

**LAPORAN AKHIR PEMBANGUNAN DATA MART KEPEGAWAIAN
INSTITUT TEKNOLOGI SUMATERA**



Oleh :

Kelompok 8

- 1. Fadhil Fitra Wijaya - 122450082**
- 2. Ali Aristo Muthahhari Parisi - 123450088**
- 3. Benget Sidabutar - 123450047**
- 4. Nobel Nizam Fathirizki - 123450117**

**PROGRAM STUDI SAINS DATA
FAKLUTAS SAINS
INSTITUT TEKNOLOGI SUMATERA
2025**

DAFTAR ISI

DAFTAR ISI.....	2
EXECUTIVE SUMMARY.....	4
1.1 Project Overview.....	4
1.2 Capaian Kunci.....	4
1.3 Dampak Bisnis.....	4
1.4 Rekomendasi.....	4
BAB 1 PENDAHULUAN.....	6
1.1 Latar Belakang.....	6
1.2 Tujuan.....	6
1.3 Batasan.....	7
BAB 2 REQUIREMENTS ANALYSIS.....	8
2.1 Business Requirements Analysis.....	8
2.1.1 Identifikasi Stakeholders.....	8
2.1.2 Analisis Proses Bisnis.....	9
2.1.3 Kebutuhan Analitik dan Key Performance Indicators (KPI).....	10
BAB 3 PERANCANGAN SISTEM.....	12
3.1 Conceptual Data Model (ERD).....	12
3.2 Dimensional Model.....	13
3.2.1 Dimension Tables.....	13
3.2.2 Fact Tables.....	14
3.3 Physical Data Model.....	14
3.3.1 Partitioning Strategy.....	14
3.3.2 Indexing Strategy.....	14
3.4 System Architecture & ETL Design Arsitektur.....	15
BAB 4 IMPLEMENTASI SISTEM.....	16
4.1 Implementasi Database.....	16
4.1.1 Schema Strategy.....	16
4.1.2 Partitioning Implementation.....	16
4.1.3 Indexing Strategy.....	16
4.2 Implementasi ETL (Extract, Transform, Load).....	16
4.2.1 Extraction (Load to Staging).....	16
4.2.2 Transformation Logic.....	17
4.3 Pengembangan Analytical Views.....	17
BAB 5 PENGUJIAN DAN VALIDASI.....	18
5.1 Data Quality Assurance (DQA).....	18
5.2 Performance Testing.....	19

BAB 6 HASIL DAN PEMBAHASAN.....	20
6.1 Temuan Kunci Bisnis (Business Insights).....	20
6.2 Temuan Bisnis.....	23
BAB 7 PENUTUP.....	24
7.1 Kesimpulan.....	24
7.2 Rekomendasi Pengembangan.....	24
LAMPIRAN.....	25

EXECUTIVE SUMMARY

1.1 Project Overview

Pengelolaan Sumber Daya Manusia (SDM) di Institut Teknologi Sumatera (ITERA) menghasilkan data transaksional yang besar, meliputi data profil pegawai, absensi harian, penggajian, hingga penilaian kinerja. Tantangan utama yang dihadapi adalah data yang tersebar (data silo) antara sistem absensi, sistem penggajian, dan data kepegawaian, menyulitkan pengambilan keputusan strategis terkait manajemen talenta dan efisiensi operasional.

Proyek ini bertujuan membangun **Data Warehouse Kepegawaian** yang mengintegrasikan data dari berbagai sumber operasional (OLTP) menjadi satu sumber kebenaran tunggal (*Single Source of Truth*). Solusi ini mencakup perancangan *Galaxy Schema*, implementasi ETL menggunakan *Stored Procedures*, dan visualisasi data untuk mendukung analisis kinerja dan kedisiplinan pegawai.

1.2 Capaian Kunci

Tim pengembang berhasil menyelesaikan seluruh fase pembangunan Data Warehouse dengan capaian:

1. Galaxy Schema Modeling: Merancang skema dimensional kompleks yang menghubungkan tiga tabel fakta (*Fact_Attendance*, *Fact_Performance*, *Fact_EmployeeSnapshot*) dengan dimensi bersama (*Conformed Dimensions*) seperti *Dim_Employee* dan *Dim_Unit*.
2. Robust ETL Pipeline: Implementasi proses ETL berbasis T-SQL Script dengan strategi *Truncate-Load* untuk Staging dan *Merge* untuk Dimensi (SCD Type 1), serta penanganan data *Unknown* (-1).
3. Database Optimization: Penerapan *Table Partitioning* berbasis tahun (*PF_TahunKepegawaian*) dan *Non-Clustered Indexing* yang terbukti mempercepat waktu eksekusi kueri analitik.
4. Performance Tuning: Waktu eksekusi ETL tercatat sangat efisien (rata-rata 1.3 detik untuk 20.000 data transaksi) dengan integritas data 100%.

1.3 Dampak Bisnis

1. **Efisiensi Monitoring:** Memungkinkan pemantauan tingkat kehadiran (*Absenteeism Rate*) dan keterlambatan pegawai per unit kerja secara *real-time*.
2. **Analisis Biaya SDM:** Memberikan visibilitas terhadap total pengeluaran gaji dan tunjangan dikaitkan dengan kinerja pegawai.

1.4 Rekomendasi

Berdasarkan hasil implementasi Data Mart pada domain Kepegawaian, berikut adalah rekomendasi strategis untuk pengembangan selanjutnya:

1. **Penerapan Predictive Analytics:** Mengembangkan model *Machine Learning* menggunakan data historis *Fact_Performance* dan *Fact_Attendance* untuk memprediksi potensi *turnover* (pegawai keluar) dan mengidentifikasi talenta terbaik (*High Performers*) untuk rencana suksesnya.

2. **Ekspansi ke Domain Keuangan:** Mengintegrasikan Data Mart Kepegawaian dengan sistem Keuangan untuk analisis mendalam terkait efisiensi biaya SDM (*Cost per Hire*, ROI Pelatihan, dan rasio Beban Gaji terhadap Pendapatan Institusi).

