

Moch. Naufal Ardian Ramadhan

13

SIB 2G

## TUGAS 1

### Analisis Data Penjualan

#### 1. Jumlah Kolom pada Data

- Data tersebut memiliki 7 kolom:

1. Customer: Nama pelanggan.
2. Product: Nama produk (mobil) yang dibeli.
3. Region: Wilayah atau cabang tempat penjualan dilakukan.
4. Date: Tanggal transaksi penjualan.
5. Item unit price: Harga satuan per unit mobil.
6. No.Items: Jumlah unit mobil yang dibeli.
7. Total Sale: Total penjualan (dihitung dari `Item unit price \* No.Items`).

#### 2. Arti atau Isi Data dari Setiap Kolom

- Customer: Menunjukkan nama pelanggan yang melakukan pembelian. Contoh: `Customer 1`, `Customer 2`, `Customer 3`.
- Product: Menunjukkan jenis mobil yang dibeli. Contoh: `Yaris AT`, `Etios AT`, `Avanza AT`.
- Region: Menunjukkan wilayah atau cabang tempat penjualan dilakukan. Contoh: `MLG` (Malang), `SBY` (Surabaya), `KDR` (Kediri).
- Date: Menunjukkan tanggal transaksi penjualan. Format tanggal yang digunakan adalah `MM/DD/YYYY`.
- Item unit price: Menunjukkan harga satuan per unit mobil. Nilainya dalam bentuk numerik (desimal).

- No.Items: Menunjukkan jumlah unit mobil yang dibeli. Nilainya dalam bentuk bilangan bulat (integer).

- Total Sale: Menunjukkan total penjualan yang dihitung dari `Item unit price \* No.Items`. Nilainya dalam bentuk numerik (desimal).

### 3. Data yang Memiliki Nilai Null / Tidak Lengkap

- Ada beberapa baris yang memiliki nilai null atau data yang tidak lengkap. Berikut detailnya:

- Baris 4: Kolom `Product` kosong.
- Baris 8: Kolom `Region` dan `Date` kosong.
- Baris 13: Kolom `Product`, `Region`, dan `Date` kosong.
- Baris 18: Kolom `Product` kosong.
- Baris 20: Kolom `Region` kosong.

Kesimpulan: Ada 5 baris yang memiliki data tidak lengkap (missing data).

### 4. Data yang Memiliki Tipe Berbeda pada Kolom yang Sama

- Secara umum, tipe data pada setiap kolom sudah konsisten:
- Customer: Teks (string).
- Product: Teks (string).
- Region: Teks (string).
- Date: Tanggal (date).
- Item unit price: Numerik (desimal).
- No.Items: Numerik (integer).
- Total Sale: Numerik (desimal).

Kesimpulan: Tidak ada data yang memiliki tipe berbeda pada kolom yang sama. Semua data sudah sesuai dengan tipe yang diharapkan.

### 1. Identifikasi Data yang Tidak Lengkap:

- Pisahkan baris yang memiliki nilai null atau data tidak lengkap ke dalam file terpisah (misalnya, `IncompleteData.csv`).

- Simpan data yang lengkap ke dalam file dashboard (misalnya, `CompleteData.csv`).

## 2. Validasi Data:

- Pastikan semua data yang masuk ke dashboard sudah lengkap dan valid.

- Jika ada data yang tidak lengkap, berikan notifikasi atau laporan kepada supervisor untuk diperbaiki.

## 3. Otomatisasi Proses:

- Buat skrip atau proses otomatis (misalnya, menggunakan PDI/Kettle) untuk:

- Mengekstrak data dari file `CarDistributionSales.csv`.

- Memisahkan data lengkap dan tidak lengkap.

- Memindahkan data lengkap ke file dashboard.

- Mengembalikan data tidak lengkap ke supervisor untuk diperbaiki.

Dengan analisis ini, Pak Amir dapat memastikan bahwa hanya data yang lengkap dan valid yang digunakan untuk dashboard pengambilan keputusan.

