



PORTFOLIO

dwmhr

SELEKSI MANDIRI 2020

DEWI MAHARANI



Data Warehouse|Data Lakehouse|Data Visualization

Background

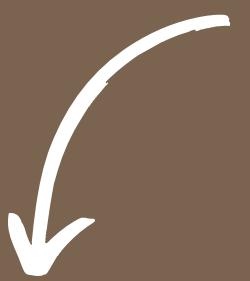


Proyek ini mengimplementasikan proses ETL (Extract, Transform, Load) untuk membangun Gudang Data yang menyimpan data pelamar seleksi mandiri setiap tahun. Data yang diekstraksi meliputi nilai akhir, jumlah pelamar per program studi, asal pelamar, dan profil orang tua. Data ditransformasikan untuk memastikan konsistensi dan struktur yang tepat, kemudian dimuat ke dalam Gudang Data untuk analisis lebih lanjut. Proses ETL ini bertujuan untuk mengelola data secara efisien dan mendukung pengambilan keputusan berbasis data.

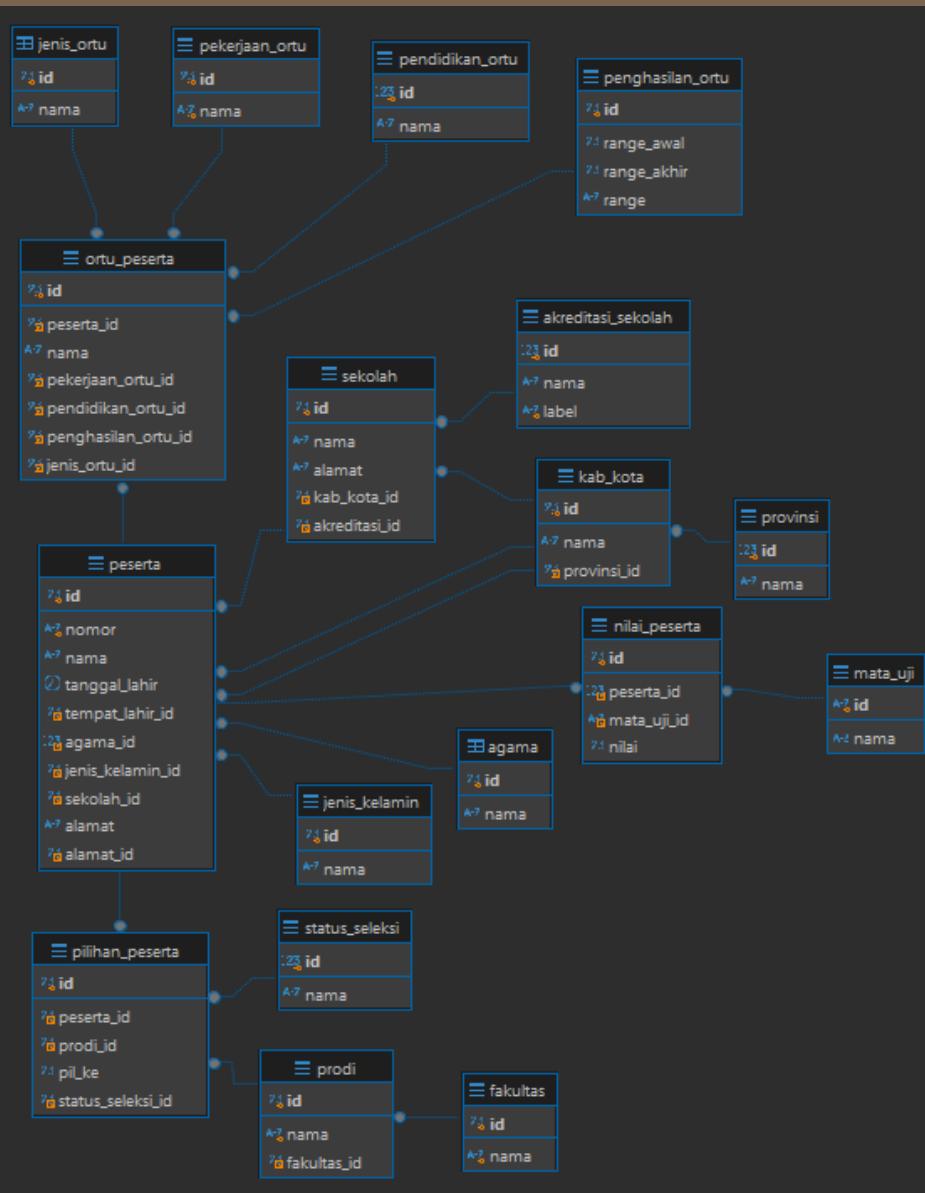


Alat yang digunakan





Analisis Struktur Database Seleksi Mandiri 2020



seleksi_mandiri_2020

Database "seleksi_mandiri_2020" menyimpan data pendaftaran dan seleksi mahasiswa jalur mandiri di Universitas XYZ, terdiri dari 18 tabel. Tabel utama seperti Peserta, Nilai_Peserta, Prodi, dan Pilihan_Peserta menyimpan data pendaftar, nilai ujian, dan status seleksi, sementara tabel lain seperti Agama, Fakultas, dan Jenis Kelamin memberikan informasi tambahan.

Integrasi Data Lakehouse untuk Seleksi Mandiri 2020



A	B	C	E	F	G	H	I	J	K	L
1	idreal	kodepumdesa	kdebp:provinsi	kabupaten	kecamatan	desa	tahun	turunan t nilai	turunan variabel	
2	R8005501	5315122002	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Tueng	2020	52.397.274		
3	R8005501	5315122003	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Sompang Kolang	2020	4.805.392		
4	R8005501	5315122004	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Wajur	2020	51.783.245		
5	R8005501	5315122005	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Kolang	2020	54.736.985		
6	R8005501	5315122006	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Golo Riwu	2020	51.268.889		
7	R8005501	5315122007	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Compong Kules	2020	41.881.616		
8	R8005501	5315122008	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Tengku	2020	4.760.394		
9	R8005501	5315122009	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Golo Wedong	2020	52.161.532		
10	R8005501	5315122010	NUSA TENGGARA TIMUR	MANGGARAI BARA	Kuwus Barat	Ranggu	2020	54.873.095		