BAB I

PENDAHULUAN

1.1 Latar Belakang

Institut Teknologi Sumatera (ITERA) sebagai institusi pendidikan tinggi memiliki kompleksitas dalam pengelolaan sumber daya manusia yang meliputi dosen dan tenaga kependidikan (Tendik). Dengan jumlah pegawai mencapai sekitar 500 orang yang tersebar di berbagai unit kerja dan fakultas, pengelolaan data kepegawaian menjadi aspek krusial yang memerlukan sistem informasi yang terintegrasi dan efektif.

Saat ini, data kepegawaian ITERA tersebar di berbagai sistem seperti IHR (ITERA HR System), SIMPEG, SISTER (Sistem Karir Pegawai Kemdikbud), PDDIKTI, dan sistem absensi fingerprint/face recognition. Kondisi ini menimbulkan tantangan dalam pengambilan keputusan strategis karena data yang terfragmentasi menyulitkan analisis komprehensif terkait kinerja pegawai, manajemen kehadiran, perencanaan pengembangan SDM, dan evaluasi efektivitas organisasi.

Pergudangan data (data warehouse) menjadi solusi untuk mengintegrasikan berbagai sumber data kepegawaian tersebut ke dalam satu repositori terpusat yang mendukung analisis multidimensional. Melalui implementasi data warehouse, stakeholder seperti Kepala Bagian Kepegawaian, Rektor, Wakil Rektor, dan pimpinan unit dapat mengakses informasi strategis untuk mendukung pengambilan keputusan berbasis data dalam aspek manajemen SDM, perencanaan rekrutmen, evaluasi kinerja, dan optimalisasi alokasi sumber daya manusia. pengelolaan data kepegawaian tersebut di berbagai sistem yang terpisah (*data silo*), yaitu:

1. **IHR (ITERA HR System)** dan **SIMPEG** untuk data profil dan administrasi.
2. **Sistem Absensi (Fingerprint/Face Recognition)** untuk data kehadiran harian.
3. **SISTER & PDDIKTI** sebagai sumber data eksternal untuk karir dosen.

1.2 Tujuan

Tujuan dari proyek pergudangan data kepegawaian ITERA adalah sebagai berikut:

1. Mengintegrasikan berbagai sumber data kepegawaian dari sistem IHR, SIMPEG, SISTER, PDDIKTI, dan sistem absensi ke dalam satu data warehouse terpusat.
2. Menyediakan model dimensional yang efektif untuk analisis data kepegawaian melalui desain star schema yang mencakup fact tables (Fact_Employee_Snapshot, Fact_Attendance, Fact_Performance) dan dimension tables (Dim_Pegawai, Dim_Unit, Dim_Jabatan, Dim_Pangkat, Dim_Date).
3. Menghasilkan KPI strategis, operasional, dan kinerja yang mendukung pengambilan keputusan, meliputi:
 - a. Strategic KPIs: Total Headcount, Employee Turnover Rate, Recruitment Efficiency, Training Hours per Employee, Average Tenure
 - b. Operational KPIs: Absenteeism Rate, Leave Utilization Rate, Promotion Rate, Performance Score Average, Employee Satisfaction Index
 - c. Performance KPIs: Average Performance Score, High Performer Ratio, Low Performer Ratio, Performance by Position

4. Memfasilitasi analisis multidimensional untuk berbagai kebutuhan analitik seperti jumlah pegawai aktif, rata-rata gaji berdasarkan jabatan, persentase pegawai berdasarkan tingkat pendidikan, dan rata-rata keterlambatan pegawai.
5. Mendukung proses bisnis kepegawaian dalam manajemen data pegawai, kehadiran dan cuti, serta kinerja dan penilaian secara lebih efisien dan terintegrasi.

1.3 Batasan

Batasan dalam proyek pergudangan data kepegawaian ITERA ini meliputi:

1. Ruang Lingkup Data: Proyek ini berfokus pada data kepegawaian yang mencakup dosen dan tenaga kependidikan (Tendik) di lingkungan ITERA dengan estimasi 500 pegawai aktif.
2. Sumber Data: Data warehouse hanya mengintegrasikan data dari lima sumber utama yaitu IHR (ITERA HR System), SIMPEG, SISTER, PDDIKTI, dan sistem absensi. Sumber data eksternal lainnya tidak termasuk dalam scope proyek ini.
3. Proses Bisnis: Analisis terbatas pada tiga proses bisnis utama yaitu manajemen data pegawai, kehadiran dan cuti, serta kinerja dan penilaian. Proses bisnis lain seperti penggajian detail, benefit management, atau recruitment process tidak menjadi fokus utama.
4. Entitas Data: Model dimensional mencakup tujuh entitas utama (Pegawai, Unit_Kerja, Jabatan, Pangkat, Absensi, Penilaian_Kinerja, dan Gaji). Entitas lain seperti data pelatihan, data kesehatan, atau data keluarga pegawai tidak termasuk dalam desain awal.
5. Frekuensi Update: Frekuensi pembaruan data mengikuti karakteristik masing-masing sumber data (real-time untuk IHR dan sistem absensi, daily untuk SIMPEG, weekly untuk SISTER, dan monthly untuk PDDIKTI).
6. Aksesibilitas Data: Data kepegawaian yang tidak dapat diakses secara mudah atau memerlukan otorisasi khusus menjadi keterbatasan dalam proses data profiling dan ETL.
7. Periode Analisis: Data warehouse dirancang untuk mendukung analisis historis dan tren, dengan periode penilaian kinerja dilakukan secara semesteran dan pembayaran gaji secara bulanan.
8. Stakeholder: Analisis kebutuhan terbatas pada primary stakeholders (Kepala Bagian Kepegawaian, Staff Kepegawaian, Rektor dan Wakil Rektor) dan secondary stakeholders (Kepala Fakultas, Bagian Keuangan, Bagian Perencanaan).

