Nama: **Rico Renaldy** **NIM:** **065001800021**		**MODUL 6** **Nama Dosen:** **Ir. Teddy Siswanto, MMSi**
Hari/Tanggal: **Kamis, 18 April 2024**	**Praktikum Data Warehouse**	**Nama Asisten Labratorium:** **1. Nia Suhernawati - 065002100005** **2. Exchell Sabatino Geraldi Ointu - 065002100007**

Merge 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.

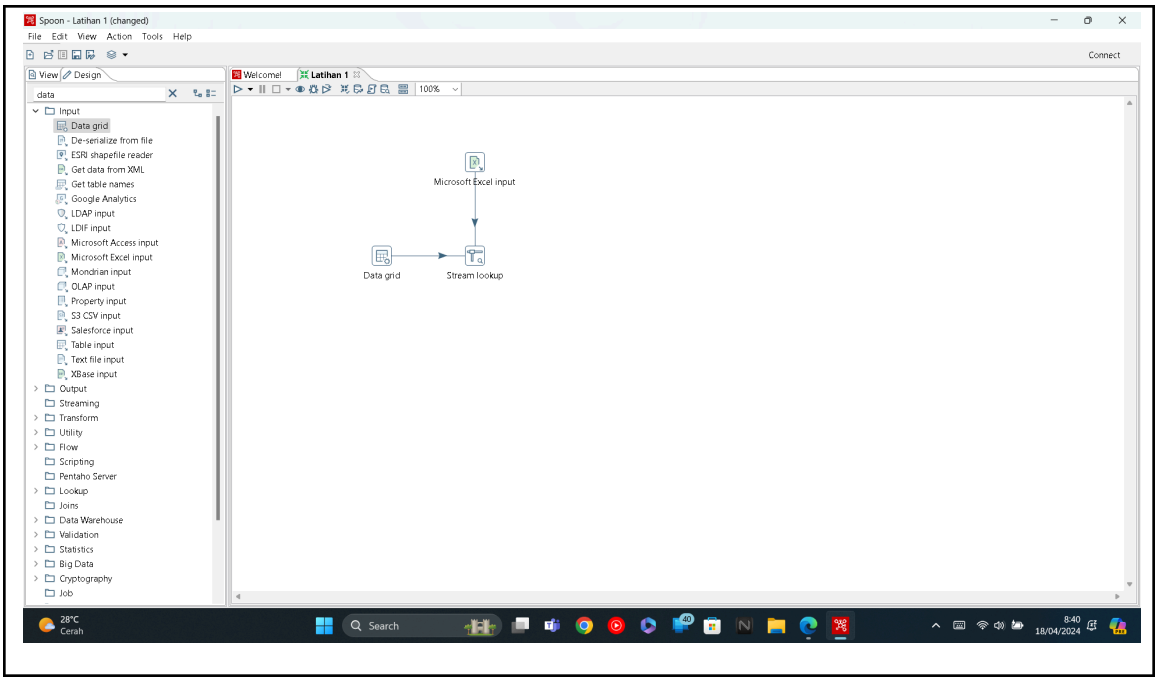
2. Alat dan Bahan

Hardware : Laptop/PC
Software : Spoon Pentaho from Hitachi Vantara

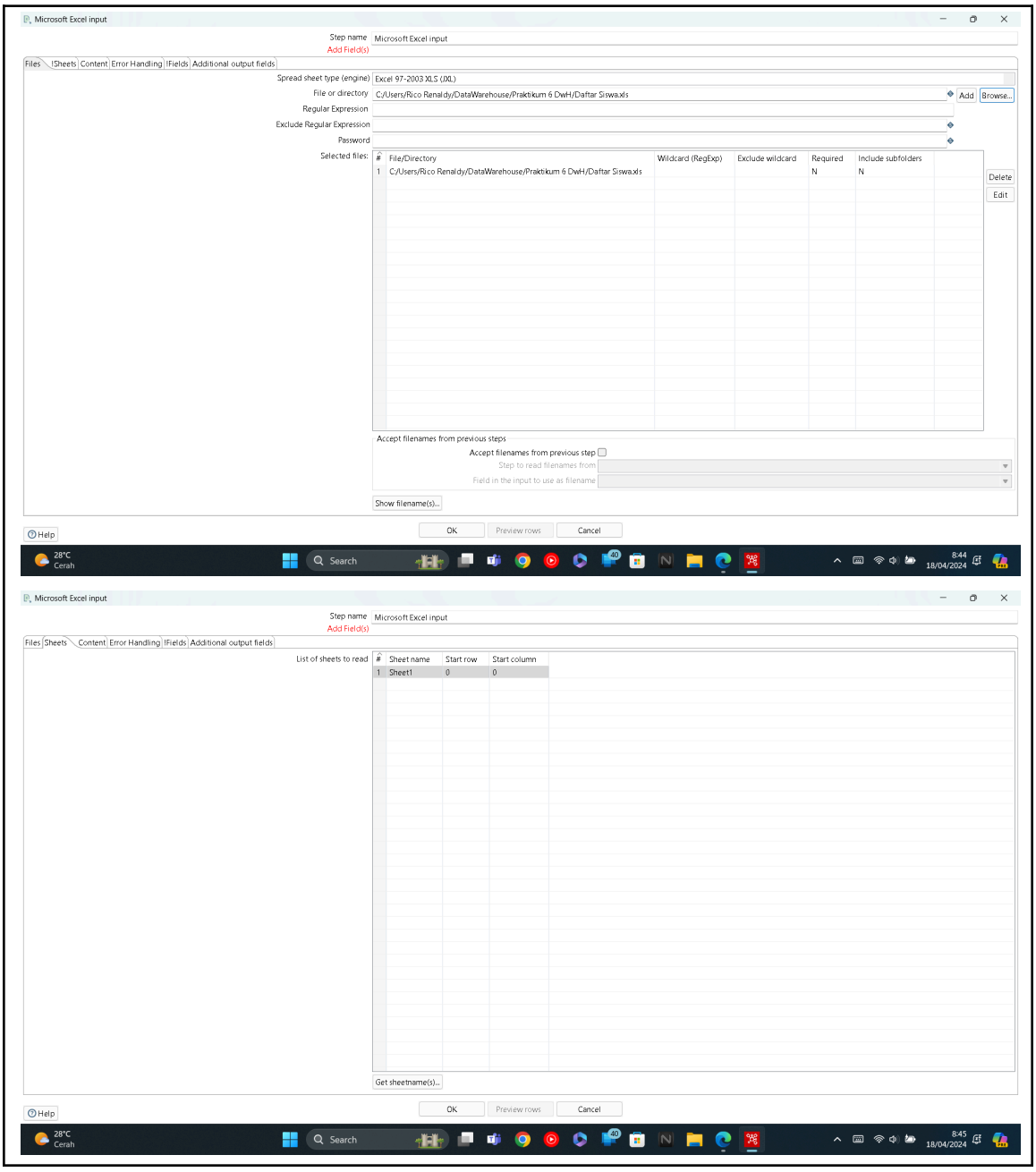
3. Elemen Kompetensi

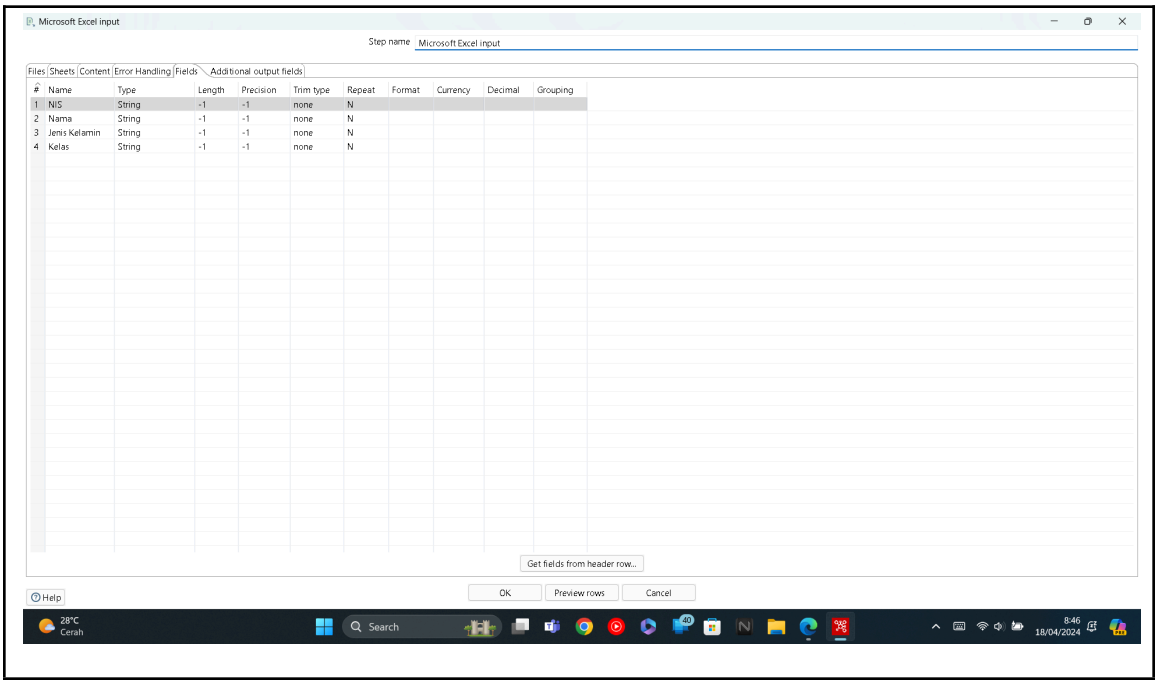
a. Latihan pertama – Melakukan penggabungan data menggunakan Stream Lookup

