IMPLEMENTASI DATA WAREHOUSE MENGGUNAKAN PENTAHO UNTUK MEMBUAT DASHBOARD VISUALISASI KINERJA PENJUALAN PADA TOKO BESI DAN BAJA RINGAN NATA JAYA PERKASA

Dewi Khuzaimah[1], Sri Wahyuni[2], M. Andres[3]
[1)]Program Studi Sistem Informasi, Universitas Panca Sakti Bekasi
E-mail : hayyandewi@gmail.com
[2)] Program Studi Informatika, Universitas Panca Sakti Bekasi
E-mail: sriyuni82sw@gmail.com
[3)]Program Studi Sistem Informasi, Universitas Panca Sakti Bekasi
E-mail: andresmuhammad3@gmail.com

ABSTRAK

Pencatatan penjualan yang terpisah pisah membuat pemilik toko sulit memperoleh gambaran kinerja usaha secara cepat. Kondisi tersebut ditemukan pada Toko Besi dan Baja Ringan Nata Jaya Perkasa, tempat data barang, data penjualan, dan data pembelian dicatat pada berkas yang berbeda sehingga penggabungannya masih dikerjakan secara manual dan laporan baru tersedia beberapa hari setelah periode penjualan berakhir. Penelitian ini mengimplementasikan data warehouse bermodel skema bintang dengan proses Extract, Transform, Load yang diotomatisasi menggunakan Pentaho Data Integration, kemudian menyajikan hasilnya pada dashboard visualisasi kinerja penjualan. Perancangan mengikuti siklus hidup pengembangan gudang data, mulai dari analisis kebutuhan bisnis, audit sumber data, pemodelan dimensi, pembangunan alur ETL, penyusunan dashboard, sampai pengujian. Rancangan awal yang menuntut petugas menggabungkan data secara manual direvisi menjadi tiga transformasi terpisah yang membaca tiga berkas sumber dan dikendalikan oleh satu job, sehingga penanda jenis mutasi serta nama dan satuan barang diisi oleh sistem, bukan diketik ulang. Pengujian menunjukkan seluruh berkas sumber terbaca tanpa galat dengan jumlah baris dibaca sama dengan jumlah baris ditulis pada setiap transformasi, yaitu 10 baris data barang, 14 baris penjualan, dan 16 baris pembelian. Tabel fakta terisi 30 baris dengan 452 satuan barang keluar dan 1.675 satuan barang masuk sehingga saldo stok berjumlah 1.223 satuan, seluruh rangkaian selesai kurang dari satu detik, dan jumlah baris tetap sama pada setiap pengulangan proses. Penelitian ini menyimpulkan bahwa otomatisasi ETL beserta dashboard yang dihasilkan menggantikan penggabungan data manual dan mempercepat penyajian informasi penjualan bagi pemilik toko.

Kata kunci : data warehouse, Pentaho Data Integration, ETL, skema bintang, dashboard kinerja penjualan

1. PENDAHULUAN

Data transaksi merupakan aset penting bagi usaha perdagangan karena menjadi dasar penilaian kinerja penjualan dan pengambilan keputusan pemilik usaha. Nilai data tersebut baru muncul apabila catatan yang tersebar dapat disatukan, diseragamkan, dan disajikan dalam bentuk yang mudah dibaca. Data warehouse hadir sebagai penyimpanan terintegrasi yang berorientasi subjek dan menyimpan data historis untuk keperluan analisis, terpisah dari basis data operasional yang melayani transaksi harian (Inmon, 2005; Chaudhuri & Dayal, 1997).

Penelitian ini dilakukan pada Toko Besi dan Baja Ringan Nata Jaya Perkasa, usaha perdagangan bahan bangunan di Kota Tegal dengan volume transaksi harian yang cukup tinggi. Hasil observasi dan wawancara dengan pemilik toko menunjukkan bahwa pencatatan tidak berada dalam satu berkas gabungan, melainkan terpisah menurut kegiatannya, yaitu pencatatan data barang, pencatatan penjualan, dan pencatatan pembelian. Ketiga sumber tersebut memiliki format kolom yang berbeda, sehingga penggabungannya masih dikerjakan secara manual dengan menyalin data, menyeragamkan format tanggal, dan menyamakan kode barang satu per satu. Cara kerja seperti itu memakan waktu, bergantung pada ketelitian petugas, dan membuat laporan kinerja penjualan baru tersedia beberapa hari setelah periode penjualan berakhir. Padahal pemilik toko membutuhkan informasi yang lebih cepat untuk menentukan barang yang perlu ditambah maupun strategi penjualan berikutnya.

Persoalan data yang terfragmentasi pada dasarnya merupakan persoalan integrasi data yang telah lama dibahas dalam bidang gudang data. Kimball dan Ross (2013) memperkenalkan pemodelan dimensional dengan skema bintang, yaitu satu tabel fakta yang dikelilingi beberapa tabel dimensi, sebagai rancangan yang menyederhanakan kueri analisis. Golfarelli dan Rizzi (2009) serta Vaisman dan Zimanyi (2022) menegaskan bahwa rancangan dimensional tersebut menjadi fondasi penyajian informasi ringkas bagi pengguna non teknis. Pada sisi pemuatan data, Vassiliadis (2009) memaparkan bahwa proses Extract, Transform, Load merupakan bagian yang paling banyak menyerap usaha dalam pembangunan gudang data, sehingga otomatisasinya menentukan keberlanjutan sistem. Wixom dan Watson (2001) menemukan bahwa mutu data dan dukungan pengguna merupakan faktor penentu keberhasilan penerapan gudang data.

