

ETL Implementation for Centralizing Academic Data in an Educational Bootcamp

Implementasi ETL untuk Sentralisasi Data Akademik Perusahaan Bootcamp

Nathania Santa Nigel Simbolon¹, Didik Kurniawan², Rahman Taufik³

^{1,2,3} Ilmu Komputer, Universitas Lampung, Indonesia

¹nathania.santa21@students.unila.ac.id, ^{2*}didikunila@gmail.com, ³
rahman.taufik@fmipa.unila.ac.id

*: Penulis korespondensi (corresponding author)

Informasi Artikel

Received: December 2025

Revised: January 2026

Accepted: January 2026

Published: February 2026

Abstract

Inefficiencies in decentralized data management have become a critical issue for bootcamp companies. This study explores the implementation of an Extract, Transform, Load (ETL) process to centralize academic data at PT. Hacktivate Teknologi Indonesia (Hacktiv8) in order to enhance learning process efficiency. With the rapid growth of disorganized academic data, an integrated automation system is required to monitor and evaluate bootcamp participants' progress. The ETL process employed in this study collects data from various Google Spreadsheets, transforms it into a normalized structure, and stores it in a centralized data warehouse using Google Cloud Platform. The research follows the 4-Steps-Kimball methodology to design the data warehouse schema and evaluates data quality based on the ISO/IEC 25012 standard. Data quality assessment is conducted through two approaches: (1) technical validation using Great Expectations shows key attribute accuracy of 98.44%, consistency of 97%, and completeness of 100%, despite 5–7% data loss due to the removal of null values in instructor columns to maintain referential integrity; (2) stakeholder feedback (N=10) collected via a Likert scale (1–5) yields an average score of 4/5, with the highest ratings in data consistency and centralized structure. The evaluation results indicate that the implemented ETL system is highly efficient and positively received by users, although suggestions were made for improvements in system documentation and adaptability. This study is expected to

contribute significantly to data-driven decision-making in non-formal education environments.

Keywords: Data warehouse, ETL, Google Cloud Platform; ISO/IEC 25012

Kata kunci: Data warehouse, ETL, Google Cloud Platform, ISO/IEC 25012

Abstrak

Inefisiensi pengelolaan data yang tidak terpusat menjadi masalah pada Perusahaan *bootcamp*. Penelitian ini membahas implementasi Extract, Transform, Load (ETL) pada sentralisasi data akademik di PT. Hacktivate Teknologi Indonesia (Hacktiv8) untuk meningkatkan efisiensi proses pembelajaran. Pertambahan data yang pesat dalam data akademik yang tidak terorganisir, diperlukan sistem otomatis yang terintegrasi untuk memonitor dan mengevaluasi perkembangan peserta *bootcamp*. Proses ETL yang diterapkan mengumpulkan data dari berbagai Google Spreadsheet, mentransformasikannya menjadi bentuk normal, dan menyimpannya dalam data warehouse yang terpusat menggunakan Google Cloud Platform. Penelitian ini menggunakan metodologi 4-steps-Kimball untuk mendesain skema data warehouse dan melakukan pengujian kualitas data berdasarkan standar ISO/IEC 25012. Evaluasi kualitas data mengadopsi standar ISO/IEC 25012 melalui dua pendekatan: (1) validasi teknis dengan Great Expectations menunjukkan akurasi atribut kunci 98,44%, konsistensi 97%, dan kelengkapan 100%, meskipun 5-7% data hilang akibat penghapusan nilai null pada kolom instruktur demi menjaga integritas referensial; (2) kuesioner stakeholder (N=10) menggunakan skala Likert 1-5 menghasilkan skor rata-rata 4/5, dengan aspek konsistensi data dan struktur terpusat meraih nilai tertinggi. Hasil evaluasi menunjukkan bahwa sistem ETL yang diimplementasikan sangat efisien, dengan penilaian positif dari pengguna, meskipun terdapat beberapa saran untuk peningkatan, terutama dalam hal dokumentasi dan adaptabilitas sistem. Penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam proses pengambilan keputusan berbasis data di environment pendidikan non-formal.

1. Pendahuluan

Peran teknologi informasi dalam meningkatkan efisiensi dan efektivitas birokrasi maupun dunia usaha telah menjadi fokus penelitian terkini [1]. Pengembangan birokrasi digital, melalui pemanfaatan teknologi informasi dan pengolahan data, dinilai sebagai langkah strategis untuk memperkuat sistem administrasi [2]. Data, sebagai elemen kunci dalam

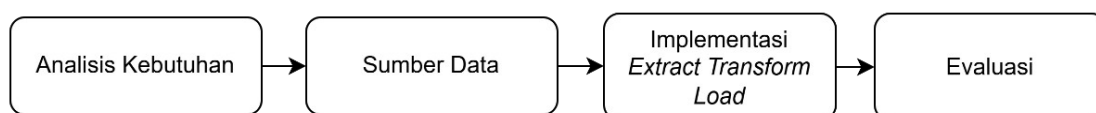
teknologi informasi, tidak hanya mendukung pengambilan keputusan bisnis tetapi juga berperan dalam analisis multidimensi untuk mengekstrak nilai informasinya [3]. Aktivitas analisis data ini tidak terbatas pada organisasi komersial, melainkan juga relevan di sektor pendidikan, termasuk dalam konteks evaluasi proses pembelajaran [4].

Hactivate Teknologi Indonesia (Hacktiv8) merupakan penyelenggara *bootcamp* pendidikan non-formal, menghadapi masalah dalam pengelolaan data akademik program Full-Time Data Science (FTDS). Data siswa dan aktivitas pembelajaran saat ini tersebar di berbagai Google Spreadsheet yang disimpan pada Google Drive berdasarkan folder angkatan dan lokasi cabang, mengakibatkan proses pelaporan, pemantauan, dan evaluasi menjadi tidak efisien [5]. Metode manual yang digunakan rentan terhadap human error dan menghambat analisis mendalam untuk pengambilan keputusan berbasis data. Sentralisasi data melalui proses Extract, Transform, Load (ETL) menjadi solusi krusial untuk mengintegrasikan sumber data terpisah ke dalam data warehouse yang terstruktur, sehingga memudahkan analisis dan mengurangi redundansi [6]. Berdasarkan penelitian Berliantara, implementasi ETL juga memerlukan penjadwalan yang efektif guna memastikan pembaruan data real-time, khususnya di lingkungan pendidikan yang membutuhkan respons cepat [7]. Studi Kabakchieva (2015) menunjukkan bahwa penerapan Business Intelligence (BI) mampu menyajikan indikator statistik dan prediksi untuk meningkatkan kualitas pembelajaran[8]. Sementara Broman dan Woo menekankan pentingnya penyimpanan data yang akurat dalam mengurangi kesalahan analisis [9].

Penelitian ini berfokus mengimplementasikan proses Extract, Transform, Load (ETL) dalam sentralisasi data untuk mengintegrasikan informasi siswa dan aktivitas akademik yang tersimpan dalam spreadsheet menggunakan pendekatan Kimball dan konsep data warehouse di Google Cloud Platform[10]. Data warehouse adalah database relasional yang dibangun khusus dan berjalan pada perangkat keras khusus baik di tempat atau di cloud [11]. Dengan menerapkan metodologi ini, diharapkan dapat meningkatkan efisiensi dalam penyediaan data yang akurat dan real-time, serta mempermudah pemantauan performa siswa dalam program *bootcamp* melalui data yang tersentralisasi.