1. Buat transformation baru dengan struktur seperti pada gambar dibawah ini.

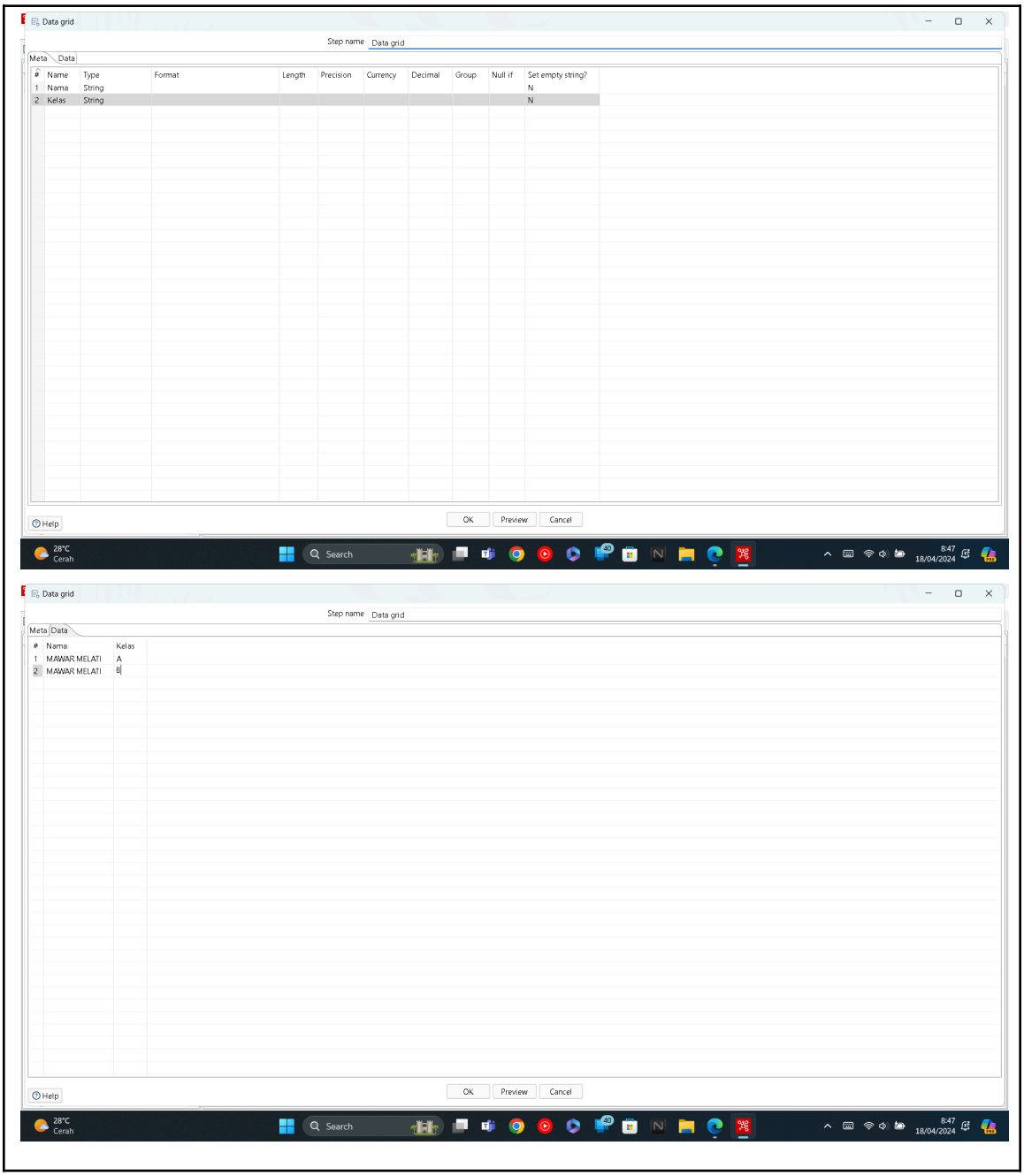


2. Microsoft Excel Input.

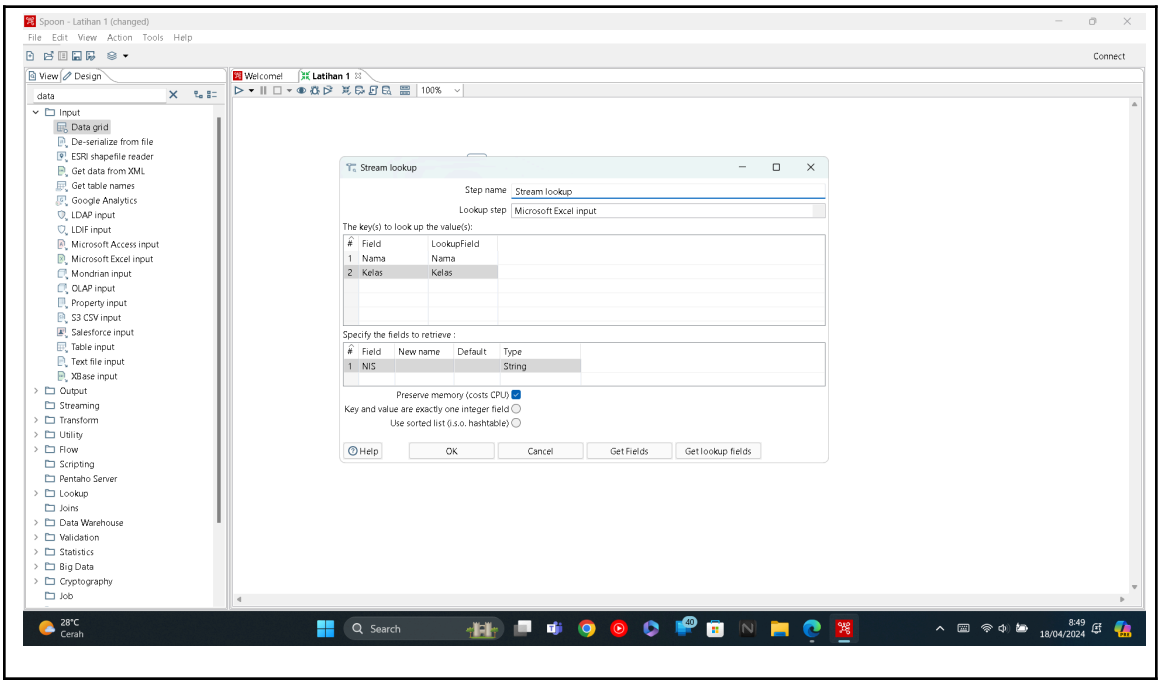

Microsoft Excel input
Step name
Microsoft Excel input
Add Field(s)
Files
Sheets
Content
Error Handling
Fields
Additional output fields
Spread sheet type (engine)
Excel 97-2003 XLS (JXL)
File or directory
C:/Users/Rico Renaldy/DataWarehouse/Praktikum 6 DwH/Daftar Siswa.xls
Add
Browse...
Regular Expression
Exclude Regular Expression
Password
Selected files:
File/Directory Wildcard (RegExp) Exclude wildcard Required Include subfolders
1 C:/Users/Rico Renaldy/DataWarehouse/Praktikum 6 DwH/Daftar Siswa.xls N N
Delete
Edit
Accept filenames from previous steps
Accept filenames from previous step
Step to read filenames from
Field in the input to use as filename
Show filename(s)...
Help
OK
Preview rows
Cancel
28°C
Cerah
Search
8:44
18/04/2024
Microsoft Excel input
Step name
Microsoft Excel input
Add Field(s)
Files
Sheets
Content
Error Handling
Fields
Additional output fields
List of sheets to read
Sheet name Start row Start column
1 Sheet1 0 0
Get sheetname(s)...
Help
OK
Preview rows
Cancel
28°C
Cerah
Search
8:45
18/04/2024