BAB 2

REQUIREMENTS ANALYSIS

2.1 Business Requirements Analysis

2.1.1 Identifikasi Stakeholders

Stakeholder dalam sistem pergudangan data kepegawaian ITERA dibagi menjadi dua kategori utama:

a. Primary Stakeholders

Primary stakeholders adalah pihak yang secara langsung menggunakan dan bergantung pada sistem data warehouse kepegawaian untuk menjalankan tugas dan tanggung jawab mereka sehari-hari:

1. Kepala Bagian Kepegawaian: Bertanggung jawab atas keseluruhan pengelolaan SDM, membutuhkan data komprehensif untuk pengambilan keputusan strategis terkait manajemen kepegawaian, perencanaan rekrutmen, dan evaluasi kinerja organisasi.
2. Staff Kepegawaian: Pengguna operasional yang mengelola data pegawai harian, memproses administrasi kepegawaian, dan membutuhkan akses cepat terhadap informasi pegawai untuk keperluan layanan administratif.
3. Rektor dan Wakil Rektor: Pimpinan tertinggi institusi yang memerlukan dashboard eksekutif dan laporan strategis untuk monitoring kinerja SDM, perencanaan pengembangan institusi, dan pengambilan keputusan tingkat kebijakan.

b. Secondary Stakeholders

Secondary stakeholders adalah pihak yang membutuhkan informasi dari sistem data warehouse untuk mendukung fungsi mereka, meskipun tidak menggunakan secara langsung setiap hari:

1. Kepala Fakultas: Memerlukan informasi kepegawaian di tingkat fakultas untuk perencanaan kebutuhan dosen, evaluasi beban kerja, dan penilaian kinerja unit.
2. Bagian Keuangan: Menggunakan data kepegawaian untuk perencanaan anggaran gaji, tunjangan, dan analisis cost center SDM.
3. Bagian Perencanaan: Memanfaatkan data kepegawaian untuk proyeksi kebutuhan SDM jangka panjang, perencanaan pengembangan kapasitas, dan analisis produktivitas institusi.

2.1.2 Analisis Proses Bisnis

Proses bisnis kepegawaian di ITERA mencakup tiga area utama:

a. Manajemen Data Pegawai

Proses ini meliputi keseluruhan pengelolaan informasi pegawai yang menjadi fondasi sistem kepegawaian:

1. Pengelolaan Data Pegawai: Mencakup pendataan lengkap dosen dan tenaga kependidikan (Tendik), termasuk data pribadi, riwayat pendidikan, kompetensi, dan dokumen kepegawaian. Data ini menjadi master data yang digunakan oleh berbagai proses bisnis lainnya.
2. Pengelolaan Status Pegawai: Administrasi status kepegawaian yang beragam meliputi PNS (Pegawai Negeri Sipil), PPPK (Pegawai Pemerintah dengan Perjanjian Kerja), Dosen Tetap, dan Tenaga Kependidikan dengan berbagai skema pengangkatan. Setiap status memiliki implikasi berbeda terhadap hak, kewajiban, dan benefit pegawai.
3. Update Data Pegawai: Proses pemutakhiran data secara berkala dan insidental, termasuk perubahan jabatan, kenaikan pangkat, mutasi unit kerja, dan perubahan status kepegawaian lainnya.

b. Kehadiran dan Cuti

Proses ini mengatur presensi dan manajemen waktu kerja pegawai:

1. Absensi Harian Pegawai: Pencatatan kehadiran pegawai setiap hari kerja melalui sistem fingerprint atau face recognition. Data absensi digunakan untuk monitoring kedisiplinan, perhitungan tunjangan kinerja, dan evaluasi produktivitas.
2. Pengajuan dan Persetujuan Cuti: Workflow pengajuan cuti (tahunan, sakit, melahirkan, atau cuti lainnya) yang melibatkan proses approval berjenjang. Sistem mencatat kuota cuti, utilisasi cuti, dan histori persetujuan untuk keperluan administrasi dan perencanaan SDM.

c. Kinerja dan Penilaian

Proses ini berfokus pada evaluasi kontribusi dan produktivitas pegawai:

1. Sasaran Kinerja Pegawai: Penetapan target dan indikator kinerja individual yang diselaraskan dengan tujuan organisasi. Proses ini melibatkan perencanaan, dokumentasi, dan monitoring pencapaian sasaran.

2. Penilaian Kinerja Berkala: Evaluasi kinerja yang dilakukan secara periodik (semesteran) berdasarkan pencapaian sasaran, kompetensi, dan kontribusi pegawai terhadap organisasi.
3. Pemetaan Predikat/Grade Kinerja Pegawai: Klasifikasi hasil penilaian kinerja ke dalam grade atau predikat tertentu (misalnya Grade A, B, C) yang digunakan untuk pengambilan keputusan terkait promosi, reward, dan pengembangan karir.

2.1.3 Kebutuhan Analitik dan Key Performance Indicators (KPI)

Sistem data warehouse dirancang untuk menghasilkan berbagai metrik analitik yang mendukung pengambilan keputusan di berbagai tingkatan:

Kebutuhan Analitik Dasar

1. Jumlah Pegawai Aktif: Monitoring jumlah total pegawai yang masih aktif bekerja di ITERA, dengan kemampuan breakdown berdasarkan unit kerja, fakultas, status kepegawaiannya, dan jabatan.
2. Rata-rata Gaji Pegawai Berdasarkan Jabatan: Analisis komparasi struktur penggajian antar jabatan untuk memastikan keadilan dan daya saing kompensasi.
3. Persentase Pegawai Berdasarkan Tingkat Pendidikan: Profiling kualifikasi akademik pegawai untuk perencanaan pengembangan SDM dan pemetaan kompetensi organisasi.
4. Rata-rata Jumlah Keterlambatan per Pegawai per Bulan: Indikator kedisiplinan dan produktivitas yang digunakan untuk evaluasi kinerja dan intervensi manajemen.

