



IMPLEMENTASI ARSITEKTUR DATA WAREHOUSE DAN DASHBOARD BUSINESS INTELLIGENCE UNTUK ANALISIS PERFORMA PENJUALAN RITEL ELEKTRONIK DI MYANMAR MENGGUNAKAN PENTAHO DAN TABLEAU

Taqiuddin Ahmad Al Aufa¹⁾, Adriano Femaz Rivaldy²⁾, Farel Ega Nurroyan³⁾, Muhammad Fawwaz Dhiaulhaq Hud⁴⁾, Amalia Anjani Arifiyanti⁵⁾

- ¹⁾ Sistem Informasi, Fakultas Ilmu Komputer, UPN Veteran Jawa Timur, Surabaya, Indonesia
Email: 23082010135@student.upnjatim.ac.id
- ²⁾ Sistem Informasi, Fakultas Ilmu Komputer, UPN Veteran Jawa Timur, Surabaya, Indonesia
Email: 23082010139@student.upnjatim.ac.id
- ³⁾ Sistem Informasi, Fakultas Ilmu Komputer, UPN Veteran Jawa Timur, Surabaya, Indonesia
Email: 23082010143@student.upnjatim.ac.id
- ⁴⁾ Sistem Informasi, Fakultas Ilmu Komputer, UPN Veteran Jawa Timur, Surabaya, Indonesia
Email: 23082010127@student.upnjatim.ac.id
- ⁵⁾ Sistem Informasi, Fakultas Ilmu Komputer, UPN Veteran Jawa Timur, Surabaya, Indonesia
Email: amalia_anjani.fik@upnjatim.ac.id

Abstract

This study aims to implement a Data Warehouse architecture and a Business Intelligence dashboard to analyze the electronic retail sales performance in Myanmar. Raw data sourced from the Electronic Retail Dataset (2020–2023) has high complexity and poor quality (dirty data). The methods used include building a data pipeline with PostgreSQL as a staging area, the Extract, Transform, Load (ETL) process using Pentaho Data Integration, Star Schema modeling, and interactive visualization using Tableau via a Virtual Datamart. The results show that the ETL process successfully cleaned data anomalies and produced an integrated schema. The resulting dashboard reveals strategic insights, such as the revenue dominance of "Laptop" products versus the high sales volume of "Desks", and the market's reliance on physical B2B channels. Furthermore, spatial disparities were identified, with Yangon and Bago serving as revenue centers, while other regions require strategic intervention. This implementation proves that the integration of Pentaho and Tableau can transform complex transactional data into visual information that effectively supports data-driven business decision-making.

Keywords: Data Warehouse, Business Intelligence, ETL, Pentaho, Tableau.

Abstrak

Penelitian ini bertujuan untuk mengimplementasikan arsitektur *Data Warehouse* dan *dashboard Business Intelligence* guna menganalisis performa penjualan ritel elektronik di Myanmar. Data mentah yang bersumber dari Electronic Retail Dataset (2020–2023) memiliki kompleksitas tinggi dan kualitas yang rendah (*dirty data*). Metode yang digunakan meliputi pembangunan data pipeline dengan PostgreSQL sebagai *staging area*, proses *Extract, Transform, Load* (ETL) menggunakan Pentaho Data Integration, pemodelan Star Schema, dan visualisasi interaktif menggunakan Tableau melalui *Virtual Datamart*. Hasil penelitian menunjukkan bahwa proses ETL berhasil membersihkan anomali data dan menghasilkan skema yang terintegrasi. *Dashboard* yang dihasilkan mengungkap wawasan strategis, seperti dominasi pendapatan dari produk "Laptop" namun tingginya volume penjualan "Desk", serta ketergantungan pasar pada saluran B2B fisik. Selain itu, teridentifikasi ketimpangan spasial di mana Yangon dan Bago menjadi pusat pendapatan, sementara wilayah lain memerlukan intervensi strategis. Implementasi ini membuktikan bahwa integrasi Pentaho dan Tableau mampu mengubah data transaksi yang kompleks menjadi informasi visual yang mendukung pengambilan keputusan bisnis berbasis data secara efektif.

Kata Kunci: *Data Warehouse*, *Business Intelligence*, ETL, Pentaho, Tableau.



PENDAHULUAN

Di era digital saat ini, volume data yang dihasilkan dari berbagai aktivitas manusia dan proses bisnis mengalami pertumbuhan yang sangat pesat. Data tidak lagi dipandang sebagai sekadar catatan historis, melainkan telah menjadi aset strategis yang memiliki nilai tinggi bagi organisasi dalam mendukung proses pengambilan keputusan. Kemampuan untuk mengolah data mentah menjadi informasi yang akurat, relevan, dan mudah dipahami menjadi faktor penting dalam menghasilkan keputusan yang efektif dan berbasis data (Dharsika, 2025).

