


Nama: Muhammad Rasya Hadziq NIM: 065001900033	 Praktikum Data Warehouse	MODUL 12 Nama Dosen: Ir. Teddy Siswanto, MMSi
Hari/Tanggal: Jum'at, 8 Juli 2022		Nama Asisten Labratorium: 1. Azhar Rizki Zulma 065001900001 2. Nadiya Amanda Rizkania 064001900003

Proyek Akhir Praktikum IV

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.

ETL Gabungan

ETL adalah singkatan dari Extract, Transform, Load. Sesuai namanya, Extract, Transform, Load merupakan proses yang meliputi ekstrak data dari beberapa sistem sumber, yang berlanjut kepada transformasi data (proses penghitungan, pembulatan, dan sebagainya) hingga akhirnya data ditampung (load) ke dalam sistem data warehouse. Data-data yang semula bercampur diolah menjadi data tunggal yang konsisten sehingga mudah untuk dilakukan analisis ketika dibutuhkan nantinya.

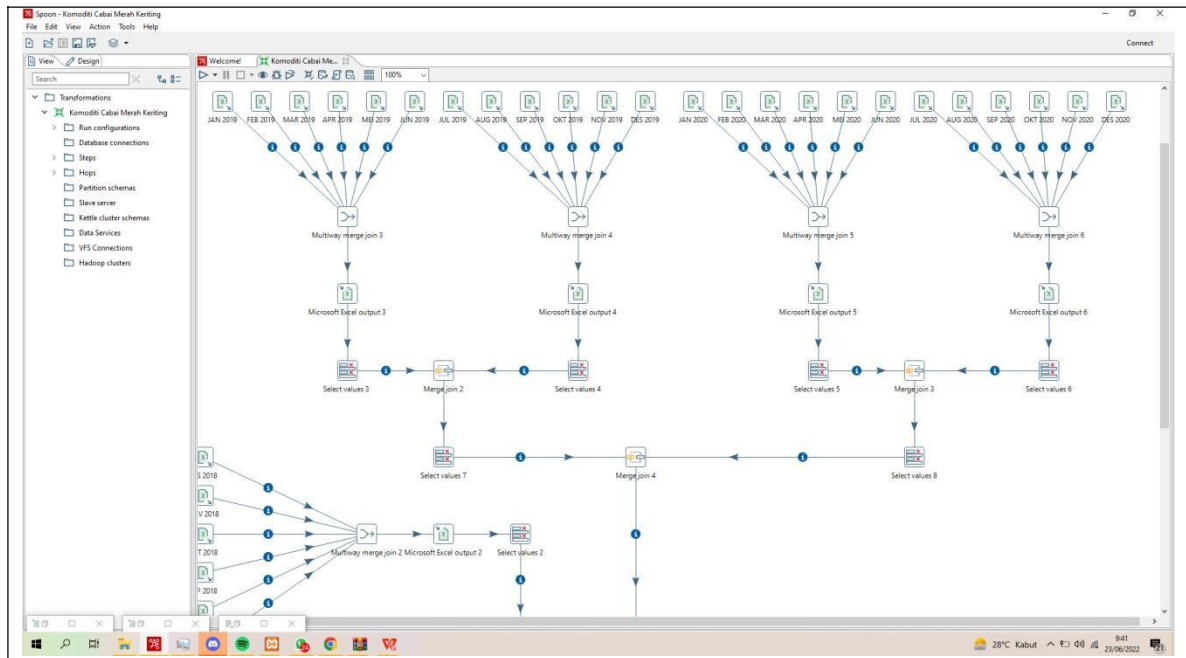
ETL menjadi pondasi untuk data analytics dan machine learning. Melalui sejumlah pengaturan, ETL dapat membersihkan dan mengorganisir data sesuai dengan yang diinginkan. Data laporan bulanan pun akan dapat dengan mudah dianalisis.



Berikut ini adalah langkah dalam melakukan ETL Gabungan menggunakan Spoon Pentaho:

1. Struktur

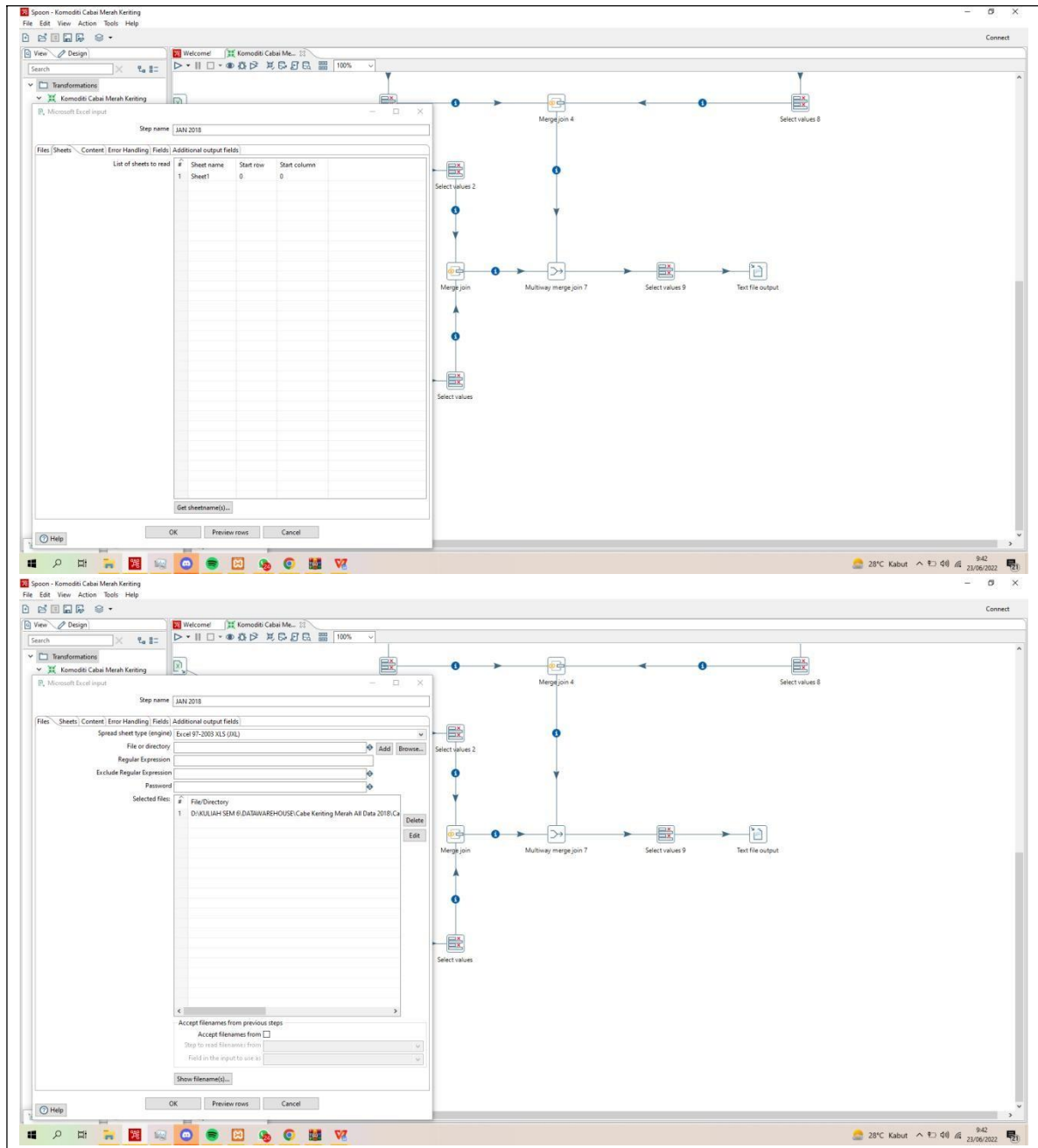
Buka Pentaho Spoon Data Integration lalu buat struktur seperti pada gambar dibawah ini lalu simpan dengan format nama ***Komoditi Cabai Merah Keriting.ktr***