2. Metode/Perancangan

Tahapan metode dalam penelitian ini disajikan dalam bentuk diagram pada gambar 1 di bawah ini.



Gambar 1. Flow Diagram Tahapan Penelitian

2.1. Analisis Kebutuhan

Tahapan analisis kebutuhan dalam penelitian ini dilakukan menggunakan observasi. Proses ini mencakup pemilihan, pengecekan, dan identifikasi jenis dan struktur data yang diperlukan untuk mengetahui jenis dan jenis analisis yang diperlukan. Observasi dilakukan untuk mengidentifikasi detail kebutuhan data untuk kebutuhan monitoring laporan analisis proses

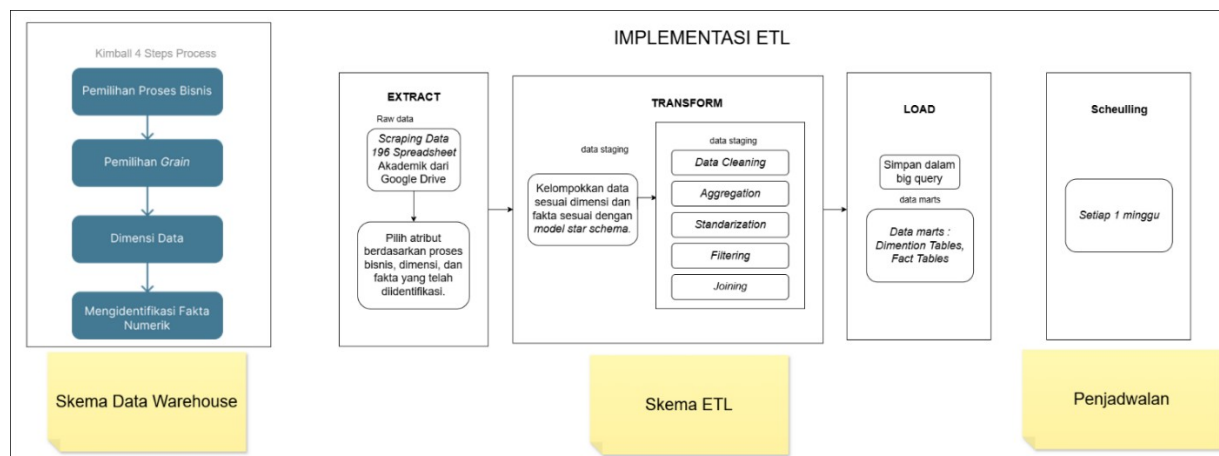
akademik di *bootcamp*, meliputi pengelolaan sheet penilaian, pembagian buddy untuk setiap murid, log *mentoring*, jadwal *bootcamp*, dan jadwal terkait dengan aktivitas peserta maupun instruktur.

2.2. Sumber Data

Sumber data diperoleh dari operasional akademik Full Time Data Science Hacktiv8. Data tersebut merupakan data aktivitas peserta *bootcamp* dari Tahun 2020-2024. Data meliputi penilaian pembelajaran peserta *bootcamp* yang terdapat pada spreadsheet dari salah satu sheet bernama 'cummulative'.

2.3. Implementasi ETL

Pada subbab ini dibahas mengenai implementasi proses ETL (Extract, Transform, Load) yang dilakukan dalam penelitian ini. Implementasi tersebut mencakup perancangan skema *data mart* di data warehouse, proses ETL itu sendiri, serta penjadwalan proses ETL agar dapat berjalan secara otomatis dan terjadwal. Pengumpulan data pada penelitian ini dilanjutkan menggunakan *four-step Kimball design process* untuk pengembangan data warehouse, seperti pada gambar 2.



Gambar 2. Diagram Implementasi ETL

2.3.1. Skema Data Warehouse

Penelitian ini menggunakan pendekatan *bottom-up Kimball* melalui metodologi empat langkah yang terdiri dari pemilihan proses bisnis, penentuan grain tabel fakta, pemilihan dimensi seperti waktu dan tanggal, serta identifikasi fakta numerik yang konsisten dengan granularitas yang telah dipilih. Proses perancangan data warehouse mengikuti alur *four-step Kimball Methodology* yang bertujuan untuk memastikan struktur model multidimensi yang terukur, sistematis, dan relevan dalam mendukung pengambilan keputusan berbasis data [4]. Tahap pertama dimulai dengan memilih proses bisnis yang akan dimodelkan, dengan mempertimbangkan tujuan analisis dan ketersediaan data. Selanjutnya, proses dilanjutkan dengan menentukan grain, yaitu tingkat detail representasi dari tabel fakta yang akan merepresentasikan pengukuran proses bisnis yang dipilih. Setelah itu, ditentukan dimensi-dimensi yang akan mendeskripsikan pengukuran dalam tabel fakta, seperti dimensi waktu dan tanggal yang umum digunakan. Selanjutnya, proses yang dilakukan adalah identifikasi terhadap fakta numerik, yaitu ukuran-ukuran yang akan dimasukkan dalam tabel fakta, yang harus selaras dan konsisten dengan tingkat granularitas yang telah ditentukan.

2.3.2. Skema ETL dan Penjadwalan

Proses ekstraksi data akademik dilakukan secara otomatis menggunakan script Python untuk mengakses sumber data berbentuk spreadsheet berada di Google Drive “Academic Operations”, dengan nama *worksheet* pada spreadsheet penilaian harian (*Cummulative*), *Mentoring Log*, spreadsheet *Teachers Simulation*. Format spreadsheet berbeda-beda tergantung angkatan program *bootcamp* tersebut.

Data Science Fulltime Program Phase 1 Mentoring Log						
No	Nama	Buddy	Week 1		Week 2	
			Mentoring 1	Mentoring 2	Mentoring 1	Mentoring 2
1	Ahmad Dani Rifai	Elita				
2	Gracia Valerine	Elita				
3	Muhammad Fiqih Al-ayubi	Elita				
4	Ryan Trisnadi	Elita				
5	Syihabuddin Ahmad Satisma	Rachman				
6	Vicky Belario	Rachman				
7	Vincent Kaunang	Rachman				
8		0	0			
9		0	0			
10		0	0			
11		0	0			

Gambar 3. Worksheet Mentoring Log

Gambar 3 merupakan contoh raw data dari spreadsheet yang menjadi data source untuk beberapa *data mart* yaitu fakta siswa *batch* lama.

