

# **PEMBANGUNAN DATA MART PLAFORM SATU DATA ITERA**

---

**Tugas besar matakuliah  
Pergudangan Data - SD25-310007**

# Daftar Isi

**Ringkasan Eksekutif**

**Pengujian dan Validasi**

**Pendahuluan**

**Deployment dan Operasional**

**Kebutuhan analisis**

**Hasil Dan pembahasan**

**Design model**

**Kesimpulan dan pengembangan selanjutnya**

**Implementasi ETL**

**Lampiran**

# EXECUTIVE SUMMARY



## PROJECT OVERVIEW

Portal Satu Data ITERA, sebuah platform data terpadu yang menyediakan akses ke dataset dan informasi institusi. Sistem ini dirancang untuk mendukung analisis penggunaan portal, pemantauan kualitas data, dan penyediaan wawasan bagi pengelola portal serta manajemen ITERA.



## CAPAIAN KUNCI

Perancangan model dim 4 fakta dan 6 dimensi  
ETL SCD type 1 dan Type 2  
Implementasi RBAC  
Interactive Dashboard



## DAMPAK BISNIS

Pemantauan akses realtime  
Analisis kualitas data  
Identifikasi Data yang populer  
dan tren pencarian  
Pelaporan metrik



## Rekomendasi

Implementasi integrasi langsung (Slakad, HRIS)  
Pengembangan dasboard mobile  
Penambahan fitur prediktif pada machine learning

Minggu ke 7

# Latar Belakang Masalah

Saat ini pengelolaan data di banyak perguruan tinggi masih menghadapi tantangan klasik yaitu data silo dimana kondisi ini data tersimpan secara terpisah di berbagai sistem/unit tanpa standar tunggal dalam definisi, struktur, format, maupun dokumentasi metadata hal ini menyebabkan rekapitulasi data yang lama, perbedaan hasil laporan unit, rendahnya interoperabilitas dan sulitnya memperoleh data institusi yang konsisten sebagai basis analitik dan kebijakan.

Seiring berjalannya waktu, ITERA membuat inovasi baru untuk membentuk inisiatif platform SATU DATA ITERA yang direpresentasikan untuk membangun kolaboratif antar unit di ITERA agar data institusi terintegrasi oleh 1 platform, memiliki standard format dan definisi yang sama antar domain, interoperabel lintas sistem dan unit dan kemudahan mengakses informasi dataset institusi dengan melihat dashboard insight

Untuk mewujudkan keefektifan dari platform ini, pembangunan data warehouse dan data mart institusi menjadi elemen paling krusial dengan mengintegrasikan menggunakan cloud provider seperti microsoft azure ITERA dapat menyiapkan fondasi implementasi data institusi dan analitik berskala open source menuju campus-wide data governance. Diharapkan untuk meningkatkan efisiensi manajerial, mengurangi kesalahan laporan dan proses manual, menyediakan data riset, dan membangun ekosistem data institusi yang modern, terpercaya dan proaktif.

# Tujuan Proyek

1

Memahami proses bisnis  
Portal Satu Data ITERA dan  
kebutuhan analitik  
pengelola  
portal

4

Mengimplementasikan  
data warehouse dengan  
proses ETL yang andal

2

Mengidentifikasi sumber data  
portal dan metrik penggunaan  
yang relevan

5

Mengembangkan  
dashboard analitik untuk  
visualisasi data

3

Merancang model dimensional  
untuk pelacakan akses dataset,  
kualitas data, dan perilaku  
pengguna

6

Mendokumentasikan  
desain dan implementasi  
secara komprehensif

# Metodologi



## Misi 1

• Desain konseptual dan logikal data mart



## Misi 2

• Implementasi basis data dan proses ETL



## Misi 3

• Dashboard, keamanan, dan pengujian

## Ruang lingkup

- **Analisis kebutuhan bisnis untuk portal satu data itera**
- **perencangan ERD untuk entitas portal**
- **perencangan skema bintang**
- **implementasi basis data menggunakan SQL server**
- **Pengembangan proses ETL**
- **Pembuatan Dashboard powerBI**
- **Pengujian dan validasi sistem**
- **Dokumentasi operasional dan pemeliharaan**

# Analisis Kebutuhan

## Identifikasi Stakeholders

Tim pengelola Portal Satu Data

Manajemen ITERA

Pengelola data (Data Stewards)

