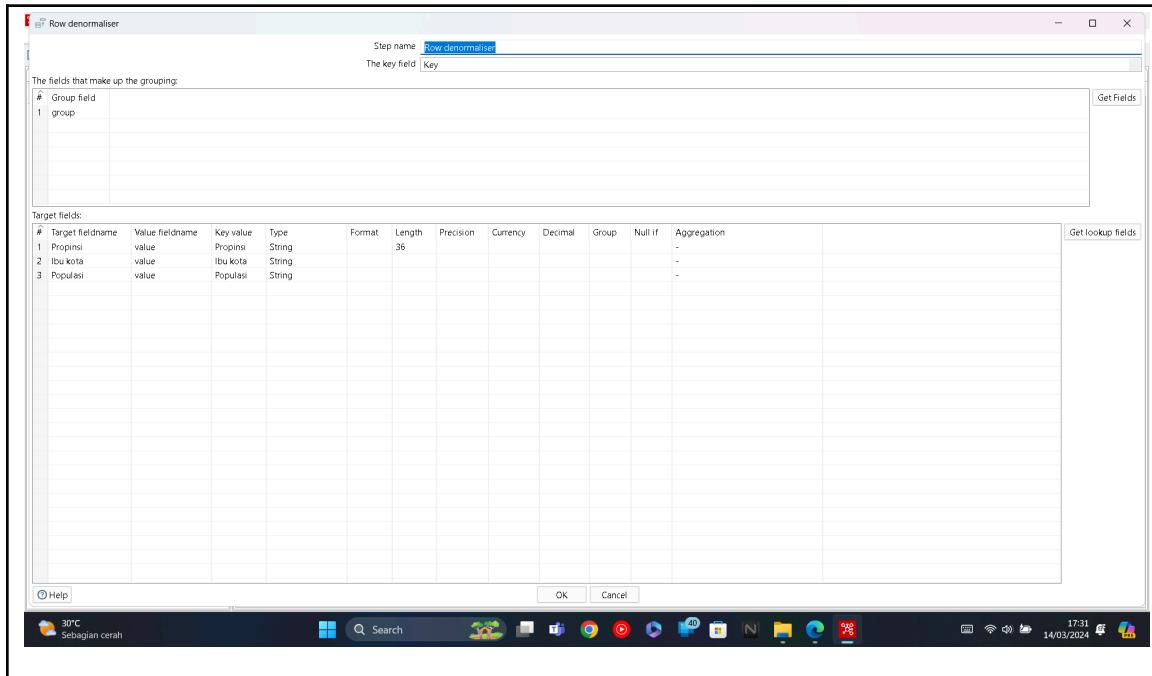
4. Konfigurasikan bagian Regex Evaluation seperti gambar dibawah ini lalu setelah itu klik OK.



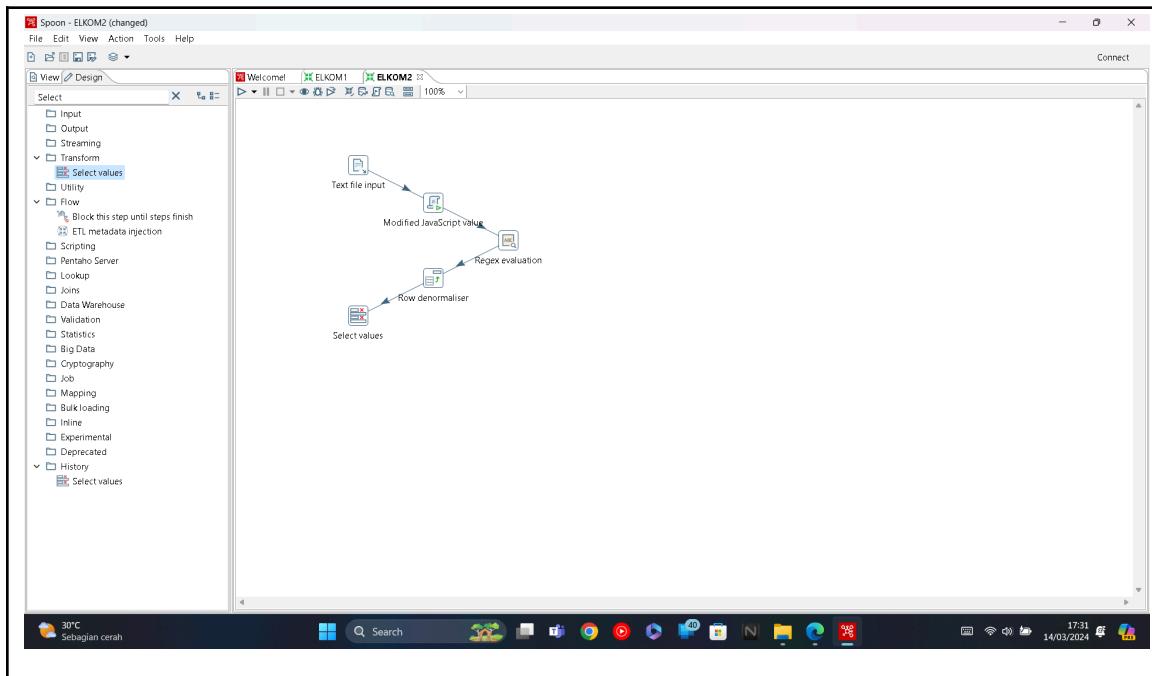
5. Setelah itu cari dan tambahkan Row denormaliser. Lalu klik 2x pada Row denormaliser yang sudah kita tambahkan.



