

Eksplorasi dan Visualisasi Data Transaksi Online Retail Untuk Mendukung Pengambilan Keputusan Bisnis

Dzaki Al Hafiz¹, Muhammad Haris Hermanto^{1*}, Anton Setiawan¹, Muhammad Said Hasibuan¹

¹ Institut Informatika dan Bisnis Darmajaya, Bandar Lampung
dzaki.2421211014@mail.darmajaya.ac.id, muhammad.2321210010@mail.darmajaya.ac.id,
anton.2421211001@mail.darmajaya.ac.id, msaid@darmajaya.ac.id

Keywords:

*exploratory data analysis,
data visualization,
interactive dashboard,
retail transactions*

ABSTRACT

The growth of e-commerce generates large and complex transaction data volumes requiring in-depth analysis for business decision-making support. This research applies Exploratory Data Analysis (EDA) approach to the Online Retail UCI dataset (541,909 transactions, 2010-2011) using Python on Google Colab for data cleaning, yielding 392,692 valid records. Analysis focuses on time trends of transactions and revenue, top 7 products by quantity and revenue, and sales distribution across regions through interactive Looker Studio dashboard. Findings reveal seasonal transaction patterns, differences between high-volume vs high-revenue products, and United Kingdom dominance (80% revenue). The interactive dashboard with time and country filters enables flexible data exploration for promotion strategy, inventory management, and market expansion decisions. This research proves data visualization effectiveness as a retail online business decision support system.

Kata Kunci

*exploratory data analysis,
visualisasi data,
dashboard interaktif,
transaksi retail*

ABSTRAK

Perkembangan e-commerce menghasilkan volume data transaksi yang besar dan kompleks, memerlukan analisis mendalam untuk mendukung pengambilan keputusan bisnis. Penelitian ini menerapkan pendekatan Exploratory Data Analysis (EDA) terhadap dataset Online Retail UCI (541.909 transaksi, 2010-2011) menggunakan Python di Google Colab untuk pembersihan data, dihasilkan 392.692 data valid. Analisis difokuskan pada tren waktu transaksi dan pendapatan, top 7 produk berdasarkan quantity dan revenue, serta distribusi wilayah penjualan melalui dashboard interaktif Looker Studio. Hasil menunjukkan pola musiman transaksi, perbedaan produk high-volume vs high-revenue, dan dominasi United Kingdom (80% revenue). Dashboard dengan filter interaktif mendukung eksplorasi data fleksibel untuk strategi promosi, pengelolaan stok, dan ekspansi pasar. Penelitian ini membuktikan efektivitas visualisasi data sebagai sistem pendukung keputusan bisnis retail online.

Korespondensi Penulis:

Muhammad Haris Hermanto,
Institut Informatika Dan Bisnis Darmajaya,
Bandar Lampung, Indonesia
Email: muhammad.2321210010@mail.darmajaya.ac.id

1. PENDAHULUAN

Perkembangan teknologi informasi dan sistem perdagangan digital telah mendorong peningkatan aktivitas transaksi pada sektor *online retail*. (Dandasi, 2025) Data transaksi yang dihasilkan dari aktivitas tersebut terus bertambah dalam jumlah dan kompleksitas, sehingga memerlukan pengelolaan dan analisis yang tepat agar dapat memberikan informasi yang bernilai bagi pengambilan keputusan bisnis. (Wang et al., 2024) Tanpa proses analisis yang memadai, data transaksi hanya menjadi arsip penyimpanan yang belum dimanfaatkan secara optimal.

