

LAPORAN EVALUASI AKHIR SEMESTER DATA LAKEHOUSE



OLEH:

INGGRIT RISMAULI SIAHAAN (5026211012)

RETHA NOVIANTY SIPAYUNG (5026211028)

IVANA LITA (5026211058)

DLH (B)

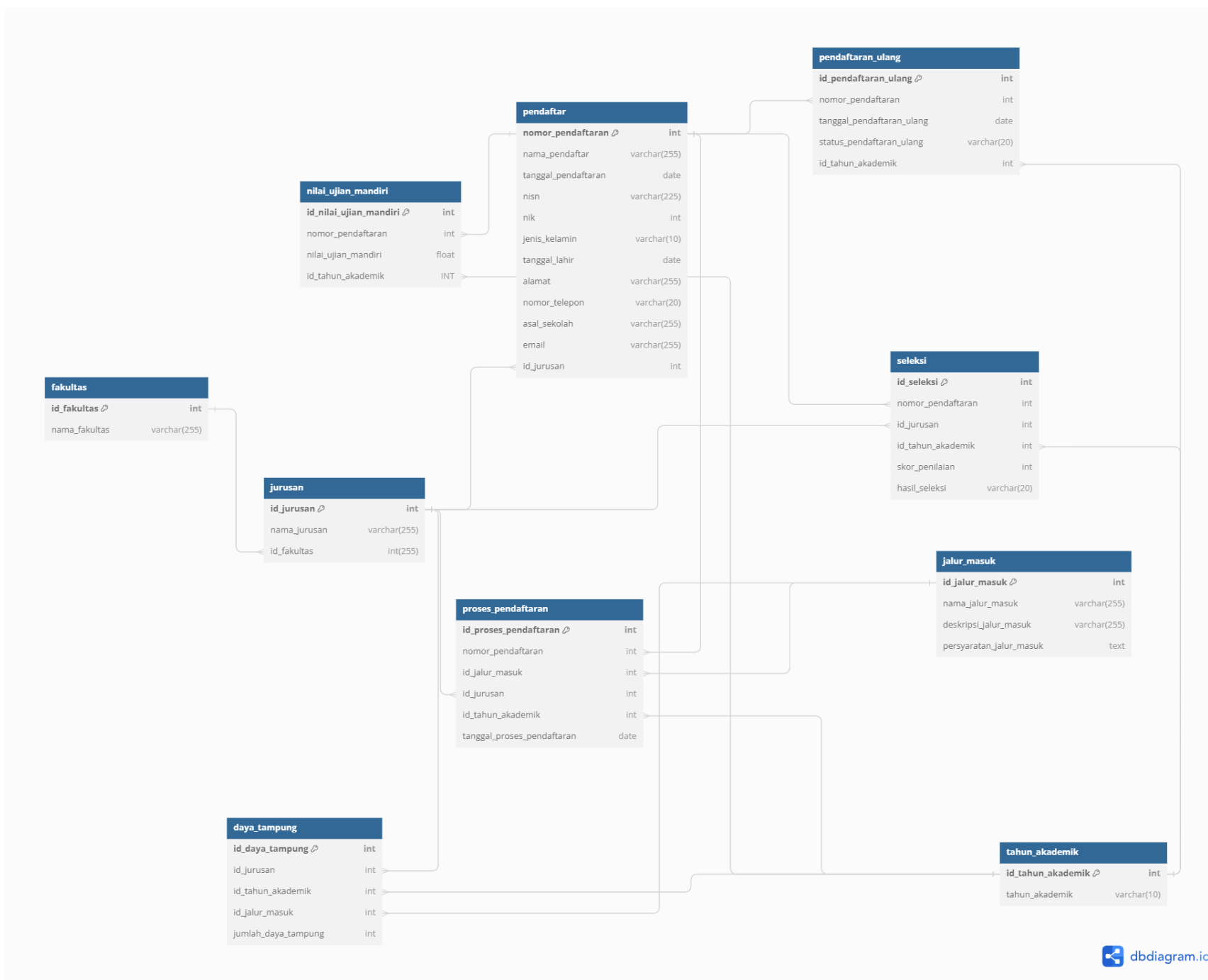
Desain & Implementasi Data Warehouse.....	3
a. Menjelaskan Proses Bisnis.....	3
b. Mendeklarasikan Granularitas.....	9
c. Menentukan Dimensi.....	10
d. Menentukan Fakta.....	10
e. Menghasilkan Tabel pada RDBMS.....	12
Desain & Implementasi ETL.....	13
a. Data & OLTP Profiling.....	13
Membuat Tabel OLTP.....	14
Import daya_tampung.....	18
Import tahun_akademik.....	18
Import jurusan.....	19
Import jalur_masuk.....	19
Import pendaftaran ulang.....	19
Import nilai_ujian_mandiri.....	20
b. High Level ETL Design.....	27
c. Source to Target Mapping.....	31
d. ETL one time historical dan ETL incremental (Teknologi : Pentaho).....	33
Contoh pemanfaatan dari Data Warehouse.....	47
a. Menjawab pertanyaan Who, Where, When, What.....	47
b. Menjawab pertanyaan Why.....	48
Desain & Implementasi Dashboard.....	49
a. Rancangan awal dashboard.....	49
b. Implementasi Dashboard.....	56

Desain & Implementasi Data Warehouse

a. Menjelaskan Proses Bisnis

College Admission Data Warehouse di Institut Teknologi Sepuluh Nopember (ITS) menjadi pondasi utama untuk mengelola dan menganalisis data terkait penerimaan mahasiswa baru. Sebagai sebuah entitas penting di ITS, data warehouse ini dirancang untuk menyimpan, mengorganisir, dan menyediakan akses yang efisien terhadap data pendaftaran mahasiswa. Penting untuk dicatat bahwa dalam penjelasan ini kami menggunakan data dummy, yang merupakan data tiruan yang dibuat untuk keperluan ilustrasi dan pembelajaran. Data dummy ini tidak mencerminkan data sebenarnya dari College Admission di ITS, namun digunakan sebagai contoh untuk menjelaskan konsep dan proses yang terlibat dalam proyek akhir ini.

Sebelum memahami lebih dalam tentang proses bisnis yang terdapat dalam College Admission ITS, berikut kami tampilkan OLTP (Online Transaction Processing) yang akan mempermudah pembaca untuk memahami setiap tahapan yang terlibat dalam proyek akhir ini. OLTP ini memberikan gambaran singkat mengenai apa yang terjadi dalam pengelolaan pendaftaran mahasiswa baru di ITS.



link OLTP : https://dbdiagram.io/d/OLTP-FP-DLH_revised-1-65708bfa56d8064ca08307c0

Proses Bisnis	Deskripsi
Penentuan Daya Tampung	<p>Proses ini dimulai tepat sebelum masa penerimaan mahasiswa baru perguruan tinggi untuk tahun ajaran baru dilaksanakan. Daya tampung ditentukan dengan mempertimbangkan antara jumlah mahasiswa dalam setiap program studi dengan kapasitas saran dan prasarananya, dosen dan tenaga kependidikan, serta layanan dan sumber daya pendidikan lainnya. Sementara untuk penentuan alokasi daya tampung di setiap jalur seleksinya mengikuti target yang ditetapkan oleh kemendikbud dalam Permendikbudriset No. 48 Tahun 2022 , yaitu:</p> <ul style="list-style-type: none"> - Jalur Undangan, sebanyak 20% untuk setiap program studi - Jalur Tes, karena ITS merupakan PTN-BH maka daya tampung minimal sebanyak 40% untuk setiap program studi - Jalur Mandiri, karena ITS merupakan PTN-BH maka daya tampung maksimal sebanyak 50% untuk setiap program studi <p>Sejak tahun 2021 hingga saat ini, target tersebut masih sama sesuai ketetapan kemendikbud. Sementara itu, untuk program studi (khusus diploma dan sarjana) di ITS sendiri memiliki sedikit perubahan dalam 3 tahun terakhir. Daya tampung di atas akan menjadi acuan untuk penerimaan mahasiswa baru ITS dari masing-masing jalur seleksi untuk setiap program studinya. Proses ini berakhir dengan diumumkannya hasil penentuan daya tampung untuk menjadi acuan informasi bagi calon pendaftar untuk menentukan program studinya.</p> <p>Sumber:</p> <p>https://www.its.ac.id/ppid/wp-content/uploads/sites/68/2020/09/Nomor-20-Tahun-2020-Panduan-Penerimaan-Mahasiswa-Baru-Program-Sarjana-S1-Melalui-SBMPTN-TA-2020-2021.pdf</p> <p>https://www.its.ac.id/id/kuliah-di-its/program-studi/program-diploma/</p> <p>https://www.its.ac.id/id/kuliah-di-its/program-studi/program-sarjana/</p>
Pendaftaran	<p>Terdapat 3 jalur pendaftaran untuk masuk ke ITS, diantaranya jalur undangan, jalur tes dan mandiri</p> <ul style="list-style-type: none"> • Jalur undangan: Proses ini diawali dengan siswa SMA/SMK yang berstatus <i>eligible</i> untuk mengikuti jalur undangan daftar ke portal seleksi mahasiswa baru yaitu website LTMPT, menggunakan NISN, NPSN, dan tanggal lahir. Kemudian, siswa melakukan aktivasi akun menggunakan email, dan login ke akun yang telah dibuat. Selanjutnya, siswa memverifikasi data dirinya yang sudah ada sebelumnya. Apabila masih terdapat data diri yang belum di isi maka diisikan dengan benar. Setelah data diri lengkap maka lanjut ke pemilihan program studi dan pengisian kebutuhan data

seperti portofolio dan prestasi jika ada. Lalu, ketika semua data sudah dirasa valid maka mahasiswa dapat melakukan finalisasi pendaftaran sehingga data tidak dapat diubah lagi. Kartu registrasi siswa akan otomatis dibuat dan siswa wajib mencetak kartu tersebut untuk keperluan administrasi nantinya.



Sumber: https://drive.google.com/file/d/1Vac_VYBNWIXjThbyjiIALC568mGfKHUd/view

- Jalur tes: Jalur tes bersama melalui nilai UTBK. Siswa harus membuat akun di portal seleksi mahasiswa baru, yaitu website LTMPPT (apabila belum ada akun). Setelah registrasi akun dan login, mahasiswa harus melengkapi biodata diri. Selanjutnya siswa dapat melakukan pendaftaran UTBK dengan memilih maksimal 2 program studi, perguruan tinggi, mengunggah portofolio (apabila ada), memilih pusat UTBK dan membayar biaya UTBK. Siswa kemudian akan mengikuti tes UTBK dengan mata uji yang sesuai dengan syarat pendaftaran prodi pada pusat UTBK yang dipilih
- Jalur mandiri: Proses ini diawali dengan calon mahasiswa melakukan pembuatan akun pendaftaran pada website myITS Admission (memerlukan NIK, Foto diri, KTP). Hanya siswa yang lulusan 3 tahun terakhir periode pendaftaran yang dapat mendaftar. Kemudian setelah mendaftarkan email sebagai myITS ID dan password, siswa dapat melakukan login dan memilih jalur seleksi mandiri (terdapat 4 kategori jalur mandiri di ITS, yakni kemitraan, mandiri umum, kedokteran & prestasi). Siswa akan mengisi formulir pendaftaran online sesuai dengan petunjuk yang ada (khusus untuk jalur mandiri prestasi akan menginput bukti prestasi), memilih program studi sesuai dengan kategori saintek/ soshum/ desain maksimal 2 pilihan, dan membayar biaya pendaftaran untuk mendapatkan kartu pendaftaran. Peserta jalur mandiri