Otomatisasi proses tersebut dapat dijalankan dengan Pentaho Data Integration, perangkat sumber terbuka yang menyusun alur pemindahan data melalui antarmuka visual tanpa banyak penulisan kode. Casters dkk. (2010) menjelaskan bahwa alur kerja pada Pentaho dibangun dari transformation sebagai satuan pengolahan baris data dan job

Copyright@2026 FST PSU Bekasi

sebagai pengendali urutan pelaksanaan, sedangkan Bouman dan van Dongen (2009) menunjukkan pemakaiannya bersama basis data relasional untuk kebutuhan pelaporan. Hasil pemuatan data kemudian disajikan pada dashboard, yaitu penyajian visual sejumlah indikator kinerja utama pada satu tampilan agar kondisi usaha dapat dinilai sekali pandang (Few, 2013).

Sebagian besar kajian gudang data menempatkan sumber data pada basis data operasional berskala besar, sedangkan usaha kecil menyimpan catatannya pada berkas terpisah menurut kegiatan dengan format yang tidak seragam. Kondisi tersebut menuntut rancangan ETL yang mampu membaca beberapa berkas sumber sekaligus tanpa penggabungan manual lebih dahulu, dan celah itulah yang dituju penelitian ini.

Berdasarkan uraian tersebut, penelitian ini bertujuan merancang arsitektur data warehouse bermodel skema bintang yang mengintegrasikan data barang, data penjualan, dan data pembelian pada Toko Besi dan Baja Ringan Nata Jaya Perkasa, mengimplementasikan proses Extract, Transform, Load otomatis menggunakan Pentaho Data Integration, serta membangun dashboard visualisasi kinerja penjualan agar pemilik toko dapat memantau indikator utama secara cepat. Ruang lingkup penelitian dibatasi pada ketiga sumber data tersebut beserta pelaporan bagi pemilik toko, sedangkan pemakaian sistem untuk pemasukan transaksi harian berada di luar cakupan.

2. LANDASAN TEORI

2.1. Data Warehouse dan OLAP

Data warehouse merupakan kumpulan data yang berorientasi subjek, terintegrasi, memiliki dimensi waktu, dan tidak berubah, yang digunakan untuk mendukung pengambilan keputusan (Inmon, 2005). Pemisahan antara basis data operasional yang menangani transaksi harian dan gudang data yang menangani analisis dilakukan agar kueri agregasi tidak membebani sistem yang melayani kegiatan usaha (Chaudhuri & Dayal, 1997). Ponniah (2010) menjelaskan bahwa pengolahan analitis daring atau Online Analytical Processing bekerja di atas data yang telah diringkas menurut dimensi tertentu, sehingga pertanyaan mengenai penjualan menurut waktu maupun barang dapat dijawab tanpa menelusuri seluruh catatan transaksi.

2.2. Pemodelan Dimensi dan Skema Bintang

Pemodelan dimensional menyusun data analitis menjadi tabel fakta yang menyimpan ukuran kuantitatif dan tabel dimensi yang menyimpan konteks pembacaan ukuran tersebut. Bentuk yang paling banyak dipakai adalah skema bintang, yaitu satu tabel fakta yang dikelilingi tabel dimensi melalui kunci pengganti atau surrogate key (Kimball & Ross, 2013). Golfarelli dan Rizzi (2009) menyebut kesederhanaan bentuk tersebut sebagai keunggulan utamanya karena jumlah penggabungan tabel dalam kueri menjadi sedikit. Vaisman dan Zimanyi (2022) menambahkan bahwa pemakaian kunci pengganti memisahkan identitas baris pada gudang data dari kode yang dipakai sistem sumber, sehingga perubahan kode pada sistem sumber tidak merusak riwayat data.

2.3. Extract, Transform, Load

Extract, Transform, Load merupakan rangkaian proses pemindahan data dari sistem sumber menuju gudang data. Tahap extract mengambil data dari sumber, tahap transform membersihkan dan menyeragamkan data, dan tahap load memuat data ke tabel tujuan. Vassiliadis (2009) mencatat bahwa tahapan tersebut menyerap porsi terbesar usaha pembangunan gudang data karena harus menangani perbedaan format, nilai kosong, dan duplikasi. Kimball dkk. (2008) menekankan bahwa pemuatan tabel dimensi perlu dijalankan sebelum tabel fakta agar setiap baris transaksi menemukan pasangan dimensinya.

2.4. Pentaho Data Integration

Pentaho Data Integration merupakan perangkat lunak ETL sumber terbuka yang menyusun alur pengolahan data melalui rangkaian langkah pada kanvas visual. Casters dkk. (2010) menjelaskan dua satuan kerja utamanya, yaitu transformation yang memproses aliran baris data dan job yang mengatur urutan pelaksanaan beserta penanganan galat. Langkah yang dipakai pada penelitian ini mencakup pembacaan berkas teks bernilai terpisah koma, penambahan nilai tetap melalui Add constants, pencarian data pada tabel melalui Database lookup, serta pemuatan bersyarat melalui Insert / Update. Bouman dan van Dongen (2009) serta dokumentasi resminya (Hitachi Vantara, 2023) menunjukkan bahwa rangkaian langkah tersebut memadai untuk membangun gudang data berskala kecil sampai menengah tanpa penulisan kode program.