Salah satu pendekatan yang dapat digunakan untuk memahami karakteristik data transaksi adalah *Exploratory Data Analysis* (EDA). (Maresti et al., n.d.) Pendekatan ini bertujuan untuk mengeksplorasi data secara deskriptif guna mengidentifikasi pola, tren, dan anomali melalui teknik visualisasi data. Visualisasi data memiliki peran penting dalam menyajikan informasi secara ringkas dan intuitif, sehingga memudahkan pemangku kepentingan dalam memahami kondisi bisnis dan mendukung proses pengambilan keputusan berbasis data. (Turkmen, 2022)

Dalam konteks bisnis online retail, visualisasi data transaksi dapat dimanfaatkan untuk menganalisis tren penjualan dari waktu ke waktu, mengidentifikasi produk dengan kinerja penjualan terbaik, serta memahami distribusi penjualan berdasarkan wilayah. (Tipnis, 2026) Informasi tersebut sangat berguna bagi pihak manajemen dalam merumuskan strategi bisnis, seperti perencanaan promosi, pengelolaan persediaan, dan penentuan fokus pasar. Oleh karena itu, pemanfaatan visualisasi data tidak hanya berfungsi sebagai alat analisis, tetapi juga sebagai sistem pendukung keputusan bisnis. (Dandasi, 2025)

Penelitian sebelumnya telah memanfaatkan dataset *Online Retail* dari *UCI Machine Learning Repository* untuk berbagai analisis data transaksi retail. (Martua et al., n.d.) menyatakan bahwa "penelitian ini bertujuan untuk mengevaluasi potensi penggunaan algoritma *FP-Growth* dalam *Market Basket Analysis* dengan memanfaatkan dataset ritel daring dari *UCI*, guna menghasilkan wawasan strategis dalam penataan produk dan penerapan strategi pemasaran seperti cross-selling dan bundling".

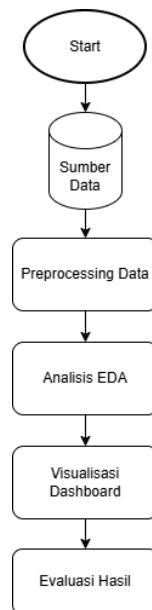
Selain itu, (Awaliyah et al., 2024) menjelaskan bahwa penggunaan dataset pelanggan ritel online, di mana algoritma *K-means* digunakan sebagai metode utama untuk segmentasi pelanggan. Hasil analisis menunjukkan bahwa *quantity* dan *unit price* memiliki pengaruh signifikan terhadap perilaku pelanggan ritel online". Penelitian (Abdur Rasyid Asruddin, n.d.) *Komparasi Model Klasifikasi Data Online Retail dengan Seleksi Fitur*, untuk komparasi model klasifikasi pada data transaksi belanja online," yang mendukung pendekatan EDA untuk keputusan bisnis.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk melakukan eksplorasi dan visualisasi data transaksi online retail menggunakan pendekatan *Exploratory Data Analysis* (EDA). (Utomo & Lubna, n.d.) Visualisasi data disajikan dalam bentuk dashboard interaktif untuk memberikan gambaran yang komprehensif mengenai pola transaksi, karakteristik produk, dan distribusi wilayah penjualan. (Chakma et al., 2025) Hasil penelitian ini diharapkan dapat memberikan dukungan informasi yang relevan bagi pengambilan keputusan bisnis serta menjadi referensi dalam pemanfaatan visualisasi data pada sektor online retail.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode *Exploratory Data Analysis* (EDA). Pendekatan ini bertujuan untuk menggambarkan pola, tren, dan karakteristik data transaksi penjualan melalui visualisasi data interaktif tanpa melakukan pemodelan prediktif.

Dataset yang digunakan dalam penelitian ini merupakan dataset *Online Retail* diambil dari sumber <https://archive.ics.uci.edu/datasets/?search=Online+Retail>, yang berisi data transaksi penjualan online pada periode Desember 2010 hingga Desember 2011. Chen, D. (2015). Dataset awal terdiri dari 541.909 baris data transaksi dengan beberapa atribut utama, antara lain *InvoiceNo*, *StockCode*, *Description*, *Quantity*, *InvoiceDate*, *UnitPrice*, *CustomerID*, dan *Country*.