	<p>umum, kemitraan dan kedokteran akan mengikuti tes SMITS dengan mata uji yang sesuai dengan syarat pendaftaran prodi</p>
Seleksi	<ul style="list-style-type: none"> <p>Jalur undangan: Proses ini diawali dengan kampus menerima data pendaftaran siswa dari LTMPT. Data tersebut kemudian diverifikasi, untuk memastikan bahwa data calon mahasiswa yang diinput di LTMPT adalah valid dan akurat. Apabila sudah valid, maka dilanjutkan dengan tahap penilaian dan seleksi prestasi akademik siswa, di mana aspek penilaian pada aktivitas ini adalah</p> <ul style="list-style-type: none"> ❖ Nilai rata-rata rapor semester 1 sampai semester 5, bobot 70% ❖ Nilai mata pelajaran tertentu, bobot 20% ❖ Nilai rapor semester 1 sampai semester 3, bobot 10% <p>Nilai mata pelajaran tertentu yang digunakan dalam seleksi prestasi akademik adalah sebagai berikut:</p> <p>Prodi Saintek</p> <ul style="list-style-type: none"> ❖ Matematika ❖ Fisika ❖ Kimia ❖ Biologi <p>Prodi Soshum</p> <ul style="list-style-type: none"> ❖ Bahasa Indonesia ❖ Bahasa Inggris ❖ Matematika <p>Setelah itu dilakukan juga penilaian terhadap prestasi non akademik dengan memberi pembobotan tambahan untuk masing-masing prestasi yang telah diinput. Setelah penilaian selesai maka dilanjutkan dengan mengurutkan nilai masing-masing mahasiswa sesuai program studinya, mahasiswa teratas dalam rentang jumlah daya tampung program studi akan dinyatakan lulus. Pada akhirnya, daftar nama siswa yang lulus seleksi diumumkan.</p> <p>Jalur Tes: Proses seleksi dalam jalur Tes Bersama Masuk PTN, seperti di ITS, terdiri dari beberapa tahap penting. Tahap pertama adalah verifikasi data dan kelayakan pendaftar oleh LTMPT, yang mencakup pengecekan data pribadi, nilai rapor, dan prestasi akademik. Kelayakan pendaftar harus memenuhi persyaratan umum dan khusus yang ditetapkan oleh institusi tujuan. Selanjutnya, calon mahasiswa akan mengikuti Ujian Tulis Bersama yang diselenggarakan oleh LTMPT. Tes Bersama terdiri dari dua sesi Tes Potensi Skolastik (TPS) dan satu sesi Tes Kompetensi Akademik (TKA). TPS</p>

	<p>digunakan untuk mengukur kemampuan penalaran umum, pengetahuan dan pemahaman umum, serta kemampuan kuantitatif, sementara TKA mengukur kemampuan akademik sesuai dengan program studi yang dipilih oleh calon mahasiswa. Setelah Tes Bersama selesai, ITS akan menetapkan nilai minimal yang harus dicapai oleh calon mahasiswa untuk dapat lolos seleksi. Nilai minimal ini ditentukan berdasarkan pertimbangan kualitas calon mahasiswa yang diinginkan oleh ITS. Selanjutnya, ITS akan melakukan pemeringkatan calon mahasiswa berdasarkan nilai UTBK, dengan tujuan untuk menentukan urutan calon mahasiswa yang akan diterima di ITS. Hasil seleksi kemudian diumumkan melalui laman resmi ITS, dan calon mahasiswa yang dinyatakan lolos seleksi dapat melanjutkan dengan proses daftar ulang di institusi tersebut. Dengan demikian, proses seleksi ini merupakan tahap krusial dalam penentuan penerimaan calon mahasiswa ke PTN, dan mengikuti prosedur yang telah ditentukan menjadi langkah penting bagi mereka yang ingin melanjutkan pendidikan di institusi seperti ITS.</p> <ul style="list-style-type: none"> ● Jalur mandiri: Terdapat 4 jenis seleksi mandiri di ITS, yaitu: <ul style="list-style-type: none"> → Mandiri Prestasi (https://www.its.ac.id/admission/sarjana/sp/) → Mandiri Umum (https://www.its.ac.id/admission/sarjana/sm/) → Mandiri Kemitraan (https://www.its.ac.id/admission/sarjana/sk/) → Mandiri Umum Kedokteran (https://www.its.ac.id/admission/sarjana/sm-kedokteran/) <p>Untuk seleksi mandiri dimulai ketika siswa yang telah mendaftar melakukan tahap lanjutan. Untuk jenis seleksi mandiri umum, kemitraan, dan umum kedokteran siswa diarahkan untuk melakukan tes online TKA SMITS, sementara jenis mandiri prestasi tidak perlu mengikut tes online TKA. Meskipun demikian seleksi jenis prestasi khusus untuk hafalan kitab tetap harus mengikuti verifikasi prestasi.</p> <p>Setelah tes dilakukan, maka pihak ITS akan melakukan penilaian. Hasil penilaian tersebut kemudian akan diumumkan kepada calon mahasiswa. Selanjutnya calon mahasiswa akan melakukan tes kesehatan untuk keperluan beberapa jurusan. Apabila lolos maka mahasiswa dinyatakan diterima di kampus ITS dan diarahkan untuk melakukan daftar ulang.</p>
--	---

Pendaftaran Ulang	Calon mahasiswa yang lulus dari masing-masing jalur seleksi ITS wajib melakukan pengisian data pendaftaran di SIPMABA. Data yang perlu diisi antara lain data diri, pendidikan, dan portofolio. Setelah calon mahasiswa melakukan pengisian data pendaftaran, pihak ITS akan memvalidasi data tersebut. Hasil validasi data akan digunakan untuk menentukan UKT yang harus dibayarkan oleh calon mahasiswa. Mahasiswa yang telah dinyatakan lulus dan telah ditentukan UKT-nya wajib melakukan pembayaran UKT. Pembayaran UKT dapat dilakukan secara online atau offline. Setelah melakukan pembayaran UKT, mahasiswa akan mendapatkan LoA yang berisi NRP, prodi, dan fakultas. LoA ini merupakan tanda bahwa mahasiswa tersebut telah resmi menjadi mahasiswa baru ITS.
-------------------	---

b. Mendeklarasikan Granularitas

Proses Bisnis	Granularitas	Alasan
Penentuan daya tampung	1 baris tabel fakta untuk setiap program studi setiap tahunnya per jalur seleksi	Tabel daya tampung akan memberikan informasi dari daya tampung masing-masing jalur seleksi untuk setiap program studi. Dikarenakan, ITS memiliki perubahan kuota penerimaan mahasiswa baru setiap tahunnya maka tahun juga termasuk granularitas.
Pendaftaran	1 baris tabel fakta untuk setiap pendaftaran calon mahasiswa per jalur masuk dan jurusan yang dipilih pada tahun akademik tersebut	Tabel pendaftaran akan memberikan informasi dari isian form pendaftaran calon mahasiswa pada tahun akademik pendaftaran per jalur masuk dan jurusan yang dipilih.
Seleksi	1 baris tabel fakta untuk proses seleksi pendaftaran calon mahasiswa per jalur masuk dan jurusan yang dipilih pada tahun akademik tersebut	Tabel seleksi akan memberikan informasi dari hasil seleksi calon mahasiswa pada tahun akademik pendaftaran untuk masing-masing jalur masuk dan jurusan yang dipilih.
Pendaftaran Ulang	1 baris tabel fakta untuk setiap calon mahasiswa yang diterima per jalur masuk dan jurusan yang dipilih pada tahun akademik tersebut	Tabel pendaftaran ulang akan memberikan informasi daftar mahasiswa yang berhak melakukan daftar ulang dan status pendaftaran ulangnya.

c. Menentukan Dimensi

Proses Bisnis	Dimensi
Penentuan daya tampung	dimensi_jurusan, dimensi_jalur_masuk, dimensi_tahun_akademik, dimensi_fakultas
Pendaftaran	dimensi_jurusan, dimensi_pendaftar, dimensi_jalur_masuk, dimensi_tahun_akademik, dimensi_fakultas
Seleksi	dimensi_jurusan, dimensi_jalur_masuk, dimensi_tahun_masuk, dimensi_pendaftar, dimensi_fakultas
Daftar ulang	dimensi_jurusan, dimensi_tahun_akademik, dimensi_jalur_masuk, dimensi_pendaftar, dimensi_fakultas

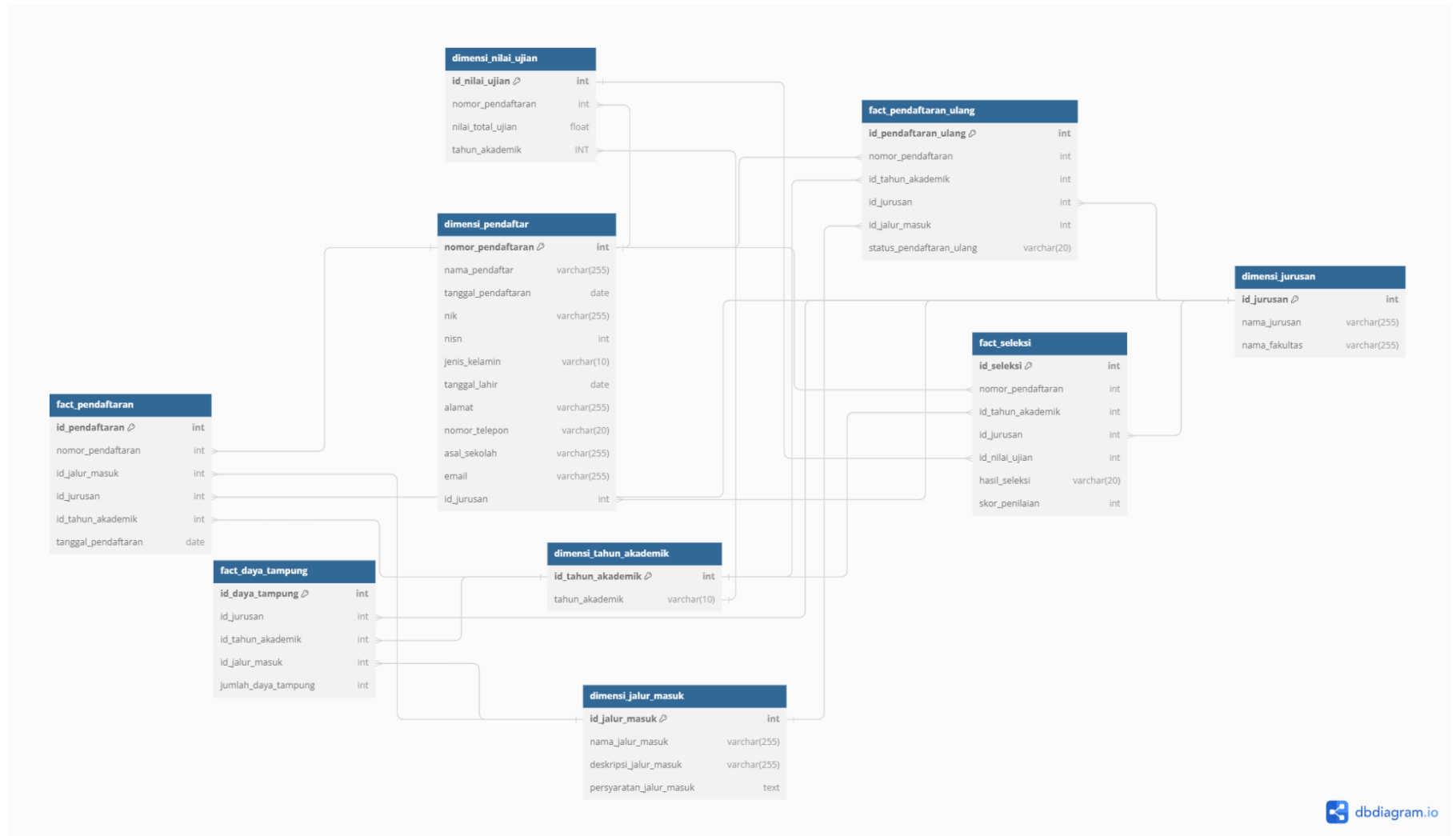
d. Menentukan Fakta

Proses Bisnis	Fakta
Penentuan daya tampung	jumlah_daya_tampung
Pendaftaran	tanggal_pendaftaran
Seleksi	hasil_seleksi, skor_penilaian
Daftar ulang	status_pendaftaran_ulang

Berikut merupakan hasil analisis dengan menggunakan *Bus-Matrix Analysis*:

Business Process	Dimension Tables					
	dimensi_jurusan	dimensi_fakultas	dimensi_jalur_masuk	dimensi_tahun_akademik	dimensi_pendaftar	dimensi_nilai
Penentuan daya tampung	X	X	X	X		
Pendaftaran	X	X	X	X	X	
Seleksi	X	X	X	X	X	X
Daftar ulang	X	X	X	X	X	

