

## Jobsheet 1 Data Warehouse

### Tugas 1

1. **Berapa jumlah kolom pada data tersebut?**
  - Customer
  - Product
  - Region
  - Date
  - Item unit price
  - No. Items
  - Total Sale
2. **Apa arti atau isi data dari setiap kolom yang ada?**
  - **Customer:** Nama atau ID pelanggan yang melakukan pembelian.
  - **Product:** Jenis mobil yang dibeli.
  - **Region:** Lokasi tempat pembelian dilakukan.
  - **Date:** Tanggal transaksi pembelian.
  - **Item unit price:** Harga satuan dari produk yang dibeli.
  - **No. Items:** Jumlah unit produk yang dibeli.
  - **Total Sale:** Total nilai penjualan, biasanya dihitung sebagai (Item unit price \* No. Items).
3. **Adakah data yang memiliki nilai null / data yang tidak lengkap?**

Ya, terdapat beberapa baris yang tampak tidak lengkap, misalnya:

  - Baris ke-5 (Customer 3,,MLG,7/24/2019,168.95,2,337.9), di mana terdapat kolom kosong setelah "Customer 3," yang mungkin seharusnya berisi nama produk.
  - Baris ke-14 (Customer 1,,,340.95,7,2386.65), yang kehilangan informasi di kolom "Product" dan "Region".
4. **Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?**
  - Beberapa kolom seharusnya memiliki tipe data numerik (misalnya "Item unit price", "No. Items", dan "Total Sale"), tetapi ada kemungkinan beberapa nilai dituliskan dalam format string jika terjadi kesalahan input.
  - Kolom "Date" memiliki format tanggal, tetapi ada kemungkinan beberapa nilai ditulis dalam format yang tidak seragam, seperti 4/16/2019 vs 8/21/2019.
  - Kolom "Customer" terkadang memiliki angka atau karakter yang tidak seragam, misalnya ada baris dengan hanya "Customer 1," yang tampak tidak lengkap.

## Tugas 2

1. Apakah data hasil eksekusi sesuai dengan data aslinya?
  - o Hasil data eksekusi sama dengan data aslinya
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?

```
2025/02/20 14:51:39 - Spoon - Transformation opened.
2025/02/20 14:51:39 - Spoon - Launching transformation [Transformation 1]...
2025/02/20 14:51:39 - Spoon - Started the transformation execution.
2025/02/20 14:51:39 - Transformation 1 - Dispatching started for transformation [Transformation 1]
2025/02/20 14:51:39 - Input Data CSV.0 - Header row skipped in file 'C:\Users\rbhag\Downloads\CarDistributionSales.csv'
2025/02/20 14:51:39 - Input Data CSV.0 - Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)
2025/02/20 14:51:39 - Spoon - The transformation has finished!!
```

3. Perhatikan gambar di bawah ini! Apa yang dimaksud dengan I,O,R,W,U,E?
  - o Pada log proses extract, terdapat kode berikut:
  - o 2024/02/17 06:51:48 - Input Data CSV.0 - Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)

Keterangan dari masing-masing huruf:

- **I (Input)** = 20 → Jumlah baris data yang diambil dari sumber.
  - **O (Output)** = 0 → Jumlah baris data yang dikirim ke output (mungkin belum diproses ke tujuan akhir).
  - **R (Rejected)** = 0 → Jumlah baris data yang ditolak karena kesalahan atau aturan validasi.
  - **W (Written)** = 19 → Jumlah baris data yang berhasil ditulis atau diproses ke output.
  - **U (Updated)** = 0 → Jumlah baris data yang diperbarui (jika prosesnya melibatkan pembaruan data).
  - **E (Error)** = 0 → Jumlah error yang terjadi selama proses extract.
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	Waktu Eksekusi	Processor	RAM	Storage	VGA
Nimas	0,1s	AMD Ryzen 3 5300U	16,0 GB	475 GB	AMD Radeon Graphics
Daffa	0,1s	AMD A9-9425 RADEON R5	8,0 GB	1,2 TB	AMD RADEON (TM) R5 Graphics
Dhevina	0,0s	12th Gen Intel(R) Core(TM) i5 - 1235U 1.30 GHz	8,00 GB	152 GB	Inter(R) UHD Graphics