Meskipun demikian, data yang diperoleh dari berbagai sumber seringkali memiliki kualitas yang kurang memadai untuk langsung digunakan dalam proses analisis. Permasalahan seperti format data yang tidak seragam, nilai yang hilang (*missing values*), duplikasi data, serta ketidakkonsistenan atribut masih sering ditemukan pada data mentah (Branchris, 2026). Oleh karena itu, diperlukan proses *data engineering* dan *data analytics* yang terstruktur untuk memastikan kualitas data yang digunakan dalam analisis. Salah satu pendekatan yang umum diterapkan adalah pembangunan *data pipeline* yang mencakup proses integrasi data, penyimpanan data pada staging area, transformasi data ke dalam *data warehouse*, hingga penyajian informasi melalui visualisasi data.

Implementasi *data warehouse* memerlukan proses pembersihan dan integrasi data yang baik agar informasi yang dihasilkan memiliki tingkat akurasi yang tinggi. Pentaho Data Integration dapat dimanfaatkan untuk mengotomatisasi proses *Extract, Transform, Load* (ETL), sedangkan Tableau digunakan untuk menyajikan hasil analisis dalam bentuk dashboard interaktif. Melalui visualisasi yang efektif, data yang kompleks dapat disajikan menjadi informasi yang lebih mudah dipahami sehingga mampu mendukung proses pengambilan keputusan oleh pemangku kepentingan.

Penelitian ini menggunakan Electronic Retail Dataset yang bersumber dari Kaggle. Dataset tersebut dipilih karena merepresentasikan aktivitas transaksi pada industri ritel elektronik yang memiliki karakteristik data yang kompleks dan dinamis. Selain itu, dataset ini mengandung berbagai permasalahan kualitas data yang umum ditemukan pada lingkungan bisnis nyata sehingga relevan untuk diterapkan dalam pembangunan *data warehouse* dan *business intelligence*.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan arsitektur *data warehouse* dan *dashboard business intelligence* guna menganalisis performa penjualan ritel elektronik di Myanmar. Implementasi dilakukan melalui proses integrasi data ke PostgreSQL, pembangunan staging area, proses ETL menggunakan Pentaho Data Integration, perancangan *data warehouse* dengan pendekatan *Star Schema*, serta visualisasi data menggunakan Tableau. Hasil penelitian diharapkan dapat memberikan gambaran mengenai penerapan *data pipeline* yang terstruktur dalam mengolah data transaksi ritel elektronik menjadi informasi yang bernilai serta mendukung pengambilan keputusan bisnis berbasis data. Adapun ruang lingkup penelitian dibatasi pada penggunaan Electronic Retail Dataset yang diperoleh

dari Kaggle dengan fokus analisis pada performa penjualan produk, wilayah penjualan, saluran penjualan, metode pembayaran, dan karakteristik pelanggan.

TINJAUAN PUSTAKA

Business Intelligence

Business Intelligence (BI) merupakan sekumpulan teknologi, aplikasi, dan proses yang digunakan untuk mengubah data mentah menjadi informasi yang bernilai guna mendukung pengambilan keputusan dalam organisasi. Melalui *Business Intelligence*, perusahaan dapat mengumpulkan, mengelola, dan menganalisis data sehingga menghasilkan informasi yang lebih akurat untuk meningkatkan efektivitas dan efisiensi operasional bisnis (Tumini & Subekti, 2023). Seiring meningkatnya volume data pada berbagai sektor industri, termasuk *e-commerce*, penerapan *Business Intelligence* menjadi semakin penting untuk membantu organisasi mengidentifikasi pola, tren, dan peluang bisnis dari data yang dimiliki (Anardani et al., 2023). Berbagai platform *Business Intelligence* saat ini tersedia dengan karakteristik yang berbeda, seperti kemudahan penggunaan, kemampuan visualisasi, integrasi data, serta aksesibilitas. Oleh karena itu, pemilihan platform *Business Intelligence* yang tepat menjadi faktor penting dalam mendukung proses analisis data dan pengambilan keputusan yang efektif (Julfia et al., 2024).

Data Warehouse

Data warehouse merupakan basis data relasional yang dirancang untuk mendukung kebutuhan analisis dan pengambilan keputusan, bukan untuk memproses transaksi operasional sehari-hari. *Data warehouse* umumnya menyimpan data historis yang berasal dari proses transaksi maupun sumber data lainnya sehingga memungkinkan integrasi data dalam satu repositori terpusat. Selain itu, *data warehouse* memiliki karakteristik utama, yaitu terintegrasi (*integrated*), berorientasi pada subjek (*subject-oriented*), memiliki dimensi waktu (*time-variant*), dan bersifat permanen (*non-volatile*). Dengan karakteristik tersebut, *data warehouse* mampu menyediakan informasi yang lebih terstruktur dan mendukung proses analisis secara efektif sehingga dapat membantu organisasi dalam pengambilan keputusan yang lebih tepat (Hasan & Febriandirza, 2021).