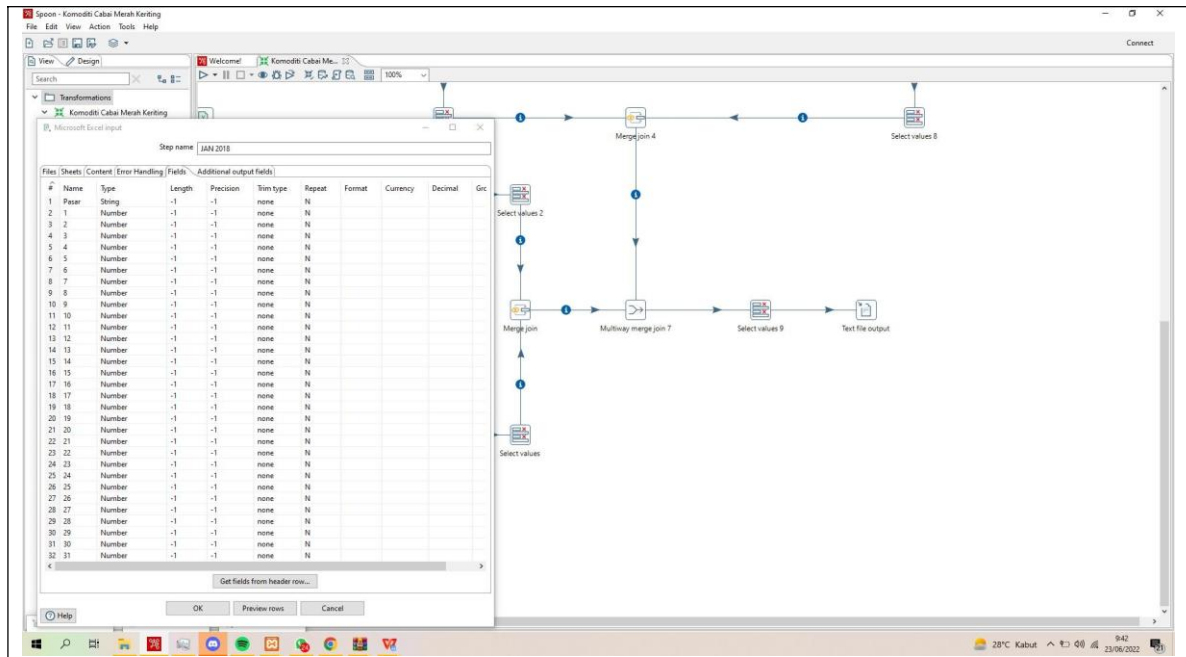


2. Excel Input

Berikut ini adalah konfigurasi pada step *Microsoft Excel Input*, lakukan pada Jan 2018 – Des 2020 (Files, Sheets, Fields)