### Tugas 3

1. **Apa perbedaan isi data output dilihat dari isi file CSV dan file Excel?**  
File CSV merupakan hasil awal sebelum dilakukan transformasi data, sedangkan file excel sudah melewati proses pembersihan data
2. **Jelaskan apa yang terjadi pada proses *Filter rows*!**
  - o Proses *Filter rows* dalam ETL (Extract, Transform, Load) digunakan untuk menyaring baris data berdasarkan kondisi tertentu.
  - o Data yang tidak memenuhi kriteria akan dibuang atau tidak diteruskan ke langkah berikutnya.
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	Waktu Eksekusi	Processor	RAM	Storage	VGA
Nimas	0,1s	AMD Ryzen 3 5300U	16,0 GB	475 GB	AMD Radeon Graphics
Daffa	0,12s	AMD A9-9425 RADEON R5	8,0 GB	1,2 TB	AMD RADEON (TM) R5 Graphics
Dhevina	0,6s	12th Gen Intel(R) Core(TM) i5 - 1235U 1.30 GHz	8,00 GB	152 GB	Inter(R) UHD Graphics

STUDI KASUS

Data Input Kependudukan

CSV file input

Step name

Input File CSV

Filename

D:\Bhagas\Data Warehouse\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		IDR	.	,
2	NIK	Integer	#	15	0	IDR	.	,
3	Tempat Lahir	String		11		IDR	.	,
4	Tanggal Lahir	Date	dd-MM-yyyy			IDR	.	,
5	Jenis Kelamin	String		9		IDR	.	,
6	Alamat	String		28		IDR	.	,
7	Kota	String		11		IDR	.	,
8	RT/RW	String		7		IDR	.	,
9	Kelurahan	String		18		IDR	.	,

