


| | | |
|---|--|---|
| Nama: Adam Zannuba |  Praktikum Data Warehouse | MODUL 12 |
| NIM: 065001900022 | | Nama Dosen: Ir. Teddy Siswanto, MMSi |
| Hari/Tanggal: 23 Juni 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 analisis 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.

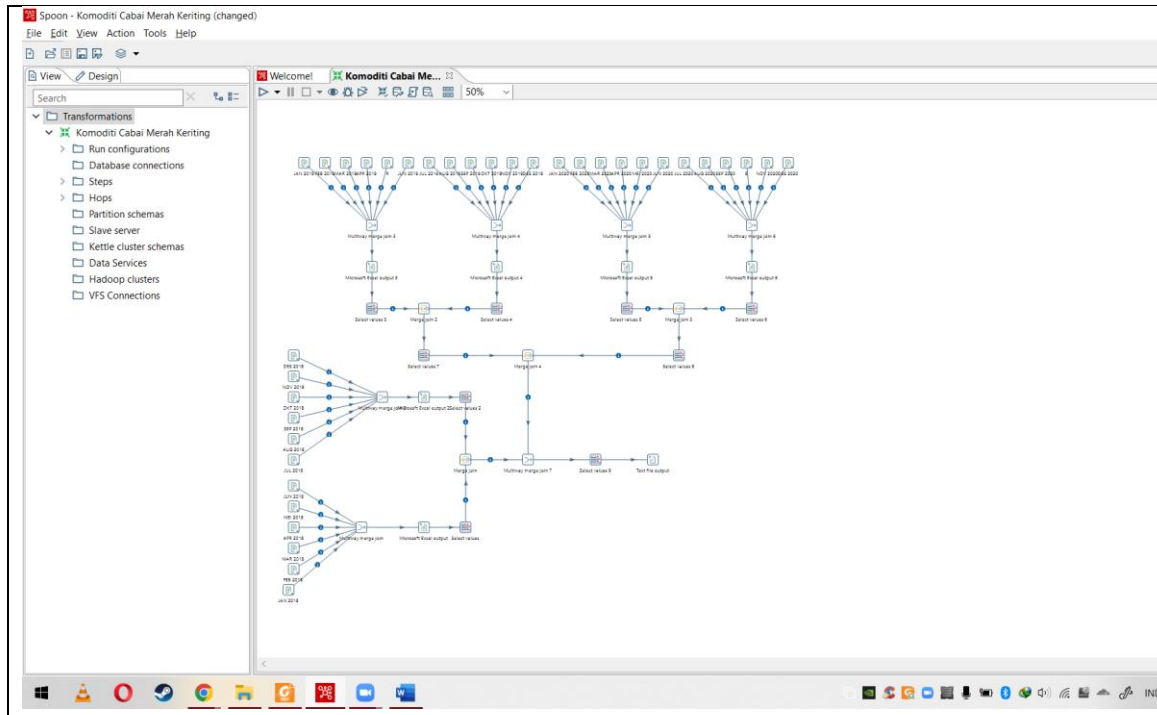
ETL menjadi pondasi untuk data analytics dan machine learning. Melalui sejumlah pengaturan, ETL dapat membersihkan dan mengorganisir data sesuai dengan yang diinginkan. Data laporan bulanan pun akan dapat dengan mudah dianalisis.



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

1. Struktur

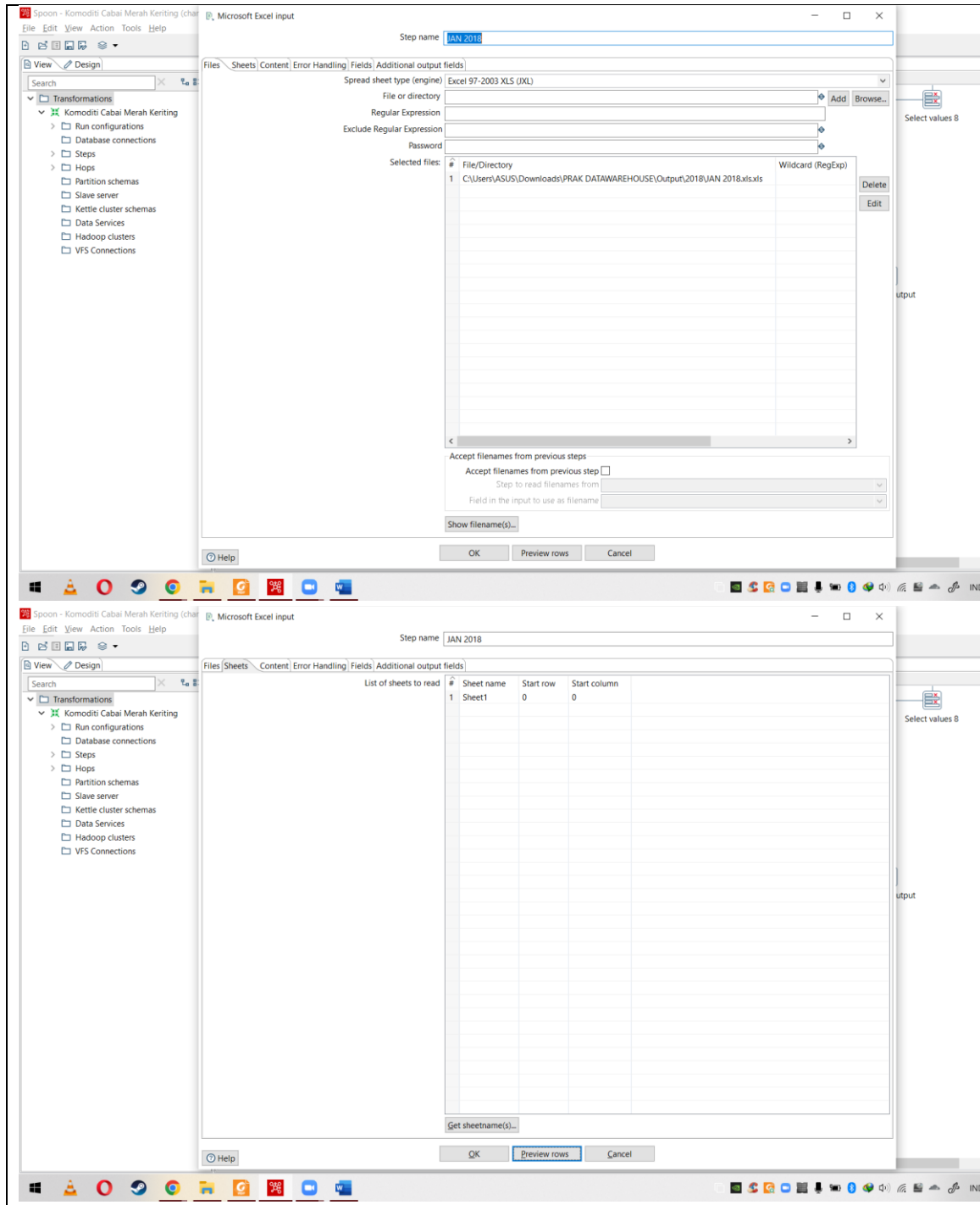
Buka Pentaho Spoon Data Integration lalu buat struktur seperti pada gambar dibawah ini lalu simpan dengan format nama ***Komoditi Cabai Merah Keriting.ktr***

