LAPORAN PRAKTIKUM DATA WAREHOUSE JOBSHEET 1



OLEH:

LOVELYTA SEKARAYU KRISDIYANTI KELAS 2B SIB / 11 (2341760081)

PROGRAM STUDI D-IV SISTEM INFORMASI BISNIS JURUSAN TEKNOLOGI INFORMASI POLITEKNIK NEGERI MALANG

Jl. Soekarno Hatta No. 9 Jatimulyo, Kec. Lowokwaru, Kota Malang, Jawa Timur 65141

TUGAS 1

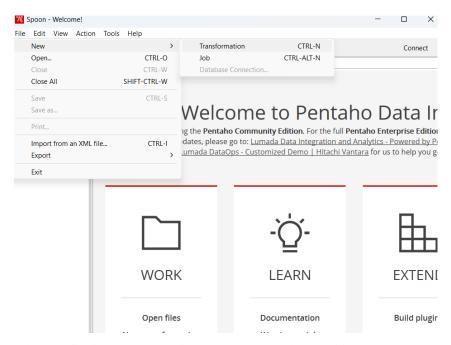
Pak Amir adalah Manager Sales Astro Mobil, yang bergerak di bidang distributor mobil wilayah Jawa Timur. Pak Amir meminta data penjualan dari beberapa cabang kepada supervisor. Data tersebut akan digunakan untuk membuat suatu Dashboard pengambilan keputusan. Dikarenakan proses permintaan ini dilakukan 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.

Analisalah data tersebut!

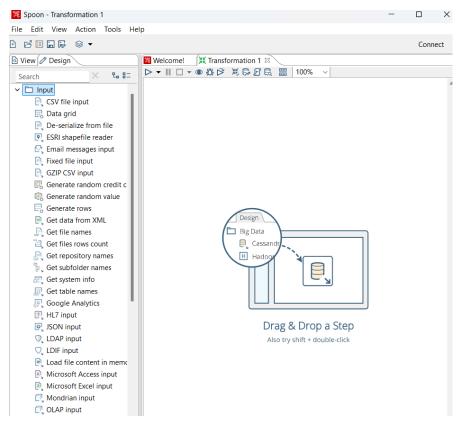
- 1. Berapa jumlah kolom pada data tersebut?
 - Terdapat 7 kolom yaitu customer, product, region, date, item unit price, no. items, total sale
- 2. Apa arti atau isi data dari setiap kolom yang ada?
 - Customer = menunjukkan nama/ID pelanggan yang melakukan transaksi
 - Product = jenis produk yang dibeli 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?
 - Iya, terdapat beberapa kolom yang memiliki nilai kosong pada kolom Product, Region, dan Date.
- 4. Adakah data yang memiliki tipe yang berbeda dengan data lainnya pada kolom yang sama?
 - Tidak ada, semua memiliki tipe yang sama

A. Pengambilan Data (Extract)

1. Buka lembar kerja Transformation melalui toolbar File

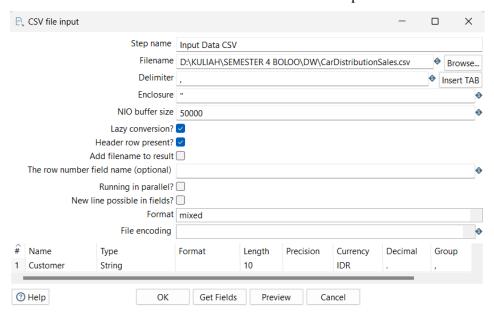


 Cari lah objek CSV file input pada Design Area. Drag and drop objek tersebut menuju Work Area

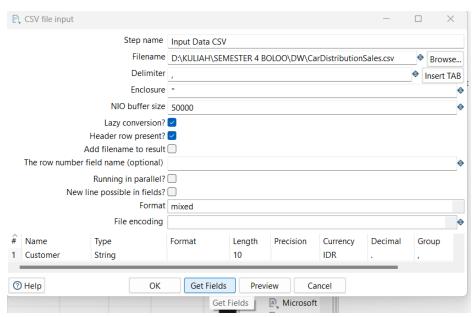


- 3. Double-click pada objek CSV file input hingga muncul jendela konfigurasinya.
- 4. Ubah nama Step name menjadi Input Data CSV, (penamaan digunakan untuk identifikasi saja, selanjutnya penamaan boleh menyesuaikan).