Pengguna Eksternal

## Analisis Proses Bisnis

Pengelolaan dataset

Akses dan download dataset

Manajemen Kualitas Data

Pencarian dan penemuan

## Kebutuhan analitik Fungsional

Dapat merekam setiap akses dataset

Dapat menghitung metrik dan mengagregasi metrik

Dapat menyediakan dashboard untuk melihat analisis tren pola

## Non Fungsional

Kinerja dan ketersediaan

Keamanan

Skalabilitas

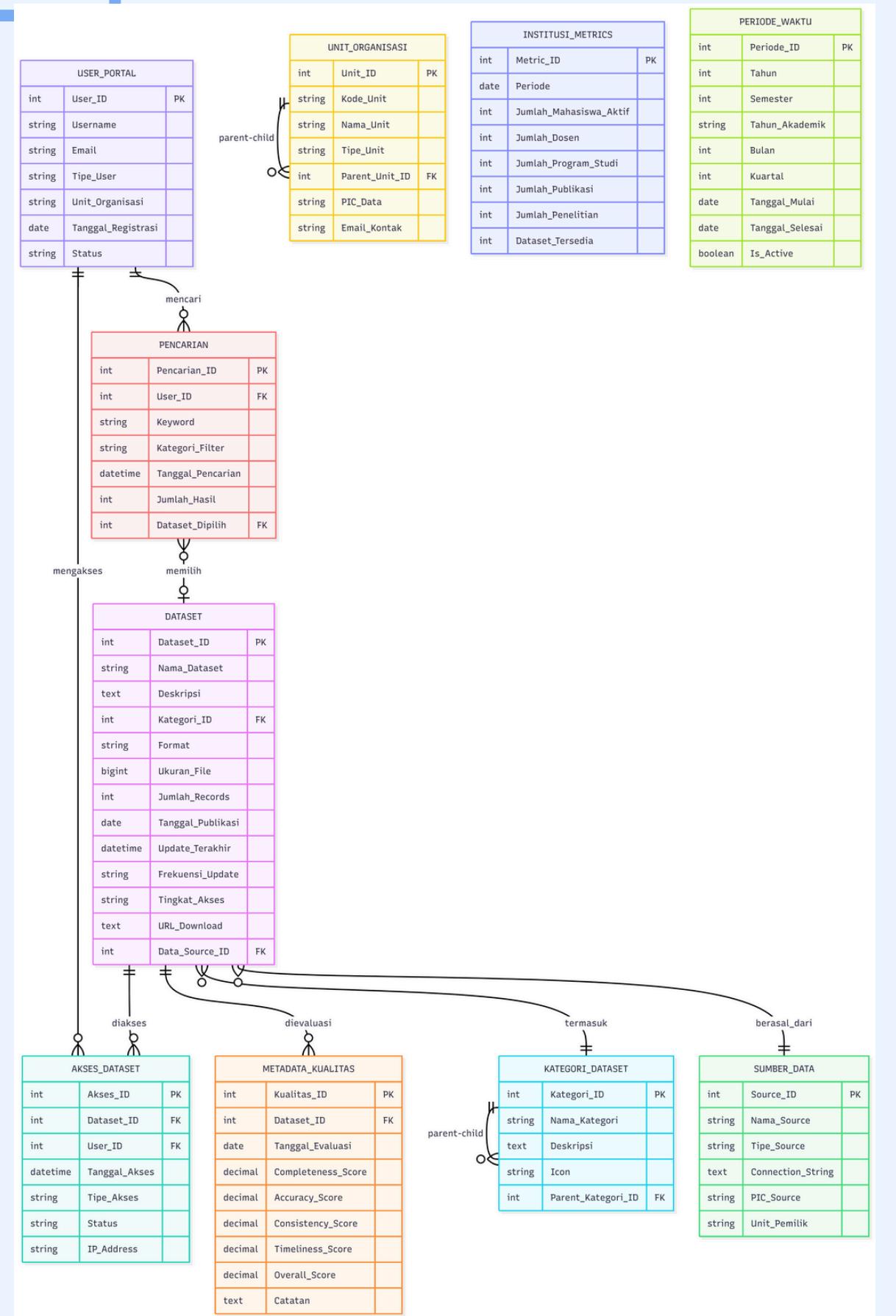
# Requirements Analysis

Sumber Data	Tipe	Volume	Frekuensi	Kualitas
SatuData_API	REST API- /JSON	~1K dataset	Real-time	Tinggi
Access_Log	Server Log	~50K bari- s/bln	Real-time	Tinggi
Search_Query_Log	Server Log	~10K bari- s/bln	Real-time	Tinggi
Dataset_Metadata	Basis Data	~700 dataset	Harian	Tinggi
User_Registry	Basis Data	~5K penggu- na	Harian	Tinggi