Gambar 1. Diagram alur penelitian

Analisis data dalam penelitian ini dilakukan menggunakan pendekatan *Exploratory Data Analysis* (EDA) dengan memanfaatkan teknik visualisasi data. Dataset yang telah melalui proses pembersihan selanjutnya digunakan untuk membangun sebuah dashboard visualisasi interaktif menggunakan Looker Studio. (Lutfi et al., 2023) Visualisasi data difokuskan pada beberapa aspek utama, yaitu analisis tren waktu untuk mengamati pola jumlah transaksi dan pendapatan dari waktu ke waktu, analisis produk dengan menampilkan tujuh produk teratas berdasarkan jumlah penjualan dan nilai pendapatan, serta analisis distribusi wilayah untuk mengidentifikasi negara-negara dengan kontribusi pendapatan tertinggi. Dashboard disusun dalam satu halaman dan dilengkapi dengan filter interaktif, seperti filter rentang waktu dan filter negara, guna mendukung proses eksplorasi data secara fleksibel dan memudahkan pengguna dalam memahami pola dan karakteristik data transaksi. (Sapitri, 2025)

Secara umum, alur penelitian diawali dengan pengumpulan dataset transaksi penjualan online yang digunakan sebagai sumber data penelitian. Selanjutnya, dataset tersebut melalui tahap pembersihan data menggunakan Google Colaboratory untuk memastikan kualitas dan validitas data yang dianalisis. (Arun Kumar Mishra et al., 2024) Dataset hasil pembersihan kemudian disimpan dan digunakan pada tahap berikutnya, yaitu analisis eksploratif dan visualisasi data. Proses analisis dilakukan dengan pendekatan *Exploratory Data Analysis* (EDA) melalui pembangunan dashboard visualisasi menggunakan Looker Studio. Tahap akhir penelitian adalah interpretasi hasil visualisasi data yang diperoleh untuk mendukung penyusunan pembahasan dan penarikan kesimpulan penelitian. (Widi Hastomo et al., 2022).

3. HASIL DAN ANALISIS

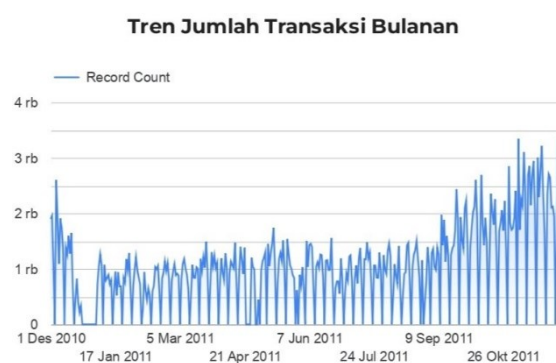
3.1. Analisis Tren Transaksi dan Pendapatan

Tahap pertama sebelum dilakukan proses analisis dan visualisasi data, dataset terlebih dahulu melalui tahap pembersihan data (*data cleaning*) untuk meningkatkan kualitas dan validitas data yang digunakan. Proses pembersihan data dilakukan menggunakan bahasa pemrograman Python pada platform Google Colaboratory, dengan beberapa tahapan utama, yaitu menghapus data transaksi yang memiliki nilai *Quantity* dan *UnitPrice* tidak valid (≤ 0), menghapus data dengan nilai *CustomerID* yang kosong karena tidak dapat merepresentasikan transaksi pelanggan secara valid, serta menghapus data duplikat untuk menghindari penghitungan transaksi ganda. Pada tahap ini juga menambahkan fitur baru yaitu *revenue* didapatkan dari *Quantity* dikalikan dengan *UnitPrice*. Setelah seluruh proses pembersihan data dilakukan, jumlah data transaksi berkurang dari 541.909 data awal menjadi 392.692 data transaksi yang dinyatakan valid dan selanjutnya digunakan dalam tahap analisis dan visualisasi data. (Agus & Hasibuan, 2025) Tabel 1 menunjukkan struktur dataset yang telah dilakukan tahap *preprocessing data*.

