

LAPORAN PRAKTIKUM JOBSHEET 1  
DATA WAREHOUSE



Disusun Oleh:  
Niriza Lailaumi Hidayat (2341760072)  
SIB 2B

SISTEM INFORMASI BISNIS  
JURUSAN TEKNOLOGI INFORMASI  
POLITEKNIK NEGERI MALANG

2025

## Data Penjualan

1	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
2	Customer 1	Yaris AT	MLG	3/13/2019	340.95	6	2045.7
3	Customer 1	Etios AT	SBY	4/16/2019	799.95	1	799.95
4	Customer 3	Avanza AT	MLG	6/18/2019	168.95	7	1182.65
5	Customer 3		MLG	7/24/2019	168.95	2	337.9
6	Customer 1	Avanza AT	SBY	8/21/2019	168.95	9	1520.55
7	Customer 2	Etios AT	KDR	9/7/2019	799.95	7	5599.65
8	Customer 3	Etios AT	MLG	9/7/2019	799.95	8	6399.6
9	Customer 2	Avanza AT			168.95	15	2534.25
10	Customer 1	Avanza AT	SBY	9/14/2019	168.95	15	2534.25
11	Customer 1	Etios AT	SBY	9/25/2019	799.95	2	1599.9
12	Customer 1	Avanza AT	SBY	1/21/2020	168.95	14	2365.3
13	Customer 1	Yaris AT	MLG	2/1/2020	340.95	2	681.9
14	Customer 1				340.95	7	2386.65
15	Customer 1	Etios AT	SBY	5/30/2020	799.95	9	7199.55
16	Customer 3	Avanza AT	MLG	6/7/2020	168.95	10	1689.5
17	Customer 3	Yaris AT	MLG	6/9/2020	340.95	4	1363.8
18	Customer 2		KDR	8/7/2020	799.95	14	11199.3
19	Customer 2	Avanza AT	KDR	8/8/2020	168.95	12	2027.4
20	Customer 1	Avanza AT		8/12/2020	168.95	4	675.8

### Tugas 1

- Berapa jumlah kolom pada data tersebut?
  - Kolom pada data berjumlah 7 yaitu kolom Customer, Product, Region, Date, Item unit price, No. Items, dan Total Sale.
- Apa arti atau isi data dari setiap kolom yang ada?
  - Customer: berisi data pelanggan yang melakukan pembelian.
  - Product: berisi data produk yang dibeli pelanggan.
  - Region: berisi data lokasi atau wilayah tempat transaksi terjadi.
  - Date: berisi data tanggal transaksi dilakukan dalam format MM/DD/YYYY.
  - Item unit price: berisi data harga satuan produk yang dibeli.
  - No. Items: berisi data jumlah produk yang dibeli dalam transaksi.
  - Total Sale: berisi data total penjualan dalam transaksi (Item unit price x No. Items)
- Adakah data yang memiliki nilai null/data yang tidak lengkap?
  - Ada, data yang memiliki nilai null atau tidak lengkap antara lain:
    - Data Product pada baris ke 4.

- b. Data Region dan Date pada baris ke 8.
  - c. Data Product, Region, dan Date pada baris ke 13.
  - d. Data Product pada baris ke 17.
  - e. Data Region pada baris ke 19.
4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?
  - Kolom Customer: semua data memiliki tipe data yang sama yaitu teks/String.
  - Kolom Product: beberapa baris pada kolom ini bernilai null sehingga dapat menyebabkan inkonsistensi tipe data.
  - Kolom Region: data seharusnya berisi kode wilayah (MLG, SBY, KDR), tetapi ada beberapa baris kosong yang menyebabkan ketidaksesuaian tipe data.
  - Kolom Date: data memiliki format tanggal MM/DD/YYYY, namun terdapat kemungkinan bahwa beberapa data memiliki format DD/MM/YYYY dan terdapat data yang bernilai null/kosong.
  - Kolom Item unit price, No. Items, dan Total Sale: semua data sama berbentuk desimal.

#### A. Pengambilan Data (Extract)

The screenshot shows the 'Spoon - Transformation 1 (changed)' window with the 'CSV file input' step selected. The configuration is as follows:

- Step name: Input Data CSV
- Filename: D:\Kuliah\CarDistributionSales.csv
- Delimiter: ,
- 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? ☐
- Format: mixed
- File encoding:

At the bottom, there is a table with the following columns: #, Name, Type, Format, Length, Precision, Currency, Decimal, Group, Trim type. The first row shows the value '1' in the '#' column.

Buttons at the bottom: Help, OK, Get Fields, Preview, Cancel.

CSV file input

Step name: Input Data CSV

Filename: D:\Kuliah\CarDistributionSales.csv [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? ☐

Format: mixed

File encoding:

Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim
Customer	String		10		\$	.	,	none
Product	String		9		\$	.	,	none
Region	String		3		\$	.	,	none
Date	Date	MM/dd/yyyy			\$	.	,	none
Item unit price	Number	##	6	2	\$	.	,	none
No.Items	Integer	#	15	0	\$	.	,	none
Total Sale	Number	##	7	2	\$	.	,	none

[Help](#) [OK](#) [Get Fields](#) [Preview](#) [Cancel](#)

Spoon - Transformation 1

File Edit View Action Tools Help

Connect

View

Search

Transformation 1

Input Data CSV

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics

2025/02/26 21:24:06 - Spoon - Transformation opened.

2025/02/26 21:24:06 - Spoon - Launching transformation [Transformation 1]...

2025/02/26 21:24:06 - Spoon - Started the transformation execution.

2025/02/26 21:24:07 - Transformation 1 - Dispatching started for transformation [Transform

2025/02/26 21:24:07 - Input Data CSV.0 - Header row skipped in file 'D:\Kuliah\CarDistributi

2025/02/26 21:24:07 - Input Data CSV.0 - Finished processing (I=20, O=0, R=0, W=19, U=0,

2025/02/26 21:24:07 - Spoon - The transformation has finished!!