Tabel 1: Sumber Data Portal Satu Data ITERA

# Perancangan Sistem

## Model Konseptual (ERD)

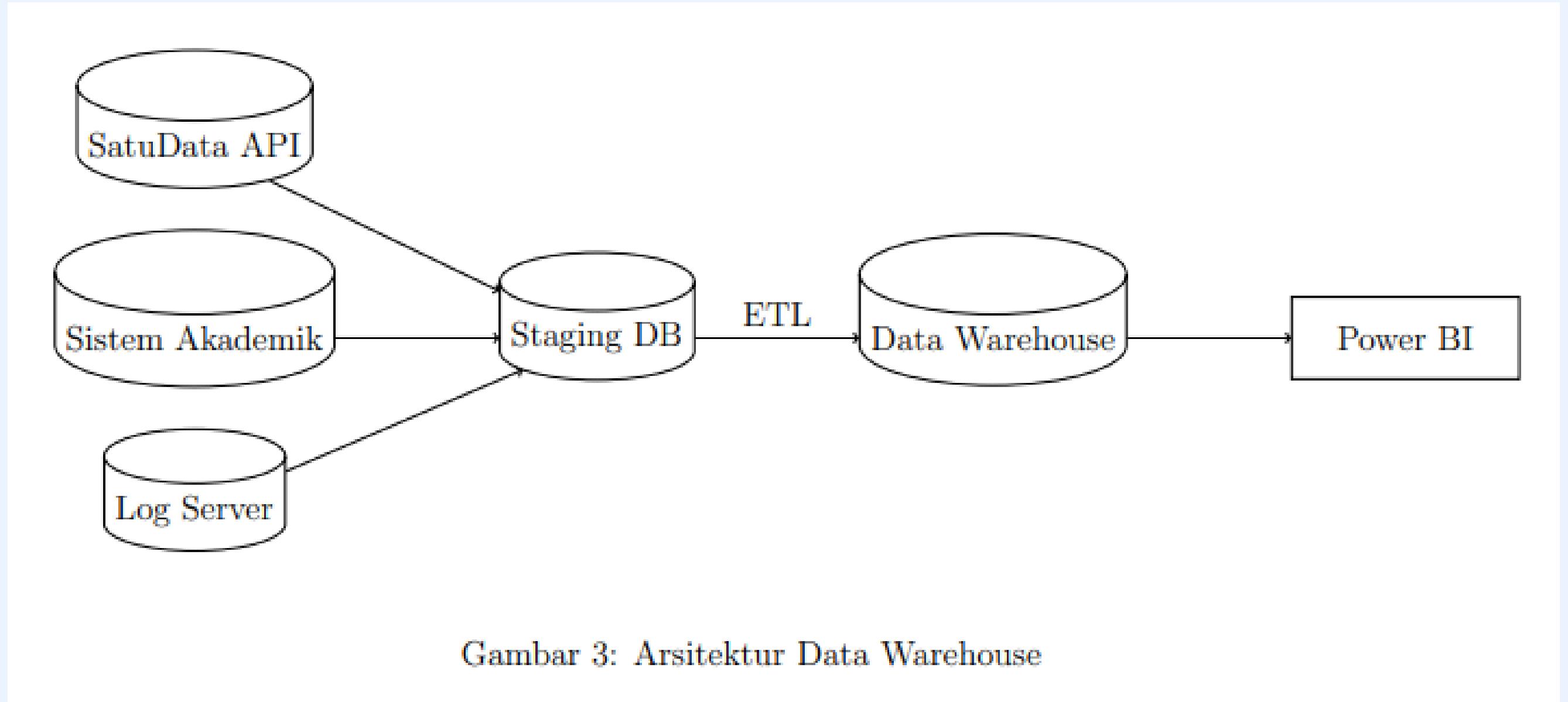


Mengidentifikasi 8 entitas utama dalam ekosistem portal. Fokus pada hubungan antara Pengguna, Dataset, dan Log Aktivitas.

- User - melakukan pencarian - memilih dataset
- User - mengakses dataset
- Dataset - memiliki kategori, metadata kualitas, dan sumber data
- Organisasi - mengelola pengguna dan dataset
- Dataset & aktivitasnya - digunakan untuk menghasilkan metrik institusi
- Periode waktu - memetakan semua aktivitas dan metrik dalam konteks temporal

