

Analisis Penjualan Pada Coffee Shop X Dengan Metode Exploratory Data Analysis Dan Metode Visualisasi Menggunakan Tableau

Sarah Devi Aritonang, Indra Kelana Jaya¹, Yolanda Rumapea²
Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, 9 September 2026
Revised, 25 September 2026
Accepted, 4 Oktober 2026

Keywords:

Kopi,
Tableau,
Exploratory Data
Analysis,
Python,
Analisis Data.

ABSTRAK

Penelitian ini membahas penerapan metode *Exploratory Data Analysis* dan visualisasi data dengan Tableau. Dengan persaingan industri kopi yang semakin ketat, pemahaman yang mendalam tentang pola penjualan dan perilaku konsumen menjadi sangat penting. EDA digunakan untuk mengidentifikasi dan mengeksplorasi karakteristik data, mengidentifikasi pola, tren, serta anomali yang mungkin mempengaruhi performa penjualan. Proses EDA ini mencakup pemeriksaan data penjualan berdasarkan produk, kategori, waktu, dan lokasi. Selanjutnya, Tableau digunakan sebagai alat visualisasi untuk mempermudah interpretasi hasil analisis dan penyajian informasi dalam bentuk grafik interaktif. Hasil penelitian menunjukkan pola penjualan yang bervariasi tergantung pada waktu dan kategori produk, serta adanya tren musiman yang signifikan pada penjualan kopi tertentu. Visualisasi yang dihasilkan memungkinkan pihak manajemen untuk mengidentifikasi produk-produk dengan performa tinggi serta momen-momen puncak penjualan, sehingga dapat mendukung pengambilan keputusan strategis untuk meningkatkan penjualan dan menyesuaikan strategi pemasaran. Analisis ini memberikan kontribusi dalam menyediakan wawasan berbasis data yang akurat untuk meningkatkan efisiensi dan efektivitas operasional di Coffee Shop X.

This is an open access article under the [CC BY-SA](#) license.



Penulis Koresponden:

Sarah Devi Aritonang,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: saraharitonang34@gmail.com

1. PENDAHULUAN

Exploratory Data Analysis (EDA) merupakan pendekatan yang dinilai fleksibel dan sangat cocok untuk mengidentifikasi serta menyelidiki fenomena yang muncul selama proses penelitian empiris. Menganalisis kinerja usaha coffee shop yang berguna untuk menunjang usaha tersebut merupakan suatu tantangan bagi pengusaha bisnis retail ini. Dalam analisis, seseorang mengidentifikasikan pola, hubungan, atau makna di balik informasi yang diberikan untuk memperoleh wawasan atau solusi. Berkembang pesatnya teknologi di era digital ini, para pemimpin bisnis dan pemasar tertarik dengan perubahan perilaku konsumen dan tuntutan pasar. Pada beberapa tahun belakangan *Industry kafe dan coffee shop* mengalami pertumbuhan yang amat signifikan[1]. Sehingga banyak konsumen yang berpindah dari satu *coffee shop* ke lainnya dikarenakan banyak referensi-referensi tempat baru. Selaku pemilik usaha harus mampu membuat baik rancangan atau keputusan-keputusan sebaik mungkin. Kegiatan ini menunjukkan upaya-upaya dalam menganalisis

apa yang sedang terjadi pada usaha *coffee shop*, seperti jumlah penjualan produk yang ada di *Coffee Shop X*, pendapatan *Coffee Shop X*, dan pertumbuhan pendapatan perbulannya. Metode ini pernah dalam jurnalnya yang berjudul “Implementasi *Exploratory Data Analysis* Untuk Analisis dan Visualisasi Data Penderita Stroke Kalimantan Selatan Menggunakan Platform *Tableau*”. jumlah warga yang mengidap penyakit stroke di Kalimantan Selatan tahun 2020 sampai dengan 2022[2]. Penggunaan metode *Exploratory Data Analysis* dan metode visualisasi data memudahkan saat menganalisis fenomena-fenomena yang terjadi. Pada penelitian ini ditemukan fakta yaitu, Dalam penelitian ini ditemukan bahwa terdapat tren perubahan total kasus baru maupun lama di berbagai wilayah Kalimantan Selatan selama periode 2020 hingga 2022. Di tahun 2021, terdapat penurunan signifikan di hampir seluruh daerah. Kota Banjarmasin mencatat total kasus stroke tertinggi tahun 2020, yakni sebanyak 5.249 kasus, namun jumlah ini menurun drastis menjadi 938 (2021) dan terus berkurang di tahun 2022. Meski demikian, beberapa wilayah seperti Kabupaten Hulu Sungai Selatan, Tabalong, dan Tanah Bumbu justru menunjukkan peningkatan jumlah kasus stroke pada tahun 2021. Meskipun demikian, secara keseluruhan, tahun 2022 mencatat penurunan kasus stroke di seluruh wilayah Kalimantan Selatan [3]. Dapat diambil kesimpulan bahwasany umumnya terdapat penurunan jumlah kasus baru serta kematian karena stroke di Kalimantan Selatan selama periode 2020 hingga 2022. Merujuk pada permasalahan yang telah dipaparkan sebelumnya, maka disusun sebuah penelitian dengan judul “Analisis Penjualan Pada *Coffee Shop X* dengan Metode *Exploratory Data Analysis* dan Metode Visualisasi Menggunakan *Tableau*”. Penelitian ini diharapkan berguna sebagai dasar agar pemilik gerai *coffee Shop X* dalam mengkaji kinerja gerai dan membuat keputusan atau kebijakan untuk memaksimalkan penjualan di *Coffee Shop X*.

