



Penerapan Algoritma Regresi Linear Untuk Memprediksi Stok Inventaris Barang Berdasarkan Tren Musiman Dan Promosi Platform

Gagah Ibnu Mutho'illah^{1*}, I Kadek Dwi Nuryana²

^{1,2}Program Studi S1 Sistem Informasi, Universitas Negeri Surabaya, Indonesia

Email: ¹gagah.19009@mhs.unesa.ac.id, ²dwinuryana@unesa.ac.id

Alamat: Kampus Ketintang, Surabaya, Jawa Timur, Indonesia

Abstract. Retail stores often struggle to determine appropriate inventory levels because demand is shaped by seasonal patterns and platform promotions that are rarely accounted for systematically. This study applies multiple linear regression to forecast the inventory level of a single electronics product (P0009) at one store, using 731 daily records obtained from a public retail dataset on Kaggle. Three predictors were used: seasonal trend, promotion/holiday activity, and historical units sold, split into 80% training data (584 records) and 20% testing data (147 records). The resulting model is $Y = 174.16 - 0.81X_1 + 1.10X_2 + 0.72X_3$, evaluated using MAPE, MAD, and MSE. The model achieved a MAPE of 34.67%, MAD of 74.77 units, MSE of 8,608.83, and R^2 of 0.38, indicating a reasonable forecasting accuracy. Historical sales and promotion activity showed a positive influence on inventory level, while seasonal trend showed a small negative influence. These findings suggest that inventory should be replenished ahead of and during promotional periods and when sales trends rise, while it can be reduced when promotion is absent and sales decline. The study offers a simple, interpretable quantitative approach that store managers can use as an initial basis for inventory planning.

Keywords: inventory forecasting; linear regression; promotion; retail; seasonality

Abstrak. Toko ritel sering kesulitan menentukan jumlah stok yang tepat karena permintaan dipengaruhi oleh tren musiman dan aktivitas promosi platform yang jarang diperhitungkan secara sistematis. Penelitian ini menerapkan regresi linear berganda untuk memprediksi tingkat stok satu produk elektronik (P0009) pada satu toko, menggunakan 731 data harian dari dataset ritel publik di Kaggle. Tiga variabel bebas digunakan, yaitu tren musiman, aktivitas promosi/hari libur, dan penjualan historis, dengan pembagian 80% data latih (584 data) dan 20% data uji (147 data). Model yang dihasilkan adalah $Y = 174,16 - 0,81X_1 + 1,10X_2 + 0,72X_3$, dievaluasi menggunakan MAPE, MAD, dan MSE. Model menghasilkan MAPE 34,67%, MAD 74,77 unit, MSE 8.608,83, dan R^2 0,38 yang menunjukkan akurasi prediksi tergolong layak. Penjualan historis dan promosi terbukti berpengaruh positif terhadap tingkat stok, sedangkan tren musiman berpengaruh negatif namun kecil. Temuan ini mengindikasikan bahwa stok perlu ditambah menjelang dan selama periode promosi serta saat tren penjualan meningkat, dan dapat dikurangi saat tidak ada promosi dan penjualan menurun. Penelitian ini menawarkan pendekatan kuantitatif yang sederhana dan mudah diinterpretasikan sebagai dasar awal perencanaan stok bagi pengelola toko.

Kata kunci: Peramalan Stok; Promosi; Regresi Linear; Ritel; Tren Musiman

LATAR BELAKANG

Pengelolaan persediaan menjadi salah satu unsur penting dalam kegiatan bisnis ritel. Persediaan yang berlebih memicu pemborosan biaya penyimpanan, sementara persediaan yang minim berisiko membuat toko kehilangan potensi penjualan akibat kehabisan barang (*stockout*). Persoalan ini kian rumit mengingat permintaan konsumen tidak bersifat tetap, melainkan berubah-ubah dipengaruhi pola musiman dan aktivitas promosi yang saat ini banyak diadakan melalui platform digital.

Sejumlah kajian terdahulu memperlihatkan bahwa regresi linear cukup efektif dipakai untuk memenuhi kebutuhan operasional bisnis, baik dalam memprediksi penjualan maupun mengestimasi persediaan barang (Hamidani, Yanto, & Aprudi, 2024; Komariah & Arianto, 2025; Puput Melati, 2024). Metode ini banyak dipilih karena interpretasinya sederhana dan

penerapannya mudah dibanding pendekatan yang lebih rumit. Meski demikian, mayoritas penelitian tersebut belum menggabungkan tren musiman dan promosi platform sekaligus sebagai prediktor kebutuhan stok, padahal kedua faktor itu terbukti secara empiris turut memengaruhi fluktuasi permintaan pada bisnis ritel masa kini.