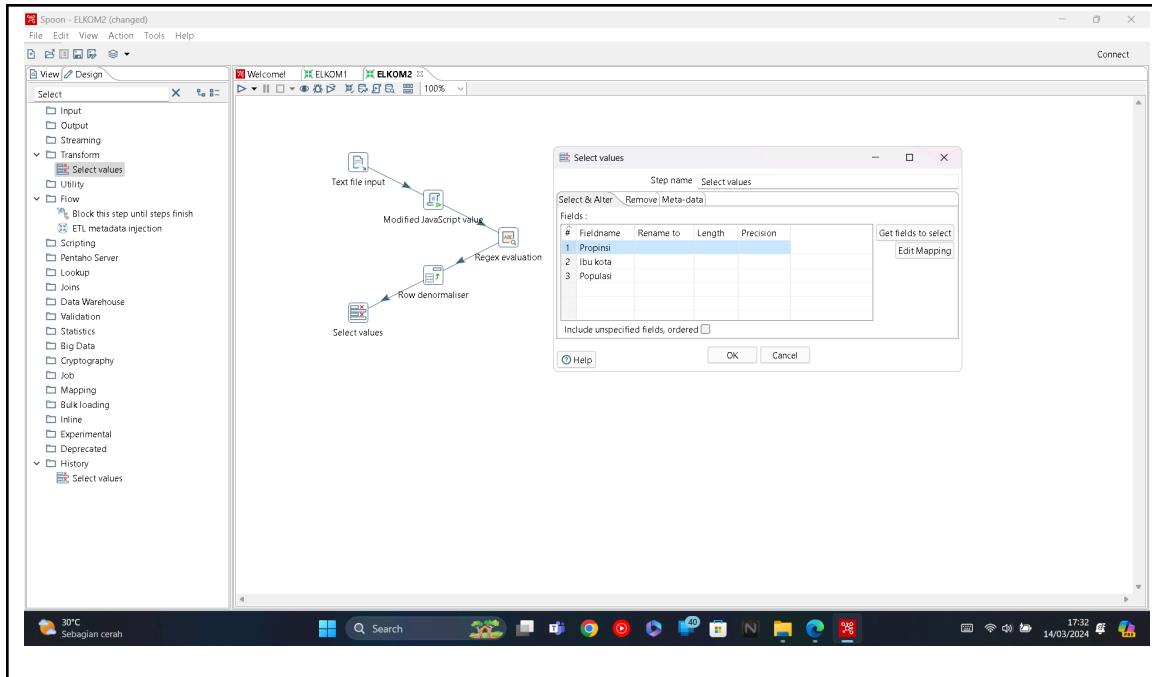
6. Konfigurasikan Row normaliser seperti pada gambar dibawah ini lalu klik OK.



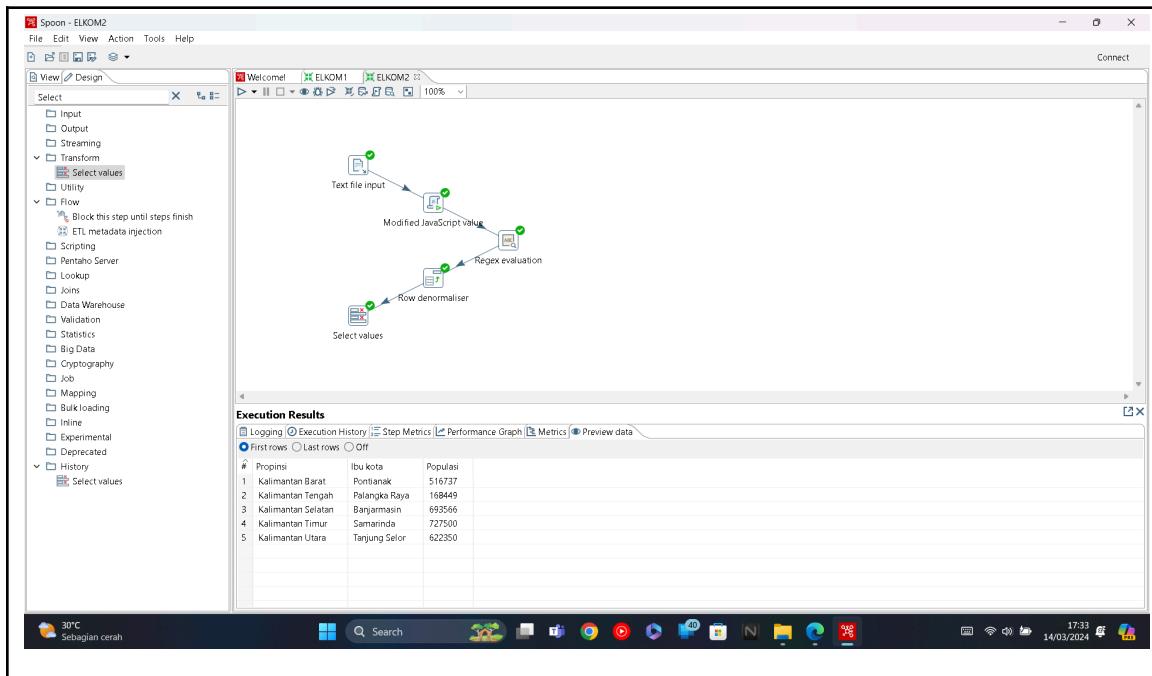
7. Setelah itu cari dan pilih Select values. Dan klik 2x pada Select values yang sudah ditambahkan.