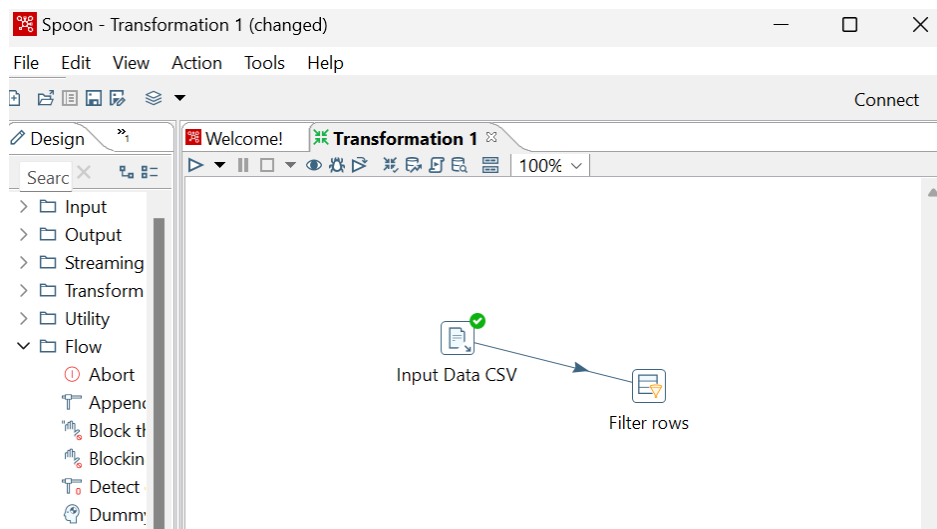
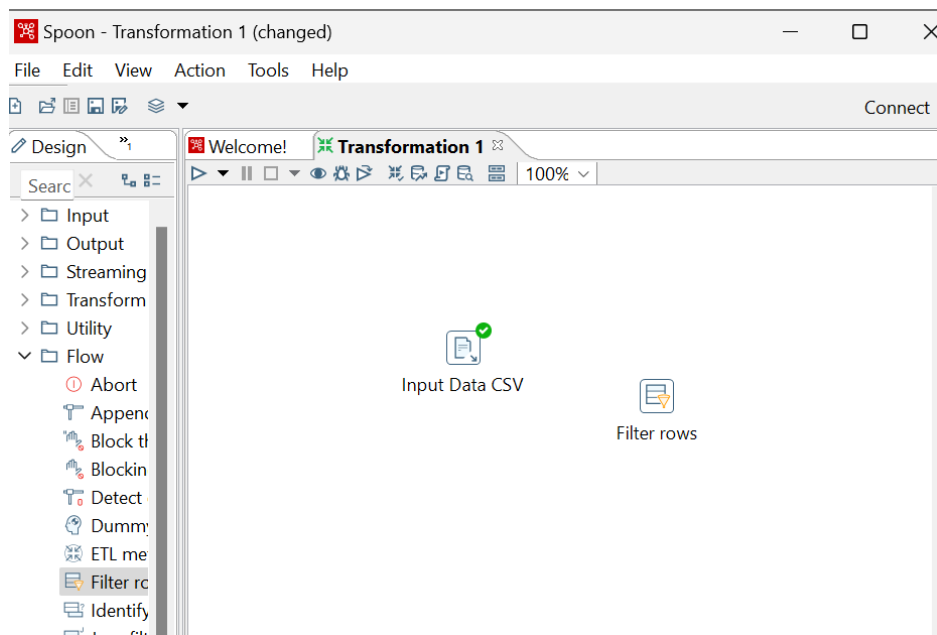
## Tugas 2

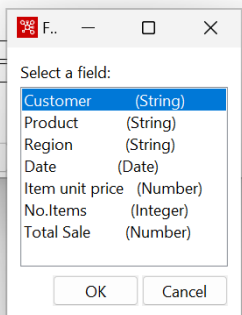
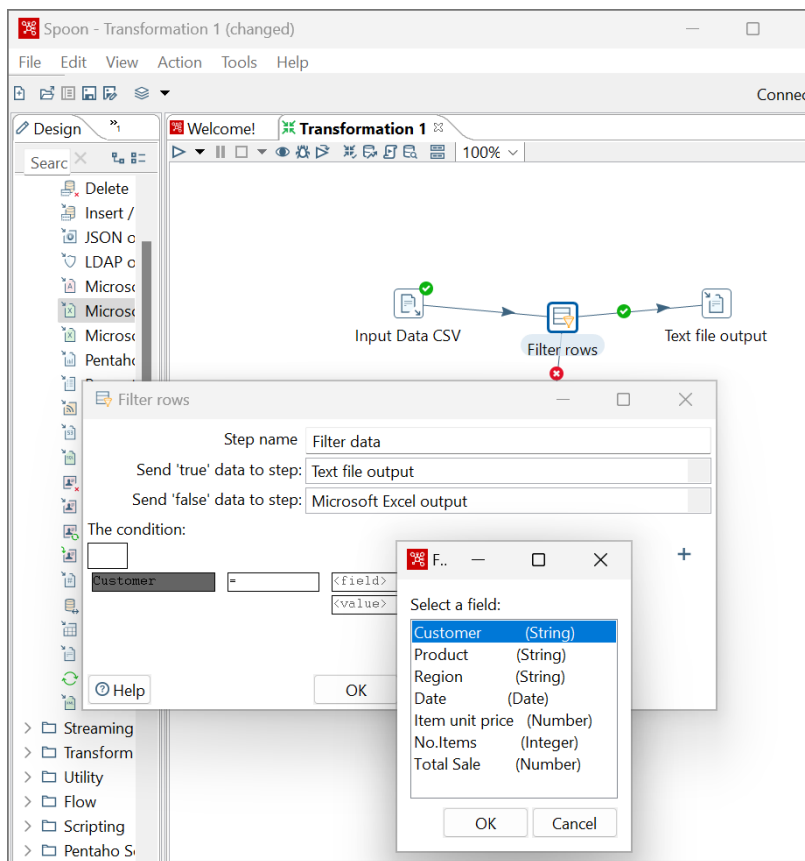
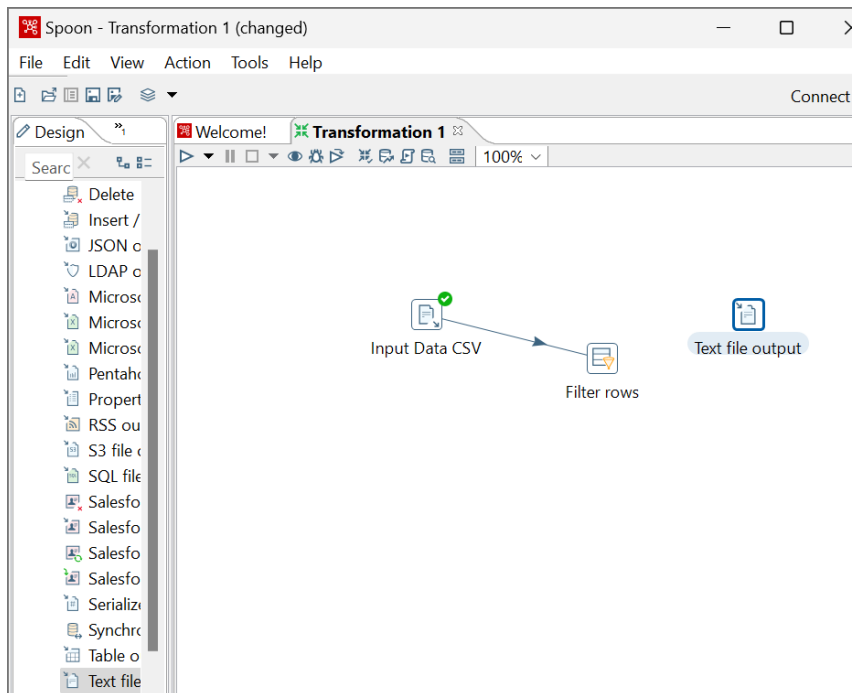
- Apakah data hasil eksekusi sesuai dengan data aslinya?
  - Ya, data hasil eksekusi sesuai.
- PDI Spoon melakukan proses extract, perhatikan pada **Tab Logging** di **Execution Results Area**, langkah-langkah apa sajakah yang dilakukan PDI Spoon untuk melakukan extract data?