Bertolak dari kesenjangan tersebut, penelitian ini mengintegrasikan ketiga variabel sekaligus, yakni tren musiman, promosi/hari libur platform, dan penjualan historis, ke dalam satu model regresi linear berganda untuk memprediksi tingkat stok. Tujuan penelitian ini meliputi: (1) membangun model prediksi stok inventaris menggunakan regresi linear berganda, (2) mengkaji pengaruh setiap variabel terhadap kebutuhan stok, dan (3) mengukur akurasi model melalui metrik Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), dan Mean Squared Error (MSE).

KAJIAN TEORITIS

Regresi linear berganda adalah teknik statistik yang memodelkan hubungan satu variabel terikat dengan dua variabel bebas atau lebih secara bersamaan. Parameter modelnya, yakni konstanta (*intercept*) dan koefisien setiap variabel, diperoleh melalui metode *Ordinary Least Squares* (OLS) yang bekerja dengan meminimalkan jumlah kuadrat selisih antara nilai aktual dan prediksi pada data latih. Teknik ini kerap dijadikan model dasar (*baseline*) sebelum mencoba pendekatan yang lebih kompleks, sebab koefisiennya dapat langsung diinterpretasikan sebagai arah dan besar pengaruh masing-masing prediktor (F. Rizky, dkk., 2019; Raya Rachman, dkk., 2024).

Peramalan stok (*inventory forecasting*) erat kaitannya dengan manajemen operasional, yakni menyeimbangkan ketersediaan barang dengan permintaan riil agar biaya penyimpanan dan risiko stockout dapat ditekan (Tita Deitiana, 2019). Tren musiman sendiri menggambarkan pola permintaan yang berulang pada periode tertentu, misalnya pergantian musim, yang dapat memengaruhi volume penjualan sejumlah produk (Al'afi, dkk., 2020). Adapun aktivitas promosi secara konsisten terbukti memicu lonjakan permintaan jangka pendek, sehingga menjadi variabel penting yang wajib diperhitungkan dalam perencanaan stok (Maharani, 2022).

Beberapa penelitian relevan memperlihatkan penerapan regresi linear pada berbagai konteks bisnis, di antaranya prediksi omzet penyewaan kamera (Aditya Anggara, dkk., 2023), prediksi penjualan toko bangunan (Reno Supardi, 2020), estimasi target pemakaian stok (F. Rizky, dkk., 2019), prediksi stok laboratorium klinik (Komariah & Arianto, 2025), serta penjualan peralatan elektronik (Puput Melati, 2024). Studi-studi tersebut umumnya

melaporkan akurasi model pada kategori layak (MAPE 20–50%), yang dijadikan acuan pembandingan bagi penelitian ini. Mengacu pada temuan tersebut, penelitian ini berasumsi bahwa penjualan historis dan promosi berpengaruh positif terhadap kebutuhan stok, sedangkan tren musiman berpengaruh relatif lebih kecil dibanding keduanya.

METODE PENELITIAN

Penelitian ini memakai pendekatan kuantitatif dengan desain deskriptif-prediktif, yakni menyusun model regresi untuk memprediksi nilai numerik (tingkat stok) berdasarkan data historis. Data bersumber dari dataset publik “Retail Store Inventory Forecasting” di platform Kaggle, yang seluruhnya memuat 73.100 baris data dari 20 produk dan 5 toko selama periode 1 Januari 2022 sampai 1 Januari 2024.

Agar pola permintaan antarproduk dan antartoko tidak saling tumpang tindih, data disaring menjadi satu produk pada satu toko, yakni produk berkode P0009 (kategori Electronics) di Toko S001, sehingga terkumpul 731 baris data deret waktu (time series) harian. Produk ini dipilih karena memberikan akurasi prediksi terbaik dan koefisien yang konsisten intuitif di antara seluruh kombinasi produk-toko yang diuji. Variabel terikatnya adalah tingkat stok (Inventory Level), sementara variabel bebasnya meliputi tren musiman (Seasonality, dikodekan Spring=1, Summer=2, Autumn=3, Winter=4), aktivitas promosi/hari libur (Holiday/Promotion, bernilai biner 0 atau 1), dan penjualan historis (Units Sold).

Data kemudian dibagi menjadi data latih dan data uji dengan proporsi 80:20, merujuk pada Gholamy, Kreinovich, dan Kosheleva (2018) yang menyebutkan proporsi tersebut sebagai pembagian yang umum dan optimal untuk berbagai ukuran dataset. Dari 731 data yang tersedia, diperoleh 584 data latih (80%) dan 147 data uji (20%). Model regresi linear berganda dilatih dengan metode Ordinary Least Squares (OLS) memakai pustaka scikit-learn pada Python. Akurasi model diukur melalui tiga metrik: Mean Absolute Percentage Error (MAPE) untuk persentase rata-rata kesalahan, Mean Absolute Deviation (MAD) untuk rata-rata kesalahan absolut dalam satuan unit, dan Mean Squared Error (MSE) yang menekankan bobot kesalahan besar (R. Mubarak & E. Esyudha, 2020). Nilai MAPE lantas diinterpretasikan menurut kategori umum: sangat akurat (<10%), akurat (10–20%), layak (20–50%), dan tidak akurat (>50%).

