

JOBSHEET 1

DATA WAREHOUSE SEDERHANA & PDI SPOON

Nama : Dahniar Davina

Kelas : SIB – 2A

No Absen : 09

TUGAS 1

Pak Amir adalah Manager Sales Astro Mobil, yang bergerak di bidang distributor mobil wilayah Jawa Timur. Pak Amir meminta data penjualan dari beberapa cabang kepada supervisor. Data tersebut akan digunakan untuk membuat suatu Dashboard pengambilan keputusan. Dikarenakan proses permintaan ini dilakukan secara berulang setiap hari setelah jam kantor maka, Pak Amir akan membuat proses untuk menarik data yang ada pada file milik supervisor tersebut secara otomatis. Namun, kadang data tersebut tidak lengkap. Sehingga Pak Amir akan mengambil data yang lengkap saja dan mengembalikan data yang tidak lengkap.

Analisa lah data tersebut!

1. Berapa jumlah kolom pada data tersebut?
 - Ada 7 kolom
2. Apa arti atau isi data dari setiap kolom yang ada?
 - Customer = menunjukkan nama/ID pelanggan yang melakukan transaksi
 - Product = jenis produk yang dibeli oleh pelanggan
 - Region = wilayah/lokasi tempat terjadinya proses transaksi
 - Date = tanggal terjadinya proses transaksi dalam format MM/DD/YYYY
 - Item unit price = harga per unit produk yang terjual
 - No items = jumlah unit produk yang dibeli dalam satu kali transaksi
 - Total Sale = total penjualan produk dalam 1 kali transaksi
3. Adakah data yang memiliki nilai null / data yang tidak lengkap?
 - Ada, 3 kolom product, 3 kolom region, 2 kolom date
4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?
 - Tidak ada, semua memiliki tipe yang sama

TUGAS 2

1. Apakah data hasil eksekusi sesuai dengan data aslinya?
 - Ya, secara keseluruhan isi data sama, hanya tampilan kolom Date pada Pentaho lebih lengkap yaitu ada keterangan waktu/timestamp.
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?
 - Setelah file extract tersimpan
 - Tab transformasi terbuka
 - Menampilkan tab transformasi
 - Mulai mengeksekusi file pada tab transformasi
 - Tab transformasi mulai melakukan proses transformasi extract data
 - Proses input data file CarDistribution dengan melewati header Row pada file
 - Proses input data selesai
 - Dan proses 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/Loaded) = 20**

Jumlah total data atau record yang dimasukkan atau dibaca dari file CSV.

➤ **O (Output/Processed Successfully) = 0**

Jumlah data yang berhasil diproses dan ditulis sebagai output. Dalam kasus ini, nilainya 0, yang berarti tidak ada data yang keluar atau disimpan setelah diproses.

➤ **R (Rejected) = 0**

Jumlah data yang ditolak karena kesalahan atau ketidaksesuaian format. Nilai 0 menunjukkan bahwa tidak ada data yang ditolak.

➤ **W (Warnings) = 19**

Jumlah record yang diproses dengan peringatan. Meskipun tidak menyebabkan kegagalan, ada sesuatu yang perlu diperhatikan.

➤ **U (Unknown/Skipped) = 0**

Jumlah record yang dilewati atau tidak dapat diidentifikasi dalam pemrosesan.

➤ E (Errors) = 0

Jumlah record yang gagal diproses karena kesalahan fatal. Dalam kasus ini, tidak ada error.

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	Processor	RAM	Storage	VGA	CPU
DAHNIAR	1detik	Intel(R) Core(TM) i5- 6300U	8 GB	256 GB	Intel(R) HD Graphics 520	CPU @ 2.40GHz 2.50 GHz
YONANDA	1detik	Intel(R) Core(TM) i5- 12450H	16 GB	512 GB	NVIDIA GeForce RTX 3050	12 CPU ~2.0GHZ
AQILAH	1detik	AMD RYZEN 5 5600H	8 GB	512 GB	AMD Radeon™ Graphics	AMD RYZEN 5 5600H with Radeon Graphics
REISHI	1detik	Intel(R) Core™ i7-7820HQ	16 GB	256 GB	NVIDIA GeForce 940MX	2.90Ghz

TUGAS 3

1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?