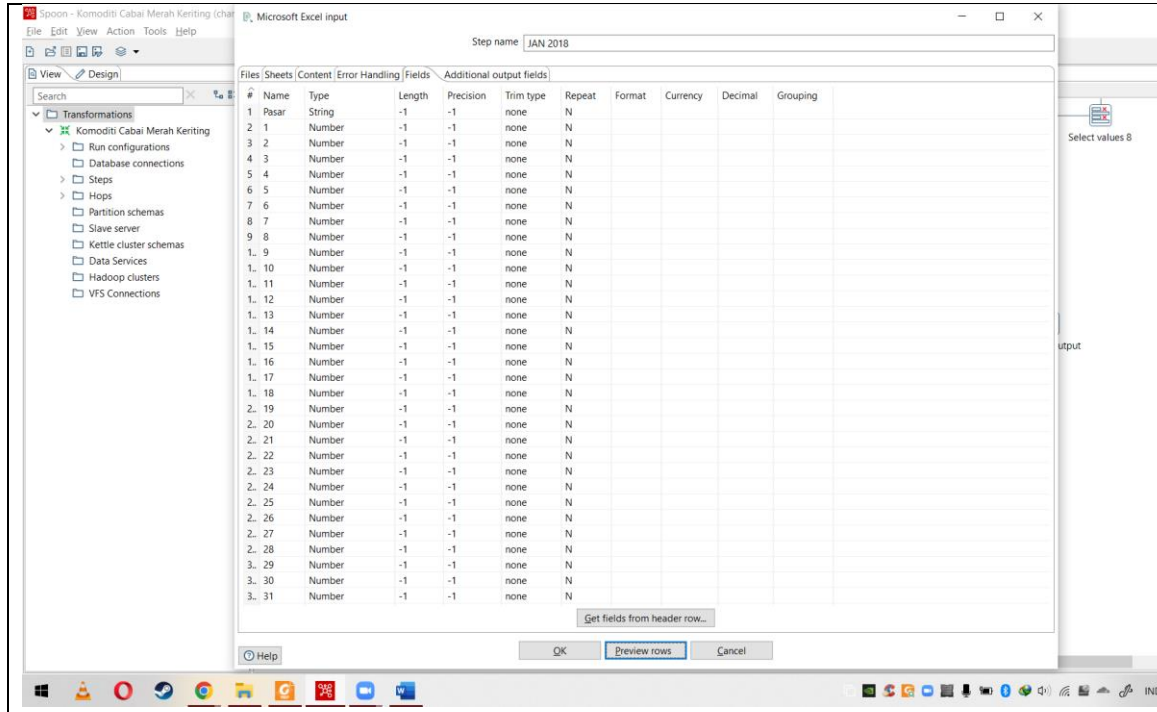


2. Excel Input

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

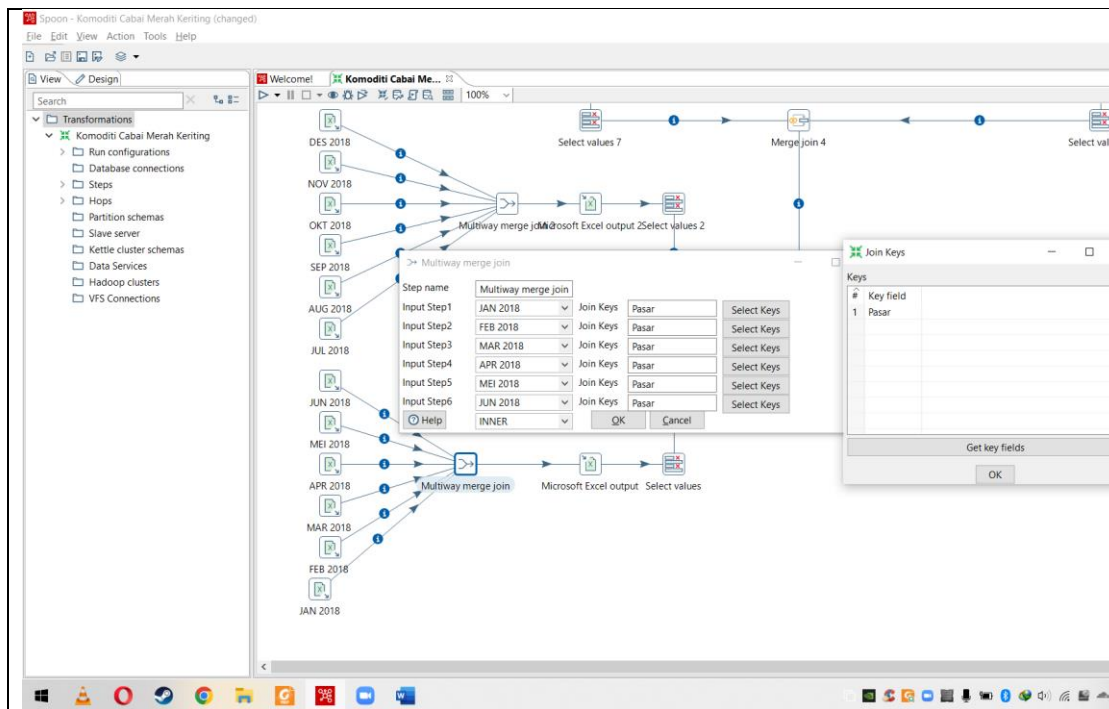


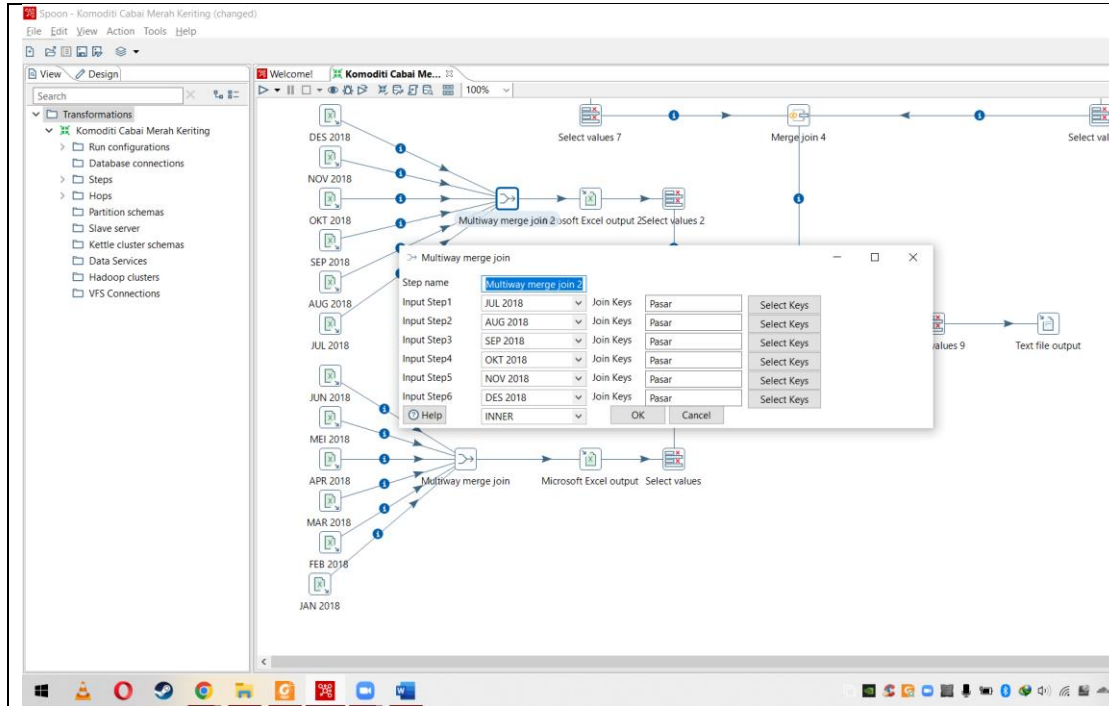




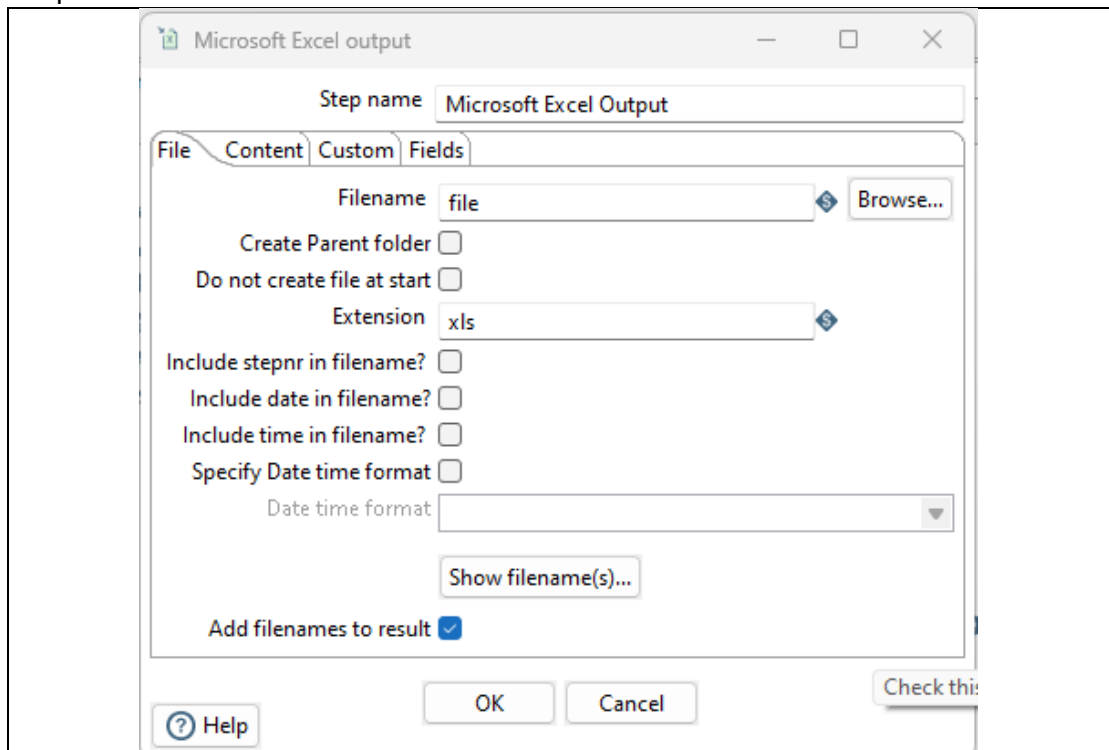
3. Multiway Merge Join

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



The image displays two screenshots of the Apache NiFi 'Select values' dialog box, showing the process of selecting fields for a data transformation step.