5. Pilih lokasi file CarDistributionSales.csv melalui Browse pada filename.

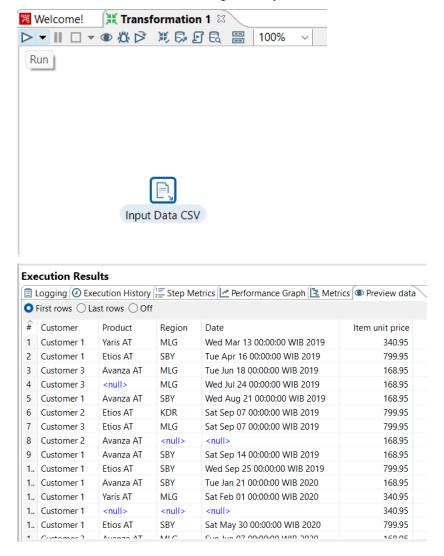


- 6. Jika data telah dipilih pastikan pada bagian Delimiter sesuai dengan kondisi pemisah data pada file yang dipilih. (Pada kesempatan ini adalah comma ", ").
- 7. Tekan Get Fields untuk mengambil judul pada setiap kolom, pastikan nama-nama kolom telah sesuai



						File encoding					
#	Name	Туре	Format	Length	Precision	Currency	Decimal	Group	Trim type		
1	Customer	String		10		IDR		,	none		
2	Product	String		9		IDR		,	none		
3	Region	String		3		IDR		,	none		
4	Date	Date	MM/dd/yyyy			IDR		,	none		
5	Item unit price	Number	#,#	6	2	IDR		,	none		
6	No.Items	Integer	#	15	0	IDR		,	none		
7	Total Sale	Number	#.#	7	2	IDR		,	none		

- 8. Tutup jendela konfigurasi tersebut
- Tekan tombol "Run" pada pojok kanan atas Area Kerja, makan akan muncul Execution Result
 Area. Pilih tab Prieview Data pada Execution Result Area. Jika proses yang dilakukan benar
 maka data akan muncul sesuai dengan aslinya



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?
 - Memulai Ekesekusi Transformasi
 - Membuka koneksi ke sumber data
 - Membaca data
 - Menampilkan jumlah data yang di extract
- 3. Perhatikan gambar dibawah ini! Apa yang dimaksud dengan I,O,R,W,U,E?

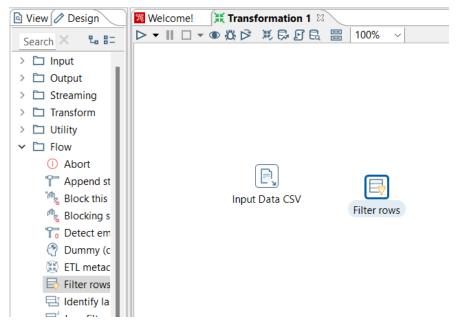
2024/02/17 06:51:48 - Input Data CSV.0 - Finished processing (I=20, O=0, R=0, W=19, U=0, E=0)

- I (Input): Jumlah total data atau record yang dimasukkan atau dibaca dari file CSV = 20
- O (Output): Jumlah data yang ditulis ke output = 0
- R (Rejected): Jumlah data yang ditolak karena kesalahan atau ketidaksesuaian format = 0
- W (Warning): Jumlah record yang diproses dengan peringatan = 19
- U (Unknown): Jumlah record yang dilewati atau tidak dapat diidentifikasi dalam pemrosesan = 0
- E (Errors): Jumlah record yang gagal diproses karena kesalahan fatal = 0
- 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
						Eksekusi
Lovelyta	AMD Ryzen 5	8,00 GB	1 TB	AMD	AMD Ryzen 5	0.01
	3500U with			Radeon(TM)	3500U with	
	Radeon Vega			Vega 8	Radeon Vega	
	Mobile Gfx			Graphics	Mobile Gfx	
	2.10 GHz					
Babby	Intel(R)	4,00 GB	99,4 GB	NVIDIA	Intel(R)	0.01
	Core(TM) i3-			GeForce	Core(TM) i3-	
	6006U CPU			920MX	6006U CPU	
	@ 2.00GHz			Intel(R) HD	@ 2.00GHz	
	1.99 GHz			Graphics	1.99 GHz	
				520		