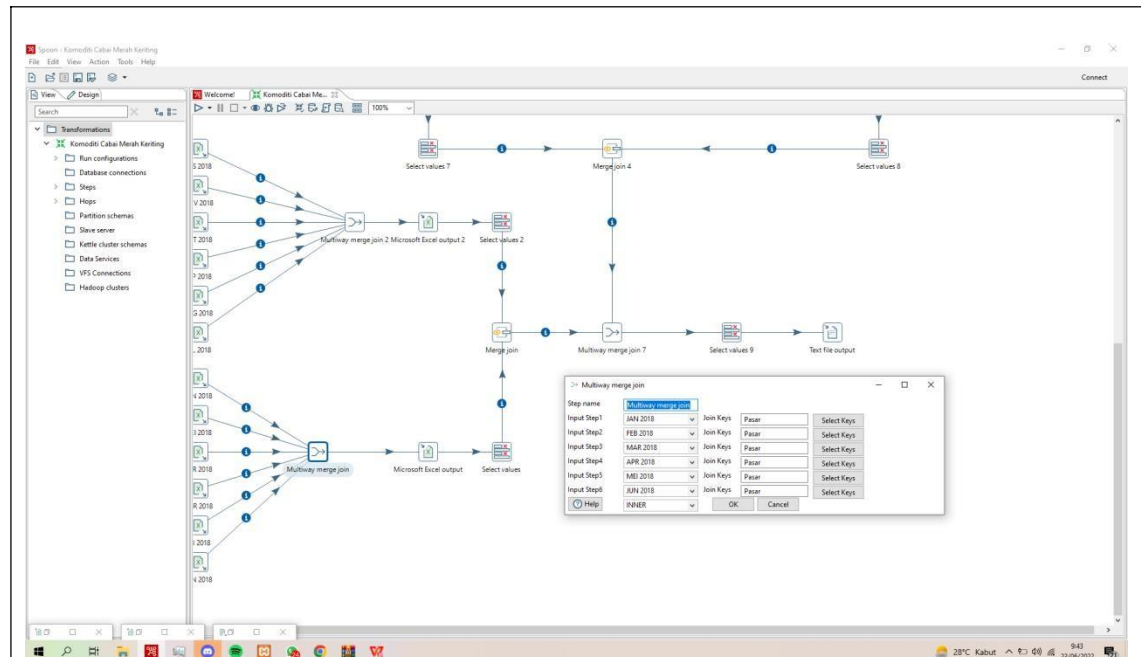






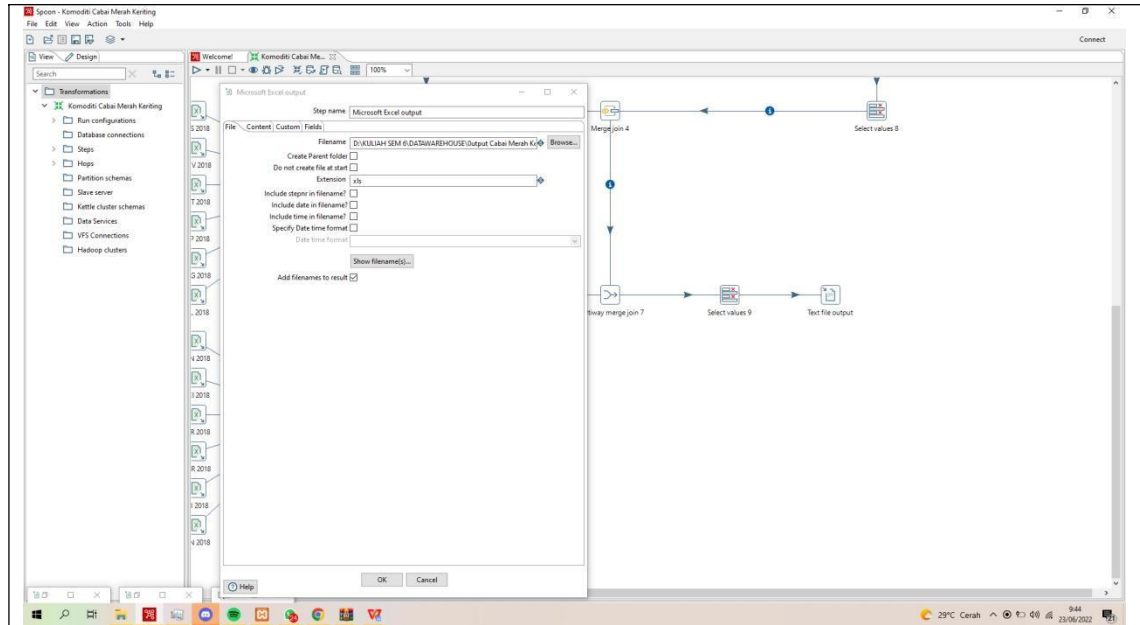
3. Multiway Merge Join

- Step selanjutnya adalah melakukan penggabungan data untuk menggabungkan data agar lebih mudah dalam mengolahnya. Step yang digunakan ialah Multiway Merge Join berikut ini konfigurasi, lakukan pada Multiway Merge Join – Multiway Merge Join 6



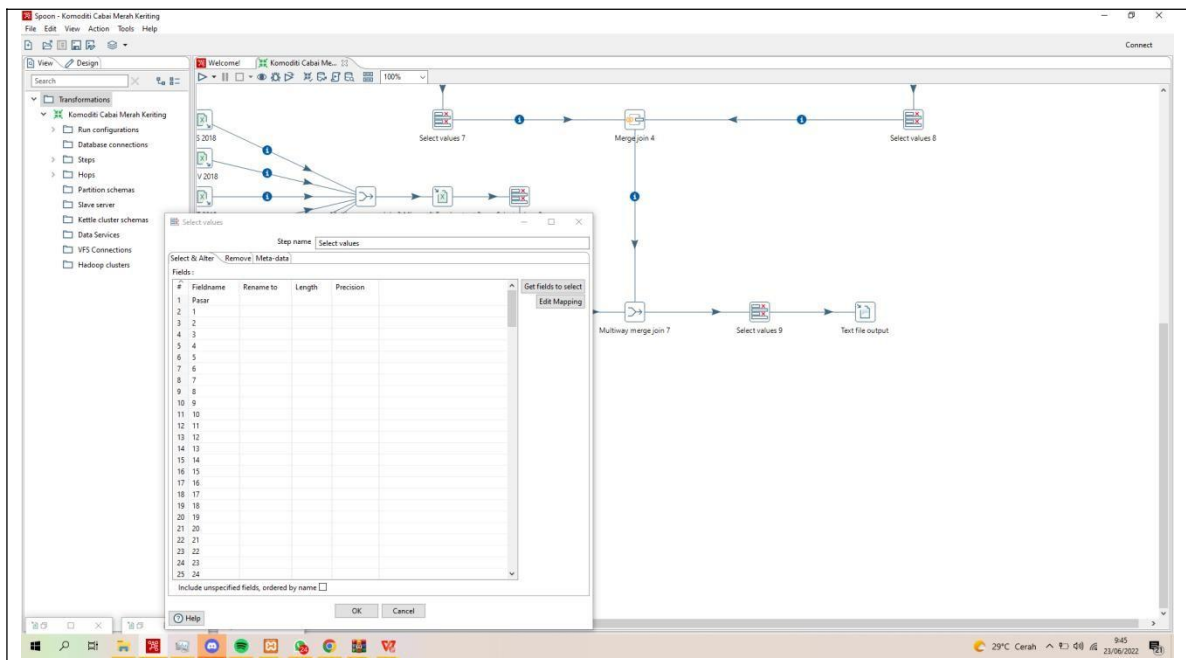
- Kemudian pada Microsoft Excel Output setelah Multiway Merge Join, konfigurasi tetap dibiarkan default. Lakukan pada Microsoft Excel Output – Microsoft Excel Output 6