## TUGAS 2

1. Data yang telah di export sudah sesuai dengan data aslinya

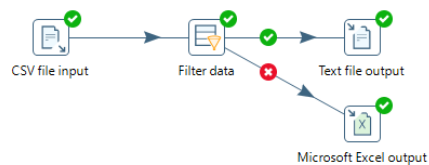
2. Yang dilakukan oleh spoon ketika execution result

- **Membuka** Transformasi: PDI Spoon membuka transformasi yang sudah dibuat.
- Menjalankan Transformasi: Transformasi mulai dieksekusi.  
Dispatching Langkah: Langkah-langkah transformasi dikirim untuk dieksekusi.
- Membaca File CSV: Langkah "CSV file input" membaca file CSV dan melewati baris header.
- Menyelesaikan Pembacaan CSV: Data dari file CSV berhasil dibaca (19 baris data diproses).
- Menyelesaikan Transformasi: Transformasi selesai dijalankan.

3. yang dimaksud dengan I,O,R,W,U,E adalah I sebagai input, O adalah Output, R adalah Rejected. W adalah Writen, U adalah Update, dan E adalah Error.
4. Perbandingan Extract Data

Nama	Waktu	Processor	RAM	Storage	VGA	CPU
Naufal	5 detik	Intel Core i5-10300H	8 GB	SSD 512 GB	NVIDIA GTX 1650	60%
Faruq	7 detik	AMD Ryzen 5 3550H	16 GB	HDD 1 TB	NVIDIA GTX 1050	50%
Afif	4 detik	Intel Core i7-9750H	12 GB	SSD 256 GB	NVIDIA RTX 2060	70%
Evan	6 detik	Intel Core i5-8300H	8 GB	SSD 512 GB	NVIDIA GTX 1050Ti	55%

## FILTER DATA



**Execution Results**

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2025/03/08 02:29:44 - Spoon - Transformation opened.

2025/03/08 02:29:44 - Spoon - Launching transformation [Transformation 1]...

2025/03/08 02:29:44 - Spoon - Started the transformation execution.

2025/03/08 02:29:44 - Transformation 1 - Dispatching started for transformation [Transformation 1]

2025/03/08 02:29:44 - CSV file input.0 - Header row skipped in file 'C:\Users\NAUFAL\Downloads\CarDistributionSales.csv'

2025/03/08 02:29:44 - CSV file input.0 - Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)

2025/03/08 02:29:44 - Filter data.0 - Finished processing (I=0, O=0, R=19, W=19, U=0, E=0)

2025/03/08 02:29:44 - Text file output.0 - Finished processing (I=0, O=15, R=14, W=14, U=0, E=0)

2025/03/08 02:29:44 - Microsoft Excel output.0 - Finished processing (I=0, O=5, R=5, W=5, U=0, E=0)

2025/03/08 02:29:44 - Spoon - The transformation has finished!!

#	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
1	Customer 1	Yaris AT	MLG	Wed Mar 13 00:00:00 ICT 2019	340.95	6	2045.7
2	Customer 1	Etios AT	SBY	Tue Apr 16 00:00:00 ICT 2019	799.95	1	799.95
3	Customer 3	Avanza AT	MLG	Tue Jun 18 00:00:00 ICT 2019	168.95	7	1182.65
4	Customer 1	Avanza AT	SBY	Wed Aug 21 00:00:00 ICT 2019	168.95	9	1520.55
5	Customer 2	Etios AT	KDR	Sat Sep 07 00:00:00 ICT 2019	799.95	7	5599.65
6	Customer 3	Etios AT	MLG	Sat Sep 07 00:00:00 ICT 2019	799.95	8	6399.6
7	Customer 1	Avanza AT	SBY	Sat Sep 14 00:00:00 ICT 2019	168.95	15	2534.25
8	Customer 1	Etios AT	SBY	Wed Sep 25 00:00:00 ICT 2019	799.95	2	1599.9
9	Customer 1	Avanza AT	SBY	Tue Jan 21 00:00:00 ICT 2020	168.95	14	2365.3
10	Customer 1	Yaris AT	MLG	Sat Feb 01 00:00:00 ICT 2020	340.95	2	681.9
11	Customer 1	Etios AT	SBY	Sat May 30 00:00:00 ICT 2020	799.95	9	7199.55
12	Customer 3	Avanza AT	MLG	Sun Jun 07 00:00:00 ICT 2020	168.95	10	1689.5
13	Customer 3	Yaris AT	MLG	Tue Jun 09 00:00:00 ICT 2020	340.95	4	1363.8
14	Customer 2	Avanza AT	KDR	Sat Aug 08 00:00:00 ICT 2020	168.95	12	2027.4