2.5. Dashboard dan Indikator Kinerja

Dashboard adalah penyajian visual informasi terpenting yang dibutuhkan untuk mencapai satu atau beberapa tujuan, disusun pada satu tampilan agar dapat dipantau sekali pandang (Few, 2013). Indikator kinerja utama yang lazim dipantau pada usaha perdagangan meliputi jumlah barang masuk, jumlah barang keluar, dan sisa stok. Sharda dkk. (2020) serta Chaudhuri dkk. (2011) menempatkan dashboard sebagai lapisan penyajian pada rantai business intelligence yang bermula dari integrasi data, sehingga mutu tampilannya bergantung pada mutu proses pemuatan data di belakangnya.

3. RANCANGAN SISTEM DAN APLIKASI

3.1. Tahapan Penelitian

Penelitian ini merupakan penelitian terapan dengan pendekatan rekayasa perangkat lunak yang mengikuti siklus hidup pengembangan gudang data (Kimball dkk., 2008). Tahapan dibagi menjadi enam fase yang saling berkesinambungan sebagaimana disajikan pada Tabel 3.1. Data dikumpulkan melalui observasi pencatatan yang

Copyright@2026 FST PSU Bekasi

berjalan di toko, wawancara dengan pemilik toko untuk memastikan indikator yang dibutuhkan, serta penelaahan berkas sumber untuk memetakan format kolomnya.

Tabel 3.1 Tahapan Penelitian dan Keluarannya

No	Tahapan	Keluaran
1	Analisis kebutuhan bisnis	Daftar indikator yang dipantau
2	Audit sumber data	Peta kolom dan format tiga berkas
3	Pemodelan dimensi	Rancangan skema bintang
4	Perancangan alur ETL	Tiga transformasi dan satu job
5	Pembangunan dashboard	Tampilan indikator dan tabel
6	Pengujian dan evaluasi	Hasil uji dan indikator kelayakan

3.2. Sumber Data dan Rancangan Skema Bintang

Audit sumber data menunjukkan tiga berkas yang dipakai toko, yaitu data_barang.csv berisi sepuluh baris data master barang, data_penjualan.csv berisi empat belas baris transaksi keluar, dan data_pembelian.csv berisi enam belas baris transaksi masuk pada periode 7 sampai 13 Juli 2026. Gudang data dirancang dengan skema bintang yang berpusat pada tabel fakta fact_stock_mutation, didampingi tabel dimensi dim_barang dan dim_waktu. Tabel fakta menyimpan ukuran kuantitatif berupa jumlah satuan barang beserta penanda jenis mutasi, sedangkan atribut nama barang dan satuan disimpan pada tabel dimensi agar tidak diulang pada setiap baris transaksi (Kimball & Ross, 2013).

Pemetaan kolom dari ketiga berkas sumber ke tabel tujuan disajikan pada Tabel 3.2. Dua kolom pada tabel fakta tidak berasal dari berkas sumber, yaitu kolom jenis_mutasi yang diisi nilai tetap sesuai berkas asalnya serta kolom nama_barang dan satuan yang diambil dari tabel dimensi melalui pencarian pada basis data. Rancangan tersebut memindahkan pengisian keterangan barang dari petugas kepada sistem.

Tabel 3.2 Pemetaan Kolom Berkas Sumber ke Tabel Tujuan

Berkas sumber	Kolom tujuan	Cara pengisian
data_barang.csv	dim_barang.kode_barang	Kunci Insert / Update
data_barang.csv	dim_barang.nama_barang dan satuan	Diperbarui bila berubah
data_penjualan.csv	fact_stock_mutation.dim_barang dan dim_waktu	Salinan langsung
data_penjualan.csv	fact_stock_mutation.jumlah_qty	Salinan langsung
data_penjualan.csv	fact_stock_mutation.jenis_mutasi	Nilai tetap Keluar
data_pembelian.csv	fact_stock_mutation.jenis_mutasi	Nilai tetap Masuk
dim_barang	fact_stock_mutation.nama_barang dan satuan	Hasil Database lookup

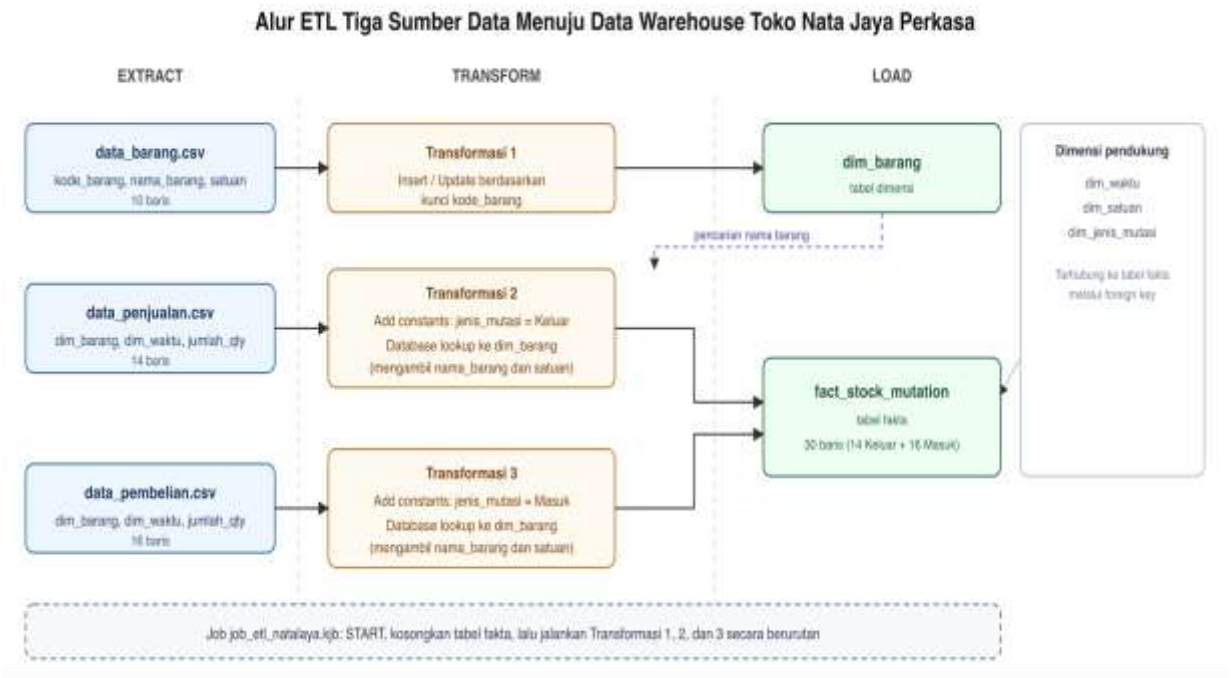
3.3. Rancangan Alur ETL Tiga Sumber Data