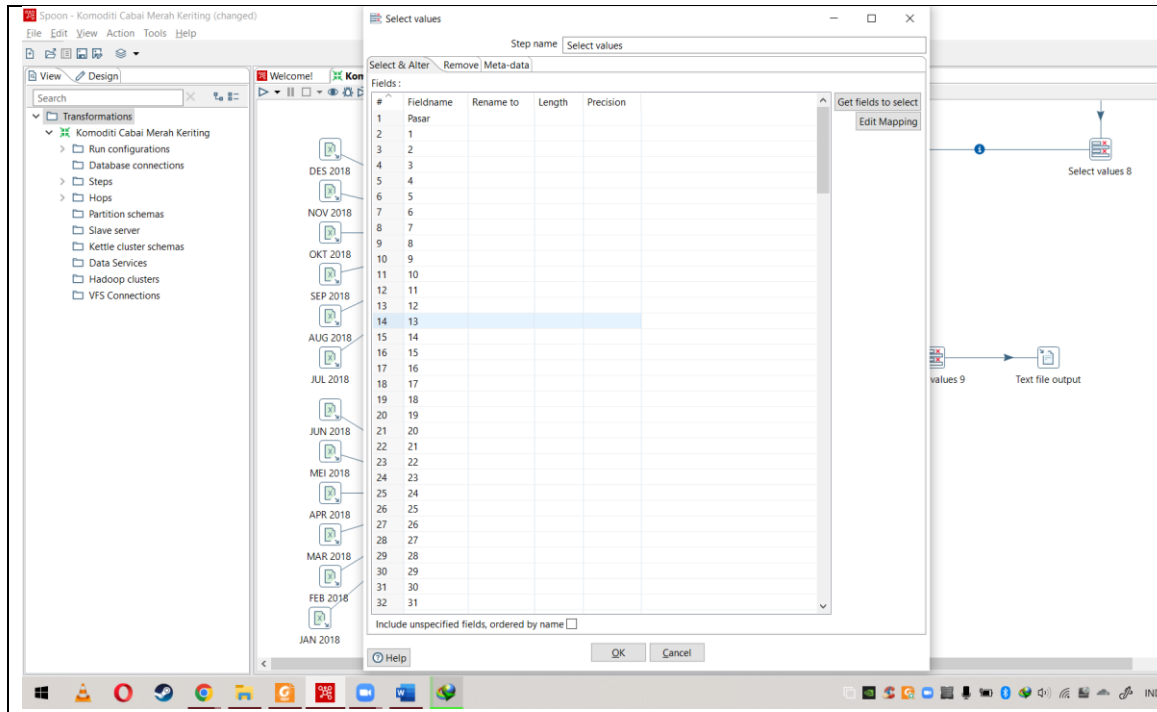
Top Screenshot: The 'Select values' dialog box is open, showing a list of fields. The 'Pasar_1' field is selected. The dialog box includes a 'Get fields to select' button and an 'Edit Mapping' button. The background shows a NiFi workflow with a 'Select values 8' step and a 'Text file output' step.

Bottom Screenshot: The 'Select values' dialog box is open, showing a list of fields. The '11' field is selected. The dialog box includes a 'Get fields to select' button and an 'Edit Mapping' button. The background shows a NiFi workflow with a 'Select values 8' step and a 'Text file output' step.

| # | Fieldname | Rename to | Length | Precision |
|----|-----------|-----------|--------|-----------|
| 1. | 15 | | | |
| 1. | 16 | | | |
| 1. | 17 | | | |
| 1. | 18 | | | |
| 2. | 19 | | | |
| 2. | 20 | | | |
| 2. | 21 | | | |
| 2. | 22 | | | |
| 2. | 23 | | | |
| 2. | 24 | | | |
| 2. | 25 | | | |
| 2. | 26 | | | |
| 2. | 27 | | | |
| 2. | 28 | | | |
| 3. | 29 | | | |
| 3. | 30 | | | |
| 3. | 31 | | | |
| 3. | Pasar_1 | | | |
| 3. | 1,1 | | | |
| 3. | 2,1 | | | |
| 3. | 3,1 | | | |
| 3. | 4,1 | | | |
| 3. | 5,1 | | | |
| 3. | 6,1 | | | |
| 4. | 7,1 | | | |
| 4. | 8,1 | | | |
| 4. | 9,1 | | | |
| 4. | 10,1 | | | |
| 4. | 11,1 | | | |
| 4. | 12,1 | | | |
| 4. | 13,1 | | | |
| 4. | 14,1 | | | |

| # | Fieldname | Rename to | Length | Precision |
|----|-----------|-----------|--------|-----------|
| 12 | 11 | | | |
| 13 | 12 | | | |
| 14 | 13 | | | |
| 15 | 14 | | | |
| 16 | 15 | | | |
| 17 | 16 | | | |
| 18 | 17 | | | |
| 19 | 18 | | | |
| 20 | 19 | | | |
| 21 | 20 | | | |
| 22 | 21 | | | |
| 23 | 22 | | | |
| 24 | 23 | | | |
| 25 | 24 | | | |
| 26 | 25 | | | |
| 27 | 26 | | | |
| 28 | 27 | | | |
| 29 | 28 | | | |
| 30 | 29 | | | |
| 31 | 30 | | | |
| 32 | 31 | | | |
| 33 | 1,1 | | | |
| 34 | 2,1 | | | |
| 35 | 3,1 | | | |
| 36 | 4,1 | | | |
| 37 | 5,1 | | | |
| 38 | 6,1 | | | |
| 39 | 7,1 | | | |
| 40 | 8,1 | | | |
| 41 | 9,1 | | | |
| 42 | 10,1 | | | |
| 43 | 11,1 | | | |