8. Setelah itu klik get values dan hapus variabel yang tidak diperlukan dan sisakan variabel dibawah ini setelah itu klik OK. Setelah itu Jalankan/running

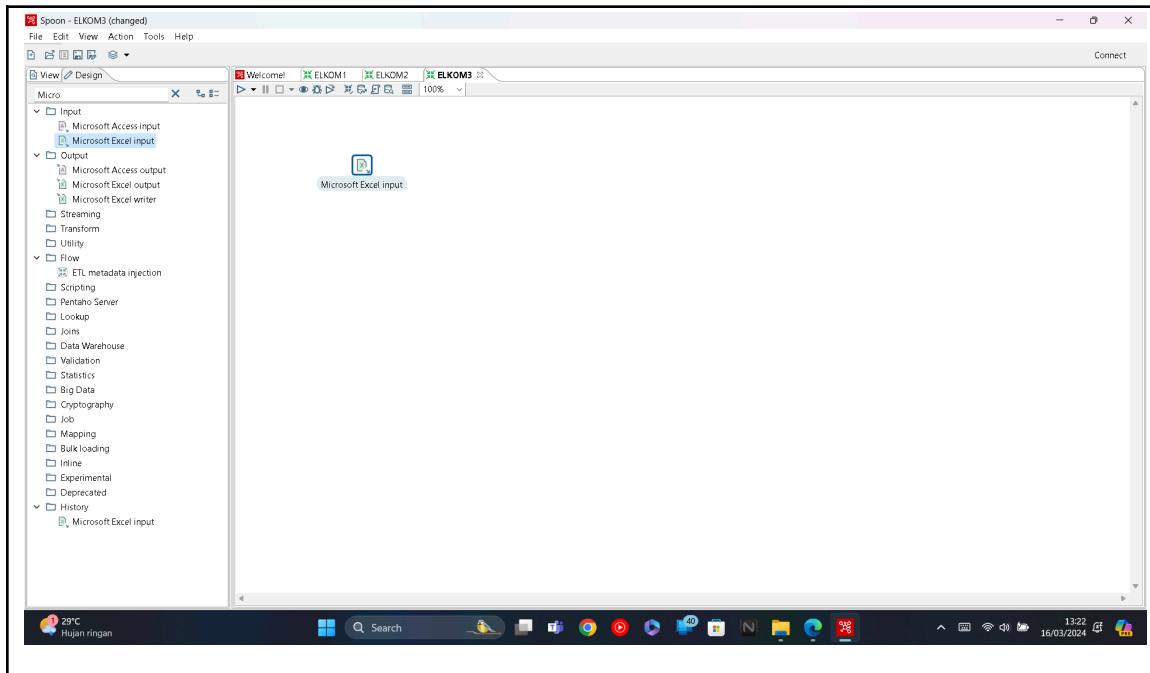


9. Berikut hasil setelah sudah dijalankan dan di running.



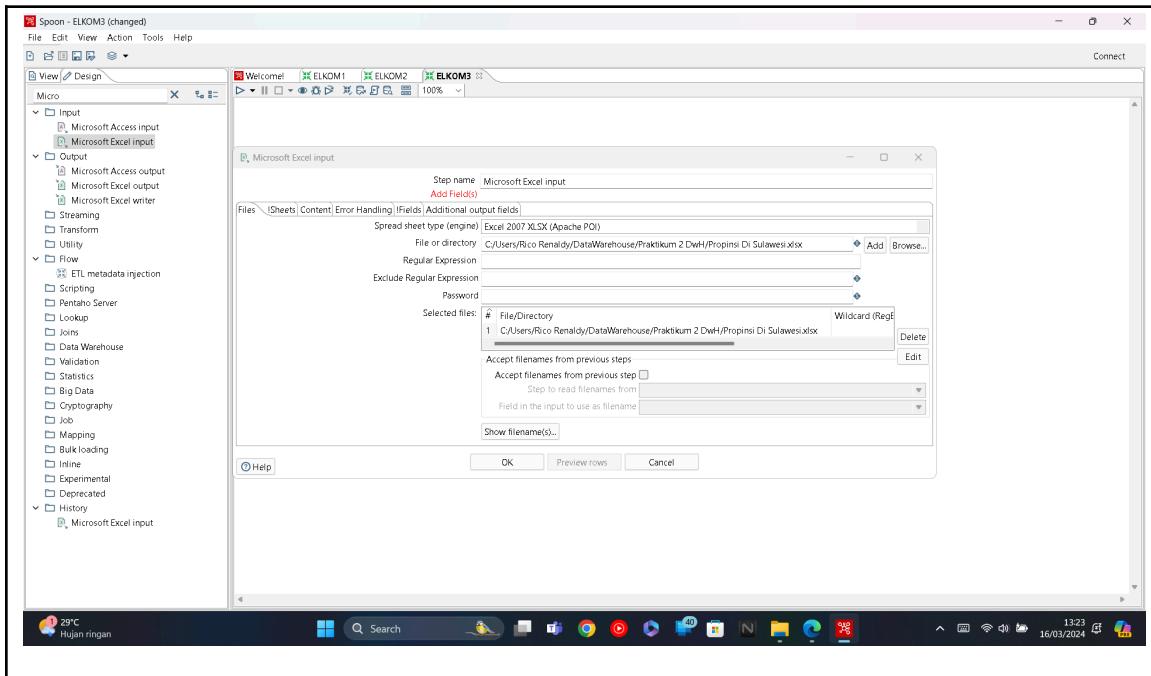
c. Latihan Ketiga – Menginput File Microsoft Excel

