

# Analisis Data Penjualan Menggunakan Teknik Pembersihan Data dan Visualisasi Dengan Python Pada Dataset CSV

Miftahul Jannah<sup>1</sup>, Nurliana<sup>2</sup>, Septinur Selase<sup>3</sup>

<sup>1,2</sup>Program Studi Teknik Informatika, Universitas Asahan

Jl. Jend. Ahmad Yani, Kisaran

<sup>1</sup>miftahstudy@gmail.com, <sup>2</sup>nsinaga420@gmail.com, <sup>3</sup>septinurselasesepti@gmail.com

**Abstract** - Unstructured or difficult-to-analyze data poses potential risks to businesses. Therefore, data visualization plays a critical role in presenting information in a more comprehensible and communicative manner. This research aims to generate valid strategic information through the analysis and visualization of sales data, identify significant trends such as purchasing patterns based on time, region, and product categories, and provide easily interpretable charts and tables to support effective decision-making by stakeholders. The research workflow consists of several stages: data collection, data processing, development of a data analysis system, implementation of the data analysis system, and finally, system testing. The outcomes include a cleaned CSV file and visualized results in the form of charts and tables, which can be utilized as strategic information to support business decision-making.

**Keywords** - Dataset, Visualization, Python.

**Abstrak** - Data yang tidak terstruktur atau sulit dianalisis berpotensi menimbulkan risiko bagi bisnis. Oleh karena itu, visualisasi data memiliki peran yang penting untuk menyajikan informasi dari data dengan cara yang lebih mudah dipahami dan komunikatif. Penelitian ini bertujuan untuk menghasilkan informasi strategis yang valid melalui analisis dan visualisasi data penjualan, mengidentifikasi tren signifikan seperti pola pembelian berdasarkan waktu, wilayah, dan kategori produk, serta menyediakan grafik dan tabel yang mudah dipahami untuk mendukung pengambilan keputusan yang efektif oleh pemangku kepentingan. Tahapan kerja penelitian ada tahap pengumpulan data, pengolahan data, pengembangan sistem analisis data, implementasi sistem analisis data, dan terakhir menguji sistem. Hasil dan pembersihan data berupa file csv yang telah dibersihkan dan hasil dari visualisasi berupa grafik serta tabel dimana hal ini dapat dimanfaatkan sebagai informasi strategis dalam mendukung pengambilan keputusan bisnis.

**Kata Kunci** - Dataset, Visualisasi, Python.

## 1. PENDAHULUAN

Di era digital saat ini, data menjadi aset yang sangat berharga bagi berbagai bisnis karena dapat memberikan pemahaman mendalam tentang perilaku pelanggan, tren pasar, dan kinerja organisasi secara keseluruhan. Akan tetapi, data yang tidak terstruktur atau sulit dianalisis berpotensi menimbulkan risiko bagi bisnis. Oleh karena itu, visualisasi data dan pembersihan data memiliki peran yang penting untuk menyajikan informasi dari data dengan cara yang lebih mudah dipahami dan komunikatif [1]. Sistem Informasi Manajemen visualisasi data telah berkembang dari sekadar penyedia data menjadi alat penting yang membantu manajer merencanakan, mengorganisir, membersihkan, dan memantau operasi bisnis [2].

Seiring dengan pertumbuhan perusahaan yang semakin pesat, jumlah dan jenis data yang dimiliki juga semakin beragam. Salah satu jenis data yang memiliki peran penting bagi perusahaan adalah data penjualan. Aktivitas penjualan memiliki peran yang signifikan karena keuntungan atau kerugian yang diperoleh melalui penjualan menjadi salah satu faktor yang membentuk kinerja dan nilai perusahaan secara keseluruhan. Perusahaan dapat meningkatkan

keuntungan dengan melakukan penjualan melalui berbagai metode, baik secara tunai maupun kredit [3]. Salah satu strategi yang efektif untuk mengatasi berbagai tantangan bisnis adalah dengan mengoptimalkan proses pengambilan keputusan. Pengambilan keputusan yang tepat, didukung oleh analisis data yang akurat dan dilakukan pada waktu yang tepat, dapat membantu mengenali peluang pasar, memahami perilaku konsumen, meningkatkan efisiensi operasional, serta menghadirkan pengalaman yang memuaskan bagi pelanggan. Dalam konteks ini, pendekatan *Business Intelligence* (BI) memiliki relevansi yang tinggi untuk mendukung proses visualisasi dan pembersihan data yang lebih efektif serta berbasis data [4].

