

## TUGAS 1

	A	B	C	D	E	F	G
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
21							

Analisa lah data tersebut!

- Berapa jumlah kolom pada data tersebut?  
= terdapat 7 kolom yakni customer, product, region, date, item unit price, no.items dan total sale
- Apa arti atau isi data dari setiap kolom yang ada?  
= menunjukkan data penjualan product oleh beberapa customer di daerah Malang, Surabaya dan Kediri
- Adakah data yang memiliki nilai null / data yang tidak lengkap?  
= ada yakni ada pada kolom product ada 3, kolom region ada 3, date ada 2 yang dimana data yang memiliki nilai null tersebut berbeda-beda baris dan penempatannya
- Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?  
= tidak ada

## TUGAS 2

- Apakah data hasil eksekusi sesuai dengan data aslinya?  
= ya, data sudah sesuai semua

CSV file input

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

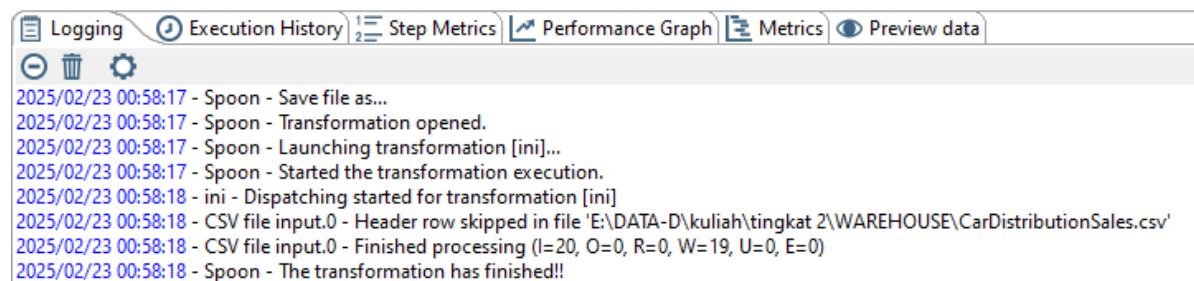
#	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
1	Customer 1	Yaris AT	MLG	Wed Mar 13 00:00:00 WIB 2019	340.95	6	2045.7
2	Customer 1	Etios AT	SBY	Tue Apr 16 00:00:00 WIB 2019	799.95	1	799.95
3	Customer 3	Avanza AT	MLG	Tue Jun 18 00:00:00 WIB 2019	168.95	7	1182.65
4	Customer 3	<null>	MLG	Wed Jul 24 00:00:00 WIB 2019	168.95	2	337.9
5	Customer 1	Avanza AT	SBY	Wed Aug 21 00:00:00 WIB 2019	168.95	9	1520.55
6	Customer 2	Etios AT	KDR	Sat Sep 07 00:00:00 WIB 2019	799.95	7	5599.65
7	Customer 3	Etios AT	MLG	Sat Sep 07 00:00:00 WIB 2019	799.95	8	6399.6
8	Customer 2	Avanza AT	<null>	<null>	168.95	15	2534.25
9	Customer 1	Avanza AT	SBY	Sat Sep 14 00:00:00 WIB 2019	168.95	15	2534.25
10	Customer 1	Etios AT	SBY	Wed Sep 25 00:00:00 WIB 2019	799.95	2	1599.9
11	Customer 1	Avanza AT	SBY	Tue Jan 21 00:00:00 WIB 2020	168.95	14	2365.3
12	Customer 1	Yaris AT	MLG	Sat Feb 01 00:00:00 WIB 2020	340.95	2	681.9

2. 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?

= berikut prosesnya

- Dilakukan save file pada folder
- Pembukaan transformasi
- Menjalankan transformasi
- Ekseskusi transformasi
- Pengiriman proses
- Membaca file csv
- Transformasi selesai

#### Execution Results



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)

= berikut penjelasannya

- I (Input Count)** → **20** = Jumlah total baris yang dibaca dari sumber data (dalam hal ini, file CSV)
- O (Output Count)** → **0** = Jumlah total baris yang dikirim ke langkah berikutnya dalam transformasi (jika ada). Dalam log ini, nilainya 0, mungkin karena tidak ada langkah output setelah membaca file.
- R (Read Count)** → **0** = Jumlah baris yang dibaca dari sumber data lain yang bukan bagian dari langkah ini. Nilai 0 berarti langkah ini hanya membaca data dari file CSV tanpa membaca tambahan dari sumber lain.
- W (Write Count)** → **19** = Jumlah baris yang berhasil ditulis setelah transformasi. Di sini, 19 baris berhasil diproses dan disimpan, kemungkinan setelah mengabaikan baris header.
- U (Update Count)** → **0** = Jumlah baris yang diperbarui dalam suatu proses. Nilai 0 menunjukkan tidak ada pembaruan data yang dilakukan.
- E (Error Count)** → **0** = Jumlah baris yang mengalami kesalahan dalam proses. Nilai 0 berarti tidak ada kesalahan yang terjadi selama transformasi.

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.