1. Buat transformation sheet baru dan cari serta tambahkan Microsoft Excel Input. Setelah itu klik 2x pada Microsoft Excel Input yang sudah ditambahkan tadi.

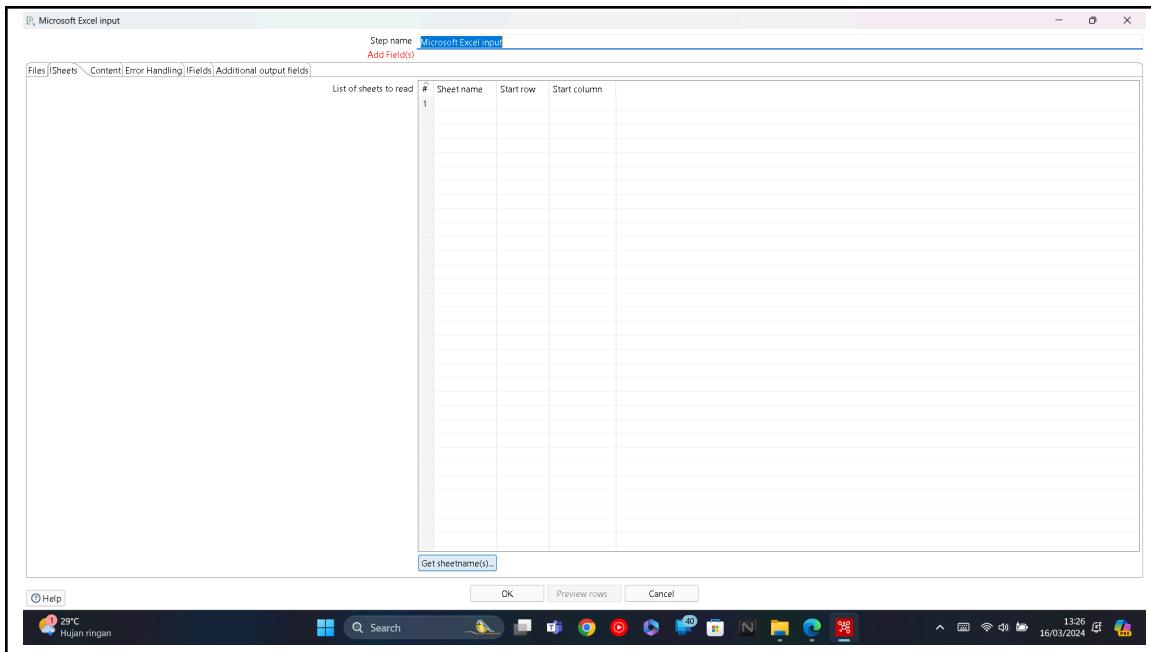


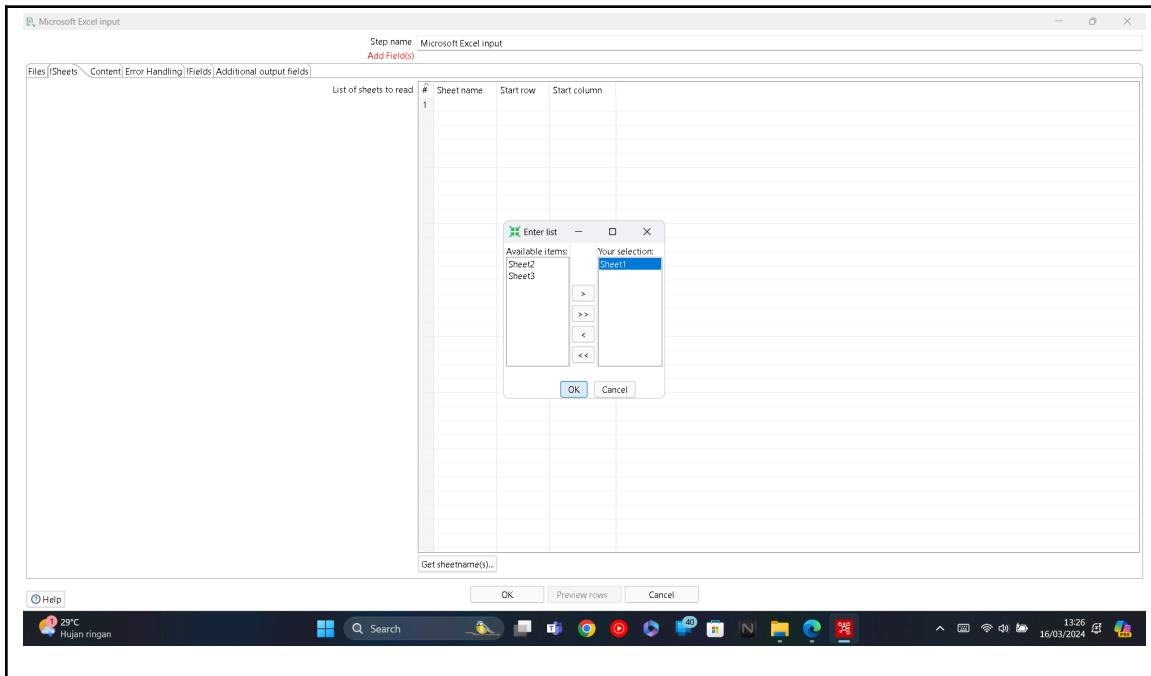
2. Browse serta add data yang ingin ditambahkan. Selanjutnya buka tab !Sheet.



