



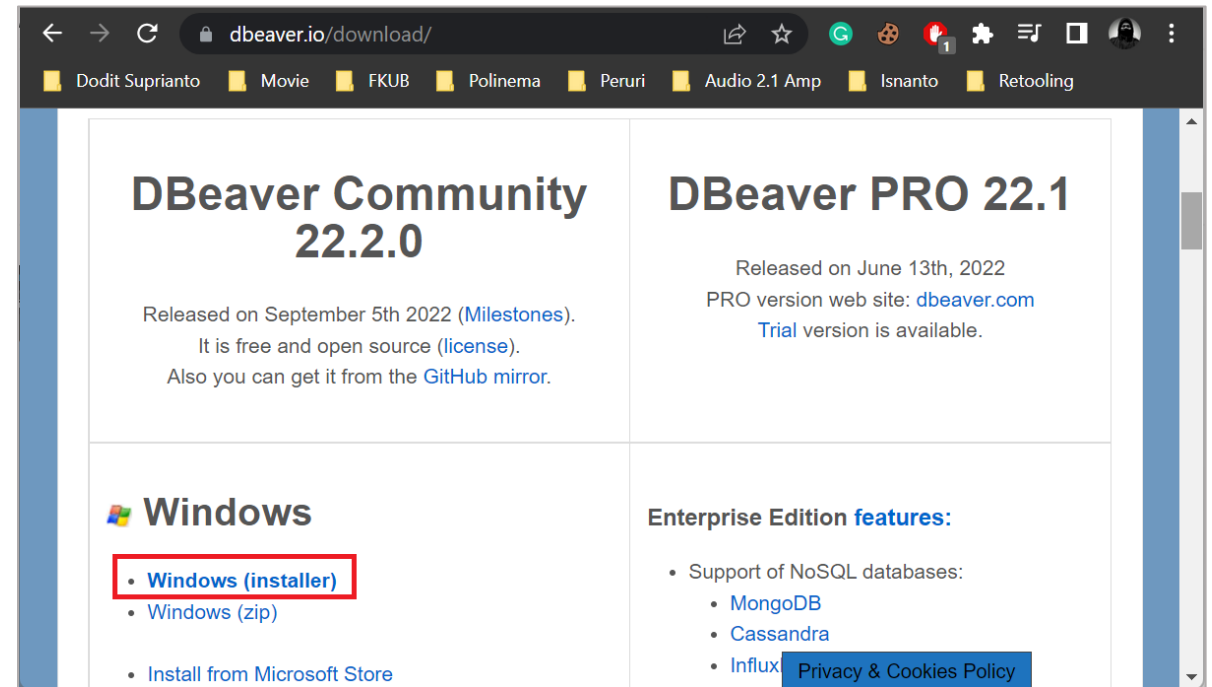
Koneksi Pentaho Ke Database MySQL

Tujuan

- Mentransformasikan file CSV ke database MySQL
- Menarik data yang bersumber dari database MySQL.
- Pentaho membutuhkan driver connection agar Pentaho dapat berkomunikasi dengan database MySQL.
- Membuat file transformasi data dari MySQL ke versi Pentaho.