Rancangan alur ETL ditunjukkan pada Gambar 3.1. Proses dibangun dari tiga transformation yang masing masing menangani satu berkas sumber. Transformation 01_load_dim_barang memuat data barang ke tabel dimensi dengan langkah Insert / Update berkunci kode_barang, sehingga barang lama diperbarui dan barang baru ditambahkan. Transformation 02_load_penjualan menandai setiap baris dengan jenis mutasi Keluar, mengambil nama barang dan satuan dari tabel dimensi, lalu memuatnya ke tabel fakta. Transformation 03_load_pembelian bekerja dengan alur yang sama untuk penanda jenis mutasi Masuk. Job job_etl_natalaya berperan sebagai pengendali yang mengosongkan tabel fakta lebih dahulu, kemudian memanggil ketiga transformation secara berurutan agar tabel dimensi selalu siap sebelum baris transaksi dimuat dan pengulangan proses tidak menghasilkan data ganda (Kimball dkk., 2008).

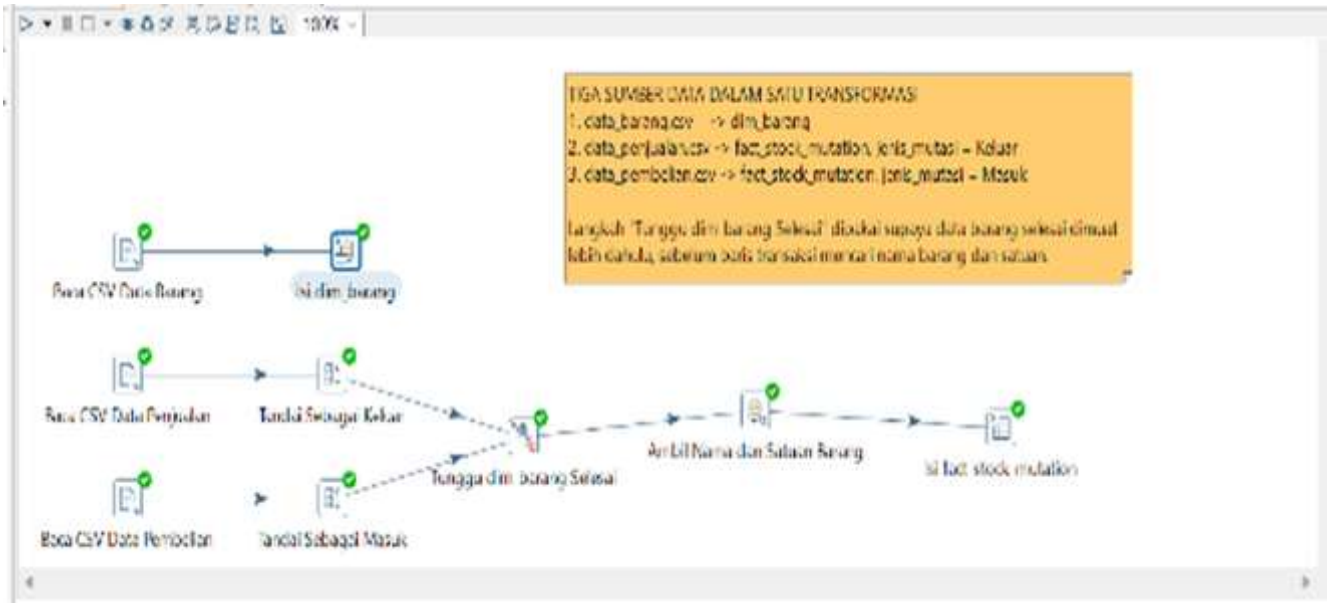
3.4. Lingkungan dan Rancangan Pengujian

Implementasi dijalankan pada satu komputer dengan Pentaho Data Integration sebagai perangkat ETL, basis data MySQL sebagai penyimpan gudang data yang diakses melalui phpMyAdmin, serta dashboard berbasis web sebagai lapisan penyajian. Pengujian dirancang pada empat hal, yaitu keterbacaan seluruh berkas sumber, kesesuaian jumlah baris yang dibaca dengan jumlah baris yang ditulis pada setiap transformation, konsistensi jumlah baris tabel fakta ketika job dijalankan berulang, serta perilaku sistem ketika satu kode barang baru ditambahkan tanpa disertai baris transaksi.