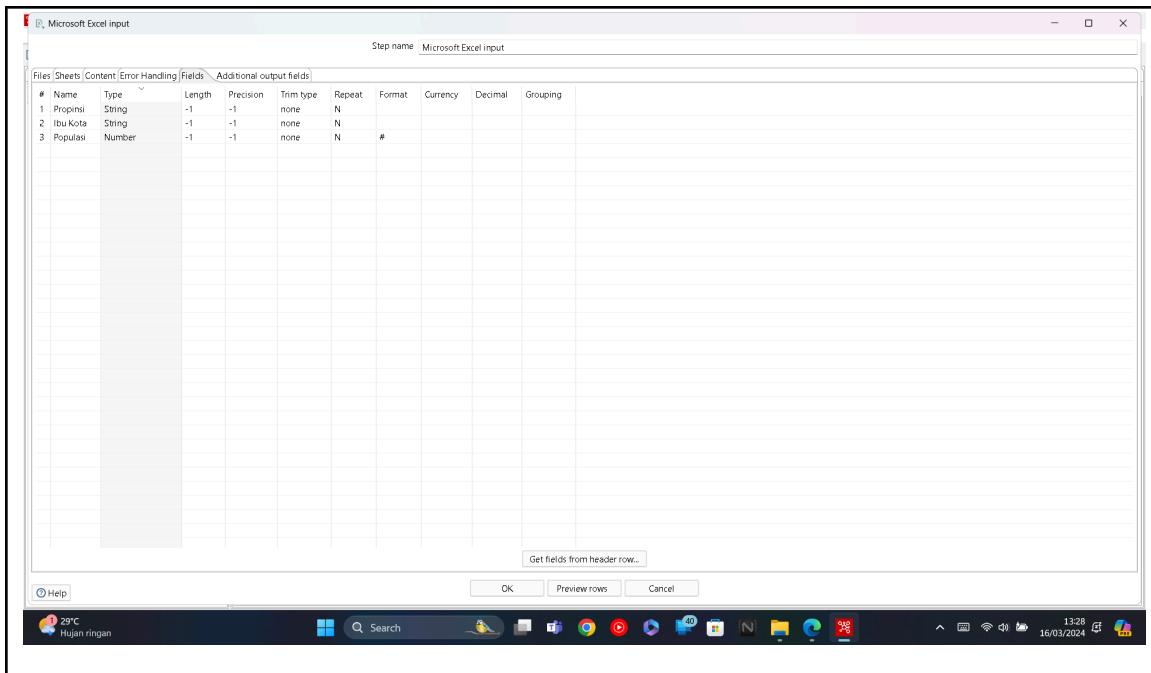


3. Klik get sheetsname lalu tambahkan Sheet yang berisi data (Sheet1) dan jika sudah klik OK dan pindah ke tab fields.

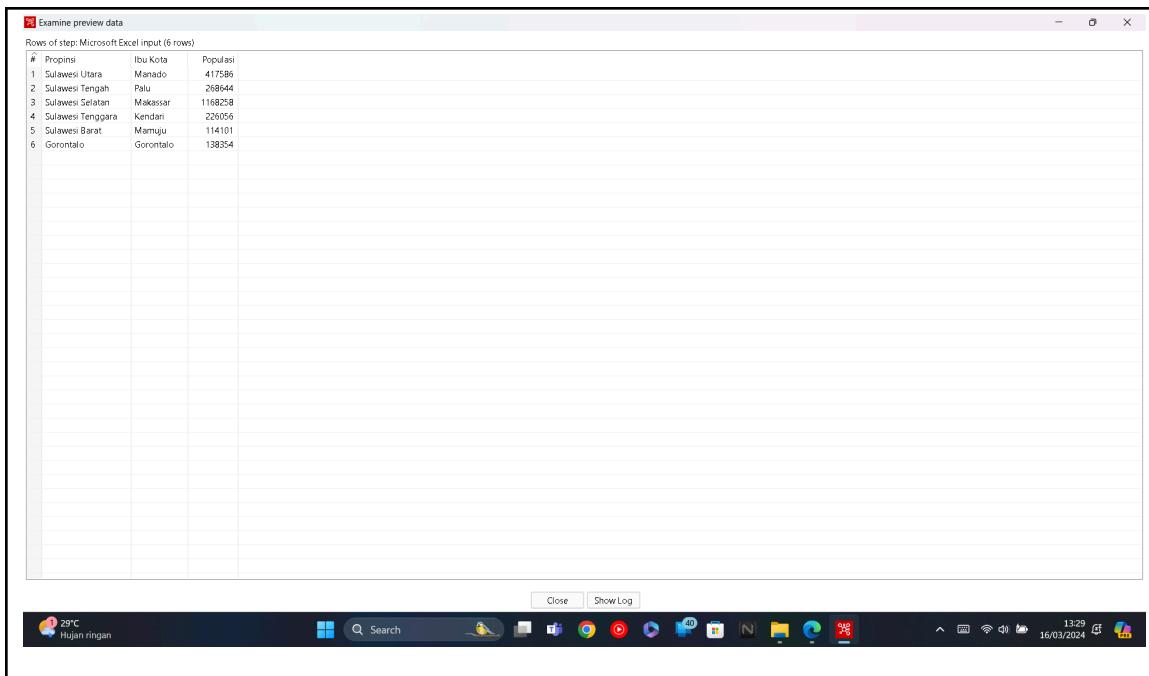




4. Lalu klik get fields from header row. Seperti pada gambar dibawah ini. Lalu setelah itu klik Preview rows.

