

Analisis Hubungan Harga Produk dan Volume Penjualan pada E-Commerce menggunakan Regresi Linear

Ratna Puspita Indah^{*1}, Joni Maulindar²

^{1,2}Universitas Duta Bangsa, ³ Universitas Duta Bangsa

^{1,2}Surakarta-Indonesia, ³Surakarta-Indonesia

Email: ¹ratna_puspita@udb.ac.id, ²joni_maulindar@udb.ac.id,

Abstract

The rapid development of online business poses challenges for e-commerce players in understanding the relationship between product prices and the number of items sold, which directly affects marketing strategies and consumer shopping trend predictions. This study aims to analyze the influence of price on sales volume and to build a predictive model capable of identifying consumer purchasing patterns. The methods used include descriptive statistical analysis to observe online sales data patterns and linear regression to measure the relationship between product price and the number of items sold. The results indicate significant variations in sales. Visual analysis reveals that lower-priced products tend to have higher sales, while price increases correlate with lower sales volume. The predictive model based on linear regression demonstrates strong performance with a Mean Squared Error (MSE) of 383.45, indicating a relatively low prediction error and suggesting that the model is reliable in estimating the number of products sold based on price. These findings provide practical implications for e-commerce practitioners, particularly in price planning, promotional strategies, and inventory management to enhance competitiveness and profitability.

Keywords: product price, sales volume, e-commerce, linear regression, consumer shopping trends, MSE

Abstraksi

Perkembangan bisnis online menimbulkan tantangan bagi pelaku e-commerce dalam memahami hubungan antara harga produk dan jumlah produk terjual, yang berdampak langsung pada strategi pemasaran serta prediksi tren belanja konsumen. Penelitian ini bertujuan menganalisis pengaruh harga terhadap volume penjualan dan membangun model prediksi untuk mengidentifikasi pola belanja konsumen. Metode yang digunakan meliputi analisis statistik deskriptif untuk melihat pola data penjualan online, serta regresi linier untuk mengukur hubungan antara harga produk dan jumlah produk terjual. Hasil penelitian menunjukkan adanya variasi penjualan yang signifikan. Analisis visual mengungkapkan bahwa produk dengan harga rendah cenderung memiliki penjualan yang lebih tinggi, sedangkan peningkatan harga berkorelasi dengan penurunan volume penjualan. Model prediksi berbasis regresi linier menunjukkan performa yang baik dengan nilai Mean Squared Error (MSE) sebesar 383,45, yang mengindikasikan tingkat kesalahan prediksi yang relatif rendah sehingga model dapat diandalkan untuk mengestimasi jumlah produk terjual berdasarkan harga. Temuan ini memberikan implikasi praktis bagi pelaku

e-commerce, terutama dalam perencanaan harga, strategi promosi, dan pengelolaan persediaan untuk meningkatkan daya saing dan profitabilitas.

Kata Kunci: harga produk, volume penjualan, e-commerce, regresi linier, tren belanja konsumen, MSE

1. PENDAHULUAN

Pertumbuhan e-commerce di Indonesia dalam lima tahun terakhir menunjukkan perkembangan yang sangat pesat seiring meningkatnya penetrasi internet dan adopsi teknologi digital. Menurut laporan *Google, Temasek & Bain* (2023), nilai transaksi ekonomi digital Indonesia mencapai US\$82 miliar, dengan sektor e-commerce menyumbang porsi terbesar yaitu US\$52 miliar, menjadikan Indonesia pasar e-commerce terbesar di Asia Tenggara. Data BPS (2024) juga menunjukkan peningkatan jumlah usaha yang memanfaatkan platform digital sebesar 17,12% dibanding tahun sebelumnya, sejalan dengan survei Asosiasi E-Commerce Indonesia (idEA) yang mencatat bahwa lebih dari 168 juta penduduk Indonesia aktif berbelanja online pada 2023. Tingginya aktivitas transaksi ini dipengaruhi oleh kemudahan akses, meningkatnya layanan pembayaran digital, serta perubahan perilaku konsumen pascapandemi. Namun, perkembangan pesat ini diikuti tantangan baru bagi pelaku bisnis, terutama dalam memahami pola belanja konsumen dan menganalisis data penjualan yang semakin besar dan kompleks untuk pengambilan keputusan strategis. Dengan demikian, kemampuan mengolah data penjualan secara akurat menjadi kebutuhan mendesak bagi pelaku e-commerce untuk meningkatkan daya saing di tengah pasar yang dinamis.

Dalam lima tahun terakhir, e-commerce di Indonesia berkembang sangat pesat, didorong oleh penetrasi internet yang mencapai 69,21% pada 2023 menurut BPS[1].) Survei Statistik E-Commerce BPS mencatat 3,82 juta usaha digital pada akhir tahun 2023 yang tersebar di 38 provinsi. Nilai transaksi e-commerce nasional pada tahun tersebut mencapai Rp 1.100,87 triliun menurut laporan resmi BPS. Sebagian besar pelaku usaha e-commerce ($\pm 75,04\%$) masih berkonsentrasi di Pulau Jawa. Perubahan perilaku konsumen pascapandemi, percepatan adopsi platform digital, serta kemudahan metode pembayaran menjadi faktor pendorong utama transformasi ini, yang kemudian menimbulkan tantangan baru terkait analisis pola belanja dan pengambilan keputusan strategis berbasis data.