Microsoft Excel input
Step name
Microsoft Excel input
Files
Sheets
Content
Error Handling
Fields
Additional output fields
Name Type Length Precision Trim type Repeat Format Currency Decimal Grouping
1 NIS String -1 -1 none N
2 Nama String -1 -1 none N
3 Jenis Kelamin String -1 -1 none N
4 Kelas String -1 -1 none N
Get fields from header row...
Help
OK
Preview rows
Cancel
28°C
Cerah
Search
8:46
18/04/2024

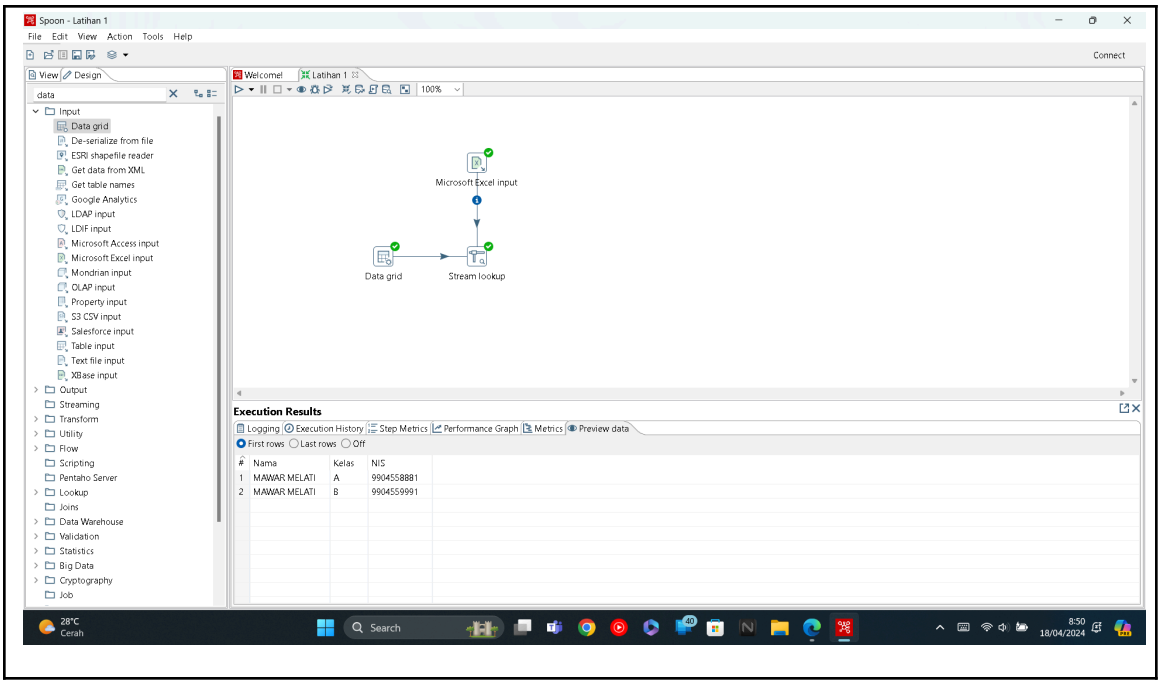
3. Data Grid.