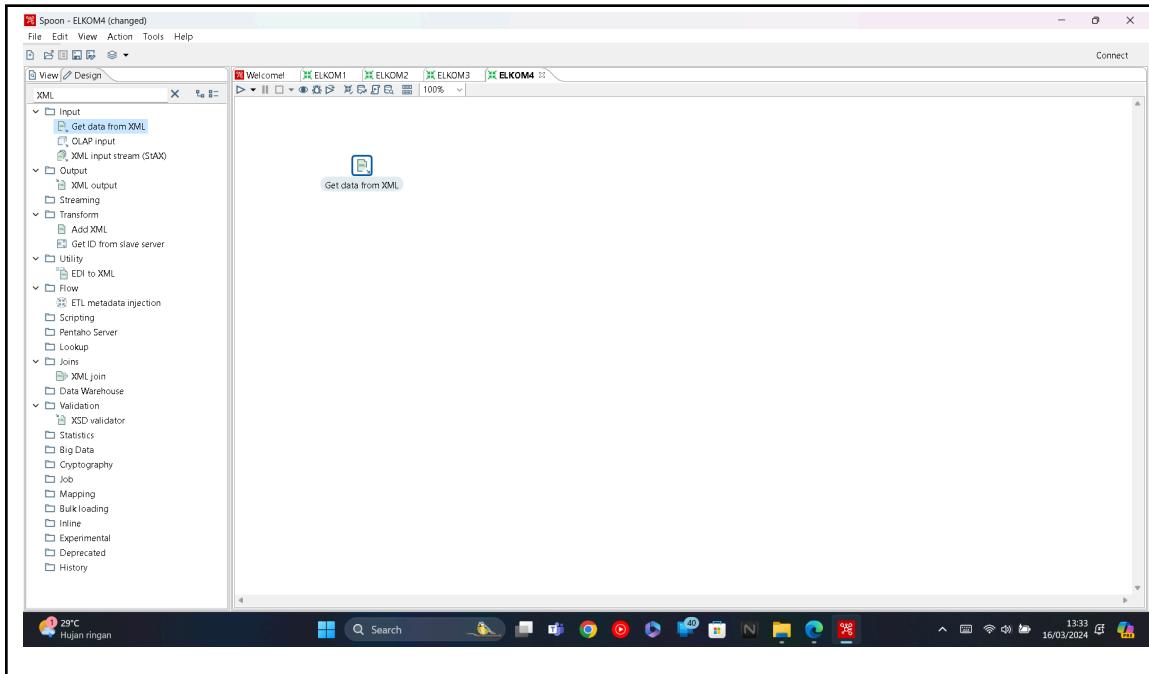


5. Berikut tampilan preview rows yang artinya data berhasil diinput. Jika sudah klik close dan klik OK.

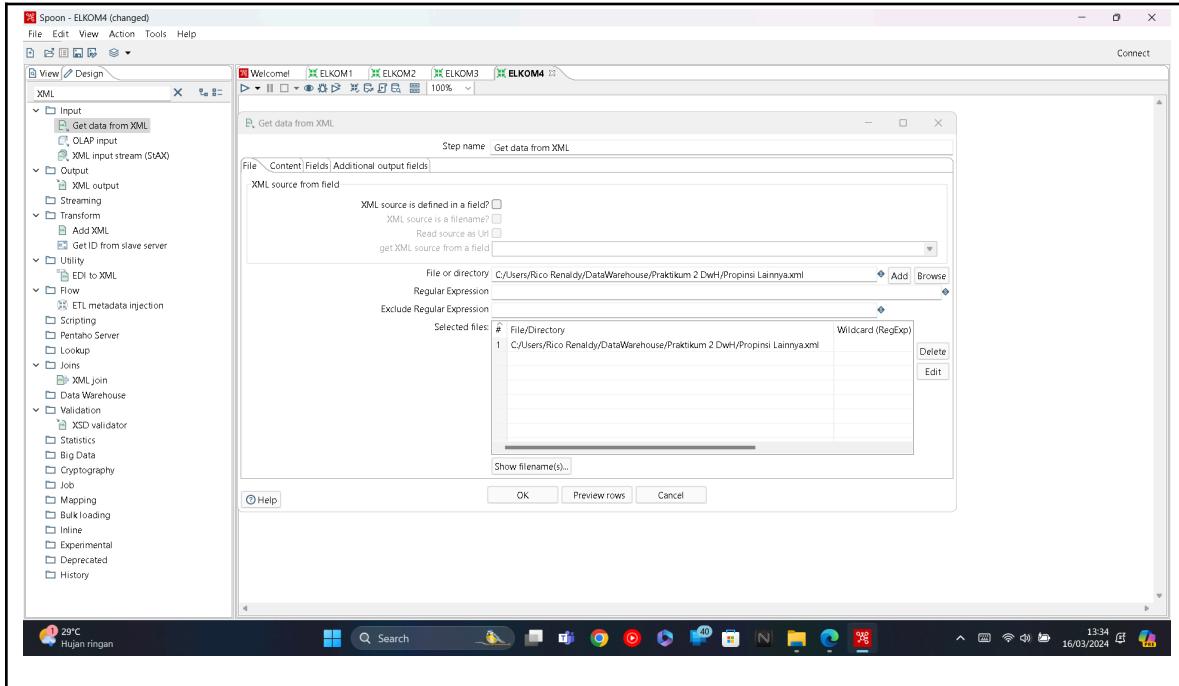


d. Latihan Keempat – Menginput File XML

1. Buat transformation sheet baru dan cari serta tambahkan Get Data From XML pada transformation. Setelah itu klik 