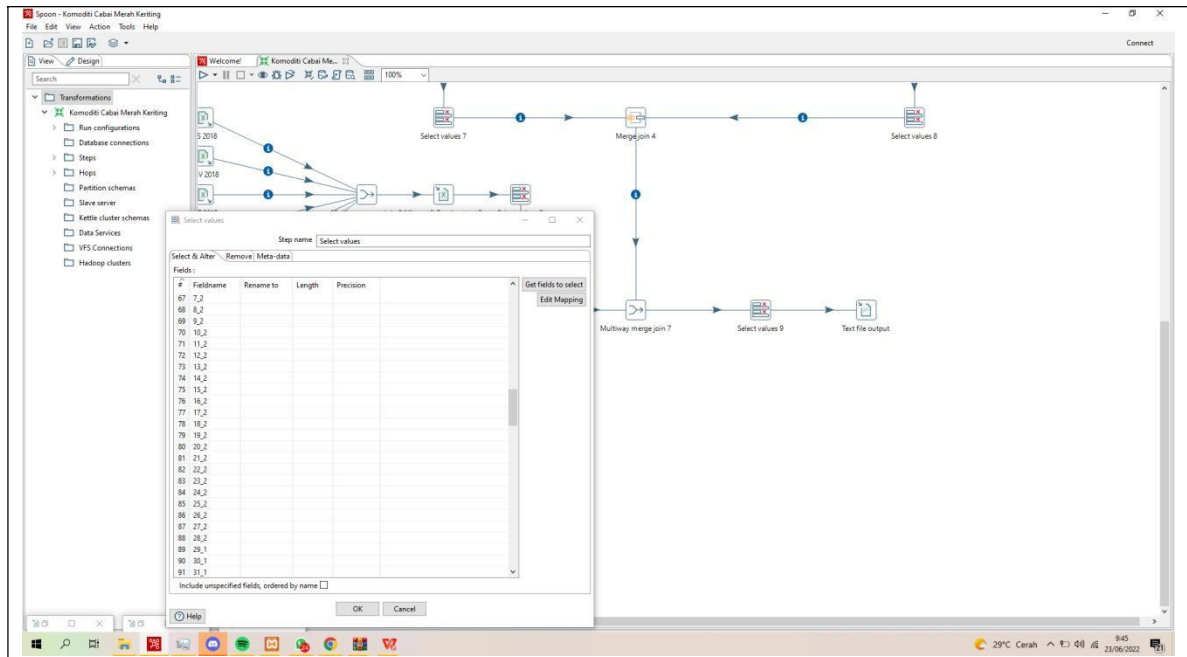




4. Select Values

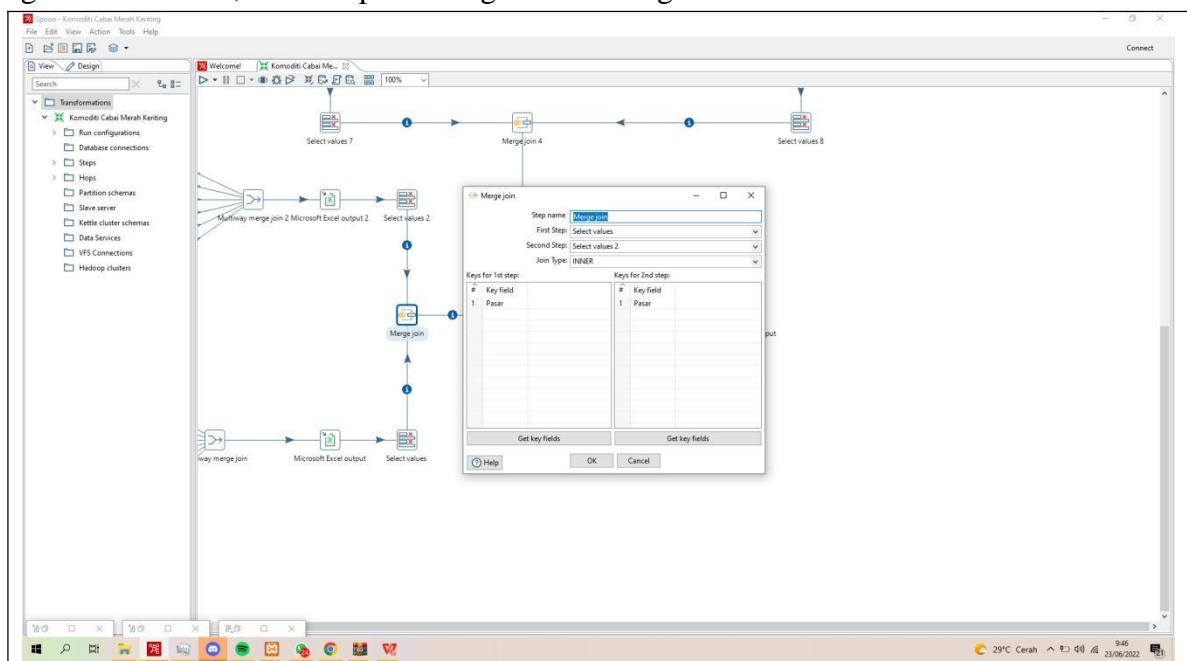
Step selanjutnya ialah menghapus beberapa variabel *pasar* setelah dilakukan proses penggabungan. Variabel tersebut dihilangkan dan hanya disisakan pada bagian atas saja. Contoh konfigurasinya seperti pada gambar dibawah ini, lakukan pada Select Values – Select Values 6