Extract, Transform, Load (ETL)

Extract, Transform, Load (ETL) merupakan proses utama dalam pembangunan *data warehouse* yang berfungsi untuk memindahkan data dari sumber operasional menuju lingkungan *data warehouse*. Tahapan *extract* dilakukan untuk mengambil data dari berbagai sumber, *transform* digunakan untuk membersihkan, memvalidasi, dan menyesuaikan format data sesuai kebutuhan analisis, sedangkan *load* merupakan proses memasukkan data yang telah ditransformasi ke dalam *data warehouse*. Proses ETL memiliki peran penting dalam menjaga kualitas, konsistensi, dan integritas data yang digunakan untuk analisis (Ghita & Trisminingsih, 2021)

Implementasi ETL juga memungkinkan dilakukannya proses data cleansing untuk menangani permasalahan



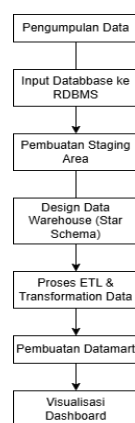
seperti data duplikat, inkonsistensi format, serta nilai yang hilang. Dengan demikian, data yang tersimpan pada *data warehouse* menjadi lebih valid dan siap digunakan untuk kebutuhan pelaporan maupun *business intelligence*. Salah satu perangkat lunak yang banyak digunakan dalam implementasi ETL adalah Pentaho Data Integration karena mampu mengotomatisasi proses transformasi data secara efektif.

Visualisasi Data

Visualisasi data merupakan proses transformasi informasi kompleks menjadi representasi grafis intuitif yang berfungsi sebagai jembatan kognitif antara data mentah dan pengambilan keputusan strategis, terutama dalam era *big data* di mana metode penyajian tradisional tidak lagi memadai. Efektivitas visualisasi sangat bergantung pada pemilihan teknik yang tepat sesuai dengan karakteristik dan dimensi data, seperti penggunaan *line charts* atau *heat maps* untuk data temporal (tren waktu), *treemaps* dan *donut charts* untuk data proporsional (komposisi kategori), serta *scatter plots* dan pemetaan geospasial untuk data relasional. Dalam konteks *Business Intelligence*, penerapan teknik visualisasi yang interaktif dan sesuai dengan jenis data ini memungkinkan analisis untuk tidak hanya menyajikan informasi, tetapi juga mengungkap pola tersembunyi, korelasi, serta tren dinamis yang pada akhirnya mendukung perumusan wawasan bisnis.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan *data engineering* dan *business intelligence* dengan metode eksperimental berbasis pengembangan sistem (*system development*). Pendekatan ini mengadopsi siklus hidup pengembangan *Data Warehouse* menggunakan metodologi *dimensional modeling* (Kimball) yang berfokus pada pembuatan *Star Schema* untuk mendukung analisis bisnis yang cepat dan intuitif. Penelitian dimulai dari tahap pengumpulan data mentah, ingestasi ke dalam *Relational Database Management System* (RDBMS), pembersihan dan transformasi data menggunakan *Extract, Transform, Load* (ETL), perancangan *Data Warehouse*, hingga visualisasi data menjadi dashboard interaktif. Alur penelitian secara keseluruhan digambarkan pada Gambar 1.



Gambar 1. Diagram Alur Penelitian

Pengumpulan Data

Tahap pengumpulan data dilakukan dengan memanfaatkan dataset publik yang bersumber dari platform Kaggle, yaitu Electronic Retails Dataset (somesomeall, 2026). Dataset ini dipilih karena merepresentasikan kompleksitas transaksi ritel elektronik di Myanmar pada periode 2020–2023, yang mencakup 13 kolom variabel seperti informasi waktu, detail produk, metrik penjualan, wilayah, saluran penjualan, dan metode pembayaran. Karakteristik khusus dari dataset ini adalah kondisi *dirty data* yang mengandung *missing values*, inkonsistensi format, dan *outliers* (Branchris, 2026). Kondisi ini sangat relevan untuk menguji efektivitas proses *data cleansing* dan rekayasa fitur dalam membangun *data pipeline* yang robust sebelum masuk ke tahap *Data Warehouse*.

Transformasi dan Pemodelan Data

Transformasi dan pemodelan data merupakan inti dari penelitian ini, yang dibagi menjadi tiga sub-tahapan utama:

1. Ingestasi dan Staging Area

Data mentah dalam format CSV diingestasi ke dalam PostgreSQL menggunakan antarmuka DBeaver. Data disimpan terlebih dahulu ke dalam staging area (skema staging) dengan tipe data teks untuk mencegah kegagalan impor akibat anomali format. Staging area berfungsi sebagai zona penampungan sementara yang menjaga integritas data sumber asli sebelum dilakukan transformasi.

2. Proses ETL dengan Pentaho Data Integration

Proses *Extract, Transform, Load* (ETL) dilakukan menggunakan Pentaho Data Integration (PDI) untuk menangani isu kualitas data. Tahap transform mencakup penanganan *missing values* (mengganti nilai null dengan "Unknown" atau 0), normalisasi teks, dan konversi tipe data. Pendekatan ini sejalan dengan penelitian Dharsika (2025) yang menyatakan bahwa otomatisasi ETL menggunakan Pentaho sangat efektif dalam menjamin kualitas dan konsistensi data sebelum masuk ke lingkungan analitik.