Copyright@2026 FST PSU Bekasi



Gambar 3.1
Rancangan Alur ETL Tiga Sumber Data



Gambar 4.1
Implementasi Transformation pada Pentaho Data Integration

4. HASIL DAN PEMBAHASAN

4.1. Revisi Rancangan Proses ETL

Rancangan awal proses ETL hanya memiliki satu berkas sumber berupa data_mutasi_stok.csv yang kolom jenis mutasi, nama barang, dan satuannya sudah tertulis lengkap. Rancangan tersebut menuntut petugas menggabungkan data penjualan dan pembelian secara manual sebelum proses dijalankan, sehingga ketelitian sistem masih bergantung pada ketelitian manusia. Penelusuran ulang terhadap pencatatan di toko menunjukkan bahwa data justru tersimpan terpisah menurut kegiatannya, sehingga rancangan direvisi agar membaca ketiga berkas sumber sekaligus. Perbandingan kedua rancangan disajikan pada Tabel 4.1.

Copyright@2026 FST PSU Bekasi

Tabel 4.1 Perbandingan Rancangan ETL Sebelum dan Sesudah Revisi

Aspek	Sebelum revisi	Sesudah revisi
Berkas sumber	Satu berkas gabungan	Tiga berkas terpisah
Jenis mutasi	Diketik petugas	Nilai tetap dari Add constants
Nama dan satuan	Diketik ulang tiap baris	Hasil Database lookup
Jumlah alur	Satu transformation	Tiga transformation satu job
Risiko kesalahan	Tinggi	Rendah
Kesesuaian kondisi toko	Rendah	Tinggi

Selain bentuk tiga transformation tersebut, disiapkan pula bentuk alternatif berupa satu transformation yang memuat ketiga sumber data pada satu kanvas. Urutan proses pada bentuk ini dijaga oleh langkah Block this step until steps finish yang menahan baris transaksi sampai pemuatan tabel dimensi selesai. Implementasi bentuk alternatif tersebut ditunjukkan pada Gambar 4.1, dan kedua bentuk menghasilkan isi tabel fakta yang sama.

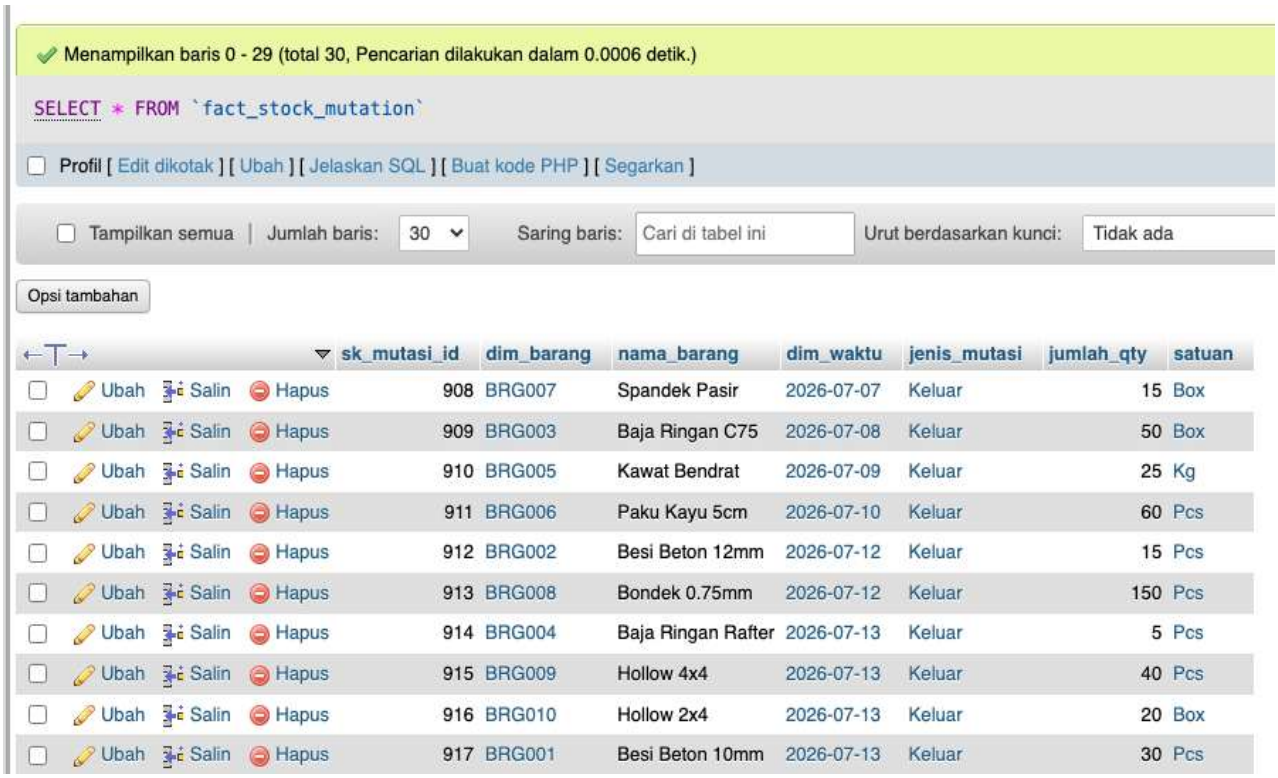
4.2. Hasil Pengujian Proses ETL

Pengujian dilakukan dengan menjalankan job pengendali, lalu mencatat jumlah baris yang dibaca dan yang ditulis pada setiap transformation. Hasilnya disajikan pada Tabel 4.2 Jumlah baris dibaca sama dengan jumlah baris ditulis pada seluruh transformation, sehingga tidak ada baris yang hilang selama proses berlangsung.