Pengolahan data yang baik dapat meminimalkan kesalahan data dan juga dapat mengidentifikasi kebutuhan masa depan yang berperan penting dalam pengembangan data. Analisis data dapat dilakukan dengan berbagai cara, yaitu menggunakan ilmu data atau metode lainnya. Menggambarkan pembersihan data secara nyata, baik yang berbentuk tabel, grafik batang, grafik pie, grafik garis, peta, dan diagram, disebut visualisasi data [5].

Maka dari itu, diperlukan analisis dan perancangan teknologi informasi yang dapat

mengatasinya, yaitu dengan membangun sebuah penelitian melalui analisis kebutuhan perusahaan yang kemudian dituangkan ke dalam aplikasi pengolahan data untuk laporan penjualan. Aplikasi ini dapat memberikan wawasan dan informasi tambahan kepada manajemen guna mengambil keputusan terkait laporan penjualan secara *real time*. Hal ini dilakukan untuk mengurangi atau menghilangkan potensi risiko kesalahan dalam mengambil keputusan terhadap strategi bisnis perusahaan. Salah satu fungsi statistik yang dilakukan adalah meringkas atau melakukan analisis dan visualisasi data penjualan menggunakan *Python* pada *dataset CSV* rangkuman data, sehingga data tersebut menghasilkan informasi yang mudah dipahami oleh banyak orang [3].

Sektor penjualan, termasuk makanan, pakaian, dan penginapan, semakin mengadopsi aplikasi sebagai platform utama. Banyak pengguna kini bergantung pada aplikasi untuk berkomunikasi, memberikan ulasan, dan berbagi pengalaman. Hal ini memungkinkan pengguna produk di platform tersebut untuk memberikan pendapat serta penilaian terhadap produk yang telah mereka gunakan [6].

Oleh karena itu, analisis data dan pengolahan data menjadi sangat penting bagi sektor penjualan. Pengolahan data satu per satu secara manual adalah tugas yang sulit. Namun, tugas-tugas yang berat seperti pemrosesan data dapat disederhanakan dengan bantuan komputer dan kecerdasan buatan. Penumpukan data akan terjadi seiring dengan semakin banyaknya jumlah data yang tersedia. Pengolahan data yang baik dapat meminimalkan kesalahan data dan juga mengidentifikasi kebutuhan masa depan yang berperan penting dalam pengembangan usaha. Analisis data dapat dilakukan dengan berbagai cara, yaitu menggunakan ilmu data atau metode lainnya. Menggambarkan data secara nyata, baik yang berbentuk tabel, grafik batang, grafik pie, grafik garis, peta, maupun diagram, disebut visualisasi data [5].

Pada visualisasi data dan pembersihan data, *Python* merupakan salah satu bahasa pemrograman berbasis objek. Bahasa pemrograman ini dapat dijalankan pada beberapa platform perangkat lunak melalui berbagai sistem operasi. Selain itu, bahasa pemrograman *Python* mendukung berbagai pustaka (*library*) yang menyediakan fungsi analisis data, pemrosesan data, fungsi *machine learning*, serta visualisasi data [7].

*Python* juga menyediakan berbagai paket untuk visualisasi dan pembersihan data yang memungkinkan pengguna menampilkan informasi dengan cara yang menarik dan mudah dipahami melalui berbagai jenis grafik dan bagan interaktif. Fitur ini menjadikan *Python* sebagai salah satu alat yang penting, tidak hanya dalam analisis data tetapi juga dalam berbagai bidang ilmu data lainnya. Dengan memanfaatkan *Python*, pemahaman terhadap data menjadi lebih mudah dan intuitif bagi penggunanya [5].