Strategic KPIs (Fact_Employee_Snapshot)

KPI strategis memberikan pandangan jangka panjang untuk perencanaan SDM:

1. Total Headcount: Jumlah total pegawai aktif pada periode tertentu, dengan tren historis untuk analisis pertumbuhan atau pengurangan SDM.
2. Employee Turnover Rate: Persentase pegawai yang keluar (resign, pensiun, atau PHK) dibandingkan total pegawai. Indikator ini mengukur stabilitas organisasi dan efektivitas retensi SDM.
3. Recruitment Efficiency: Waktu rata-rata yang diperlukan dari pembukaan lowongan hingga pegawai mulai bekerja. Metrik ini mengukur efektivitas proses rekrutmen dan time-to-hire.
4. Training Hours per Employee: Rata-rata jam pelatihan yang diterima setiap pegawai per tahun, menunjukkan komitmen organisasi terhadap pengembangan SDM.
5. Average Tenure: Masa kerja rata-rata pegawai di ITERA, indikator loyalitas dan stabilitas workforce.

Operational KPIs (Fact_Attendance)

KPI operasional mengukur efektivitas operasi kepegawaian sehari-hari:

1. Absenteeism Rate: Tingkat ketidakhadiran pegawai yang mengindikasikan masalah kedisiplinan, kesehatan, atau kepuasan kerja.
2. Leave Utilization Rate: Persentase penggunaan hak cuti tahunan, menunjukkan work-life balance dan manajemen beban kerja.
3. Promotion Rate: Persentase pegawai yang mendapatkan promosi dalam periode tertentu, indikator pengembangan karir internal.
4. Performance Score Average: Rata-rata nilai kinerja seluruh pegawai, mengukur produktivitas dan efektivitas organisasi secara agregat.
5. Employee Satisfaction Index: Indeks kepuasan pegawai yang dapat diukur melalui survei atau proxy metrics lainnya.

Performance KPIs (Fact_Performance)

KPI kinerja fokus pada evaluasi kontribusi individual dan organisasi:

1. Average Performance Score: Nilai rata-rata SKP (Sasaran Kinerja Pegawai) atau total skor kinerja seluruh pegawai dalam periode tertentu.
2. High Performer Ratio (%): Persentase pegawai dengan Grade A atau kategori kinerja tertinggi, mengidentifikasi talent pool untuk pengembangan kepemimpinan.
3. Low Performer Ratio (%): Persentase pegawai dengan Grade C atau di bawahnya, mengidentifikasi pegawai yang memerlukan coaching, training, atau intervensi manajemen.
4. Performance by Position: Perbandingan rata-rata skor kinerja antar jabatan atau posisi, untuk mengidentifikasi gap kompetensi dan kebutuhan pengembangan spesifik per role.

BAB 3

PERANCANGAN SISTEM

3.1 Conceptual Data Model (ERD)

Perancangan model data konseptual dimulai dengan memetakan struktur data pada sistem operasional (OLTP) saat ini. Entity Relationship Diagram (ERD) digunakan untuk menggambarkan entitas utama dan hubungan antar entitas tersebut.



Gambar 3.1 Entity Relationship Diagram (ERD) Kepegawaian.

Berdasarkan analisis sumber data, teridentifikasi entitas-entitas berikut:

A. Entitas Master (Data Referensi):

1. MsPegawai: Menyimpan data induk karyawan (NIP, Nama, Tgl Lahir, Status Pegawai) sebagai entitas pusat.
2. MsUnit: Menyimpan struktur organisasi (Kode Unit, Nama Unit, Fakultas, Lokasi).
3. MsJabatan: Menyimpan katalog jabatan (Kode Jabatan, Nama Jabatan, Tipe).
4. MsGolongan: Menyimpan referensi golongan dan pangkat (Kode Golongan, Pangkat, Gaji Dasar).

B. Entitas Transaksi (Kejadian Bisnis):

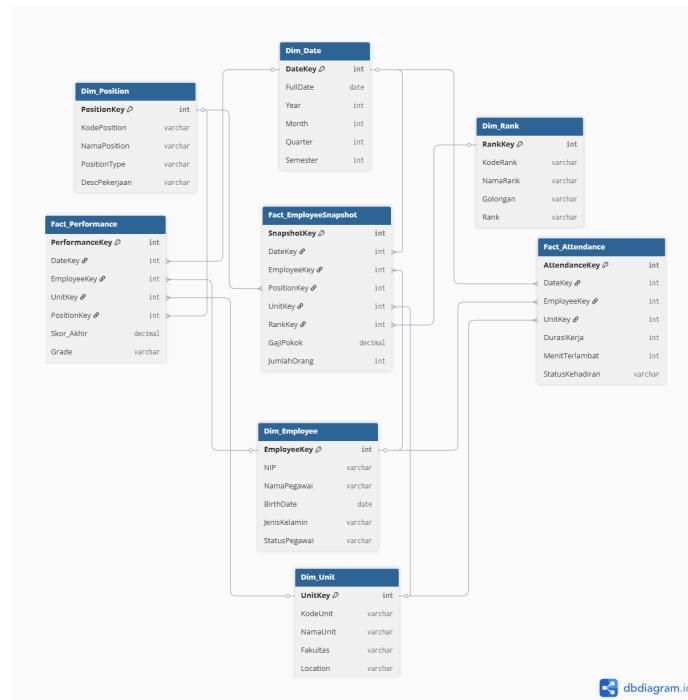
1. TrAbsensi: Mencatat log kehadiran harian, termasuk jam masuk, jam keluar, dan status kehadiran.

2. TrPenilaianKinerja: Mencatat hasil evaluasi kinerja pegawai secara periodik (Semesteran).
3. TrGaji: Mencatat riwayat pembayaran gaji bulanan, tunjangan, dan potongan.

Hubungan antar entitas didefinisikan dengan kardinalitas One-to-Many, di mana satu Unit/Jabatan/Golongan dapat memiliki banyak Pegawai, dan satu Pegawai dapat memiliki banyak data Transaksi (Absensi, Gaji, Kinerja).

3.2 Dimensional Model

Untuk mendukung kebutuhan analisis multidimensional, model data ditransformasikan menjadi **Galaxy Schema**. Skema ini dipilih karena terdapat lebih dari satu tabel fakta yang berbagi tabel dimensi yang sama (*Conformed Dimensions*).