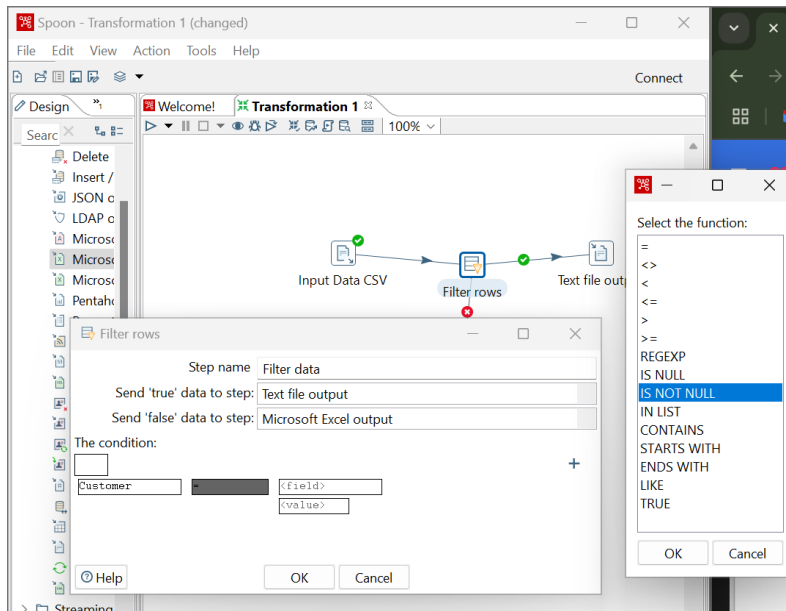
- Langkah-langkah yang dilakukan PDI Spoon untuk melakukan ekstrak data adalah
  1. Membuka transformasi
  2. Menjalankan transformasi
  3. Melakukan proses ekstrak dari file CSV
  4. Memproses data dari sumber
  5. Transformasi selesai
- 3. Perhatikan gambar dibawah ini! Apa yang dimaksud dengan I,O,R,W,U,E ?  
 2024/02/17 06:51:48 - Input Data CSV.0 - Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)
  - I (Input) = 20 → Data yang dibaca dari file sebanyak 20 baris.
  - O (Output) = 0 → Tidak ada data yang dikirim keluar (jika ada kesalahan atau filter).
  - R (Rejected) = 0 → Tidak ada data yang ditolak.
  - W (Written) = 19 → 19 baris data berhasil diekstrak dan ditulis ke tahap selanjutnya.
  - U (Updated) = 0 → Tidak ada data yang diperbarui.
  - E (Error) = 0 → Tidak ada kesalahan yang terjadi.
- 4. Berapa waktu yang dibutuhkan untuk melakukan eksekusi Extract data ini? Bandingkan dengan 3 teman yang lain, identifikasi pula spesifikasi perangkat yang digunakan (Processor, RAM, Storage, VGA, CPU). Bandingkan dalam bentuk tabel

Nama	Processor	RAM	Storage	VGA	CPU	Waktu
Niriza	AMD Ryzen™ 3 7320U	8.00 GB	1 TB	AMD Radeon Graphics	AMD Ryzen™ 3 7320U	0.1s
Maharani	AMD Athlon 300U with Radeon Vega Mobile Gfx	4.00 GB	202 GB	AMD Radeon™ Vega 3 Graphics	AMD Athlon 300U with Radeon Vega Mobile Gfx	0.4s
Aqueena	Intel(R) Core™ i7-10750H CPU @ 2.60GHz 2.59 GHz	8.00 GB	237 GB	NVIDIA GeForce RTX 3050 Ti	Intel(R) Core™ i7-10750H CPU @ 2.60GHz 2.59 GHz	0.0s
Nimas	AMD Ryzen 3 5300U	16.00 GB	475 GB	AMD Radeon Graphics	AMD Ryzen 3 5300U	0.1S