11	R8005501	5316012002	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Labolewa	2020	59.651.516		
12	R8005501	5316012004	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Tedamude	2020	53.728.538		
13	R8005501	5316012006	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Nggolonio	2020	561.692		
14	R8005501	5316012008	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Olaia	2020	55.599.299		
15	R8005501	5316012010	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Aerama	2020	57.066.143		
16	R8005501	5316012012	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Marapokot	2020	60.392.709		
17	R8005501	5316012013	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Tonggurambang	2020	65.252.957		
18	R8005501	5316012014	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Waekokak	2020	60.621.337		
19	R8005501	5316012015	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Ngegedhawe	2020	47.824.332		
20	R8005501	5316012016	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Tadakisa	2020	57.356.653		
21	R8005501	5316012017	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Nggolombay	2020	64.922.951		
22	R8005501	5316012018	NUSA TENGGARA TIMUR	NAGEKEO	Aesesha	Nangadhero	2020	54.908.649		
23	R8005501	5316022002	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Podenura	2020	56.707.757		
24	R8005501	5316022003	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Tonggo	2020	53.332.428		
25	R8005501	5316022004	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Wokodekororo	2020	56.486.086		
26	R8005501	5316022005	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Riti	2020	49.526.593		
27	R8005501	5316022006	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Kotakeo	2020	55.777.421		
28	R8005501	5316022007	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Pagomogo	2020	49.877.329		
29	R8005501	5316022008	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Degalea	2020	56.432.317		
30	R8005501	5316022009	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Ulupulu	2020	47.140.327		
31	R8005501	5316022010	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Bidoa	2020	50.651.209		
32	R8005501	5316022011	NUSA TENGGARA TIMUR	NAGEKEO	Nangaroro	Utetoto	2020	4.669.138		