Tabel 1. Daftar Kolom Dataset Baru

No	Nama Kolom	Tipe	Jumlah Data	Jumlah NULL
1	InvoiceNo	Object	392692	-
2	StockCode	Object	392692	-
3	Description	Object	392692	-
4	Quantity	Int64	392692	-
5	InvoiceDate	Datetime64[ns]	392692	-
6	UnitPrice	Float64	392692	-
7	CustomerID	Float64	392692	-
8	Country	Object	392692	-
9	Revenue	Float64	392692	-

Berdasarkan hasil visualisasi tren waktu, jumlah transaksi penjualan online menunjukkan pola fluktuatif sepanjang periode pengamatan. (Sasha Safira Nuraini et al., 2025) Pada beberapa bulan tertentu terlihat peningkatan jumlah transaksi yang cukup signifikan, yang mengindikasikan adanya pola musiman (*seasonality*) dalam aktivitas penjualan. (Najib & Stefany, 2024) Pola ini mencerminkan bahwa perilaku konsumen tidak bersifat konstan, melainkan dipengaruhi oleh waktu tertentu dalam satu tahun. Grafik transaksi bulanan dapat dilihat pada gambar 2.

**Gambar 2.** Tren jumlah transaksi bulanan

Selanjutnya, analisis tren pendapatan menunjukkan pola yang tidak selalu sejalan dengan jumlah transaksi. Meskipun pada beberapa periode terjadi peningkatan jumlah transaksi, peningkatan tersebut tidak selalu diikuti oleh kenaikan pendapatan yang sebanding. (Homepage et al., n.d.) Hal ini menunjukkan bahwa tingginya volume transaksi tidak secara otomatis menghasilkan pendapatan yang lebih besar, karena nilai transaksi juga dipengaruhi oleh harga satuan produk yang dijual. Grafik pendapatan bulanan dapat dilihat pada gambar 3.

**Gambar 3.** Tren pendapatan bulanan

3.2. Analisis Produk Terlaris

Analisis produk dilakukan dengan menampilkan tujuh produk teratas berdasarkan jumlah penjualan dan nilai pendapatan. Hasil visualisasi menunjukkan bahwa terdapat perbedaan antara produk dengan kuantitas penjualan tertinggi dan produk dengan kontribusi pendapatan tertinggi. Beberapa produk memiliki jumlah penjualan yang tinggi, namun menghasilkan pendapatan yang relatif lebih rendah,

sedangkan produk lain dengan jumlah penjualan lebih sedikit justru memberikan kontribusi pendapatan yang lebih besar.



Gambar 4. Produk dengan pendapatan tertinggi

Perbedaan tersebut mengindikasikan bahwa harga satuan produk berperan penting dalam menentukan kontribusi pendapatan. Dengan demikian, strategi penjualan tidak hanya perlu berfokus pada peningkatan volume penjualan, tetapi juga pada optimalisasi produk dengan nilai jual yang lebih tinggi.



Gambar 5. Produk terlaris

3.3. Analisis Distribusi Wilayah Penjualan

Analisis distribusi wilayah menunjukkan bahwa kontribusi pendapatan penjualan tidak tersebar secara merata di seluruh negara. Visualisasi tujuh negara dengan pendapatan tertinggi memperlihatkan adanya satu negara yang mendominasi pendapatan penjualan dibandingkan negara lainnya. Kondisi ini menunjukkan adanya konsentrasi pasar pada wilayah tertentu, yang kemungkinan dipengaruhi oleh jumlah pelanggan, tingkat permintaan, serta intensitas aktivitas penjualan di wilayah tersebut. (Sulistyo et al., 2025).



Gambar 6. Negara dengan pendapatan tertinggi

Meskipun demikian, negara-negara lain tetap menunjukkan kontribusi pendapatan yang signifikan dan memiliki potensi untuk dikembangkan lebih lanjut. Informasi ini dapat menjadi dasar bagi pengambilan keputusan strategis dalam menentukan fokus pasar dan perluasan wilayah penjualan.

3.4. Pembahasan Umum

Secara keseluruhan, hasil visualisasi data melalui dashboard interaktif mampu memberikan gambaran yang komprehensif mengenai pola penjualan, karakteristik produk, dan distribusi wilayah transaksi. Pemanfaatan visualisasi data dalam pendekatan *Exploratory Data Analysis* (EDA) mempermudah proses identifikasi pola dan anomali dalam data transaksi penjualan. (Velásquez-Henao, 2025) Selain itu, penggunaan filter interaktif memungkinkan eksplorasi data secara fleksibel berdasarkan rentang waktu dan negara tertentu, sehingga mendukung analisis deskriptif yang lebih mendalam dan informatif.