5. Merge Join

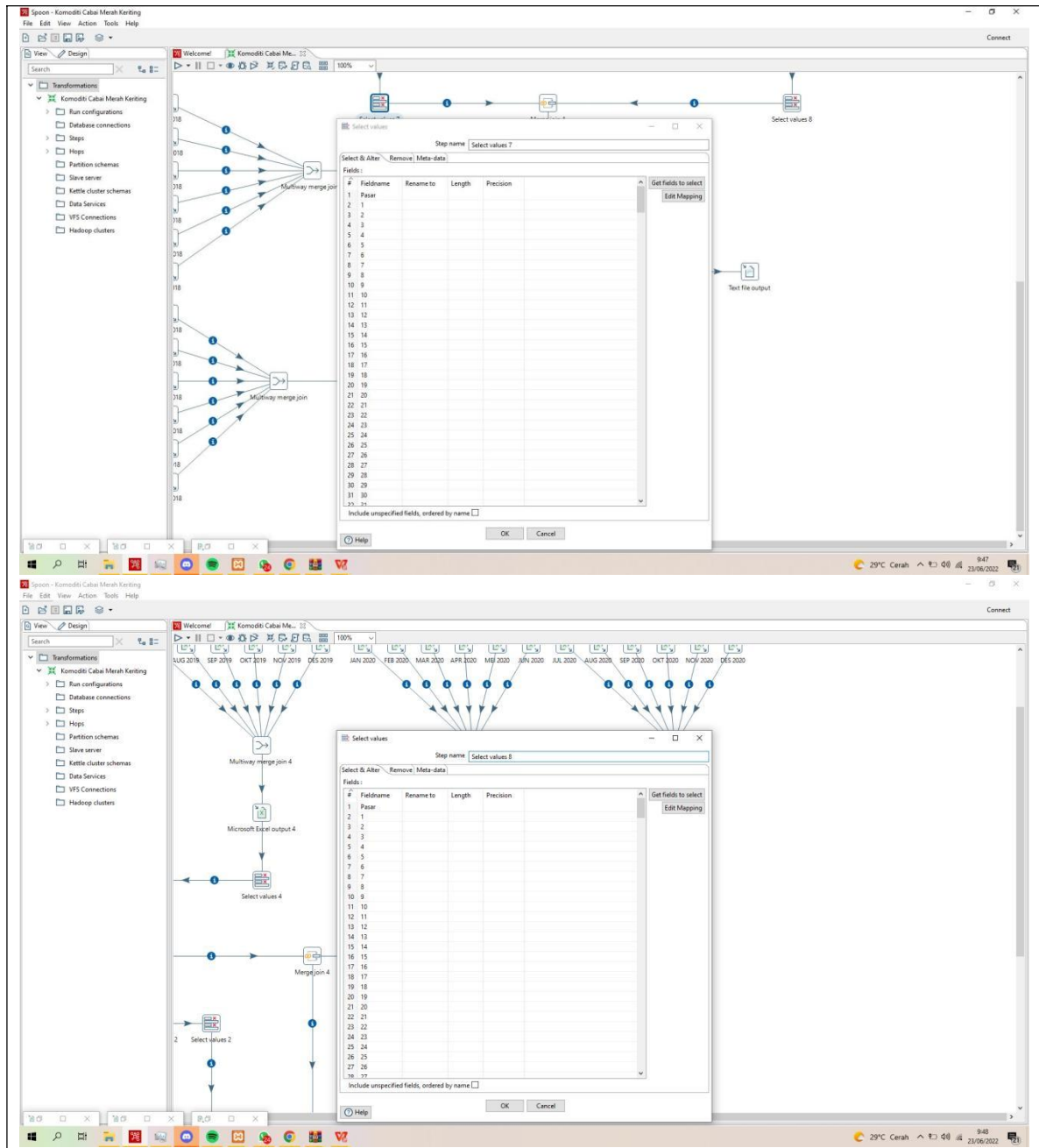
Step selanjutnya ialah melakukan penggabungan dari dua semester data yang sebelumnya sudah di bersihkan menggunakan Select Values. Konfigurasinya dapat dilihat pada gambar dibawah, lakukan pada Merge Join – Merge Join 3



6. Select Values

Lalu hapus variabel pasar dari hasil penggabungan dan sisakan variabel pasar pada list paling atas data seperti pada langkah 4 yang telah dilakukan sebelumnya. Lakukan pada Select Values 7 & Select Values 8