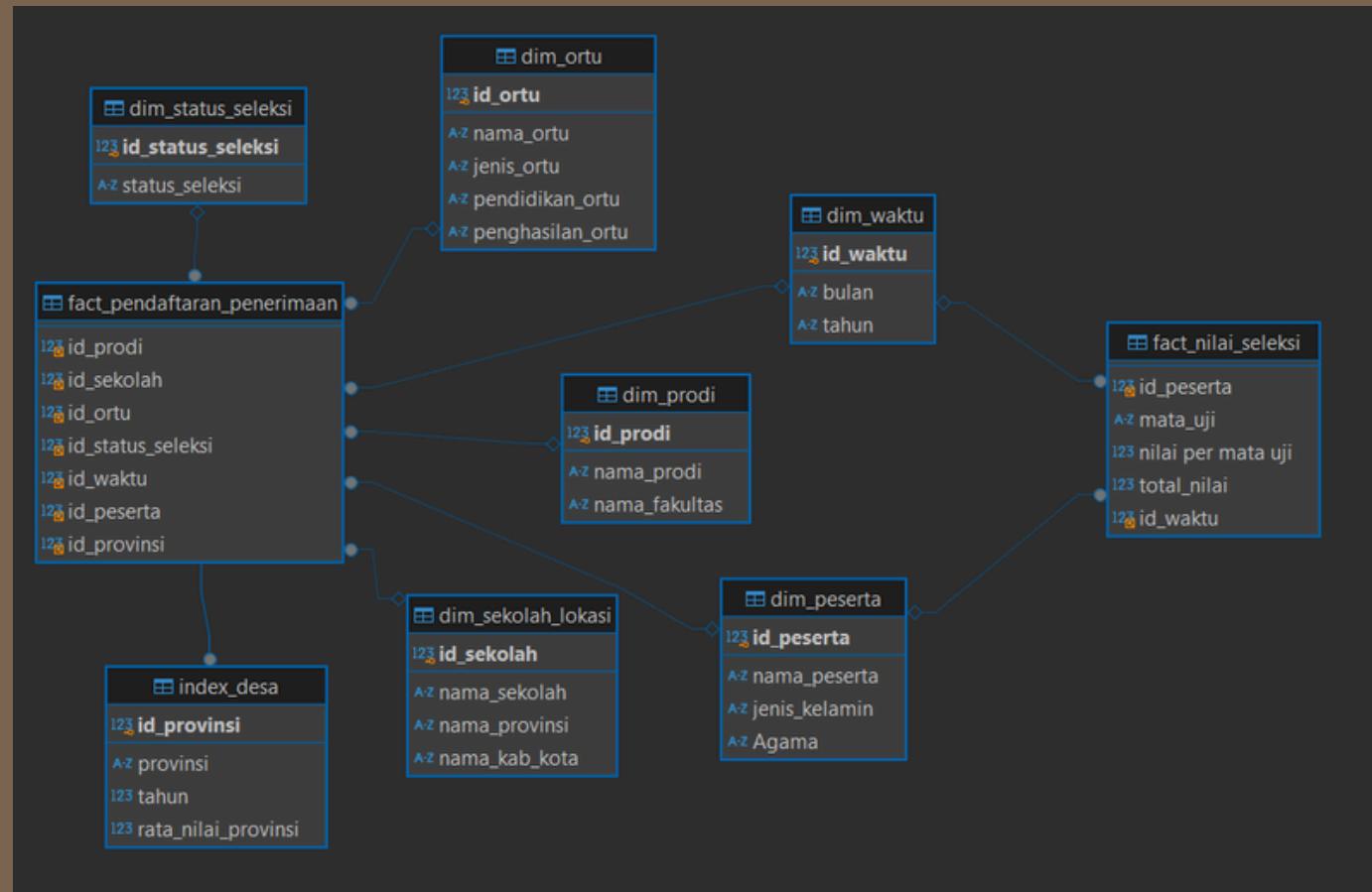
A	B	C	D
1 provinsi	tahun	IDProv	rataNilaiProv
2 Aceh		2020	11 52.46
3 Bali		2020	51 67.53
4 Banten		2020	36 57.79
5 Bengkulu		2020	17 52.47
6 DI Yogyak		2020	34 69.61
7 Gorontalo		2020	75 59.95
8 Jambi		2020	15 55.06
9 Jawa Bara		2020	32 64.07
10 Jawa Teng		2020	33 63.21
11 Jawa Timu		2020	35 62.19
12 Kalimanta		2020	61 49.16
13 Kalimanta		2020	63 56.05
14 Kalimanta		2020	62 49.93

=VLOOKUP(D2;Sheet1!\$B\$1:\$C\$34;2;0)

=MID(B7;7;LEN(B7)-5)

Ditambahkan data dari Satu Data Indonesia, yang kemudian diolah di Excel untuk mendapatkan data yang relevan sesuai kebutuhan analisis. Data hasil olahan ini diintegrasikan dengan database Seleksi Mandiri 2020.