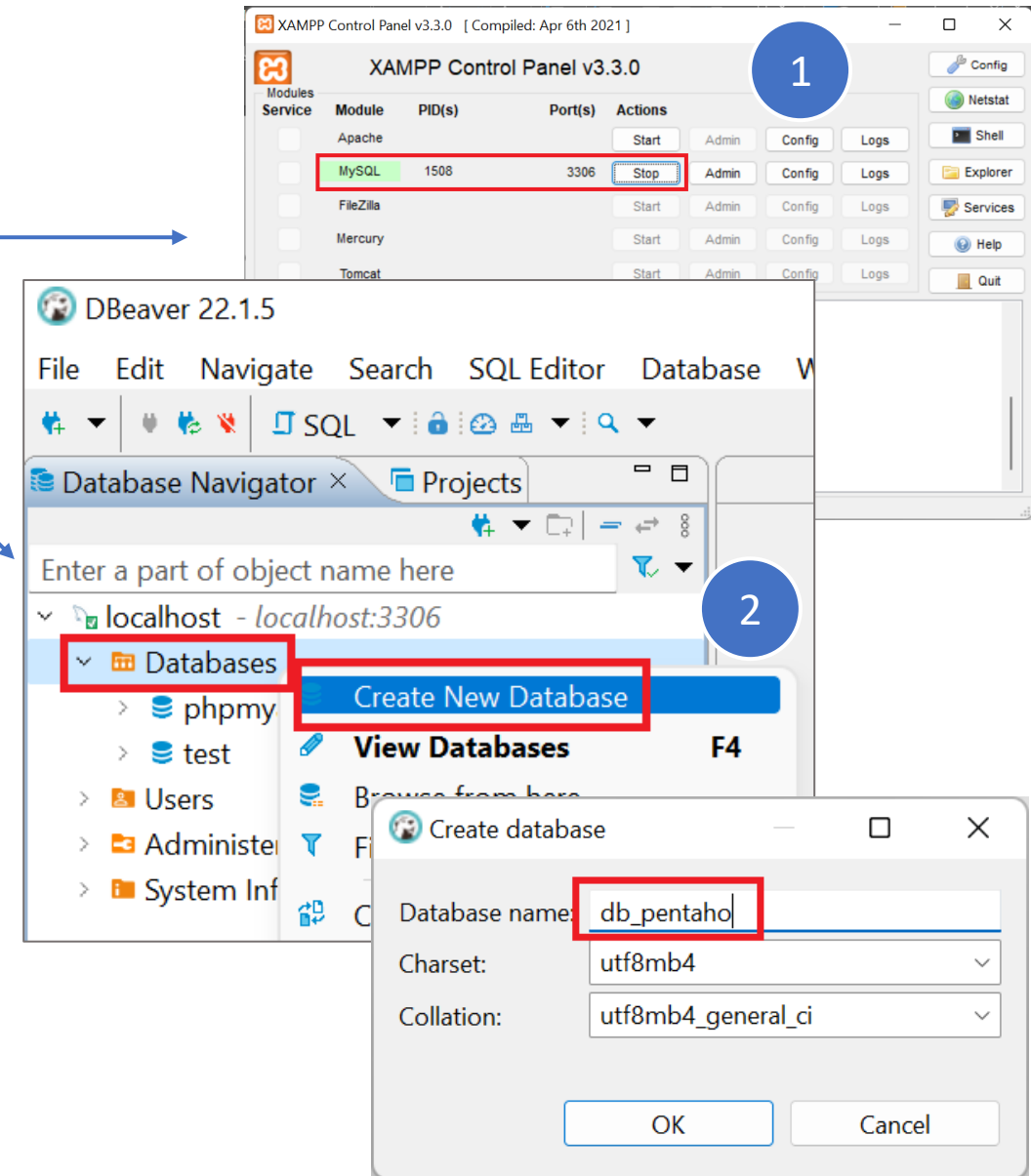
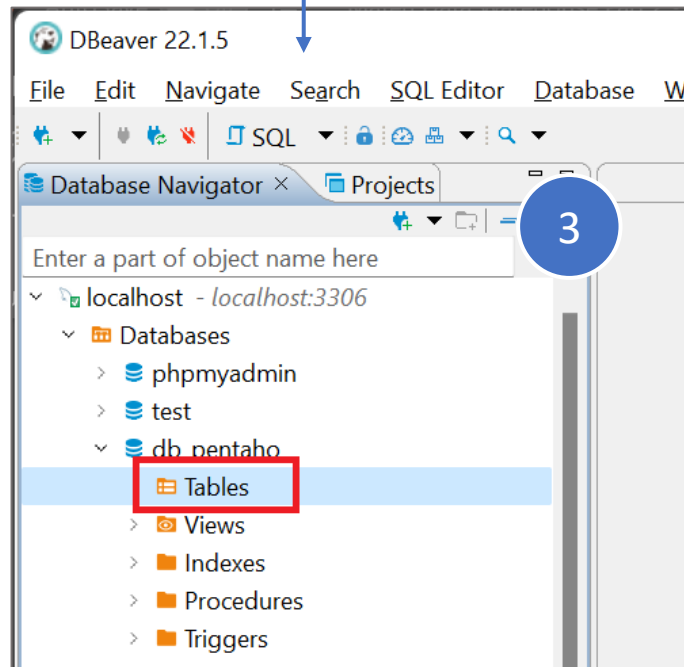
MySQL Database + Tools DBeaver

- Pastikan computer telah terinstall MySQL Server dan Service Running (bisa menggunakan XAMPP)
- Gunakan database tool management software seperti DBeaver atau lainnya (https://dbeaver.io/files/dbeaver-ce-latest-x86_64-setup.exe)



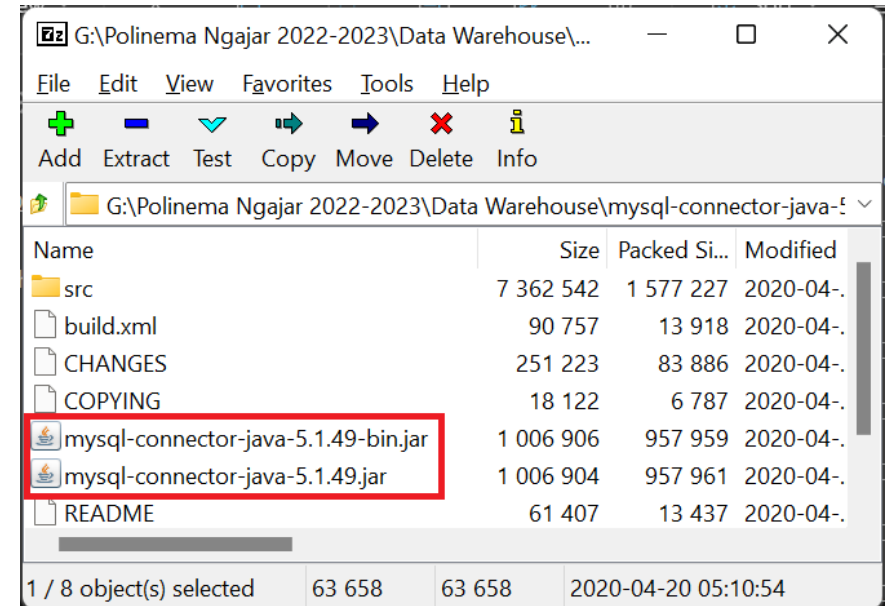
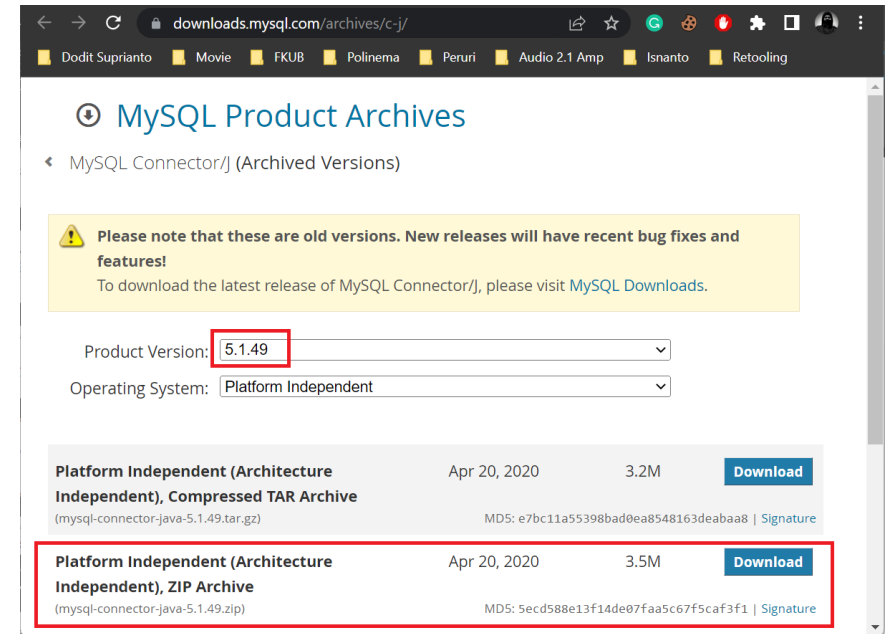
Create Database