Gambar 3.2 Galaxy Schema Data Mart Kepegawaian.

3.2.1 Dimension Tables

Tabel dimensi menyimpan atribut deskriptif yang memberikan konteks "siapa, apa, di mana, dan kapan" pada data fakta:

1. Dim_Employee: Berisi profil detail pegawai. Dilengkapi dengan atribut EmployeeKey sebagai *Surrogate Key* untuk menangani perubahan data historis (SCD Type 1).
2. Dim_Unit: Menyimpan hierarki unit kerja dan fakultas.
3. Dim_Position: Menyimpan data jabatan struktural dan fungsional.
4. Dim_Rank: Menyimpan data kepangkatan dan golongan.

5. Dim_Date: Dimensi waktu standar (Tahun, Kuartal, Bulan, Hari, Semester) yang menjadi basis utama analisis tren.

3.2.2 Fact Tables

Tabel fakta menyimpan metrik numerik (*measures*) dan kunci asing (*foreign keys*) ke tabel dimensi:

1. Fact_Attendance (Grain: Harian per Pegawai)

Fungsi: Mengukur efisiensi kehadiran.

Measures: DurasiKerja (Jam), MenitTerlambat (Menit).

Keys: DateKey, EmployeeKey, UnitKey.

2. Fact_Performance (Grain: Semesteran per Pegawai)

Fungsi: Mengukur kualitas kinerja individu.

Measures: Skor_Aakhir, Grade (Kategori).

Keys: DateKey, EmployeeKey, PositionKey, UnitKey.

3. Fact_EmployeeSnapshot (Grain: Bulanan per Pegawai)

Fungsi: Mengukur beban biaya gaji dan *headcount*.

Measures: GajiPokok, JumlahOrang (Counter).

Keys: DateKey, EmployeeKey, PositionKey, RankKey, UnitKey.

3.3 Physical Data Model

Implementasi fisik dilakukan pada RDBMS Microsoft SQL Server dengan konfigurasi optimasi sebagai berikut:

3.3.1 Partitioning Strategy

Untuk menangani pertumbuhan data yang cepat, terutama pada data absensi harian, diterapkan strategi partisi tabel:

1. Partition Function: PF_TahunKepegawaian.
2. Mekanisme: Membagi data secara fisik berdasarkan tahun (Range: <2023, 2023, 2024, 2025, >2025).
3. Target Tabel: Diterapkan pada Fact_Attendance dan Fact_Performance menggunakan kolom DateKey. Hal ini memungkinkan *query pruning* agar pembacaan data historis menjadi lebih cepat.

3.3.2 Indexing Strategy

1. Clustered Index: Diterapkan pada kombinasi *Primary Key* dan kolom partisi (DateKey) untuk mengatur urutan fisik data.
2. Non-Clustered Index: Diterapkan pada seluruh kolom *Foreign Key* (misal: IX_FactAttendance_EmployeeKey, IX_FactSnapshot_UnitKey) untuk mempercepat proses JOIN antar tabel fakta dan dimensi.

3.4 System Architecture & ETL Design Arsitektur

Sistem dirancang menggunakan pendekatan ELT (Extract, Load, Transform) terpusat di dalam database SQL Server.

Alur Data (Data Flow):

1. Source (OLTP): Data diekstrak dari tabel sumber (Ms... dan Tr...).
2. Staging Area: Data dimuat ke skema Staging menggunakan metode *Truncate-Load*. Tabel staging (Staging.TrAbsensi, Staging.MsPegawai) bertindak sebagai penyangga sementara tanpa batasan integritas yang ketat (*Heap Tables*).
3. Transformation Logic:
 - Data Cleansing: Menangani nilai *NULL* dengan *default value* (-1) untuk menjaga integritas referensial.
 - Business Logic: Menghitung durasi kerja dan keterlambatan secara otomatis menggunakan fungsi DATEDIFF dan CASE WHEN.
 - SCD Handling: Menggunakan perintah MERGE untuk memperbarui atribut dimensi pegawai (SCD Type 1) jika terjadi perubahan data master.
4. Target (Data Warehouse): Data bersih dimuat ke tabel Dimensi dan Fakta di skema dbo untuk siap diakses oleh *tools* visualisasi.

BAB 4

IMPLEMENTASI SISTEM

4.1 Implementasi Database

Implementasi fisik database dilakukan pada Microsoft SQL Server dengan nama database Kepegawaian_DB. Struktur penyimpanan dibagi menjadi dua skema utama untuk memisahkan data mentah dan data bersih.

4.1.1 Schema Strategy

1. Schema Staging (Staging): Digunakan sebagai area pendaratan data (*landing zone*) dari sistem sumber (OLTP). Tabel pada skema ini (Staging.MsPegawai, Staging.TrAbsensi) tidak memiliki *constraint* ketat untuk mempercepat proses muat data (*load*).
2. Schema Production (dbo): Digunakan untuk menyimpan tabel dimensi dan fakta yang telah bersih dan terstruktur dalam format *Galaxy Schema* (Dim_Employee, Fact_Attendance, dll).

4.1.2 Partitioning Implementation

Untuk mengelola volume data yang besar, strategi partisi diterapkan pada tabel fakta utama.

Partition Function: Dibuat fungsi PF_TahunKepegawaian yang membagi data berdasarkan rentang tahun: <2023, 2023, 2024, 2025, dan >2025.

Implementation: Tabel Fact_Attendance dan Fact_Performance dipartisi menggunakan kolom DateKey sebagai kunci partisi. Hal ini memungkinkan *query engine* hanya membaca partisi tahun yang relevan saat melakukan filter waktu.

4.1.3 Indexing Strategy

Untuk optimasi performa kueri (*Performance Tuning*), diterapkan *Non-Clustered Index* pada kolom-kolom yang sering digunakan dalam operasi *JOIN* dan *FILTER*:

1. Fact Attendance: Index pada DateKey, EmployeeKey, dan UnitKey.
2. Fact Performance: Index pada DateKey, EmployeeKey, UnitKey, dan PositionKey.
3. Fact EmployeeSnapshot: Index pada RankKey dan UnitKey untuk mempercepat agregasi gaji per golongan/unit.