2. METODE PENELITIAN

2.1 Exploratory Data Analysis

Exploratory Data Analysis (EDA) merupakan proses guna menelaah serta mengerti data serta menggali wawasan atau karakteristik penting yang terkandung di dalamnya. Secara umum, EDA dibagi menjadi 2 pendekatan, meliputi analisis berbasis visual (grafis) serta analisis tanpa visualisasi (non-grafis) [4]. EDA memiliki peran krusial sebagai langkah awal dalam memahami rumusan masalah serta mengeksplorasi hubungan antar fitur dalam data sebelum menerapkan model. Tahapan ini penting dilakukan sebelum memulai proses machine learning atau pemodelan statistik, agar model yang dikembangkan dapat lebih tepat sasaran dan relevan dengan permasalahan yang dianalisis serta menyediakan konteks yang diperlukan untuk merepresentasikan hasil dengan benar, dan dapat diterapkan pada konteks bisnis.

2.1.1 Pengidentifikasian Tujuan Analisis

Identifikasi tujuan analisis merupakan Langkah awal yang sangat penting dalam proses eksplorasi data. Tahap ini berfokus pada memahami masalah kebutuhan utama yang ingin dijawab melalui data, sekaligus menentukan arah eksplorasi yang spesifik dan relevan, proses ini melibatkan beberapa aspek, yaitu:

1. Memahami konteks bisnis atau penelitian. Penting untuk memahami tujuan organisasi. Konteks ini mencakup seperti Masalah utama yang hadapi, informasi yang dibutuhkan, dan dampak potensial dari analisis terhadap pengambilan keputusan
2. Merumuskan pertanyaan utama. Setelah memahami konteks bisnis, Langkah berikutnya adalah merumuskan pertanyaan spesifik dan dapat dijawab menggunakan data
3. Menentukan arah eksplorasi data. Setelah pertanyaan dirumuskan, Langkah selanjutnya adalah menentukan variabel-variabel yang relevan serta metode yang akan digunakan untuk menjawabnya

2.1.2 Data Cleaning

Data cleaning ialah tahapan penting dalam pengolahan data yang bertujuan untuk menghilangkan data yang bersifat noise, tidak konsisten, maupun tidak relevan, agar hasil analisis menjadi lebih akurat dan dapat diandalkan [5]. Proses ini dilakukan agar meningkatkan kualitas data sehingga hasil menjadi lebih akurat. Adapun tahapan-tahapan dalam *data cleaning* sebagai berikut:

1. Menangani *Missing Values*. Proses ini bertujuan untuk mencari data yang kosong atau hilang dalam dataset. Kemudian data yang kosong tersebut diisi dengan hasil dari metode tertentu seperti *mean*, *median*, *modus* atau estimasi lain.
2. Menghapus duplikasi data. Dalam proses ini penulis akan memeriksa baris data yang berulang atau duplikat. Ini bertujuan agar hasil analisis nantinya tidak mengalami kesalahan.
3. Koreksi kesalahan data. Proses ini bertujuan untuk menemukan kesalahan dalam format atau nilai data.
4. Menstandarkan format data. Memastikan format data sesuai dengan nilai dari data tersebut. Contohnya seperti tanggal harus diisi dengan format DD-MM-YYYY