Data Science Fulltime Program Phase 1 Attendance												
Nama	Buddy	Week 1						Week 2				
		Day 3		Day 4		Day 5		Day 1		Day 2		Day 3
		AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	PM
Ahmad Dani Rifai	Elita	H	H	H	H	H			H	H	H	
Gracia Valerine	Elita	H	H	H	H	H			H	H	H	
Muhammad Fiqih Al-ayubi	Elita	H	H	A	H	H			H	H	H	
Ryan Trisnadi	Elita			S	S	H			H	H	H	
Syihabuddin Ahmad Satisma	Rachman	H	H	H	H	H			H	H	H	
Vicky Belario	Rachman	H	H	H	H	H			H	H	H	
Vincent Kaunang	Rachman	H	H	H	H	I			H	H	H	
	0		0									

Gambar 4. Worksheet Absensi

Pada *worksheet* ‘absensi’ akan dilakukan *scraping* menggunakan script python *new_absensi.py* untuk dijalankan extract absensi peserta current *batch* untuk jenis kehadiran kelas biasa (kelas teknikal). Selanjut akan dilakukan *scraping* untuk jenis kehadiran. Kolom yang digunakan untuk ekstrasi data kehadiran *mentoring* adalah Nama, Buddy, Week 1-4, *Mentoring* 1-2 setiap minggunya.

	A	B	C	D	E	F	G	H	I	J	K
300											
301											
302	Week 1										
303	Monday, May 13, 2024		Tuesday, May 14, 2024		Wednesday, May 15, 2024		Thursday, May 16, 2024		Friday, May 17, 2024		
304	Morning Afternoon		Morning Afternoon		Morning Afternoon		Morning Afternoon		Morning Afternoon		
305	Batch 17 HCK Phase 0		Batch 16 HCK Phase 1		Batch 15 HCK Phase 2		Batch 14 HCK Phase 0		Batch 13 HCK Phase 1		Batch 12 HCK Phase 2
306	Hana Hana		Riki Riki		Yuda Yuda		Lis Lis		Hana Hana		Elita
307	Lis Lis		Elita Elita		Riki Riki		Hana Hana		Rachman Rachman		Irfan
308	Irfan Irfan		Lis Lis		Elita Elita		Rachman Rachman		Yuda Yuda		Bach
309	Yuda Yuda		Hana Hana		Danu Danu		Elita Elita		Irfan Irfan		Dar
310	Elita Elita		Riki Riki		Vincent Vincent		Vincent Vincent		Ali Ali		Vinc
311	Phase 1		Phase 2		Phase 0		Phase 1		Phase 2		Phase 0
312	Ali Ali		Yosef Yosef		Yosef Yosef		Vincent Vincent		Vincent Vincent		Ali
313	Batch 5 BSD Phase 2		Batch 4 SBY Phase 1		Batch 3 SBY Phase 0		Batch 2 SBY Phase 2		Batch 1 SBY Phase 1		Batch 0 SBY Phase 0
314	Dyco Dyco		Yosef Yosef		Dyco Dyco		Dyco Dyco		Dyco Dyco		Dyco
315	Yosef Yosef		Elita Elita		Fahmi Fahmi		Hana Hana		Yosef Yosef		Dyco
316	Standby Jira		S3B Offline		S3B Online		S3B Offline		S3B Online		S3B Offline
317	Riki Riki		Lis Lis		Yosef Yosef		Ali Ali				
318	HCK HCK		SBY SBY		BSD BSD						
319	Yosef Yosef		Ali Ali								
320											
321	Week 2										
322	Monday, May 20, 2024		Tuesday, May 21, 2024		Wednesday, May 22, 2024		Thursday, May 23, 2024		Friday, May 24, 2024		
323	Morning Afternoon		Morning Afternoon		Morning Afternoon		Morning Afternoon		Morning Afternoon		
324	Batch 17 HCK Phase 0		Batch 16 HCK Phase 1		Batch 15 HCK Phase 2		Batch 14 HCK Phase 0		Batch 13 HCK Phase 1		Batch 12 HCK Phase 2
325	Riki Riki		Yuda Yuda		Riki Riki		Lis Lis		Lis Lis		Irfan
326	Hana Hana		Riki Riki		Elita Elita		Lis Lis		Lis Lis		Elita
327	Yuda Yuda		Irfan Irfan		Yuda Yuda		Lis Lis		Lis Lis		Bach

Gambar 5. Worksheet Teachers Simulations

Proses extract dilakukan dengan cara *scraping* tiap cell pada *worksheet* pada gambar 5 dengan mengambil nilai tiap cell yang berisi nama instruktur, *week*, *batch*, *location*, yang berada pada kolom A-B. Proses yang sama menggunakan script python. Python adalah bahasa pemrograman serbaguna yang dapat dijalankan di hampir semua tempat, ia cukup fleksibel dan gesit untuk memberikan kemampuan dalam menciptakan solusi yang sangat khusus untuk mengatasi masalah unik yang terkait dengan big data[12]

Transformasi yang dilakukan adalah membersihkan data dari nilai null dengan mengubah data ke tipe yang tepat. Salah satu contoh transformasi ini adalah mengubah string data yang diambil menjadi tipe numerik dan tanggal. Bagian dari proses standarisasi adalah mengubah atribut tanggal menjadi format yyyy/mm/dd. Ini juga mencakup menempatkan catatan dalam format standar dan mengurutkan data kehadiran pada tanggal tertentu (*filtering*). Selanjutnya, untuk menghitung total persen dari skor penilaian, beberapa kolom digabungkan, seperti yang terlihat dalam tabel *batch* saat ini. Untuk atribut "nama", terutama, huruf pertama setiap kata harus diubah menjadi huruf kapital. Skema data dapat diubah dan normalisasi diterapkan pada tahap transformasi berikutnya.

Pada tahapan load, semua data yang akan diproses untuk kebutuhan *data mart*, penyimpanan data operasional, data warehouse perusahaan akan disimpan ke menggunakan layanan yang tersedia di Google Cloud Platform (GCP) yaitu bigquery. Tools berfungsi sebagai pengelola data warehouse dan big data. Penyimpanan data disesuaikan dengan skema data warehouse yang dirancang sebelumnya.

2.4. Evaluasi

Evaluasi bertujuan mengukur kualitas dan efektivitas data hasil proses ETL sebelum analisis dan pelaporan [4] Dalam penelitian ini, evaluasi dilakukan melalui dua pendekatan [13]. Pertama, validasi kualitas data menggunakan tools Great Expectations untuk memeriksa konsistensi tipe data, format tanggal, dan validasi karakter (contoh: nama tanpa simbol khusus) [14]. Kedua, kuesioner dengan responden stakeholder (tim akademik dan instruktur) menggunakan skala Likert 1-5 yang mengadopsi standar ISO/IEC 25012:2008 guna mengukur efisiensi proses ETL dalam konteks operasional [15]. Terdapat 10 pertanyaan dengan dua

kategori pertanyaan yaitu penilaian terhadap kualitas data (P1-P10) dan efisiensi proses (P6-P10). Setiap kategori akan dirata-rata berdasarkan tiap responden. Hal ini dilakukan untuk menjadi aspek penilaian efisiensi proses. Kemudian dihitung total skor Likert sebagai berikut:

SS : Sangat Setuju, diberi nilai 5

S : Setuju, diberi nilai 4

RG : Cukup Setuju, diberi nilai 3

TS : Tidak Setuju, diberi nilai 2

ST : Sangat Tidak Setuju, diberi nilai 1

Kualitas data = $(3 + 2 + 4 + 3 + 2) / 5 = 2,80$. Nilai tersebut diperoleh berdasarkan jawaban responden pada kuesioner. Maka dari itu responden tersebut memperoleh nilai 2,80 dan nilai tersebut nantinya perlu dibulatkan menjadi 3, begitupun pada variabel metrik lain. Berdasarkan perhitungan skala likert dilakukan dengan langkah-langkah, menghitung total jawaban kuesioner kemudian menghitung total minimal. Pada penelitian ini skor 1 menjadi minimal yaitu dan total pertanyaan pada kuesioner sebanyak 10. Perhitungan skor maksimal dilakukan dengan cara:

$$\text{Total min} = \text{skor minimal skala likert} * \text{total pertanyaan}$$

$$\text{Total min} = 1 * 10 = 10$$

Perhitungan total maksimal pada penelitian ini skor maksimal yaitu 5. Perhitungan skor maksimal dilakukan dengan cara:

$$\text{Total maks} = \text{skor maksimal skala likert} * \text{total pertanyaan}$$

$$\text{Total maks} = 5 * 10 = 50$$

Untuk menentukan rentang dilakukan perhitungan:

$$\text{Range} = \text{Total maks} - \text{Total min}$$

$$\text{Range} = 50 - 10 = 40$$

Menghitung nilai panjang kelas interval (p). Untuk menghitung nilai interval dilakukan perhitungan:

$$p = 40/5$$

$$p = 8$$

Pada penelitian ini terdapat 5 kelas skala dengan kategori sangat setuju, setuju, cukup setuju, kurang setuju, dan tidak setuju. Maka dari itu dengan panjang kelas interval (p) = 8 dan dengan dimulai batas bawah 10. Proses hitung rentang p dikurangi dengan 1. Hal ini memastikan bahwa ada "batas atas" dari suatu kategori tidak tumpang tindih dengan batas bawah kategori berikutnya, maka distribusi kelas dapat dilihat pada tabel 1.

Tabel 1. Distribusi Kelas

No.	Perhitungan Batas Atas	p	Kategori
1.	$10 + 8 - 1 = 17$	$10 - 17$	Sangat Tidak Setuju
2.	$18 + 8 - 1 = 25$	$18 - 25$	Kurang Setuju

3.	$26 + 8 - 1 = 33$	$26 - 33$	Cukup Setuju
4.	$34 + 8 - 1 = 41$	$34 - 41$	Setuju
5.	$42 + 8 - 1 = 50$	$42 - 50$	Sangat Setuju

Hasil kategori berdasarkan contoh pengisian kuesioner variabel kualitas data, total skor penilaian responden x jika dijumlahkan menghasilkan 31, yang muncul dari skor kumulatif pada kategori kualitas data (14) dan efisiensi proses (17). Total skor ini berada pada interval (p) kategori 'cukup setuju'. Hal ini menunjukkan bagaimana responden menilai keakuratan dan efisiensi sistem yang digunakan keandalan datanya dan seberapa baik kerjanya. Selanjutnya, setiap kategori yang diperoleh setiap responden akan dilihat dari persentase tiap kategori untuk semua responden. Evaluasi selanjutnya diperlukan untuk menilai setiap indikator pengukuran dengan menghitung rata-ratanya sehingga akan diperoleh kesimpulan kualitas data dan efisiensi dengan adanya implementasi ETL pada sentralisasi data dengan skema data warehouse yang dirancang [16].

Metodologi ini sangat penting untuk menggambarkan pemahaman yang lebih luas tentang kualitas informasi dan *workflow* setelah implementasi Extract, Transform, Load (ETL) untuk penggabungan data melalui skema data warehouse yang sudah ada. Dengan demikian, penilaian hasil akan membantu dalam menentukan keberhasilan sistem dalam memenuhi kebutuhan pengguna dan menentukan bagian-bagian yang perlu ditingkatkan.

3. Hasil dan Pembahasan

3.1. Analisis Kebutuhan

Hasil observasi mendapatkan beberapa informasi penting yang dikumpulkan adalah sebagai berikut :

Tabel 2. Analisis Kebutuhan

No.	Kebutuhan	Keterangan
1.	Informasi <i>Batch</i> dan Akademik	- Detail <i>batch</i> peserta, tanggal mulai, dan kesesuaian dengan kalender akademik. - Penentuan buddy untuk setiap fase (berdasarkan dokumen penilaian atau pembagian buddy).
2.	Progres dan Evaluasi	- Nilai tugas per fase dan perkembangan belajar mingguan. - Status akhir peserta per fase (Withdraw/Passed/Failed/Cuti). - Kehadiran kelas teknikal dan <i>mentoring</i>. - Poin kesempatan dan keputusan yang diambil peserta. - Informasi <i>batch</i>/fase saat peserta mengulang.

3.2. Sumber Data

Sumber data berdasarkan kebutuhan berjumlah 196 buah Google Spreadsheet dengan format kolom berbeda. Selanjutnya penelitian ini memerlukan jadwal instruktur mengajar yang ada pada database operasional bernama “Teachers Simulation” yang memiliki format spreadsheet

3.3. Implementasi ETL

Pemodelan data warehouse pada implementasi ETL ini dirancang menggunakan metode Kimball's *four-step* dimensional design process menghasilkan ditampilkan dalam tabel 3.

3.3.1. Skema Data Warehouse

Berikut ini adalah hasil pemodelan data warehouse yang dirancang menggunakan metode Kimball's *four-step* dimensional design ditampilkan dalam tabel 3.

Tabel 3. Pemodelan Data Warehouse Kimball

No	Tahap	Hasil
1	Proses Bisnis	- Mengetahui progres belajar peserta <i>current batch</i> - Mengkaji perkembangan belajar peserta <i>past batch</i>. - Mengetahui informasi detail aktivitas mengajar instruktur seperti kegiatan <i>mentoring</i> dan <i>technical class</i> yang diadakan instruktur. - Menyediakan detail kehadiran setiap hari pada setiap <i>batch</i>. - Membandingkan jumlah dan status peserta baik yang aktif maupun alumni dari keseluruhan <i>batch</i>.
2	Grain	- Grain progres belajar peserta <i>past batch</i> dan <i>current batch</i> untuk setiap evaluasi mingguan seorang siswa dalam satu <i>batch</i> pada satu minggu tertentu. - Grain aktivitas mengajar instruktur per instruktur setiap hari - Grain kehadiran <i>current batch bootcamp</i> adalah per siswa setiap hari
3	Dimensi	- Tabel Dimensi: - Tabel <i>dim_student</i> - Tabel <i>dim_instructor</i> - Tabel <i>dim_batch</i> - Tabel <i>dim_ct</i> - Tabel <i>dim_week</i> - Tabel <i>dim_date</i> - Tabel <i>dim_kelulusan</i>
4	Fakta Numerikal	- Tabel fakta progres belajar peserta <i>bootcamp current batch</i> - Tabel fakta progres belajar peserta <i>bootcamp past_batch</i> - Tabel fakta kehadiran tiap peserta <i>current batch bootcamp</i> - Tabel fakta aktivitas mengajar instruktur - Tabel fakta peserta aktif dan alumni (<i>all_batch</i>)

