

BAB I

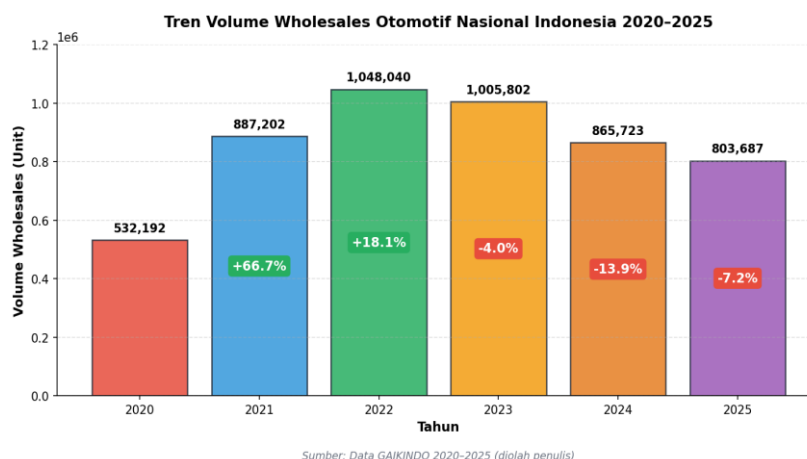
PENDAHULUAN

1.1 Latar Belakang

Salah satu sektor industri yang memiliki kontribusi signifikan terhadap perekonomian nasional Indonesia adalah industri otomotif. Industri ini menghasilkan data penjualan kendaraan dalam jumlah besar yang terus bertambah setiap tahun dan memiliki nilai strategis untuk mendukung analisis pasar serta pengambilan keputusan berbasis data [1], [2]. Sebagai asosiasi resmi industri otomotif Indonesia, GAIKINDO secara berkala mempublikasikan data penjualan kendaraan nasional yang mencakup kategori *wholesales*, *retail sales*, produksi, ekspor, dan impor kendaraan [3], [5]. Dinamika volume penjualan otomotif nasional selama periode 2020 hingga 2025 menunjukkan bahwa data historis kendaraan terus bertambah setiap tahun dan memiliki potensi besar untuk dianalisis lebih lanjut menggunakan pendekatan *Business Intelligence* dan *Big Data Analytics*.

Permasalahan konkret yang ditemukan pada data GAIKINDO meliputi: (1) perbedaan format kolom antarperiode, di mana sheet *wholesales* tahun 2020 memiliki struktur kolom yang berbeda dengan sheet *wholesales* tahun 2024 meskipun memuat jenis data yang sama; (2) inkonsistensi penamaan merek, seperti penulisan "TOYOTA" (huruf kapital semua) pada satu sheet dan "Toyota" (title case) pada sheet lainnya, serta variasi penulisan "MITSUBISHI MOTORS" dan "MITSUBISHI FUSO" yang perlu dibedakan; (3) tidak adanya kolom turunan analitik seperti pertumbuhan *year-over-year* (YoY), *market share*, dan pengelompokan segmen merek, sehingga pengguna harus menghitungnya secara manual; serta (4) data tersebar dalam lima kategori yang tidak terhubung

(*wholesales, retail sales, produksi, ekspor, impor*), sehingga analisis lintas-kategori memerlukan proses integrasi data yang kompleks.



Gambar 1.1 Tren Volume Wholesales Otomotif Nasional 2020 hingga 2025

Sumber: diolah dari data GAIKINDO (2025) [1], [5]

Gambar 1.1 menampilkan *Tren Volume Wholesales Otomotif Nasional* 2020 hingga 2025, sebagai bagian dari hasil analisis data pada penelitian ini.

Meskipun data tersebut tersedia secara publik melalui [GAIKINDO](#), data masih disajikan dalam bentuk mentah berupa *spreadsheet multi-sheet* dengan struktur yang tidak terintegrasi antarperiode maupun antarkategori. Setiap file memiliki format kolom, penamaan merek, dan struktur tabel yang berbeda sehingga menciptakan kompleksitas data (*data complexity*) dan keberagaman struktur data (*data heterogeneity*) [2], [4] Kondisi tersebut menyebabkan proses eksplorasi dan analisis data menjadi sulit dilakukan secara langsung, terutama bagi pengguna yang tidak memiliki kemampuan teknis pengolahan data.