4. KESIMPULAN

Berdasarkan hasil eksplorasi dan visualisasi data transaksi online retail menggunakan pendekatan *Exploratory Data Analysis* (EDA), (Fakhrana & Voutama, 2025) penelitian ini menunjukkan bahwa visualisasi data mampu memberikan informasi yang relevan untuk mendukung pengambilan keputusan bisnis. Analisis tren transaksi dan pendapatan mengindikasikan adanya pola fluktuasi yang dapat dimanfaatkan sebagai dasar penentuan waktu yang tepat dalam perencanaan strategi penjualan dan promosi. Analisis produk menunjukkan perbedaan antara produk dengan jumlah penjualan tertinggi dan produk dengan kontribusi pendapatan tertinggi, sehingga dapat menjadi dasar dalam pengambilan keputusan terkait penentuan produk prioritas dan strategi harga. Selain itu, analisis distribusi wilayah memperlihatkan adanya konsentrasi pendapatan pada negara tertentu, yang dapat dimanfaatkan sebagai bahan pertimbangan dalam menentukan fokus pasar dan pengembangan wilayah penjualan. Dengan demikian, eksplorasi dan visualisasi data transaksi berperan sebagai alat pendukung keputusan bisnis yang berbasis pada pemahaman data secara komprehensif.

REFERENSI

- [1] M. Abdur Rasyid Asruddin, "Komparasi model klasifikasi data online retail dengan seleksi fitur," unpublished manuscript, n.d. [Online]. Available: <http://archive.ics.uci.edu/ml/datasets/Online+Retail>
- [2] I. Agus and M. Hasibuan, "Customer segmentation using an enhanced RFM-K-means framework on the online retail dataset," *International Journal of Informatics and Information Systems*, vol. 8, no. 4, pp. 221–233, 2025, doi: 10.47738/ijis.v8i4.289.
- [3] A. K. Mishra, M. Sinha, and S. K. Jha, "Data analytics for visualization of business insights for an online retail store using Python," *International Journal of Management & Entrepreneurship Research*, vol. 6, no. 10, pp. 3314–3320, 2024, doi: 10.51594/ijmer.v6i10.1636.
- [4] D. A. Awaliyah, B. Prasetyo, R. Muzayanah, and A. D. Lestari, "Optimizing customer segmentation in online retail transactions through the implementation of the K-means clustering algorithm," *Scientific Journal of Informatics*, vol. 11, no. 2, pp. 539–548, 2024, doi: 10.15294/sji.v11i2.6137.
- [5] S. Chakma, G. Chakma, and R. Chakma, "Statistical insights and exploratory data analysis for e-commerce sales data: A collaborative filtering and recency-based recommendation system," *International Journal of Business and Management*, vol. 20, no. 5, p. 258, 2025, doi: 10.5539/ijbm.v20n5p258.
- [6] V. V. Dandasi, "Data warehousing and business intelligence technologies in retail: A comparative analysis of predictive models for performance optimization," *Journal of Business Intelligence and Data Analytics*, vol. 2, no. 2, pp. 1–8, 2025, doi: 10.55124/jbid.v2i2.251.
- [7] I. F. Fakhrana and A. Voutama, "Analisis tren penjualan dan strategi harga di e-commerce menggunakan exploratory data analysis," *Jurnal Ilmu Komputer dan Bisnis*, vol. 16, no. 1, pp. 154–164, 2025, doi: 10.47927/jikb.v16i1.976.