# Perancangan Sistem

## Arsitektur Data Warehouse

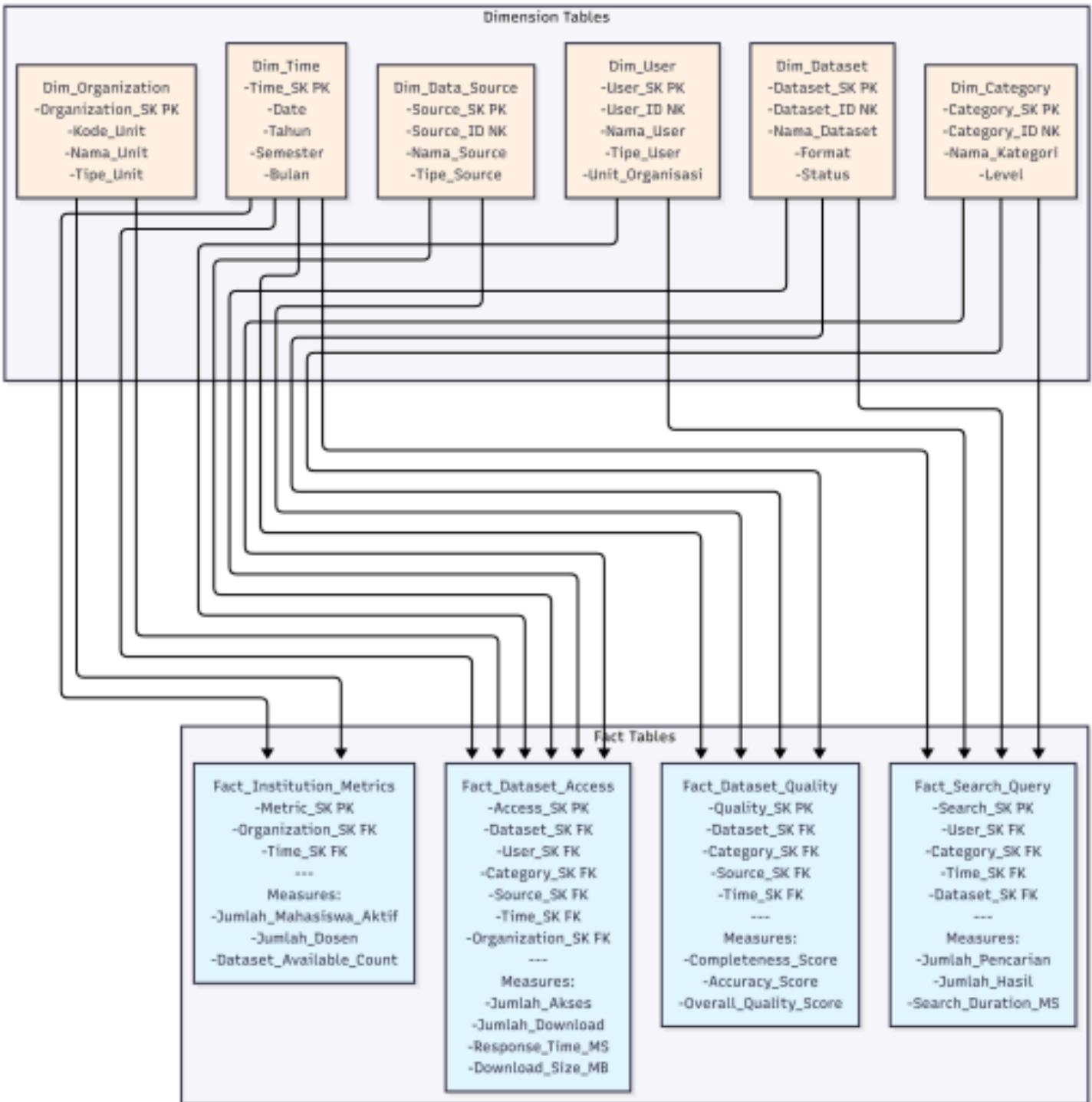


Gambar 3: Arsitektur Data Warehouse

- Sumber: API, Sistem Akademik, Log Server.
- ETL: Ekstraksi dan transformasi data.
- Storage: SQL Server 2019 (Star Schema).
- Output: Power BI Dashboard.

# Perancangan Sistem

## Model Dimentional



- 4 Tabel Fakta: Access, Quality, Search Query, Institution Metrics .
- SCD Type 2: Diterapkan pada Dimensi Dataset untuk melacak perubahan data.