Data GAIKINDO juga tersebar dalam banyak sheet dan kategori yang berbeda sehingga proses integrasi data historis menjadi kurang efisien apabila dilakukan secara manual menggunakan *spreadsheet* konvensional. Keberagaman struktur ini menuntut proses penyatuan data yang konsisten sebelum analisis dapat

dilakukan. Tanpa proses integrasi yang sistematis, informasi penting seperti tren lintas tahun dan perbandingan antar kategori sulit diperoleh secara cepat dan akurat.

Permasalahan tersebut menimbulkan kesenjangan antara ketersediaan data publik dan kemampuan pemanfaatan data oleh pengguna akhir (*data usability gap*). Pengguna yang ingin memperoleh informasi seperti tren penjualan kendaraan, *market share* merek, maupun pertumbuhan penjualan tahunan masih harus melakukan proses pengolahan data secara manual menggunakan *spreadsheet* konvensional [1], [5]. Selain membutuhkan waktu yang relatif lama, proses manual juga memiliki keterbatasan dalam menangani data historis *multi-periode* dan analisis *multidimensi* secara efisien [4], [6].

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan *Big Data Analytics* dan *Business Intelligence* (BI) yang mampu melakukan proses integrasi, *Transformasi*, dan analisis data secara lebih sistematis [2], [6]. Dalam penelitian ini, proses pengolahan data dilakukan menggunakan *framework CRISP-DM* yang terdiri dari tahap *selection*, *preprocessing*, *Transformation*, *Data Mining*, dan *interpretation/Evaluation* [7]. Pengolahan data dilakukan menggunakan Jupyter Notebook dengan bahasa pemrograman *Python* serta pustaka *Pandas* dan *NumPy* untuk mendukung integrasi data *multi-sheet* dan analisis *multidimensi* berbasis *Online Analytical Processing* (OLAP).

Sistem yang dikembangkan diharapkan meringankan pekerjaan dua pihak sekaligus. Bagi instansi (GAIKINDO dan pemangku kepentingan industri), sistem mengotomasi proses integrasi dan penyajian data yang selama ini dilakukan manual dan berulang setiap bulan, sehingga menghemat waktu serta mengurangi risiko kesalahan. Bagi publik (masyarakat umum, mahasiswa, dan pelaku usaha), sistem menyediakan informasi penjualan otomotif yang terstruktur dan mudah dipahami tanpa memerlukan keterampilan teknis pengolahan data, sehingga data publik yang tersedia dapat benar-benar dimanfaatkan.

Beberapa penelitian terdahulu telah membahas penerapan *Business Intelligence* pada data penjualan, namun sebagian besar masih berfokus pada lingkup internal perusahaan dan belum membahas integrasi data publik *multi-periode* secara komprehensif [5], [6]. Urgensi penelitian ini terletak pada adanya kesenjangan pemanfaatan data (*data usability gap*) antara ketersediaan data publik GAIKINDO dan kemampuan masyarakat untuk memanfaatkannya. Data penjualan otomotif nasional sesungguhnya bersifat terbuka, namun nilainya belum dapat dirasakan secara luas karena format mentahnya hanya dapat diolah oleh pengguna dengan keterampilan teknis tertentu. Padahal, informasi seperti tren penjualan, pangsa pasar merek, dan percepatan transformasi menuju kendaraan listrik sangat dibutuhkan oleh akademisi, pelaku usaha pendukung otomotif, maupun masyarakat umum untuk pengambilan keputusan dan kajian. Tanpa sistem yang mampu mengintegrasikan dan menyajikan data tersebut secara otomatis, proses analisis akan terus dilakukan manual, memakan waktu, rawan kesalahan interpretasi, dan tidak dapat direproduksi. Oleh karena itu, dibutuhkan suatu sistem *Business Intelligence* yang mengubah data mentah *multi-periode* menjadi informasi terstruktur, akurat, dan mudah diakses publik secara berkelanjutan. Kondisi inilah yang menjadikan penelitian ini relevan dan mendesak untuk dilakukan. Oleh karena itu, penelitian ini berfokus pada perancangan dan implementasi sistem *Business Intelligence* untuk mengolah data historis penjualan otomotif nasional GAIKINDO tahun 2020 hingga 2025 menjadi *dataset* analitik yang lebih terstruktur, mudah dipahami, dan dapat dimanfaatkan publik secara lebih efektif.