## B. Filter Data (Transform) dan Pengemasan data (Load)







Customer IS NOT NULL

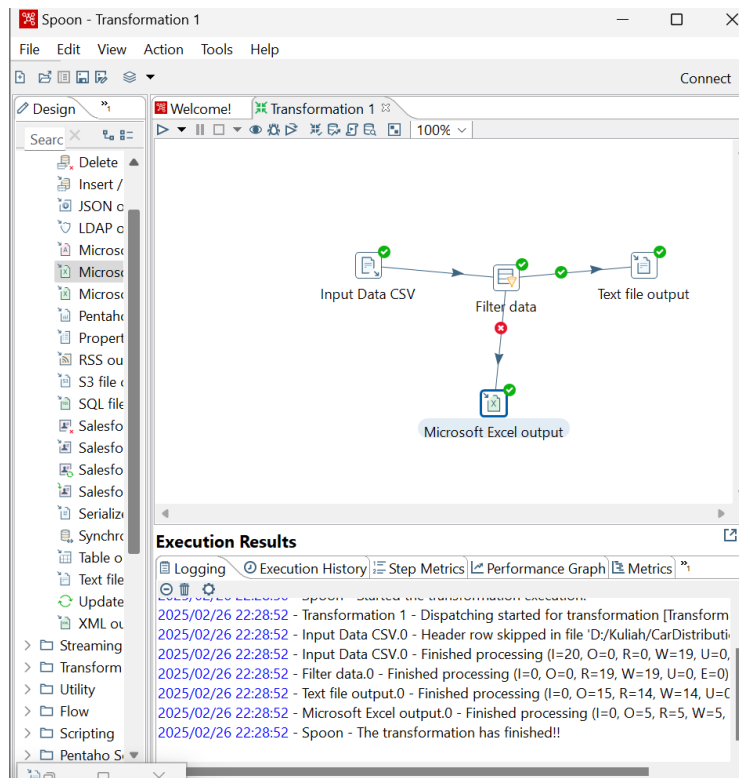
AND

```
(
    Product IS NOT NULL
    AND
    (
        Region IS NOT NULL
        AND
        (
            Date IS NOT NULL
            AND
            (
                Item unit price IS NOT NULL
                AND
                (
                    No.Items IS NOT NULL
                    AND
                    Total Sale IS NOT NULL
                )
            )
        )
    )
)
```





[illegible]



Data failedSalesCarDistribution:

failedSalesCarDistribution - Compatibility							
File Home Insert Page Layout Formulas Data Review View Help							
<div> <div>Paste</div> <div>Clipboard</div> </div> <div> <div>Arial</div> <div>10</div> <div>A<sup>+</sup> A<sup>-</sup></div> <div>B I U</div> <div>Font</div> </div> <div> <div>Alignment</div> </div>							
A1	Customer						
	A	B	C	D	E	F	G
1	Customer	Product	Region	Date	Item unit p	No.Items	Total Sale
2	Customer		MLG	#####	0168.95	2.00	00337.90
3	Customer	Avanza AT			0168.95	15.00	02534.25
4	Customer				0340.95	7.00	02386.65
5	Customer		KDR	#####	0799.95	14.00	11199.30
6	Customer	Avanza AT		#####	0168.95	4.00	00675.80
7							

Data outputSalesCarDistribution:

```

outputSalesCarDistribution.txt
File Edit View

Customer;Product;Region;Date;Item unit price;No.Items;Total Sale
Customer 1;Varis AT ;MLG;3/13/2019;340.9;6;2045.7
Customer 1;Etios AT ;SBY;4/16/2019;800;1;800
Customer 3;Avanza AT;MLG;6/18/2019;168.9;7;1182.7
Customer 1;Avanza AT;SBY;8/21/2019;168.9;9;1520.5
Customer 2;Etios AT ;KDR;9/7/2019;800;7;5599.6
Customer 3;Etios AT ;MLG;9/7/2019;800;8;6399.6
Customer 1;Avanza AT;SBY;9/14/2019;168.9;15;2534.2
Customer 1;Etios AT ;SBY;9/25/2019;800;2;1599.9
Customer 1;Avanza AT;SBY;1/21/2020;168.9;14;2365.3
Customer 1;Varis AT ;MLG;2/1/2020;340.9;2;681.9
Customer 1;Etios AT ;SBY;5/30/2020;800;9;7199.6
Customer 3;Avanza AT;MLG;6/7/2020;168.9;10;1689.5
Customer 3;Varis AT ;MLG;6/9/2020;340.9;4;1363.8
Customer 2;Avanza AT;KDR;8/8/2020;168.9;12;2027.4