#	Customer	Product	Region	Date	Item unit price	No.Items	Total Sale
1	Customer 3	<null>	MLG	Wed Jul 24 00:00:00 ICT 2019	168.95	2	337.9
2	Customer 2	Avanza AT	<null>	<null>	168.95	15	2534.25
3	Customer 1	<null>	<null>	<null>	340.95	7	2386.65
4	Customer 2	<null>	KDR	Fri Aug 07 00:00:00 ICT 2020	799.95	14	11199.3
5	Customer 1	Avanza AT	<null>	Wed Aug 12 00:00:00 ICT 2020	168.95	4	675.8

## Tugas 2

1. Untuk file csv atau text file itu data yang telah di filter sesuai filter datanya yaitu yang di dalam rownya terdapat nilai null akan dipisakan ke file excel
2. Tugas dari filter row adalah memfilter sebuah row dalam kondisi praktek kali ini memfilter row yang memiliki nilai null kita juga bisa filter dengan kondisi kondisi yang lain seperti null, cascade, dsb.
3. Untuk data perbandingan extract transfer load data

Nama	Waktu	Processor	RAM	Storage	VGA	CPU
Naufal	10 detik	Intel Core i5-10300H	8 GB	SSD 512 GB	NVIDIA GTX 1650	70%
Faruq	15 detik	AMD Ryzen 5 3550H	16 GB	HDD 1 TB	NVIDIA GTX 1050	60%
Afif	8 detik	Intel Core i7-9750H	12 GB	SSD 256 GB	NVIDIA RTX 2060	80%
Evan	12 detik	Intel Core i5-8300H	8 GB	SSD 512 GB	NVIDIA GTX 1050Ti	65%