- Aktifkan service DB MySQL
- Buat database 'db_pentaho' dengan Dbeaver.
- Biarkan **table** kosong karena akan diimport dari file CSV luar.



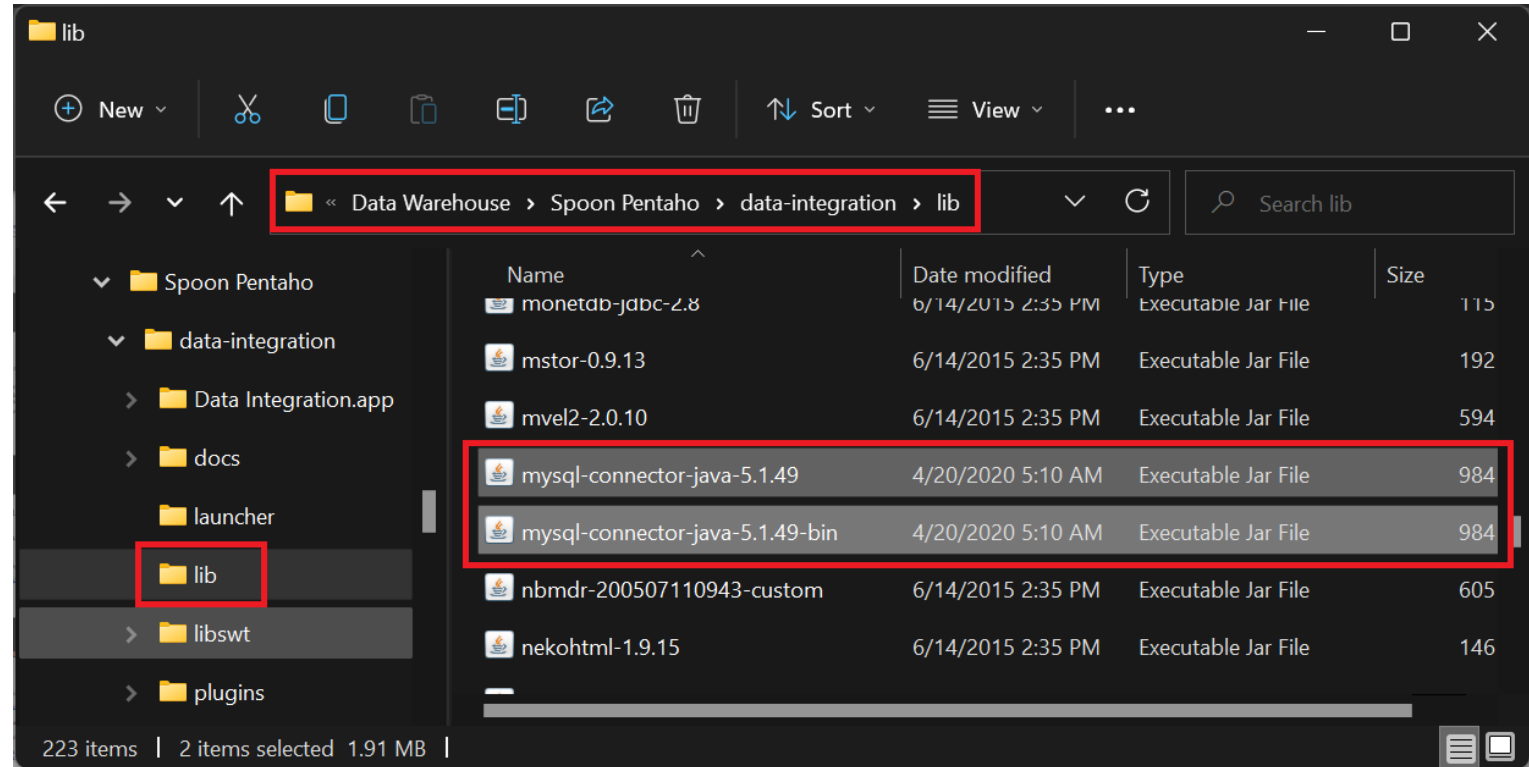
MySQL Connector

- Download MySQL Connector Versi 5.1.49 for Java
(<https://downloads.mysql.com/archives/c-j/>)
- Unzip **mysql-connector-java-5.1.49.zip**