2x pada Get Data From XML yang sudah ditambahkan.

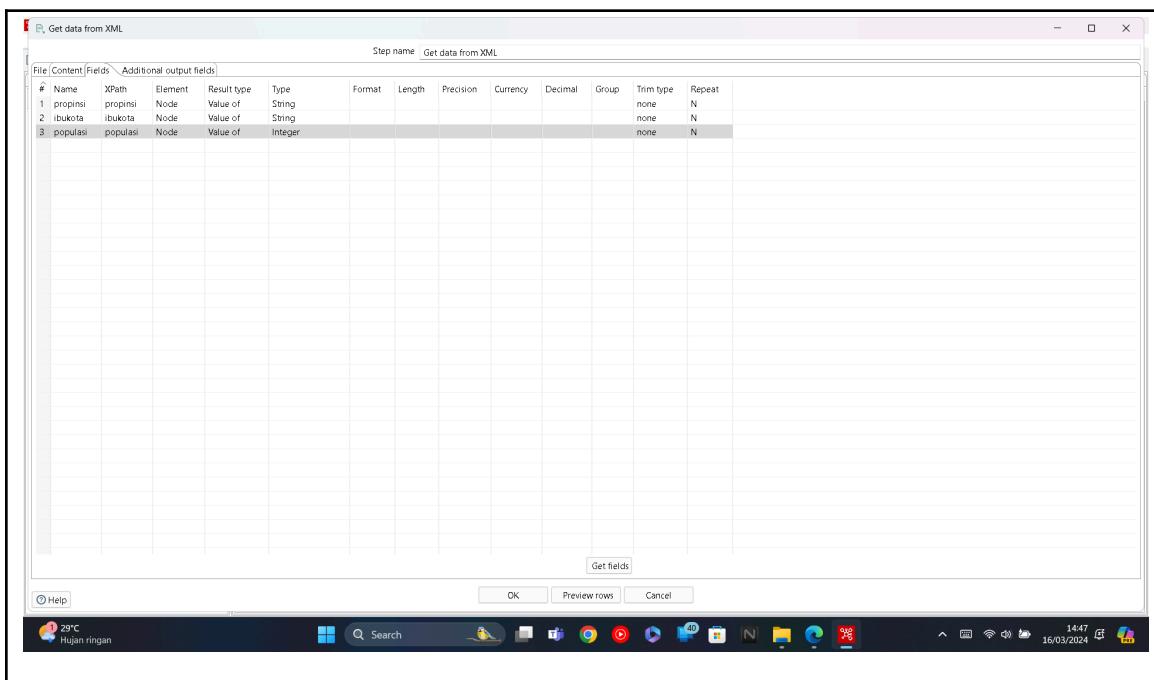


2. Browse dan cari data file XML yang ingin diinput lalu add atau tambahkan, setelah itu buka tab Fields.



3. Lalu klik Get Fields dan isi beberapa kolom data yang kosong seperti pada gambar dibawah. Lalu selanjutnya klik Preview rows.