Help

OK

Get Fields

Preview

Cancel

First rows

Last rows

Off

#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Stat
1	Budi Santoso	3503010101010001	Jakarta	Tue Jan 01 00:00:00 WIB 1980	Laki-laki	Jl. Merdeka No. 123	Jakarta	001/002	Menteng	Jakarta Pusat	Islam	Mer
2	Siti Rahayu	3503010101010002	Bandung	Tue May 15 00:00:00 WIB 1990	Perempuan	Jl. Cendrawasih No. 45	Bandung	003/004	Cikutra	Bandung Wetan	Kristen	Belu
3	Joko Susilo	3503010101010003	Surabaya	Fri Oct 10 00:00:00 WIB 1975	Laki-laki	Jl. Diponegoro No. 67	Surabaya	005/006	Gubeng	Surabaya Pusat	Katolik	Mer
4	Maya Dewi	3503010101010004	Yogyakarta	Sat Aug 20 00:00:00 WIB 1988	Perempuan	Jl. Malioboro No. 8	Yogyakarta	007/008	Sosromenduran	Yogyakarta Utara	Hindu	Belu
5	Adi Pratama	3503010101010005	Semarang	Sun Mar 05 00:00:00 WIB 1995	Laki-laki	Jl. Pemuda No. 21	Semarang	009/010	Pandananran	Semarang Tengah	Buddha	Belu
6	Rini Setiawan	3503010101010006	Malang	Sun Dec 12 00:00:00 WIB 1982	Perempuan	Jl. Simpang Borobudur No. 15	Malang	011/012	Dinoyo	Malang Selatan	Kristen	Mer
7	Andi Wijaya	3503010101010007	Makassar	Sun Jun 25 00:00:00 WIB 1978	Laki-laki	Jl. Jenderal Sudirman No. 10	Makassar	013/014	Barombong	Makassar Selatan	Islam	Mer
8	Desi Purnama	3503010101010008	Palembang	Thu Apr 08 00:00:00 WIB 1993	Perempuan	Jl. Demang Lebar Daun No. 20	Palembang	015/016	Alang-Alang Lebar	Palembang Selatan	Kristen	Belu
9	Fajar Prabowo	3503010101010009	Medan	Tue Sep 30 00:00:00 WIB 1986	Laki-laki	Jl. Sudirman No. 50	Medan	017/018	Sel Sikambing	Medan Timur	Hindu	Mer
1.	Rina Sari	3503010101010010	Denpasar	Fri Jul 17 00:00:00 WIB 1992	Perempuan	Jl. Gatot Subroto No. 30	Denpasar	019/020	Deuh Puri Kauh	Denpasar Utara	Buddha	Belu
1.	Ahmad Subhan	3503010101010011	Surakarta	Wed Jul 03 00:00:00 WIB 1991	Laki-laki	Jl. Pahlawan No. 55	Surakarta	021/022	Jebres	Surakarta Timur	Islam	Belu
1.	Dian Pertiwi	3503010101010012	Maluku	Tue Dec 10 00:00:00 WIB 1985	Perempuan	Jl. Pattimura No. 10	Ambon	023/024	Baguala	Ambon	Kristen	Mer
1.	Rudi Hartono	3503010101010013	Pontianak	Mon May 28 00:00:00 WIB 1973	Laki-laki	Jl. Ahmad Yani No. 30	Pontianak	025/026	Pontianak Barat	Pontianak Kota	Hindu	Mer
1.	Lina Fitriani	3503010101010014	Manado	Mon Aug 15 00:00:00 WIB 1994	Perempuan	Jl. Wolter Monginsidi No. 17	Manado	027/028	Mapangget	Manado Selatan	Kristen	Belu
1.	Agus Santoso	3503010101010015	Padang	Fri Apr 02 00:00:00 WIB 1982	Laki-laki	Jl. Sudirman No. 20	Padang	029/030	Padang Barat	Padang	Islam	Mer
1.	Rina Agustina	3503010101010016	Bengkulu	Mon Nov 19 00:00:00 WIB 1990	Perempuan	Jl. Imam Bonjol No. 45	Bengkulu	031/032	Ratu Samban	Bengkulu Tengah	Kristen	Mer
1.	Bambang Kusumo	3503010101010017	Banjarmasin	Thu Sep 25 00:00:00 WIB 1980	Laki-laki	Jl. A. Yani No. 60	Banjarmasin	033/034	Banjarmasin Tengah	Banjarmasin	Hindu	Mer
1.	Nia Dewi Lestari	3503010101010018	Manokwari	Sun Jun 14 00:00:00 WIB 1987	Perempuan	Jl. Diponegoro No. 40	Manokwari	035/036	Manokwari Barat	Manokwari	Kristen	Belu
1.	Doni Setiawan	3503010101010019	Palu	Mon Mar 07 00:00:00 WIB 1977	Laki-laki	Jl. Soekarno-Hatta No. 70	Palu	037/038	Palu Barat	Palu	Islam	Mer
2.	Maya Wulandari	3503010101010020	Jayapura	Sat Oct 23 00:00:00 WIB 1993	Perempuan	Jl. Proklamasi No. 80	Jayapura	039/040	Kota Jayapura	Jayapura	Kristen	Belu

Data Masyarakat Mahasiswa

Text file output

Step nameMasyarakatMahasiswa

FileContentFields

FilenameD:\Bhagas\Data Warehouse\outputMahasiswaBrowse...

Pass output to servlet☐

Create Parent folder☒

Do not create file at start☐

Accept file name from field?☐

File name field

Extensiontxt

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)...

Add filenames to result☒

HelpOKCancel

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

Terdapat 3 Value dalam Data Masyarakat Mahasiswa

Data Masyarakat Dokter

Text file output

Step nameMasyarakatDokter

FileContentFields

FilenameD:\Bhagas\Data Warehouse\Jobsheet1\outputDokterBrowse...

Pass output to servlet☐

Create Parent folder☒

Do not create file at start☒

Accept file name from field?☐

File name field

Extensiontxt

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)...

Add filenames to result☒

HelpOKCancel

Execution Results

LoggingExecution HistoryStep MetricsPerformance GraphMetricsPreview data

First rowsLast rowsOff

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

Terdapat 2 Value dalam Data Masyarakat Dokter

Putera Bhagaswara R  
2341760136 / SIB2B / 19

## Data Masyarakat PNS

Text file output

Step name

File Content Fields

Filename  Browse...

Pass output to servlet ☐

Create Parent folder ☒

Do not create file at start ☐

Accept file name from field? ☐

File name field

Extension

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)...