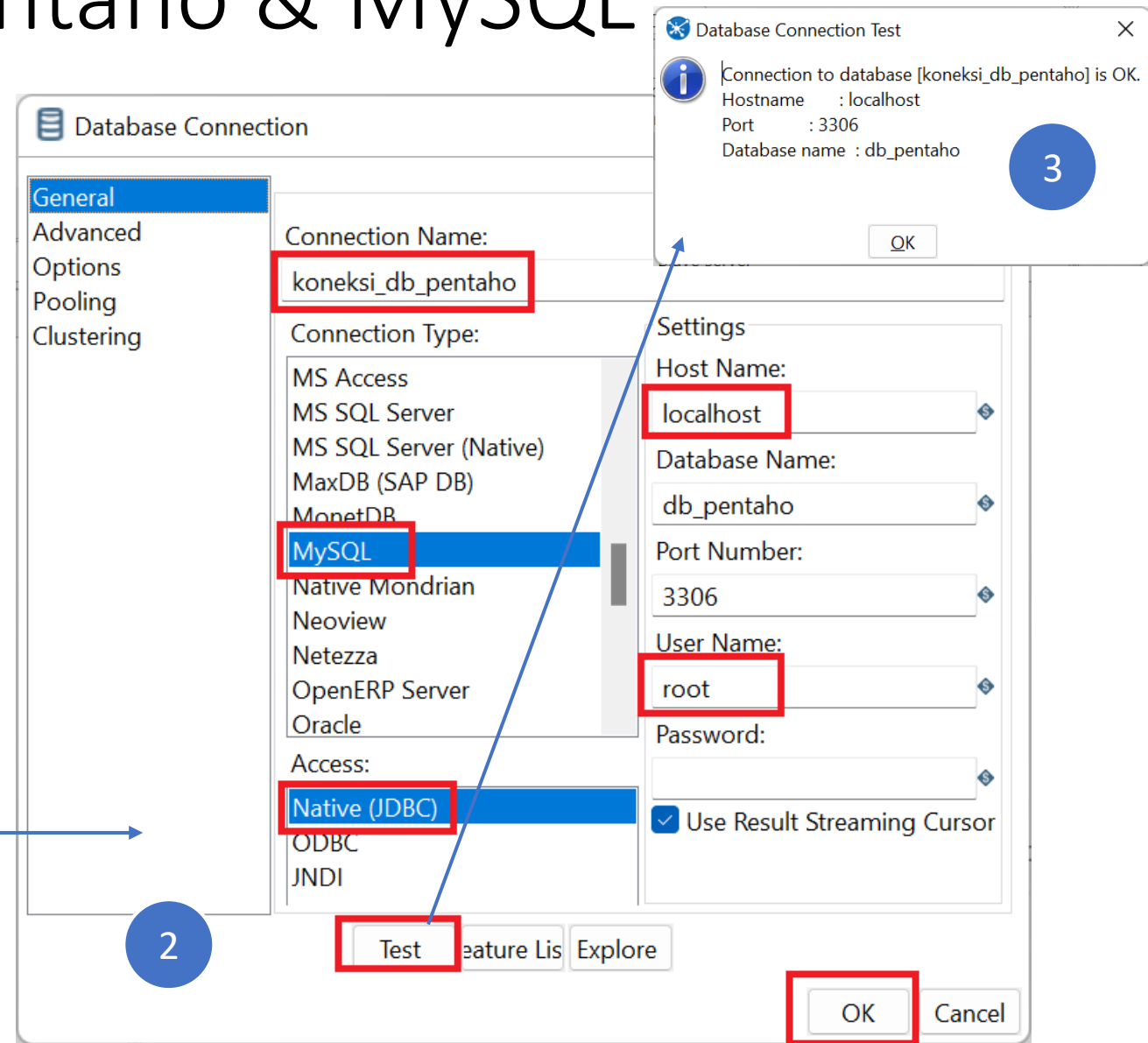
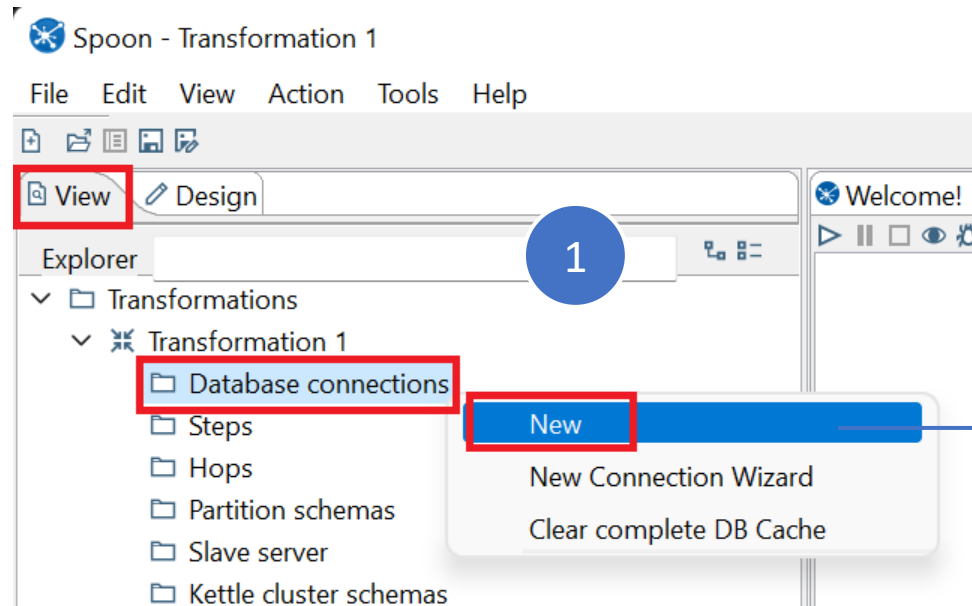
MySQL Connector