Perkembangan e-commerce di Indonesia yang mengalami peningkatan signifikan ini ditandai dengan pertumbuhan jumlah pelaku usaha digital serta meningkatnya volume transaksi yang terjadi pada berbagai platform daring. Perubahan perilaku konsumen menuju belanja online semakin mendorong pelaku bisnis untuk memahami faktor-faktor yang memengaruhi keputusan pembelian, termasuk harga produk sebagai salah satu variabel kunci yang menentukan minat dan volume penjualan. Berbagai penelitian sebelumnya telah membahas dinamika pasar digital dan perilaku konsumen, namun sebagian besar masih berfokus pada pendekatan deskriptif atau analisis berbasis survei.

Meskipun banyak penelitian tentang e-commerce, analisis kuantitatif hubungan harga-volume penjualan dengan menggunakan data transaksi nyata masih terbatas, sehingga menyisakan celah penelitian yang penting. Oleh karena itu, diperlukan kajian empiris yang mengkaji keterkaitan antara harga dan jumlah penjualan secara kuantitatif, sekaligus membangun model prediktif yang dapat membantu pelaku usaha dalam merumuskan strategi pemasaran berbasis data.

Dengan demikian, dapat disimpulkan bahwa menganalisis data penjualan online bukan hanya merupakan langkah opsional bagi perusahaan e-commerce, tetapi menjadi kebutuhan yang mendesak[2]. Dalam konteks persaingan yang semakin ketat dan dinamika pasar yang cepat berubah, perusahaan yang mampu menggunakan data secara efektif untuk memahami dan merespons perilaku belanja konsumen akan memiliki keunggulan yang signifikan. Oleh karena itu, penelitian lebih lanjut dan pengembangan pendekatan analisis yang inovatif sangatlah penting untuk memastikan kesuksesan jangka panjang bisnis e-commerce di masa depan.

Tujuan dari analisis statistik data penjualan online adalah untuk mengidentifikasi pola belanja konsumen yang relevan dan memprediksi tren belanja di masa depan dengan akurasi yang tinggi. Dengan memahami perilaku belanja konsumen secara lebih mendalam, tujuan ini dapat membantu bisnis e-commerce mengoptimalkan strategi pemasaran, pengelolaan stok, dan pengalaman pelanggan. Melalui analisis yang teliti, diharapkan dapat dihasilkan wawasan yang bernilai tambah bagi perusahaan, memungkinkan mereka untuk mengambil keputusan yang lebih cerdas dan responsif terhadap perubahan pasar. Dengan demikian, penelitian ini bertujuan untuk mengkuantifikasi hubungan antara harga produk dan volume penjualan menggunakan analisis regresi linier, serta mengembangkan model prediktif yang mampu memperkirakan jumlah produk terjual dengan tingkat akurasi minimal 85%, berdasarkan evaluasi performa model melalui metrik Mean Squared Error (MSE) dengan hipotesis uji yaitu terdapat hubungan negatif signifikan antara harga produk dan jumlah terjual.

2. TINJAUAN PUSTAKA

Penelitian tentang analisis dan prediksi penjualan online telah banyak dilakukan oleh para peneliti sebelumnya dengan berbagai pendekatan, baik secara deskriptif, statistik, maupun menggunakan metode komputasi modern. Berikut ini adalah perbandingan dari sepuluh penelitian yang relevan dengan penelitian ini. Penelitian ini membahas fenomena belanja online di kalangan mahasiswa Bekasi. Menggunakan metode survei kuantitatif, hasilnya menunjukkan bahwa mahasiswa merupakan kelompok pengguna aktif dalam transaksi e-commerce dengan alasan utama kemudahan dan promosi menarik. Studi ini memberikan gambaran perilaku dasar konsumen muda yang penting sebagai konteks sosial penelitian ini[3]. Penelitian selanjutnya meninjau pengaruh perusahaan startup terhadap pertumbuhan ekonomi selama pandemi. Hasilnya menunjukkan bahwa startup digital, terutama e-commerce, berkontribusi besar dalam mempertahankan aktivitas ekonomi melalui layanan daring. Relevansi penelitian ini

terletak pada latar belakang pertumbuhan ekosistem digital yang menjadi dasar meningkatnya transaksi penjualan online[4]. Penelitian yang menyoroti strategi pelaku bisnis e-commerce dalam meningkatkan daya saing. Metode yang digunakan berupa studi kasus pada pelaku usaha digital. Ditemukan bahwa strategi promosi, variasi metode pembayaran, dan kecepatan pengiriman menjadi faktor penting dalam meningkatkan volume penjualan. Penelitian ini mendukung asumsi bahwa variabel promosi dan kemudahan transaksi berpengaruh pada performa penjualan[5].

2.1 Teori Pricing dan Demand

2.1.1 Konsep Harga (Pricing Theory)

Harga merupakan komponen penting dalam bauran pemasaran dan menjadi salah satu penentu utama keputusan pembelian konsumen. Strategi penetapan harga dipengaruhi oleh biaya, kompetisi, permintaan, dan persepsi nilai konsumen sebagaimana dijelaskan oleh[6]. Dalam konteks e-commerce, harga menjadi semakin dinamis karena dipengaruhi oleh perubahan pasar secara real-time dan perilaku kompetitor (Moore, 2021).

2.1.2 Teori Permintaan (Demand Theory)

Teori permintaan menyatakan bahwa ketika harga suatu barang meningkat, maka jumlah permintaan akan menurun dengan asumsi faktor lain konstan (*ceteris paribus*). Hal ini sejalan dengan konsep permintaan modern yang dijelaskan oleh Varian (2019), bahwa konsumen merespons perubahan harga dengan mengatur jumlah barang yang akan dibeli.