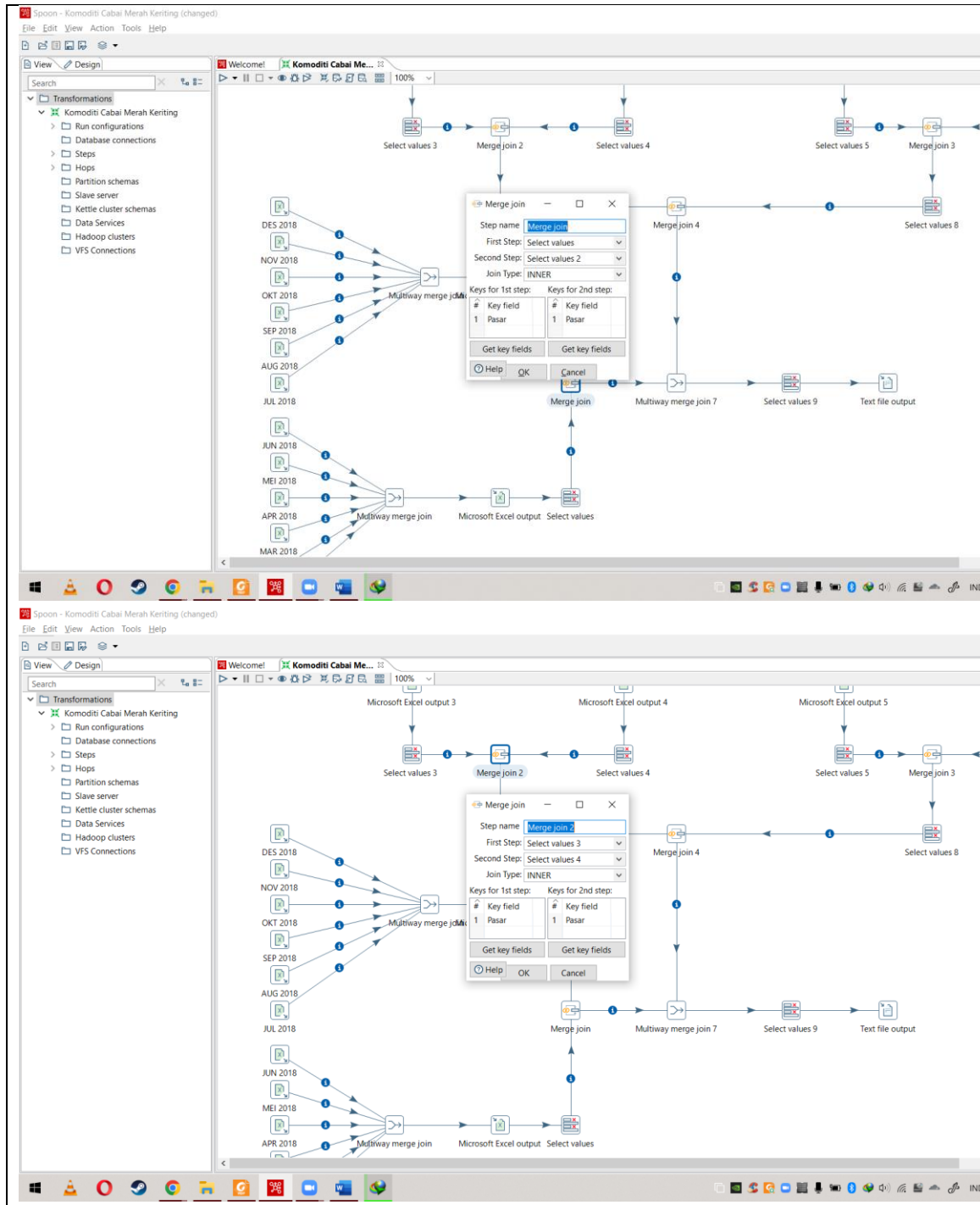


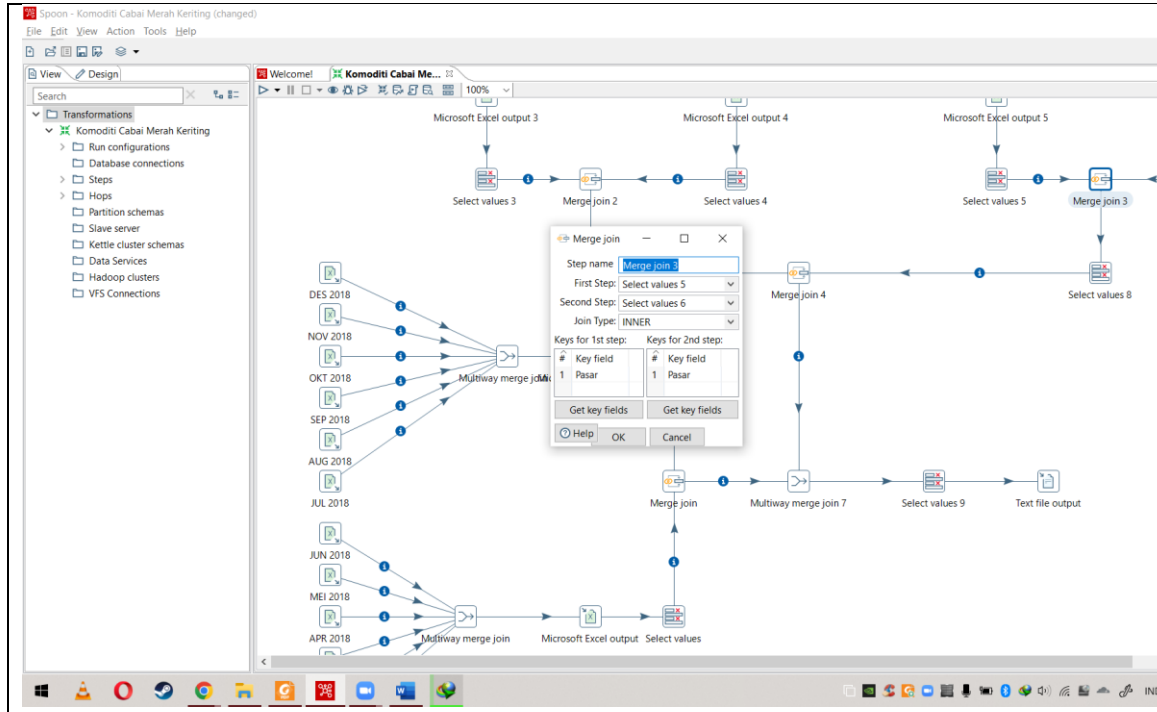


5. Merge Join

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

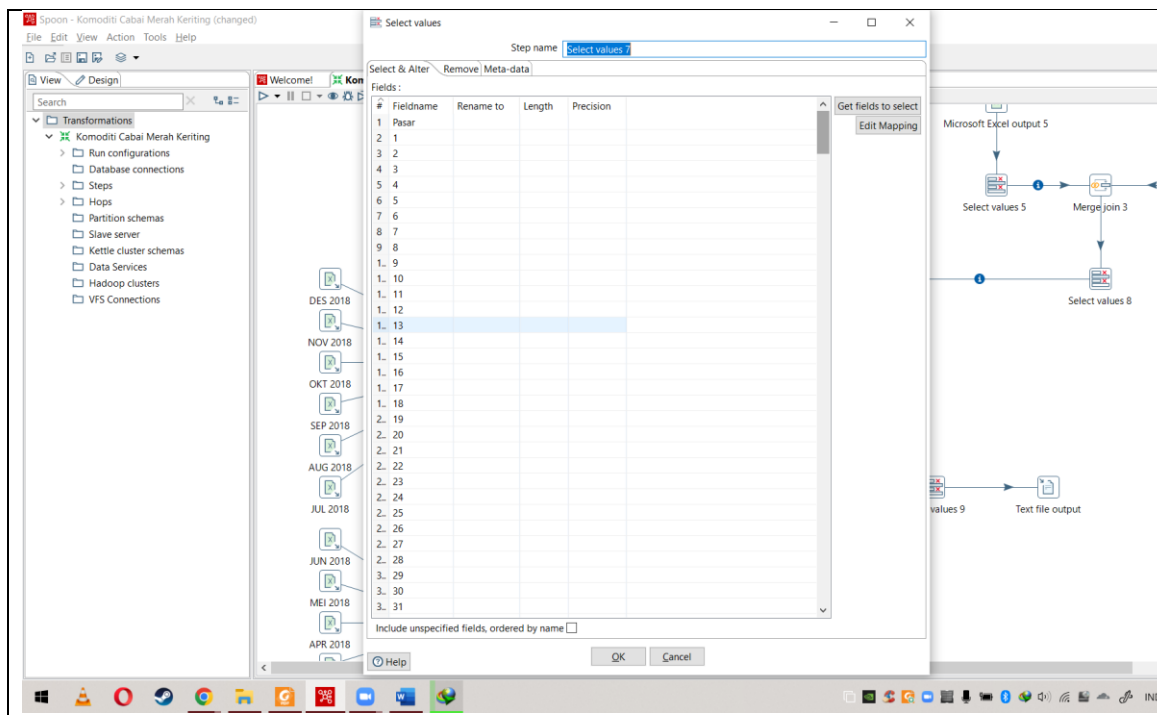






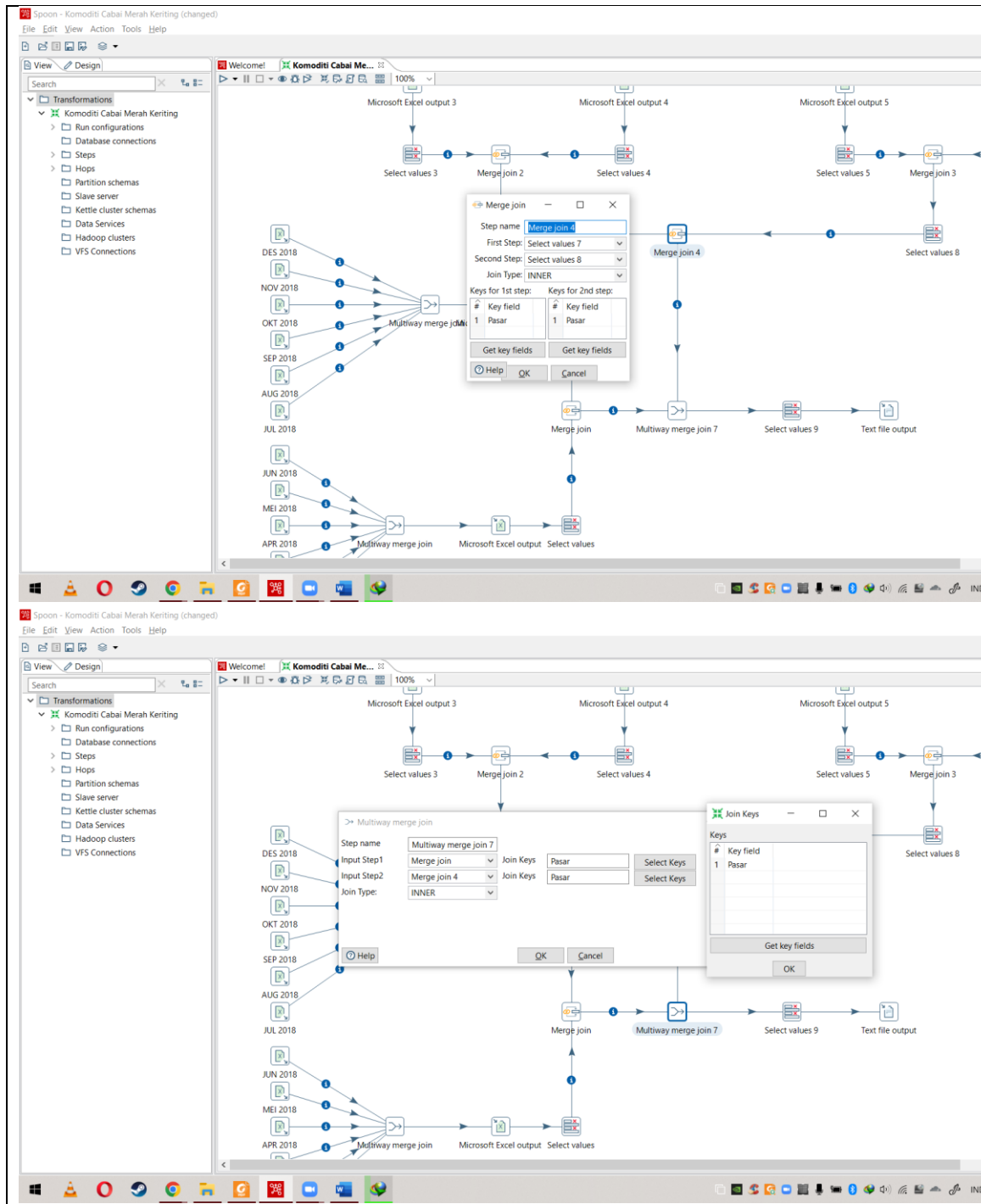
6. Select Values

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



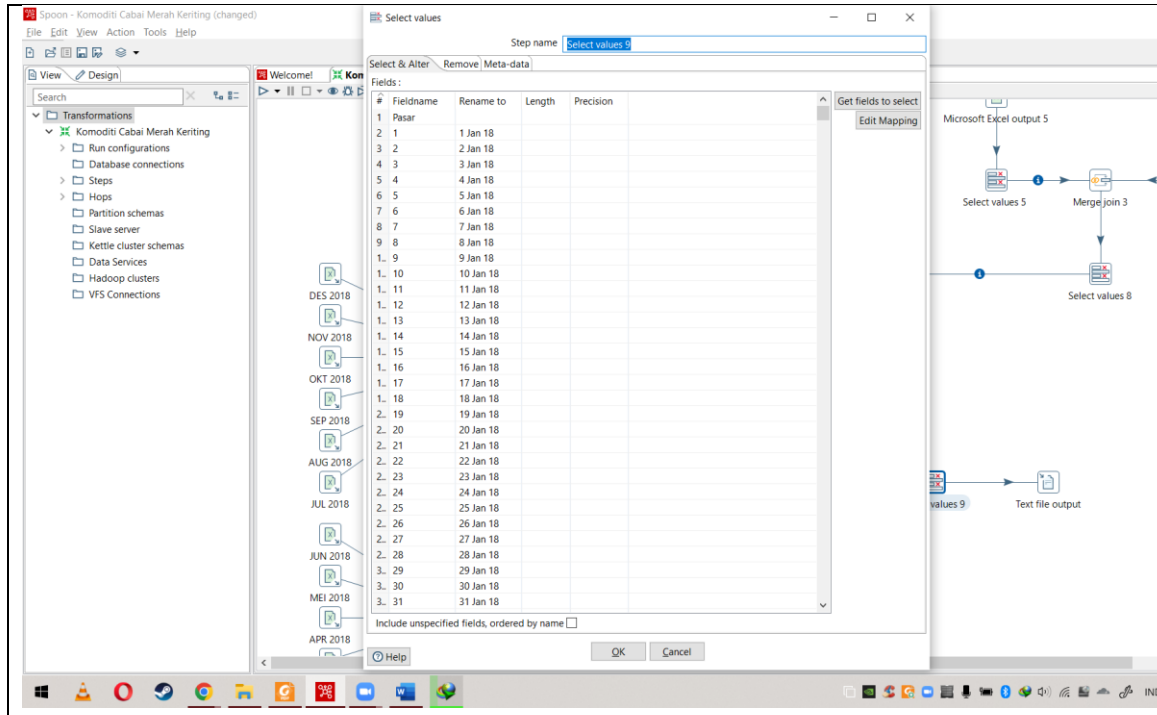
7. Merge Join & Multiway Merge Join

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

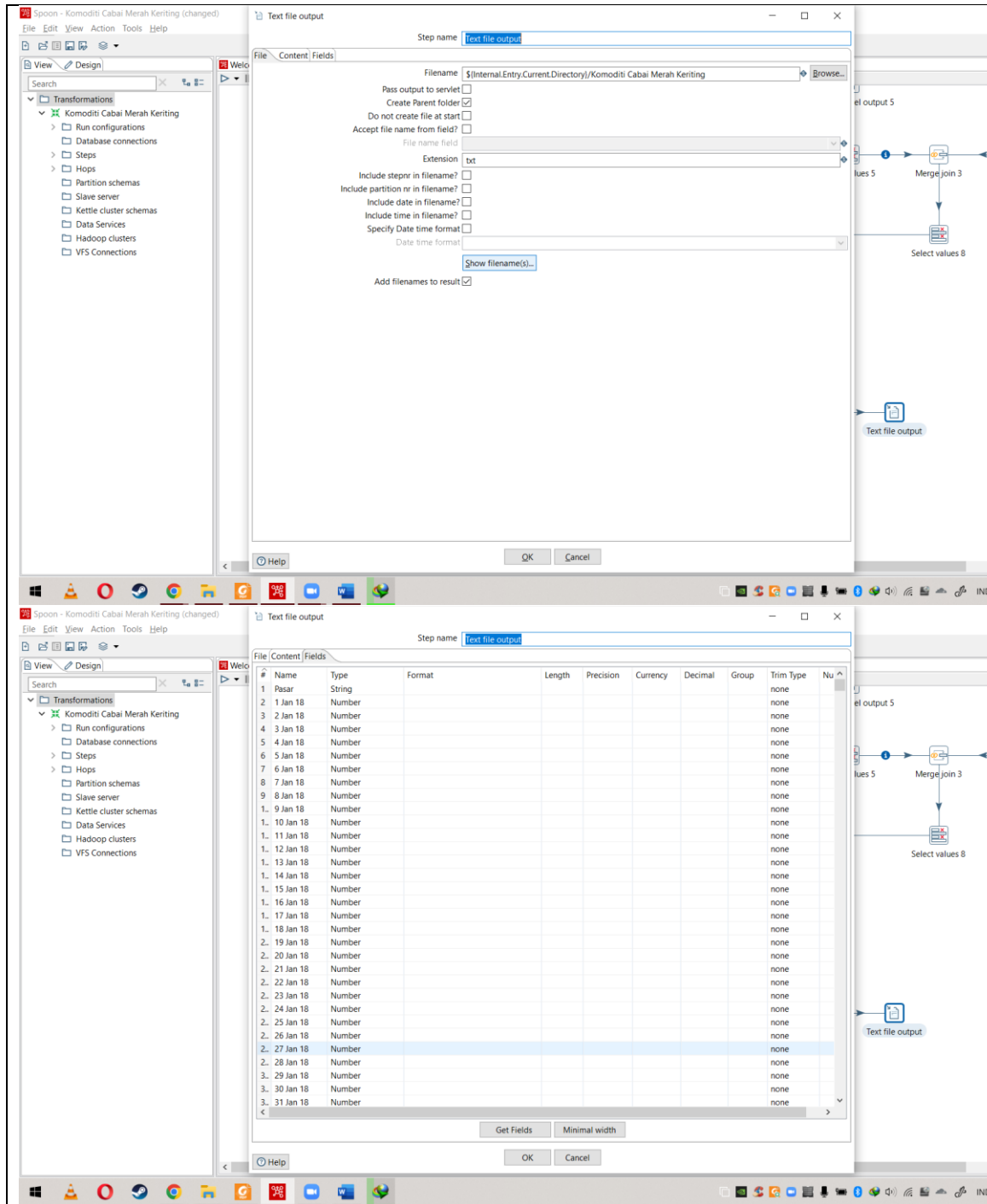
Step selanjutnya ialah menghapus variabel pasar dari hasil penggabungan dan sisakan variabel pasar pada list paling atas data seperti pada langkah 4 dan langkah 6 yang telah dilakukan sebelumnya lalu rename setiap data dengan format *DD-MMM-YY* contoh seperti 19-Des-20 selengkapnya dapat dilihat pada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



Jurusan Teknik Informatika & Sistem Informasi
Fakultas Teknologi Industri – Universitas Trisakti