3. Dimensional Modeling (Star Schema)

Pemodelan data dilakukan menggunakan pendekatan *Star Schema* yang terdiri dari satu tabel fakta (*fact_sales*) dan enam tabel dimensi (*dim_date*, *dim_product*, *dim_region*, *dim_sales_channel*, *dim_payment_method*, dan *dim_customer_type*). Penggunaan surrogate key dan pemisahan antara tabel fakta (metrik transaksi) dan tabel dimensi (konteks bisnis) bertujuan untuk mengoptimalkan performa query analitik dan memudahkan proses agregasi data (Hasan & Febriandirza, 2021).

Visualisasi dan Analisis Bisnis

Tahap akhir dari metodologi ini adalah penyajian informasi melalui *Business Intelligence*. Alih-alih membuat tabel fisik baru yang redundan, penelitian ini mengimplementasikan *Virtual Datamart* menggunakan fitur *Database View* pada PostgreSQL (*vw_fact_sales_flat*). *View* ini melakukan join permanen antara tabel fakta dan dimensi, menyajikan data dalam format flat table yang denormalisasi.

Data dari *Virtual Datamart* kemudian dihubungkan ke Tableau untuk membangun dashboard interaktif. Pemilihan teknik visualisasi seperti *line chart* untuk tren temporal, *bubble chart* untuk analisis spasial, dan *heatmap* untuk



perilaku pelanggan dilakukan untuk menjembatani kesenjangan kognitif antara data mentah dan pengambilan keputusan strategis. Integrasi antara SQL untuk manajemen datamart dan Tableau untuk visualisasi terbukti mampu menghasilkan dashboard yang agile dan mendukung decision support system secara efektif (Akter & Mamun, 2023).

HASIL DAN PEMBAHASAN

Arsitektur Data Warehouse

Implementasi arsitektur *Data Warehouse* pada proyek ini menggunakan pendekatan *Star Schema* yang terdiri dari satu tabel fakta (*fact_sales*) dan enam tabel dimensi (*dim_date*, *dim_product*, *dim_region*, *dim_sales_channel*, *dim_payment_method*, dan *dim_customer_type*). Pendekatan ini dipilih karena kemampuannya dalam menyederhanakan struktur data relasional yang kompleks menjadi format yang optimal untuk query analitik dan agregasi data.

Sebagai lapisan abstraksi antara basis data fisik dan alat visualisasi, dibangunlah sebuah *Virtual Datamart* berupa *Database View* bernama *dwh.vw_fact_sales_flat*. Berbeda dengan *Materialized View*, pendekatan virtual ini memastikan bahwa setiap pembaruan data melalui proses ETL di Pentaho akan langsung tercermin di Tableau tanpa memerlukan proses refresh manual. Struktur ini berhasil menjembatani kesenjangan antara normalisasi database PostgreSQL dengan kebutuhan denormalisasi data untuk pelaporan bisnis yang cepat.

penilaian pelanggan, pencilaan data berupa *record* dengan kuantitas negatif yang mengindikasikan retur atau kesalahan pencatatan, serta format tanggal yang tidak seragam yang dapat menghambat analisis deret waktu.

Untuk membangun tabel dimensi kategorikal, pendekatan transformasi berfokus pada standardisasi dan retensi data tanpa menghilangkan konteks transaksi. Pada dimensi produk, atribut diekstrak dan diurutkan menggunakan step Sort Rows sebelum dimuat ke tabel target, di mana *record* dengan nilai produk yang kosong secara tegas difilter dan dibuang karena nama produk merupakan entitas identitas yang mutlak. Sebaliknya, pada dimensi wilayah, saluran penjualan, metode pembayaran, dan tipe pelanggan, strategi penanganan nilai hilang tidak dilakukan dengan pembuangan baris, melainkan dengan mengisi nilai kosong menggunakan label "*Unknown*" melalui step If field value is null. Pendekatan ini mencegah hilangnya data transaksi akibat atribut pendukung yang tidak tercatat. Selanjutnya, normalisasi teks dilakukan menggunakan step Replace in string dan fungsi kapitalisasi untuk menyamakan variasi penulisan, serta step Value Mapper untuk mengelompokkan kategori saluran penjualan guna menyederhanakan analisis.

Berdiferensiasi dari dimensi kategorikal yang diekstrak dari data transaksi, tabel dimensi waktu dibangun menggunakan pendekatan generatif melalui kueri native PostgreSQL. Fungsi *generate_series()* dimanfaatkan untuk menghasilkan deret tanggal kontinu yang mencakup seluruh rentang waktu dataset tanpa celah (*gap*), bahkan untuk tanggal di mana tidak terjadi transaksi sama sekali. Atribut temporal analitis seperti hari dalam seminggu, nama bulan, kuartal, tahun, serta penanda akhir pekan dikomputasi langsung di level *database* menggunakan fungsi *EXTRACT()* dan *TO_CHAR()*. Strategi pemindahan logika komputasi ke sisi *database* ini secara signifikan meminimalkan beban pemrosesan pada aplikasi ETL dan menjamin kontinuitas analisis waktu yang akurat.