1.2 Rumusan Masalah

Berdasarkan uraian tersebut, permasalahan utama penelitian ini dapat ditegaskan secara lebih spesifik sebagai berikut. Data penjualan otomotif GAIKINDO memang tersedia secara publik, namun tersebar dalam 38 berkas dan lebih dari 100 sheet yang tidak saling terhubung, dengan tiga karakteristik masalah yang saling memperkuat: (1) heterogenitas struktur, yaitu format kolom dan

susunan tabel yang berbeda antarperiode maupun antarkategori sehingga tidak dapat langsung digabungkan; (2) inkonsistensi penamaan merek, yang menyebabkan satu merek tercatat dalam beberapa bentuk penulisan dan menghasilkan agregasi yang keliru apabila diproses apa adanya; serta (3) ketiadaan indikator turunan analitik seperti pertumbuhan year-over-year, market share, dan segmentasi merek, sehingga setiap pengguna harus menghitungnya sendiri secara manual. Akibatnya, untuk memperoleh satu informasi sederhana sekalipun, misalnya tren penjualan lintas tahun atau pangsa pasar suatu merek, pengguna harus melakukan integrasi dan pembersihan data secara berulang yang memakan waktu, rawan kesalahan, dan tidak dapat direproduksi. Inti permasalahannya bukan pada ketersediaan data, melainkan pada rendahnya keterpakaian data (data usability gap): jarak antara data publik yang sudah tersedia dan bentuk pengetahuan yang benar-benar dapat digunakan untuk pemantauan dan pengambilan keputusan. Permasalahan inilah yang menjadi dasar perumusan masalah penelitian.

Berdasarkan latar belakang di atas, masalah utama penelitian ini adalah data penjualan otomotif nasional GAIKINDO yang tersebar dalam banyak berkas dan kategori dengan struktur tidak konsisten, sehingga sulit diintegrasikan dan dianalisis oleh publik. Rumusan masalah penelitian dirumuskan sebagai berikut:

1. Bagaimana cara mengintegrasikan data historis penjualan otomotif nasional GAIKINDO tahun 2020 hingga 2025 yang tersebar dalam banyak berkas dan kategori menjadi satu *dataset* terstruktur yang konsisten dan siap dianalisis menggunakan kerangka kerja *CRISP-DM* (CRISP-DM)?
2. Bagaimana merancang sistem *Business Intelligence* yang mampu mentransformasi data mentah penjualan otomotif menjadi informasi yang mudah dipahami publik melalui *dashboard* analitik?
3. Bagaimana hasil implementasi *dashboard* publik dapat menyajikan informasi tren penjualan, perbandingan merek, dan distribusi pasar otomotif Indonesia kepada publik?

4. Bagaimana dinamika transformasi pasar otomotif nasional menuju elektrifikasi (*Electric Vehicle/EV*) selama 2020 hingga 2025 dapat diidentifikasi dan dianalisis secara deskriptif berdasarkan data GAIKINDO?

1.3 Batasan Masalah

Agar penelitian ini lebih terfokus dan dapat diselesaikan secara optimal, ditetapkan batasan-batasan sebagai berikut:

1. Data yang digunakan merupakan data historis resmi dari GAIKINDO untuk periode Januari 2020 hingga Desember 2025, mencakup kategori *wholesales*, *retail sales*, produksi, ekspor *CBU*, dan impor *CBU*.
2. Kerangka kerja yang digunakan adalah kerangka hibrida yang memadukan SDLC-Waterfall, CRISP-DM, serta analitik OLAP (deskriptif).
3. *Deployment Dashboard* dilakukan menggunakan platform *GitHub Pages* sebagai web statis yang dapat diakses publik secara gratis.
4. Analitik yang dilakukan bersifat deskriptif (*descriptive analytics*), yaitu menyajikan informasi tren, perbandingan, dan distribusi melalui agregasi data dan visualisasi berbasis OLAP. Analitik prediktif dan preskriptif berada di luar ruang lingkup penelitian.
5. Analisis transformasi pasar kendaraan listrik (*EV*) bersifat deskriptif berdasarkan data historis GAIKINDO dan diposisikan sebagai temuan turunan (*secondary finding*) dari analisis tren penjualan otomotif nasional, bukan fokus utama yang terpisah.
6. Faktor kebijakan pemerintah yang relevan (misalnya insentif PPnBM dan insentif kendaraan listrik) digunakan hanya sebagai konteks deskriptif dalam interpretasi tren penjualan, bukan sebagai variabel yang dimodelkan secara kuantitatif.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan penelitian disusun berkorespondensi satu-satu dengan keempat rumusan masalah revisi:

1. Mengintegrasikan dan mengolah data historis penjualan otomotif nasional GAIKINDO tahun 2020 hingga 2025 menjadi satu *dataset* terstruktur yang konsisten menggunakan kerangka kerja CRISP-DM.
2. Merancang dan mengimplementasikan sistem *Business Intelligence* yang mentransformasi data mentah menjadi informasi yang mudah dipahami publik melalui *dashboard* analitik.
3. Menyajikan informasi tren penjualan, perbandingan antar merek, dan distribusi pasar otomotif Indonesia melalui *dashboard* publik berbasis web.
4. Mengidentifikasi dan menganalisis secara deskriptif dinamika transformasi pasar otomotif nasional menuju elektrifikasi (*EV*) berdasarkan data GAIKINDO 2020 hingga 2025.

1.4.2 Manfaat Penelitian

Manfaat penelitian ini dibedakan menjadi dua, yaitu manfaat teoritis dan manfaat praktis.

1. Manfaat Teoritis

Penelitian ini memberikan kontribusi pada pengembangan kajian *Business Intelligence* dan analitik data di Indonesia, khususnya pada penerapan arsitektur BI empat lapisan (*data source*, *ETL*, *Data Warehouse*, dan *Dashboard*) untuk pengolahan data publik berskala besar. Selain itu, penelitian ini memperkaya literatur akademik mengenai integrasi data *multi-sumber* menggunakan pendekatan *ETL* dan *dimensional Modeling*.

2. Manfaat Praktis

Bagi pemangku kepentingan industri otomotif, *Dashboard* yang dihasilkan menyediakan akses informasi tren penjualan, distribusi *market share*, dan dinamika *Transformasi* pasar EV yang ringkas dan mudah dipahami. Bagi akademisi dan peneliti, *dataset* terstruktur yang tersimpan dalam *Data Warehouse* dapat dijadikan dasar penelitian lanjutan di bidang analitik data otomotif. Bagi pelaku usaha kecil-menengah di sektor pendukung otomotif, informasi tren penjualan dapat dimanfaatkan untuk perencanaan bisnis yang lebih berbasis data. Bagi peneliti selanjutnya, menyediakan *dataset* granular GAIKINDO yang telah diproses dan didokumentasikan sebagai basis penelitian lanjutan.

1.5 Sistematika Penulisan

Untuk memberikan gambaran yang sistematis dan terstruktur mengenai keseluruhan isi penelitian, penulisan skripsi ini disusun dalam lima bab dengan rincian sebagai berikut:

1. **Bab I Pendahuluan** membahas latar belakang penelitian, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.
2. **Bab II Landasan Teori** menguraikan penelitian terdahulu yang relevan, teori-teori pendukung yang berkaitan dengan *Business Intelligence*, *Big Data Analytics*, *Data Warehouse*, *Dashboard* analitik, *OLAP*, dan industri otomotif Indonesia, serta *framework CRISP-DM* dan *tools/perangkat lunak* yang digunakan dalam penelitian.
3. **Bab III Metodologi Penelitian** membahas gambaran umum objek penelitian, metode penelitian yang mencakup alur penelitian dan tahapan *CRISP-DM*, teknik pengumpulan data, serta teknik analisis data yang diterapkan.
4. **Bab IV Analisis dan Hasil Penelitian** memaparkan analisa masalah, perancangan sistem, pengujian dan implementasi, serta hasil dan diskusi

yang mencakup penerapan keenam fase *CRISP-DM* dan analisis *Dashboard* yang dihasilkan.

5. **Bab V Simpulan dan Saran** menyajikan simpulan yang menjawab seluruh rumusan masalah dan saran untuk pengembangan penelitian selanjutnya.