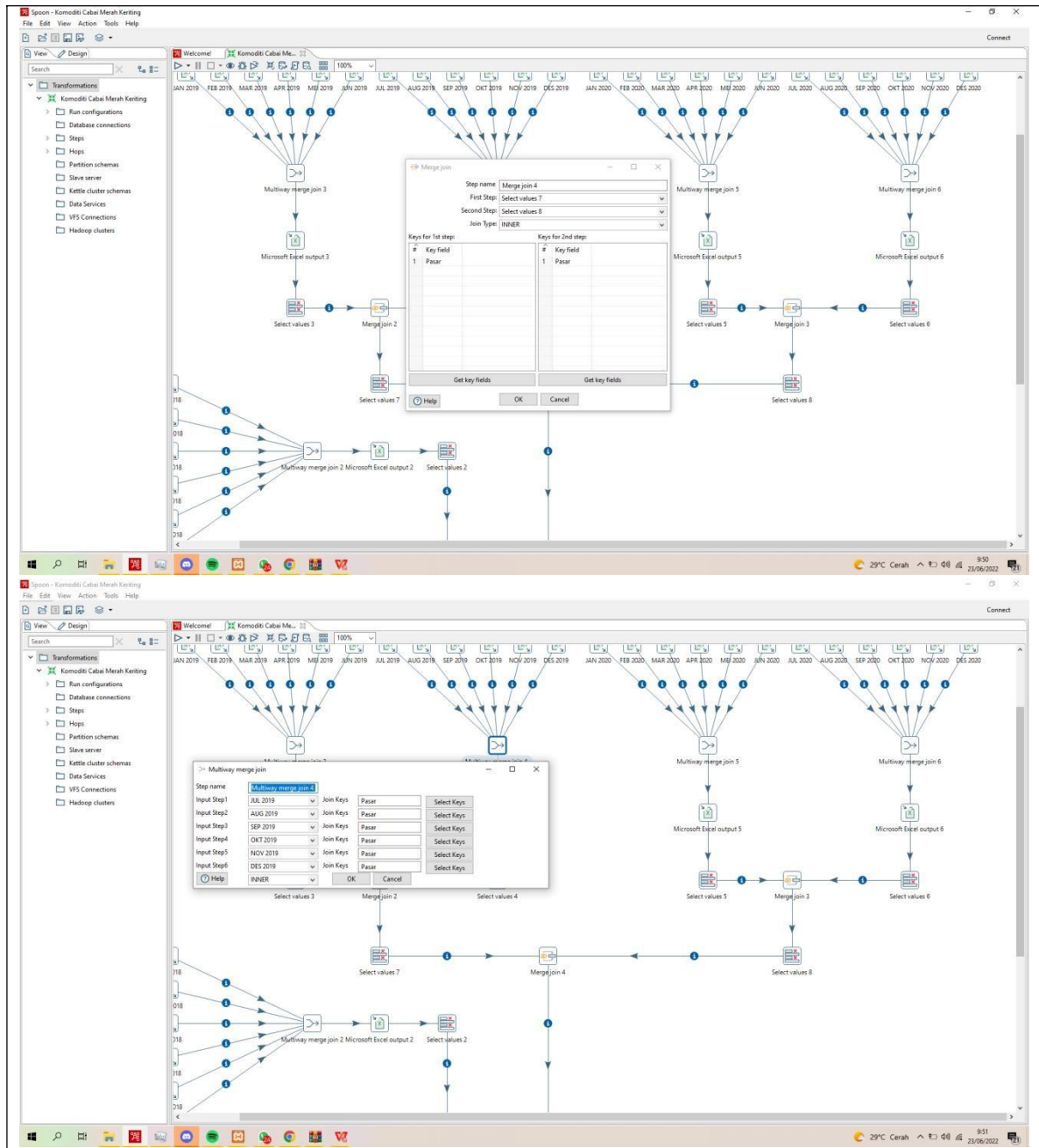




7. Merge Join & Multiway Merge Join

Step selanjutnya ialah melakukan Merge Join dan setelahnya melakukan Multiway Merge Join dengan konfigurasi sebagai berikut, lakukan pada Merge Join 4 & Multiway Merge Join 7