Danica	Intel(R)	16.0 GB	237 GB	Intel(R)	Intel(R)	0.01
	Core(TM) i5-			UHD	Core(TM) i5-	
	8365U CPU			Graphics	8365U CPU	
	@ 1.60GHz			620	@ 1.60GHz	
	1.90GHz				1.90GHz	
Dhynar	AMID	4,00 GB	1 TB 5400	AMD	AMID	0.01
	Ryzen 3		rpm 2.5"	Radeon TM	Ryzen 3	
	3250U with		SATA	Graphics	3250U with	
	Radeon		Hard Drive		Radeon	
	Graphics				Graphics	
	2600 MHz 2				2600 MHz 2	
	Core(s), 4				Core(s), 4	
	Logical				Logical	
	Processor(s)				Processor(s)	

- B. Filter Data (Transform) dan Pengemasan Data (Load)
 - 1. Carilah objek Filter rows pada Design Area. Drag and drop objek tersebut pada Work Area

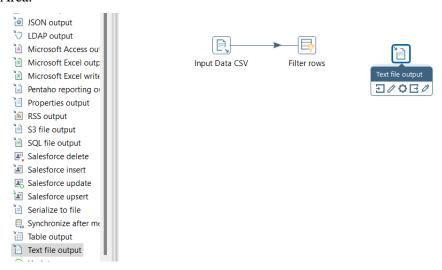


2. Hubungkan output pada Input Data CSV menuju Filter rows. Pilih koneksi Main output of step saat menghubungkan kedua objek tersebut.

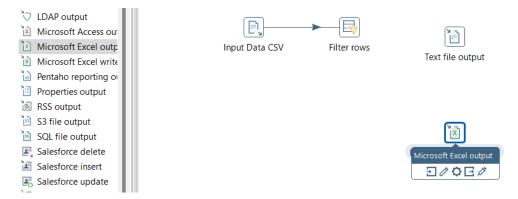
Pada tahap ini Input Data CSV dihubungkan oleh konektor pada Filter rows memiliki makna bahwa setelah proses Input Data CSV dilakukan maka proses selanjutnya adalah Filter rows.



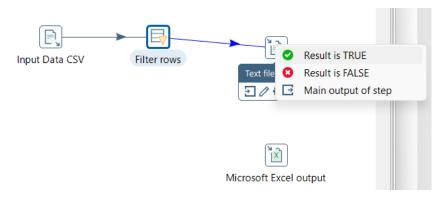
 Carilah objek Text file output pada Design Area. Drag and drop objek tersebut pada Work Area.



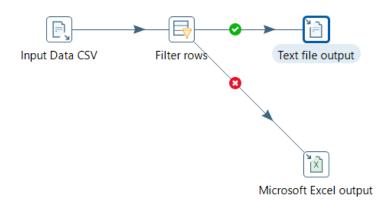
4. Carilah objek Microsoft excel output pada Design Area. Drag and drop objek tersebut pada Work Area.



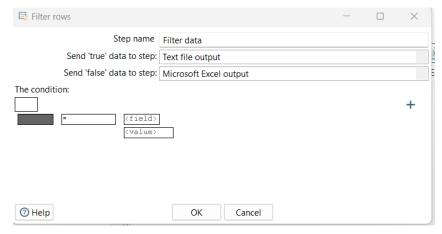
5. Hubungkan filter rows dengan text file output menggunakan konektor output dari filter rows. Pilih Result is true sebagi jenis konektornya.



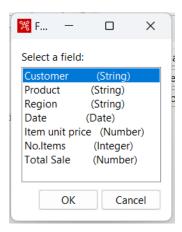
6. Hubungkan filter rows dengan microsoft excel output menggunakan konektor output dari filter rows. Pilih Result is false sebagi jenis konektornya.



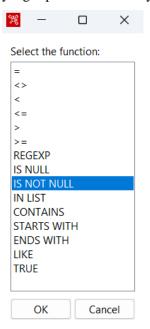
- 7. Double click pada filter rows hingga muncul jendela konfigurasi. ubah step name menjadi Filter data.
- 8. Cek bagian send 'true' data to step menjadi Text file output dan cek bagian Send 'false' data to step menjadi Microsoft Excel output.
- 9. Klik <Field> pada bagian condition