```

### Tugas 3

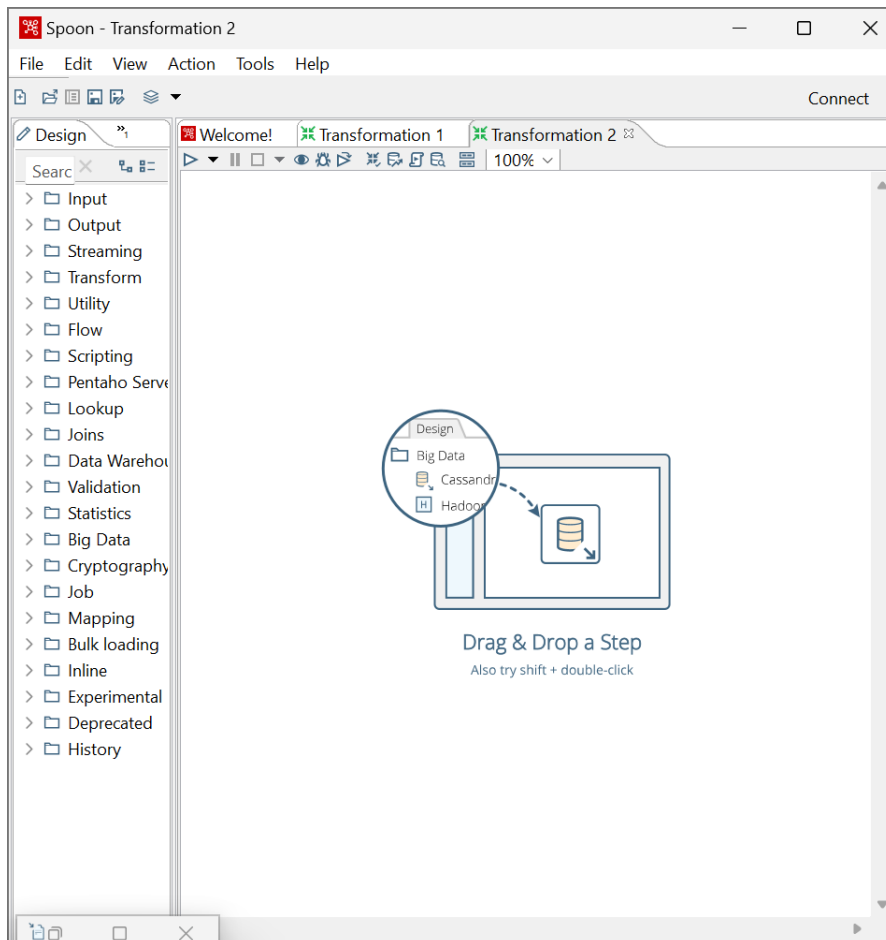
1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?
  - Isi data output pada file csv merupakan data penjualan yang lengkap dan berhasil, sedangkan data output pada file excel berisi data penjualan yang tidak lengkap/null dan gagal.
2. Jelaskan apa yang terjadi pada proses filter rows
  - Filter rows berfungsi sebagai penyaring data berdasarkan kriteria tertentu untuk memastikan hanya data yang relevan yang akan diproses dalam transformasi. Filter rows bekerja dengan proses seperti berikut:
    1. Menerima data masukan
    2. Mengevaluasi kondisi filter
    3. Memisahkan data
    4. Menghasilkan output
3. Berapa waktu yang dibutuhkan untuk melakukan eksekusi Extract Transfer Load data ini? Bandingkan dengan 3 teman yang lain, identifikasi pula spesifikasi perangkat yang digunakan (Processor, RAM, Storage, VGA, CPU). bandingkan dalam bentuk tabel

Nama	Processor	RAM	Storage	VGA	CPU	Waktu
Niriza	AMD Ryzen™ 3 7320U	8.00 GB	1 TB	AMD Radeon Graphics	AMD Ryzen™ 3 7320U	0.3s
Maharani	AMD Athlon	4.00 GB	202 GB	AMD Radeon™	AMD Athlon	0.4s

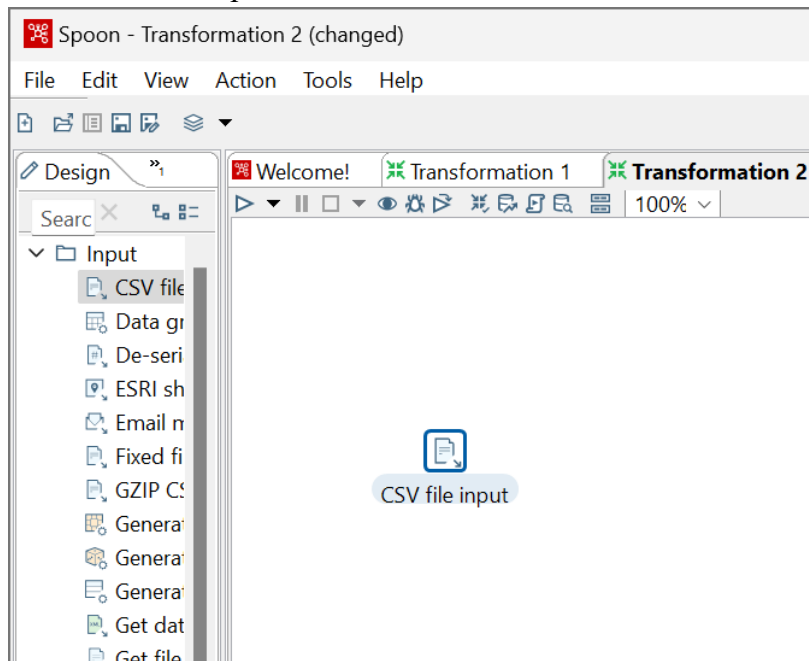
	300U with Radeon Vega Mobile Gfx			Vega 3 Graphics	300U with Radeon Vega Mobile Gfx	
Aqueena	Intel(R) Core™ i7- 10750H CPU @ 2.60GHz 2.59 GHz	8.00 GB	237 GB	NVIDIA GeForce RTX 3050 Ti	Intel(R) Core™ i7- 10750H CPU @ 2.60GHz 2.59 GHz	0.3s
Nimas	AMD Ryzen 3 5300U	16.00 GB	475 GB	AMD Radeon Graphics	AMD Ryzen 3 5300U	0.4s

## Studi Kasus

### 1. Membuat file transformation baru



## 2. Menambahkan input file csv



### 3. Konfigurasi CSV file input

CSV file input

Step name: Input Data CSV Konoha

Filename: D:\Kuliah\dataDummyKependudukan.csv 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? ☐

Format: mixed

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1									

Help OK Get Fields Preview Cancel

CSV file input

Step name: Input Data CSV Konoha

Filename: D:\Kuliah\dataDummyKependudukan.csv 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? ☐