Gambar 2: Dimensional Model Portal Satu Data ITERA

# Implementasi

## Proses ETL & Basis Data

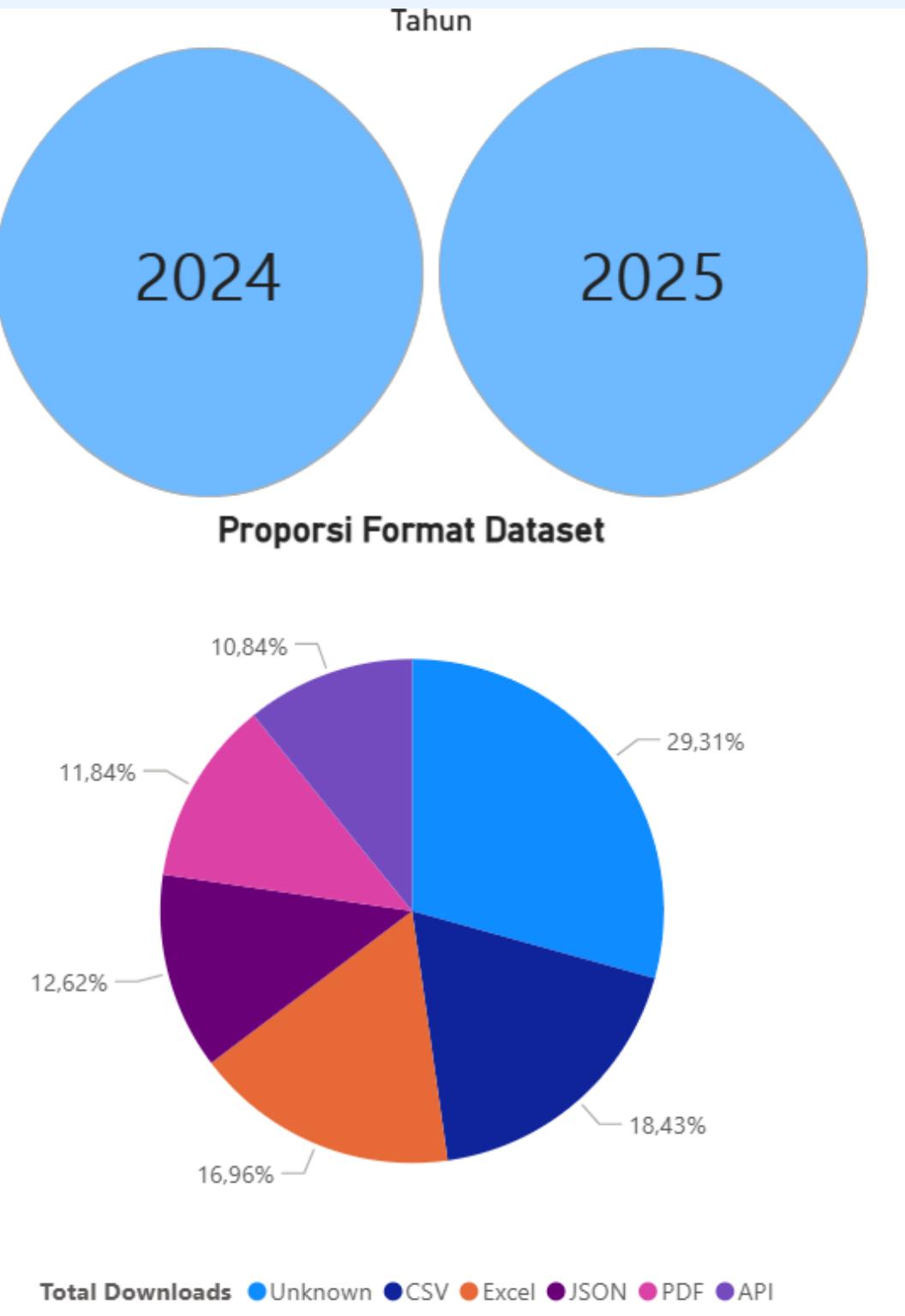
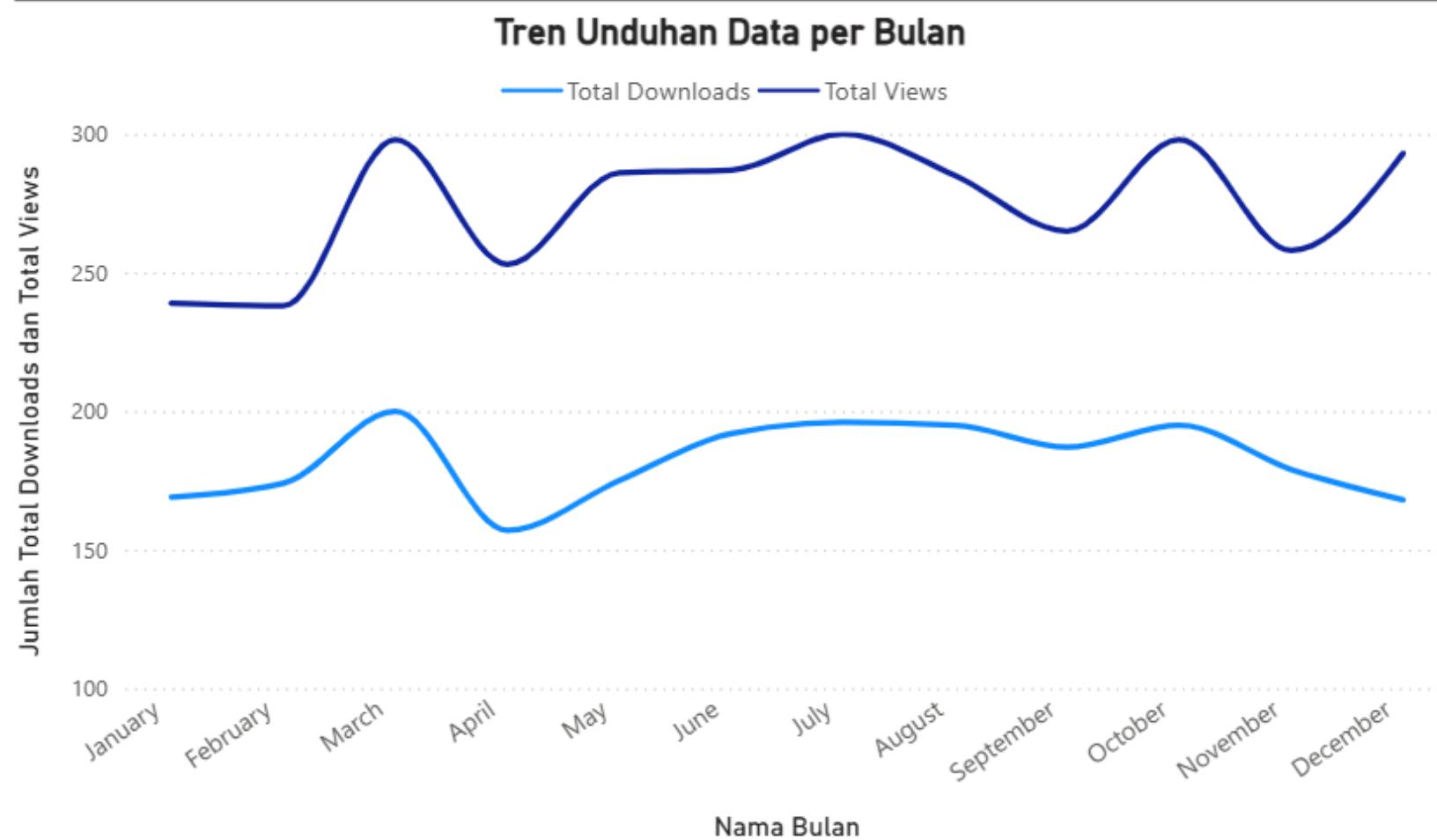
Proses	Frekuensi	Jadwal
ETL Harian	Senin-Sabtu	02:00 WIB
ETL Mingguan	Minggu	03:00 WIB
Refresh Penuh	Bulanan	Minggu ke-1, 01:00 WIB

- Otomasi: Menggunakan Stored Procedure `usp_Master_ETL`.
- Jadwal: Refresh harian setiap jam 02:00 pagi.
- Keamanan: Role-Based Access Control (RBAC).