10. Pilih field yang diinginkan untuk digunakan sebagai kondisi. Pada kesempatan ini pilih Customer, kemudian klik OK.

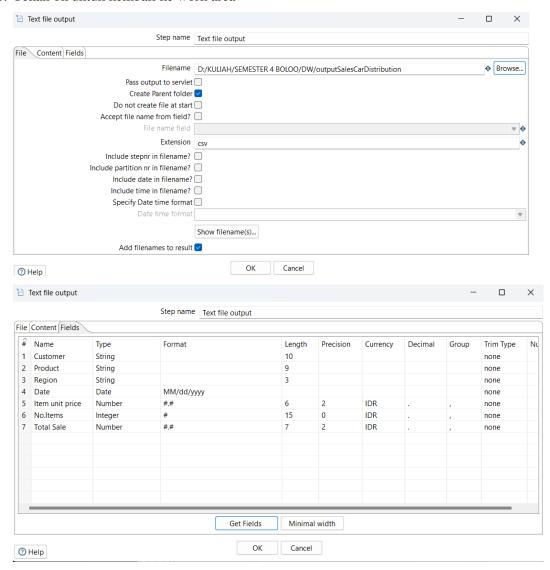


11. Pada bagian value pilih IS NOT NULL kemudian tekan OK , hal ini dilakukan karena data yang dipilih adalah data yang tidak Null

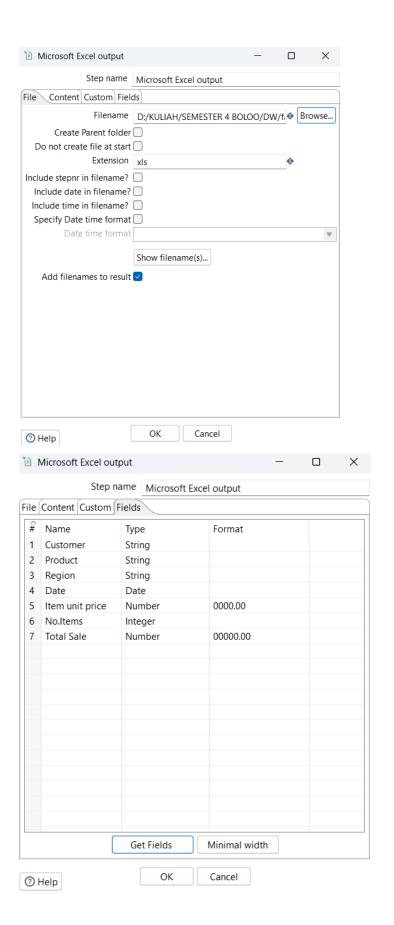


- 12. Tekan tombol + untuk menambah kondisi.
- 13. Pilih statement yang masih null untuk menambah kondisi.
- 14. Ulangi proses nomor 9,10,11 dengan mengganti semua field yang ada. hingga statement menjadi seperti pada gambar dibawah ini.
- 15. Gunakan operator AND untuk menghubung
- 16. Jika semua field sudah masuk dalam kondisi maka tekan OK.
- 17. Double-click pada text file output, pilih lokasi file untuk menyimpan hasil output file dengan menekan tombol browse. beri nama outputSalesCarDistribution.
- 18. Ubah extension menjadi csv.
- 19. Pada tab Fields, tekan tombol Get fields untuk mengambil field data.

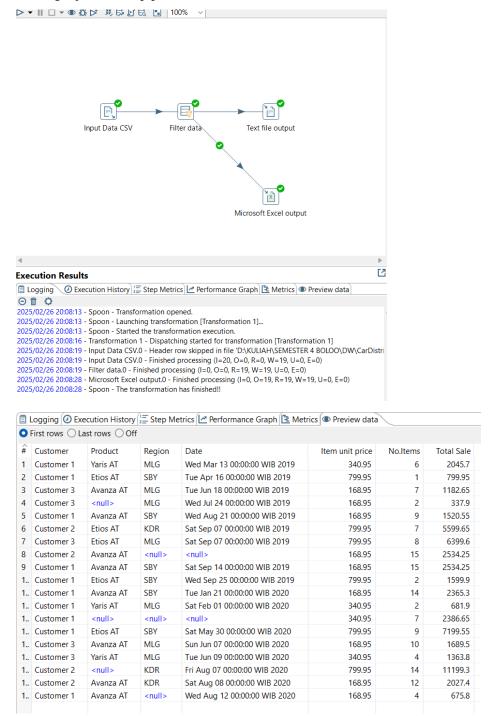
20. Tekan ok untuk kembali ke work area



- 21. Hal yang sama dilakukan pada konfigurasi Microsoft Excel output, Double-click pada objek Microsoft Excel output hingga muncul jendela konfigurasi.
- 22. Pilih lokasi file untuk menyimpan file output dengan menekan tombol browse. Beri nama failedSalesCarDistribution. untuk extension tetap dengan extension xls.
- 23. Pada tab fields tekan tombol Get fields untuk mengambil field data.
- 24. Tekoan ok untuk kembali ke work area.