Tabel 4.2 Hasil Pengujian Proses ETL Tiga Sumber Data

Transformation	Dibaca	Ditulis	Status
01_load_dim_barang	10	10	Berhasil
02_load_penjualan	14	14	Berhasil
03_load_pembelian	16	16	Berhasil
job_etl_natalaya	40	40	Berhasil

Catatan log eksekusi menunjukkan keempat langkah job, yaitu pengosongan tabel fakta beserta ketiga transformation, seluruhnya bernilai true tanpa satu pun galat dan rangkaian selesai dalam waktu kurang dari satu detik. Setelah job dijalankan, tabel fakta terisi tiga puluh baris dengan rincian empat belas baris bertanda Keluar sebanyak 452 satuan dan enam belas baris bertanda Masuk sebanyak 1.675 satuan, sehingga saldo stok keseluruhan berjumlah 1.223 satuan pada periode 7 sampai 13 Juli 2026. Isi tabel fakta hasil pemuatan tersebut ditunjukkan pada Gambar 4.2.


Menampilkan baris 0 - 29 (total 30, Pencarian dilakukan dalam 0.0006 detik.)

```
SELECT * FROM `fact_stock_mutation`
```

Profil [Edit dikotak] [Ubah] [Jelaskan SQL] [Buat kode PHP] [Segarkan]

Tampilkan semua | Jumlah baris: 30 Saring baris: Cari di tabel ini Urut berdasarkan kunci: Tidak ada

Opsi tambahan

	sk_mutasi_id	dim_barang	nama_barang	dim_waktu	jenis_mutasi	jumlah_qty	satuan
Ubah Salin Hapus	908	BRG007	Spandek Pasir	2026-07-07	Keluar	15	Box
Ubah Salin Hapus	909	BRG003	Baja Ringan C75	2026-07-08	Keluar	50	Box
Ubah Salin Hapus	910	BRG005	Kawat Bendrat	2026-07-09	Keluar	25	Kg
Ubah Salin Hapus	911	BRG006	Paku Kayu 5cm	2026-07-10	Keluar	60	Pcs
Ubah Salin Hapus	912	BRG002	Besi Beton 12mm	2026-07-12	Keluar	15	Pcs
Ubah Salin Hapus	913	BRG008	Bondek 0.75mm	2026-07-12	Keluar	150	Pcs
Ubah Salin Hapus	914	BRG004	Baja Ringan Rafter	2026-07-13	Keluar	5	Pcs
Ubah Salin Hapus	915	BRG009	Hollow 4x4	2026-07-13	Keluar	40	Pcs
Ubah Salin Hapus	916	BRG010	Hollow 2x4	2026-07-13	Keluar	20	Box
Ubah Salin Hapus	917	BRG001	Besi Beton 10mm	2026-07-13	Keluar	30	Pcs

Gambar 4.2
Isi Tabel Fakta fact_stock_mutation Hasil ETL

Pengujian tambahan dilakukan dengan menambahkan satu kode barang baru pada data_barang.csv tanpa menambah baris transaksi. Barang tersebut muncul pada tabel dimensi dan pada tabel master di dashboard dengan saldo nol, sedangkan tabel fakta tidak berubah. Hasil tersebut menunjukkan bahwa pemisahan jalur pemuatan data master dan data transaksi berjalan sesuai rancangan.

4.3. Isi Gudang Data dan Saldo Stok

Kolom sk_mutasi_id pada tabel fakta berperan sebagai kunci pengganti yang memastikan setiap baris transaksi memiliki identitas unik, sedangkan kolom dim_barang dan dim_waktu menjadi penghubung ke tabel dimensi. Saldo stok setiap barang dihitung dari selisih jumlah masuk dan jumlah keluar pada tabel fakta, sehingga satu tabel dimensi dapat dipakai bersama oleh data penjualan maupun data pembelian. Hasil perhitungannya disajikan pada Tabel 4.3.

Tabel 4.3 Saldo Stok Hasil Perhitungan Tabel Fakta

Kode	Nama barang	Masuk	Keluar	Saldo
BRG 001	Besi Beton 10mm	240	30	210
BRG 002	Besi Beton 12mm	170	27	143
BRG 003	Baja Ringan C75	185	50	135
BRG 004	Baja Ringan Rafter	105	5	100
BRG 005	Kawat Bendrat	50	40	10
BRG 006	Paku Kayu 5cm	110	60	50
BRG 007	Spandek Pasir	30	25	5
BRG 008	Bondek 0.75mm	560	150	410
BRG 009	Hollow 4x4	135	40	95
BRG 010	Hollow 2x4	90	25	65

Angka pada Tabel 4.3 dinyatakan dalam satuan masing masing barang sebagaimana tercatat pada tabel dimensi dim_barang, yaitu Pcs, Box, atau Kg.

4.4. Dashboard Visualisasi Kinerja Penjualan

Hasil pemuatan data disajikan pada dashboard seperti terlihat pada Gambar 4.3. Indikator kinerja utama berupa total barang masuk, total barang keluar, dan sisa stok dihitung langsung dari tabel fakta, sehingga pemilik toko dapat menilai kondisi usaha tanpa membuka basis data. Dashboard juga menampilkan jumlah baris dan waktu perubahan terakhir setiap berkas sumber sebagai penanda keterlacakan, memisahkan tampilan data master dari data transaksi, serta menyediakan penyaringan menurut jenis mutasi dan tanggal. Penyajian sekali pandang seperti ini sejalan dengan prinsip rancangan dashboard yang dikemukakan Few (2013).