# Implementasi

## Dashboard Analitik

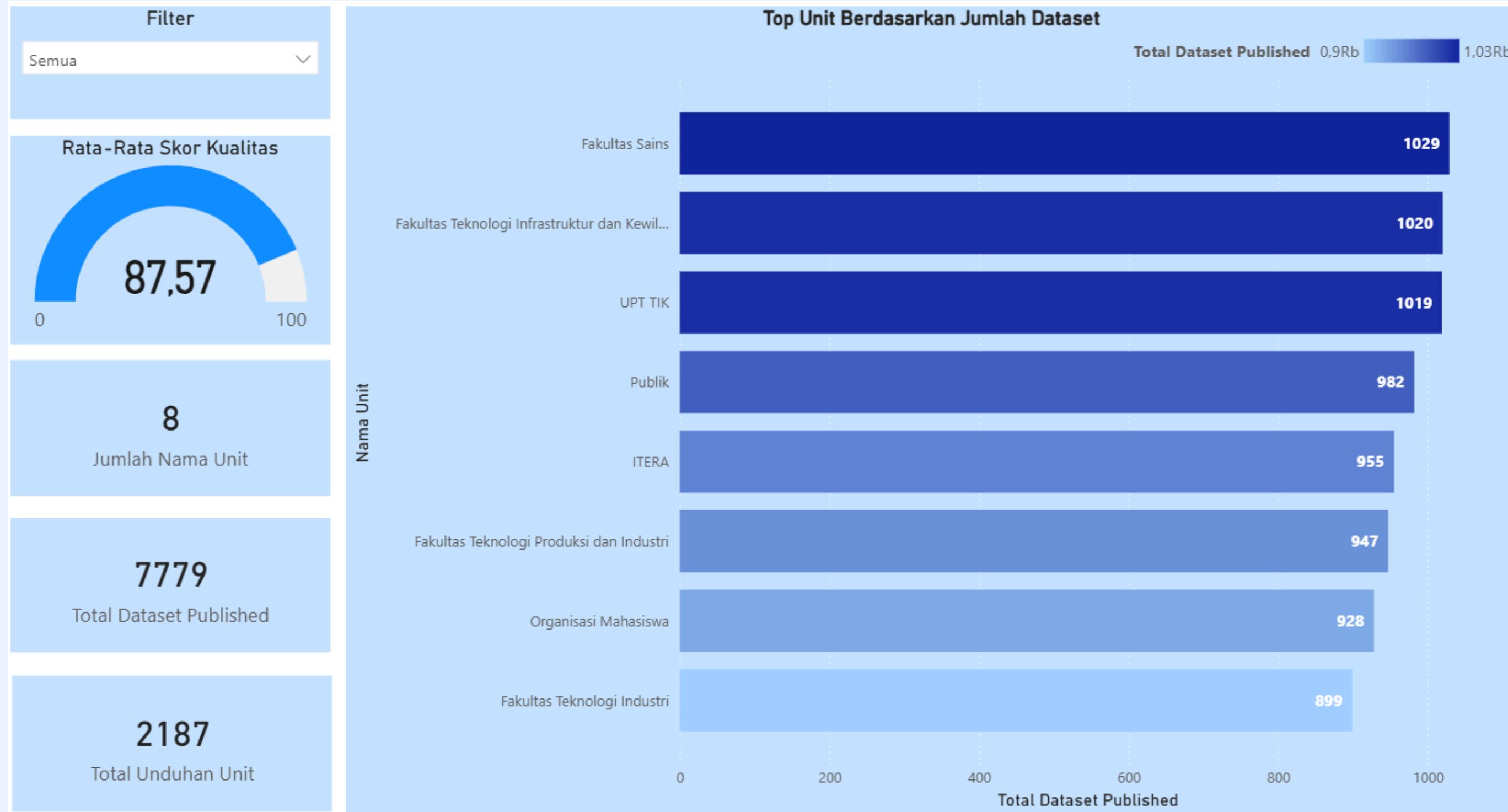
>Nama Dataset	Kategori Nama	Format	Total Downloads	Total Views	
Dataset ITERA Fasilitas No.176	Fasilitas	CSV	26	12	
Dataset ITERA Keuangan No.111	Keuangan	JSON	20	20	
Dataset Popularity Dipilih	o.66	Keuangan	Unknown	19	18
	o.165	Keuangan	PDF	18	14
Dataset ITERA Penelitian No.106	Penelitian	Unknown	18	14	
Dataset ITERA Penelitian No.166	Penelitian	Unknown	18	19	
Dataset ITERA Akademik No.97	Akademik	CSV	17	10	
Dataset ITERA Fasilitas No.151	Fasilitas	PDF	17	21	
Dataset ITERA Fasilitas No.181	Fasilitas	Excel	17	18	
Dataset ITERA Kepegawaian No.152	Kepegawaian	PDF	17	25	
<b>Total</b>			<b>2187</b>	<b>3300</b>	