A	B	C	D	E	F	G
Customer	Product	Region	Date	Item unit p	No.Items	Total Sale
Customer		MLG	#####	0168,95	2,00	00337,90
Customer	Avanza AT			0168,95	15,00	02534,25
Customer				0340,95	7,00	02386,65
Customer		KDR	#####	0799,95	14,00	11199,30
Customer	Avanza AT		#####	0168,95	4,00	00675,80

File excel

- Jumlah data pada file excel lebih sedikit, karena hanya berisi file dengan data yang bernilai NULL
- Tampilannya pun berbeda, jika excel menggunakan sekat cell untuk memisah tiap jenis data yang ada dan berbentuk menyerupai table.

```
Customer;Product;Region;Date;Item unit price;No.Items;Total Sale
Customer 1;Yaris 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;Yaris 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;Yaris AT ;MLG;6/9/2020;340.9;4;1363.8
Customer 2;Avanza AT;KDR;8/8/2020;168.9;12;2027.4
```

File csv

- Sedangkan pada file csv, jumlah data lebih banyak, karena berisi file dengan data yang tidak memiliki nilai NULL
 - Untuk tampilannya pun berbeda, menggunakan (;) titik koma untuk memisahkan tiap jenis data pada satu baris dan berbentuk tulisan biasa.
2. Jelaskan apa yang terjadi pada proses Filter rows!
- Pada proses filter rows terjadi proses penyaringan dari data yang telah diinputkan ke text file input jika dalam kondisi TRUE (tidak memiliki nilai NULL didalamnya) dan otomatis tersaring ke Microsoft excel output jika dalam kondisi FALSE (memiliki nilai NULL didalamnya).
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 table

Nama	Waktu	Processor	RAM	Storage	VGA	CPU
DAHNIAR	1detik	Intel(R) Core(TM) i5- 6300U	8 GB	256 GB	Intel(R) HD Graphics 520	CPU @ 2.40GHz 2.50 GHz
YONANDA	2detik	Intel(R) Core(TM) i5- 12450H	16 GB	512 GB	NVIDIA GeForce RTX 3050	12 CPU ~2.0GHZ
AQILAH	2detik	AMD RYZEN 5 5600H	8 GB	512 GB	AMD Radeon™ Graphics	AMD RYZEN 5 5600H with Radeon Graphics

REISHI	1detik	Intel(R) Core™ i7-7820HQ	16 GB	256 GB	NVIDIA GeForce 940MX	2.90Ghz
--------	--------	-----------------------------	-------	--------	----------------------------	---------

STUDI KASUS

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 karena 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 emmbantu mengambil keputusan.

Pak Pascanowo mendapatakan data penduduk yang akan pindah ke wilayahnya melalui link: https://github.com/dik4rizky/datasources/blob/82821b8ec341176f979956c84b79720d4aea_d012/dataDummyKependudukan.csv

Bantulah Pak Pascanowo untuk memishkan 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.

➤ Masyarakat Mahasiswa

☒ First rows ☐ Last rows ☐ Off

[illegible]

➤ Masyarakat Dokter

#	Nama	NIK	Tempat Lahir	tanggal Lahir	Jenis Kelam...	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perka...	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-Alan...	Palembang Selat...	Kristen	Belum Meni...	Dokter
2	Bambang Kusumo	3503010101010017	Banjarmasin	Thu Sep 25 00:00:00 ICT 1980	Laki-laki	Jl. A. Yani No. 60	Banjarmasin	033/034	Banjarmasi...	Banjarmasin	Hindu	Menikah	Dokter