2.1.3 Data Exploration

Teknik *Preprocessing*, seperti pemilihan fitur dan sampel serta proses diskretisasi, memegang peranan penting dalam menyederhanakan dataset berukuran besar. Penyederhanaan ini memungkinkan proses pembelajaran jadi cepat juga akurat, sekaligus menghasilkan struktur data yang mudah untuk dianalisis dan dipahami [6]. Tahap ini, data dianalisis secara deskriptif untuk mengidentifikasi pola, distribusi, dan potensi masalah dalam dataset. Tahapan ini penting untuk memberikan gambaran tentang kualitas dan struktur data. Adapun Langkah-langkahnya sebagai berikut:

1. Deskripsi Dataset.
Langkah ini dilakukan untuk memahami struktur dataset. Seperti jumlah baris, jumlah kolom, nama kolom, dan tipe data
2. Statistik deskriptif
Analisis deskriptif ialah jenis analisis data berfungsi guna menguji generalisasi hasil penelitian berdasarkan pada satu sampel. Tahapan ini sangat penting karena tahapan ini dapat memberikan gambaran umum data, mendeteksi masalah dalam data, dan dasar untuk analisis lanjutan. Adapun komponen-komponen statistik deskriptif, yaitu:
 1. Mean (Rata-rata):
Rata-rata aritmatika yang dihitung dengan menjumlah semua nilai dalam dataset, terakhir yaitu membaginya dengan total data.

$$Mean = \frac{\sum_{i=1}^n x_i}{n}$$

2. Median (Nilai tengah):
Nilai tengah dari dataset yang sudah diurutkan. Jika jumlah data ganjil, median adalah nilai tengah. Jika jumlah data genap, median adalah rata-rata dua nilai tengah.
3. Modus:
Nilai yang paling sering muncul dalam dataset.
4. Standar Deviasi
Mengukur seberapa jauh data tersebar dari rata-rata.

$$SD = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}}$$

2.1.4 Laporan dan Dokumentasi

Laporan merupakan isi dari hasil proses analisis data yang disajikan dalam bentuk tulisan, visualisasi, atau representasi untuk menyampaikan temuan, *insight*, dan rekomendasi kepada pihak yang berkepentingan. Adapun tujuan laporan sebagai berikut:

1. Menyampaikan hasil analisis secara jelas dan terstruktur.
2. Memudahkan pembaca untuk memahami *insight* yang diperoleh.
3. Mendukung pengambilan keputusan berdasarkan data

Dokumentasi adalah catatan sistematis yang mencakup proses, metode, dan Langkah-langkah yang dilakukan selama analisis data, termasuk keputusan teknis, asumsi, serta referensi. Adapun tujuan dari dokumentasi sebagai berikut:

1. Menyediakan catatan untuk proses replikasi atau audit di masa mendatang
2. Meningkatkan transparansi dalam proses analisis
3. Memastikan proses analisis dapat dipahami oleh tim lain di masa depan[7].

2.1.5 Visualisasi Data

Visualisasi merupakan proses merancang gambar, diagram, atau animasi guna menyajikan informasi secara visual. Dalam pengertian lain, visualisasi ialah proses merubah data menjadi bentuk visual ataupun tabel agar karakteristik data serta hubungan antar elemen atau atribut bisa dianalisis dan disampaikan dengan lebih jelas. Visualisasi data juga dikenal sebagai salah satu teknik paling efektif dan menarik dalam eksplorasi data[8].

Visualisasi data berfungsi untuk menyajikan gambaran umum dari data dalam jumlah besar, sehingga dapat mempermudah para peneliti atau ilmuwan dalam melakukan analisis dan pengambilan kesimpulan [9]. Visualisasi data yang efektif harus disusun agar mudah dimengerti oleh ilmuwan, sehingga mereka dapat menyampaikan hasil visualisasi tersebut dengan jelas. Visualisasi data merupakan bagian yang tidak terpisahkan dari proses ilmiah[10].

Adapun tujuan-tujuan visualisasi data yaitu sebagai berikut :

1. Mengubah data kompleks menjadi informasi yang mudah dipahami.
2. Menyajikan pola dan tren tersembunyi dalam data secara visual.
3. Mempermudah pengambilan keputusan berbasis data melalui representasi visual.
4. Mengkomunikasikan *insight* dari data secara efektif dan intuitif.
5. Menyederhanakan analisis data untuk meningkatkan pemahaman dan aksi.