- Ekstrak 2 file berikut di folder '**\data-integration\lib**':
 - **mysql-connector-java-5.1.49.zip**
 - **mysql-connector-java-5.1.49-bin.zip**



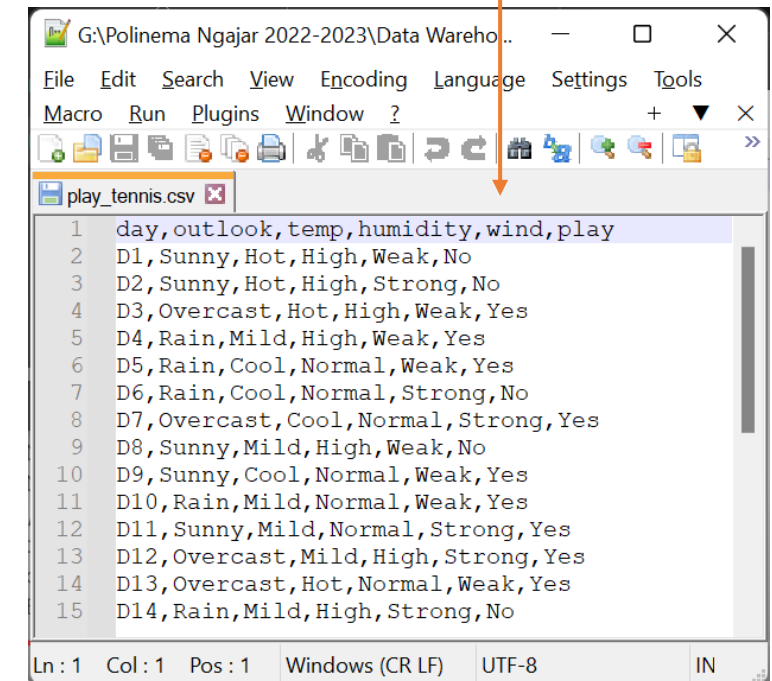
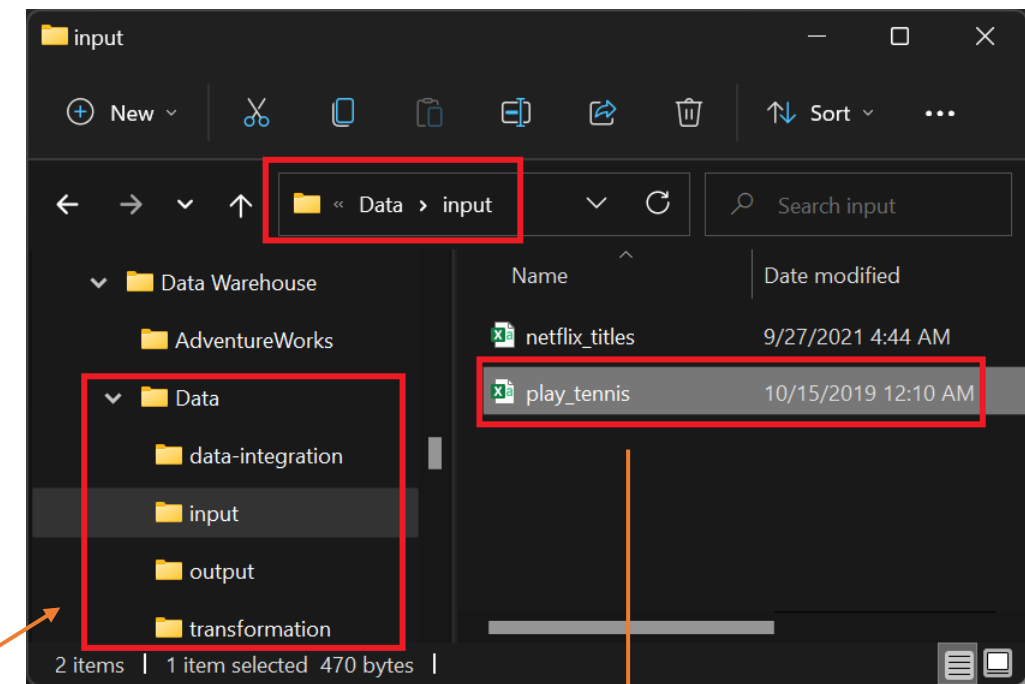
Membuat Koneksi Pentaho & MySQL

- Buka menu **File > new > Transformation**
- Pilih tab **View > Database Connections > New.**
- Jika ERROR, uninstall JDK11 dan install JDK8