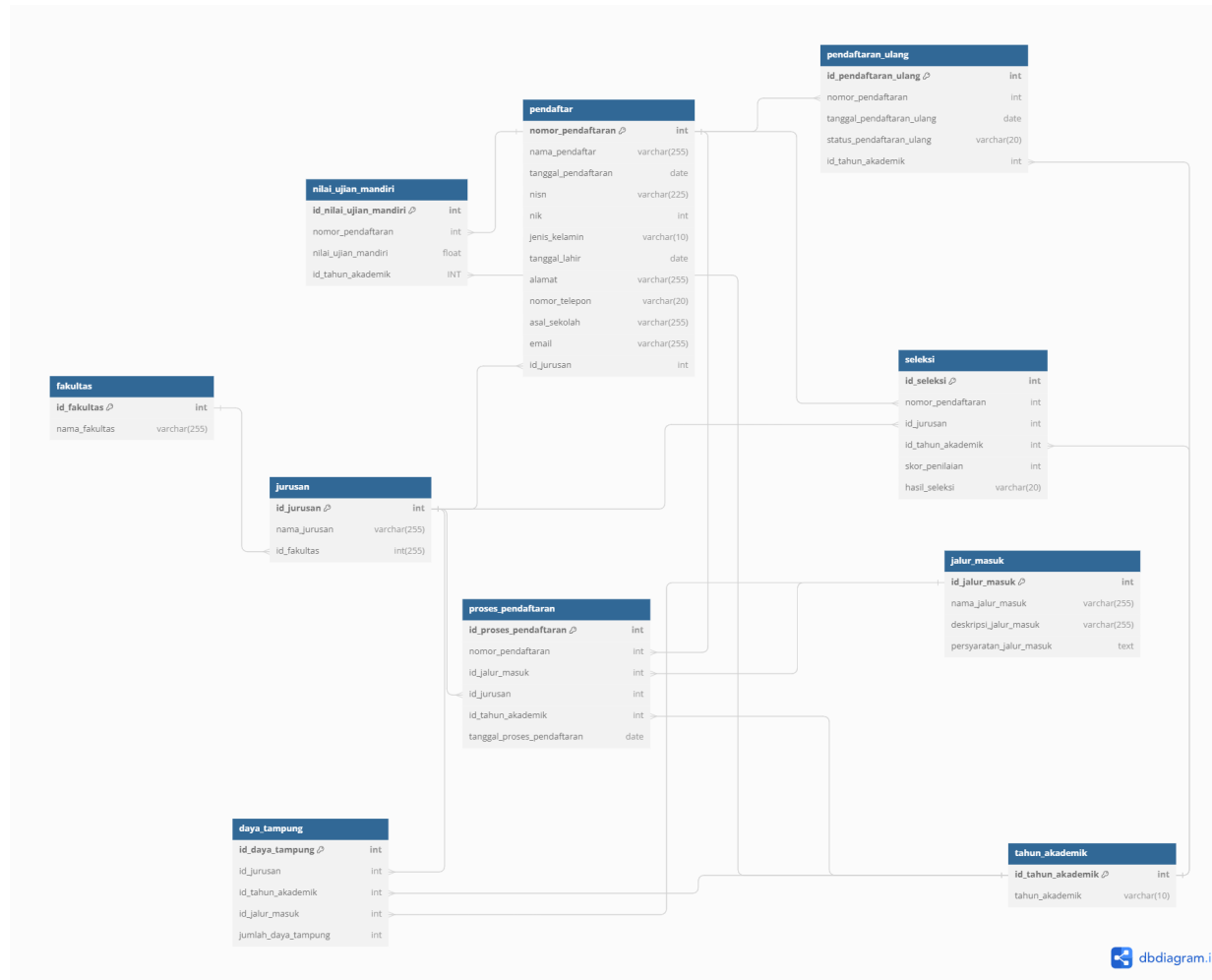
e. Menghasilkan Tabel pada RDBMS



Link OLAP : https://dbdiagram.io/d/OLAP-FP-DLH_revised1-65708bef56d8064ca08306a3

Desain & Implementasi ETL

a. Data & OLTP Profiling



Membuat Tabel OLTP

```
CREATE TABLE [pendaftar] (  
    [nomor_pendaftaran] int PRIMARY KEY,  
    [nama_pendaftar] varchar(255),  
    [tanggal_pendaftaran] date,  
    [nispn] varchar(225),  
    [nik] int,  
    [jenis_kelamin] varchar(10),  
    [tanggal_lahir] date,  
    [alamat] varchar(255),  
    [nomor_telepon] varchar(20),  
    [asal_sekolah] varchar(255),  
    [email] varchar(255)  
)  
GO
```

```
CREATE TABLE [jalur_masuk] (  
    [id_jalur_masuk] int PRIMARY KEY,  
    [nama_jalur_masuk] varchar(255),  
    [deskripsi_jalur_masuk] varchar(255),  
    [persyaratan_jalur_masuk] text  
)  
GO
```

```
CREATE TABLE [nilai_ujian_mandiri] (  
    [id_nilai_ujian_mandiri] int PRIMARY KEY,  
    [nomor_pendaftaran] int,  
    [nilai_ujian_mandiri] float  
)  
GO
```

```
CREATE TABLE [jurusan] (  
    [id_jurusan] int PRIMARY KEY,  
    [nama_jurusan] varchar(255),  
    [id_fakultas] int  
)  
GO
```

```
CREATE TABLE [tahun_akademik] (  
    [id_tahun_akademik] int PRIMARY KEY,  
    [tahun_akademik] varchar(10)  
)  
GO
```

```
CREATE TABLE [proses_pendaftaran] (  
    [id_proses_pendaftaran] int PRIMARY KEY,  
    [nomor_pendaftaran] int,  
    [id_jalur_masuk] int,  
    [id_jurusan] int,  
    [id_tahun_akademik] int,  
    [tanggal_proses_pendaftaran] date  
)  
GO
```

```
CREATE TABLE [seleksi] (  
    [id_seleksi] int PRIMARY KEY,  
    [nomor_pendaftaran] int,  
    [id_jurusan] int,  
    [id_tahun_akademik] int,  
    [skor_penilaian] int,  
    [hasil_seleksi] varchar(20)
```

```
)  
GO
```

```
CREATE TABLE [pendaftaran_ulang] (  
    [id_pendaftaran_ulang] int PRIMARY KEY,  
    [nomor_pendaftaran] int,  
    [tanggal_pendaftaran_ulang] date,  
    [status_pendaftaran_ulang] varchar(20),  
    [id_tahun_akademik] int  
)  
GO
```

```
CREATE TABLE [daya_tampung] (  
    [id_daya_tampung] int PRIMARY KEY,  
    [id_jurusan] int,  
    [id_tahun_akademik] int,  
    [id_jalur_masuk] int,  
    [jumlah_daya_tampung] int  
)  
GO
```

```
CREATE TABLE [fakultas] (  
    [id_fakultas] int PRIMARY KEY,  
    [nama_fakultas] varchar(255)  
)  
GO
```

```
ALTER TABLE [nilai_ujian_mandiri] ADD FOREIGN KEY ([nomor_pendaftaran]) REFERENCES [pendaftar] ([nomor_pendaftaran])  
GO
```

```
ALTER TABLE [jurusan] ADD FOREIGN KEY ([id_fakultas]) REFERENCES [fakultas] ([id_fakultas])
```


GO

```
ALTER TABLE [proses_pendaftaran] ADD FOREIGN KEY ([nomor_pendaftaran]) REFERENCES [pendaftar] ([nomor_pendaftaran])
GO
```

```
ALTER TABLE [proses_pendaftaran] ADD FOREIGN KEY ([id_jalur_masuk]) REFERENCES [jalur_masuk] ([id_jalur_masuk])
GO
```

```
ALTER TABLE [proses_pendaftaran] ADD FOREIGN KEY ([id_jurusan]) REFERENCES [jurusan] ([id_jurusan])
GO
```

```
ALTER TABLE [proses_pendaftaran] ADD FOREIGN KEY ([id_tahun_akademik]) REFERENCES [tahun_akademik] ([id_tahun_akademik])
GO
```

```
ALTER TABLE [seleksi] ADD FOREIGN KEY ([nomor_pendaftaran]) REFERENCES [pendaftar] ([nomor_pendaftaran])
GO
```

```
ALTER TABLE [seleksi] ADD FOREIGN KEY ([id_jurusan]) REFERENCES [jurusan] ([id_jurusan])
GO
```

```
ALTER TABLE [seleksi] ADD FOREIGN KEY ([id_tahun_akademik]) REFERENCES [tahun_akademik] ([id_tahun_akademik])
GO
```

```
ALTER TABLE [pendaftaran_ulang] ADD FOREIGN KEY ([nomor_pendaftaran]) REFERENCES [pendaftar] ([nomor_pendaftaran])
GO
```

```
ALTER TABLE [pendaftaran_ulang] ADD FOREIGN KEY ([id_tahun_akademik]) REFERENCES [tahun_akademik] ([id_tahun_akademik])
GO
```

```
ALTER TABLE [daya_tampung] ADD FOREIGN KEY ([id_jurusan]) REFERENCES [jurusan] ([id_jurusan])
GO
```

```
ALTER TABLE [daya_tampung] ADD FOREIGN KEY ([id_tahun_akademik]) REFERENCES [tahun_akademik] ([id_tahun_akademik])
GO
```

```
ALTER TABLE [daya_tampung] ADD FOREIGN KEY ([id_jalur_masuk]) REFERENCES [jalur_masuk] ([id_jalur_masuk])
GO
```

```
ALTER TABLE [pendaftar] ADD FOREIGN KEY ([email]) REFERENCES [pendaftar] ([jenis_kelamin])
GO
```

Selanjutnya dilakukan import data mockup ke dalam SSMS database OLTPpowerdlh

Import daya_tampung

```
INSERT INTO [daya_tampung] VALUES (8000001,1000000021,1001,42101,12),
(8000002,1000000022,1001,42101,12),
(8000003,1000000023,1001,42101,12),
(8000004,1000000024,1001,42101,12),
...
(8000412,1000000064,1003,42303,28),
(8000413,1000000065,1003,42303,28),
(8000414,1000000066,1003,42303,10);
```

Import tahun_akademik

```
INSERT INTO tahun_akademik (id_tahun_akademik, tahun_akademik) VALUES
('1001', '2021'),
('1002', '2022'),
```

('1003', '2023');

Import jurusan

```
INSERT INTO jurusan VALUES
    (1000000021,'fisika',5600001),
    (1000000022,'matematika',5600001),
    (1000000023,'statistika',5600001),
    ...
    (1000000064,'statistika bisnis',5600007),
    (1000000065,'teknologi kedokteran',5600008),
    (1000000066,'kedokteran',5600008);
```

Import jalur_masuk

```
INSERT INTO jalur_masuk VALUES
    (42101,N'SNMPTN 2021',N'Seleksi Nasional Masuk Perguruan Tinggi Negeri 2021',N'1. Warga Negara Indonesia (WNI) yang memiliki Nomor Induk Kependudukan (NIK). 2. Merupakan siswa SMA/MA/SMK kelas terakhir (kelas XII) pada tahun 2021 yang memiliki prestasi unggul. 3. Memiliki akun LTMPPT dan NISN yang terdaftar di PDSS. 4. Memiliki nilai rapor semester 1 sampai dengan 5 yang telah diisikan di PDSS. 5. Wajib mengunggah PORTOFOLIO bagi yang memilih program studi Bidang Seni & Olahraga. 6. Memiliki prestasi akademik dan memenuhi persyaratan yang ditentukan oleh masing-masing PTN Akademik, PTN Vokasi, atau PTKIN.'),
    ...
    (42303,N'Mandiri 2023',N'Seleksi Mandiri ITS 2023',N'Seleksi penerimaan calon mahasiswa baru melalui jalur Seleksi Mandiri Umum dapat diikuti oleh pendaftar dari SMA/MA/SMK sejenis lulusan tahun 2021, 2022 dan 2023. Seleksi Mandiri ITS akan didasarkan pada nilai tes TKA SMITS tahun 2023.');
```

Import pendaftaran ulang

```
INSERT INTO [pendaftaran_ulang] VALUES (11200001,42160277,1001,N'Sudah',N'3/17/2021'),
    (11200002,42160507,1001,N'Sudah',N'3/18/2021'),
```

```
(11200003,42160093,1001,N'Sudah',N'3/18/2021'),
(11200004,42160231,1001,N'Sudah',N'3/20/2021'),
...
(11200447,42160455,1001,N'Sudah',N'3/28/2021'),
(11200448,42160041,1001,N'Sudah',N'3/22/2021'),
(11200449,42160547,1001,N'Sudah',N'3/15/2021');
```

Import nilai_ujian_mandiri

```
INSERT INTO `nilai_ujian_mandiri` VALUES (32000001,42161651,'72.13',1001),
(32000002,42161652,'82.64',1001),
(32000003,42161653,'57.28',1001),
(32000004,42161654,'49.27',1001),
...
(32004498,42363332,'31.41',1003),
(32004499,42363333,'71.92',1003),
(32004500,42363334,'50.52',1003);
```

1. OLTP Profiling pendaftar

Nama Tabel: pendaftar					
Total baris: 4500			Total kolom: 11		
Primary Key: 1			Foreign Key: 1	Non Key: 10	
Kolom	Tipe Data	PK/FK	Null Allowed	Min Value	Max Value

nomor_pendaftaran	int	PK	No	-	-
nama_pendaftar	varchar(50)	-	No	-	-
tanggal_pendaftaran	date	-	No	5/6/2021	6/30/2023
jenis_kelamin	int	-	No	1	2
tanggal_lahir	date	-	No	2/1/2001	9/28/2005
alamat	varchar(50)	-	No	-	-
nomor_telepon	varchar(50)	-	No	-	-
asal_sekolah	varchar(50)	-	No	-	-
email	varchar(50)	-	No	-	-
nik	varchar(50)	-	No	-	-
nisan	varchar(50)	-	No	-	-
id_jurusan	int	FK	Yes	1000000021	1000000066