Format: mixed

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group
1	Nama	String		16		\$	.	,
2	NIK	Integer	#	15	0	\$	.	,
3	Tempat Lahir	String		11		\$	.	,
4	Tanggal Lahir	Date	dd-MM-yyyy			\$	.	,
5	Jenis Kelamin	String		9		\$	.	,
6	Alamat	String		28		\$	.	,
7	Kota	String		11		\$	.	,
8	RT/RW	String		7		\$	.	,
9	Kelurahan	String		18		\$	.	,
1.	Kecamatan	String		17		\$	.	,
1.	Agama	String		7		\$	.	,
1.	Status Perkawinan	String		13		\$	.	,
1.	Pekerjaan	String		14		\$	.	,

Help OK Get Fields Preview Cancel

#### 4. Menambahkan Switch Case

Spoon - Transformation 2 (changed)

File Edit View Action Tools Help

Design

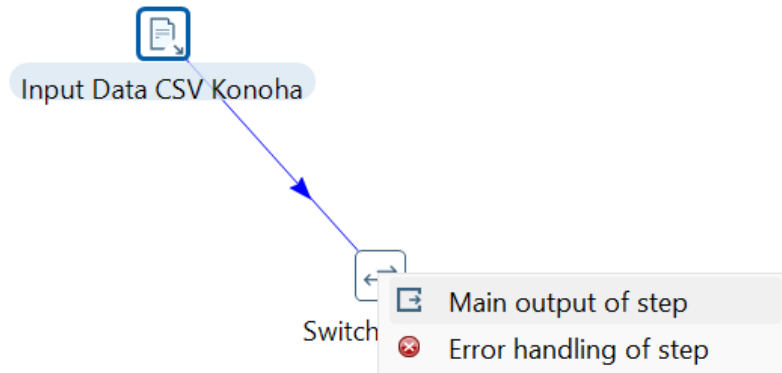
Search

- > Input
- > Output
- > Streaming
- > Transform
- > Utility
- ✓ Flow
  - Abort
  - Append
  - Block t
  - Blockin
  - Detect
  - Dumm
  - ETL me
  - Filter rc
  - Identify
  - Java filt
  - Job exe
  - Prioritiz

Input Data CSV Konoha

Switch / case

5. Menghubungkan output Input Data CSV Konoha menuju Switch Case



## 6. Konfigurasi Switch Case

[illegible]

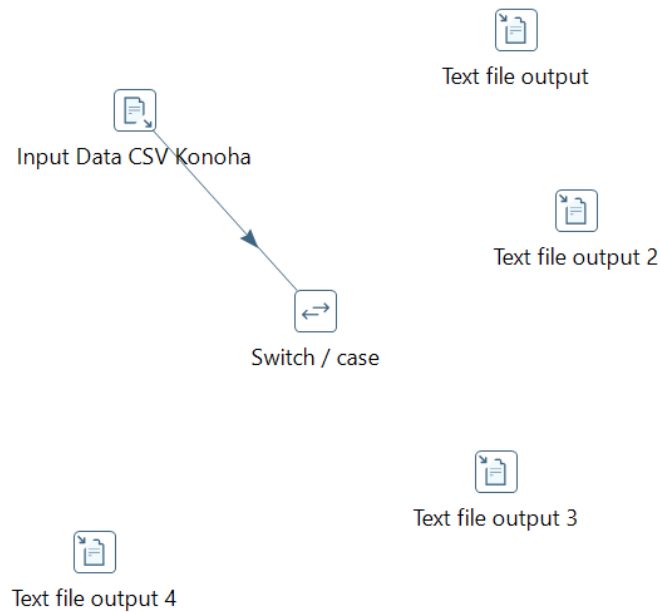
Field to switch: Pekerjaan

Case values:

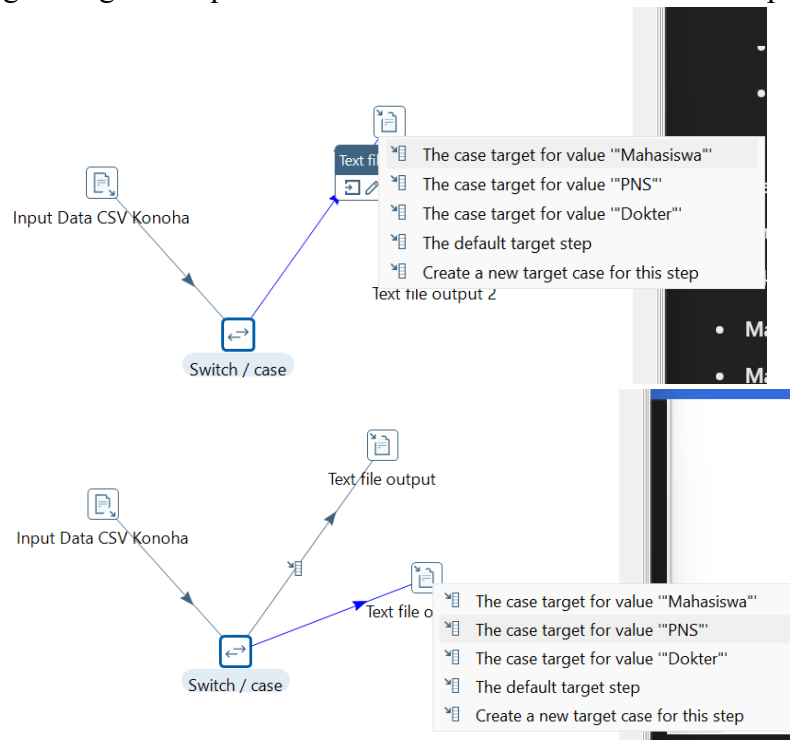
- a. Mahasiswa  $\rightarrow$  Output ke Text file output