Dalam beberapa tahun terakhir, *Python* telah menjadi salah satu bahasa pemrograman yang paling banyak dipilih untuk analisis data dan pengembangan aplikasi berbasis data. Dengan berbagai pustaka dan modul yang tersedia, *Python* mendukung penerapan algoritma pembersihan data, analisis, serta visualisasi dengan cara yang lebih efektif dan fleksibel. Kelebihan *Python* dalam menangani pembersihan dan visualisasi data serta kemudahan penggunaannya menjadikannya pilihan yang populer dalam berbagai proyek berbasis data. Dalam studi ini, kami berfokus pada analisis data penjualan dengan pendekatan pembersihan dan visualisasi data menggunakan *Python* dalam konteks klasifikasi data. Penelitian ini bertujuan untuk menghasilkan informasi strategis yang valid melalui analisis dan visualisasi data penjualan, mengidentifikasi tren signifikan seperti pola pembelian berdasarkan waktu, wilayah, dan kategori produk, serta menyediakan grafik dan tabel yang mudah dipahami untuk mendukung pengambilan keputusan yang efektif oleh pemangku kepentingan [8].

Penelitian ini bertujuan untuk menganalisis data penjualan dengan memanfaatkan teknik pembersihan data dan visualisasi menggunakan *Python* pada *dataset CSV*. Visualisasi data memiliki peran yang penting dalam memahami tren penjualan, perilaku pelanggan, serta pola transaksi yang dapat digunakan untuk meningkatkan performa bisnis. Penggunaan metode visualisasi dengan bantuan *Python* dan penerapan teknik pembersihan data terbukti efektif dalam mengolah serta mempresentasikan informasi penting yang terkandung dalam *dataset*. Penelitian ini juga memanfaatkan layanan *Google Collab* sebagai alat untuk mendukung proses visualisasi dan analisis data. *Google Collab* adalah platform berbasis komputasi awan yang dikembangkan oleh Google, yang menawarkan kemudahan serta fleksibilitas bagi berbagai aktivitas penelitian dan pengolahan data [1].

Berdasarkan penelitian terdahulu dengan judul Analisa Dan Visualisasi Data Penjualan Menggunakan Exploratory Data Analysis Pada PT. Telkominfra menghasilkan sistem informasi yang mampu melakukan pembersihan data, analisis EDA, serta menyajikan hasil dalam bentuk grafik untuk memberikan wawasan kepada manajemen mengenai kontrak yang perlu dievaluasi. Selain itu, penerapan sistem ini dapat mengurangi waktu pengolahan data dari 6 jam menjadi hanya 30 menit. Pada penelitian ini dilakukan dengan dataset yang berbeda dan menampilkan visualisasi data berdasarkan tren [3].

## II. METODE PENELITIAN

Peran data penjualan sangatlah penting dalam pengambilan keputusan bisnis. Penelitian ini bertujuan untuk menganalisis data penjualan dengan melakukan pembersihan data dan visualisasi menggunakan pemrograman *Python* berbasis *dataset CSV*. Proses ini

akan menghasilkan informasi yang valid dan krusial untuk mendukung pengambilan keputusan strategis.

Deskripsi Data:

Penelitian ini menggunakan data yang berasal dari file CSV berupa informasi dari penjualan produk selama tahun 2019 yang diambil dari sumber *dataset* Kaggle yang dapat diakses [disini](#). Kolom yang ada antara lain:

**Order ID:** ID unik untuk setiap transaksi.  
**Product:** Nama produk yang dibeli.  
**Quantity Ordered:** Jumlah produk yang dipesan.  
**Price Each:** Harga per unit produk.  
**Order Date:** Tanggal dan waktu pesanan dilakukan.  
**Purchase Address:** Alamat pengiriman produk.  
**Total Price:** Total harga untuk setiap produk ( $Quantity\ Ordered \times Price\ Each$ ).

Proses Pembersihan Data:

- Menghapus baris kosong dan data yang tidak relevan.
- Mengonversi kolom 'Order Date' menjadi tipe *datetime*.
- Menghitung kolom 'Sales' jika tidak ada dengan menggunakan rumus:  

$$Sales = Quantity\ Ordered \times Price\ Each$$

Teknik Analisis:

1. Penjualan Berdasarkan Bulan  
Analisis dilakukan untuk mengetahui pola penjualan berdasarkan bulan selama tahun 2019. Informasi bulan diekstrak dari kolom 'Order Date', lalu data dikelompokkan berdasarkan bulan untuk menghitung total penjualan. Hasilnya divisualisasikan menggunakan diagram batang untuk menunjukkan tren penjualan bulanan.
2. Penjualan Berdasarkan Wilayah Kota  
Analisis ini bertujuan untuk mengetahui kota dengan penjualan tertinggi. Nama kota diekstrak dari kolom 'Purchase Address', kemudian data dikelompokkan berdasarkan kota untuk menghitung total penjualan. Hasilnya divisualisasikan dalam diagram batang untuk memudahkan perbandingan antar kota.
3. Penjualan Berdasarkan Jam  
Analisis dilakukan untuk mengetahui tren penjualan berdasarkan waktu dengan mengekstrak informasi jam dari kolom 'Order Date'.
4. Penjualan Berdasarkan Kategori  
Jika data kategori produk tersedia, analisis dilakukan untuk mengetahui produk mana yang memiliki penjualan tertinggi.
5. Pasangan Produk Terbanyak  
Menganalisis pasangan produk yang sering dibeli bersama dengan menggunakan kombinasi produk dalam setiap transaksi.

Alat yang Digunakan:

- *Python Libraries*:

- *pandas*: Untuk manipulasi dan analisis data.
- *matplotlib*: Untuk visualisasi data.
- *itertools* dan *collections.Counter*: Untuk analisis kombinasi produk.

Tahapan Kerja Penelitian

#### 1. Pengumpulan Data

Tahap pertama adalah mengumpulkan data yang akan digunakan dalam penelitian. Data dikumpulkan melalui penelitian perpustakaan (*library research*), penelitian lapangan (*field research*), serta analisis *dataset* yang tersedia. *Dataset* utama berupa data penjualan dalam format CSV yang diperoleh dari sumber *dataset* Kaggle.

#### 2. Pengolahan Data

Data yang telah dikumpulkan diolah dengan tujuan mengelompokkan dan membersihkan data. Proses ini mencakup validasi data untuk memastikan konsistensi serta penanganan data yang hilang atau tidak relevan. Pengolahan ini dilakukan untuk mempersiapkan *dataset* agar siap dianalisis lebih lanjut.

#### 3. Pengembangan Sistem Analisis Data

Tahapan ini melibatkan perancangan sistem analisis data menggunakan *Integrated Development Environment (IDE)* yaitu *Visual Studio Code (VSCoDe)*. Proses ini melibatkan pembuatan skrip *Python* untuk membaca *dataset*, membersihkan data, dan memvisualisasikan hasil analisis menggunakan *library* seperti *pandas*, *matplotlib*, *itertools*, dan *collections.Counter*.

#### 4. Implementasi Sistem Analisis Data

Implementasi dilakukan dengan menjalankan skrip *Python* yang telah dirancang pada *dataset* penjualan. Tahap ini mencakup pengaplikasian teknik pembersihan data serta pembuatan visualisasi untuk menggambarkan pola atau tren penjualan.

#### 5. Menguji Sistem

Tahap pengujian dilakukan untuk mengevaluasi akurasi dan keandalan sistem analisis data. *Dataset* yang berbeda digunakan untuk menguji efektivitas proses pembersihan data dan visualisasi. Hasil dari pengujian ini digunakan untuk mengidentifikasi dan memperbaiki kelemahan sistem sebelum hasil akhir dianalisis dan disajikan.

### III. HASIL DAN PEMBAHASAN

Dari penelitian yang telah dilakukan menghasilkan Analisis tren penjualan. Untuk lebih jelasnya dapat dilihat pada gambar berikut :

Hasil Pembersihan Data

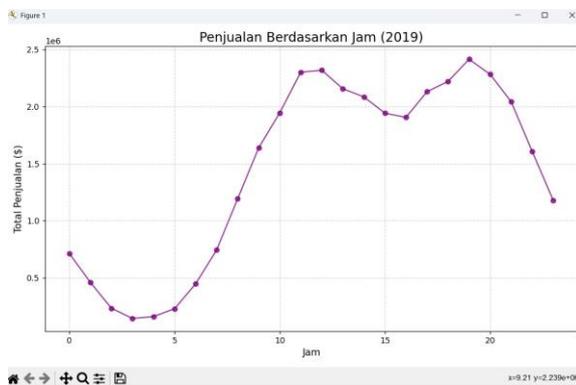
Teknik pembersihan data yang dilakukan pada *dataset Monthly Sales Data of sell of electronics equipment across USA in 2019* berupa file CSV yang dilakukan