2. OLTP Profiling proses_pendaftaran

Nama Tabel: proses_pendaftaran					
Total baris: 4500			Total kolom: 6		
Primary Key: 1			Foreign Key: 4	Non Key: 1	
Kolom	Tipe Data	PK/FK	Null Allowed	Min Value	Max Value

id_proses_pendaftaran	int	PK	No	-	-
nomor_pendaftaran	int	FK	No	-	-
id_jalur_masuk	int	FK	No	-	-
id_jurusan	int	FK	No	-	-
id_tahun_akademik	int	FK	No	-	-
tanggal_proses_pendaftaran	date	-	No	6/5/2021	7/9/2023

3. OLTP Profiling seleksi

Nama Tabel: seleksi					
Total baris: 4500			Total kolom: 6		
Primary Key: 1			Foreign Key: 4	Non Key: 1	
Kolom	Tipe Data	FK	Null Allowed	Min Value	Max Value
id_seleksi	int	PK	No	-	-
id_jurusan	int	FK	No	-	-
id_tahun_akademik	int	FK	No	-	-
nomor_pendaftaran	int	FK	No	-	-
skor_penilaian	int	FK	No		

hasil_seleksi	varchar(20)	-	No	-	-
---------------	-------------	---	----	---	---

4. OLTP Profiling pendaftaran_ulang

Nama Tabel: pendaftaran_ulang					
Total baris: 8440			Total kolom: 5		
Primary Key: 1			Foreign Key:2	Non Key: 2	
Kolom	Tipe Data	FK	Null Allowed	Min Value	Max Value
id_pendaftaran_ulang	int	PK	No	-	-
nomor_pendaftaran	int	FK	No	-	-
id_tahun_akademik	date	FK	No	-	-
tanggal_pendaftaran_ulang	varchar(50)	-	Yes	3/3/2023	7/22/2023
status_pendaftaran_ulang	int	-	No	-	-

5. OLTP Profiling fakultas

Nama Tabel: fakultas	
Total baris: 8	Total kolom: 2

Primary Key: 1			Foreign Key:0	Non Key: 1	
Kolom	Tipe Data	FK	Null Allowed	Min Value	Max Value
id_fakultas	int	PK	No	-	-
nama_fakultas	varchar(255)	-	No	-	-

6. OLTP Profiling jalur_masuk

Nama Tabel: jalur_masuk					
Total baris: 9			Total kolom: 4		
Primary Key: 1			Foreign Key: 0	Non Key: 3	
Kolom	Tipe Data	FK	Null Allowed	Min Value	Max Value
id_jalur_masuk	int	PK	No	-	-
nama_jalur_masuk	varchar(255)	-	No	-	-
deskripsi_jalur_masuk	varchar(255)	-	No	-	-
persyaratan_jalur_masuk	text	-	No	-	-

7. OLTP Profiling nilai_ujian_mandiri

Nama Tabel: nilai_ujian_mandiri					
Total baris: 4500			Total kolom: 3		
Primary Key: 1			Foreign Key: 1	Non Key: 1	
Kolom	Tipe Data	PK/FK	Null Allowed	Min Value	Max Value
id_nilai_ujian_mandiri	int	PK	No	-	-
nomor_pendaftaran	int	FK	No	-	-
nilai_ujian_mandiri	float	-	No	-	-

8. OLTP Profiling tahun_akademik

Nama Tabel: tahun_akademik					
Total baris: 3			Total kolom: 2		
Primary Key: 1			Foreign Key: 0	Non Key: 1	
Kolom	Tipe Data	FK	Null Allowed	Min Value	Max Value
id_tahun_akademik	int	PK	No	-	-
tahun_akademik	int	-	No	2021	2023

9. OLTP Profiling daya_tampung

Nama Tabel: daya_tampung					
Total baris: 414			Total kolom: 5		
Primary Key: 1			Foreign Key: 3	Non Key: 1	
Kolom	Tipe Data	PK/FK	Null Allowed	Min Value	Max Value
id_daya_tampung	int	PK	No	-	-
id_jurusan	int	FK	No	-	-
id_tahun_akademik	int	FK	No	-	-
id_jalur_masuk	int	FK	No	-	-
jumlah_daya_tampung	int	-	No	10	30

10. OLTP Profiling jurusan

Nama Tabel: jurusan

Total baris: 46			Total kolom: 3		
Primary Key: 1			Foreign Key: 1	Non Key: 1	
Kolom	Tipe Data	FK	Null Allowed	Min Value	Max Value
id_jurusan	int	PK	No	-	-
nama_jurusan	varchar(255)	-	No	-	-
id_fakultas	int	FK	No	-	-

b. High Level ETL Design

Level table ke table , antar OLTP ke Star Schema yang kalian sudah desain

Proses Bisnis	Source (OLTP)	Target (OLAP)
Penentuan daya tampung	<ul style="list-style-type: none"> - daya_tampung.id_daya_tampung - daya_tampung.id_jurusan - daya_tampung.id_tahun_akademik - daya_tampung.id_jalur_masuk - daya_tampung.jumlah_daya_tampung 	fact_daya_tampung
	<ul style="list-style-type: none"> - jurusan.id_jurusan - jurusan.nama_jurusan - fakultas.id_fakultas 	dimensi_jurusan
	<ul style="list-style-type: none"> - tahun_akademik.id_tahun_akademik - tahun_akademik.tahun_akademik 	dimensi_tahun_akademik
	<ul style="list-style-type: none"> - jalur_masuk.id_jalur_masuk - jalur_masuk.nama_jalur_masuk - jalur_masuk.deskripsi_jalur_masuk 	dimensi_jalur_masuk

	<ul style="list-style-type: none"> - jalur_masuk.persyaratan_jalur_masuk 	
Pendaftaran	<ul style="list-style-type: none"> - proses_pendaftaran.id_proses_pendaftaran - proses_pendaftaran.nomor_pendaftaran - proses_pendaftaran.id_jalur_masuk - proses_pendaftaran.id_jurusan - proses_pendaftaran.id_tahun_akademik - proses_pendaftaran.tanggal_proses_pendaftaran 	fact_pendaftaran
	<ul style="list-style-type: none"> - jurusan.id_jurusan - jurusan.nama_jurusan - fakultas.id_fakultas 	dimensi_jurusan
	<ul style="list-style-type: none"> - tahun_akademik.id_tahun_akademik - tahun_akademik.tahun_akademik 	dimensi_tahun_akademik
	<ul style="list-style-type: none"> - jalur_masuk.id_jalur_masuk - jalur_masuk.nama_jalur_masuk - jalur_masuk.deskripsi_jalur_masuk - jalur_masuk.persyaratan_jalur_masuk 	dimensi_jalur_masuk
	<ul style="list-style-type: none"> - pendaftar.nomor_pendaftaran - pendaftar.nama_pendaftar - pendaftar.tanggal_pendaftaran - pendaftar.nisn - pendaftar.nik - pendaftar.jenis_kelamin - pendaftar.tanggal_lahir - pendaftar.alamat - pendaftar.nomor_telepon - pendaftar.asal_sekolah - pendaftar.email 	dimensi_pendaftar
Seleksi	<ul style="list-style-type: none"> - seleksi.id_seleksi - seleksi.id_proses_pendaftaran 	fact_seleksi

	<ul style="list-style-type: none"> - seleksi.skor_penilaian - pendaftar.nomor_pendaftaran - jurusan.id_jurusan - tahun_akademik.id_tahun_akademik - jalur_masuk.id_jalur_masuk 	
	<ul style="list-style-type: none"> - jurusan.id_jurusan - jurusan.nama_jurusan - fakultas.id_fakultas 	dimensi_jurusan
	<ul style="list-style-type: none"> - tahun_akademik.id_tahun_akademik - tahun_akademik.tahun_akademik 	dimensi_tahun_akademik
	<ul style="list-style-type: none"> - jalur_masuk.id_jalur_masuk - jalur_masuk.nama_jalur_masuk - jalur_masuk.deskripsi_jalur_masuk - jalur_masuk.persyaratan_jalur_masuk 	dimensi_jalur_masuk
	<ul style="list-style-type: none"> - pendaftar.nomor_pendaftaran - pendaftar.nama_pendaftar - pendaftar.tanggal_pendaftaran - pendaftar.nisn - pendaftar.nik - pendaftar.jenis_kelamin - pendaftar.tanggal_lahir - pendaftar.alamat - pendaftar.nomor_telepon - pendaftar.asal_sekolah - pendaftar.email 	dimensi_pendaftar
Daftar ulang	<ul style="list-style-type: none"> - pendaftaran_ulang.id_pendaftaran_ulang - pendaftaran_ulang.id_seleksi - pendaftaran_ulang.tanggal_pendaftaran_ulang - pendaftaran_ulang.status_pendaftaran_ulang - pendaftar.nomor_pendaftaran 	fact_pendaftaran_ulang

	<ul style="list-style-type: none"> - jurusan.id_jurusan - tahun_akademik.id_tahun_akademik - jalur_masuk.id_jalur_masuk 	
	<ul style="list-style-type: none"> - jurusan.id_jurusan - jurusan.nama_jurusan - fakultas.id_fakultas 	dimensi_jurusan
	<ul style="list-style-type: none"> - tahun_akademik.id_tahun_akademik - tahun_akademik.tahun_akademik 	dimensi_tahun_akademik
	<ul style="list-style-type: none"> - jalur_masuk.id_jalur_masuk - jalur_masuk.nama_jalur_masuk - jalur_masuk.deskripsi_jalur_masuk - jalur_masuk.persyaratan_jalur_masuk 	dimensi_jalur_masuk
	<ul style="list-style-type: none"> - pendaftar.nomor_pendaftaran - pendaftar.nama_pendaftar - pendaftar.tanggal_pendaftaran - pendaftar.nisn - pendaftar.nik - pendaftar.jenis_kelamin - pendaftar.tanggal_lahir - pendaftar.alamat - pendaftar.nomor_telepon - pendaftar.asal_sekolah - pendaftar.email 	dimensi_pendaftar

c. Source to Target Mapping

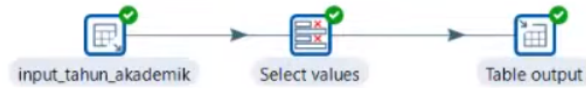
- Level attribute per table tujuan

Source Table	Level Atribut	Target Table
dimensi_jurusan	id_jurusan	fact_daya_tampung
		fact_pendaftaran
		fact_seleksi
		fact_pendaftaran_ulang
dimensi_jalur_masuk	id_jalur_masuk	fact_daya_tampung

		fact_pendaftaran
		fact_seleksi
		fact_pendaftaran_ulang
dimensi_tahun_akademik	id_tahun_akademik	fact_daya_tampung
		fact_pendaftaran
		fact_seleksi
		fact_pendaftaran_ulang
dimensi_pendaftar	nomor_pendaftaran	fact_pendaftaran
		fact_seleksi
		fact_pendaftaran_ulang
dimensi_fakultas	id_fakultas	dimensi_jurusan
dimensi_nilai_ujian	id_nilai_ujian	dimensi_pendaftar

d. ETL one time historical dan ETL incremental (Teknologi : Pentaho)

A. Dimensi tahun akademik



1. Menginput Tabel Tahun Akademik (Source: OLTPpowerdlh)

Table input

Step name: input_tahun_akademik

Connection: OLTPpowerdlh

SQL

```
SELECT
  id_tahun_akademik
, tahun_akademik
FROM dbo.tahun_akademik
```

Line 5 Column 0

Store column info in step meta ☐

Enable lazy conversion ☐

Replace variables in script? ☐

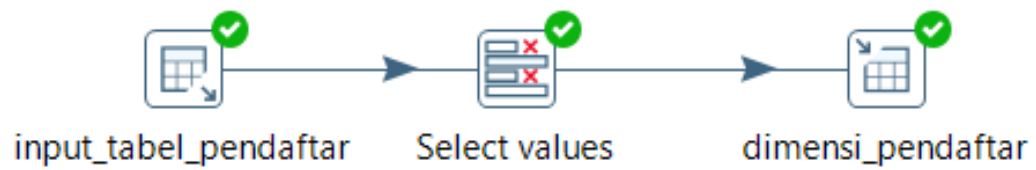
Insert data from step

Execute for each row? ☐

Limit size: 0

Help OK Preview Cancel

2. Memilih kolom yang akan diinput, yaitu kolom id_tahun_akademik & tahun_akademik



1. Menginput Tabel Tahun Pendaftar (Source: OLTPpowerdlh)

Table input

Step name: Table input_pendaftar

Connection: OLTPpowerdlh

SQL

```
SELECT
, nomor_pendaftaran
, nama_pendaftar
, tanggal_pendaftaran
, jenis_kelamin
, tanggal_lahir
, alamat
, nomor_telepon
, asal_sekolah
, email
, nik
, nisan
, id_jurusan
FROM dbo.pendaftar
```

Line 12 Column 6

Store column info in step meta data ☐

Enable lazy conversion ☐

Replace variables in script? ☐

Insert data from step

Execute for each row? ☐

Limit size 0

Help OK Preview Cancel

2. Memilih kolom yang akan diinput

Table output

Step name: Table output

Connection: OLTPpowerdlh

Target schema: dbo

Target table: dimensi_pendaftar

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☐

Main options / Database fields

Partition data over tables: ☐

Partitioning field:

Partition data per month: ☐

Partition data per day: ☐

Use batch update for inserts: ☒

Is the name of the table defined in a field?: ☐

Field that contains name of table:

Store the tablename field: ☒

Return auto-generated key: ☐

Name of auto-generated key field:

Help OK Cancel SQL

C. Dimensi jalur masuk

1. Menginput Tabel Jalur Masuk (Source: OLTPpowerdlh)

Table input

Step name: input_tabel_jalur_masuk

Connection: OLTPpowerdlh

Get SQL select statement...