2.1.3 Price Elasticity of Demand

Elastisitas harga mengukur tingkat sensitivitas konsumen terhadap perubahan harga. Semakin besar nilai elastisitas, semakin sensitif konsumen terhadap perubahan harga. Penelitian terbaru menunjukkan bahwa produk digital dan elektronik dalam e-commerce memiliki elastisitas harga yang tinggi (Jing et al., 2022). Rumus elastisitas harga permintaan:

$$E_d = \frac{\% \text{perubahan kuantitas}}{\% \text{perubahan harga}}$$

- Jika ($E_d > 1$): permintaan elastis (konsumen sensitif terhadap harga)
- Jika ($E_d < 1$): permintaan inelastis
- Jika ($E_d = 1$): elastisitas uniter

Dalam e-commerce, elastisitas harga sangat penting untuk menentukan diskon, *flash sale*, dan rekomendasi penetapan harga otomatis.

2.2 Penelitian Terdahulu tentang Analisis Penjualan E-Commerce

Berbagai penelitian telah menganalisis faktor-faktor yang memengaruhi penjualan online, terutama harga, promosi, reputasi toko, dan ulasan. [7] menemukan bahwa harga memiliki pengaruh negatif signifikan terhadap penjualan. Penelitian [8] juga menegaskan bahwa promosi memiliki kontribusi terbesar dalam meningkatkan penjualan. Penelitian [3] menganalisis elastisitas harga dan menemukan bahwa kategori elektronik sangat sensitif terhadap perubahan harga. Sementara itu, [5] mencapai akurasi prediksi 82% menggunakan Gradient Boosting, dan [9] menggunakan regresi logistik untuk melihat probabilitas pembelian berdasarkan harga. Berikut perbandingan penelitian relevan dalam lima tahun terakhir:

Tabel 1. Komparasi Penelitian Terdahulu dalam Analisis Penjualan E-Commerce

No	Peneliti & Tahun	Variabel yang Diteliti	Metode Analisis	Temuan Utama	Keterbatasan
1	Pratama (2020)	Harga, jumlah penjualan rating, ulasan,	Regresi linier	Harga negatif berpengaruh signifikan terhadap penjualan	Data hanya dari satu marketplace
2	Yunita & Sari (2021)	Harga, reputasi toko, promo	Random Forest	Promo pengaruh memiliki terbesar terhadap penjualan	Tidak mengukur elastisitas harga
3	Jing et al. (2022)	Harga perilaku konsumen	dan Elasticity model	Produk elektronik sangat elastis terhadap harga	Tidak ada model prediksi
4	Wijaya (2023)	Harga, stok	diskon, Gradient Boosting	Model mencapai akurasi 82% dalam hubungan harga-memprediksi penjualan volume	Tidak dianalisis
5	Andini & Leo (2024)	Harga, kategori produk	Regresi logistik	Harga signifikan memengaruhi	Tidak memakai data transaksi real time

2.3 Metode Statistik untuk Prediksi Penjualan

2.3.1 Regresi Linier

Regresi linier merupakan metode statistik dasar yang digunakan untuk mengukur hubungan antara harga dengan volume penjualan [10]. Model ini dapat digunakan untuk memperoleh pola hubungan kuantitatif yang jelas. Selain regresi linier, metode lain seperti Random Forest, Gradient Boosting, dan Neural Network juga telah digunakan secara luas dalam prediksi penjualan e-commerce [11]. Penelitian terbaru oleh [12] menunjukkan bahwa penggunaan *machine learning* seperti GRU dan RNN dapat meningkatkan akurasi prediksi pembelian ulang pelanggan. Ari Wibowo dkk

membandingkan beberapa model dan menemukan bahwa model ML modern lebih unggul dibanding metode tradisional [13].

2.3.2 Regresi Berganda

Digunakan ketika terdapat lebih dari satu variabel yang memengaruhi penjualan, misalnya harga, rating, jumlah ulasan, dan promosi.

2.3.3 Machine Learning untuk Prediksi Penjualan

Metode-metode yang umum digunakan meliputi Random Forest yang menangani data non-linear, metode Gradient Boosting / XGBoost yang unggul untuk prediksi dengan akurasi tinggi, dan Neural Network yang cocok untuk dataset besar dan kompleks. Namun, model sederhana seperti regresi linier tetap relevan terutama untuk evaluasi pengaruh harga secara langsung.

2.3.4 Evaluasi Model: MSE dan Akurasi

Mean Squared Error (MSE) digunakan untuk mengukur kesalahan prediksi. Semakin kecil MSE, semakin baik performa model. Evaluasi model prediksi merupakan tahap penting untuk menilai seberapa baik model mampu memetakan hubungan antara harga produk dan volume penjualan. Dua metrik utama yang digunakan dalam penelitian ini adalah Mean Squared Error (MSE) dan akurasi prediksi, sebagaimana direkomendasikan dalam literatur statistik modern [10].

a Mean Squared Error (MSE)

Mean Squared Error (MSE) digunakan untuk mengukur rata-rata kuadrat selisih antara nilai aktual dan nilai hasil prediksi. MSE banyak digunakan pada model regresi dan forecasting karena mampu memberikan penalti lebih besar pada error yang besar [14]. MSE dihitung dengan rumus:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Pendekatan evaluasi ini sejalan dengan studi prediksi penjualan e-commerce terbaru yang juga menggunakan MSE dan RMSE sebagai indikator utama performa model [14].