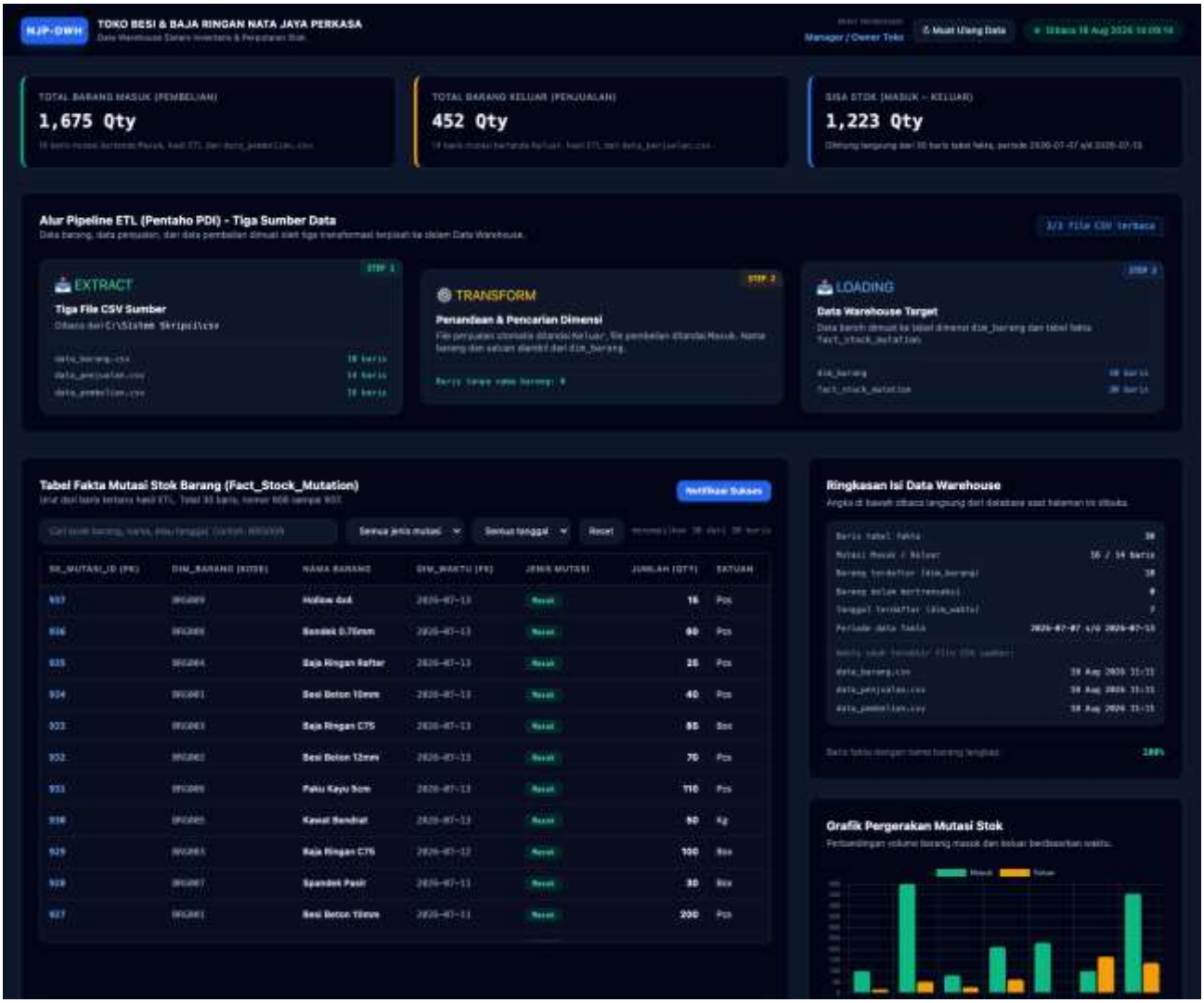
4.5. Kelayakan Model

Kelayakan model dinilai dari keberhasilan pembacaan seluruh sumber data, kelengkapan hasil transformasi, konsistensi hasil ketika proses diulang, kecepatan proses, serta kemudahan pemakaian. Rangkuman penilaiannya disajikan pada Tabel 4.4.

Tabel 4.4 Indikator Kelayakan Model

Indikator	Hasil evaluasi	Status
Kelengkapn sumber data	Tiga dari tiga berkas terbaca	Sangat layak
Kelengkapan transformasi	Seluruh 30 baris fakta lengkap keterangannya	Sangat layak
Konsistensi pengulangan	Jumlah baris tetap 30 setiap kali job dijalankan	Sangat layak
Kecepatan proses	Rangkaian ETL selesai kurang dari satu detik	Sangat layak
Kemudahan pemakaian	Proses dijalankan dengan satu perintah	Layak

Copyright@2026 FST PSU Bekasi



Gambar 4.3
Dashboard Visualisasi Hasil Gudang Data

4.6. Pembahasan

Hasil pengujian menunjukkan bahwa persoalan data yang terfragmentasi pada toko tidak diselesaikan dengan menambah pekerjaan penggabungan, melainkan dengan memindahkan pekerjaan tersebut kepada sistem. Penanda jenis mutasi yang sebelumnya diketik petugas kini diisi nilai tetap sesuai berkas asalnya, sedangkan nama barang dan satuan diambil dari tabel dimensi. Perubahan itu menghilangkan sumber kesalahan penulisan yang selama ini paling sering muncul, sekaligus menegaskan manfaat pemisahan atribut keterangan ke tabel dimensi sebagaimana dianjurkan pada pemodelan dimensional (Kimball & Ross, 2013).