SQL

```
SELECT
id_jalur_masuk
, nama_jalur_masuk
, deskripsi_jalur_masuk
, persyaratan_jalur_masuk
FROM dbo.jalur_masuk
```

Line 7 Column 0

Store column info in step meta: ☐

Enable lazy conversion: ☐

Replace variables in script?: ☐

Insert data from step:

Execute for each row?: ☐

Limit size: 0

Help OK Preview Cancel

2. Memilih kolom yang akan diinput

[illegible]

- Memasukkan hasil transformasi ke dalam table output dimensi_jalur_masuk (Target: OLAPpowerdlh)

Table output

Step name: dimensi_jalur_masuk

Connection: OLAPpowerdlh

Target schema: dbo

Target table: dimensi_jalur_masuk

Commit size: 1000

Truncate table: ☒

Ignore insert errors: ☐

Specify database fields: ☒

Main options Database fields

Fields to insert:

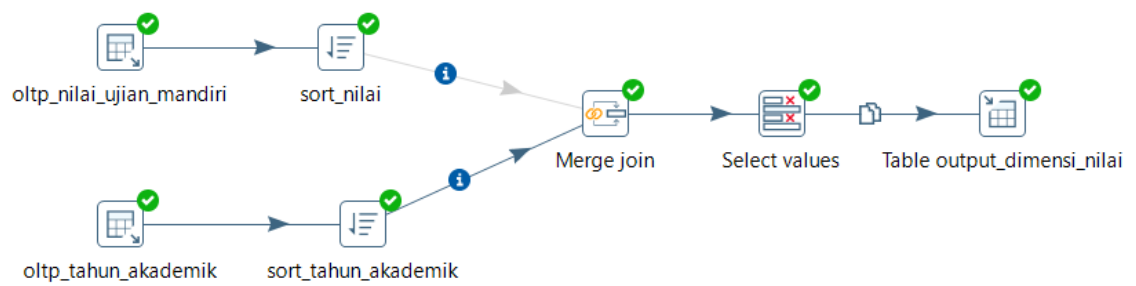
#	Table field	Stream field	
1	id_jalur_mas...	id_jalur_masuk	
2	nama_jalur_...	nama_jalur_m...	
3	deskripsi_jal...	deskripsi_jalur...	
4	persyaratan...	persyaratan_ja...	

Get fields

Enter field mapping

Help OK Cancel SQL

D. Dimensi nilai



1. Input nilai ujian mandiri dari tabel nilai_ujian_mandiri yang ada pada database OLTPpowerdlh dan tahun akademik dari tabel tahun_akademik yang ada pada database OLTPpowerdlh

Table input

Step name

oltp_nilai_ujian_mandiri

Connection

OLTPpowerdlh

Edit...

New...

Wizard...

SQL

Get SQL select statement...

```

SELECT
  id_nilai_ujian_mandiri
, nomor_pendaftaran
, nilai_ujian_mandiri
, id_tahun_akademik
FROM dbo.nilai_ujian_mandiri

```

Line 7 Column 0

Store column info in step meta

Enable lazy conversion

Replace variables in script?

Insert data from step

Execute for each row?

Limit size

Help

OK

Preview

Cancel

Sort rows

Step name:

Sort directory:

TMP-file prefix:

Sort size (rows in memory):

Free memory threshold (in %):

Compress TMP Files? ☐

Only pass unique rows? (verifies keys only) ☐

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted
1	id_tahun_akademik	Y	N	N	0	N

3. Merge join hasil dari yang sudah disort kemudian pilih keys nya berdasarkan id_tahun_akademik

Merge join

Step name:

First Step:

Second Step:

Join Type:

Keys for 1st step:

#	Key field
1	id_tahun_...

Keys for 2nd step:

#	Key field
1	id_tahun_...

4. Select values yang akan dimasukkan ke dalam tabel output

Select values

Step name:

Select & Alter

Fields:

#	Fieldname	Rename to	Length	Precision
1	id_nilai_ujian_mandiri			
2	nomor_pendaftaran			
3	nilai_ujian_mandiri			
4	tahun_akademik			

Include unspecified fields, ordered by ☐

5. Out put nya nanti berupa tabel dimensi_nilai yang ada pada database OLAPpowerdlh

Table output

Step name: Table output_dimensi_nilai

Connection: OLAPpowerdlh

Target schema: dbo

Target table: dimensi_nilai

Commit size: 1000

Truncate table: ☒

Ignore insert errors: ☐

Specify database fields: ☒

Main options | Database fields

Fields to insert:

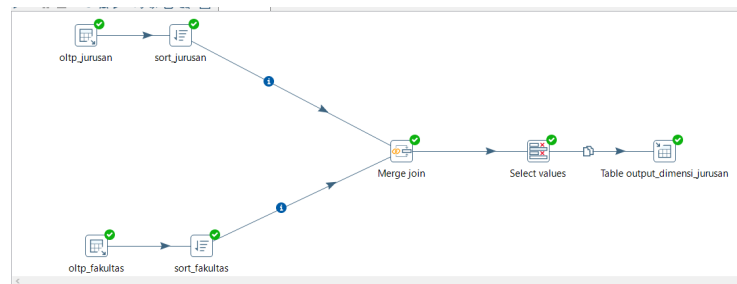
#	Table field	Stream field	Field name
1	id_nilai_ujian...	id_nilai_ujian...	
2	nomor_pen...	nomor_penda...	
3	nilai_ujian...	nilai_ujian_ma...	
4	tahun_akad...	tahun_akade...	

Get fields

Enter field mapping

OK Cancel SQL

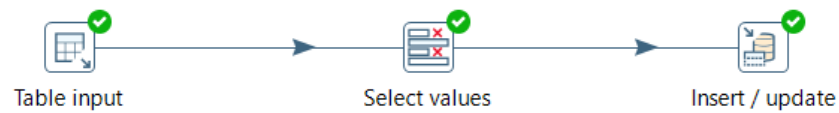
E. Dimensi jurusan



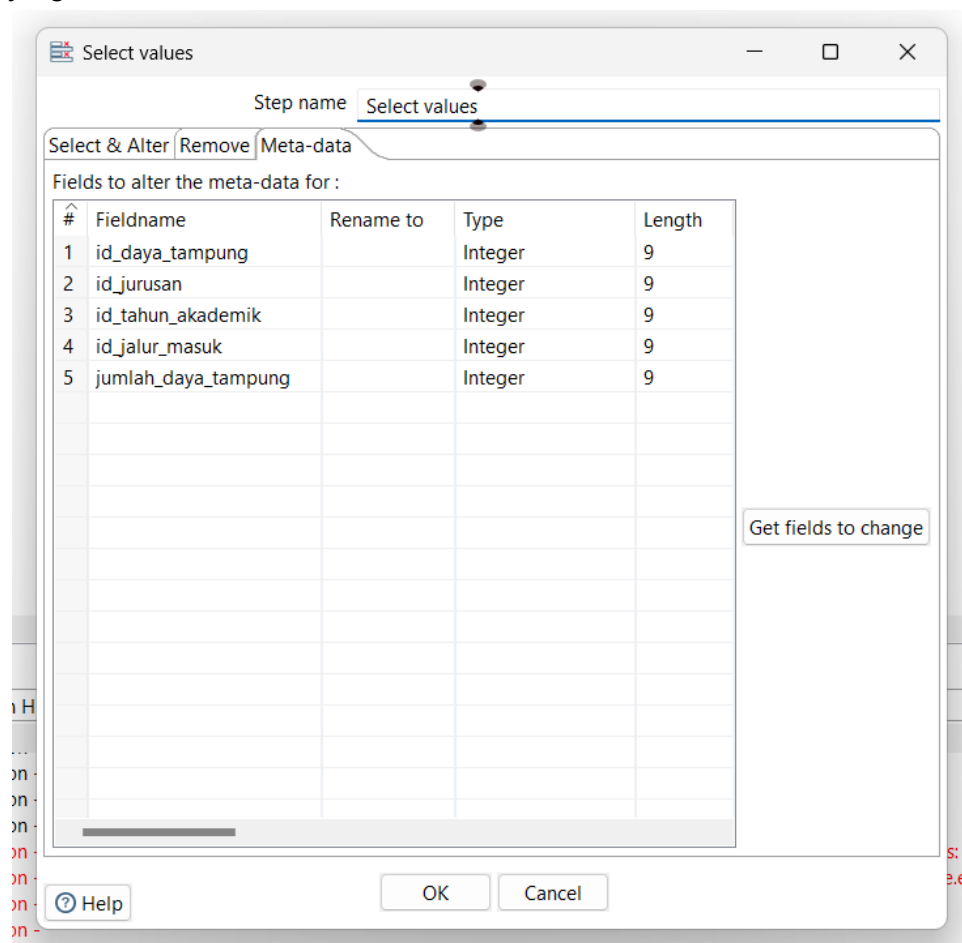
1. Input nilai ujian mandiri dari tabel jurusan yang ada pada database OLTPpowerdlh dan tahun akademik dari tabel fakultas yang ada pada database OLTPpowerdlh
2. Sort kedua tabel berdasarkan id_fakultas
3. Merge join hasil dari yang sudah disort kemudian pilih keys nya berdasarkan id_fakultas
4. Select values yang akan dimasukkan ke dalam tabel output
5. Out put nya nanti berupa tabel dimensi_jurusan yang ada pada database OLAPpowerdlh

F. Fact Daya Tampung

1. Masukan tabel dari OLTPpowerdlh yang akan diolah menjadi fact_daya_tampung di OLAP



2. Select value dari tabel yang diinput sebelum nya. Pilih sesuai dengan star schema yang akan dibuat



3. Untuk hasil output nya akan menggunakan insert/update agar ketika ada modify suatu saat nanti bisa tersimpan dengan otomatis

Insert / update

Step name: Insert / update

Connection: OLAPpowerdlh [Edit... New... Wizard...]

Target schema: dbo [Browse...]

Target table: fact_daya_tampung [Browse...]