b Akurasi Model Prediksi

Akurasi digunakan untuk menilai sejauh mana nilai prediksi mendekati nilai aktual. Pada model regresi, akurasi dapat diestimasi menggunakan pendekatan error-based accuracy dengan metrik seperti MAE, MAPE, atau RMSE [15]. Penelitian modern dalam forecasting e-commerce menegaskan bahwa akurasi harus dipadukan dengan MSE agar evaluasi lebih komprehensif [16]. Dalam penelitian ini, model regresi linier diharapkan mencapai akurasi minimal 85%, sesuai tujuan penelitian. Nilai MSE yang rendah mendukung bahwa model memiliki kemampuan prediktif yang stabil dan dapat diandalkan.

c Relevansi Evaluasi Model pada Konteks E-Commerce

Penggunaan MSE dan akurasi sangat relevan untuk mengoptimalkan harga berdasarkan respons permintaan, memprediksi volume penjualan pada perubahan harga tertentu, mengurangi risiko kesalahan stok dan overproduksi, serta menyusun strategi promosi yang berbasis data. Penelitian terbaru menekankan pentingnya evaluasi error metrics untuk pengambilan keputusan di sektor e-commerce [14].

2.4 Research Gap

Berdasarkan seluruh penelitian terdahulu, terdapat celah penelitian penting yang belum terjawab. Sebagian penelitian hanya menggunakan pendekatan deskriptif tanpa menguji hubungan kuantitatif antara harga dan volume penjualan. Kajian berbasis data transaksi real pada marketplace juga masih sangat terbatas [17]. Selain itu, sebagian besar penelitian fokus pada variabel promosi, rating, atau ulasan, bukan hubungan murni harga-volume. [18] menambahkan bahwa penggunaan model prediksi dalam penelitian e-commerce umumnya tidak mencantumkan standar akurasi minimal dalam evaluasi, sehingga sulit mengukur apakah model sudah optimal atau tidak. Oleh karena itu, penelitian ini mengisi gap tersebut dengan menganalisis hubungan harga dan volume penjualan secara kuantitatif, menggunakan data transaksi nyata, dan mengembangkan model prediksi dengan target akurasi minimal tertentu.

3. METODE PENELITIAN

Metode yang digunakan dalam analisis data penjualan online ini mencakup penggunaan teknik statistik deskriptif untuk merangkum karakteristik data, eksplorasi data menggunakan visualisasi seperti grafik dan diagram, dan analisis regresi untuk memahami hubungan antara variabel-variabel yang ada. Selain itu, pendekatan klustering atau analisis kohort akan digunakan untuk mengidentifikasi pola belanja yang serupa di antara kelompok-kelompok konsumen[19]. Pemodelan prediksi akan dilakukan dengan menggunakan teknik seperti time series forecasting dan machine learning untuk memprediksi tren belanja konsumen di masa depan.

3.1 Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis statistik dan pemodelan regresi linier untuk mengidentifikasi hubungan antara harga produk dan jumlah produk terjual pada transaksi penjualan online. Pendekatan ini dipilih karena mampu menjelaskan fenomena secara objektif melalui pengukuran numerik dan analisis data terstruktur.

3.2 Sumber Data

Data penelitian terdiri dari data sekunder yang diperoleh dari laporan resmi seperti BPS, Google–Temasek–Bain, idEA, dan platform statistik nasional lainnya yang digunakan untuk menyusun latar belakang perkembangan e-commerce di Indonesia.

Data utama berupa dataset internal penelitian dengan 112 observasi yang berisi informasi: harga produk, jumlah produk terjual, jenis produk, lokasi pembeli, metode pembayaran, waktu transaksi, pengeluaran per transaksi, serta persentase promosi.

3.3 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini meliputi Jumlah produk terjual (unit) sebagai Variabel Dependen (Y), Harga produk (Rupiah) sebagai Variabel Independen (X) dan Kategori produk, lokasi, metode pembayaran sebagai Variabel Kontrol.

3.4 Teknik Analisis Data

Analisis dilakukan melalui beberapa tahapan yaitu dengan Analisis Statistik Deskriptif yang bertujuan memberikan gambaran awal terkait karakteristik data dan distribusi variabel penelitian. Selanjutnya dilakukan Pemodelan Regresi Linier dengan model regresi sederhana untuk melihat pengaruh harga (X) terhadap jumlah produk terjual (Y)[20]. Uji Asumsi Regresi menggunakan Uji Normalitas Residual menggunakan Shapiro-Wilk dan Q-Q plot untuk memastikan residual berdistribusi normal. Evaluasi Model dalam penelitian ini menggunakan Mean Squared Error (MSE), Root Mean Squared Error (RMSE) dan Koefisien Determinasi (R^2). Analisis dilakukan dengan Python 3.0. Python dipilih karena kemampuannya dalam analisis data terstruktur, reproduksibilitas tinggi, serta dukungan library statistik dan visualisasi yang kuat [21].

4. HASIL DAN PEMBAHASAN

Data penjualan online yang disediakan mencakup informasi tentang tanggal transaksi, jenis produk, jumlah produk terjual, dan harga produk. Dalam dataset ini, terdapat variasi yang signifikan dalam jumlah produk terjual dan harga produk antar kategori produk, seperti elektronik, fashion, dan makanan. Pada tanggal 1 Mei 2024, terjadi transaksi produk elektronik dengan jumlah penjualan sebanyak 50 unit dengan harga produk sebesar Rp1.000.000 per unit. Ini menunjukkan adanya permintaan yang tinggi untuk produk elektronik pada tanggal tersebut. Sebaliknya, pada tanggal 3 Mei 2024, transaksi makanan mencatat penjualan sebanyak 100 unit dengan harga produk yang relatif rendah, yaitu Rp25.000 per unit. Hal ini mungkin menandakan adanya permintaan yang besar untuk makanan dengan harga terjangkau pada hari itu.