Gambar Dataset Popularity

# Implementasi

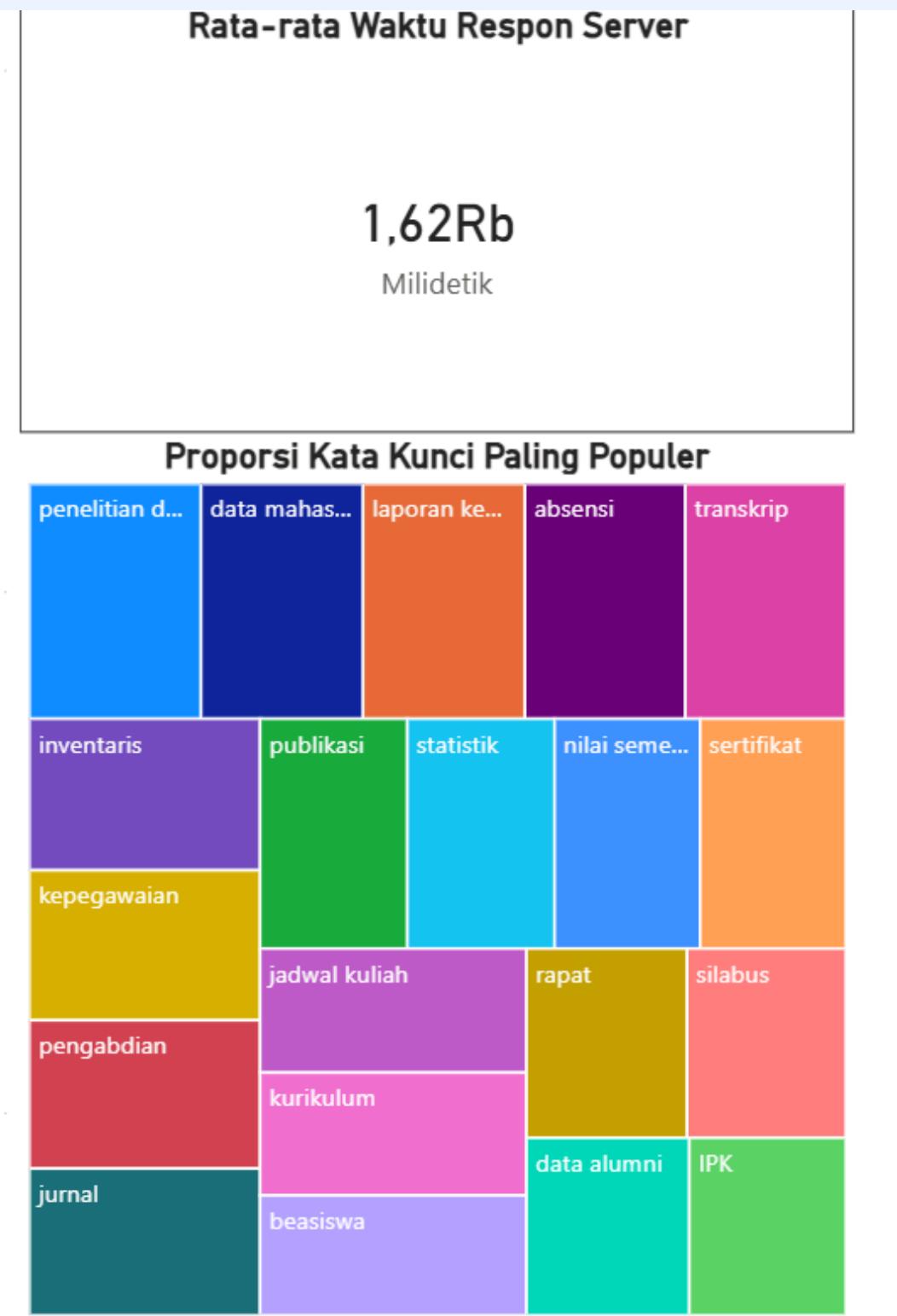
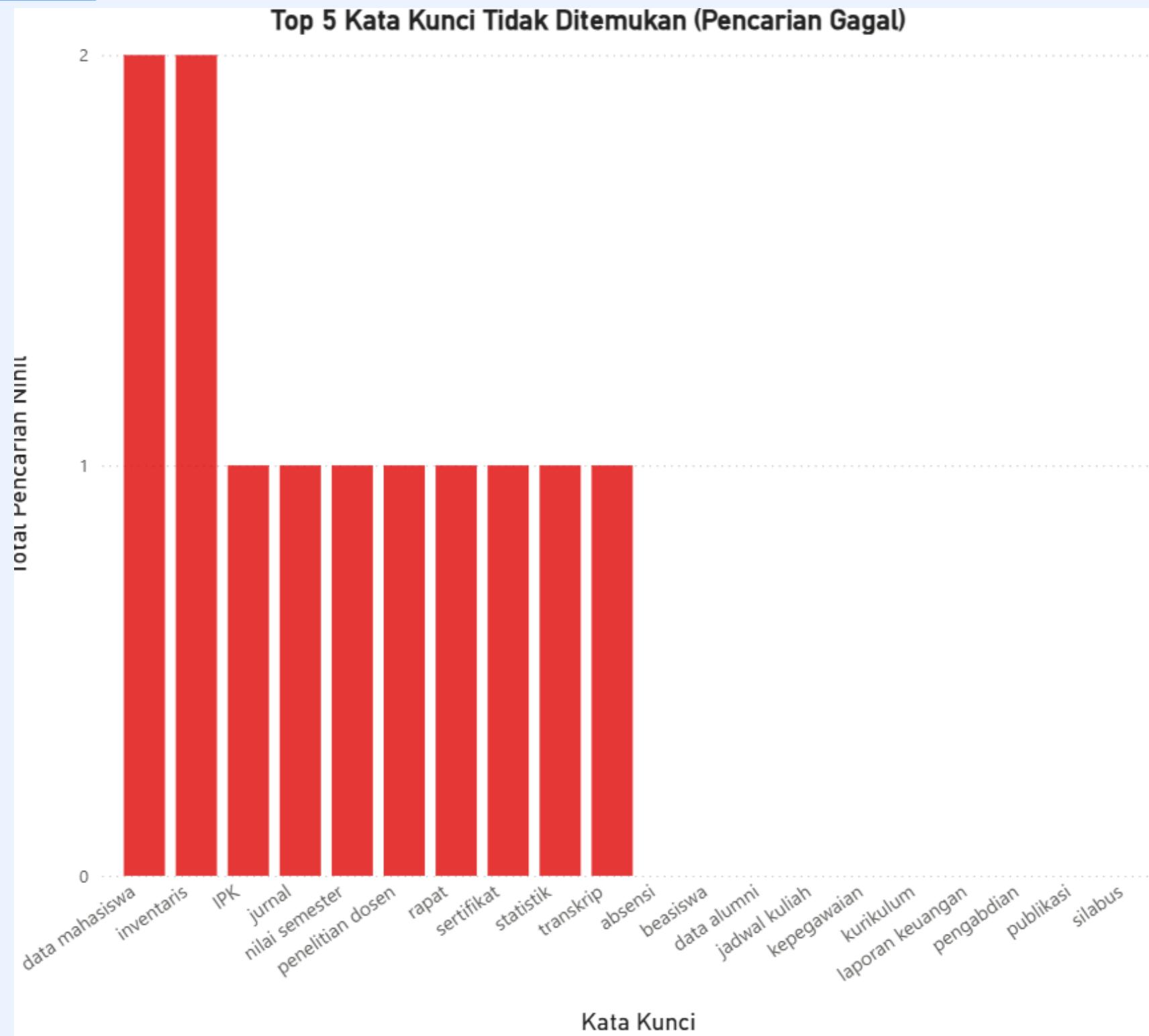
## Dashboard Analitik