4. Stream lookup (Get fields, Get lookup fields & hapus variabel yang tidak digunakan).

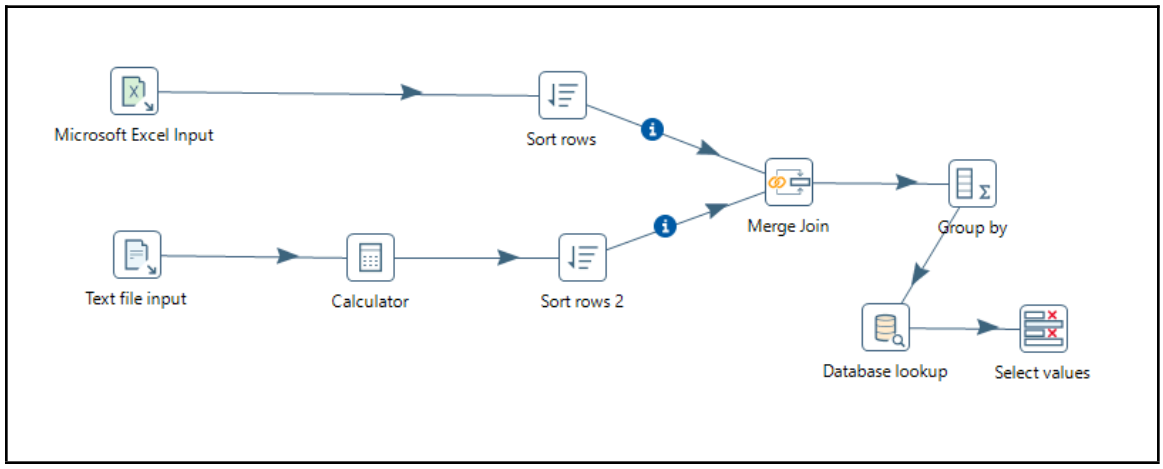


5. Ketika dijalankan maka outputnya akan seperti gambar berikut ini.

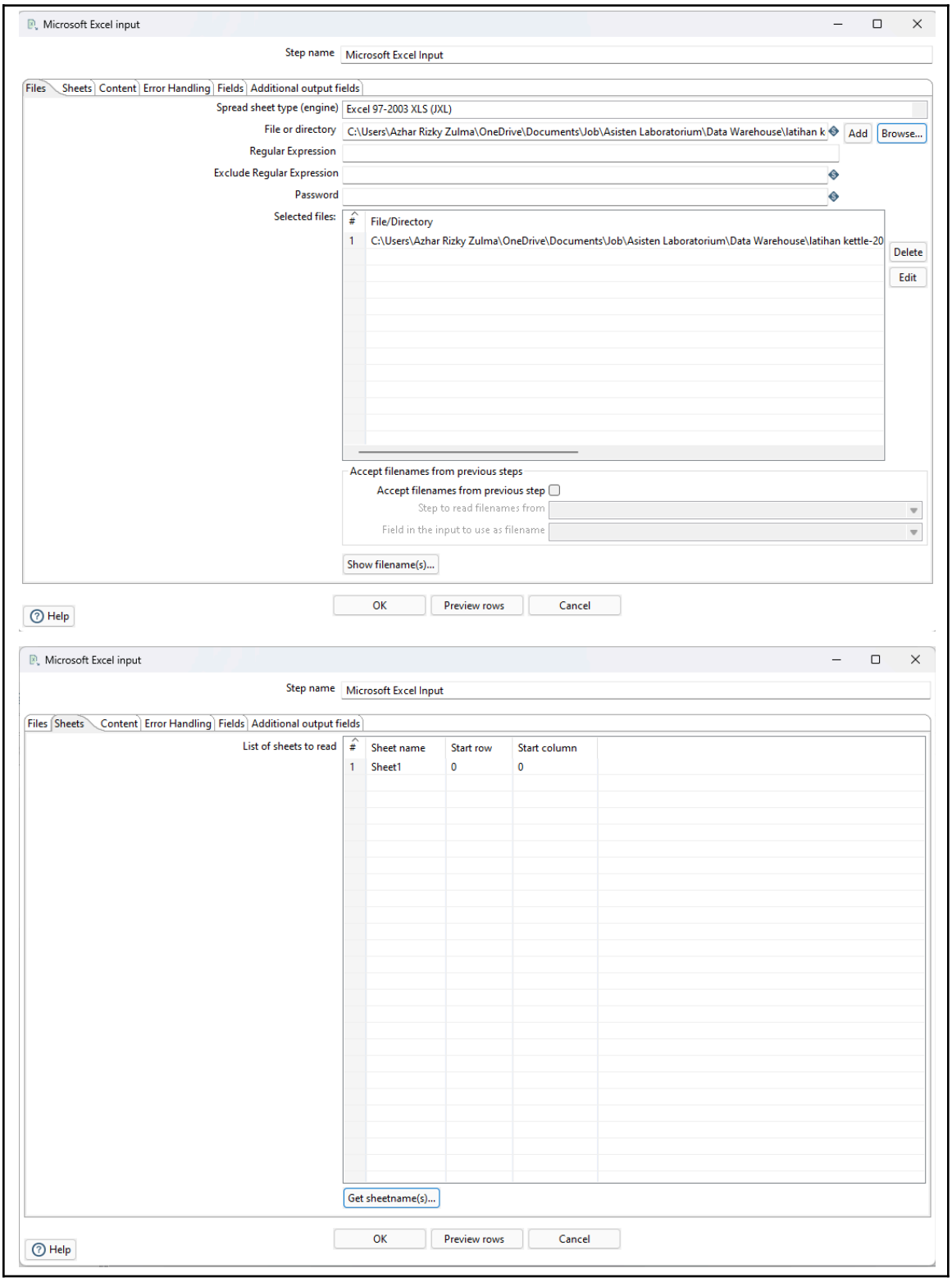


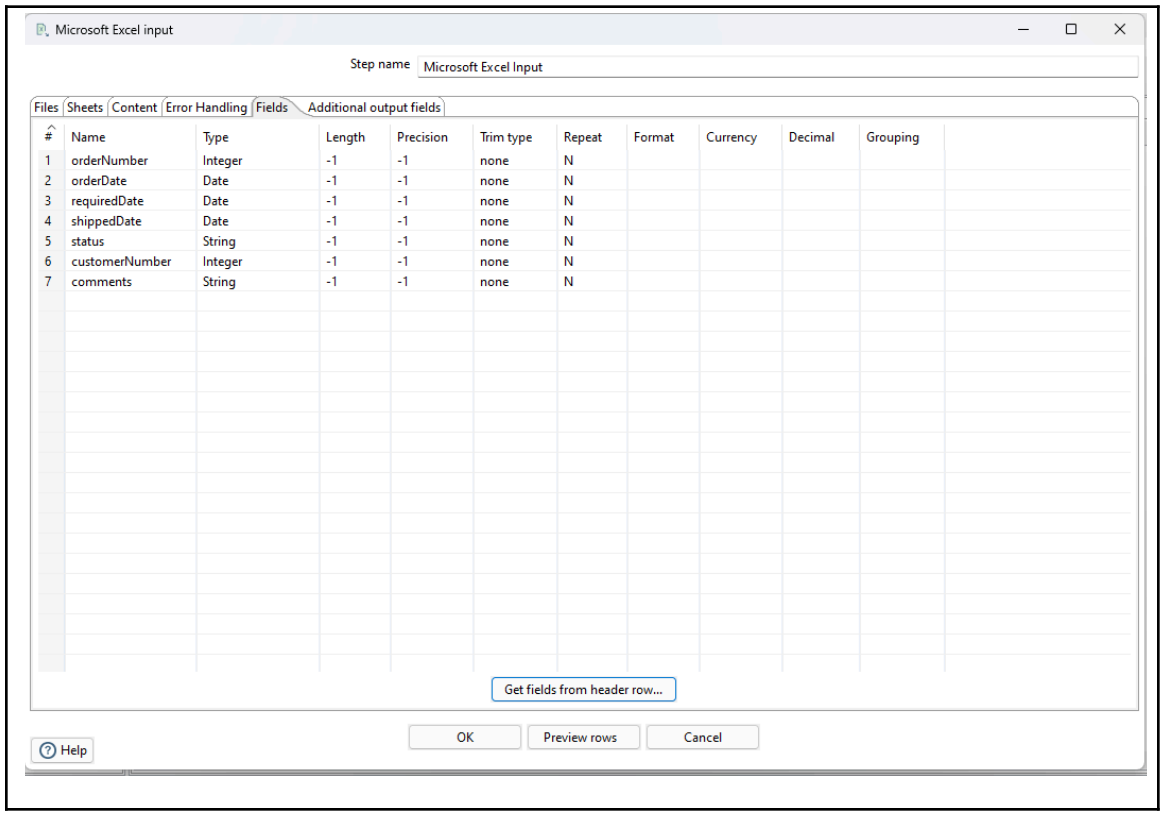
b. Latihan Kedua – Penggabungan dengan Merge dan Database Keylookup