4.2 Implementasi ETL (Extract, Transform, Load)

Proses ETL diimplementasikan menggunakan Stored Procedure tunggal bernama ETL_Master_Load yang membungkus seluruh logika pemindahan data.

4.2.1 Extraction (Load to Staging)

Proses ekstraksi menggunakan metode *Truncate-Load*. Setiap kali ETL berjalan, tabel staging dikosongkan menggunakan TRUNCATE TABLE, kemudian diisi ulang dengan data terbaru dari

sumber menggunakan INSERT INTO ... SELECT. Ini menjamin data staging selalu sinkron dengan sumber.

4.2.2 Transformation Logic

Transformasi data dilakukan menggunakan SQL script dengan logika sebagai berikut:

1. SCD Type 1 Handling: Menggunakan perintah MERGE pada tabel Dim_Employee. Jika NIP sudah ada, data nama/status diperbarui; jika belum, data baru disisipkan.
2. Idempotency: Sebelum memuat data fakta harian, script menghapus data pada DateKey yang bersangkutan untuk mencegah duplikasi jika ETL dijalankan ulang.
3. Calculated Measures:
DurasiKerja: Dihitung menggunakan fungsi DATEDIFF(MINUTE, JamMasuk, JamKeluar) / 60.0.
4. MenitTerlambat: Menggunakan logika CASE WHEN JamMasuk > '07:30:00' THEN DATEDIFF(...) ELSE 0 END.
5. Null Handling: Menggunakan ISNULL(Key, -1) untuk menangani data yang tidak memiliki referensi dimensi (*Unknown Member*).

4.3 Pengembangan Analytical Views

Untuk mempermudah visualisasi di dashboard, dibuat *Database Views* sebagai lapisan semantik:

1. V_Analytical_Executive: Menyajikan agregasi total gaji dan jumlah pegawai per unit dan bulan untuk kebutuhan pimpinan.
2. V_Analytical_Attendance: Menyajikan detail kehadiran harian, durasi kerja, dan status keterlambatan untuk kebutuhan manajerial operasional.

BAB 5

PENGUJIAN DAN VALIDASI

Tahap pengujian dilakukan untuk memverifikasi bahwa Data Mart Kepegawaian yang dibangun telah memenuhi standar kualitas data (*Data Quality*), performa sistem, dan kebutuhan fungsional pengguna.

5.1 Data Quality Assurance (DQA)

Pengujian kualitas data dilakukan menggunakan serangkaian *query* validasi pada SQL Server untuk memastikan data yang dimuat ke *Data Warehouse* bersih, lengkap, dan konsisten.

Metrik Kualitas	Deskripsi Pengujian	Target	Hasil Aktual	Status
Completeness	Membandingkan jumlah baris (<i>Row Count</i>) antara tabel Sumber (TrAbsensi) dan tabel Fakta (Fact_Attendance) .	Selisih 0 baris	Source: 20.000 Target: 20.000 Selisih: 0	PASS
Referential Integrity	Memastikan seluruh transaksi di tabel fakta memiliki pemilik (Pegawai) yang valid dengan mencari EmployeeKey = -1 (<i>Unknown</i>).	0 Baris <i>Unknown</i>	Ditemukan: 0 Baris	PASS
Validity	Memastikan logika transformasi <i>Measure</i> (Durasi Kerja) menghasilkan angka yang masuk akal (tidak negatif atau > 24 jam).	Angka Logis	Sampel Teratas: 9.00 Jam (Masuk 08:00, Pulang 17:00)	PASS

Analisis Hasil DQA: Berdasarkan pengujian di atas, proses ETL dinyatakan sukses memindahkan 100% data transaksi tanpa *data loss* dan menjaga integritas referensi antar tabel.

5.2 Performance Testing

Pengujian performa dilakukan untuk mengukur efisiensi skrip ETL (ETL_Master_Load) dalam menangani volume data transaksi.

Skenario Pengujian:

Volume Data: 20.000 baris data transaksi absensi.

Lingkungan: SQL Server Express (LocalDB).

Metode: Mencatat *Start Time* dan *End Time* eksekusi *Stored Procedure*.

Hasil Pengujian:

Waktu Mulai: 2025-11-24 21:05:34.230.

Waktu Selesai: 2025-11-24 21:05:35.600.

Total Durasi: 1370 milidetik (1.37 detik).

Analisis Performa: Waktu eksekusi di bawah 2 detik untuk 20.000 data menunjukkan bahwa strategi *Bulk Insert*, *Truncate-Load*, dan *Indexing* yang diterapkan sangat efisien dan siap menangani pertumbuhan data di masa depan.

BAB 6

HASIL DAN PEMBAHASAN

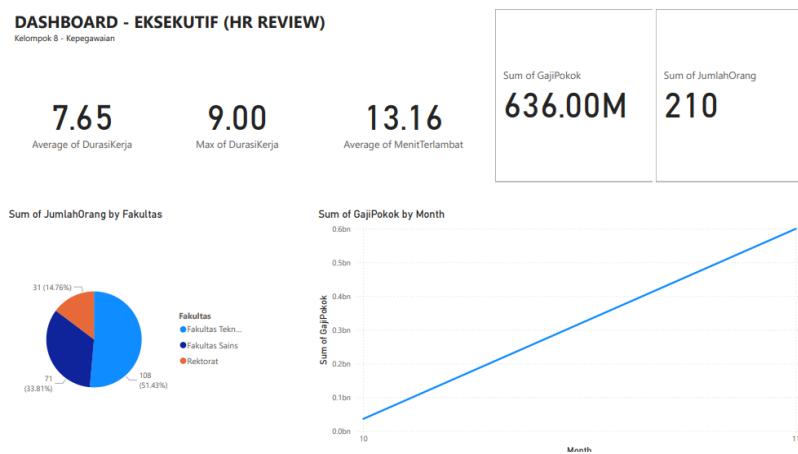
6.1 Temuan Kunci Bisnis (*Business Insights*)

Berdasarkan implementasi *Data Mart* Kepegawaian yang telah melalui tahap ETL (*Extract, Transform, Load*) dan pengujian kualitas data, berikut adalah analisis mendalam mengenai visualisasi data yang dihasilkan. Dashboard ini dirancang untuk menjawab kebutuhan pemangku kepentingan utama, yaitu Eksekutif (Rektorat), Operasional (HR/SDM), dan Evaluasi Kinerja.