➤ Masyarakat PNS

#	Nama	NIK	Tempat L...	Tanggal Lahir	Jenis Kel...	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawin...	Pekerjaan
1	Andi Wijaya	35030101010100...	Makassar	Sun Jun 25 00:00:00 ICT 1978	Laki-laki	Jl. Jenderal Sudirman No. 10	Makassar	013/014	Barombong	Makassar Selat...	Islam	Menikah	PNS
2	Agus Santo...	35030101010100...	Padang	Fri Apr 02 00:00:00 ICT 1982	Laki-laki	Jl. Sudirman No. 20	Padang	029/030	Padang Barat	Padang	Islam	Menikah	PNS

➤ Masyarakat Lainnya

#	Nama	NIK	Tempat...	Tanggal Lahir	Jenis Kela...	Alamat	Kota	RT/RW	Kelurahan	Kecamatan	Agama	Status Perkawin...	Pekerjaan
1	Budi Santoso	3503010101010...	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	3503010101010...	Surabaya	Fri Oct 10 00:00:00 ICT 1975	Laki-laki	Jl. Diponegoro No. 67	Surabaya	005/006	Gubeng	Surabaya Pusat	Katolik	Menikah	Pegawai Swasta
3	Maya Dewi	3503010101010...	Yogyakarta...	Sat Aug 20 00:00:00 ICT 1988	Perempu...	Jl. Malioboro No. 8	Yogyakarta...	007/008	Sosromenduran	Yogyakarta Utara	Hindu	Belum Menikah	Guru
4	Rini Setiawan	3503010101010...	Malang	Sun Dec 12 00:00:00 ICT 1982	Perempu...	Jl. Sempang Borobudur No. 15	Malang	011/012	Dinoyo	Malang Selatan	Kristen	Menikah	Wiraswasta
5	Fajar Prabowo	3503010101010...	Medan	Tue Sep 30 00:00:00 ICT 1986	Laki-laki	Jl. Sudirman No. 50	Medan	017/018	Sei Siambing	Medan Timur	Hindu	Menikah	Wiraswasta
6	Rina Sari	3503010101010...	Depas...	Fri Jul 17 00:00:00 ICT 1992	Perempu...	Jl. Gatot Subroto No. 30	Depas...	019/020	Daik Puri Kasih	Depas Utara	Buddha	Belum Menikah	Freelancer
7	Ahmad Subh...	3503010101010...	Surakar...	Wed Jul 03 00:00:00 ICT 1991	Laki-laki	Jl. Pahlawan No. 55	Surakarta	021/022	Jebres	Surakarta Timur	Islam	Belum Menikah	Wiraswasta
8	Dian Pertiwi	3503010101010...	Maluku	Tue Dec 10 00:00:00 ICT 1985	Perempu...	Jl. Pattimura No. 10	Ambon	023/024	Baguaa	Pompon	Kristen	Menikah	Guru
9	Rudi Hartono	3503010101010...	Pontia...	Mon May 28 00:00:00 ICT 1...	Laki-laki	Jl. Ahmad Yani No. 30	Pontiank	025/026	Pontianak Barat	Pontianak Kota	Hindu	Menikah	Wiraswasta
10	Lina Fitriani	3503010101010...	Manado	Mon Aug 15 00:00:00 ICT 1...	Perempu...	Jl. Wolter Monginsidi No. 17	Manado	027/028	Mapangnet	Manado Selatan	Kristen	Belum Menikah	Akuntan
11	Rina Agustina	3503010101010...	Bengk...	Mon Nov 19 00:00:00 ICT 1...	Perempu...	Jl. Imam Bonjol No. 45	Bengkulu	031/032	Ratu Samban	Bengkulu Tengah	Kristen	Menikah	Wiraswasta
12	Nia Dewi Lest...	3503010101010...	Manok...	Sun Jun 14 00:00:00 ICT 1987	Perempu...	Jl. Diponegoro No. 40	Manok...	035/036	Manokwari Bar...	Manokwari	Kristen	Belum Menikah	Penulis
13	Doni Setiawan	3503010101010...	Palu	Mon Mar 07 00:00:00 ICT 19...	Laki-laki	Jl. Soekarno-Hatta No. 70	Palu	037/038	Palu Barat	Palu	Islam	Menikah	Wiraswasta