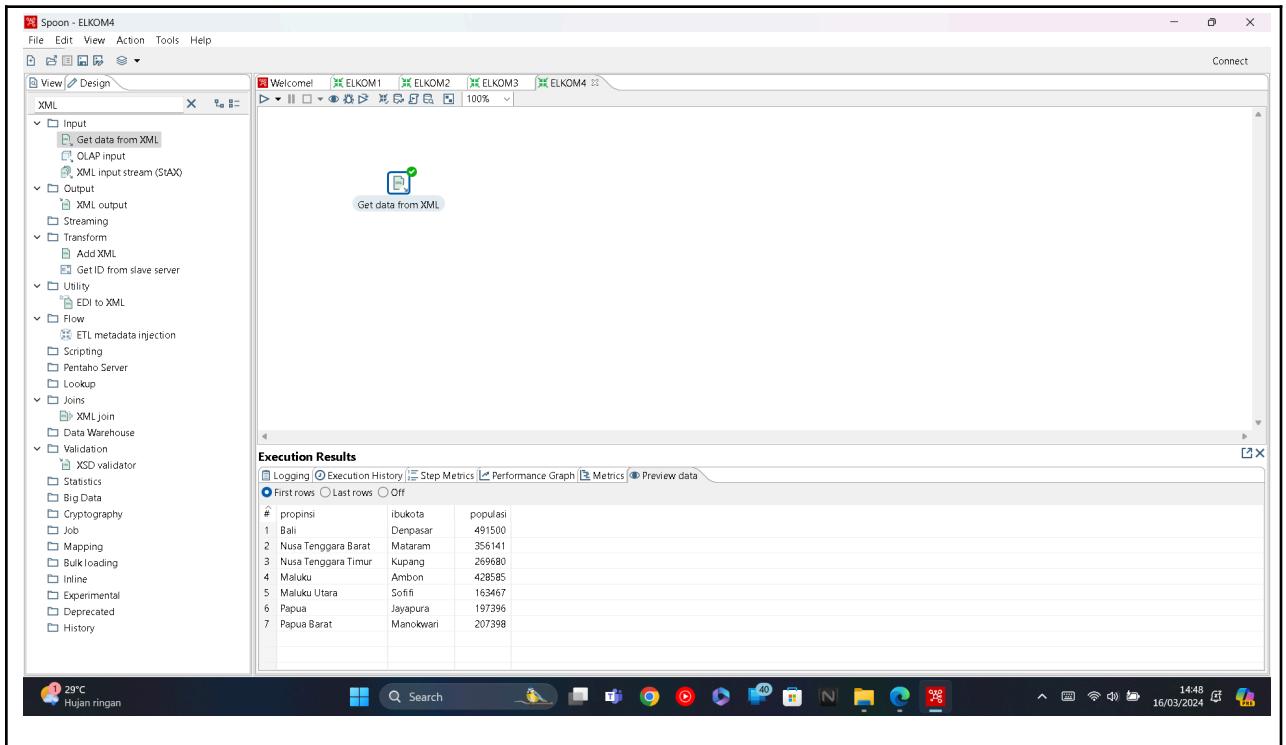




4. Jika data yang diinputkan sudah berhasil ditampilkan maka klik close dan setelah itu klik OK, lalu setelah itu di Run.

Rows of step: Get data from XML (7 rows)			
#	propinsi	ibukota	populasi
1	Bali	Denpasar	491500
2	Nusa Tenggara Barat	Mataaram	356141
3	Nusa Tenggara Timur	Kupang	269680
4	Maluku	Ambon	428585
5	Maluku Utara	Sofifi	163467
6	Papua	Jayapura	197396
7	Papua Barat	Manokwari	207395





4. Soal Latihan

Soal:

1. Apa yang dimaksud Input Data?
2. Apa yang dimaksud data XML dan kenapa XML dijadikan sebagai format penyimpanan data?

Jawaban:

1. Input data merujuk pada informasi atau nilai yang dimasukkan ke dalam suatu sistem, program komputer, atau proses untuk pengolahan lebih lanjut. Input data bisa berupa teks, angka, gambar, suara, atau bentuk data lainnya yang relevan dengan konteksnya.
2. XML adalah singkatan dari eXtensible Markup Language. Ini adalah bahasa markup yang dirancang untuk menyimpan dan mengirimkan data terstruktur. XML dijadikan format penyimpanan data karena beberapa alasan:

- Keterbacaan manusia: XML menggunakan tag-tag yang mudah dibaca oleh manusia, membuatnya mudah dipahami dan dikelola tanpa perlu alat khusus.
- Keterstrukturkan data: XML memungkinkan penggunaan hierarki yang jelas untuk data, memfasilitasi organisasi dan representasi yang sistematis.



- Keterinteroperasian: XML adalah standar terbuka yang banyak didukung oleh berbagai platform dan bahasa pemrograman. Ini membuatnya cocok untuk pertukaran data antar sistem yang berbeda.
- Kemampuan ekstensibilitas: XML memungkinkan pengguna untuk mendefinisikan struktur data mereka sendiri menggunakan Document Type Definition (DTD) atau XML Schema, yang memungkinkan adaptasi dengan kebutuhan aplikasi yang berbeda.

5. Kesimpulan

- a. Dalam pelaksanaan praktikum Data Warehouse, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
- b. Kita dapat mengetahui bahwa input data merupakan langkah krusial dalam proses pengolahan data di Spoon. Dengan input data yang akurat dan tepat, kita dapat memastikan bahwa analisis dan transformasi data yang dilakukan di dalam Data Warehouse dapat berjalan dengan baik dan menghasilkan keluaran yang relevan dan bermakna. Oleh karena itu, penting untuk memahami sumber data yang digunakan, format data yang dibutuhkan, dan proses input yang sesuai untuk memastikan keberhasilan dalam pembangunan dan pengelolaan Data Warehouse.

6. Cek List (✓)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	✓	
2.	Latihan Kedua	✓	
3.	Latihan Ketiga	✓	
4.	Latihan Keempat	✓	

7. Formulir Umpam Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	20 Menit	1
2.	Latihan Kedua	20 Menit	1
3.	Latihan Ketiga	20 Menit	1



4.	Latihan Keempat	20 Menit	1
----	-----------------	----------	---

Keterangan:

1. Menarik
2. Baik
3. Cukup
4. Kurang