Pemakaian job sebagai pengendali urutan terbukti penting. Pemuatan tabel dimensi yang selalu berjalan sebelum tabel fakta membuat setiap baris transaksi menemukan pasangan dimensinya, sehingga tidak ada baris fakta yang kehilangan nama barang. Pengosongan tabel fakta pada awal rangkaian juga menjaga jumlah baris tetap sama pada setiap pengulangan, yang menjadi dasar sifat dapat diulang pada proses ETL (Vassiliadis, 2009). Temuan ini sejalan dengan anjuran Kimball dkk. (2008) mengenai urutan pemuatan gudang data dan memperlihatkan bahwa anjuran tersebut tetap berlaku pada skala usaha kecil.

Dibandingkan praktik sebelumnya, penyajian informasi menjadi lebih cepat karena pemilik toko tidak lagi menunggu penggabungan berkas secara manual. Perbaikan tersebut senada dengan temuan Wixom dan Watson (2001) bahwa mutu data menentukan manfaat gudang data yang dirasakan pengguna, serta dengan penempatan dashboard sebagai lapisan akhir rantai business intelligence yang bertumpu pada integrasi data di belakangnya (Chaudhuri dkk., 2011; Sharda dkk., 2020).

Keterbatasan penelitian perlu dicatat. Pengujian dilakukan pada volume data satu pekan yang relatif kecil, sehingga kecepatan proses kurang dari satu detik belum dapat digeneralisasi untuk volume data beberapa tahun. Proses ETL juga masih dijalankan secara manual dan

Copyright@2026 FST PSU Bekasi

gudang data belum memuat dimensi pelanggan maupun wilayah, sehingga analisis kinerja penjualan masih terbatas pada dimensi barang dan waktu.

5. KESIMPULAN DAN SARAN

Data warehouse bermodel skema bintang dengan proses Extract, Transform, Load otomatis menggunakan Pentaho Data Integration berhasil diimplementasikan pada Toko Besi dan Baja Ringan Nata Jaya Perkasa. Rancangan yang semula bertumpu pada satu berkas gabungan direvisi menjadi tiga transformation yang membaca data barang, data penjualan, dan data pembelian, dan dikendalikan oleh satu job sehingga urutan pemuatan tabel dimensi dan tabel fakta selalu terjaga.

Pengujian menunjukkan seluruh berkas sumber terbaca tanpa galat dengan jumlah baris dibaca sama dengan jumlah baris ditulis pada setiap transformation, tabel fakta terisi tiga puluh baris dengan 452 satuan barang keluar dan 1.675 satuan barang masuk sehingga saldo stok berjumlah 1.223 satuan, rangkaian proses selesai kurang dari satu detik, dan jumlah baris tetap sama pada setiap pengulangan. Dashboard yang dibangun menyajikan indikator kinerja beserta keterlacakan sumber data, sehingga pemilik toko memperoleh informasi penjualan tanpa menunggu penggabungan berkas secara manual.

Penelitian selanjutnya disarankan menambahkan penjadwalan otomatis agar proses ETL berjalan tanpa dipicu secara manual, memperluas gudang data dengan dimensi pelanggan dan wilayah untuk memperdalam analisis kinerja penjualan, serta menguji sistem pada volume data beberapa tahun agar kinerja proses pada data berukuran besar dapat diukur.

DAFTAR PUSTAKA

Bouman, R., & van Dongen, J. (2009). Pentaho solutions: Business intelligence and data warehousing with Pentaho and MySQL. Indianapolis: Wiley Publishing.

Casters, M., Bouman, R., & van Dongen, J. (2010). Pentaho Kettl solutions: Building open source ETL solutions with Pentaho Data Integration. Indianapolis: Wiley Publishing.

Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. ACM SIGMOD Record, 26(1), 65-74.

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. Communications of the ACM, 54(8), 88-98.

Few, S. (2013). Information dashboard design: Displaying data for at-a-glance monitoring (2nd ed.). Burlingame: Analytics Press.

Golfarelli, M., & Rizzi, S. (2009). Data warehouse design: Modern principles and methodologies. New York: McGraw-Hill.

Hitachi Vantara. (2023). Pentaho Data Integration documentation. Santa Clara: Hitachi Vantara.

Inmon, W. H. (2005). Building the data warehouse (4th ed.). Indianapolis: Wiley Publishing.

Kimball, R., & Ross, M. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling (3rd ed.). Indianapolis: John Wiley & Sons.

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2008). The data warehouse lifecycle toolkit (2nd ed.). Indianapolis: John Wiley & Sons.

Ponniah, P. (2010). Data warehousing fundamentals for IT professionals (2nd ed.). Hoboken: John Wiley & Sons.

Sharda, R., Delen, D., & Turban, E. (2020). Analytics, data science, and artificial intelligence: Systems for decision support (11th ed.). Harlow: Pearson Education.

Vaisman, A., & Zimanyi, E. (2022). Data warehouse systems: Design and implementation (2nd ed.). Berlin: Springer.

Vassiliadis, P. (2009). A survey of extract-transform-load technology. International Journal of Data Warehousing and Mining, 5(3), 1-27.

Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. MIS Quarterly, 25(1), 17-41.

Copyright@2026 FST PSU Bekasi