Name	Processor	RAM	Storage	VGA	CPU	Time
Kartika	Intel Core i3-6006U 2.0 GHz	8 GB	535 GB	Intel HD dan NVIDIA GeForce 920MX 2GB	Intel Core i3-6006U	1 s
Dini	AMD Ryzen 7 5800HS with Radeon Graphics 3.20 GHz	16 GB	475 GB	AMD Radeon Graphics	AMD Ryzen 7 5800HS with Radeon Graphics	1 s
Nimas	AMD Ryzen 3 5300U	16 GB	475 GB	AMD Radeon Graphics	AMD Ryzen 3 5300U with Radeon Graphics	1 s
Sufyan	I7 gen 9	16 GB	256 GB	GTX 1650	I7 gen 9	4 s

### TUGAS 3

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?  
 = perbedaan di kedua file tersebut adalah jika excel lebih terstruktur dan sudah pisah sesuai kolom-kolomnya sedangkan file csv masih berantakan dan perlu di pisah lagi sesuai kolomnya
- a. File excel

	A	B	C	D	E	F	G
1	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
2	Customer 3		MLG	7/24/2019 0:00	0168.95	2.00	00337.90
3	Customer 2	Avanza AT			0168.95	15.00	02534.25
4	Customer 1				0340.95	7.00	02386.65
5	Customer 2		KDR	8/7/2020 0:00	0799.95	14.00	11199.30
6	Customer 1	Avanza AT		8/12/2020 0:00	0168.95	4.00	00675.80
7							

b. File csv

	A	B	C	D	E	F	G
1	Customer;Product;Region;Date;Item unit price;No.Items;Total Sale						
2	Customer 1;Yaris AT ;MLG;3/13/2019;340.9;6;2045.7						
3	Customer 1;Etios AT ;SBY;4/16/2019;800;1;800						
4	Customer 3;Avanza AT;MLG;6/18/2019;168.9;7;1182.7						
5	Customer 1;Avanza AT;SBY;8/21/2019;168.9;9;1520.5						
6	Customer 2;Etios AT ;KDR;9/7/2019;800;7;5599.6						
7	Customer 3;Etios AT ;MLG;9/7/2019;800;8;6399.6						
8	Customer 1;Avanza AT;SBY;9/14/2019;168.9;15;2534.2						
9	Customer 1;Etios AT ;SBY;9/25/2019;800;2;1599.9						
10	Customer 1;Avanza AT;SBY;1/21/2020;168.9;14;2365.3						

2. Jelaskan apa yang terjadi pada proses Filter rows!

= Langkah "**Filter rows**" menyaring data dengan memastikan semua kolom **tidak kosong (NULL)**.

- **Data lengkap** (tidak ada NULL) = dikirim ke **Text file output**

- **Data tidak lengkap** (ada NULL) = dikirim ke **Microsoft Excel output**

Tujuannya untuk **memisahkan data valid dan tidak valid** sebelum diproses lebih lanjut.

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

Name	Processor	RAM	Storage	VGA	CPU	Time
Kartika	Intel Core i3-6006U 2.0 GHz	8 GB	535 GB	Intel HD dan NVIDIA GeForce 920MX 2GB	Intel Core i3-6006U	0 s
Dini	AMD Ryzzen 7 5800HS with Radeon Graphics 3.20 GHz	16 GB	475 GB	AMD Radeon Graphics	AMD Ryzzen 7 5800HS with Radeon Graphics	1 s
Nimas	AMD Ryzen 3 5300U	16 GB	475 GB	AMD Radeon Graphics	AMD Ryzen 3 5300U with Radeon Graphics	4 s

## STUDY CASE

Pak Pascanowo seorang ketua RW di Ibu Kota Nusabangsa (IKN). Terdapat 20 orang dari negara Konoha yang akan pindah dan berdomisili di IKN. Tata letak pemukiman pada IKN telah dikondisikan agar berkelompok disesuaikan dengan lokasi pekerjaannya agar akses terjangkau.

Sebagai contoh, mahasiswa akan dikumpulkan dan berdomisili di wilayah edukasi yang akses dekat dengan kampus-kampus perguruan tinggi. Sedangkan dokter akan berdomisili dekat dengan rumah sakit atau layanan kesehatan dan PNS akan berdomisili dekat dengan perkantoran dan layanan publik. Selain itu, orang yang bukan dari ketiga kategori tersebut bebas memilih domisili atau tempat tinggalnya di IKN.

Pak Pascanowo akan mendapatkan data dari pemerintah pusat tentang Masyarakat Konoha yang akan pindah ke IKN. dari data tersebut Pak Pascanowo akan membagi menjadi 4 kelompok,