1. Nyalakan Apache dan MySQL pada XAMPP dan buat transformasi baru dengan struktur sebagai berikut.

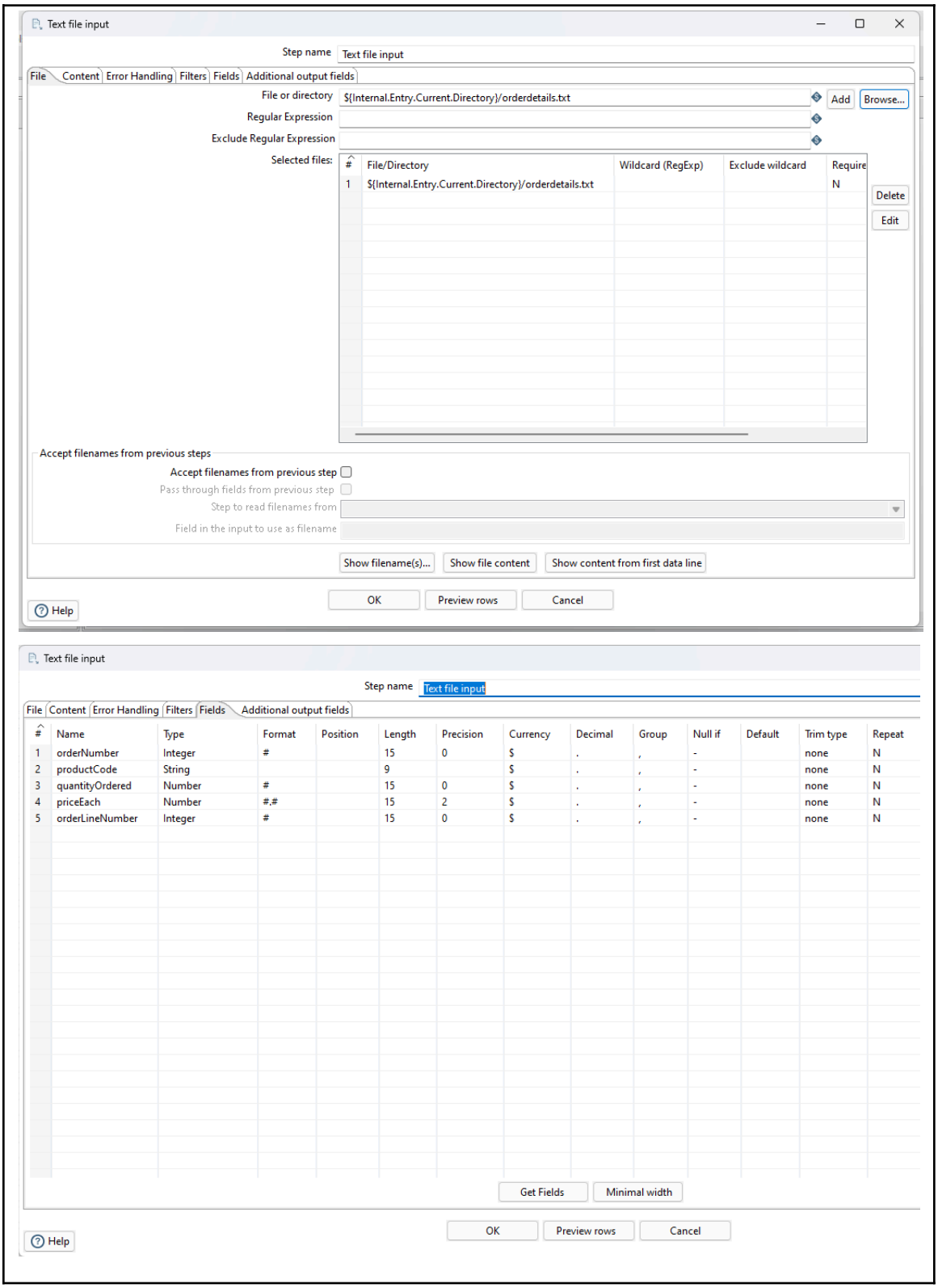


2. Microsoft Excel Input (Get sheetsname dan Get fields).

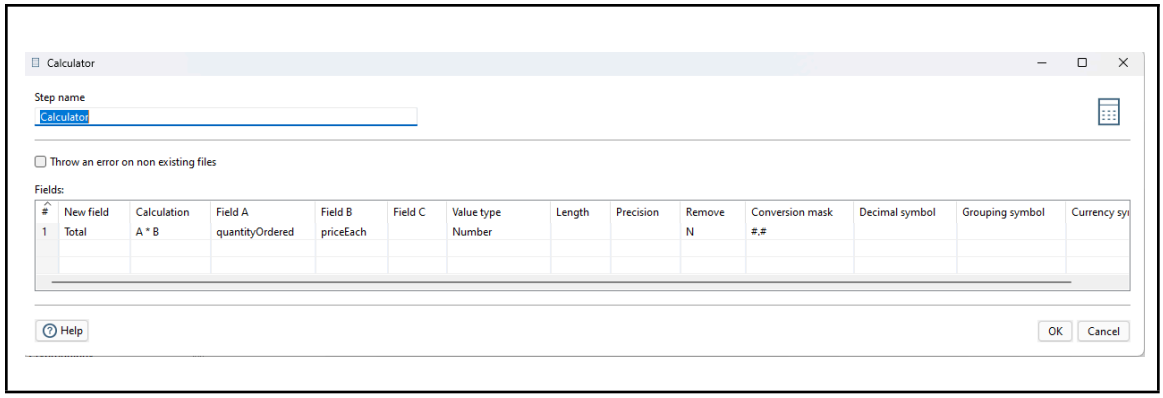

Microsoft Excel input
Step name
Microsoft Excel Input
Files
Sheets
Content
Error Handling
Fields
Additional output fields
Spread sheet type (engine)
Excel 97-2003 XLS (JXL)
File or directory
C:\Users\Azhar Rizky Zulma\OneDrive\Documents\Job\Asisten Laboratorium\Data Warehouse\latihan k
Add
Browse...
Regular Expression
Exclude Regular Expression
Password
Selected files:
File/Directory
1
C:\Users\Azhar Rizky Zulma\OneDrive\Documents\Job\Asisten Laboratorium\Data Warehouse\latihan kettle-20
Delete
Edit
Accept filenames from previous steps
Accept filenames from previous step
Step to read filenames from
Field in the input to use as filename
Show filename(s)...
Help
OK
Preview rows
Cancel
Microsoft Excel input
Step name
Microsoft Excel Input
Files
Sheets
Content
Error Handling
Fields
Additional output fields
List of sheets to read
Sheet name
Start row
Start column
1
Sheet1
0
0
Get sheetname(s)...
Help
OK
Preview rows
Cancel



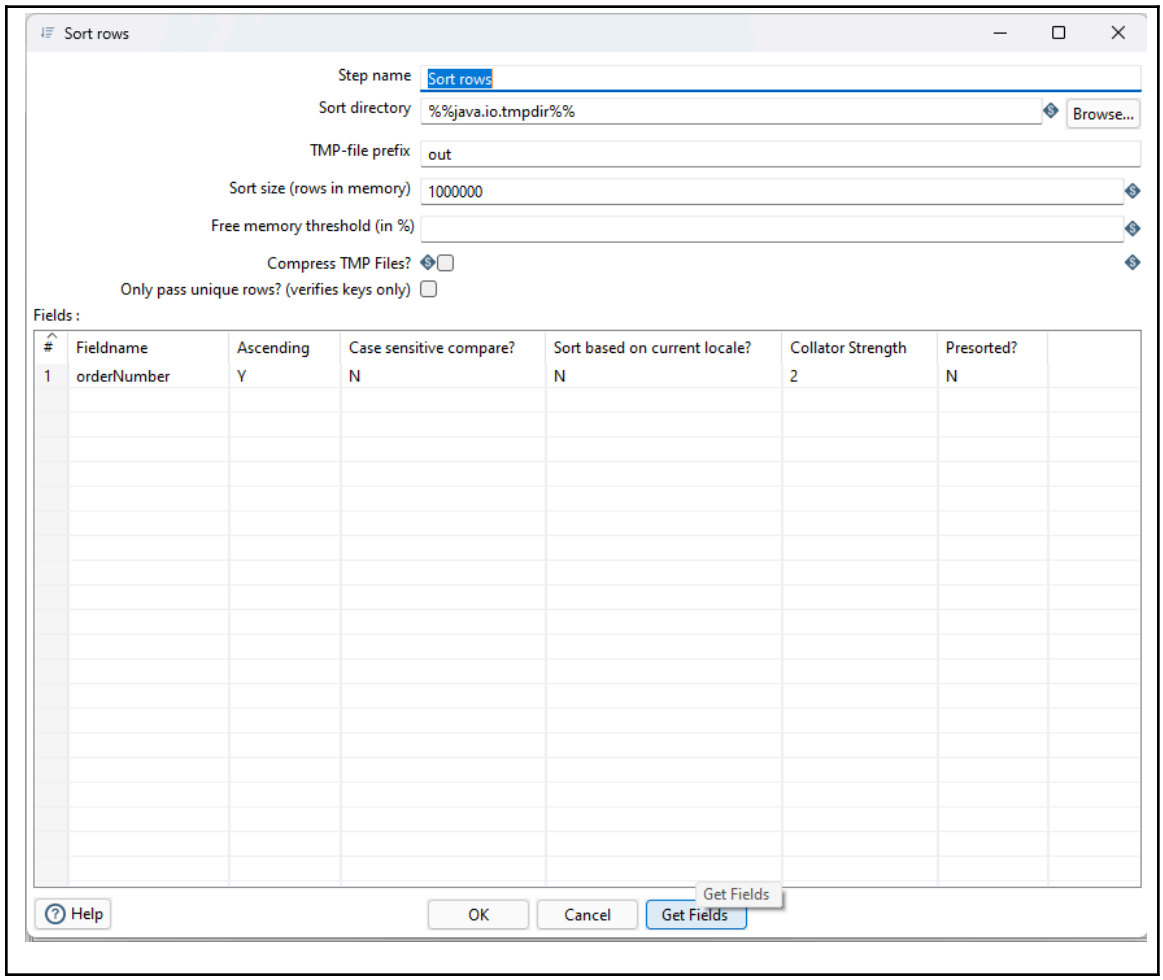
3. Text file input (Get fields pada tab fields).


Text file input
Step name
Text file input
File
Content
Error Handling
Filters
Fields
Additional output fields
File or directory
${Internal.Entry.Current.Directory}/orderdetails.txt
Add
Browse...
Regular Expression
Exclude Regular Expression
Selected files:
File/Directory Wildcard (RegExp) Exclude wildcard Require
1 ${Internal.Entry.Current.Directory}/orderdetails.txt N
Delete
Edit
Accept filenames from previous steps
Accept filenames from previous step
Pass through fields from previous step
Step to read filenames from
Field in the input to use as filename
Show filename(s)...
Show file content
Show content from first data line
OK
Preview rows
Cancel
Help
Text file input
Step name
Text file input
File
Content
Error Handling
Filters
Fields
Additional output fields
Name Type Format Position Length Precision Currency Decimal Group Null if Default Trim type Repeat
1 orderNumber Integer # 15 0 $. , - none N
2 productCode String 9 $. , - none N
3 quantityOrdered Number # 15 0 $. , - none N
4 priceEach Number #.# 15 2 $. , - none N
5 orderLineNumber Integer # 15 0 $. , - none N
Get Fields
Minimal width
OK
Preview rows
Cancel
Help