Add filenames to result ☒

Help OK Cancel

[illegible]

Terdapat 2 Value dalam Data Masyarakat PNS

## Data Masyarakat Lainnya

Text file output

Step name MasarakatLainnya

File Content Fields

File name field D:\Bhagas\Data Warehouse\output\MasarakatLainnya Browse...

Pass output to servlet ☐

Create Parent folder ☒

Do not create file at start ☐

Accept file name from field? ☐

Extension txt

Include step\mr 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)...

Add filenames to result ☒

Help OK Cancel

☑ First rows ☐ Last rows ☐ Off												
#	Nama	NIK	Tempat Lahir	Tanggal Lahir	Jenis Kelamin	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawinan
1	Budi Santoso	3503010101010001	Jakarta	Tue Jan 01 00:00:00 WIB 1980	Laki-laki	Jl. Merdeka No. 123	Jakarta	001/002	Menteng	Jakarta Pusat	Islam	Menikah
2	Joko Susilo	3503010101010002	Surabaya	Fri Oct 10 00:00:00 WIB 1975	Laki-laki	Jl. Diponegoro No. 67	Surabaya	005/006	Gubeng	Surabaya Pusat	Katolik	Menikah
3	Maya Dewi	3503010101010003	Yogyakarta	Sat Aug 10 00:00:00 WIB 1988	Perempuan	Jl. Malioboro No. 8	Yogyakarta	007/008	Sosromenduran	Yogyakarta Utara	Hindu	Belum Menikah
4	Rini Setiawan	3503010101010004	Malang	Sun Dec 12 00:00:00 WIB 1982	Perempuan	Jl. Simpang Borobudur No. 15	Malang	011/012	Dinoyo	Malang Selatan	Kristen	Menikah
5	Fajar Prabowo	3503010101010005	Medan	Tue Sep 30 00:00:00 WIB 1986	Laki-laki	Jl. Sudirman No. 50	Medan	017/018	Sei Sikambing	Medan Timur	Hindu	Menikah
6	Rina Sari	3503010101010100	Denpasar	Fri Jul 17 00:00:00 WIB 1992	Perempuan	Jl. Gatot Subroto No. 30	Denpasar	019/020	Dauh Puri Kauh	Denpasar Utara	Buddha	Belum Menikah
7	Ahmad Subhan	3503010101010111	Surakarta	Wed Jul 03 00:00:00 WIB 1991	Laki-laki	Jl. Palihanwa No. 55	Surakarta	012/022	Jebres	Surakarta Timur	Islam	Belum Menikah
8	Dian Pertiwi	3503010101010102	Maluku	Tue Dec 10 00:00:00 WIB 1985	Perempuan	Jl. Pattimura No. 13	Ambon	023/024	Baguala	Ambon	Kristen	Menikah
9	Rudi Hartono	3503010101010103	Pontianak	Mon May 28 00:00:00 WIB 1973	Laki-laki	Jl. Ahmad Yani No. 30	Pontianak	025/026	Pontianak Barat	Pontianak Kota	Hindu	Menikah
10	Lina Fitriani	3503010101010104	Manado	Mon Aug 15 00:00:00 WIB 1994	Perempuan	Jl. Wolter Monginsidi No. 17	Manado	027/028	Mampangret	Manado Selatan	Kristen	Belum Menikah
11	Rina Agustina	3503010101010106	Bengkulu	Mon Nov 19 00:00:00 WIB 1990	Perempuan	Jl. Imam Bonjol No. 45	Bengkulu	031/032	Ratu Samban	Bengkulu Tengah	Kristen	Menikah
12	Nia Dewi Lestari	3503010101010108	Manokwari	Sun Jun 14 00:00:00 WIB 1987	Perempuan	Jl. Diponegoro No. 40	Manokwari	035/036	Manokwari Barat	Manokwari	Kristen	Belum Menikah
13	Doni Setiawan	3503010101010109	Palu	Mon Mar 07 00:00:00 WIB 1987	Laki-laki	Jl. Soekarno-Hatta No. 70	Palu	037/038	Palu Barat	Palu	Islam	Menikah

Terdapat 13 Value dalam Data Masyarakat (Penulis,Wiraswasta,Freelancer,Guru,Pegawai Swasta,Akuntan)

**Output :**