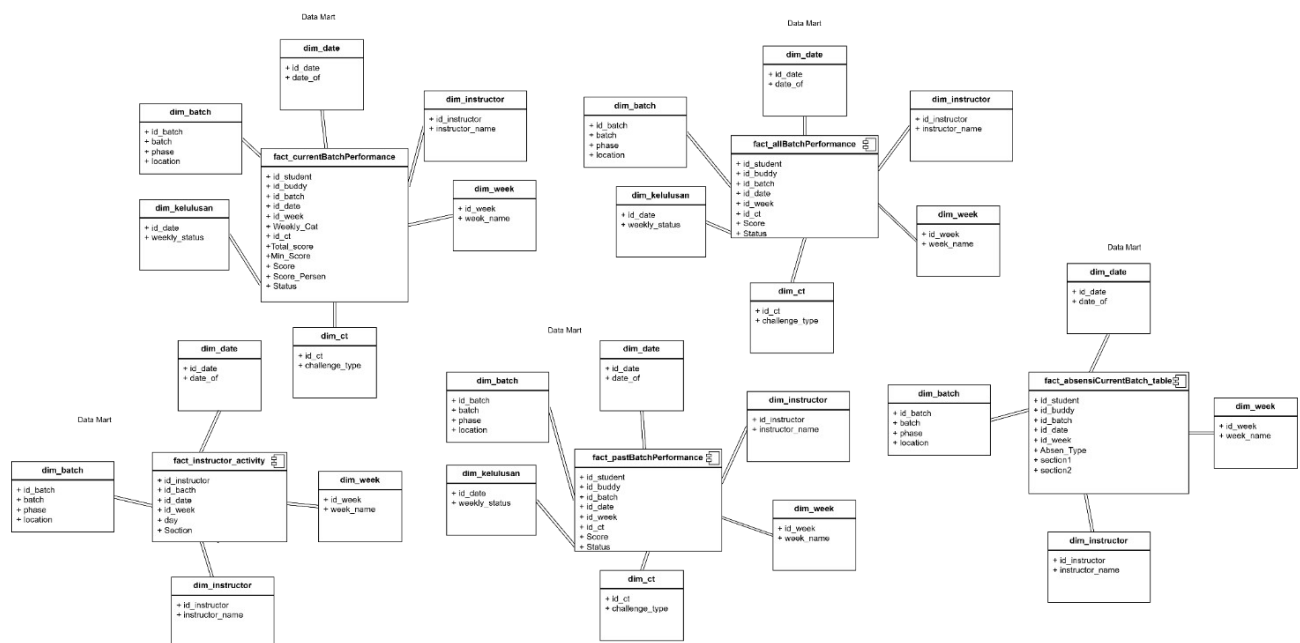
Tabel 4 digunakan untuk memetakan hubungan antara fakta (fact tables) dan dimensi (dimension tables) dalam arsitektur data warehouse adalah sebagai berikut. Tabel dimensi di bawah ini adalah bus dimension conformance matrix dan konsep dalam data warehouse yang digunakan dalam perancangan data warehouse diperkenalkan oleh Ralph Kimball.

Tabel 4. Pemodelan Bus Dimension Conformance Matrix

Grain	Dimensi	Student	Insruktur	Batch	Challenge_Type	Week	Kelulusan	Date
Progres belajar peserta <i>bootcamp past batch</i>		X	X	X	X	X	X	X
Progres belajar peserta <i>bootcamp current batch</i>		X	X	X	X	X	X	X

Kehadiran peserta bootcamp current batch	X	X	X		X		X
Aktivitas mengajar instruktur		X	X		X		X
Peserta aktif dan alumni	X	X	X	X	X	X	X

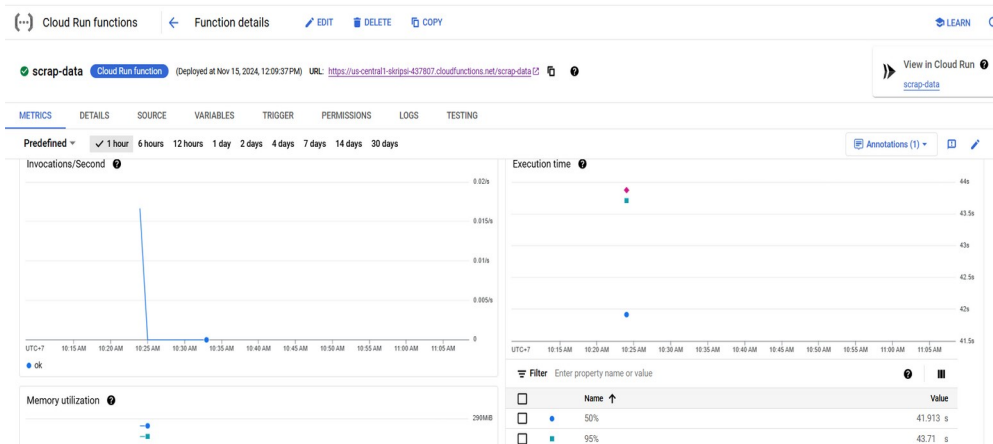
Tabel fakta numerikal mencakup atribut-atribut kunci yang berasal dari tabel dimensi terlibat yang telah dirancang pada tahapan sebelumnya. Modelling yang diterapkan menggunakan pendekatan star schema. Berikut ini adalah ilustrasi tabel fakta yang akan dibuat dalam gambar untuk masing-masing tabel fakta dengan star schema untuk masing-masing *data mart* seperti gambar 6 berikut.



Gambar 6. Skema Data Warehouse

Setelah merancang pemodelan data warehouse, dilakukan pembuatan pipeline ETL di GCP dengan mengekstrak data dari spreadsheet akademik ke staging area untuk mencegah gangguan pada transaksi database, memilih field yang relevan sesuai kebutuhan pengguna, dan membuang yang tidak diperlukan sebelum proses transformasi dan load.

Tahap ekstraksi dilakukan menggunakan Google Cloud Functions untuk mengotomasi *scraping* dan ekstraksi data dari Google Sheets yang menyimpan informasi peserta, penilaian, aktivitas instruktur, dan log kehadiran *bootcamp* setiap minggu.



Gambar 7. Proses *Extract*

Gambar 7 merupakan tampilan dari dashboard proses ETL untuk memantau performa fungsi scrap-data, termasuk frekuensi pemanggilan, waktu eksekusi (41,9–43,8 detik), dan penggunaan memori (290 MiB). Hasil ekstraksi disimpan dalam data staging sebagai lima tabel staging untuk tahap transformasi berikutnya.