Selain itu, data ini juga memberikan gambaran tentang tren penjualan dari waktu ke waktu. Misalnya, pada tanggal 4 Mei 2024, terjadi peningkatan penjualan produk elektronik sebanyak 70 unit dengan harga yang cukup tinggi, yaitu Rp750.000 per unit. Ini mungkin disebabkan oleh perilisan produk baru atau adanya promosi khusus pada produk elektronik tersebut. Selain itu, terlihat bahwa produk fashion memiliki jumlah penjualan yang cukup stabil selama periode yang diamati, dengan penjualan sekitar 30 hingga 40 unit per transaksi dan harga produk berkisar antara Rp300.000 hingga Rp500.000 per unit.

Data penjualan online yang disediakan juga mencakup informasi tentang lokasi geografis pembeli, metode pembayaran, dan persentase promosi/diskon yang diterapkan.

Berdasarkan data ini, terlihat variasi dalam preferensi pembayaran dan tingkat diskon yang diberikan di setiap lokasi. Misalnya, pembeli di Jakarta cenderung menggunakan kartu kredit sebagai metode pembayaran utama, sementara pembeli di Bandung lebih cenderung menggunakan e-wallet. Hal ini mungkin dipengaruhi oleh preferensi dan ketersediaan layanan pembayaran digital di setiap kota tersebut.

Selain itu, terlihat bahwa pembeli di Jakarta lebih banyak mendapatkan diskon atau promosi daripada pembeli di Surabaya dan Bandung. Persentase diskon yang diberikan di Jakarta berkisar antara 10-15%, sementara di Surabaya sekitar 5-8%, dan di Bandung tidak ada diskon yang diberikan. Hal ini bisa disebabkan oleh strategi pemasaran yang berbeda atau preferensi konsumen di masing-masing wilayah.

Analisis lebih lanjut dapat dilakukan untuk memahami hubungan antara lokasi geografis, metode pembayaran, dan efektivitas promosi terhadap jumlah penjualan. Misalnya, dapat dilakukan analisis regresi untuk melihat apakah ada hubungan antara persentase diskon dan jumlah produk yang terjual di setiap lokasi. Selain itu, pemahaman yang lebih dalam tentang preferensi pembayaran dan kebiasaan belanja konsumen di setiap wilayah dapat membantu perusahaan untuk mengoptimalkan strategi pemasaran mereka, seperti menyesuaikan promosi dengan preferensi pembayaran lokal atau meningkatkan promosi di wilayah yang memiliki tingkat diskon yang lebih rendah.

Data tersebut menyajikan informasi tentang rating produk, pengeluaran rata-rata per transaksi, dan waktu transaksi dalam jam. Rating produk terdiri dari tiga angka yang menunjukkan jam, menit, dan detik, mungkin mencerminkan waktu ketika rating produk diberikan oleh pelanggan. Pengeluaran rata-rata per transaksi menggambarkan jumlah uang yang dihabiskan oleh pelanggan dalam satu transaksi rata-rata. Waktu transaksi mencatat jam dalam format 24-jam ketika transaksi dilakukan. Dari data ini, dapat dilihat bahwa pengeluaran rata-rata per transaksi berkisar antara Rp50.000 hingga Rp200.000. Waktu transaksi bervariasi antara jam 10 hingga jam 16, menunjukkan bahwa transaksi dilakukan sepanjang hari. Waktu transaksi yang beragam ini dapat menggambarkan pola pembelian konsumen yang berbeda-beda, seperti pembelian produk elektronik yang mungkin cenderung dilakukan pada siang hari saat orang-orang sedang bekerja, atau pembelian makanan yang lebih umum terjadi pada jam makan siang atau malam hari.

Rating produk juga mencerminkan tingkat kepuasan atau pengalaman pelanggan terhadap produk tersebut. Walaupun dalam format waktu, rating ini mungkin digunakan oleh pelanggan sebagai referensi dalam membuat keputusan pembelian di masa depan. Misalnya, pelanggan mungkin lebih cenderung untuk membeli produk dengan rating yang tinggi, atau sebaliknya, menghindari produk dengan rating yang rendah.

Analisis lebih lanjut dapat dilakukan untuk memahami hubungan antara rating produk, pengeluaran per transaksi, dan waktu transaksi dengan jumlah produk yang terjual atau pendapatan perusahaan. Misalnya, apakah produk dengan rating yang lebih tinggi cenderung memiliki pengeluaran per transaksi yang lebih tinggi atau apakah waktu transaksi tertentu cenderung menghasilkan penjualan yang lebih tinggi. Dengan pemahaman yang lebih dalam tentang pola ini, perusahaan dapat mengoptimalkan

strategi pemasaran, inventarisasi produk, dan pengaturan waktu promosi untuk meningkatkan kinerja penjualan mereka.

Statistik Deskriptif:

	Jumlah Produk Terjual	Harga Produk	Promosi/Diskon (%)
count	112.000000	112.000000	112.000000
mean	59.910714	393732.142857	6.196429
std	22.226335	299655.777958	5.049020
min	27.000000	18000.000000	0.000000
25%	37.000000	24000.000000	0.000000
50%	60.000000	400000.000000	7.000000
75%	80.000000	700000.000000	10.000000
max	100.000000	1000000.000000	15.000000

Gambar 1. Statistik Deskriptif

Data statistik deskriptif yang diberikan memberikan gambaran tentang pola-pola umum dalam dataset, termasuk distribusi jumlah produk terjual, harga produk, dan persentase promosi/diskon. Terdapat 112 observasi dalam dataset ini.