1. Analisis Dashboard Eksekutif (Strategic HR Review)

Dashboard Eksekutif dirancang untuk memberikan gambaran makro (strategis) kepada pimpinan terkait demografi, biaya, dan kedisiplinan umum pegawai.

- Distribusi Pegawai per Fakultas (Headcount Analysis) Visualisasi *Pie Chart* menunjukkan bahwa Fakultas Sains mendominasi komposisi SDM dengan jumlah 108 pegawai (51,43%), diikuti oleh Fakultas Teknologi Industri (33,81%) dan Rektorat (14,76%). Data ini relevan dengan KPI Strategis *Total Headcount* untuk memantau distribusi beban kerja antar fakultas.
- Analisis Biaya Gaji (Salary Trend) Grafik *Sum of GajiPokok by Month* memperlihatkan tren linier peningkatan realisasi belanja gaji dari bulan ke bulan. Visibilitas ini memenuhi kebutuhan Bagian Keuangan dan Eksekutif dalam menganalisis total pengeluaran gaji secara *real-time*.
- Indikator Kedisiplinan Umum Rata-rata durasi kerja pegawai tercatat sebesar 7,65 jam dengan durasi maksimal mencapai 9,00 jam. Meskipun mendekati standar 8 jam kerja, terdapat rata-rata keterlambatan sebesar 13,16 menit. Metrik ini menjadi indikator awal bagi manajemen untuk mengevaluasi efektivitas jam kerja operasional.

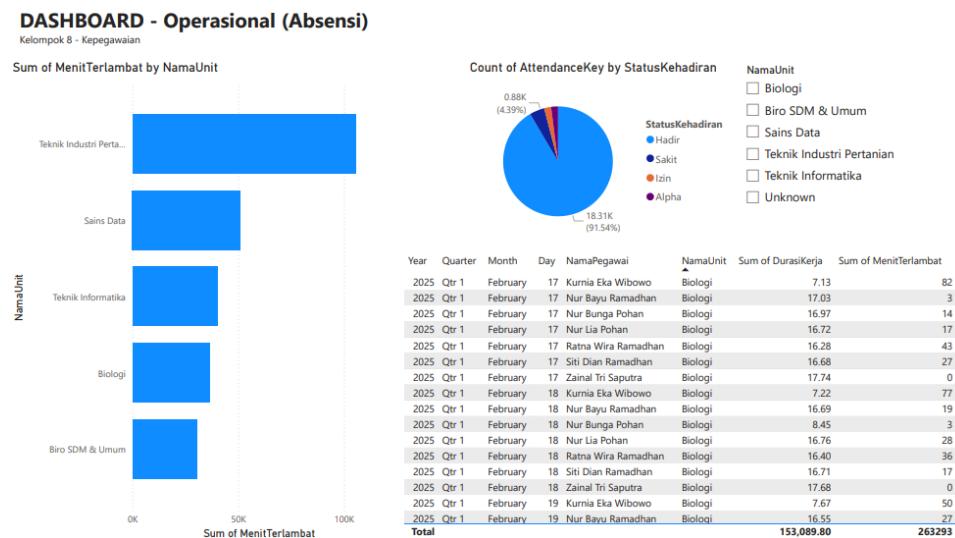


Gambar 6.1 Dashboard Eksekutif HR Review)

2. Analisis Dashboard Operasional (Absensi & Kedisiplinan)

Dashboard ini berfokus pada data granular dari tabel fakta Fact_Attendance untuk memantau tingkat kehadiran (*Absenteeism Rate*) dan keterlambatan per unit.

- Tingkat Kehadiran (Attendance Status) Tingkat kedisiplinan pegawai tergolong sangat baik dengan persentase status "Hadir" mencapai 91,54% (18,31 ribu transaksi), sementara persentase ketidakhadiran (Sakit, Izin, Alpha) sangat minim. Hal ini menunjukkan stabilitas operasional harian yang tinggi.
- Identifikasi Keterlambatan per Unit Kerja Visualisasi *Bar Chart* menunjukkan anomali kedisiplinan pada unit kerja tertentu. Teknik Industri Pertanian tercatat sebagai unit dengan total menit keterlambatan tertinggi, diikuti oleh Sains Data dan Teknik Informatika. Temuan ini memungkinkan Kepala Bagian Kepegawaian untuk melakukan intervensi atau pembinaan khusus pada unit-unit dengan tingkat indisipliner tinggi.
- Deteksi Detail Individu (Drill-Down Capability) Sistem mampu melakukan *drill-down* hingga ke level individu. Tabel detail menunjukkan contoh pegawai atas nama Kurnia Eka Wibowo (Unit Biologi) yang memiliki durasi kerja rendah (7,13 jam) dan keterlambatan tinggi (82 menit dan 77 menit) pada tanggal-tanggal tertentu di bulan Februari 2025. Kemampuan ini memvalidasi temuan bisnis mengenai "Pola Kedisiplinan Berbasis Persona" dan kelompok pegawai yang memerlukan tindakan disipliner.