menggunakan bahasa pemrograman *python* dengan hasil dari pemrograman adalah file data set CSV yang sudah dibersihkan dengan menghapus baris kosong, menghapus baris yang tidak memiliki nilai tanggal yang valid pada kolom 'Order Date', mengonversi kolom 'Order Date' ke format tanggal, menghitung kolom 'Total Price'. Setelah dilakukan pembersihan data maka data siap diproses *dataset* akan disimpan menjadi file 'Cleaned Sales Data.csv' lalu dilanjutkan ke tahap visualisasi dan pengujian.

Hasil Visualisasi Pada *Dataset Monthly Sales Data of sell of electronics equipment across USA in 2019*

A. *Penjualan Berdasarkan Jam*

Setelah analisis dilakukan berdasarkan jam, ditemukan bahwa jam-jam dengan penjualan tertinggi terjadi pada pukul 19:00. Grafik berikut menunjukkan distribusi penjualan per jam sepanjang tahun 2019:

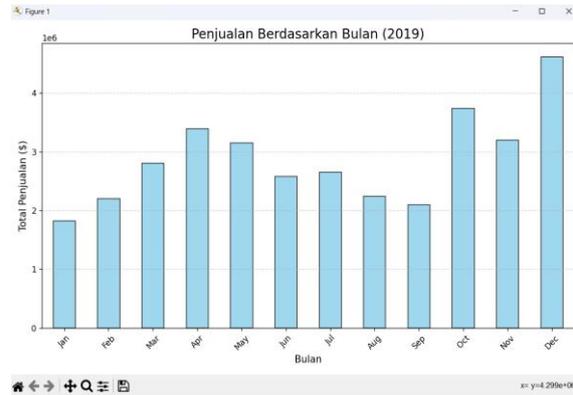


Gambar 1. Penjualan Berdasarkan Jam

Berdasarkan gambar grafik hasil dari penjualan tertinggi terjadi pada jam 19.00 dengan urutan selanjutnya dengan rentan waktu 11.00 - 12.00 hal ini menggambarkan tingkat tertinggi aktivitas penjualan berdasarkan trend waktu. Hasil ini menunjukkan bahwa pelanggan cenderung melakukan pembelian pada waktu-waktu tertentu, yang dapat membantu perusahaan dalam merencanakan promosi atau penawaran khusus pada jam-jam tersebut.

B. *Penjualan Berdasarkan Bulanan*

Setelah analisis dilakukan berdasarkan penjualan tiap bulan, ditemukan bahwa bulan dengan penjualan tertinggi terjadi pada bulan Desember. Grafik berikut menunjukkan distribusi penjualan per bulan pada tahun 2019:

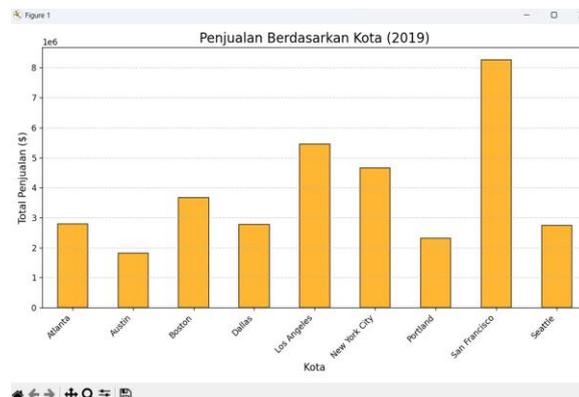


Gambar 2. Penjualan Berdasarkan Bulan

Berdasarkan gambar grafik hasil dari penjualan tertinggi terjadi pada bulan Desember. Hal ini menggambarkan tingkat tertinggi aktivitas penjualan adalah pada akhir tahun.