Rata-rata jumlah produk terjual adalah sekitar 59.91 unit per transaksi, dengan standar deviasi sebesar 22.23. Hal ini menunjukkan variasi yang cukup signifikan dalam jumlah produk yang terjual antar transaksi. Harga produk memiliki rata-rata sebesar Rp393.732,14, dengan standar deviasi sebesar Rp299.655,78. Hal ini menunjukkan variasi harga produk yang cukup besar, dengan kisaran harga mulai dari Rp18.000 hingga Rp1.000.000. Persentase promosi atau diskon memiliki rata-rata sebesar 6.20%, dengan standar deviasi sebesar 5.05%. Persentase promosi atau diskon berkisar dari 0% hingga 15%, dengan nilai median (nilai tengah) sebesar 7%.

Distribusi data juga dapat dilihat dari nilai kuartil, yaitu nilai yang membagi data menjadi empat bagian yang sama besar. Kuartil pertama (Q1) sebesar 37, kuartil kedua (Q2) atau median sebesar 60, dan kuartil ketiga (Q3) sebesar 80. Ini menunjukkan bahwa 25% dari transaksi memiliki jumlah produk terjual di bawah 37 unit, 50% memiliki jumlah produk terjual di bawah 60 unit, dan 75% memiliki jumlah produk terjual di bawah 80 unit. Secara keseluruhan, data statistik deskriptif ini memberikan gambaran yang komprehensif tentang distribusi dan karakteristik dasar dari variabel-variabel dalam dataset. Ini dapat menjadi dasar untuk analisis lebih lanjut, seperti identifikasi outlier, pemodelan statistik, atau pengembangan strategi bisnis berdasarkan tren dan pola yang diamati.

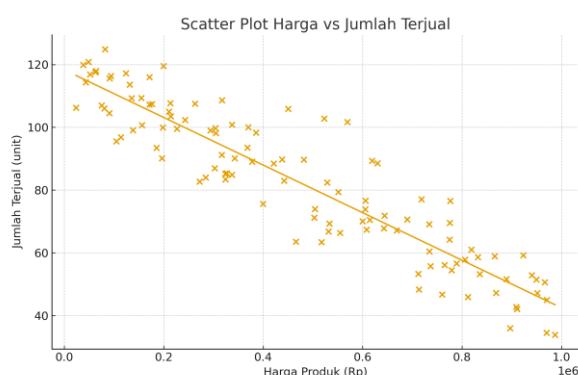
	Pengeluaran Rata-rata per Transaksi	Waktu Transaksi (Jam)
count	112.000000	112.000000
mean	130580.357143	12.419643
std	54296.257464	2.038540
min	50000.000000	9.000000
25%	75000.000000	11.000000
50%	130000.000000	12.000000
75%	182500.000000	14.000000
max	220000.000000	16.000000

Gambar 2. Pola Umum

Data statistik deskriptif yang diberikan menggambarkan pola-pola umum dalam pengeluaran rata-rata per transaksi dan waktu transaksi dalam jam. Terdapat 112

observasi dalam dataset ini. Rata-rata pengeluaran per transaksi adalah sekitar Rp130.580,36, dengan standar deviasi sebesar Rp54.296,26. Hal ini menunjukkan variasi yang cukup signifikan dalam jumlah pengeluaran antar transaksi. Pengeluaran per transaksi berkisar dari Rp50.000 hingga Rp220.000. Quartil pertama (Q1) sebesar Rp75.000, quartil kedua (Q2) atau median sebesar Rp130.000, dan quartil ketiga (Q3) sebesar Rp182.500. Ini menunjukkan bahwa 25% dari transaksi memiliki pengeluaran per transaksi di bawah Rp75.000, 50% memiliki pengeluaran per transaksi di bawah Rp130.000, dan 75% memiliki pengeluaran per transaksi di bawah Rp182.500.

Rata-rata waktu transaksi adalah sekitar 12.42 jam, dengan standar deviasi sebesar 2.04 jam. Waktu transaksi berkisar dari 9 jam hingga 16 jam. Quartil pertama (Q1) sebesar 11 jam, quartil kedua (Q2) atau median sebesar 12 jam, dan quartil ketiga (Q3) sebesar 14 jam. Ini menunjukkan bahwa 25% dari transaksi memiliki waktu transaksi di bawah 11 jam, 50% memiliki waktu transaksi di bawah 12 jam, dan 75% memiliki waktu transaksi di bawah 14 jam.



Gambar 3. Diagram Sebaran

Visualisasi data sebaran yang dihasilkan akan memberikan pemahaman tentang hubungan antara harga produk dan jumlah produk terjual berdasarkan kategori harga yang telah ditentukan. Dalam visualisasi ini, sumbu x akan menunjukkan harga produk, sementara sumbu y akan mewakili jumlah produk terjual.