Pembangunan tabel fakta memerlukan alur transformasi yang kompleks untuk mengakomodasi aturan bisnis dan pembersihan data tingkat lanjut. Tahap ini diawali dengan pembersihan dan validasi menggunakan Modified JavaScript Value untuk mem-parse format tanggal dan mengonversi persentase diskon, diikuti dengan penyaringan *record* yang memiliki kuantitas nol atau negatif untuk memastikan hanya transaksi penjualan valid yang diproses. Penanganan nilai hilang pada tabel fakta diterapkan secara kontekstual; kolom numerik seperti diskon dan kuantitas diisi dengan nilai nol agar tidak mendistorsi agregasi matematis, sementara kolom subjektif seperti penilaian pelanggan dikonversi menjadi *Not a Number* (NaN) untuk membedakan secara eksplisit antara nilai nol dan data yang tidak tersedia. Inti dari transformasi fakta ini terletak pada proses denormalisasi menggunakan serangkaian step Database Lookup yang menerjemahkan kunci natural dari data transaksi menjadi *surrogate key* yang merujuk pada *primary key* di masing-masing tabel dimensi.

Pada tahap akhir transformasi fakta, dilakukan pengayaan metrik finansial untuk menghitung indikator bisnis utama. Step Calculator dan skrip JavaScript lanjutan digunakan untuk menurunkan nilai pendapatan kotor,



Gambar 2. Desain Star Schema Data Warehouse

Evaluasi Kualitas Data dan Proses ETL

Tahap *Extract, Transform, Load* (ETL) merupakan fondasi kritis dalam pembangunan *Data Warehouse* untuk menjamin integritas, konsistensi, dan kesiapan data sebelum masuk ke tahap analisis. Proses ini dieksekusi menggunakan Pentaho Data Integration (PDI) sebagai mesin ETL utama, dengan PostgreSQL sebagai target database. Berdasarkan hasil data profiling pada lapisan staging, dataset mentah memiliki karakteristik *dirty data* yang merepresentasikan kompleksitas data dunia nyata. Beberapa anomali kualitas data yang teridentifikasi meliputi inkonsistensi format teks pada atribut kategorikal seperti wilayah dan saluran penjualan, adanya nilai hilang (*missing values*) pada kolom krusial seperti persentase diskon dan



jumlah potongan diskon, dan total penjualan bersih setelah diskon. Melalui serangkaian tahapan ETL yang terstruktur dan otomatis ini, data mentah yang awalnya tidak terstruktur dan mengandung berbagai anomali berhasil ditransformasikan menjadi skema *Star Schema* yang ternormalisasi. Hasil akhir dari pipeline ini adalah dataset analitis yang akurat, memiliki integritas referensial yang kuat, dan siap dikonsumsi oleh lapisan *Datamart* untuk keperluan visualisasi dan pelaporan bisnis.

id_sales_fact	id_fact_sales	id_time	id_dim_date	id_dim_time	id_dim_store	id_dim_product	id_dim_customer	id_dim_promotion	id_dim_channel	id_dim_device	id_dim_location	id_dim_status
3523	20.000.001	8.07	1.10	4.07	4.07	2.07	2	378.000	5	712.000	1	700.000
3524	20.000.002	1.10	2.01	3.01	3.01	1.01	11	790.000	3	167.000	1	167.000
3525	20.000.003	7.10	7.01	4.01	4.01	1.01	14	423.000	9	476.000	1	476.000
3526	20.000.004	7.10	7.01	4.01	4.01	2.01	3	423.000	9	187.000	1	187.000
3527	20.000.005	2.10	2.01	1.01	1.01	1.01	4	180.000	11	176.000	1	176.000
3528	20.000.006	7.10	7.01	4.01	4.01	4.01	5	423.000	10	281.000	1	312.000
3529	20.000.007	2.10	2.01	3.01	4.01	2.01	2	120.000	15	212.000	1	200.000
3530	20.000.008	2.10	2.01	4.01	3.01	4.01	2	200.000	15	176.000	1	200.000
3531	20.000.009	1.10	1.01	1.01	1.01	4.01	15	101.000	10	1.601.000	1	1.676.000
3532	20.000.010	8.07	8.01	2.01	4.01	4.01	13	423.000	9	4.126.000	1	4.126.000
3533	20.000.011	2.10	2.01	1.01	2.01	2.01	12	300.000	8	2.020.000	1	2.020.000
3534	20.000.012	4.10	4.01	1.01	1.01	3.01	7	2.000.000	15	1.400.000	1	4.000.000
3535	20.000.013	8.07	8.01	3.01	4.01	4.01	5	423.000	15	2.404.250	1	3.120.000
3536	20.000.014	1.10	1.01	4.01	2.01	3.01	19	790.000	8	16.420.000	1	14.250.000
3537	20.000.015	7.10	7.01	1.01	1.01	1.01	7	423.000	5	476.000	1	476.000
3538	20.000.016	7.10	7.01	1.01	1.01	1.01	6	750.000	10	4.550.000	1	4.800.000
3539	20.000.017	1.10	1.01	1.01	1.01	4.01	8	790.000	5	5.700.000	1	6.000.000
3540	20.000.018	1.10	1.01	2.01	1.01	4.01	8	790.000	5	5.700.000	1	6.000.000
3541	20.000.019	7.10	7.01	2.01	1.01	2.01	11	423.000	9	487.500	1	487.500
3542	20.000.020	8.07	8.01	2.01	4.01	2.01	14	423.000	10	7.431.000	1	8.700.000
3543	20.000.021	4.10	4.01	1.01	1.01	2.01	18	2.000.000	15	30.000.000	1	30.000.000
3544	20.000.022	8.07	8.01	2.01	4.01	3.01	7	423.000	5	4.194.250	1	4.194.250
3545	20.000.023	1.10	1.01	4.01	1.01	1.01	15	1.200.000	5	17.816.000	1	18.700.000
3546	20.000.024	8.07	8.01	3.01	1.01	2.01	18	423.000	10	30.126.000	1	31.250.000
3547	20.000.025	5.10	5.01	3.01	4.01	4.01	10	1.200.000	15	18.620.000	1	19.600.000
3548	20.000.026	8.07	8.01	2.01	2.01	3.01	10	1.200.000	15	1.700.000	1	1.700.000
3549	20.000.027	4.10	4.01	1.01	1.01	1.01	8	2.000.000	8	18.000.000	1	18.000.000
3550	20.000.028	2.10	2.01	3.01	1.01	2.01	9	100.000	5	1.008.750	1	1.008.750
3551	20.000.029	6.10	6.01	2.01	1.01	2.01	14	423.000	15	7.431.000	1	8.700.000
3552	20.000.030	1.10	1.01	1.01	1.01	4.01	11	1.500.000	8	10.000.000	1	10.000.000
3553	20.000.031	4.10	4.01	3.01	2.01	4.01	9	2.000.000	15	17.000.000	1	18.000.000
3554	20.000.032	8.07	8.01	4.01	1.01	2.01	10	423.000	9	18.200.000	1	18.200.000
3555	20.000.033	4.10	4.01	1.01	4.01	4.01	8	2.000.000	10	34.200.000	1	36.000.000