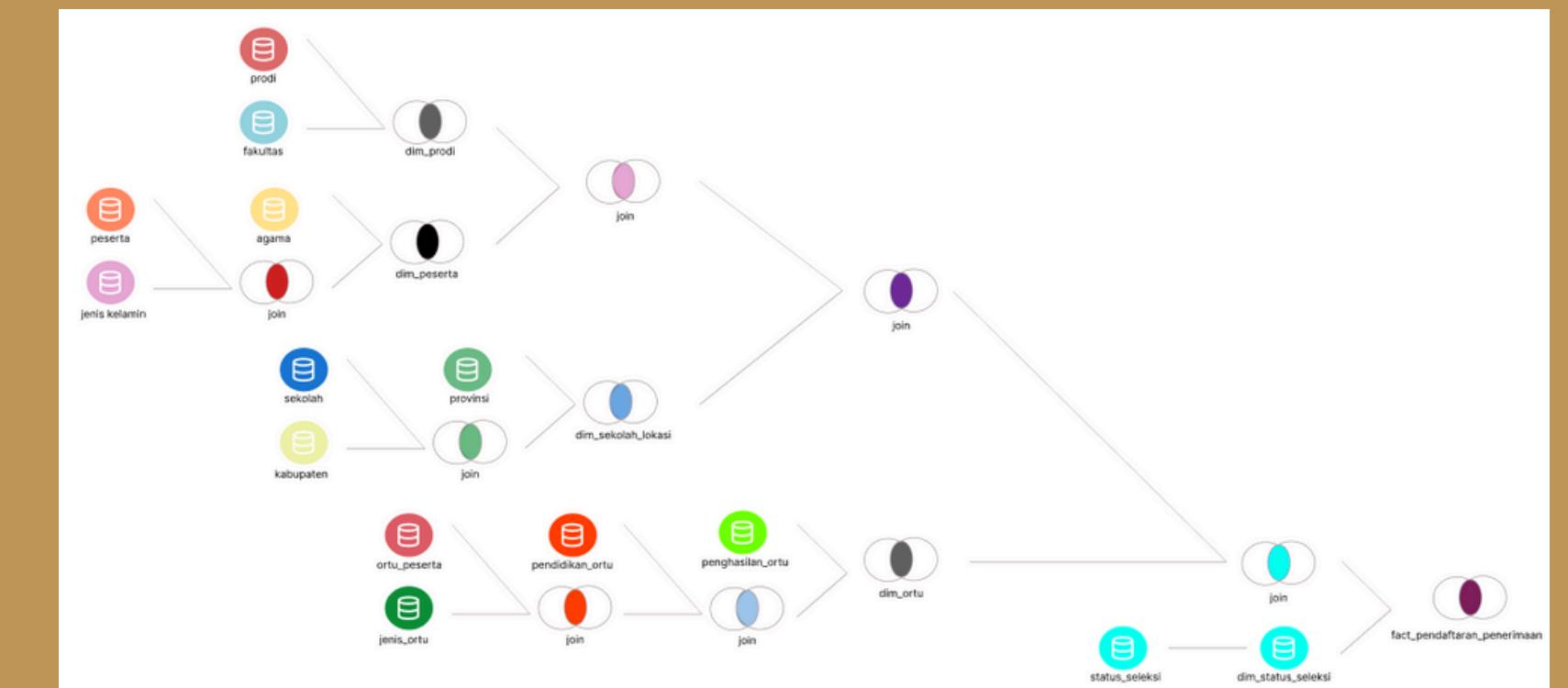
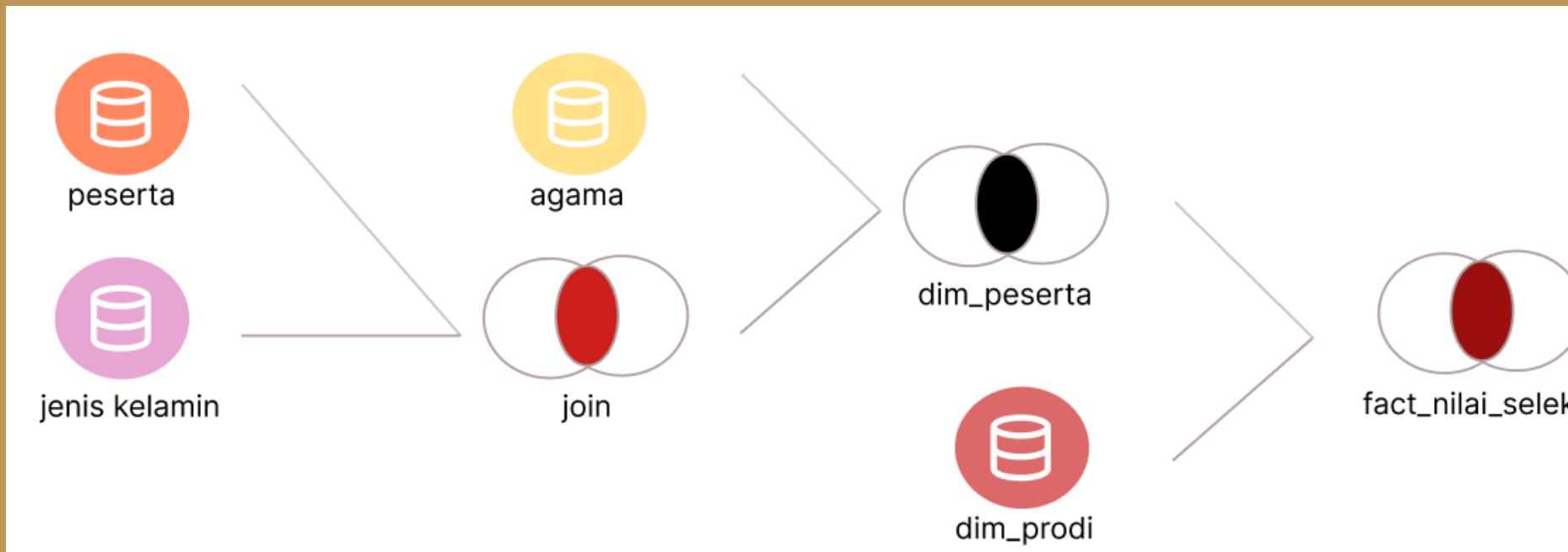
DESAIN STAR SCHEMA SELEKSI MANDIRI 2020