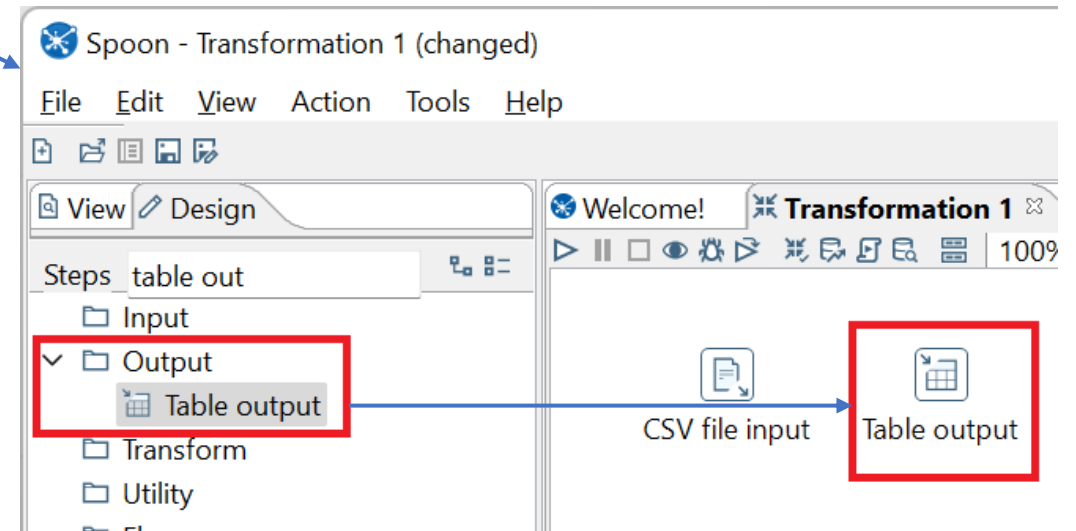
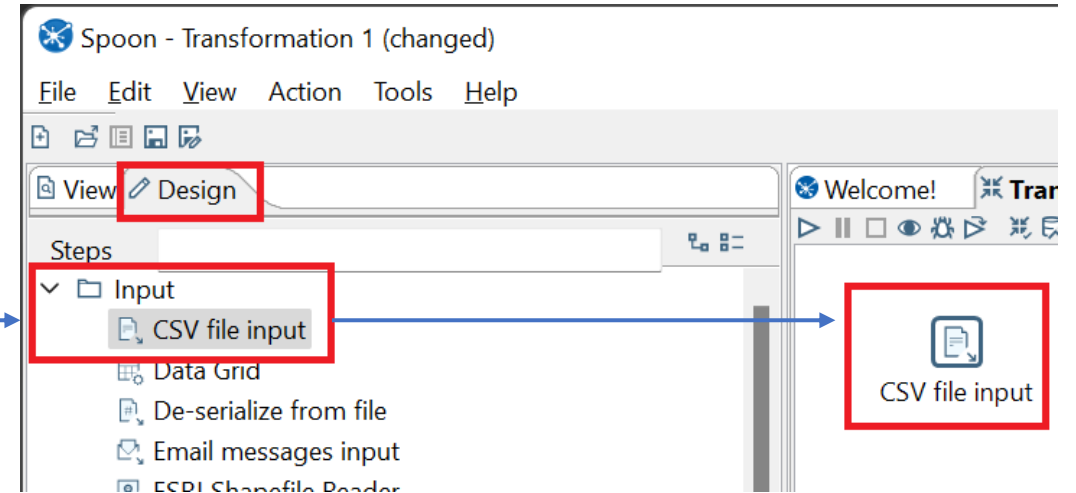
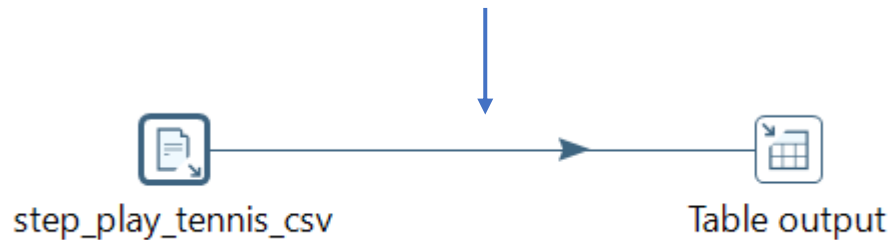
File Sumber CSV

- Download file CSV di <https://www.kaggle.com/datasets/fredericobreno/play-tennis/download?datasetVersionNumber=1>
- Uraikan filenya 'archive' > 'play_tennis' dan letakkan di folder '**Data\input**' seperti project sebelumnya.



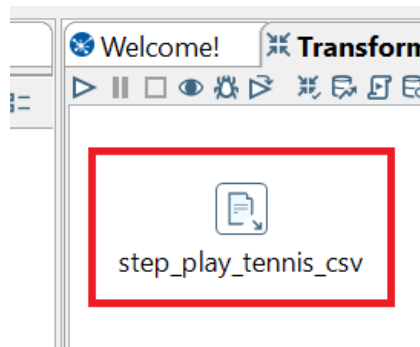
Transformation

- Buka menu **Design > Input > CSV file input** dan drag-drop ke kanan.
- Buka menu **Design > Input > CSV file input** dan drag-drop ke kanan
- Buat hubungan antara 2 step tersebut dengan garis hop.