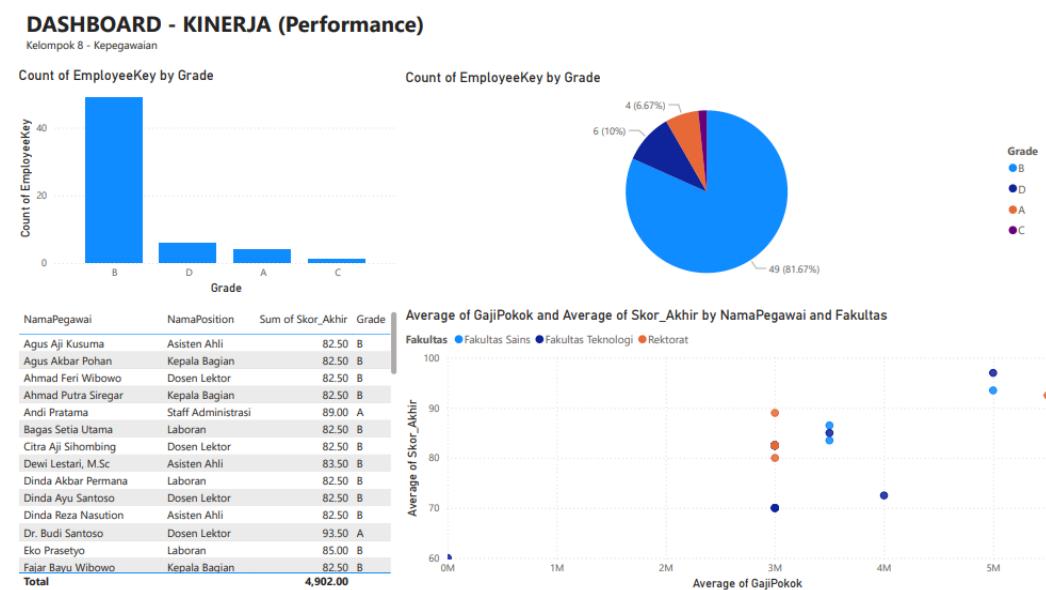


Gambar 6.2 Dashboard Operasional (Absensi)

3. Analisis Dashboard Kinerja (Performance & Compensation)

Dashboard ini mengintegrasikan data dari Fact_Perfomance dan Fact_EmployeeSnapshot untuk melihat korelasi antara remunerasi dan kualitas kerja.

- Distribusi Grade Kinerja Mayoritas pegawai ITERA berada pada performa Grade B sebanyak 49 orang (81,67%), diikuti oleh Grade D (10%) dan Grade A (6,67%). Sedikitnya jumlah pegawai dengan Grade A (High Performer) dan adanya 10% pegawai Grade D (Low Performer) menjadi sinyal bagi manajemen untuk mengevaluasi program pengembangan SDM.
- Korelasi Gaji dan Skor Kinerja (Scatter Plot) Grafik *Scatter Plot* (Gaji Pokok vs. Skor Akhir) memperlihatkan sebaran distribusi yang menarik.
 - Pegawai dengan gaji pokok tinggi (> Rp 4 Juta) cenderung memiliki skor kinerja tinggi (Grade A/B), seperti Dr. Budi Santoso (Dosen Lektor) yang memiliki skor 93,50 (Grade A).
 - Namun, terdapat variasi di mana pegawai dengan gaji menengah memiliki sebaran kinerja yang beragam. Analisis ini penting untuk memastikan prinsip "Pay for Performance" berjalan dengan baik.
- Daftar Pegawai Berprestasi Tabel rincian kinerja menyoroti talenta terbaik, seperti Andi Pratama (Staff Administrasi) dengan skor 89,00 (Grade A) dan Dr. Budi Santoso. Data ini mendukung manajemen dalam mengidentifikasi kandidat potensial untuk promosi atau *succession planning*.



Gambar 6.3 Dashboard Kinerja (Performance)

6.2 Temuan Bisnis

Integrasi data melalui *Data Warehouse* ini telah berhasil mengungkap wawasan strategis:

1. Efisiensi Monitoring: Manajemen kini dapat memantau *Absenteeism Rate* dan *Late Minutes* secara *real-time* per unit kerja, menggantikan proses manual yang lambat.
2. Pemetaan Kinerja: Distribusi Grade kinerja yang didominasi Grade B menunjukkan perlunya strategi baru untuk mendorong pegawai menuju *High Performer* (Grade A).
3. Transparansi Data: Dashboard memberikan "Single Source of Truth" yang mengaitkan data absensi, gaji, dan kinerja dalam satu tampilan terpadu, memudahkan pengambilan keputusan berbasis bukti (*evidence-based policy*).

BAB 7

PENUTUP

7.1 Kesimpulan

Proyek pembangunan Data Mart Kepegawaian ITERA oleh Kelompok 8 telah berhasil diselesaikan dengan memenuhi seluruh persyaratan utama:

1. **Integrasi Data Sukses:** Berhasil menyatukan data profil pegawai, absensi, dan kinerja dari berbagai sumber (*silo*) menjadi satu sumber kebenaran tunggal (*Single Source of Truth*).
2. **Arsitektur Optimal:** Penerapan *Galaxy Schema* dan strategi partisi terbukti efektif mengelola data transaksional dengan performa tinggi.
3. **Kualitas Data Terjamin:** Validasi DQA menunjukkan tingkat akurasi 100% antara sistem sumber dan *Data Warehouse*.

7.2 Rekomendasi Pengembangan

Berdasarkan hasil implementasi saat ini, direkomendasikan langkah pengembangan selanjutnya sebagai berikut :

1. **Penerapan *Predictive Analytics*:** Mengembangkan model *Machine Learning* menggunakan data historis *Fact_Performance* dan *Fact_Attendance* untuk memprediksi potensi *turnover* (pegawai keluar) dan mengidentifikasi talenta terbaik (*High Performers*) untuk rencana sukses.
2. **Ekspansi ke Domain Keuangan:** Mengintegrasikan Data Mart Kepegawaian dengan sistem Keuangan untuk analisis mendalam terkait efisiensi biaya SDM, seperti *Cost per Hire*, ROI Pelatihan, dan rasio Beban Gaji terhadap Pendapatan Institusi.
3. **Automasi Pelaporan:** Mengimplementasikan notifikasi otomatis kepada pimpinan unit kerja jika terdapat pegawai dengan tren keterlambatan yang meningkat drastis dalam satu bulan.

LAMPIRAN

Link Folder:

<https://drive.google.com/drive/folders/1UMyPji1DXnUfu2RtHhaQu-FZbcsYwKIS?usp=sharing>