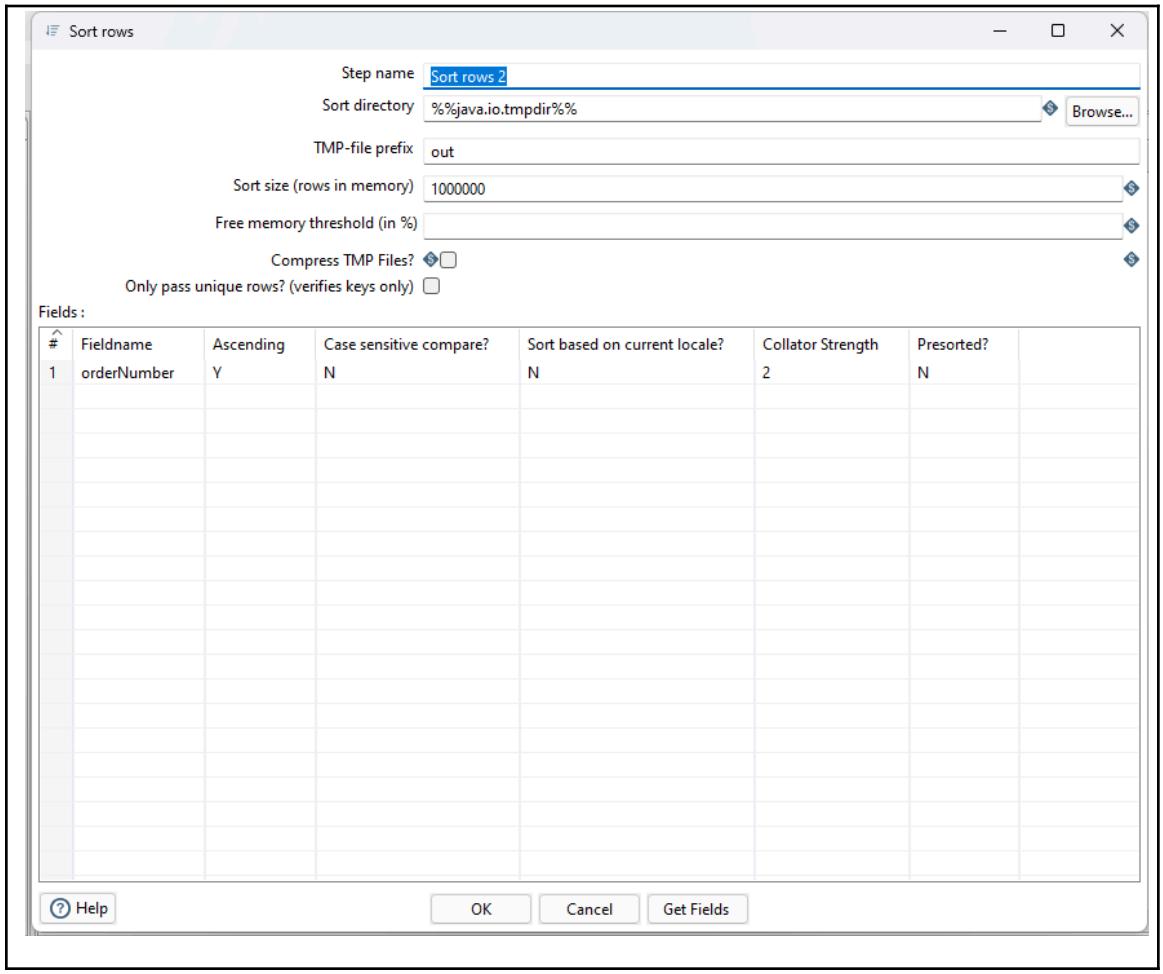
4. Calculator.



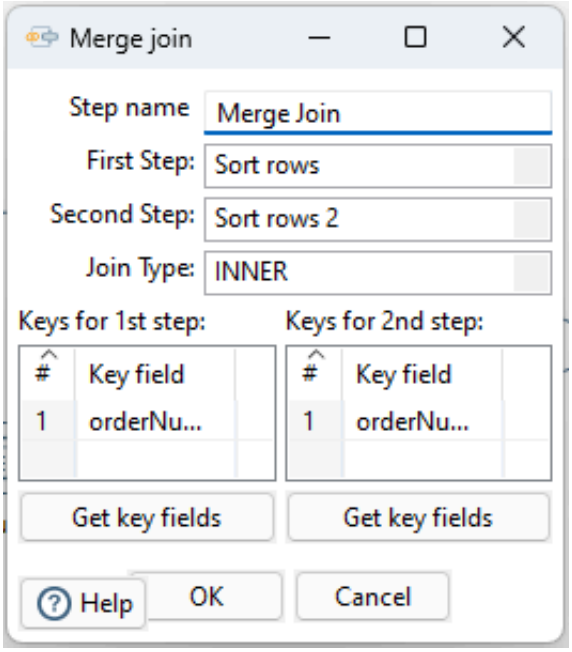
5. Sort rows (Get fields)


Sort rows
Step name
Sort rows
Sort directory
%%java.io.tmpdir%%
Browse...
TMP-file prefix
out
Sort size (rows in memory)
1000000
Free memory threshold (in %)
Compress TMP Files?
Only pass unique rows? (verifies keys only)
Fields :
Fieldname Ascending Case sensitive compare? Sort based on current locale? Collator Strength Presorted?
1 orderNumber Y N N 2 N
Help
OK
Cancel
Get Fields
Get Fields

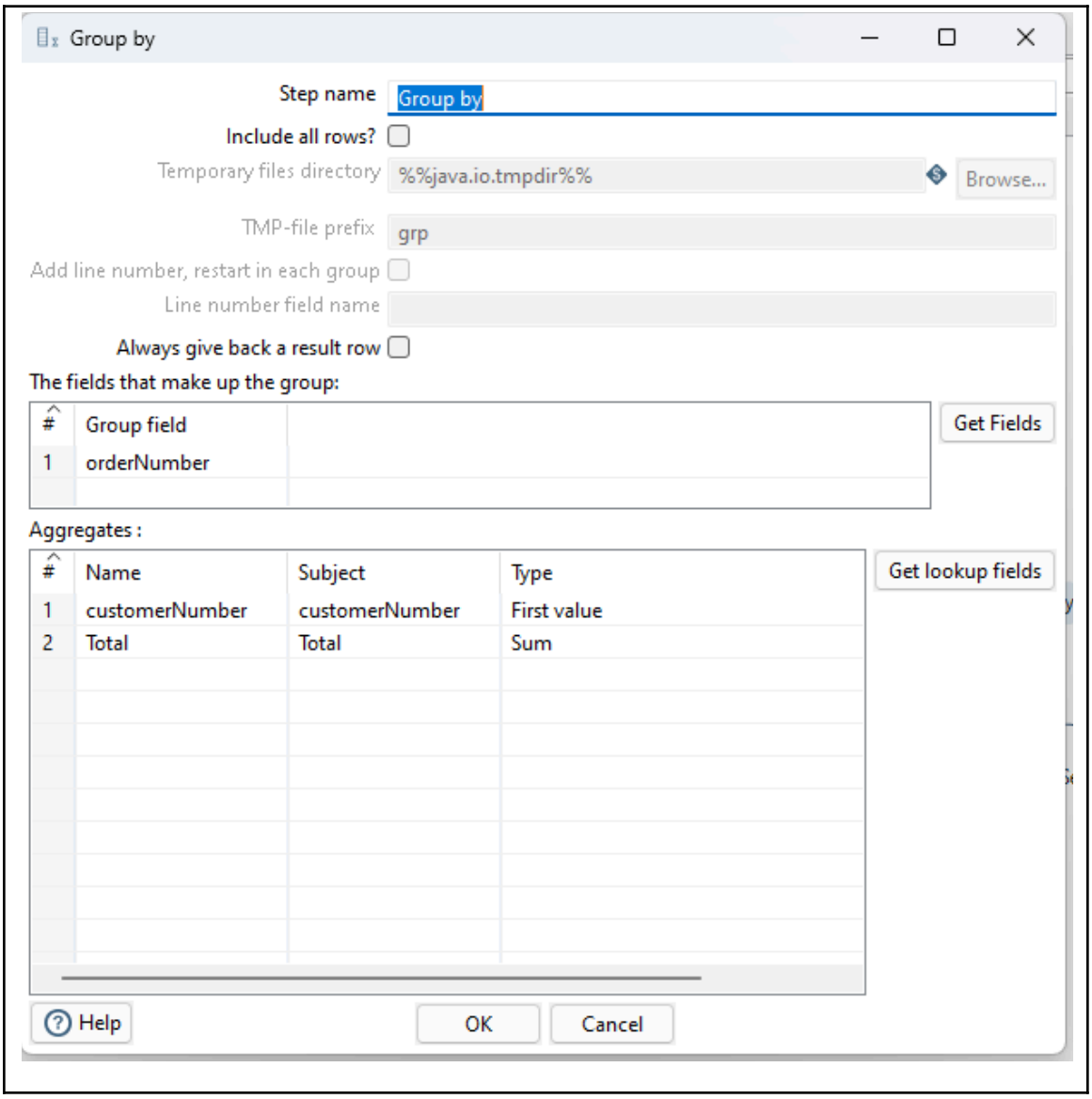
6. Sort rows 2 (Get fields 2).