Transformation

- Konfigurasi **CSV file input**



CSV Input

Step name: step_play_tennis_csv

Filename: G:\Polinema Ngajar 2022-2023\Data Warehouse\Data\input\play_ Browse...

Delimiter: , Insert TAB

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

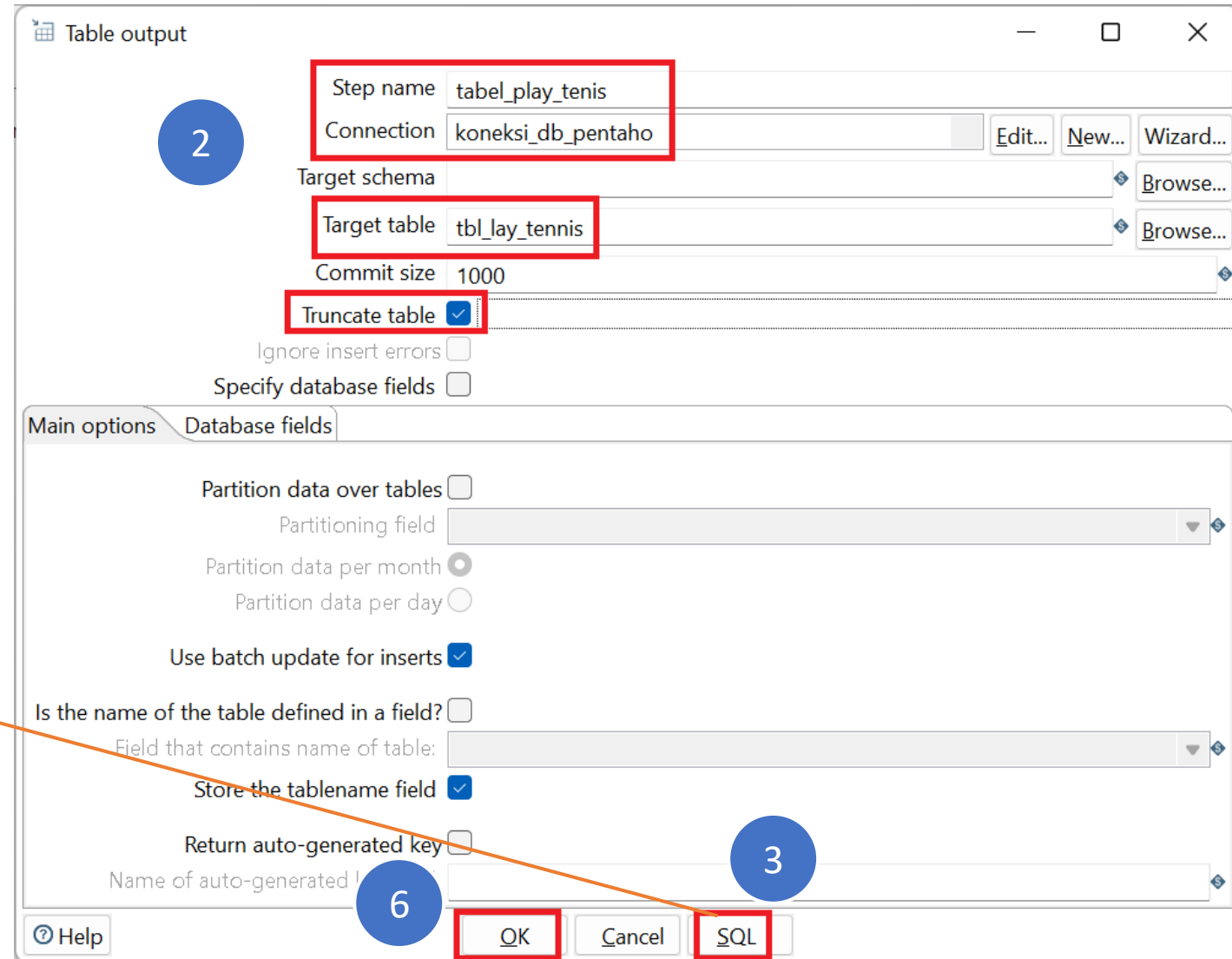
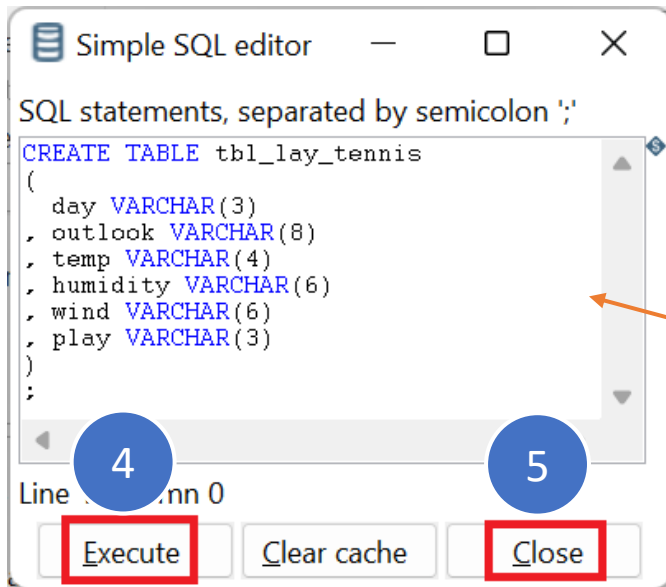
File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	day	String		3		\$.	,	none