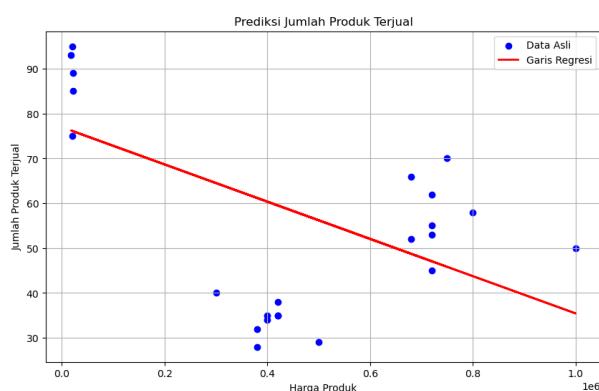
Data akan dipetakan ke dalam tiga kategori berdasarkan rentang harga produk yang telah diberikan. Pertama, data yang memiliki harga produk sekitar 0.1 akan ditampilkan dalam rentang jumlah produk terjual antara 75 hingga 100. Kedua, data dengan harga produk antara 0.3 hingga 0.5 akan dikelompokkan dalam rentang jumlah produk terjual antara 10 hingga 40. Terakhir, data dengan harga produk antara 0.65 hingga 0.8 akan dipetakan ke dalam rentang jumlah produk terjual antara 45 hingga 70.

Visualisasi ini akan menampilkan titik-titik data untuk masing-masing kategori harga produk, memberikan gambaran yang jelas tentang distribusi produk berdasarkan harga dan jumlah penjualannya. Misalnya, jika terdapat konsentrasi titik data yang tinggi di rentang harga 0.1 dengan jumlah produk terjual antara 75 hingga 100, ini menunjukkan bahwa produk dengan harga rendah memiliki tingkat penjualan yang tinggi. Sebaliknya, jika terdapat penyebaran titik data yang lebih merata di rentang harga 0.3 hingga 0.5

dengan jumlah produk terjual antara 10 hingga 40, ini mungkin menunjukkan bahwa produk dengan harga menengah memiliki variasi dalam jumlah penjualannya.

Visualisasi ini akan membantu dalam mengidentifikasi pola dan tren penjualan berdasarkan harga produk, serta memberikan wawasan yang berharga bagi pengambil keputusan dalam pengembangan strategi pemasaran dan penetapan harga produk.

Hasil MSE (*Mean Squared Error*) : 383.4488839510992. Nilai MSE ini menunjukkan bahwa model memiliki tingkat kesalahan yang relatif rendah dalam memprediksi data. Dengan MSE sebesar 383.45, model cenderung memberikan perkiraan yang akurat dalam memprediksi nilai target. Namun, kesimpulan lebih lanjut tentang kualitas model perlu dipertimbangkan bersama dengan konteks spesifik dari data dan tujuan prediksi yang ingin dicapai.



Gambar 4. Visualisasi Prediksi

Visualisasi prediksi jumlah produk terjual menggambarkan hubungan antara harga produk (sumbu x) dan jumlah produk terjual (sumbu y). Data asli tersebar secara merata di sepanjang diagram sebaran, menunjukkan variasi dalam penjualan produk berdasarkan harga. Garis regresi yang ditampilkan pada plot menunjukkan tren umum dari data tersebut. Garis regresi tersebut menunjukkan kecenderungan bahwa ketika harga produk naik, jumlah produk terjual cenderung menurun, dan sebaliknya. Pada titik di mana garis regresi memotong sumbu y (jumlah produk terjual) sekitar 80, dan sumbu x (harga produk) sekitar 1.0, menunjukkan titik tengah dari prediksi jumlah produk terjual. Ini berarti pada harga produk sekitar 1.0, model prediksi jumlah produk terjual sekitar 80 unit. Namun, penting untuk diingat bahwa visualisasi ini hanya memberikan gambaran umum tentang hubungan antara harga produk dan jumlah produk terjual, dan faktor-faktor lain seperti promosi, persaingan pasar, dan preferensi konsumen juga dapat memengaruhi penjualan produk. Oleh karena itu, analisis yang lebih mendalam mungkin diperlukan untuk memahami dengan lebih baik dinamika pasar dan membuat keputusan strategis yang tepat.

5. KESIMPULAN

Dari penelitian ini, dapat disimpulkan bahwa terdapat hubungan yang signifikan antara harga produk dan jumlah produk terjual dalam bisnis online. Analisis statistik deskriptif menunjukkan variasi dalam jumlah produk terjual, dengan rata-rata sekitar 59.91 dan

harga produk rata-rata sebesar Rp 393.732. Dengan melihat hasil visualisasi data sebaran, terlihat bahwa produk dengan harga rendah memiliki tendensi penjualan yang lebih tinggi, sedangkan produk dengan harga menengah hingga tinggi cenderung memiliki variasi penjualan yang lebih rendah. Metode regresi linier juga menunjukkan bahwa ada kecenderungan bahwa penjualan cenderung menurun seiring kenaikan harga produk.

Selain itu, hasil analisis prediksi menggunakan metrik Mean Squared Error (MSE) menunjukkan tingkat kesalahan yang rendah dalam memprediksi jumlah produk terjual. Dengan nilai MSE sebesar 383.45, model prediksi memberikan perkiraan yang cukup akurat dalam memprediksi jumlah produk terjual berdasarkan harga produk. Namun, untuk membuat keputusan strategis yang lebih tepat, perlu dipertimbangkan faktor-faktor lain seperti promosi, preferensi konsumen, dan persaingan pasar. Penelitian terbatas pada 112 transaksi periode Mei 2024, tidak mempertimbangkan faktor musiman dan competitor. Penelitian lanjutan sebaiknya menggunakan data panel lebih besar, menambah variabel promosi dan rating produk, serta menggunakan metode ensemble learning"