Commit size: 100

Don't perform any updates: ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	id_daya_tampung	=	id_daya_tampung	

[Get fields]

Update fields:

#	Table field	Stream field	Update
1	id_daya_tampung	id_daya_tampung	N
2	id_jurusan	id_jurusan	N
3	id_tahun_akademik	id_tahun_akademik	N
4	id_jalur_masuk	id_jalur_masuk	N
5	jumlah_daya_tampung	jumlah_daya_tampung	Y

[Get update fields] [Edit mapping]

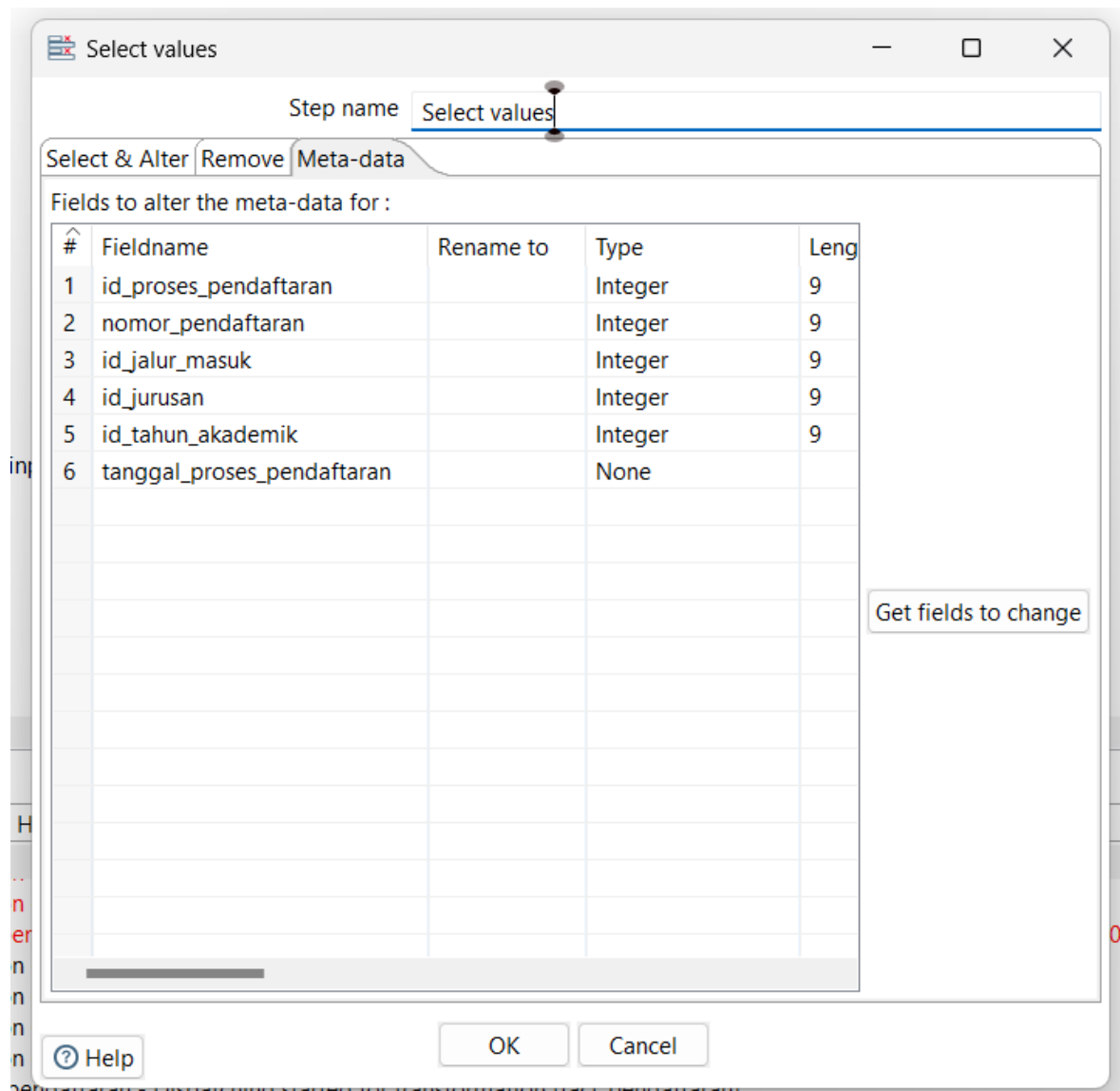
[Help] [OK] [Cancel] [SQL]

G. Fact Fact Pendaftaran

- Masukan tabel dari OLTPpowerdlh yang akan diolah menjadi fact_pendaftaran di OLAP



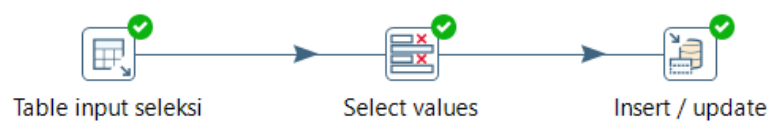
- Select value dari tabel yang diinput sebelum nya. Pilih sesuai dengan star schema yang akan dibuat



- Untuk hasil output nya akan menggunakan insert/update agar ketika ada modify suatu saat nanti bisa tersimpan dengan otomatis

H. Fact Seleksi

- Masukan tabel dari OLTPpowerdlh yang akan diolah menjadi fact_ pendaftaran di OLAP



2. Select value dari tabel yang diinput sebelum nya. Pilih sesuai dengan star schema yang akan dibuat

Step name: Select values

Select & Alter Remove Meta-data

Fields to alter the meta-data for :

#	Fieldname	Rename to	Type	Length	Pr
1	id_seleksi		Integer	9	0
2	nomor_pendaftaran		Integer	9	0
3	id_jurusan		Integer	9	0
4	id_tahun_akademik		Integer	9	0
5	skor_penilaian		Number	9	0
6	hasil_seleksi		String	20	

Get fields to change

Help OK Cancel

3. Untuk hasil output nya akan menggunakan insert/update agar ketika ada modify suatu saat nanti bisa tersimpan dengan otomatis

Insert / update

Step name: Insert / update

Connection: OLAPpowerdlh [Edit...] [New...] [Wizard...]

Target schema: dbo [Browse...]

Target table: fact_seleksi [Browse...]

Commit size: 100

Don't perform any updates: ☒

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	id_seleksi	=	id_seleksi	

[Get fields]

Update fields:

#	Table field	Stream field	Update
1	id_seleksi	id_seleksi	N
2	nomor_pendaftaran	nomor_pendaftaran	N
3	id_jurusan	id_jurusan	N
4	id_tahun_akademik	id_tahun_akademik	N
5	skor_penilaian	skor_penilaian	Y
6	hasil_seleksi	hasil_seleksi	Y

[Get update fields] [Edit mapping]

[Help] [OK] [Cancel] [SQL]

Contoh pemanfaatan dari Data Warehouse

a. Menjawab pertanyaan Who, Where, When, What

1. Who, Where, When, What:

a. Pertanyaan "Who":

- Query:

```
SELECT FROM dimensi_jurusan
WHERE nama_jurusan = 'Teknik Mesin' AND id_jalur_masuk = 42202 AND
id_tahun_akademik = 2022;
```

- Tujuan:

Mengetahui profil mahasiswa yang mendaftar di jurusan Teknik Mesin melalui jalur SBMPTN pada tahun 2022. Informasi ini berguna untuk menilai minat dan preferensi calon mahasiswa.

b. Pertanyaan "Where":

- Query:

```
SELECT nama_fakultas FROM dimensi_fakultas
WHERE id_jurusan IN (SELECT id_jurusan FROM dimensi_jurusan WHERE
nama_jurusan = 'Fisika')
AND id_tahun_akademik = 2023 AND hasil_seleksi = 'Lolos';
```

- Tujuan:

Mengetahui fakultas-fakultas asal mahasiswa yang lulus dalam jurusan Fisika pada tahun akademik 2023. Informasi ini dapat membantu dalam pengembangan program-program akademik.

c. Pertanyaan "When":

- Query:

```
SELECT COUNT() AS jumlah_pendaftar FROM dimensi_tahun_akademik
WHERE id_tahun_akademik = 2023;
```

- Tujuan:

Mengetahui jumlah pendaftar pada tahun akademik 2023. Informasi ini membantu dalam merencanakan kapasitas penerimaan dan sumber daya yang dibutuhkan.

d. Pertanyaan "What":

- Query:

```
SELECT nomor_pendaftaran, nilai_total_ujian FROM dimensi_nilai_ujian
WHERE id_jurusan = 1000000021 AND id_jalur_masuk = 42101;
```

- Tujuan:

Mengetahui nilai total ujian dari mahasiswa yang mendaftar di jurusan Fisika melalui jalur SNMPTN. Informasi ini membantu dalam mengevaluasi tingkat kualifikasi calon mahasiswa.

Dengan menjawab pertanyaan-pertanyaan tersebut, informasi yang mendalam dapat ditemukan dari Data Warehouse, mendukung pengambilan keputusan yang lebih baik, serta membantu ITS dalam menyusun strategi penerimaan mahasiswa baru yang efektif dan efisien. Pentingnya pemanfaatan data dalam konteks ini sangat memberikan dampak positif terhadap proses penerimaan mahasiswa baru di ITS.

b. Menjawab pertanyaan Why

Query untuk Menjawab "Why":

```
SELECT jalur_masuk.nama_jalur_masuk, COUNT() AS jumlah_pendaftar
FROM fakt_table
```



```

JOIN      dimensi_jalur_masuk      ON      fakt_table.id_jalur_masuk      =
dimensi_jalur_masuk.id_jalur_masuk
JOIN      dimensi_tahun_akademik    ON      fakt_table.id_tahun_akademik    =
dimensi_tahun_akademik.id_tahun_akademik
WHERE      fakt_table.id_jurusan      =      'FISIKA'      AND
dimensi_tahun_akademik.tahun_akademik = '2022'
GROUP BY jalur_masuk.nama_jalur_masuk;

```

Tujuan:

Menjawab pertanyaan mengapa jumlah pendaftar pada jurusan Fisika pada tahun akademik 2022 berbeda-beda di setiap jalur masuk. Informasi ini dapat membantu ITS dalam mengidentifikasi faktor-faktor yang memengaruhi minat calon mahasiswa pada jalur masuk tertentu. Beberapa pertimbangan yang dapat diambil meliputi kebijakan penerimaan, popularitas jurusan pada jalur tertentu, atau perubahan kondisi eksternal yang memengaruhi keputusan calon mahasiswa.

Dari hasil query di atas, kita dapat melihat bahwa jalur masuk SNMPTN memiliki jumlah pendaftar yang lebih tinggi dibandingkan dengan SBMPTN dan jalur mandiri pada jurusan Fisika pada tahun akademik 2022. Ini bisa disebabkan oleh berbagai faktor, seperti reputasi jurusan, kebijakan penerimaan, atau perubahan tren dalam preferensi calon mahasiswa. Dengan pemahaman ini, ITS dapat mengambil langkah-langkah strategis untuk meningkatkan daya tarik jurusan Fisika pada jalur-jalur tertentu atau menyesuaikan kebijakan penerimaan untuk meningkatkan minat mahasiswa pada jalur yang kurang diminati.

Desain & Implementasi Dashboard

a. Rancangan awal dashboard,

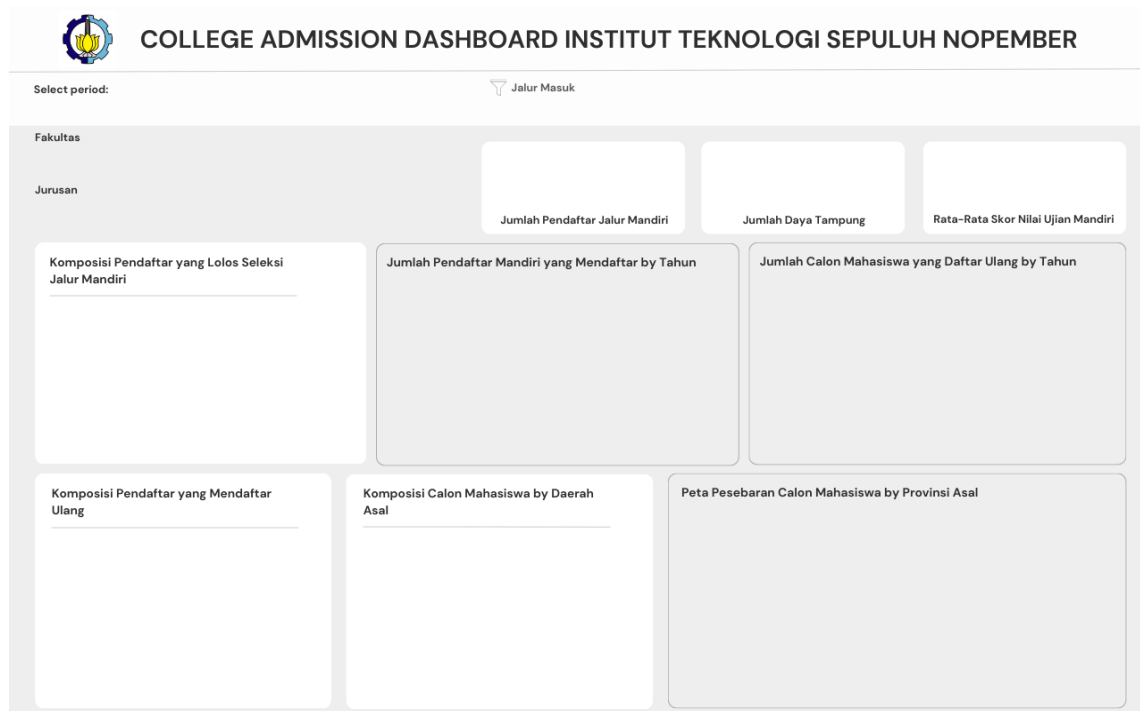
i. Tujuan dashboard

Dashboard “College Admission Institut Teknologi Sepuluh Noepmber” yang akan dibuat memiliki tujuan untuk meningkatkan pengelolaan dan pemahaman atas proses pendaftaran, penerimaan daya tampung, dan seleksi mahasiswa baru pada jalur seleksi mandiri. Hal ini akan memberikan wawasan yang lebih baik kepada tim pengelola dan pembuat keputusan untuk meningkatkan efisiensi dan efektivitas proses tersebut. Dashboard juga menyajikan visualisasi mengenai proses pendaftaran ulang dari berbagai jalur seleksi dimana hal ini akan memberikan wawasan mengenai persebaran calon mahasiswa baru ITS.