Tahap transform adalah proses memastikan data sesuai dengan kebutuhan sebelum dimasukkan ke data warehouse.

```
transform-dim-and-fact-new-batch
123 b.id_batch,
124 d.id_date,
125 w.id_week,
126 bd.Weekly_Cat,
127 ct.id_ct,
128 bd.Total_score,
129 bd.Min_Score,
130 bd.Score,
131 bd.Score_Persen,
132 bd.Status
133 FROM base_data AS bd
134 JOIN 'skripsi-437807-etl_Hacktiv_dim_Student' AS s
135 ON bd.Name = s.name_student
136 JOIN 'skripsi-437807-etl_Hacktiv_dim_Instructor' AS i
137 ON bd.Buddy = i.buddy_name
138 JOIN 'skripsi-437807-etl_Hacktiv_dim_Batch' AS b
139 ON bd.Batch = b.Batch AND bd.Phase = b.Phase AND bd.Location = b.Location
140 JOIN 'skripsi-437807-etl_Hacktiv_dim_ct' AS ct
141 ON bd.Challenge_Type = ct.challenge_type
142 JOIN 'skripsi-437807-etl_Hacktiv_dim_Date' AS d
143 ON PARSE_DATE('%d/%m/%Y', bd.Date_of_Enrollment) = PARSE_DATE('%d/%m/%Y', d.date_of_enrollment)
144 JOIN 'skripsi-437807-etl_Hacktiv_dim_kelulusan' AS k
145 ON bd.Weekly_Status = k.Weekly_Status
146 JOIN 'skripsi-437807-etl_Hacktiv_dim_Week' AS w
147 ON bd.Week = w.week_name;
```

Gambar 8. Proses Transformasi

Proses ini mencakup data cleaning (menghapus nilai null dan duplikasi), standarisasi (menyesuaikan tipe data dan format tanggal ke YYYY/MM/DD), joining (menggabungkan data dari berbagai sumber berdasarkan atribut yang sama), agregasi (menganalisis performa peserta di berbagai lokasi dan berdasarkan mentor), serta filtering (menyaring siswa dalam kategori 'safe' dan 'danger').

Standarisasi memastikan data numerik dan tanggal konsisten untuk mempermudah analisis, filtering, dan join. Sementara itu, tahap joining menghasilkan tabel fakta dengan hanya menyertakan kolom yang relevan sebelum data dimuat ke *data mart*.

Tahap load adalah tahap akhir dalam proses ETL. Data yang telah diekstrak dan ditransformasi disimpan ke dalam data warehouse. Data disusun dalam tabel fakta dan dimensi

sesuai desain Kimball. Salah satu tabel hasil load, `fact_currentBatchPerformance_tabel` yang menyimpan data performa siswa dalam *batch bootcamp*.

Row	id_stude	id_bugi	id_batch	id_date	id_week	Weekly_Cat	id_ct	Total_score	Min_Score	Score	Score_Perseg	Status
1	179	27	52	42	1	Safe	1	50.0	35.0	45.25	90.5	Safe
2	773	27	52	42	1	Safe	1	50.0	35.0	42.75	85.5	Safe
3	131	27	52	42	1	Safe	1	50.0	35.0	46.5	93.0	Safe
4	200	27	52	42	1	Safe	1	50.0	35.0	48.0	96.0	Safe
5	773	27	52	42	2	Safe	2	83.0	58.1	49.5	59.64	Danger
6	131	27	52	42	2	Safe	2	83.0	58.1	60.0	72.29	Safe
7	200	27	52	42	2	Safe	2	83.0	58.1	77.75	93.67	Safe
8	179	27	52	42	2	Safe	2	83.0	58.1	41.5	50.0	Danger
9	200	27	52	42	3	Safe	3	68.0	47.6	0.0	0.0	Danger
10	131	27	52	42	3	Safe	3	68.0	47.6	0.0	0.0	Danger
11	179	27	52	42	3	Safe	3	68.0	47.6	0.0	0.0	Danger
12	773	27	52	42	3	Safe	3	68.0	47.6	0.0	0.0	Danger
13	200	27	52	42	1	Safe	8	31.0	21.7	22.0	70.97	Safe
14	131	27	52	42	1	Safe	8	31.0	21.7	23.5	75.81	Safe
15	773	27	52	42	1	Safe	8	31.0	21.7	13.5	43.55	Danger
16	179	27	52	42	1	Safe	8	31.0	21.7	22.0	70.97	Safe

Gambar 9. Proses Load ke Data mart

Tabel ini mencakup (`id_student`), *batch* (`id_batch`), minggu pembelajaran (`id_week`), dan tanggal (`id_date`), serta metrik agregasi performa seperti total skor, skor minimum, skor aktual, dan persentase skor. Selain itu, tabel ini mengategorikan status siswa sebagai "Safe" atau "Danger", untuk menganalisis performa secara real-time dan terjadwal.

Alur kerja ETL dirancang secara *batch processing*. Scheduling dilakukan setiap hari Senin pada jam 08:00 waktu Asia/Jakarta, ditulis dengan simbol cron jobs `0 8 ** 1`.

Name	Status of last execution	Region	State	Description	Frequency	Target	Last run	Next run
scraping-the-data-hacktiv	Success	us-central1	Enabled	scraping data akademik dari g spreadsheet	0 8 * * 1 (Asia/Jakarta)	URL : https://us-central1-skripsi-437807.cloudfunctions.net/scrap-data	6 Jan 2025, 08:00:00	13 Jan 2025, 08:00:00
transform-load-data	Success	us-central1	Enabled	transformin from data staging and load to data akademik	5 8 * * 1 (Asia/Pontianak)	URL : https://workflowexecutions.googleapis.com/v1/projects/skripsi-437807/locations/us-central1/workflows/transform-data1/executions	6 Jan 2025, 08:05:00	13 Jan 2025, 08:05:00

Gambar 10. Proses Scheduling

Pada sistem ini, terdapat dua fungsi utama yang berjalan menggunakan Google Cloud Scheduler, masing-masing dijadwalkan untuk melakukan tugas secara otomatis sebagai bagian dari pipeline ETL (Extract, Transform, Load). Function pertama "scraping-the-data-hacktiv," bertugas untuk mengambil data akademik dari Google Spreadsheet menggunakan Cloud Function.

The screenshot shows the 'Workflow details' page for a workflow named 'transform-data1'. It displays a table of execution logs with columns for State, Execution ID, Workflow revision, Create time, Start time, End time, and Duration. All listed executions are in a 'Succeeded' state.

State	Execution ID	Workflow revision	Create time	Start time	End time	Duration
returnOutput Succeeded	2b3b5262-ab4f-4127-b694-31cc63a3489	000004-c8a	6 Jan 2025, 08...	6 Jan 2025, 08...	6 Jan 2025, 08...	3 minutes
returnOutput Succeeded	42b9270d-bec2-4a2e-b35f-389c2275009c	000004-c8a	30 Dec 2024, 08...	30 Dec 2024, 0...	30 Dec 2024, 0...	3 minutes
returnOutput Succeeded	8f643710-bbf9-465f-95a3-95cb1a99224	000004-c8a	23 Dec 2024, 08...	23 Dec 2024, 0...	23 Dec 2024, 0...	3 minutes
returnOutput Succeeded	4a58d845-3f71-4127-882d-ed8f8a778f	000004-c8a	16 Dec 2024, 08...	16 Dec 2024, 0...	16 Dec 2024, 0...	3 minutes
returnOutput Succeeded	90ea08e1-40a0-47b6-8982-b1a9d40aee1	000004-c8a	9 Dec 2024, 08...	9 Dec 2024, 08...	9 Dec 2024, 08...	3 minutes
returnOutput Succeeded	d11b5cc4-2861-4688-ab87-bc307e7e9931	000004-c8a	2 Dec 2024, 08...	2 Dec 2024, 08...	2 Dec 2024, 08...	3 minutes
returnOutput Succeeded	07f1d1d5-b64a-4c99-c0b9-311a78c345c	000004-c8a	29 Nov 2024, 09...	29 Nov 2024, 0...	29 Nov 2024, 0...	3 minutes
returnOutput Succeeded	3ec2be90-c115-4749-a5c5-a600b0be995	000004-c8a	29 Nov 2024, 09...	29 Nov 2024, 0...	29 Nov 2024, 0...	3 minutes
returnOutput Succeeded	b7a938bf-e925-4a4b-bed8-6191b4359045	000004-c8a	28 Nov 2024, 21...	28 Nov 2024, 2...	28 Nov 2024, 2...	3 minutes
returnOutput Succeeded	59b9da7f-4a4b-4229-8854-04fa37e0a50d	000004-c8a	28 Nov 2024, 21...	28 Nov 2024, 2...	28 Nov 2024, 2...	3 minutes