Gambar Executive Governance

# Implementasi

## Dashboard Analitik



Gambar Search Analysis

# Pengujian dan Validasi

## Hasil Pengujian Kinerja & UAT

Hasil pengujian kinerja query dashboard:

Query	Waktu Eksekusi	Status
Executive Summary	2.3 detik	Lulus
Dataset Popularity	1.8 detik	Lulus
User Activity	3.1 detik	Lulus
Data Quality Report	2.7 detik	Lulus

Performa Query: Rata-rata 2.5 detik (Target < 10 detik)  
Akurasi ETL: 100% Berhasil.

UAT: 6 Test Case (Fungsi, Data, Security) - Semua Lulus

User Acceptance Testing dilaksanakan dengan hasil:

Keterangan	Nilai
Total Test Case	6
Lulus	6
Gagal	0

# Deployment dan Operasional

## Deployment Produksi & Backup

Deployment: SQL Server 2019 pada Docker/Linux.

Backup:

- Full: Mingguan (Retensi 30 hari).
- Differential: Harian.
- Log: Per 6 jam.

Tipe Backup	Frekuensi	Waktu	Retensi
Full	Mingguan (Minggu)	02:00 WIB	30 Hari
Differential	Harian (Senin-Sabtu)	02:00 WIB	14 Hari
Transaction Log	Setiap 6 Jam	00, 06, 12, 18	7 Hari

Tabel 2: Jadwal Backup

# Hasil

Berdasarkan implementasi dan pengujian, diperoleh temuan berikut:

- Model dimensional Star Schema efektif untuk analisis portal data
- Mekanisme SCD Type 2 berhasil melacak perubahan metadata dataset
- Proses ETL dapat menangani volume data yang diharapkan
- Dashboard Power BI menyediakan visualisasi yang informatif

Data warehouse menghasilkan wawasan berikut:

- Kategori dataset Akademik adalah yang paling sering diakses
- Waktu puncak penggunaan portal adalah 09:00–11:00 dan 14:00–16:00
- 80% unduhan dilakukan oleh pengguna internal (mahasiswa dan dosen)
- Rata-rata skor kualitas dataset adalah 87.5%

## 8.3. Metrik Kinerja

Metrik	Target	Aktual
Waktu Respons Query	< 10 detik	2.5 detik
Keberhasilan ETL	> 95%	100%
Ketersediaan Sistem	> 99%	99.5%
Kelulusan UAT	100%	100%

- Waktu Respons Query: 2.5 detik (Lebih cepat dari target 10 detik).
- Keberhasilan ETL: 100% (Target > 95%).
- Ketersediaan Sistem: 99.5% (Target > 99%).

# Kesimpulan

Proyek Data Warehouse Satu Data ITERA telah berhasil mencapai tujuan utamanya melalui perancangan model dimensional yang komprehensif, implementasi basis data yang kuat, serta proses ETL yang andal dengan tingkat keberhasilan pengujian 100% dan dashboard yang informatif. Keberhasilan ini memberikan pembelajaran penting mengenai krusialnya analisis kebutuhan mendalam, pendekatan iteratif, serta dokumentasi lengkap untuk pemeliharaan jangka panjang; namun, sistem saat ini masih memiliki ruang untuk penyempurnaan mengingat integrasi sumber data yang masih bersifat simulasi, belum adanya fitur alerting otomatis untuk kegagalan ETL, serta belum tersedianya akses dashboard dalam versi mobile.

# Daftar Pustaka



ITERA Official Website. (2024).  
<https://www.itera.ac.id/>



Satu Data ITERA. (2024).  
<https://data.itera.ac.id/>



Inmon, W. H. (2005). Building the Data Warehouse (4th ed.). Wiley.



Microsoft. (2024). SQL Server Documentation.  
<https://docs.microsoft.com/en-us/sql/>



Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling (3rd ed.). Wiley.

# MARI BERDISKUSI

Silakan bertanya.  
Terbatas bagi 3 pertanyaan.

# TERIMA KASIH

---

Mohon maaf atas kesalahan dan  
kekurangan selama presentasi berlangsung.