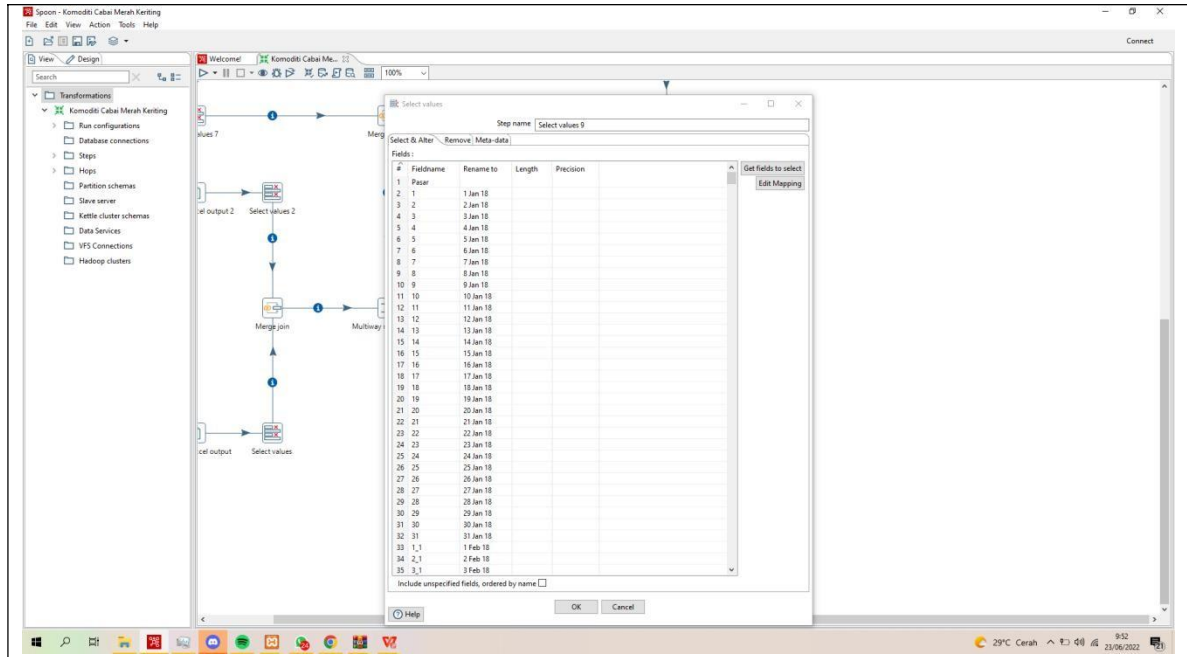




8. Select Values

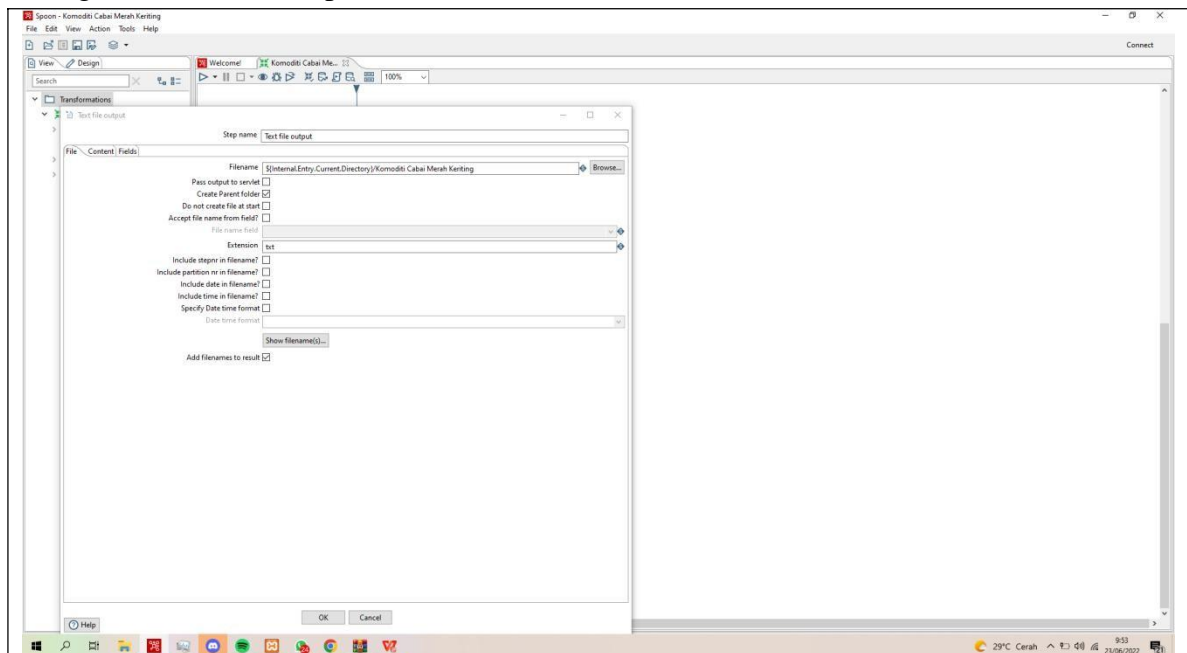
Step selanjutnya ialah menghapus variabel pasar dari hasil penggabungan dan sisakan variabel pasar pada list paling atas data seperti pada langkah 4 dan langkah 6 yang telah dilakukan sebelumnya lalu rename setiap data dengan format *DD-MMM-YY* contoh seperti 19-Des-20 selengkapnya dapat dilihat pada gambar, lakukan pada Select Values 9

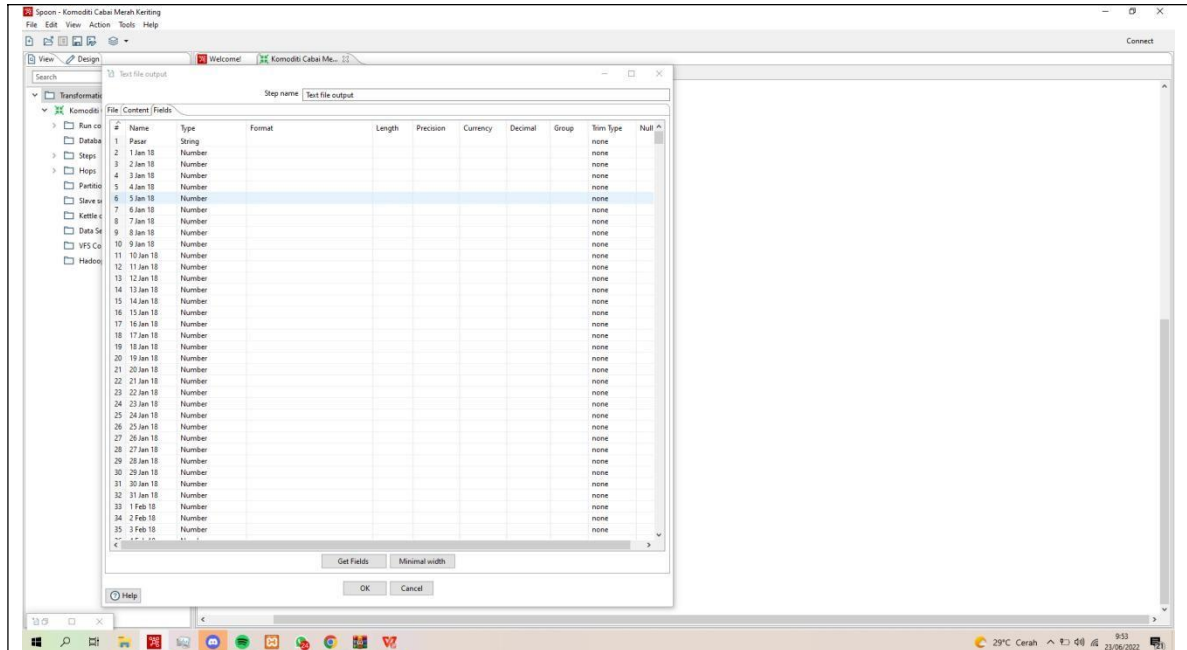




9. Text File Output

Langkah terakhir simpan data menggunakan step *Text File Output* dengan konfigurasi sebagai berikut lalu simpan dengan format nama