25. Tekan tombul Run pada pojok kiri atas Work Area. jika berhasil maka akan muncul tanda centang hijau di setiap proses.



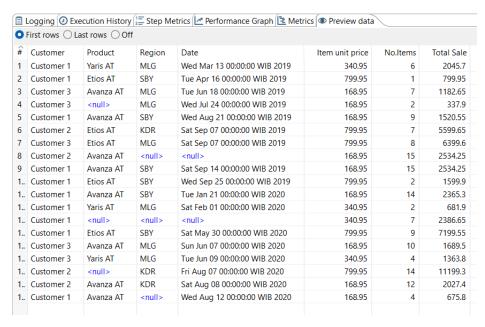
Langkah-langkah diatas merupakan proses melakukan filter data, dimana data yang lengkap akan diletakkan pada text file csv dan data yang belum lengkap akan diletakkan pada ms excel.

TUGAS 3

- 1. Apa perbedaan isi data output dilihat dari isi file csv dan file excel?
 - Excel:
 - Jumlah data lebih sedikit
 - Menggunakan sekat cell untuk memisahkan tiap jenis data yang ada dan berbentuk menyerupai table.

Δ	А	В	C	D	E	F	G
	Customer	Product	Region	Date	Item unit p	No.ltems	Total Sale
!	Customer		MLG	########	0168,95	2,00	00337,90
1	Customer:	Avanza AT			0168,95	15,00	02534,25
	Customer				0340,95	7,00	02386,65
,	Customer:		KDR	########	0799,95	14,00	11199,30
i	Customer	Avanza AT		#########	0168,95	4,00	00675,80

- CSV:
 - Jumlah data lebih banyak
 - 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 tabel

Nama	Processor	RAM	Storage	VGA	CPU	Waktu
						Eksekusi
Lovelyta	AMD Ryzen 5	8,00 GB	1 TB	AMD	AMD Ryzen 5	0.01
	3500U with			Radeon(TM)	3500U with	
	Radeon Vega			Vega 8	Radeon Vega	
	Mobile Gfx			Graphics	Mobile Gfx	
	2.10 GHz					
Babby	Intel(R)	4,00 GB	99,4 GB	NVIDIA	Intel(R)	0.01
	Core(TM) i3-			GeForce	Core(TM) i3-	
	6006U CPU			920MX	6006U CPU	
	@ 2.00GHz			Intel(R) HD	@ 2.00GHz	
	1.99 GHz			Graphics	1.99 GHz	
				520		
Danica	Intel(R)	16.0 GB	237 GB	Intel(R)	Intel(R)	0.01
	Core(TM) i5-			UHD	Core(TM) i5-	
	8365U CPU			Graphics	8365U CPU	
	@ 1.60GHz			620	@ 1.60GHz	
	1.90GHz				1.90GHz	
Dhynar	AMID	4,00 GB	1 TB 5400	AMD	AMID	0.01
	Ryzen 3		rpm 2.5"	Radeon TM	Ryzen 3	
	3250U with		SATA	Graphics	3250U with	
	Radeon		Hard Drive		Radeon	
	Graphics				Graphics	
	2600 MHz 2				2600 MHz 2	
	Core(s), 4				Core(s), 4	
	Logical				Logical	
	Processor(s)				Processor(s)	

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. darai 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 tersbut akan digunakan untuk analisa di masing-masing wilayah dan juga Dashboard untuk membantu mengambil keputusan.

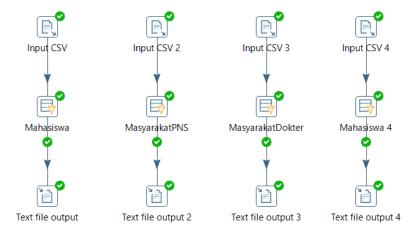
Pak Pascanowo mendapatakan data penduduk yang akan pindah ke wilayahnya melalui link: https://github.com/dik4rizky/datasources/blob/82821b8ec341176f979956c84b79720d4aead012/dataDum myKependudukan.csv

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.