C. *Penjualan Berdasarkan Kota*

Setelah analisis dilakukan berdasarkan penjualan di berbagai wilayah kota, ditemukan bahwa kota dengan penjualan tertinggi terjadi di kota San Francisco. Grafik berikut menunjukkan distribusi penjualan di beberapa kota pada tahun 2019:

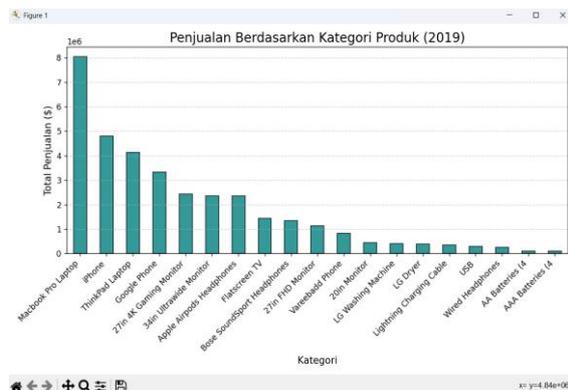


Gambar 3. Penjualan Berdasarkan Kota

Berdasarkan gambar grafik hasil dari penjualan tertinggi terjadi di San Francisco. Hal ini menggambarkan tingkat tertinggi aktivitas penjualan adalah di kota San Francisco.

D. *Penjualan Berdasarkan Kategori*

Setelah analisis dilakukan berdasarkan kategori produk, ditemukan bahwa MacBook Pro Laptop menjadi produk dengan penjualan tertinggi. Grafik berikut menunjukkan distribusi penjualan produk berdasarkan kategori sepanjang tahun 2019:



Gambar 4. Penjualan Berdasarkan Kategori Produk

Berdasarkan gambar grafik hasil dari penjualan tertinggi yaitu MacBook Pro Laptop. Hal ini menggambarkan tingkat tertinggi pembelian produk Macbook Pro Laptop. Dari hasil analisis penjualan berdasarkan kategori, produk kategori Elektronik menunjukkan total penjualan tertinggi. Ini menunjukkan bahwa produk ini adalah yang paling diminati oleh pelanggan.

E. Penjualan Berdasarkan Kombinasi Produk

Setelah analisis dilakukan berdasarkan penjualan produk yang dibeli berpasangan atau kombinasi, ditemukan bahwa pembelian produk tertinggi dibeli berpasangan adalah iPhone dan Lightning Charging Cable. Hasil pembelian 10 produk kombinasi tertinggi berdasarkan jumlah pembelian pada tahun 2019:

Tabel 1. Sepuluh Produk Kombinasi Penjualan Tertinggi

No	Produk Kombinasi	Jumlah Pembelian
1	iPhone, Lightning Charging Cable	1005 kali
2	Google Phone, USB-C Charging Cable	987 kali
3	iPhone, Wired Headphones	447 kali
4	Google Phone, Wired Headphones	414 kali
5	Vareebadd Phone, USB-C Charging Cable	361 kali
6	iPhone, Apple AirPods Headphones	360 kali
7	Google Phone, Bose SoundSport Headphones	220 kali
8	USB-C Charging Cable, Wired Headphones	160 kali
9	Vareebadd Phone, Wired Headphones	143 kali
10	Lightning Charging Cable, Wired Headphones	92 kali

Berdasarkan tabel hasil dari penjualan produk kombinasi yang dibeli bersamaan tertinggi yaitu pembelian produk iPhone dengan Lightning Charging

Cable. Hal ini menggambarkan tingkat pembelian produk yang dibeli bersamaan dengan rata-rata penjualan per order adalah \$193.30.

Hasil Dari Pengujian Akhir

Setelah dilakukan pengujian dengan mengecek data, menguji kolom 'Total Price', pengujian apakah dapat menampilkan hasil visualisasi tanpa error. Hasil yang didapat dari pengujian ini adalah pesan sebagai berikut:

1. Data tidak mengandung nilai NaN
  2. Perhitungan Total Price sudah benar
  3. Visualisasi Penjualan Bulanan berhasil dibuat
- Dengan pesan berikut maka pengujian akhir telah dilakukan dan tidak ditemukan masalah pada saat proses implementasi.