- [8] J. Faris Fadil, M. Firmansyah, W. Asriyani, and M. F. Fadil, "Klasterisasi pola transaksi online retail menggunakan metode self-organizing map," *Journal of Innovative and Creativity*, vol. 5, no. 3, n.d.
- [9] A. Lutfi *et al.*, "Drivers and impact of big data analytic adoption in the retail industry: A quantitative investigation applying structural equation modeling," *Journal of Retailing and Consumer Services*, vol. 70, 2023, doi: 10.1016/j.jretconser.2022.103129.
- [10] D. Chen, "Online retail," *UCI Machine Learning Repository*, 2015, doi: 10.24432/C5BW33.
- [11] F. A. Maresti, G. M. Anugraheni, R. A. Hargiyanto, and K. Mustaqim, "Penerapan exploratory data analysis (EDA) dan analisis recency, frequency, and monetary (RFM) untuk segmentasi pelanggan e-commerce," unpublished manuscript, n.d. [Online]. Available: <http://ejurnal.ulbi.ac.id/index.php/competitive>
- [12] Y. Martua, L. Sianipar, R. A. M. Yulianto, and E. Rilvani, "Study literature review application of market basket analysis on UCI online retail dataset for business insights," *Kohesi: Jurnal Multidisiplin Saintek*, vol. 1, no. 2, n.d., doi: 10.8734/Kohesi.v1i2.365.
- [13] M. K. Najib and E. M. Stefany, "Visualisasi data penjualan supermarket dengan Microsoft Power BI untuk menghasilkan insight dan rekomendasi," *Jurnal Sistem Informasi (TEKNOFILE)*, vol. 2, pp. 921–928, 2024.
- [14] D. Sapitri, "Visualisasi data e-commerce menggunakan Power BI dalam meningkatkan strategi pemasaran," *Jurnal Sistem Informasi (TEKNOFILE)*, vol. 3, pp. 749–759, 2025.
- [15] S. S. Nuraini, A. M. Romadhoni, M. A. F. Taqiyuddin, and M. Wicaksono, "Pemanfaatan visualisasi data dalam business intelligence untuk strategi bisnis perusahaan retail," *Jurnal Manuhara: Pusat Penelitian Ilmu Manajemen dan Bisnis*, vol. 3, no. 3, pp. 234–243, 2025, doi: 10.61132/manuhara.v3i3.1992.
- [16] G. B. Sulistyio, N. Hasan, S. Kiswati, F. Natalia, and E. Muningsih, "Segmentasi pelanggan dan optimalisasi penjualan pada data retail online berbasis model RFM," *Computer and Network Technology*, vol. 5, no. 1, pp. 16–22, 2025. [Online]. Available: <http://jurnal.bsi.ac.id/index.php/conten16>
- [17] A. Tipnis, "Framework for data analysis of retail market using machine learning algorithms," *International Journal for Research in Applied Science and Engineering Technology*, vol. 14, no. 1, pp. 703–713, 2026, doi: 10.22214/ijraset.2026.76800.
- [18] B. Turkmen, "Customer segmentation with machine learning for online retail industry," *The European Journal of Social and Behavioural Sciences*, vol. 31, no. 2, pp. 111–136, 2022, doi: 10.15405/ejsbs.316.
- [19] F. S. Utomo and Z. Lubna, "Visualisasi data penjualan untuk menentukan business insight menggunakan Metabase pada perusahaan retail global," *Sistemasi: Jurnal Sistem Informasi*, vol. 13, no. 4, n.d. [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [20] J. D. Velásquez-Henao, "Emerging trends in retail analytics: A bibliometric analysis of the last decade," *DYNA (Colombia)*, vol. 92, no. 237, pp. 16–29, 2025, doi: 10.15446/dyna.v92n237.116880.
- [21] L. Wang, T. R. A. L. Pertheban, T. Li, and L. Zhao, "Application of business intelligence based on big data in e-commerce data evaluation," *Heliyon*, vol. 10, no. 21, 2024, doi: 10.1016/j.heliyon.2024.e38768.
- [22] W. Hastomo, A. S. B. Karno, Sudjiran, D. Arif, and E. S. Moreta, "Exploratory data analysis untuk data belanja pelanggan dan pendapatan bisnis," *Infotekmesin*, vol. 13, no. 2, pp. 314–321, 2022, doi: 10.35970/infotekmesin.v13i2.1547.