Gambar 3. Tabel Fakta

Implementasi Dashboard Visualisasi

Tahap akhir dari siklus rekayasa data dalam penelitian ini adalah implementasi visualisasi menggunakan perangkat lunak Tableau. Untuk memastikan konektivitas yang efisien dan performa kueri yang optimal, Tableau tidak dihubungkan langsung ke tabel fakta dan dimensi yang ternormalisasi pada skema Data Warehouse. Sebagai gantinya, koneksi data diarahkan ke lapisan Virtual *Datamart* berupa *database view* (*vw_fact_sales_flat*) yang telah dibangun di PostgreSQL. Pendekatan ini menyajikan data dalam format flat table (tabel datar) yang telah ter-join secara logis, sehingga memudahkan proses pemetaan atribut ke dalam kanvas visualisasi tanpa memerlukan penulisan sintaks relasional yang berulang pada sisi aplikasi BI.

Dashboard yang dirancang diberi tajuk "Myanmar Electronic Retail Overview (2020-2023)" dengan tujuan utama menyediakan pemahaman komprehensif mengenai performa bisnis ritel elektronik. Antarmuka *dashboard* ini dirancang dengan prinsip interaktivitas, di mana seluruh komponen visualisasi terintegrasi dengan sebuah filter dinamis berbasis wilayah (*Region*). Fitur ini memungkinkan pengguna untuk mengisolasi dan menganalisis karakteristik pasar secara spesifik di setiap daerah, mengubah *dashboard* dari sekadar laporan statis menjadi instrumen eksplorasi data yang adaptif.

Panel paling atas *dashboard* didedikasikan untuk menampilkan *Key Performance Indicators* (KPI) yang merangkum metrik vital secara *real-time* sesuai dengan filter yang aktif. Metrik ini meliputi total pendapatan kotor (*Gross Revenue*), volume produk terjual, rata-rata persentase diskon, dan rata-rata tingkat kepuasan pelanggan. Di bawahnya, analisis fluktuasi temporal direpresentasikan melalui visualisasi *line chart* pada komponen Sales Trend yang memetakan pergerakan

pendapatan kotor per kuartal dalam mata uang MMK. Grafik ini dirancang untuk memfasilitasi identifikasi pola musiman serta titik lonjakan atau penurunan pendapatan secara historis.

Untuk mengakomodasi analisis spasial dan kategorial, digunakan visualisasi *bubble chart* pada komponen *Regional Sales*. Pada grafik ini, ukuran setiap *bubble* memproporsikan besaran kontribusi pendapatan dari masing-masing kategori produk di wilayah yang sedang dianalisis, memungkinkan manajemen untuk dengan cepat mengidentifikasi produk unggulan di setiap wilayah geografis. Selanjutnya, evaluasi efektivitas produk dilakukan menggunakan *dual-axis chart* yang mengkombinasikan *bar chart* untuk volume penjualan (kuantitas) dan *line chart* untuk nilai finansial (pendapatan kotor). Visualisasi ini secara eksplisit menonjolkan disparitas bisnis, seperti mengungkap produk yang paling laku secara unit dibandingkan dengan produk yang paling menguntungkan secara *revenue*.