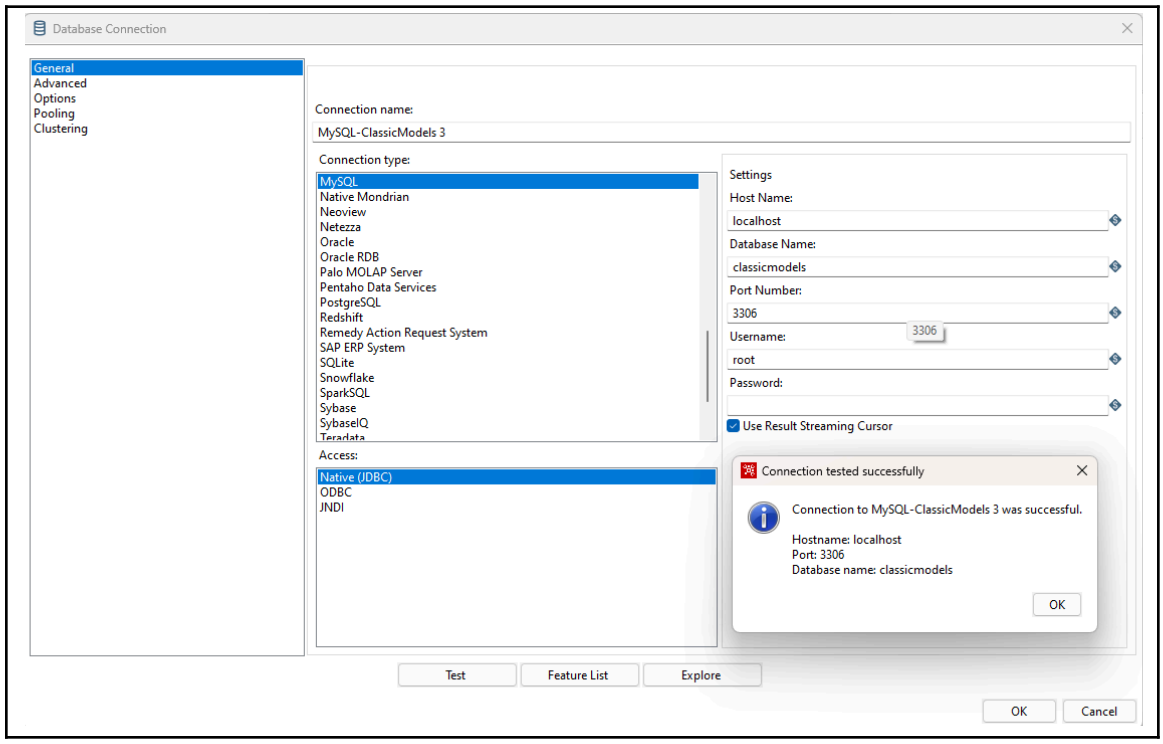
7. Merge join.

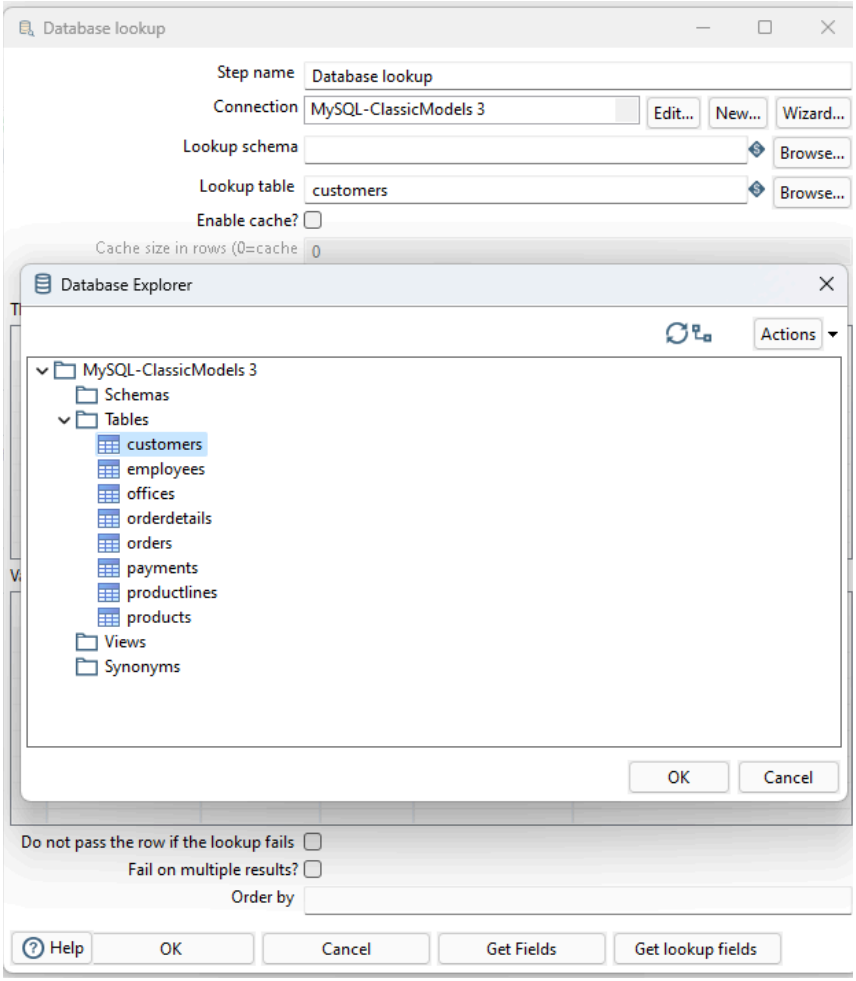


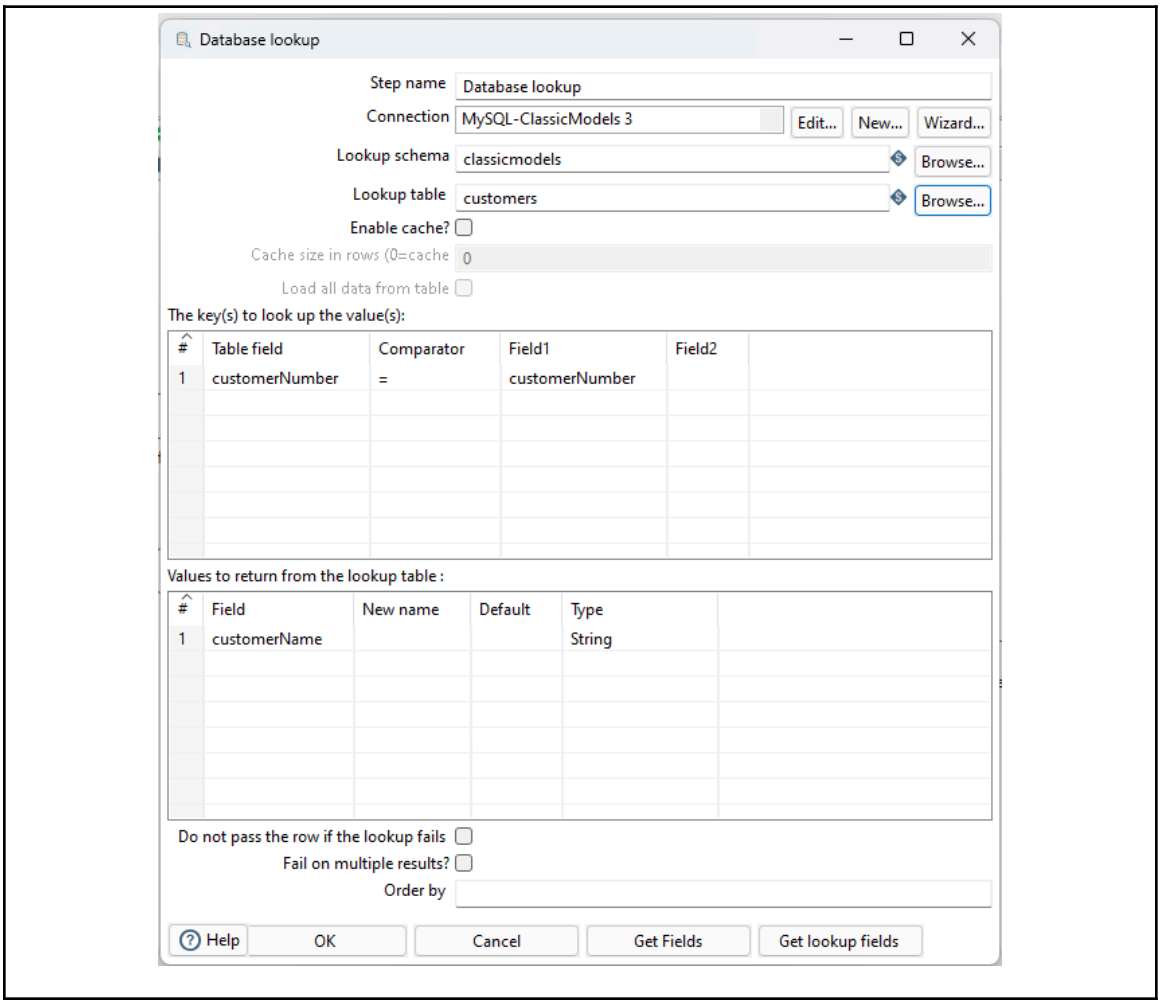
8. Group by (Get fields & Get lookup fields).