Selain itu, dengan menggunakan metode machine learning, dilakukan clustering K-Means dimana hasil cluster adalah untuk melihat range hasil nilai ujian pendaftar jalur mandiri berdasarkan kelulusan seleksi mandiri pendaftar pada data yang ada. Hasil clustering disajikan dalam bentuk visualisasi scatter plot pada halaman kedua dashboard.

ii. Definisikan kelompok visual (indikator, trend, detail)

Dashboard yang akan dibuat akan menyajikan visualisasi mulai dari komposisi, trend perkembangan pendaftar, dan persebaran calon mahasiswa. Pada bagian atas dashboard terdiri dari bagian fitur filter berdasarkan periode tahun akademik, jalur masuk, fakultas dan jurusan serta *overview card* dari jalur seleksi mandiri. Selanjutnya, pada bagian baris pertama visualisasi akan berisi gambaran pendaftar ke ITS melalui jalur seleksi mandiri yang disajikan baik dalam bentuk, sedangkan pada baris kedua visualisasi akan berisi gambaran persebaran dan komposisi calon mahasiswa baru ITS yang dapat disaring per jalur seleksi. Visualisasi yang dibuat akan menggunakan berbagai grafik visual baik donut chart, pie chart, line chart, column chart dan filled map. Berikut merupakan layout dari halaman pertama dashboard yang akan dibuat:



Untuk halaman kedua dashboard akan menggunakan scatter plot yang memvisualisasi hasil cluster dari range hasil nilai ujian pendaftar jalur mandiri berdasarkan kelulusan seleksi mandiri pendaftar pada data yang ada

iii. List dari visual per dashboard

Berikut merupakan list visualisasi dashboard yang akan dibuat dengan data pada OLAPpowerdlh

1. Jumlah Pendaftar Jalur Mandiri

Visual ini menampilkan jumlah pendaftar melalui jalur mandiri yang dapat di saring untuk setiap tahun penerimaan dalam bentuk kartu angka.

2. Jumlah Daya Tampung

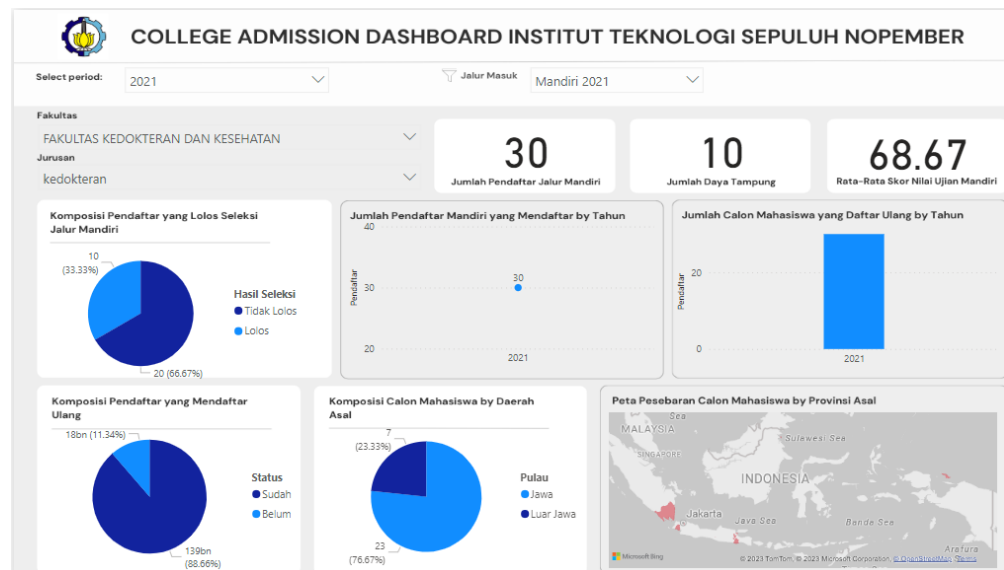
Visual dalam bentuk kartu angka yang berisi jumlah daya tampung yang dapat disaring untuk setiap tahun penerimaan, fakultas dan jurusan. Jumlah Pendaftar Jalur Mandiri dan Jumlah Daya Tampung dapat digunakan untuk mengetahui perbandingan antara jumlah pendaftar dan daya tampung jalur

mandiri. Data ini dapat digunakan untuk menilai tingkat persaingan jalur mandiri.

3. Rata-rata Skor Nilai Ujian Mandiri
Rata-rata skor nilai didapatkan dari rerata skor penilaian pada fact_seleksi yang disajikan dalam kartu angka dan dapat disaring berdasarkan tahun penerimaan
4. Komposisi Pendaftar yang Lolos Seleksi Mandiri
Visualisasi ini disajikan dalam bentuk diagram pie yang menggambarkan komposisi pendaftar yang lolos dan yang tidak lolos seleksi mandiri per tahun penerimaan. Visualisasi ini menggunakan data dari fact_seleksi untuk melihat hasil seleksi pendaftar jalur mandiri.
5. Jumlah Pendaftar Mandiri by Tahun
Dengan menggunakan data dari fact_pendaftaran yang berisi detail pendaftar jalur mandiri, visual ini menggunakan grafik garis untuk menunjukkan tren jumlah pendaftar melalui jalur mandiri dari tahun ke tahun.
6. Jumlah Calon Mahasiswa yang Daftar Ulang by Tahun
Dengan menggunakan data pada tabel fact_pendaftaran ulang, visualisasi ini menampilkan jumlah calon mahasiswa yang melakukan daftar ulang dari tahun ke tahun yang disajikan melalui grafik garis.
7. Komposisi Pendaftar yang Mendaftar Ulang
Dengan diagram pie, visual ini memperlihatkan komposisi pendaftar yang daftar ulang berdasarkan kelompok jalur seleksi dan tahun penerimaan. Visualisasi ini menggunakan fact_pendaftaran_ulang untuk melihat status pendaftaran ulang dari pendaftar yang lolos seleksi
8. Komposisi Calon Mahasiswa by Daerah Asal
Komposisi Calon Mahasiswa by Daerah Asal dapat digunakan untuk mengetahui persebaran calon mahasiswa berdasarkan daerah asal, dimana dibagi menjadi Pulau Jawa dan pulau di luar Jawa. Visual ini dapat digunakan untuk menilai kesetaraan akses pendidikan tinggi, khususnya ITS bagi calon mahasiswa dari daerah Jawa dan luar Jawa.

Data yang digunakan berasal dari dimensi pendaftar dimana alamat yang berisi kota dan provinsi pendaftar akan dipisah menjadi kolom baru dan diklasifikasikan berdasarkan pulau provinsinya.

9. Peta Persebaran Calon Mahasiswa by Provinsi Asal
Peta Persebaran Calon Mahasiswa by Provinsi Asal dapat digunakan untuk melihat persebaran calon mahasiswa secara geografis. Data ini dapat digunakan untuk menilai pemerataan akses pendidikan tinggi khususnya ITS di berbagai provinsi di Indonesia.



Berikut contoh hasil dashboard yang disaring untuk periode penerimaan tahun 2021 untuk jalur masuk mandiri. Visualisasi ini menyajikan rangkuman untuk fakultas kedokteran dan kesehatan khususnya untuk jurusan kedokteran.

iv. List visualisasi tersebut juga dijelaskan apakah menggunakan metode machine learning atau tidak

Pada sheet 2 dashboard ditampilkan hasil visualisasi penerapan machine learning dengan Clustering pada skor penilaian. Penerapan teknik clustering dalam pembuatan dashboard College Admission bertujuan untuk memberikan wawasan mendalam terkait struktur dan karakteristik pendaftar. Dengan melakukan clustering berdasarkan skor penilaian, kami berharap dapat mengidentifikasi kelompok-kelompok pendaftar yang memiliki kesamaan dalam aspek tertentu. Hal ini memungkinkan kami untuk melakukan segmentasi pendaftar, memahami pola-pola unik di antara kelompok-kelompok tersebut, dan menyelidiki perbedaan dalam keputusan seleksi. Visualisasi hasil clustering pada dashboard akan mempermudah pengguna untuk memahami dan menganalisis informasi tersebut secara lebih holistik. Sebagai hasilnya, pemangku kepentingan dalam College Admission dapat mengambil keputusan yang lebih terinformasi dan mengembangkan strategi yang sesuai untuk meningkatkan proses seleksi dan pengambilan keputusan.

Untuk tahapan Clustering dilakukan di Google Collab yang nanti setelah diolah, dataframe nya akan didownload dan ditampilkan visualisasinya di dashboard Power BI.

Berikut merupakan tahapan dari clustering nya :

1. Import library dari python yang akan digunakan dan juga import dataset. Di sini dataset yang digunakan adalah dataset dari tabel fact_seleksi. Kemudian tampilkan data dengan `.head()` untuk melihat isi data yang diimport apakah sudah sesuai apa belum

```
Clustering pada skor nilai mandiri.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[42] # Import library
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import silhouette_score

from google.colab import files
uploaded = files.upload()

seleksi_college.csv
• seleksi_college.csv(1000000000) - 216364 bytes, last modified 12/17/2023 - 100% done
Saving seleksi_college.csv to seleksi_college (1).csv

[14] data = pd.read_csv('seleksi_college.csv', sep=';')

[15] data.head()
```

	id_seleksi	id_jurusan	id_tahun_akademik	nomor_pendaftaran	skor_penilaian	hasil_seleksi
0	40500001	1000000001	1001	42161651	72.13	Lolos
1	40500002	1000000002	1001	42161652	82.64	Lolos
2	40500003	1000000003	1001	42161653	57.28	Tidak Lolos
3	40500004	1000000004	1001	42161654	49.27	Tidak Lolos
4	40500005	1000000005	1001	42161655	85.87	Lolos

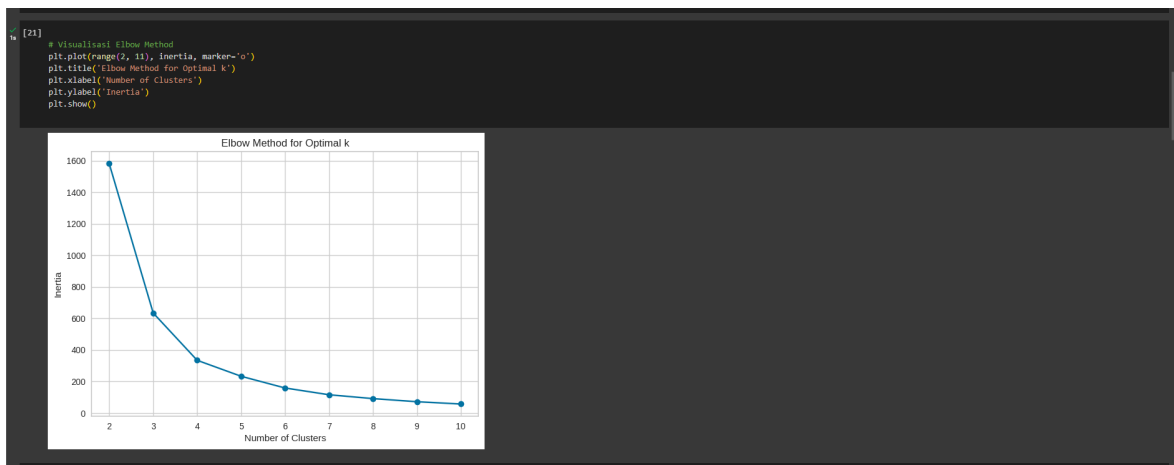
2. Dilakukan pemilihan kolom yang akan digunakan sebagai variabel bebas, yaitu 'skor_penilaian' dan dilanjutkan dengan penormalan skala fitur serta mencari nilai k nya paling optimal dengan metode Elbow

```
[18] # Pilih fitur
features = ['skor_penilaian']

[19] # Menggunakan StandardScaler untuk menormalkan skala fitur
scaler = StandardScaler()
data_scaled = scaler.fit_transform(data[features])

[20] # Menentukan jumlah cluster dengan metode Elbow
inertia = []
for n_clusters in range(2, 11):
    kmeans = KMeans(n_clusters=n_clusters, random_state=42)
    kmeans.fit(data_scaled)
    inertia.append(kmeans.inertia_)
```