Star schema ini dirancang untuk mengakses dan memproses data secara efisien, memberikan gambaran lengkap tentang nilai peserta, jumlah pendaftar per program studi, profil pendaftar, serta informasi orang tua dan status seleksi, yang akan mendukung pembuatan dashboard untuk keputusan strategis manajemen.



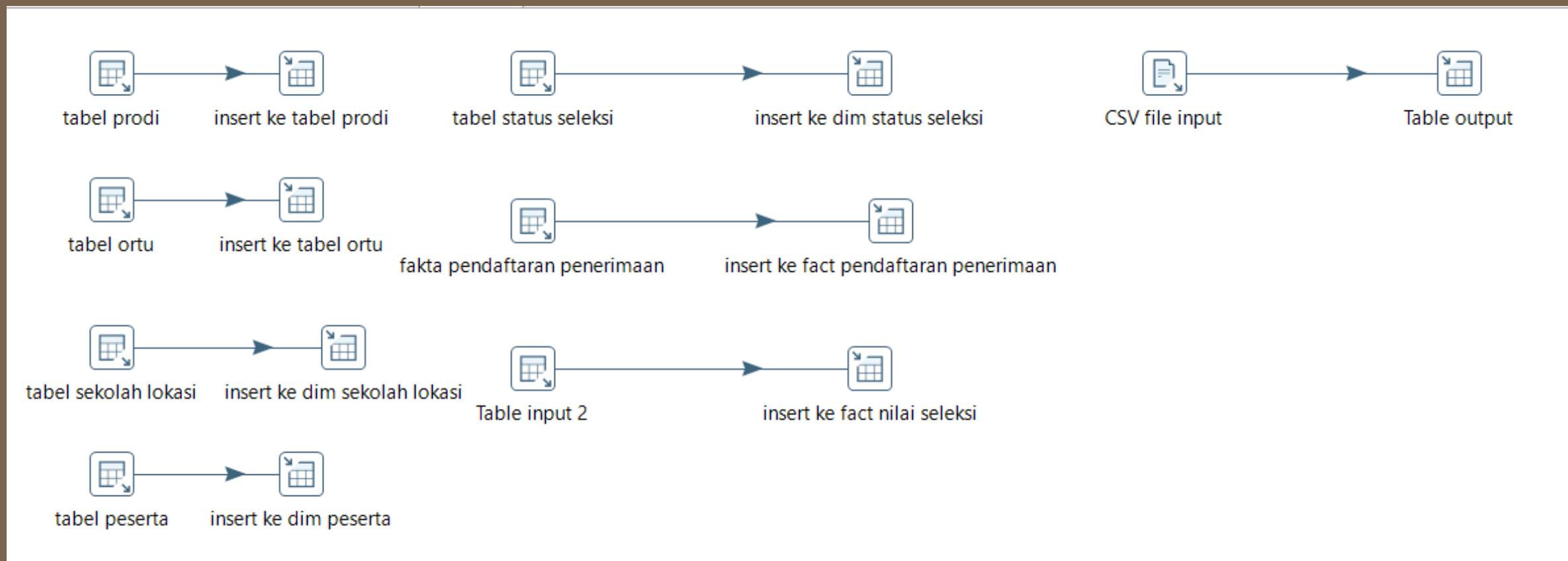
Diagram ETL Seleksi mandiri 2020



Setelah struktur database Seleksi Mandiri 2020 dirancang dan diorganisasikan ke dalam Star Schema, tahap berikutnya adalah memastikan data yang diperlukan dapat diekstraksi, diolah, dan dimuat ke dalam data warehouse.

Pada tahap ini, proses ETL menjadi kunci untuk mengintegrasikan data dari berbagai sumber. Proses ini tidak hanya bertujuan untuk mentransfer data, tetapi juga untuk membangun tabel fakta utama, yaitu `fact_pendaftaran_penerimaan` dan `fact_nilai_seleksi`, yang menjadi pusat analisis data Seleksi Mandiri 2020.

ETL dengan Pentaho



```
SELECT
    prodi.id AS id_prodi,
    prodi.nama AS nama_prodi,
    fakultas.nama AS nama_fakultas
FROM
    prodi
JOIN
    fakultas
ON
    prodi.fakultas_id = fakultas.id;
```