The screenshot displays the SAP Business Objects Data Services (BO DS) environment. The top menu bar includes 'Spoon - Komoditi Cabai Merah Keriting', 'File', 'Edit', 'View', 'Action', 'Tools', and 'Help'. Below the menu is a toolbar with icons for file operations and a 'View' dropdown set to 'Design'.

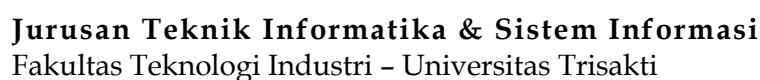
The main workspace shows a data flow diagram for a project named 'Komoditi Cabai Merah Keriting'. The diagram illustrates a multiway merge process. It features three 'Multiway merge join' nodes (labeled 3, 4, and 5) which are connected to 'Microsoft Excel output' nodes. The data sources are represented by a series of 'P' icons (likely representing partitions or files) with dates ranging from 2019 FEB to 2020 OCT. The diagram is divided into three distinct sections, each showing a similar merge and output structure.

Below the design view, the 'Execution Results' section is active. It includes tabs for 'Logging', 'Execution History', 'Step Metrics', 'Performance Graph', 'Metrics', and 'Preview data'. The 'Execution History' tab is selected, showing a table of execution results. The table has columns for a step number, a step name, and 14 columns of numerical data representing execution metrics.

| # | Pasar | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|-----------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1 | Pasar Baru Metro Atom | 35000.0 | 35000.0 | 35000.0 | 35000.0 | 35000.0 | 30000.0 | 30000.0 | 35000.0 | 30000.0 | 30000.0 | 30000.0 | 30000.0 | 30000.0 | 30000.0 |
| 2 | Pasar Cempaka Putih | 15000.0 | 15000.0 | 15000.0 | 15000.0 | 15000.0 | 15000.0 | 15000.0 | 15000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 |
| 3 | Pasar Cengkareng | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 |
| 4 | Pasar Cibubur | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 28000.0 | 28000.0 | 22000.0 | 22000.0 | 22000.0 | 24000.0 | 24000.0 | 20000.0 | 22000.0 | 24000.0 |
| 5 | Pasar Cijantung | 20000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 | 25000.0 |
| 6 | Pasar Cipete | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 22000.0 | 22000.0 | 22000.0 |
| 7 | Pasar Ciplak | 18000.0 | 18000.0 | 18000.0 | 18000.0 | 18000.0 | 18000.0 | 18000.0 | 18000.0 | 18000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 |
| 8 | Pasar Grogol | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 20000.0 | 24000.0 | 24000.0 | 24000.0 | 20000.0 |
| 9 | Pasar Jembatan Lima | 17000.0 | 17000.0 | 18000.0 | 19000.0 | 20000.0 | 15000.0 | 16000.0 | 16000.0 | 16000.0 | 16000.0 | 18000.0 | 20000.0 | 20000.0 | 20000.0 |

[illegible]

Github Repository:



5. Soal Latihan

Soal:

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

Jawaban:

- 1.
- 2.

6. Kesimpulan

- a. Dalam pengerjaan praktikum Data Warehouse, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
- b. 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 | ... |

Keterangan:

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