3. Berikut merupakan hasil dari pencarian nilai k, bisa dilihat bahwa nilai yang paling optimal adalah nilai K=5



4. Dilanjutkan dengan membuat kolom clustering dari hasil clusteringnya

```
[43] # Berdasarkan Elbow Method, pilih jumlah cluster yang optimal
optimal_clusters = 5

[38] # Melakukan k-means clustering
kmeans = KMeans(n_clusters=optimal_clusters, random_state=42)
data['cluster'] = kmeans.fit_predict(data_scaled)

data['cluster']

0      1
1      3
2      0
3      0
4      3
..
4495   3
4496   2
4497   2
4498   1
4499   0
Name: cluster, length: 4500, dtype: int32

[39] # Mapping kategori ke nilai numerik
mapping = {'lolos': 1, 'Tidak Lolos': 0}
data['hasil_seleksi_numerik'] = data['hasil_seleksi'].map(mapping)
```

5. Berikut merupakan tabel hasil dari clustering

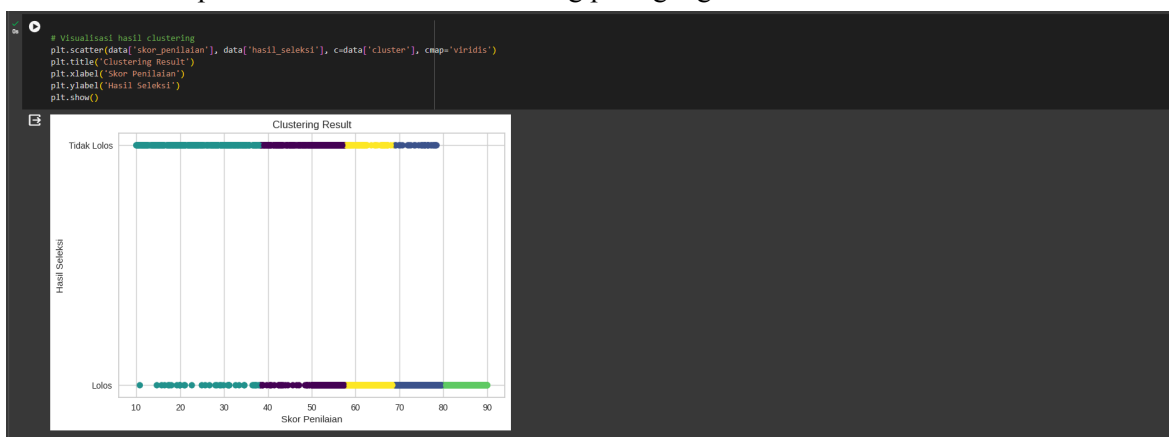
```
[39] # Mapping kategori ke nilai numerik
mapping = {'lolos': 1, 'Tidak Lolos': 0}
data['hasil_seleksi_numerik'] = data['hasil_seleksi'].map(mapping)

[40] data
```

	id_seleksi	id_jurusan	id_tahun_akademik	nomor_pendaftaran	skor_penilaian	hasil_seleksi	cluster	hasil_seleksi_numerik
0	40500001	1000000061	1001	42161651	72.13	Lolos	1	1
1	40500002	1000000062	1001	42161652	82.64	Lolos	3	1
2	40500003	1000000063	1001	42161653	57.28	Tidak Lolos	0	0
3	40500004	1000000064	1001	42161654	49.27	Tidak Lolos	0	0
4	40500005	1000000065	1001	42161655	85.87	Lolos	3	1
...
4495	40504496	1000000050	1003	42363330	85.56	Lolos	3	1
4496	40504497	1000000051	1003	42363331	27.32	Tidak Lolos	2	0
4497	40504498	1000000052	1003	42363332	31.41	Tidak Lolos	2	0
4498	40504499	1000000053	1003	42363333	71.92	Lolos	1	1
4499	40504500	1000000054	1003	42363334	50.52	Tidak Lolos	0	0

4500 rows x 8 columns

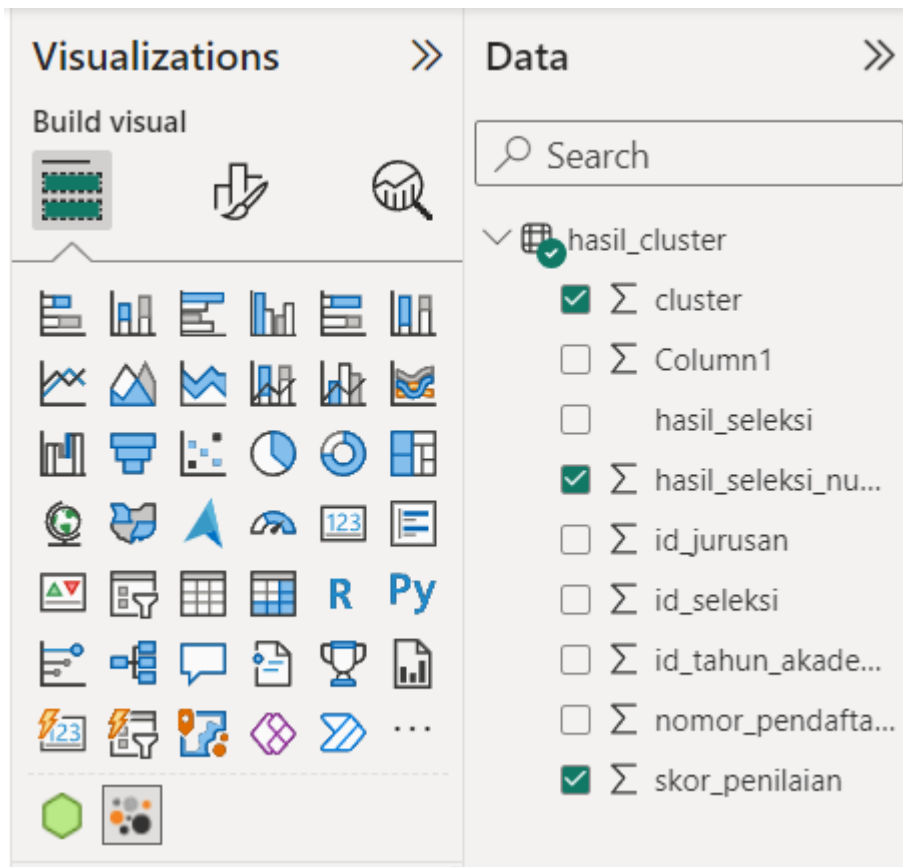
6. Ini merupakan hasil visualisasi clustering pada google collab :



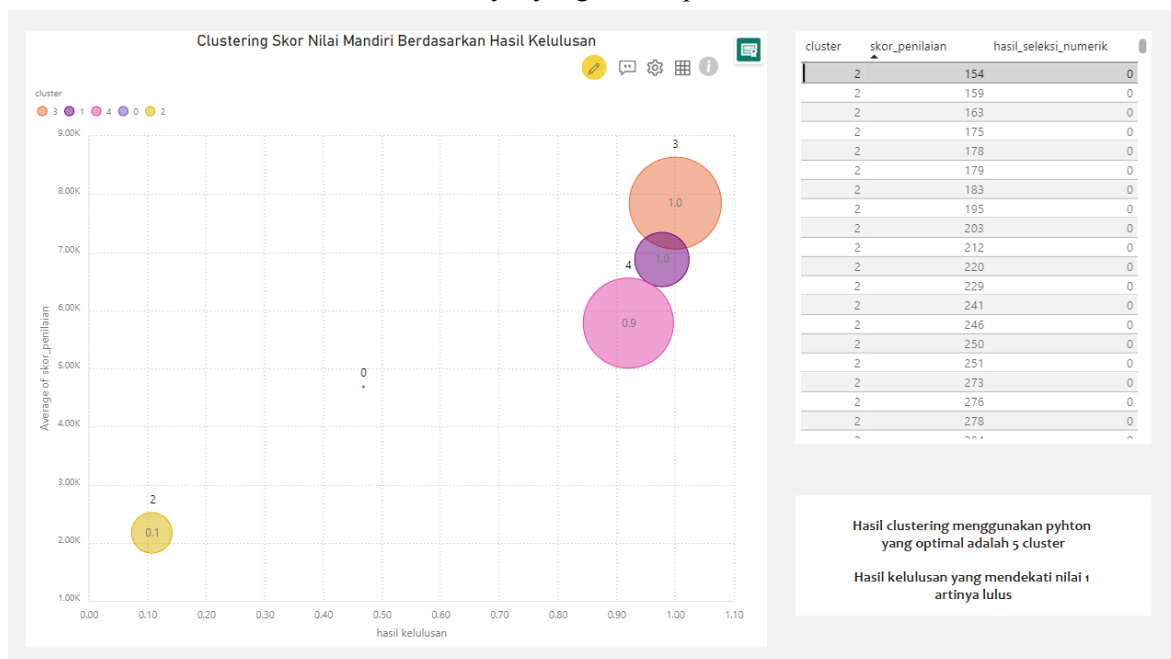
7. Untuk menyimpan tabel tersebut ke dalam bentuk format csv, dapat menggunakan kode ini dan file csv nya pun berhasil didownload

```
[41] from google.colab import files
data.to_csv('hasil_cluster.csv')
files.download('hasil_cluster.csv')
```

Tahap berikutnya adalah mengimport data tersebut ke Power BI tempat di mana dashboard akan diolah, berikut merupakan hasil impor nya :

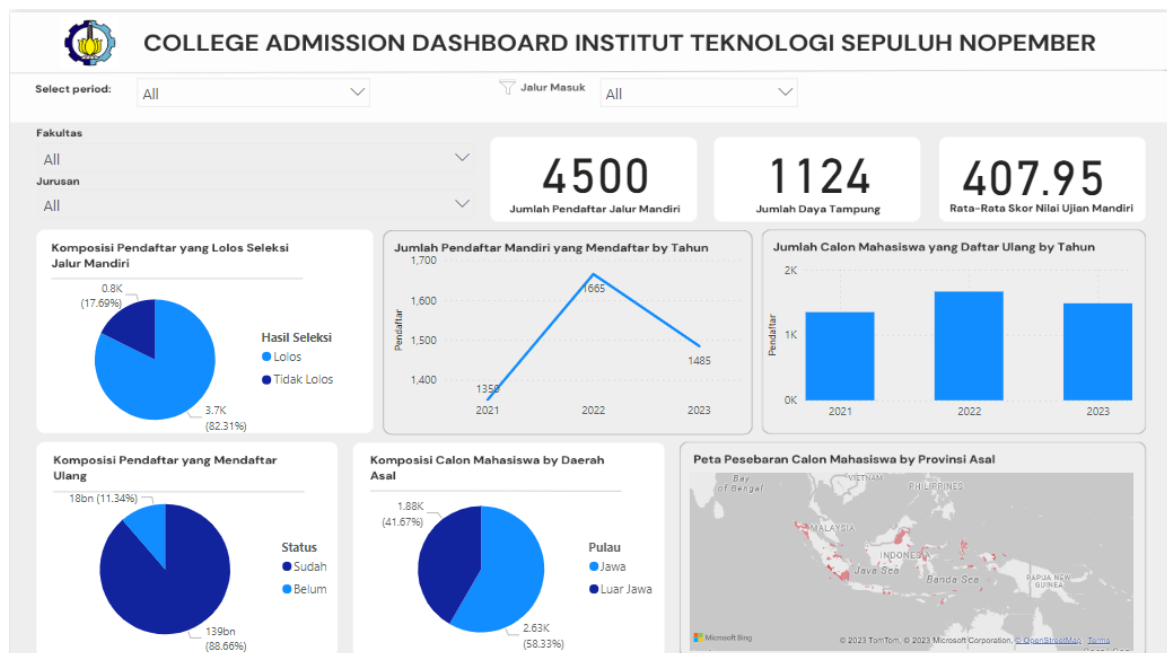


Dan ini adalah hasil visualisasi dari datanya, yang berada pada sheet kedua dashboard :



Bisa dilihat dari hasil clustering tersebut terdapat 5 kelas clustering yang ada pengolahan dataset 'fact_seleksi' dan ada juga tabel tambahan di sampingnya untuk melihat cluster, skor_penilaian dan hasil_seleksi_numerik yang mana 0 artinya tidak lulus dan 1 artinya lulus.

b. Implementasi Dashboard



Terlihat bahwa jumlah pendaftar jalur seleksi mandiri dari tahun 2021 ke 2022 mengalami peningkatan namun menurun drastis pada tahun 2023. Pada bagian komposisi calon mahasiswa berdasarkan pulau terlihat bahwa pendaftar di luar Pulau Jawa masih belum mencapai setengah. Dari kedua hal tersebut diharapkan dapat dilakukan peningkatan untuk melakukan promosi ITS ke luar pulau Jawa dan meningkatkan jumlah pendaftar pada tahun 2024.