**Data masyarakat berprofesi mahasiswa** akan diberikan ke Pak Emir Makarena selaku ketua RT wilayah pendidikan.

**Data masyarakat berprofesi PNS** akan diberikan ke Pak Yasana Laili selaku ketua RT wilayah pemerintahan

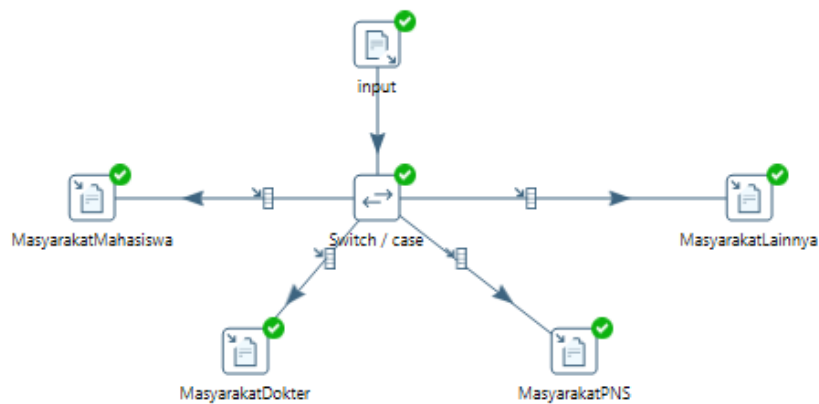
**Data masyarakat berprofesi Dokter** akan diberikan ke Pak Budi Pekerti selaku ketua RT wilayah kesehatan

data-data tersebut akan digunakan untuk analisa di masing-masing wilayah dan juga Dashboard untuk membantu mengambil keputusan.

Pak Pascanowo mendapatkan data penduduk yang akan pindah ke wilayahnya melalui link: <https://github.com/dik4rizky/datasources/blob/82821b8ec341176f979956c84b79720d4ead012/dataDummyKependudukan.csv>

Bantulah Pak Pascanowo untuk memisahkan data tersebut menjadi 4 file: MasyarakatMahasiswa, MasyarakatPNS, MasyarakatDokter, MasyarakatLainnya.

Dikarenakan proses tersebut akan repetisi setiap ada masyarakat Konoha yang berpindah ke IKN maka buatlah sistem tersebut menggunakan PDI Spoon.



### a. Mahasiswa

Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
Perempuan	Jl. Cendrawasih No. 45	Bandung	003/004	Cikutra	Bandung Wetan	Kristen	Belum Menikah	Mahasiswa
Laki-laki	Jl. Pemuda No. 21	Semarang	009/010	Pandanaran	Semarang Tengah	Buddha	Belum Menikah	Mahasiswa
Perempuan	Jl. Proklamasi No. 80	Jayapura	039/040	Kota Jayapura	Jayapura	Kristen	Belum Menikah	Mahasiswa

### b. Dokter

Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
Perempuan	Jl. Demang Lebar Daun No. 20	Palembang	015/016	Alang-Alang Lebar	Palembang Selatan	Kristen	Belum Menikah	Dokter
Laki-laki	Jl. A. Yani No. 60	Banjarmasin	033/034	Banjarmasin Tengah	Banjarmasin	Hindu	Menikah	Dokter

### d. PNS

Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
Laki-laki	Jl. Jenderal Sudirman No. 10	Makassar	013/014	Barombong	Makassar Selatan	Islam	Menikah	PNS
Laki-laki	Jl. Sudirman No. 20	Padang	029/030	Padang Barat	Padang	Islam	Menikah	PNS

### e. Lainnya

Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
Laki-laki	Jl. Merdeka No. 123	Jakarta	001/002	Menteng	Jakarta Pusat	Islam	Menikah	Wiraswasta
Laki-laki	Jl. Diponegoro No. 67	Surabaya	005/006	Gubeng	Surabaya Pusat	Katolik	Menikah	Pegawai Swasta
Perempuan	Jl. Malioboro No. 8	Yogyakarta	007/008	Sosromenduran	Yogyakarta Utara	Hindu	Belum Menikah	Guru
Perempuan	Jl. Simpang Borobudur No. 15	Malang	011/012	Dinoyo	Malang Selatan	Kristen	Menikah	Wiraswasta
Laki-laki	Jl. Sudirman No. 50	Medan	017/018	Sei Sikambing	Medan Timur	Hindu	Menikah	Wiraswasta
Perempuan	Jl. Gatot Subroto No. 30	Denpasar	019/020	Dauh Puri Kauh	Denpasar Utara	Buddha	Belum Menikah	Freelancer
Laki-laki	Jl. Pahlawan No. 55	Surakarta	021/022	Jebres	Surakarta Timur	Islam	Belum Menikah	Wiraswasta
Perempuan	Jl. Pattimura No. 10	Ambon	023/024	Baguala	Ambon	Kristen	Menikah	Guru
Laki-laki	Jl. Ahmad Yani No. 30	Pontianak	025/026	Pontianak Barat	Pontianak Kota	Hindu	Menikah	Wiraswasta
Perempuan	Jl. Wolter Monginsidi No. 17	Manado	027/028	Mapanget	Manado Selatan	Kristen	Belum Menikah	Akuntan
Perempuan	Jl. Imam Bonjol No. 45	Bengkulu	031/032	Ratu Samban	Bengkulu Tengah	Kristen	Menikah	Wiraswasta
Perempuan	Jl. Diponegoro No. 40	Manokwari	035/036	Manokwari Barat	Manokwari	Kristen	Belum Menikah	Penulis
Laki-laki	Jl. Soekarno-Hatta No. 70	Palu	037/038	Palu Barat	Palu	Islam	Menikah	Wiraswasta