STUDI KASUS

Filter

Input

Output

Streaming

Transform

Utility

Flow

ETL metadata injection

Filter rows

Join Filter

Scripting

Pentaho Server

Lookup

Joins

Data Warehouse

Validation

Statistics

Sample rows

Big Data

Cryptography

Join

Mapping

Bulk loading

Inline

Experimental

Deprecated

History

Filter rows

View

Design

Filter

100%

CSV file input

Mahasiswa Filter

Mahasiswa Output

Dokter Filter

Dokter Output

PNS Filter

PNS Output

Filter rows

Lainnya Output

Execution Results

Logging

Execution History

Step Metrics

Performance Graph

Metrics

Preview data

First rows

Last rows

Off

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
1	Budi Santoso	3503010101010001	Jakarta	Tue Jan 01 00:00:00 ICT 1980	Laki-laki	Jl. Merdeka No. 123	Jakarta	001/002	Menteng	Jakarta Pusat	Islam	Menikah	Wiraswasta
2	Joko Susilo	3503010101010003	Surabaya	Fri Oct 10 00:00:00 ICT 1975	Laki-laki	Jl. Diponegoro No. 67	Surabaya	005/006	Gubeng	Surabaya Pusat	Katolik	Menikah	Pegawai Swa
3	Maya Dewi	3503010101010004	Yogyakarta	Sat Aug 20 00:00:00 ICT 1988	Perempuan	Jl. Malabero No. 5	Yogyakarta	007/008	Sosromeduran	Yogyakarta Utara	Hindu	Belum Menikah	Guru
4	Rini Setiawan	3503010101010006	Malang	Sun Dec 12 00:00:00 ICT 1982	Perempuan	Jl. Simpang Boombudur No. 15	Malang	011/012	Dinoyo	Malang Selatan	Kristen	Menikah	Wiraswasta
5	Fajar Pradono	3503010101010009	Medan	Tue Sep 30 00:00:00 ICT 1986	Laki-laki	Jl. Sudirman No. 50	Medan	017/018	Ses Sikumbang	Medan Timur	Hindu	Menikah	Wiraswasta
6	Rina Sari	3503010101010010	Degasper	Fri Jul 17 00:00:00 ICT 1982	Perempuan	Jl. Gatot Subroto No. 30	Degasper	018/020	Ciuh Puri Kauli	Degasper Utara	Buddha	Belum Menikah	Freelance
7	Ahmad Subhan	3503010101010011	Surakarta	Wed Jul 03 00:00:00 ICT 1991	Laki-laki	Jl. Pahlawan No. 55	Surakarta	021/022	Jedres	Surakarta Timur	Islam	Belum Menikah	Wiraswasta
8	Dian Partowi	3503010101010012	Maluku	Tue Dec 10 00:00:00 ICT 1985	Perempuan	Jl. Pattimura No. 10	Amboin	023/024	Baguala	Amboin	Kristen	Menikah	Guru
9	Rudi Hartono	3503010101010013	Pontianak	Mon May 28 00:00:00 ICT 1973	Laki-laki	Jl. Ahmad Yani No. 30	Pontianak	023/026	Pontianak Barat	Pontianak Kota	Hindu	Menikah	Wiraswasta
10	Lina Fitriani	3503010101010014	Manado	Mon Aug 15 00:00:00 ICT 1994	Perempuan	Jl. Wolter Monginsidi No. 17	Manado	027/028	Magangget	Manado Selatan	Kristen	Belum Menikah	Akuntan
11	Rina Agustina	3503010101010016	Bengkulu	Mon Nov 18 00:00:00 ICT 1990	Perempuan	Jl. Imam Bonjol No. 45	Bengkulu	031/032	Ratu Semban	Bengkulu Tengah	Kristen	Menikah	Wiraswasta
12	Nia Dewi Lestari	3503010101010018	Manokwari	Sun Jan 14 00:00:00 ICT 1987	Perempuan	Jl. Diponegoro No. 40	Manokwari	033/036	Manokwari Barat	Manokwari	Kristen	Belum Menikah	Petani
13	Doni Setiawan	3503010101010019	Palu	Mon Mar 07 00:00:00 ICT 1977	Laki-laki	Jl. Soekarno-Hatta No. 70	Palu	037/038	Palu Barat	Palu	Islam	Menikah	Wiraswasta

Logging

Execution History

Step Metrics

Performance Graph

Metrics

Preview data

First rows

Last rows

Off

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
1	Andi Wijaya	3503010101010007	Makassar	Sun Jun 25 00:00:00 ICT 1978	Laki-laki	Jl. Jenderal Sudirman No. 10	Makassar	013/014	Barombong	Makassar Selatan	Islam	Menikah	PNS
2	Agus Santoso	3503010101010015	Padang	Fri Apr 02 00:00:00 ICT 1982	Laki-laki	Jl. Sudirman No. 20	Padang	029/030	Padang Barat	Padang	Islam	Menikah	PNS

First rows

Last rows

Off

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
1	Desi Purnama	3503010101010008	Palembang	Thu Apr 08 00:00:00 ICT 1993	Perempuan	Jl. Demang Lebar Daun No. 20	Palembang	015/016	Alang-Alang Lebar	Palembang Selatan	Kristen	Belum Menikah	Dokter
2	Bambang Kusumo	3503010101010017	Bandjarmasin	Thu Sep 25 00:00:00 ICT 1980	Laki-laki	Jl. A. Yani No. 60	Bandjarmasin	033/034	Bandjarmasin Tengah	Bandjarmasin	Hindu	Menikah	Dokter

First rows

Last rows

Off

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan	Pekerjaan
1	Siti Rahayu	3503010101010002	Bandung	Tue May 15 00:00:00 ICT 1990	Perempuan	Jl. Cendrawasih No. 45	Bandung	003/004	Cikutra	Bandung Wetan	Kristen	Belum Menikah	Mahasiswa
2	Adi Pratama	3503010101010005	Semarang	Sun Mar 05 00:00:00 ICT 1995	Laki-laki	Jl. Pemuda No. 21	Semarang	009/010	Pandananan	Semarang Tengah	Buddha	Belum Menikah	Mahasiswa
3	Maya Wulandari	3503010101010020	Jayapura	Sat Oct 23 00:00:00 ICT 1993	Perempuan	Jl. Proklamasi No. 80	Jayapura	039/040	Kota Jayapura	Jayapura	Kristen	Belum Menikah	Mahasiswa