Gambar 11. Dashboard Workflow Bigquery

Fungsi trigger *scraping* dan transform load Workflow pada gambar 11 memastikan bahwa data diproses urutan pipeline ETL. Adanya Workflows, sehingga pipeline ini menjadi *scalable*.

3.4. Evaluasi

Evaluasi kualitas data dan proses dilakukan melalui data validation dan kuesioner. Data validation menggunakan Great Expectations untuk mengukur akurasi, konsistensi, kelengkapan, dan keterkinian data. Pada proses implementasi ETL, hasil analisis menunjukkan adanya kehilangan data pada tahap transformasi yang terjadi pada tabel *fact_allBatchPerformance_table* dan *fact_pastBatchPerformance_table*. Hal ini disebabkan oleh beberapa record pada database asal (spreadsheet) yang mencantumkan data siswa tanpa mencatat nama instruktur pengampunya. Nilai tersebut berupa string kosong atau karakter seperti '-'. Skema data warehouse, khususnya menurut prinsip Kimball, tabel fakta harus mematuhi aturan integritas referensial. Hal ini berarti, setiap foreign key pada tabel fakta harus memiliki referensi yang valid di tabel dimensi, dan tidak boleh mengandung *null value*. Nilai null pada kolom foreign key menyebabkan fakta kehilangan konteks, sehingga tidak dapat dihubungkan dengan dimensi yang relevan, seperti *dim_Instructor*.

Row	Nama	Budd	Batch	Location	Phase	Date_of_Enro	Week	Status	id_studen	id_buddy	id_batc	id_date	id_weer
1	haider farras		002	hck	phase 0	2022-10-...	Week 2	danger	333	nuli	14	291	2
2	haider farras		002	hck	phase 0	2022-10-...	Week 3	danger	333	nuli	14	291	3
3	haider farras		002	hck	phase 0	2022-10-...	Week 3	safe	333	nuli	14	291	3
4	haider farras		002	hck	phase 0	2022-10-...	Week 4	safe	333	nuli	14	291	4
5	haider farras		002	hck	phase 0	2022-10-...	Week 1	safe	333	nuli	14	291	1
6	haider farras		002	hck	phase 0	2022-10-...	Week 1	safe	333	nuli	14	291	1
7	haider farras		002	hck	phase 0	2022-10-...	Week 2	safe	333	nuli	14	291	2
8	luthfan milzam	-	006	hck	phase 1	2023-06-...	Week 4	danger	446	nuli	65	437	4
9	luthfan milzam	-	006	hck	phase 1	2023-06-...	Week 2	danger	446	nuli	65	437	2

Gambar 12. Hasil Analisis Kehilangan Data Setelah ETL

Gambar 12 menunjukkan pada proses ETL, data dengan nilai null atau karakter '-' pada kolom instruktur dibuang untuk memastikan konsistensi dan keandalan data dalam data warehouse. Meskipun demikian, langkah ini menjaga integritas referensial, hal tersebut juga mengakibatkan perbedaan jumlah data antara database asal dan tabel fakta, yang diidentifikasi sebagai kehilangan data pada proses ETL dieksekusi.

Tabel 5. Hasil Evaluasi Standard Data Quality Dimensions ISO/IEC 25012

No	Metrik	Akurasi	Rules
1	Akurasi	90%	• Atribut (student_name) pada tabel dimensi dim_student dan (instructor_name) pada tabel harus bebas dari tanda baca. • Atribut 'Weekly_Status' pada tabel dim_kelulusan hanya berisi status kelulusan. • Atribut location hanya berisi SBY, HCK, RMT, BSD.
2	Konsistensi	97%	Atribut batch tabel dimensi dim_Batch memiliki 3 digit angka • Atribut tanggal pada tabel dim_Date memiliki format YYYY/MM/DD Jumlah row sama dengan referensi tabel sesudah dan sebelum transformation.
3	Kelengkapan	100%	Atribut bebas dari nilai null.
4	Terkini	TRUE	Atribut tanggal date min 28/06/2021 dan max (tanggal terkini)

Hasilnya menunjukkan sebagian kecil data tidak memenuhi aturan tertentu, terutama karena nilai null atau format yang tidak valid. Kuesioner digunakan untuk menilai relevansi data berdasarkan feedback pengguna. Evaluasi ini memastikan kualitas data dalam data warehouse sesuai standar ISO/IEC 25012.

Evaluasi kedua dengan pengujian dari kuesioner menghasilkan rata-rata skor 4 dari 5 skala likert, yang menunjukkan bahwa mayoritas pengguna setuju bahwa kualitas data yang dihasilkan baik. Aspek konsistensi dan struktur mendapatkan nilai tertinggi. Namun, ada beberapa aspek yang masih perlu ditingkatkan, terutama kesesuaian data dengan kebutuhan spesifik untuk *reporting*, monitoring.

Tabel 6. Distribusi Penilaian Keseluruhan Dengan Skala Likert

No.	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total Skor	Kategori
1	4	3	4	5	4	5	2	4	4	5	40	Setuju
2	4	4	4	4	4	4	4	4	4	4	40	Setuju
3	4	5	5	5	5	5	4	5	4	5	47	Sangat Setuju
4	5	4	5	5	5	5	4	4	4	5	46	Sangat Setuju
5	3	4	3	3	4	5	3	4	4	2	35	Setuju
6	4	4	5	5	4	5	3	4	5	4	43	Sangat Setuju
7	3	4	4	4	4	4	4	4	4	4	39	Setuju
8	4	3	4	4	3	5	4	3	5	3	38	Setuju
9	4	3	3	4	3	4	3	3	4	3	34	Setuju
10	3	4	4	4	4	5	4	5	4	4	41	Setuju

Tabel 31 menyajikan hasil evaluasi kualitas data yang dihasilkan oleh proses ETL berdasarkan penilaian dari 10 end user melalui 10 item pertanyaan (P1–P10) menggunakan skala Likert. Hasil evaluasi menunjukkan bahwa rata-rata skor berada pada angka 4, yang mengindikasikan persepsi positif terhadap kualitas data secara umum. Aspek konsistensi dan struktur memperoleh skor tertinggi, menandakan kepuasan tinggi terhadap kedua dimensi tersebut. Namun, aspek kesesuaian data untuk kebutuhan *reporting*, monitoring, dan analisis mendapat skor terendah (rata-rata 3), menunjukkan area yang perlu ditingkatkan. Dari sisi efisiensi

proses sebelum dan sesudah implementasi ETL, sebagian besar pengguna juga memberikan skor 4 yang semula hanya 2.8, menunjukkan bahwa proses dianggap cukup efisien. Meskipun demikian, dokumentasi ETL mendapat pertimbangan khusus karena memperoleh skor terendah dari salah satu responden, yang menunjukkan perlunya peningkatan dalam kelengkapan dan kejelasan dokumentasi teknis program.