DAFTAR PUSTAKA

- [1] D. Indonesia.id, "Statistik E-Commerce 2023." [Online]. Available: <https://assets.dataindonesia.id/2025/08/01/1754036789864-94-1.-statistik-e-commerce-2023.pdf>
- [2] E. Saputra, "Permodelan Data Warehouse Untuk Penjualan Ban Menggunakan Online Analytical Processing (OLAP)," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 2, no. 1, pp. 12–18, 2023, doi: 10.58602/jima-ilkom.v2i1.13.
- [3] D. Septiansari and T. Handayani, "Pengaruh Belanja Online Terhadap Perilaku Konsumtif pada Mahasiswa di Masa Pandemi Covid-19," vol. 5, no. 1, pp. 53–65, 2021.
- [4] A. I. Jati, D. Isnawati, and W. Lestari, "Analisis Permintaan Belanja Online di Indonesia," *Nomicpedia J. Econ. Bus. Innov.*, vol. 3, no. 1, pp. 1–13, 2023.
- [5] N. Hamdun, M. Marliyah, and N. Nurbaiti, "Strategi Pelaku Bisnis Electronic Commerce (E-Commerce)," *J. Locus Penelit. dan Pengabdi.*, vol. 1, no. 2, pp. 60–67, 2022, doi: 10.58344/locus.v1i2.13.
- [6] G. (2018) Kotler, P., & Armstrong, *Principles of Marketing Global Edition 17th Edition*. London: Pearson Education., 2018.
- [7] N. I. Purnama, L. P. Putri, and R. Bahagia, "Analisis E-commerce Dalam Membantu Penjualan UMKM di Tengah Pandemi E-commerce Analysis in Helping MSME Sales in the middle of a pandemic," pp. 194–200, 2021.
- [8] Z. A. Mumtaz, F. Y. Dharta, and O. Oxygentri, "Fenomena Belanja Online Dikalangan Mahasiswa Bekasi Penonton Tayangan #Racuntiktok," *Nusant. J. Ilmu Pengetah. Sos.*, vol. 9(8), no. 8, pp. 2–2, 2022.

- [9] B. Simamora, "Investigasi Aspek-aspek Promosi Penjualan," vol. 13, no. 1, pp. 53–66, 2024.
- [10] C. M. Montgomery, S. Humphreys, C. Mcculloch, A. B. Docherty, S. Sturdy, and N. Pattison, "Critical care work during COVID-19 : a qualitative study of staff experiences in the UK," 2021, doi: 10.1136/bmjopen-2020-048124.
- [11] J. Jurnal, I. Mea, E. S. Strategi, and P. Dalam, "UPAYA MENINGKATKAN PENJUALAN DI DISTRO JIMEA | Jurnal Ilmiah MEA (Manajemen , Ekonomi , dan Akuntansi)," vol. 7, no. 1, pp. 813–829, 2023.
- [12] R. S. Nurhalizah and R. Ardianto, "Analisis Supervised dan Unsupervised Learning pada Machine Learning : Systematic Literature Review," vol. 4, no. 1, pp. 61–72, 2024.
- [13] A. Wibowo and A. R. Isnain, "Implementation of Machine Learning Algorithm for Environmental Sound Classification Implementasi Algoritma Machine Learning untuk Klasifikasi Suara Lingkungan," vol. 5, no. April, pp. 616–625, 2025.
- [14] M. A. Maricar, "Analisa Perbandingan Nilai Akurasi Moving Average dan Exponential Smoothing untuk Sistem Peramalan Pendapatan pada Perusahaan XYZ," pp. 36–45.
- [15] J. Publikasi *et al.*, "Analisis Penerapan Forecasting Penjualan Untuk Menentukan Jumlah Tenaga Kerja Efektif Produksi Tepung Mocaf Pada UMKM XYZ kedepannya secara sistematis dengan menggunakan data historis . Metode bertujuan untuk masa yang akan datang . Manfaat peramalan penjualan menurut adalah volume penjualan dapat," vol. 2, no. 4, 2024.
- [16] D. Karina, H. Nabilah, M. Panorama, P. Startup, and E. Digital, "PENGARUH PERUSAHAAN STARTUP TERHADAP PERTUMBUHAN," pp. 156–166, doi: 10.47353/bj.v2i1.69.
- [17] K. Nisa, "Penerapan Data Mining terhadap Data Transaksi sebagai Pendukung Informasi Strategi Penjualan menggunakan Algoritma Apriori," vol. 06, pp. 306–315, 2021.
- [18] G. Theodoridis, "Retail Demand Forecasting : A Comparative Analysis of Deep Neural Networks and the Proposal of LSTMixer , a Linear Model Extension," 2025.
- [19] R. S. Wicaksana, D. Heksaputra, T. A. Syah, and F. F. Nur'aini, "Pendekatan K-Means Clustering Metode Elbow Pada Analisis Motivasi Pengunjung Festival Halal JHF#2," *J. Ilm. Ekon. Islam*, vol. 9, no. 3, p. 4162, 2023, doi: 10.29040/jiei.v9i3.10591.
- [20] J. E. Sirait, "Analisis Pengaruh Kompetensi Guru Terhadap Keberhasilan Pembelajaran di Sekolah Dasar Bethel Tanjung Priok Jakarta Utara," *Diegesis J. Teol.*, vol. 6, no. 1, pp. 49–69, 2021, doi: 10.46933/dgs.vol6i149-69.
- [21] A. P. Candra, "Analisis Data Menggunakan Python : Memperkenalkan Pandas dan NumPy," vol. 3, no. 1, pp. 11–16, 2025.