Visualisasi data hadir dalam berbagai bentuk untuk memudahkan memahami pola dan tren data. Berikut macam-macam diagram yang digunakan pada penelitian ini untuk membantu visualisasi data:

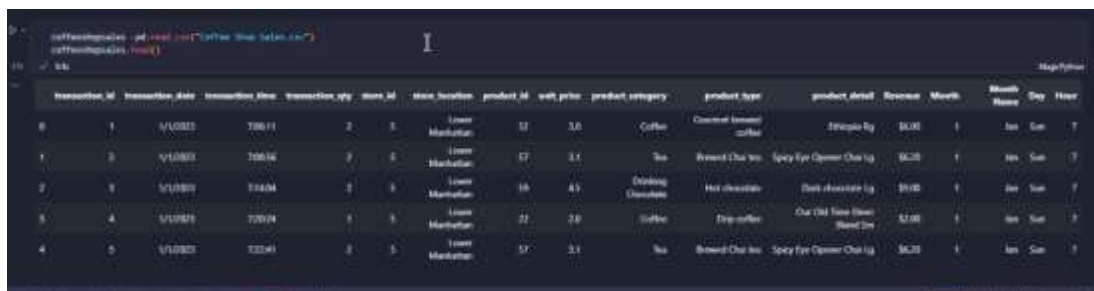
1. Diagram batang: memakai batang horizontal atau vertikal untuk mewakili data kategori. Panjang batang menunjukkan nilai atau frekuensi data
2. Diagram garis: memakai garis untuk menghubungkan titik-titik data yang mewakili nilai dari waktu ke waktu, menunjukkan tren atau perubahan nilai[11].

3. HASIL DAN PEMBAHASAN

3.1 Implementasi dan Pengujian Sistem

3.1.1 Impor Data Penjualan *Coffe Shop X*

Tahapan pertama adalah menginput data yang akan di olah menjadi sebuah *dataframe* agar dapat di proses menggunakan Bahasa pemrograman *python*. Adapun tahapannya terdapat pada gambar 1:



transaction_id	transaction_date	transaction_time	transaction_qty	store_id	store_location	product_id	unit_price	product_category	product_type	product_detail	Revenue	Month	Month Name	Day	Hour
0	1	1/1/2022	72611	2	5	Upper Manhattan	12	3.0	Coffee	Control Roasted coffee	21600	1	Jan	Sun	7
1	2	1/1/2022	72616	2	5	Upper Manhattan	17	2.4	Tea	Brewed Chai Tea	16200	1	Jan	Sun	7
2	3	1/1/2022	72614	2	5	Upper Manhattan	16	4.5	Drinking Accessories	Hot chocolate	27000	1	Jan	Sun	7
3	4	1/1/2022	72614	1	5	Upper Manhattan	22	2.0	Coffee	Triple coffee	12000	1	Jan	Sun	7
4	5	1/1/2022	72241	2	5	Upper Manhattan	17	2.4	Tea	Brewed Chai Tea	16200	1	Jan	Sun	7

Gambar 1 Tahapan import data

3.1.2 Mengidentifikasi Data

Tahapan selanjutnya adalah mengidentifikasi data untuk mengetahui informasi tentang data seperti variabel data, *missing value*, tipe data, *duplicated data*, dan jumlah data. Adapun tahapannya terdapat pada gambar 2:

```

coffee_shop_sales.info()
<class 'pandas.core.frame.DataFrame'>
Int64Index: 140116 entries, 0 to 140115
Data columns (total 16 columns):
 #   Column                Non-Null Count  Dtype  
---  --
 0   transaction_id         140116 non-null  int64  
 1   transaction_date       140116 non-null  datetime64[ns]
 2   transaction_time       140116 non-null  datetime64[ns]
 3   transaction_qty        140116 non-null  int64  
 4   store_id              140116 non-null  int64  
 5   store_location         140116 non-null  object  
 6   product_id            140116 non-null  int64  
 7   unit_price            140116 non-null  float64 
 8   product_category       140116 non-null  object  
 9   product_type           140116 non-null  object  
10   product_detail         140116 non-null  object  
11   Revenue               140116 non-null  object  
12   Month                 140116 non-null  int64  
13   Month Name            140116 non-null  object  
14   Day                   140116 non-null  object  
15   Hour                  140116 non-null  int64  
dtypes: datetime64[ns](2), float64(1), int64(6), object(7)
memory usage: 18.2+ MB

```

Gambar 2 Tahap Identifikasi Data

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan analisis yang dilakukan pada kasus di atas, maka ditarik kesimpulan berikut:

1. *Exploratory Data Analysis* (EDA) dan visualisasi data memakai *Tableau* terbukti efektif dalam menganalisis kinerja penjualan Coffee Shop X. Dari Januari hingga Juni 2023, kategori kopi menjadi yang paling dominan dengan total penjualan 58.416 unit (39,1%) dan pendapatan tertinggi sebesar Rp2.699.524,45 (52,3%). Jumlah transaksi meningkat setiap bulan, dari 17.313 di Januari hingga 35.352 di Juni, dengan puncak harian tertinggi pada 16 Juni sebanyak 1.331 transaksi. Produk terlaris secara individual adalah *Brewed Chai Tea*, mencerminkan tingginya minat konsumen terhadap produk tersebut.
2. Visualisasi data yang dihasilkan melalui *Tableau* memudahkan proses analisis, sehingga informasi dapat disampaikan dengan lebih jelas, menarik, dan mudah dipahami. Hasil visualisasi ini sangat penting untuk mengidentifikasi waktu puncak penjualan, produk dengan performa rendah, serta kontribusi masing-masing cabang terhadap total penjualan. Informasi ini bisa jadi acuan yang kuat untuk *stakeholder* dalam mengambil keputusan strategis untuk mengembangkan *coffee shop* ke arah yang lebih baik.

Selama penelitian, Disadari bahwa penelitian ini belum sepenuhnya sempurna. Dengan demikian, terdapat beberapa aspek yang perlu dipertimbangkan untuk pengembangan studi selanjutnya, yaitu:

1. Melakukan evaluasi terhadap produk-produk dengan performa penjualan rendah, seperti strategi promosi, *bundling* produk, atau pengembangan varian baru dapat menjadi solusi untuk meningkatkan daya tarik produk-produk tersebut.
2. Visualisasi *Tableau* yang telah dibuat sudah cukup informatif. Namun, ke depan dapat dikembangkan dengan menambahkan filter interaktif, seperti berdasarkan lokasi toko atau jenis pelanggan, agar analisis bisa lebih dinamis dan mendalam.

REFERENSI

- [1] Johannes Kurniawan, *Analisis Dan Visualisasi Data*, vol. 3, no. 1. 2023.
- [2] A. Singh Bangari and C. Hava Muntean, "A Comparative Evaluation of Machine Learning Models and EDA through Tableau Using CICIDS2017 Dataset MSc Research Project Data Analyticss".
- [3] F. Rozi Lubis, M. Khoiruddin Harahap, and A. Mahmud Husein, "Analisis Prediktif untuk Keputusan Bisnis : Peramalan Penjualan," *J. Data Sci. Indones.*, pp. 32–40, 2019.
- [4] M. Z. Siambaton and A. M. Husein, "Menganalisis Data Kesehatan Global : Pendekatan Analisis Data Eksplorasi Visual," *Data Sci. Indones.*, vol. 1, no. 2, pp. 41–49, 2022, doi: 10.47709/dsi.v1i2.1315.
- [5] A. A. Zulyani, A. S. Y. Irawan, and A. Jamaludin, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Tingkat Vaksinasi Pada Kecamatan Tambun Selatan," *J. Soc. Sci. Res.*, vol. 3, no. 3, pp. 7037–7050, 2023.
- [6] W. Aryunani, M. Aulia, I. Purwanto, Y. Setiani, and I. Purnama, "Visualisasi Data Student Mental Health Menggunakan Looker Studio Dengan Metode Sna," *J. Inf. dan Komput.*, vol. 12, no. 01, pp. 133–142, 2024.
- [7] R. Mayasari, B. Nugraha, A. R. Juwita, and N. Heryana, "Analisis Produktifitas Padi di Pulau Sumatera menggunakan Exploratory Data Analysis (EDA)," *J. Elektron. Sist. Inf. Unsika*, vol. 1, no. 1, pp. 17–24, 2023.
- [8] E. S. Endang Lestariningsih, Eka Ardhianto, W.T. Handoko, "Visualisasi Data Penduduk Berbasis Web Di Kelurahan," *Din. Inform.*, vol. 10, no. 1, pp. 25–32, 2018.
- [9] K. Sahoo, A. K. Samal, J. Pramanik, and S. K. Pani, "Exploratory data analysis using python," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 4727–4735, 2019, doi: 10.35940/ijitee.L3591.1081219.
- [10] A. N. Syahrudin and T. Kurniawan, "Input dan output pada bahasa pemrograman python," no. January, 2020.
- [11] M. Radhi, A. Amalia, D. R. H. Sitompul, S. H. Sinurat, and E. Indra, "Analisis Big Data Dengan Metode Exploratory Data Analysis (Eda) Dan Metode Visualisasi Menggunakan Jupyter Notebook," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 4, no. 2, pp. 23–27, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2475.