JAWABAN:



- Mahasiswa

Α	В	С	D	E	F	G	Н		J	K	L	M
Nama	NIK	Tempat La	Tanggal La	Jenis Kelar	Alamat	Kota	RT/RW	Kelurahan	Kecamata	Agama	Status Per	Pekerjaan
Siti Rahayı	3,5E+15	Bandung	########	Perempua	Jl. Cendrav	Bandung	003/004	Cikutra	Bandung V	Kristen	Belum Me	Mahasiswa
Adi Pratan	3,5E+15	Semarang	########	Laki-laki	Jl. Pemuda	Semarang	009/010	Pandanara	Semarang	Buddha	Belum Me	Mahasiswa
Maya Wul	3,5E+15	Jayapura	########	Perempua	Jl. Proklam	Jayapura	039/040	Kota Jayar	Jayapura	Kristen	Belum Me	Mahasiswa

- PNS

Α	В	C	D	E	F	G	Н	1	J	K	L	М
Nama	NIK	Tempat La	Tanggal La	Jenis Kelar	Alamat	Kota	RT/RW	Kelurahan	Kecamata	Agama	Status Per	Pekerjaan
Andi Wijay	3,5E+15	Makassar	########	Laki-laki	Jl. Jendera	Makassar	013/014	Barombon	Makassar	Islam	Menikah	PNS
Agus Santo	3,5E+15	Padang	########	Laki-laki	Jl. Sudirma	Padang	029/030	Padang Ba	Padang	Islam	Menikah	PNS

- Dokter

Α	В	С	D	E	F	G	Н	1	J	K	L	M
Nama	NIK	Tempat La	Tanggal La	Jenis Kelar	Alamat	Kota	RT/RW	Kelurahan	Kecamataı	Agama	Status Per	Pekerjaan
Desi Purna	3,5E+15	Palembang	########	Perempua	Jl. Demang	Palemban	015/016	Alang-Alan	Palembang	Kristen	Belum Me	Dokter
Bambang F	3,5E+15	Banjarmas	########	Laki-laki	Jl. A. Yani I	l Banjarmas	033/034	Banjarmas	Banjarmas	Hindu	Menikah	Dokter

- Masyarakat Lainnya

Α	В	C	D	E	F	G	H	1	J	K	L	M	N
Nama	NIK	Tempat La	Tanggal La	Jenis Kelar	Alamat	Kota	RT/RW	Kelurahan	Kecamataı	Agama	Status Perl	Pekerjaan	
Budi Santo	3,5E+15	Jakarta	########	Laki-laki	Jl. Merdek	Jakarta	001/002	Menteng	Jakarta Pu	Islam	Menikah	Wiraswast	a
Joko Susilc	3,5E+15	Surabaya	########	Laki-laki	Jl. Dipone	Surabaya	005/006	Gubeng	Surabaya F	Katolik	Menikah	Pegawai Sv	wasta
Maya Dew	3,5E+15	Yogyakarta	########	Perempua	Jl. Maliobo	Yogyakarta	007/008	Sosromen	Yogyakarta	Hindu	Belum Me	Guru	
Rini Setiaw	3,5E+15	Malang	########	Perempua	Jl. Simpang	Malang	011/012	Dinoyo	Malang Se	Kristen	Menikah	Wiraswast	a
Fajar Prabo	3,5E+15	Medan	########	Laki-laki	Jl. Sudirma	Medan	017/018	Sei Sikamb	Medan Tin	Hindu	Menikah	Wiraswast	a
Rina Sari	3,5E+15	Denpasar	########	Perempua	Jl. Gatot Si	Denpasar	019/020	Dauh Puri	Denpasar I	Buddha	Belum Me	Freelancer	
Ahmad Sul	3,5E+15	Surakarta	########	Laki-laki	Jl. Pahlawa	Surakarta	021/022	Jebres	Surakarta '	Islam	Belum Me	Wiraswast	a
Dian Pertiv	3,5E+15	Maluku	########	Perempua	Jl. Pattimu	Ambon	023/024	Baguala	Ambon	Kristen	Menikah	Guru	
Rudi Harto	3,5E+15	Pontianak	########	Laki-laki	Jl. Ahmad '	Pontianak	025/026	Pontianak	Pontianak	Hindu	Menikah	Wiraswast	a
Lina Fitriar	3,5E+15	Manado	#######	Perempua	Jl. Wolter	Manado	027/028	Mapanget	Manado S	Kristen	Belum Me	Akuntan	
Rina Agust	3,5E+15	Bengkulu	########	Perempua	Jl. Imam B	Bengkulu	031/032	Ratu Samb	Bengkulu T	Kristen	Menikah	Wiraswast	a
Nia Dewi L	3,5E+15	Manokwai	########	Perempua	Jl. Dipone	Manokwai	035/036	Manokwa	Manokwai	Kristen	Belum Me	Penulis	
Doni Setia	3,5E+15	Palu	########	Laki-laki	Jl. Soekarn	Palu	037/038	Palu Barat	Palu	Islam	Menikah	Wiraswast	а