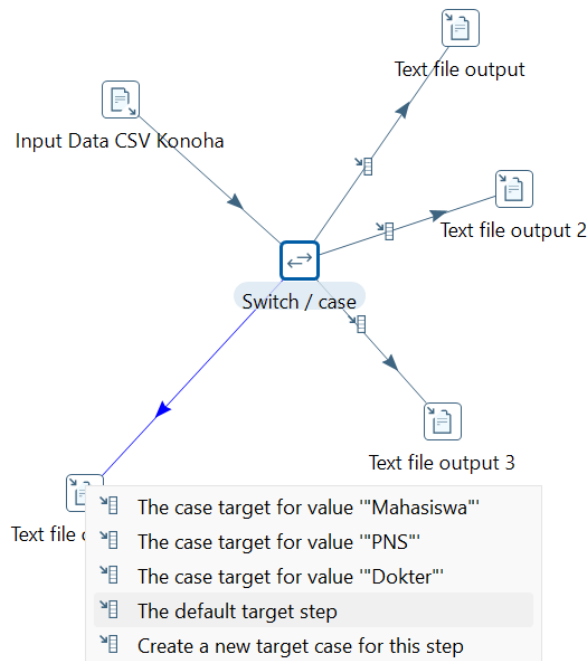
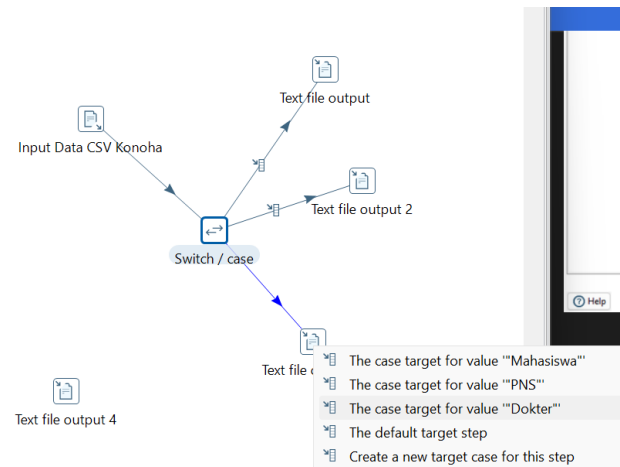


- b. PNS → Output ke Text file output 2
  - c. Dokter → Output ke Text file output 3
  - d. Default case (jika tidak cocok) → Output ke Text file output 4
7. Menambahkan 4 Text file output



8. Menghubungkan output dari Switch Case ke semua Text file output





## 9. Konfigurasi semua Text file output

Text file output

Step name: Text file output

File Content Fields

Filename: D:\Kuliah\MasyarakatMahasiswa [Browse...](#)

Pass output to servlet: ☐

Create Parent folder: ☒

Do not create file at start: ☐

Accept file name from field: ☐

File name field:

Extension: csv

Include stepnr in filename: ☐

Include partition nr in filename: ☒

Include date in filename: ☐

Include time in filename: ☐

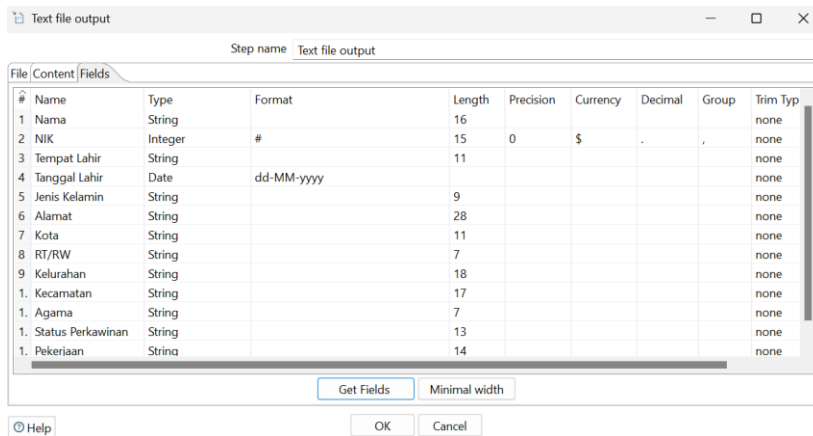
Specify Date time format: ☐

Date time format:

Show filename(s): [Show filename\(s\)...](#)

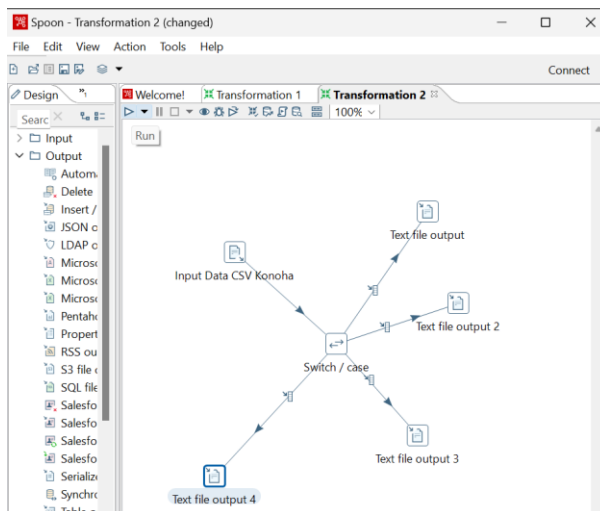
Add filenames to result: ☒

[Help](#) [OK](#) [Cancel](#)

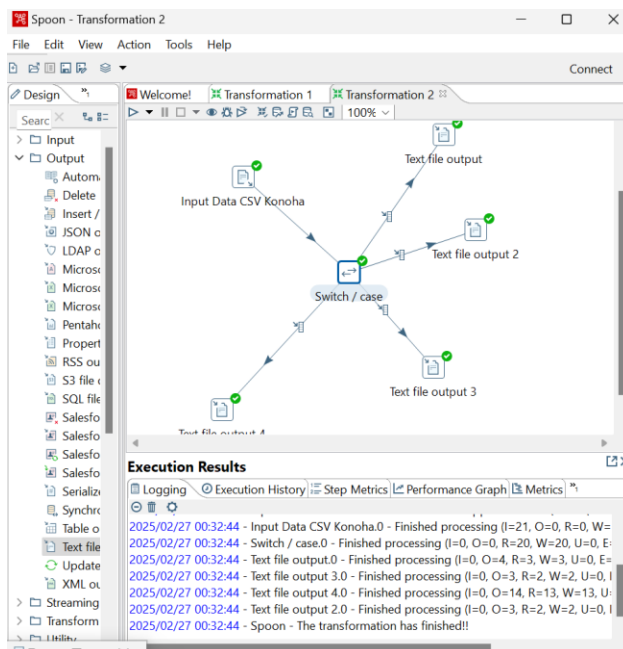


\*Langkah yang sama dilakukan ke semua Text file output

## 10. Melakukan run



## 11. Transformasi berhasil



Output:

- Data Masyarakat Mahasiswa

Nama;NIK	Tempat Lahir;Tanggal Lahir;Jenis Kelamin;Alamat;Kota;RT/RW;Kelurahan;Kecamatan;Agama;Status Perkawinan;Pekerjaan
Siti Rahayu	;3503010101010002;Bandung ;15-05-1990;Perempuan;Jl. Cendrawasih No. 45 ;Bandung ;003/004;Cikutra ;Bandung Wetan ;Kristen;Belum Menikah;Mahasiswa
Adi Pratama	;3503010101010005;Semarang ;05-03-1995;Laki-laki;Jl. Pemuda No. 21 ;Semarang ;009/010;Pandanaran ;Semarang Tengah ;Buddha ;Belum Menikah;Mahasiswa
Maya Wulandari	;3503010101010020;Jayapura ;23-10-1993;Perempuan;Jl. Proklamasi No. 80 ;Jayapura ;039/040;Kota Jayapura ;Jayapura ;Kristen;Belum Menikah;Mahasiswa

- Data MasyarakatPNS

[illegible]

- Data MasyarakatDokter

Nama;NIK;	Tempat Lahir;	Tanggal Lahir;	Jenis Kelamin;	Alamat;Kota;RT/RW;	Kelurahan;	Kecamatan;	Agama;	Status Perkawinan;	Pekerjaan
Desi Purnama ;3503010101010008;	Palembang ;08-04-1993;	Perempuan;	Jl. Demang Lebar Daun No. 20;	Palembang ;015/016;	Alang-Alang Lebar ;Palembang Selatan;	Kristen;	Belum Menikah;	Dokter	
Bambang Kusumo ;3503010101010017;	Banjarmasin;	25-09-1980;	Laki-laki;	Jl. A. Yani No. 60 ;Banjarmasin;	033/034;	Banjarmasin Tengah;	Banjarmasin ;Hindu ;Menikah ;Dokter		

- Data Masyarakat Lainnya

```

Nama;NIK;Tempat Lahir;Tanggal Lahir;Jenis Kelamin;Alamat;Kota;RT/RW;Kelurahan;Kecamatan;Agama;Status Perkawinan;Pekerjaan
Budi Santoso ;3503010101010001;Jakarta ;01-01-1980;Laki-laki;Jl. Merdeka No. 123 ;Jakarta ;001/002;Menteng ;Jakarta
Pusat ;Islam ;Menikah ;Wiraswasta
Joko Susilo ;3503010101010003;Surabaya ;10-10-1975;Laki-laki;Jl. Diponegoro No. 67 ;Surabaya ;005/006;Gubeng ;Surabaya
Pusat ;Katolik;Menikah ;Pegawai Swasta
Maya Dewi ;3503010101010004;Yogyakarta ;20-08-1988;Perempuan;Jl. Malioboro No. 8 ;Yogyakarta ;007/008;Sosromenduran ;Yogyakarta
Utara ;Hindu ;Belum Menikah;Guru
Rini Setiawan ;3503010101010006;Malang ;12-12-1982;Perempuan;Jl. Simpang Borobudur No. 15;Malang ;011/012;Dinoyo ;Malang
Selatan ;Kristen;Menikah ;Wiraswasta
Fajar Prabowo ;3503010101010009;Medan ;30-09-1986;Laki-laki;Jl. Sudirman No. 50 ;Medan ;017/018;Sei Sikambing ;Medan
Timur ;Hindu ;Menikah ;Wiraswasta
Rina Sari ;3503010101010010;Denpasar ;17-07-1992;Perempuan;Jl. Gatot Subroto No. 30 ;Denpasar ;019/020;Dauh Puri Kauh ;Denpasar
Utara ;Buddha ;Belum Menikah;Freelancer
Ahmad Subhan ;3503010101010011;Surakarta ;03-07-1991;Laki-laki;Jl. Pahlawan No. 55 ;Surakarta ;021/022;Jebres ;Surakarta
Timur ;Islam ;Belum Menikah;Wiraswasta
Dian Pertiwi ;3503010101010012;Maluku ;10-12-1985;Perempuan;Jl. Pattimura No. 10 ;Ambon ;023/024;Baguala ;Ambon ;Kristen;Menikah ;Guru
Rudi Hartono ;3503010101010013;Pontianak ;28-05-1973;Laki-laki;Jl. Ahmad Yani No. 30 ;Pontianak ;025/026;Pontianak Barat ;Pontianak
Kota ;Hindu ;Menikah ;Wiraswasta
Lina Fitriani ;3503010101010014;Manado ;15-08-1994;Perempuan;Jl. Wolter Monginsidi No. 17;Manado ;027/028;Mapanget ;Manado
Selatan ;Kristen;Belum Menikah;Akuntan
Rina Agustina ;3503010101010016;Bengkulu ;19-11-1990;Perempuan;Jl. Imam Bonjol No. 45 ;Bengkulu ;031/032;Ratu Samban ;Bengkulu
Tengah ;Kristen;Menikah ;Wiraswasta
Nia Dewi Lestari;3503010101010018;Manokwari ;14-06-1987;Perempuan;Jl. Diponegoro No. 40 ;Manokwari ;035/036;Manokwari
Barat ;Manokwari ;Kristen;Belum Menikah;Penulis
Doni Setiawan ;3503010101010019;Palu ;07-03-1977;Laki-laki;Jl. Soekarno-Hatta No. 70 ;Palu ;037/038;Palu
Barat ;Palu ;Islam ;Menikah ;Wiraswasta

```