IV. KESIMPULAN

Berdasarkan pelaksanaan penelitian, implementasi, dan pengujian, beberapa poin penting dapat disimpulkan yaitu analisis data penjualan melalui proses pembersihan data dan visualisasi dengan menggunakan Python serta dataset CSV mempermudah identifikasi pola dan tren yang relevan, sehingga mendukung pengambilan keputusan strategis dalam bisnis. Proses pembersihan data yang dilakukan berhasil meningkatkan akurasi dataset, memastikan bahwa informasi yang dihasilkan dapat diandalkan dan memiliki kualitas yang lebih baik. Hasil visualisasi data memberikan representasi yang sederhana dan mudah dipahami terkait performa penjualan, yang sangat bermanfaat bagi pengelola bisnis dalam merencanakan strategi untuk meningkatkan penjualan. Informasi strategis yang dapat dimanfaatkan dalam mengambil keputusan bisnis adalah data penjualan tertinggi yang terjadi pada pukul 19:00, bulan Desember, dan di kota San Francisco. MacBook Pro Laptop menjadi produk dengan penjualan tertinggi, sedangkan kombinasi iPhone dan Lightning Charging Cable sering dibeli bersama. Hal ini dapat dijadikan rencana bisnis dengan memfokuskan promosi dan kampanye pemasaran pada pukul 19:00, terutama pada bulan Desember. Contohnya mengadakan diskon waktu terbatas atau iklan khusus untuk meningkatkan penjualan. Utamakan pemasaran di kota San Francisco, misalnya dengan promosi lokal atau peningkatan stok produk. Tingkatkan ketersediaan dan variasi aksesoris yang relevan untuk produk populer seperti MacBook Pro Laptop. Serta memberikan promosi seperti iPhone dengan Lightning Charging Cable, dan tawarkan dengan diskon khusus untuk mendorong pembelian pasangan produk yang sering dibeli bersama. Untuk penelitian berikutnya, disarankan untuk memasukkan faktor eksternal seperti promosi, kondisi cuaca, atau pola musiman yang dapat mempengaruhi penjualan. Penambahan variabel-variabel tersebut akan memberikan analisis yang lebih menyeluruh dan

memperkaya wawasan mengenai pola belanja konsumen.

#### DAFTAR PUSTAKA

- [1] R. G. Guntara, “Visualisasi Data Laporan Penjualan Toko Online Melalui Pendekatan Data Science Menggunakan Google Colab,” *ULIL ALBAB: Jurnal Ilmiah Multidisiplin*, vol. 2, no. 6, hlm. 2091, Mei 2023.
- [2] S. Mujilahwati, “Visualisasi Data Hasil Klasifikasi Naïve Bayes dengan Matplotlib pada Python,” dalam *Prosiding SNST ke-11 Tahun 2021*, Fakultas Teknik Universitas Wahid Hasyim, hlm. 205, 2021.
- [3] A. H. Wibowo, K. Faisah, dan Y. Devianto, “Analisa dan Visualisasi Data Penjualan Menggunakan Exploratory Data Analysis Pada PT. Telkom Infra,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 3, hlm. 2292–2304, Sept. 2022.
- [4] R. Senjaya, B. N. Sari, dan I. Purnamasari, “Implementasi Business Intelligence pada Toko Smart-S dalam Membantu Proses Analisis Bisnis dengan Metode OLAP,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, hlm. 5906, Agustus 2024.
- [5] A. K. Putri dan D. I. Nur, “Penggunaan Bahasa Python untuk Analisis dan Visualisasi Data Penduduk di Desa Sumberjo, Nganjuk,” *KARYA: Jurnal Pengabdian Kepada Masyarakat*, vol. 3, no. 3, hlm. 206–217, 2023.
- [6] D. R. F. Daud, B. Irawan, dan A. Bahtiar, “Penerapan Metode Naive Bayes pada Analisis Sentimen Aplikasi McDonald’s di Google Play Store,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, hlm. 759, Feb. 2024.
- [7] C. M. Hellyana, “Implementasi Bahasa Python untuk Clusterisasi Data Penjualan Menggunakan Metode K-Means,” *Informatics and Computer Engineering Journal*, vol. 3, no. 2, hlm. 140, Agustus 2023.
- [8] P. S. Zakaria, R. Julianto, dan R. S. Bernada, “Implementasi Naive Bayes Menggunakan Python dalam Klasifikasi Data,” dalam *Buletin Ilmiah Ilmu Komputer dan Multimedia*, vol. 1, no. 1, hlm. 126–131, Mei 2023.