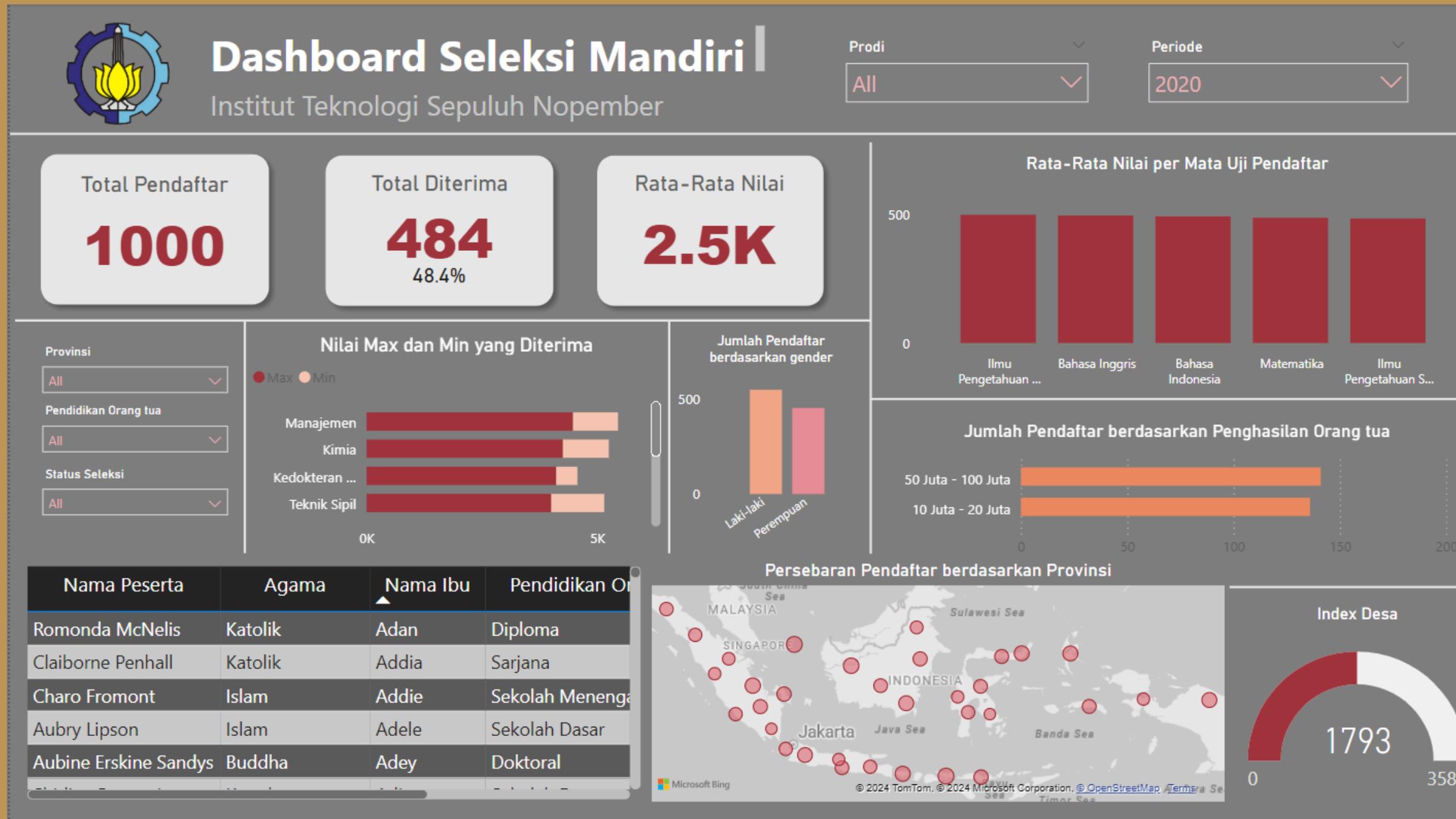
	id_ortu	nama_ortu	jenis_ortu	pendidikan_ortu	penghasilan_ortu
1	1,001	Aprilette	Ibu	Sekolah Menengah Pertama	500 Juta - 1 Miliar
2	1,002	Raychel	Ibu	Sekolah Dasar	50 Juta - 100 Juta
3	1,003	Cele	Ibu	Sekolah Menengah Atas	10 Juta - 20 Juta
4	1,004	Randee	Ibu	Sekolah Dasar	10 Juta - 20 Juta
5	1,005	Doti	Ibu	Sekolah Menengah Atas	10 Juta - 20 Juta
6	1,006	Constantine	Ibu	Diploma	0 - 1 Juta
7	1,007	Viola	Ibu	Sekolah Menengah Pertama	100 Juta - 500 Juta
8	1,008	Imojean	Ibu	Sarjana	50 Juta - 100 Juta
9	1,009	Laurette	Ibu	Diploma	100 Juta - 500 Juta
10	1,010	Janaya	Ibu	Magister	50 Juta - 100 Juta
11	1,011	Nell	Ibu	Magister	0 - 1 Juta
12	1,012	Latrina	Ibu	Sekolah Menengah Atas	0 - 1 Juta
13	1,013	Monah	Ibu	Sekolah Menengah Atas	50 Juta - 100 Juta
14	1,014	Fleur	Ibu	Diploma	20 Juta - 50 Juta
15	1,015	Danna	Ibu	Sekolah Menengah Atas	1 Juta - 5 Juta
16	1,016	Margarete	Ibu	Magister	20 Juta - 50 Juta
17	1,017				10 Juta - 20 Juta

Setelah diagram ETL dirancang, langkah selanjutnya adalah implementasi proses Extract, Transform, dan Load menggunakan Pentaho Data Integration. Dalam proses ini, Pentaho digunakan untuk menjalankan alur ETL sesuai dengan diagram yang telah dibuat.



Visualisasi Data Dengan power BI

Dashboard Seleksi Mandiri



Setelah ETL dan integrasi data, langkah berikutnya adalah visualisasi hasil analisis menggunakan Power BI. Dashboard interaktif ini menyajikan data Seleksi Mandiri 2020 secara intuitif, memudahkan pemangku kepentingan untuk menggali wawasan dan mendukung pengambilan keputusan berbasis data.



PORTFOLIO

dwmhr



**TERIMA
KASIH :)**

DEWI MAHARNI