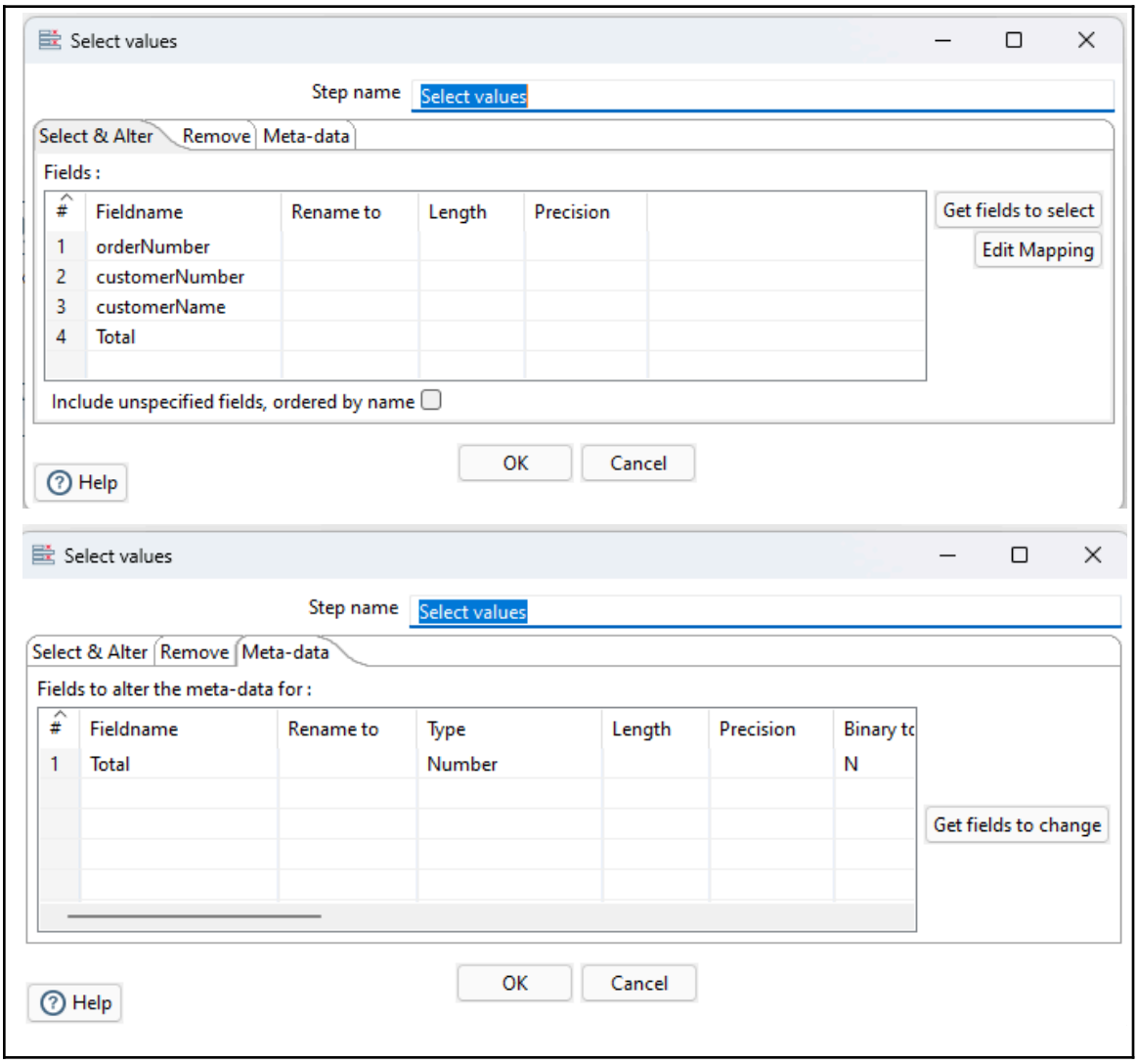
9. Database lookup (Database connection, Database explorer)


Database Connection
General
Advanced
Options
Pooling
Clustering
Connection name:
MySQL-ClassicModels 3
Connection type:
MySQL
Native Mondrian
Neoview
Netezza
Oracle
Oracle RDB
Palo MOLAP Server
Pentaho Data Services
PostgreSQL
Redshift
Remedy Action Request System
SAP ERP System
SQLite
Snowflake
SparkSQL
Sybase
SybaseIQ
Teradata
Access:
Native (JDBC)
ODBC
JNDI
Settings
Host Name:
localhost
Database Name:
classicmodels
Port Number:
3306
3306
Username:
root
Password:
Use Result Streaming Cursor
Connection tested successfully
Connection to MySQL-ClassicModels 3 was successful.
Hostname: localhost
Port: 3306
Database name: classicmodels
OK
Test
Feature List
Explore
OK
Cancel

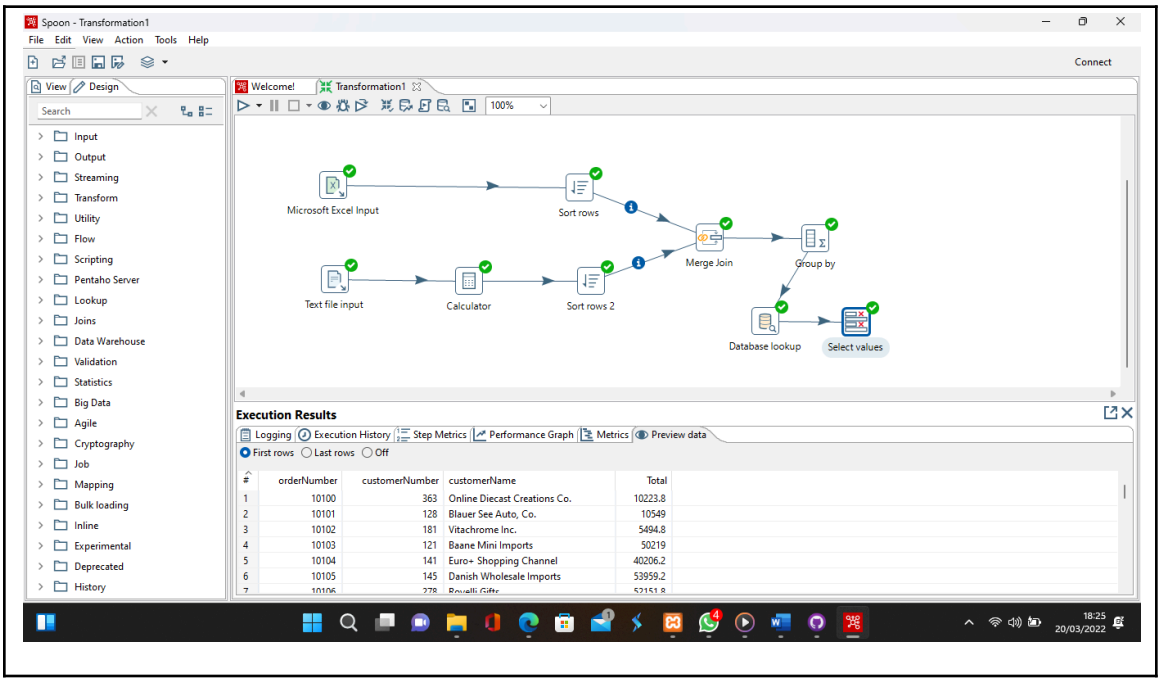

Database lookup
Step name
Database lookup
Connection
MySQL-ClassicModels 3
Edit...
New...
Wizard...
Lookup schema
Browse...
Lookup table
customers
Browse...
Enable cache?
Cache size in rows (0=cache
0
Database Explorer
Actions
MySQL-ClassicModels 3
Schemas
Tables
customers
employees
offices
orderdetails
orders
payments
productlines
products
Views
Synonyms
OK
Cancel
Do not pass the row if the lookup fails
Fail on multiple results?
Order by
Help
OK
Cancel
Get Fields
Get lookup fields



8. Select values (Get fields to select, Get fields to change).


Select values
Step name
Select values
Select & Alter
Remove
Meta-data
Fields :
Fieldname Rename to Length Precision
1 orderNumber
2 customerNumber
3 customerName
4 Total
Get fields to select
Edit Mapping
Include unspecified fields, ordered by name
OK
Cancel
Help
Select values
Step name
Select values
Select & Alter
Remove
Meta-data
Fields to alter the meta-data for :
Fieldname Rename to Type Length Precision Binary tc
1 Total Number N
Get fields to change
OK
Cancel
Help

9. Output



4. File Praktikum

Github Repository:

5. Soal Latihan

Soal:

1. Apa yang dimaksud dengan Merge Data?
2. Apa tujuan dan fungsi dari menggabungkan data?

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 (✔)

8.

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	**…**	
2.	Latihan Kedua	**…**	

9. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	… Menit	…
2.	Latihan Kedua	… Menit	…

Keterangan:

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