Terakhir, analisis perilaku konsumen dipetakan menggunakan *heatmap chart* yang menyilangkan tipe pelanggan (*Customer Type*) dengan saluran penjualan (*Sales Channel*). Intensitas warna pada matriks ini mengindikasikan konsentrasi transaksi, yang secara visual membantu para pengambil keputusan dalam mengidentifikasi kanal distribusi yang paling efektif untuk setiap segmen pelanggan. Melalui implementasi ini, kompleksitas data pada arsitektur *Star Schema* berhasil diterjemahkan menjadi narasi visual yang intuitif, sehingga siap digunakan sebagai landasan strategis bagi pemangku kepentingan.



Gambar 4. Dashboard Tableau

Analisis Insight Dashboard

Implementasi *dashboard* "Myanmar Electronic Retail Overview (2020-2023)" pada Tableau tidak hanya berfungsi sebagai alat pemantauan metrik, tetapi juga sebagai instrumen diagnostik untuk mengekstrak wawasan strategis dari data historis. Secara makro, analisis performa keseluruhan (*overall performance*) mengungkap bahwa bisnis ritel elektronik ini berhasil membukukan total pendapatan kotor sebesar K9,79 Miliar dari penjualan 14.630 unit produk selama periode empat tahun. Tren penjualan menunjukkan dinamika yang fluktuatif, namun mencatatkan titik puncak (*peak*) tertinggi pada kuartal pertama tahun 2022. Fluktuasi ini mengindikasikan bahwa



transaksi ritel elektronik di Myanmar sangat sensitif terhadap kondisi musiman atau faktor makroekonomi yang terjadi pada periode tersebut.

Evaluasi terhadap portofolio produk mengungkapkan adanya disparitas yang signifikan antara volume penjualan dan nilai finansial yang dihasilkan. Berdasarkan komponen *Product Performance*, "Laptop" teridentifikasi sebagai kontributor pendapatan terbesar bagi perusahaan, yang divisualisasikan melalui lonjakan garis pendapatan yang dominan. Namun, secara kuantitas fisik, "Desk" merupakan produk yang paling banyak diserap oleh pasar dengan total 1.938 unit terjual. Temuan ini memberikan implikasi strategis bahwa meskipun produk bernilai tinggi seperti Laptop menjadi penopang utama arus kas dan profitabilitas, produk dengan harga lebih rendah seperti Desk memiliki tingkat perputaran inventori (*inventory turnover*) yang jauh lebih cepat, sehingga memerlukan strategi manajemen stok yang berbeda.

Dari perspektif perilaku konsumen, analisis matriks pada komponen *Customer Behaviour* menunjukkan bahwa transaksi bisnis sangat didominasi oleh segmen pelanggan *Wholesale* yang melakukan pembelian secara langsung melalui saluran fisik (*Store*). Dominasi ini menegaskan bahwa meskipun tren digital sedang berkembang, model bisnis B2B (*Business-to-Business*) melalui saluran *offline* tetap menjadi tulang punggung utama distribusi barang elektronik di pasar Myanmar. Hal ini mengisyaratkan bahwa investasi pada hubungan kemitraan grosir dan ketersediaan stok di toko fisik masih menjadi prioritas yang lebih krusial dibandingkan ekspansi saluran online murni.

Ditinjau dari aspek spasial, terdapat ketimpangan performa yang mencolok antarwilayah. Yangon dan Bago beroperasi sebagai pusat gravitasi ekonomi, masing-masing menyumbang pendapatan di atas K2 Miliar. Yangon, dengan pendapatan K2,15 Miliar, mencatatkan tingkat kepuasan pelanggan tertinggi (rata-rata rating 3,2). Permintaan di ibu kota ini didominasi oleh produk "Headphones" dan "Laptop", dengan karakteristik unik di mana segmen pelanggan VIP lebih aktif berbelanja melalui saluran *Online*. Di sisi lain, wilayah Bago (K2,06 Miliar) menunjukkan lonjakan penjualan yang sangat agresif pada Q1 2022, di mana produk "Monitor" menjadi yang paling laris. Penjualan masif di Bago ini secara eksklusif didorong oleh segmen *Wholesale* melalui toko fisik, membuktikan bahwa preferensi saluran penjualan sangat bergantung pada demografi dan karakteristik regional masing-masing.

Sementara itu, wilayah seperti Mandalay, Naypyidaw, dan area tanpa identitas wilayah (*Unknown*) menunjukkan performa kelas menengah dengan pendapatan berkisar antara K1,7 Miliar hingga K2 Miliar, namun dengan pola perilaku yang sangat spesifik. Mandalay sangat bergantung pada penjualan "Laptop" dengan tingkat diskon rata-rata tertinggi (6,23%), mengindikasikan bahwa pasar di wilayah ini sangat sensitif terhadap harga. Naypyidaw menunjukkan anomali tren di mana puncak penjualan baru terjadi di pertengahan hingga akhir 2022 dengan preferensi kuat pada "Keyboard" dan "Headphones".