Help OK Get Fields Preview Cancel

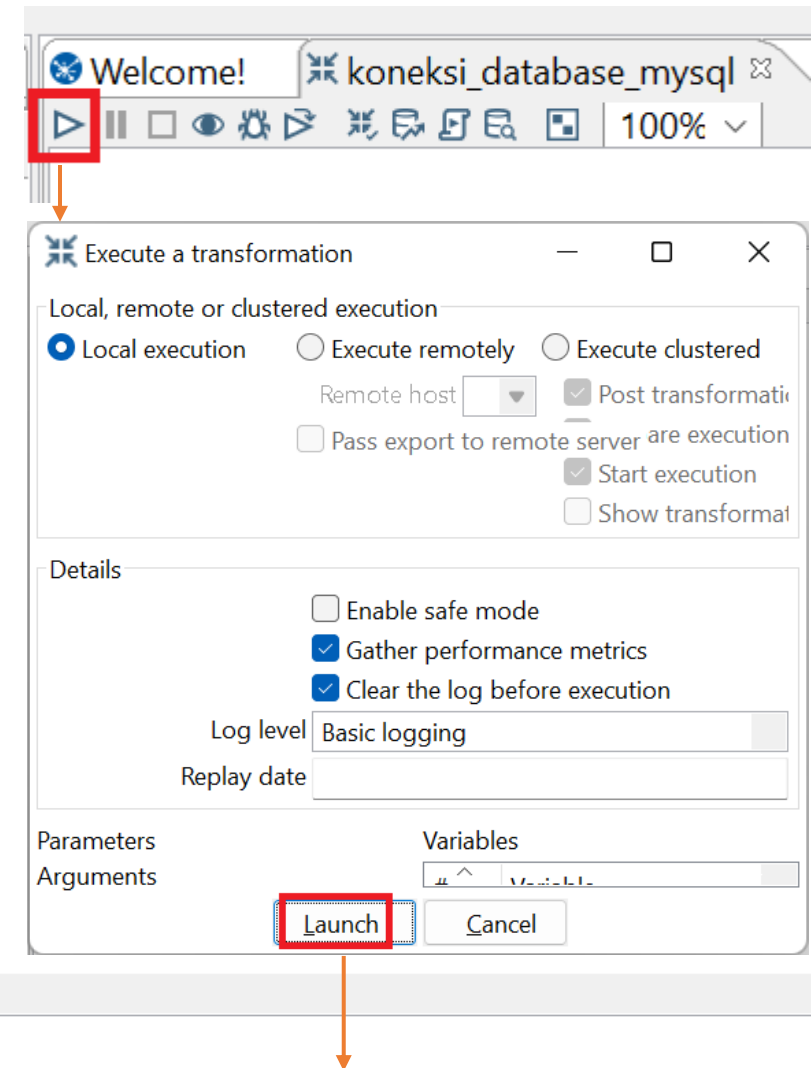
Transformation

- Konfigurasi Table Output



Transformation

- Simpan file transformation dengan nama “**koneksi_database_mysql**” di foler ‘transformation’ yang telah kita buat sebelumnya.
- Jalankan tombol **Run** dan **Launch**



Execution Results													
Execution History Logging Step Metrics Performance Graph Metrics Preview data													
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	step_play_tennis_csv	0	0	14	15	0	0	0	0	Finished	0.0s	7,500	-
2	tabel_play_tenis	0	14	14	0	14	0	0	0	Finished	0.1s	250	-

Transformation

- Buka hasil transformasi di DB MySQL dengan DBeaver.
- Refresh terlebih dahulu database '**db_pentaho**'.
- Perhatikan jumlah baris/record antara MySQL dan CSV apakah sama?

The screenshot shows the DBeaver 22.1.5 interface. In the left sidebar, the 'Database Navigator' panel shows the 'db_pentaho' database selected, with the 'tbl_lay_tennis' table highlighted. The 'tbl_lay_tennis' table is expanded, showing its structure: Columns, Constraints, Foreign Keys, References, Triggers, Indexes, and Partitions. The 'Data' tab is selected, displaying a table with 14 rows and 6 columns: day, outlook, temp, humidity, and wind. The data is as follows:

	day	outlook	temp	humidity	wind
1	D1	Sunny	Hot	High	Weak
2	D2	Sunny	Hot	High	Strong
3	D3	Overcast	Hot	High	Weak
4	D4	Rain	Mild	High	Weak
5	D5	Rain	Cool	Normal	Weak
6	D6	Rain	Cool	Normal	Strong
7	D7	Overcast	Cool	Normal	Strong
8	D8	Sunny	Mild	High	Weak
9	D9	Sunny	Cool	Normal	Weak
10	D10	Rain	Mild	Normal	Weak
11	D11	Sunny	Mild	Normal	Strong
12	D12	Overcast	Mild	High	Strong

The status bar at the bottom indicates '14 row(s) fetched - 2ms (1ms fetch), on 2022-09-10'.



Tugas

Tugas

1. Buat transformasi dari CSV ke DB MySQL
2. Buat transformasi dari JSON ke DB MySQL
3. Buat transformasi dari XML ke DB MySQL

Catatan:

- Poin 1, database MySQL sumbernya bebas silahkan dicari.
- Poin 2 & 3, file JSON dan XML dapat diperoleh dari dataset Kaggle (bebas).
- Kumpulkan di LMS