4. Kesimpulan dan Saran

Penelitian ini menghasilkan sistem pipeline ETL untuk sentralisasi data akademik menggunakan skema data warehouse yang diimplementasikan dengan bantuan Google Cloud Platform dan scheduler untuk otomatisasi proses Extract, Transform, Load. Sistem ini ditujukan untuk mengatasi fragmentasi data akademik dan mendukung efisiensi dalam monitoring, pelaporan, serta pengambilan keputusan.

Evaluasi kualitas data berdasarkan standar ISO/IEC 25012 menunjukkan bahwa data yang dihasilkan memenuhi kriteria akurasi, kelengkapan, dan konsistensi dengan tingkat akurasi atribut utama di atas 98% serta kelengkapan mendekati 100%. Penilaian pengguna melalui kuesioner juga menunjukkan respons positif dengan skor rata-rata 4 pada skala Likert, khususnya untuk aspek konsistensi dan struktur. Namun, dokumentasi proses ETL dinilai masih perlu ditingkatkan. Sebagai tindak lanjut penelitian selanjutnya, disarankan adanya dokumentasi untuk meningkatkan pemahaman terhadap sistem, perbaikan pada dokumentasi ETL, serta penanganan skenario kesalahan secara sistematis. Penelitian lanjutan diharapkan dapat mengembangkan fitur visualisasi interaktif untuk memperluas dukungan analitik yang lebih *scalable*.

Daftar Pustaka

- [1] D. Mirza, L. Suryani, L. Latip, and V. Aditiya, "Literature Riview: Peran Teknologi Informasi dalam Meningkatkan Efisiensi dan Efektivitas Birokrasi," *Jurnal Administrasi Publik dan Bisnis*, vol. 5, no. 1, pp. 51–55, Mar. 2023, doi: 10.36917/japabis.v5i1.84.
- [2] Lilik Wienarni, "Pengembangan Birokrasi Digital Di Indonesia," *JURNAL EKONOMI, SOSIAL & HUMANIORA*, vol. 1, no. Vol 1 No 02 (2019): INTELEKTIVA : Jurnal Ekonomi, Sosial dan Humaniora E-ISSN 2686-5661, pp. 24–32, 2019.
- [3] G. Agapito, C. Zucco, and M. Cannataro, "COVID-WAREHOUSE: A Data Warehouse of Italian COVID-19, Pollution, and Climate Data," *Int J Environ Res Public Health*, vol. 17, no. 15, p. 5596, Aug. 2020, doi: 10.3390/ijerph17155596.
- [4] A. A. Yulianto, "Extract Transform Load (ETL) Process in Distributed Database Academic Data Warehouse," *APTIKOM Journal on Computer Science and Information Technologies*, vol. 4, no. 2, pp. 61–68, Jul. 2019, doi: 10.11591/APTIKOM.J.CSIT.36.
- [5] I. A. A. Alsibhawi, J. B. Yahaya, and H. B. Mohamed, "Business Intelligence Adoption for Small and Medium Enterprises: Conceptual Framework," Apr. 01, 2023, *MDPI*. doi: 10.3390/app13074121.
- [6] M. Fernandez, A. Davila, and P. Angeleri, "Data quality applied to an academic business intelligence solution: Lesson learned," in *2017 IEEE Colombian Conference on*

Communications and Computing (COLCOM), IEEE, Aug. 2017, pp. 1–6. doi: 10.1109/ColComCon.2017.8088200.

- [7] A. Y. Berliantara, S. A. Wicaksono, and A. Pinandito, “Optimasi Scheduling untuk Proses Extract, Transform, Load (ETL) pada Data Warehouse Menggunakan Metode Round Robin Data Partitioning (Studi Kasus: Universitas XYZ),” 2017. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] D. Kabakchieva, “Business Intelligence Systems for Analyzing University Students Data,” *Cybernetics and Information Technologies*, vol. 15, no. 1, pp. 104–115, Mar. 2015, doi: 10.1515/cait-2015-0009.
- [9] K. W. Broman and K. H. Woo, “Data Organization in Spreadsheets,” *Am Stat*, vol. 72, no. 1, pp. 2–10, Jan. 2018, doi: 10.1080/00031305.2017.1375989.
- [10] G. Wijaya, S. Nusa, and M. Jakarta, “Perancangan Data Warehouse Nilai Mahasiswa Dengan Kimball Nine-Step Methodology,” *JURNAL INFORMATIKA*, vol. 4, no. 1, 2017.
- [11] K. U. Rehman, U. Ahmad, and S. Mahmood, “A Comparative Analysis of Traditional and Cloud Data Warehouse,” *VAWKUM Transactions on Computer Sciences*, vol. 15, pp. 34–40, Jan. 2018, doi: 10.21015/vtcs.v15i1.487.
- [12] S. Saifullah and V. A. Permadi, “Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms,” in *Proceeding - 2019 5th International Conference on Science in Information Technology: Embracing Industry 4.0: Towards Innovation in Cyber Physical System, ICSITech 2019*, Institute of Electrical and Electronics Engineers Inc., Oct. 2019, pp. 140–145. doi: 10.1109/ICSITech46713.2019.8987496.
- [13] S. P. Tahalea and A. SN, “Central Actor Identification of Crime Group using Semantic Social Network Analysis,” *Indonesian Journal of Information Systems*, vol. 2, no. 1, p. 24, 2019, doi: 10.24002/ijis.v2i1.2354.
- [14] Hermantoro, A. P. Suryotomo, A. I. Uktoro, and R. A. Renjani, “Unmanned Aerial Vehicle Application for Plantation Mapping and Automatic Oil Palm Trees Counting on Oil Palm Plantation Management,” in *International Conference on the Role of Agricultural Engineering for Sustainable Agriculture Production*, 2016, pp. 47–50.
- [15] Awang Hendrianto Pratomo, W. Kaswidjanti, and S. Mu’arifah, “Implementasi Algoritma Region of Interest (ROI) Untuk Meningkatkan Performa Algoritma Deteksi Dan Klasifikasi Kendaraan,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 7, no. 1, pp. 155–162, 2020, doi: 10.25126/jtiik.202071718.
- [16] S. P. Tahalea, “Identifikasi Peran Hero DOTA2 Menggunakan Social Network Analysis,” *TEKNOMATIKA*, vol. 12, no. 2, pp. 81–86, 2020.