Di ujung bawah spektrum, wilayah Taninthary dan Magway menunjukkan anomali berupa penetrasi pasar yang hampir nihil. Taninthary hanya membukukan K6,5 Juta dari

26 unit, sedangkan Magway mencatatkan angka paling ekstrem dengan hanya 11 unit "Monitor" senilai K6,8 Juta dan tingkat kepuasan terendah (2,0). Minimnya transaksi di kedua wilayah ini bukan sekadar fluktuasi bisnis, melainkan sebuah sinyal peringatan (*red flag*) bagi manajemen. Kondisi ini mengindikasikan bahwa ekspansi pasar di wilayah tersebut belum optimal, atau terdapat hambatan struktural dalam distribusi dan pencatatan data historis yang memerlukan investigasi operasional lebih lanjut.

KESIMPULAN

Penelitian ini telah berhasil mengimplementasikan arsitektur *Data Warehouse* dan *dashboard Business Intelligence* untuk menganalisis performa penjualan ritel elektronik di Myanmar. Melalui pendekatan *Star Schema* dan pipeline ETL yang terstruktur menggunakan Pentaho Data Integration, permasalahan kualitas data mentah (*dirty data*) seperti *missing values*, inkonsistensi format, dan *outliers* berhasil diatasi secara otomatis. Pembangunan *Virtual Datamart* berupa *Database View* pada PostgreSQL terbukti efektif menjembatani normalisasi basis data dengan kebutuhan denormalisasi, sehingga memungkinkan penyajian data yang *agile* dan *real-time* pada Tableau tanpa redundansi fisik.

Hasil visualisasi pada *dashboard "Myanmar Electronic Retail Overview (2020-2023)"* berhasil mengekstrak wawasan strategis yang komprehensif dari total pendapatan kotor sebesar K9,79 Miliar. Analisis mengungkap disparitas portofolio produk, di mana "Laptop" menjadi penyumbang pendapatan terbesar, sementara "Desk" memiliki perputaran inventori tercepat. Dari perspektif perilaku konsumen dan spasial, teridentifikasi bahwa segmen *Wholesale* melalui saluran fisik (*Store*) masih mendominasi pasar. Selain itu, terdapat ketimpangan performa antarwilayah yang signifikan; dengan Yangon dan Bago beroperasi sebagai pusat gravitasi ekonomi, sementara wilayah seperti Taninthary dan Magway menunjukkan anomali penetrasi pasar yang memerlukan investigasi operasional lebih lanjut.

Secara keseluruhan, integrasi antara Pentaho untuk rekayasa data dan Tableau untuk visualisasi tidak hanya berfungsi sebagai alat monitoring metrik historis, tetapi juga sebagai instrumen diagnostik yang kuat. *Dashboard* interaktif ini memberikan landasan strategis bagi pemangku kepentingan untuk merumuskan kebijakan manajemen stok yang adaptif, mengoptimalkan saluran distribusi B2B, serta melakukan ekspansi pasar yang lebih terarah berdasarkan karakteristik demografis dan preferensi konsumen di setiap wilayah.

DAFTAR PUSTAKA

- Akter, M., & Mamun, M. N. H. (2023). "Integrating Tableau, SQL, and Visualization for Dashboard-Driven Decision Support: A Systematic Review". *American Journal of Advanced Technology and Engineering Solutions*, 03(01), 01-30.
- Branchris. (2026). Perancangan Data Warehouse Penjualan E-Commerce untuk Analisis Tren Produk dan Brand Populer.



- Darmawan, R., & Swalaganata, G. (2025). "Analisa Komparatif Power BI dan Tableau dalam Implementasi Business Intelligence pada Brazilian E-Commerce Public Dataset by Olist". *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 8936-8944.
- Dharsika, I. G. (2025). ETL Implementation with Pentaho for Sales Data Visualization A Case Study of Lunabit Beauty Bar.
- Nayak, S. (2025). "The Role of Data Visualization Tools in Financial Decision-Making: A Comparative Analysis of Tableau, Power BI, and SSRS". *The ES Accounting and Finance*, 3(03), 282-301.
- Patel, D. B. (2023). "Comparative Analysis of Data Visualization Tools: Tableau, Power BI & Looker". *The Eastasouth Journal of Information System and Computer Science*, 01(01), 55-59.
- Putri, N. A., & Karnadi. (2025). "Implementasi Business Intelligence Untuk Analisis Data Kinerja Penjualan Menggunakan Power BI". *Bulletin of Computer Science Research*, 6(1), 20-28.
- Senduk, F. K., Waluyo, R., & Isnaini, K. N. (2025). "Analisis Data dengan Business Intelligence dan Tableau untuk Visualisasi Garis Kemiskinan Indonesia". *Sistemasi: Jurnal Sistem Informasi*, 14(3), 1122-1141.
- somesomeall. (2026). Electronic retails dataset (Versi 1) [Dataset]. Kaggle. <https://www.kaggle.com/datasets/somesomeall/electronic-retails-dataset>