2. Alat dan Bahan

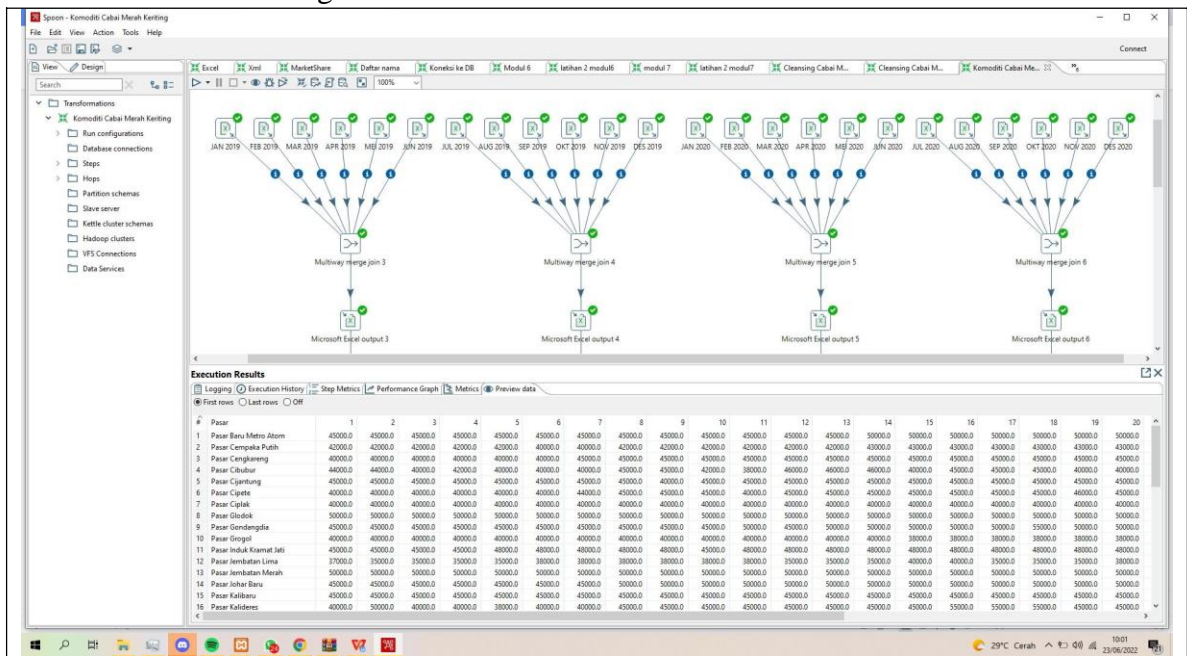
Hardware : Laptop/PC

Software : Spoon Pentaho from Hitachi Vantara

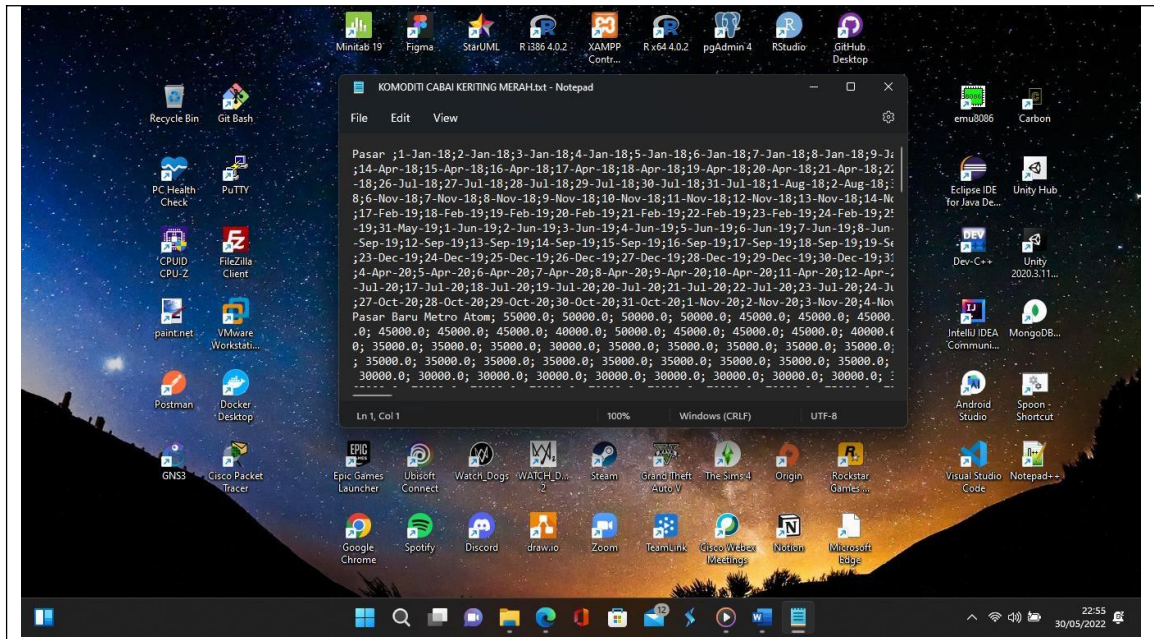
3. Elemen Kompetensi

a. Tugas Akhir Praktikum IV

1. Struktur ETL Gabungan



2. Data ETL Gabungan



4. File Praktikum

Github Repository:

<https://github.com/rasyahdz/PraktikumDWH.git>

5. Soal Latihan

Soal:

1. Apa itu ETL?
2. Mengapa ETL itu penting?

Jawaban:

1. ETL adalah singkatan dari Extract, Transform, Load. Sesuai namanya, Extract, Transform, Load merupakan proses yang meliputi ekstrak data dari beberapa sistem sumber, yang berlanjut kepada transformasi data (proses penghitungan, pembulatan, dan sebagainya) hingga akhirnya data ditampung (load) ke dalam sistem data warehouse. Data-data yang semula bercampur diolah menjadi data tunggal yang konsisten sehingga mudah untuk dilakukan analisis ketika dibutuhkan nantinya.

ETL menjadi pondasi untuk data analytics dan machine learning. Melalui sejumlah pengaturan, ETL dapat membersihkan dan mengorganisir data sesuai dengan yang diinginkan. Data laporan bulanan pun akan dapat dengan mudah dianalisis.

2. Ada sejumlah alasan mengapa Extract, Transform, Load penting untuk diterapkan dalam sebuah organisasi atau perusahaan. Alasan-alasan tersebut antara lain:



- Extract, Transform, Load membantu perusahaan menganalisis data bisnis sebagai acuan pengambilan keputusan penting.
- Database transaksional umumnya tidak dapat menjawab permasalahan bisnis yang akan sangat mudah dijawab oleh ETL.
- ETL menyediakan metode untuk memindahkan data dari berbagai sumber ke dalam satu data warehouse.
- Ketika data source berubah, data warehouse akan secara otomatis ter-update.
- Sistem Extract, Transform, Load yang didesain dan didokumentasikan dengan baik akan sangat berpengaruh terhadap tingkat kesuksesan proyek data warehouse.
- Extract, Transform, Load dapat memverifikasi aturan transformasi, agregasi, dan kalkulasi dalam sebuah data set.
- Proses Extract, Transform, Load memungkinkan hadirnya contoh perbandingan data antara sumber dan sistem target.
- Proses Extract, Transform, Load dapat melakukan transformasi pada data yang kompleks dan mungkin membutuhkan area ekstra untuk penyimpanan data.
- Berkat kemampuan pemrograman perpindahan data otomatis, Extract, Transform, Load dapat meningkatkan produktivitas data profesional tanpa membutuhkan keahlian teknis untuk menulis kode dan script.

6. Kesimpulan

- Dalam pengerjaan praktikum Data Warehouse, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
- Kita dapat mengetahui...

7. Cek List (✓)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Tugas Akhir Praktikum IV	✓	

8. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Tugas Akhir Praktikum IV	Menit	Baik

Keterangan:

- Menarik
- Baik
- Cukup
- Kurang