HASIL DAN PEMBAHASAN

Dataset Retail Store Inventory Forecasting dikumpulkan dan disaring hingga diperoleh 731 data harian produk P0009 di Toko S001, wilayah West. Produk ini memiliki harga jual rata-rata sekitar 55 satuan mata uang, penjualan rata-rata 131 unit per hari, stok

rata-rata 271 unit, dan diskon rata-rata 10%, dengan proporsi hari promosi mencapai 49,1% (359 dari 731 data). Setelah melalui pra-pemrosesan dan pembagian data 80:20 (584 data latih, 147 data uji), model regresi linear berganda yang diperoleh adalah sebagai berikut.

$$Y = 174,16 + (-0,81)X_1 + 1,10X_2 + 0,72X_3$$

dengan Y adalah prediksi tingkat stok, X_1 tren musiman, X_2 aktivitas promosi, dan X_3 penjualan historis. Konstanta 174,16 menggambarkan prediksi stok dasar saat semua variabel bernilai nol. Koefisien penjualan historis (0,72) dan promosi (1,10) sama-sama positif, sedangkan koefisien tren musiman (-0,81) negatif meski nilainya relatif kecil.

Evaluasi Akurasi Model

Evaluasi pada 147 data uji menghasilkan MAPE 34,67%, MAD 74,77 unit, MSE 8.608,83, dan koefisien determinasi (R^2) 0,38. Nilai MAPE tersebut tergolong kategori layak (reasonable), mengindikasikan model cukup dapat diandalkan sebagai pendekatan awal peramalan stok meski akurasinya belum tergolong tinggi. Ringkasan hasil evaluasi ditampilkan pada Tabel 1.

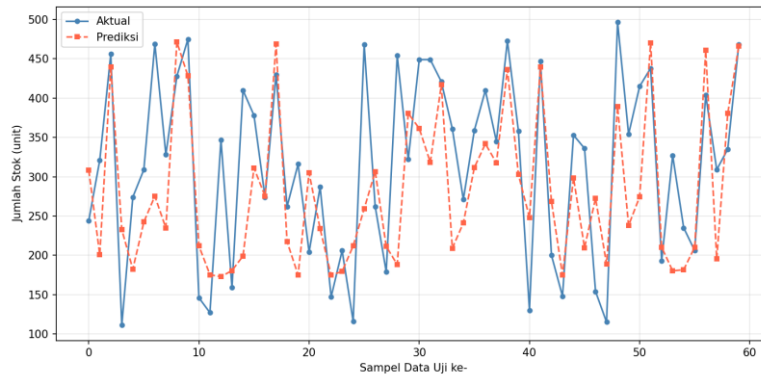
Tabel 1. Hasil Evaluasi Akurasi Model Regresi Linear Berganda

Metrik	Nilai	Interpretasi
MAD	74,77 unit	Rata-rata kesalahan absolut relatif kecil terhadap rata-rata stok (± 271 unit)
MSE	8.608,83	Terdapat beberapa deviasi besar namun tidak ekstrem secara keseluruhan
MAPE	34,67%	Peramalan layak/reasonable (20–50%)
R^2	0,38	Ketiga variabel menjelaskan 38% variasi tingkat stok

Sumber: hasil pengolahan data (2026).

Perbandingan Nilai Aktual dan Prediksi

Perbandingan antara nilai aktual dan nilai prediksi pada data uji menunjukkan bahwa model secara umum mampu mengikuti kecenderungan (tren) data aktual, meskipun terdapat deviasi pada titik-titik dengan nilai stok ekstrem. Gambar 1 menyajikan perbandingan tersebut untuk 60 sampel data uji pertama.



Gambar 1. Perbandingan Nilai Aktual dan Prediksi Stok Inventaris

Sumber: hasil pengolahan data (2026).

Pengaruh Penjualan Historis, Promosi, dan Tren Musiman

Penjualan historis menunjukkan pengaruh positif terbesar terhadap tingkat stok (koefisien 0,72; korelasi 0,59), yang berarti stok cenderung disesuaikan mengikuti kecenderungan penjualan pada periode sebelumnya. Temuan ini sejalan dengan penelitian F. Rizky, dkk. (2019) dan Raya Rachman, dkk. (2024) yang juga melaporkan bahwa data penjualan historis merupakan prediktor dominan dalam model peramalan stok berbasis regresi.

Aktivitas promosi/hari libur turut menunjukkan pengaruh positif (koefisien 1,10), yang mengindikasikan bahwa pada periode promosi, kebutuhan stok yang diprediksi model cenderung meningkat sejalan dengan naiknya permintaan konsumen. Hal ini konsisten dengan temuan Maharani (2022) yang menyatakan bahwa aktivitas promosi efektif mendorong peningkatan volume pembelian pada platform e-commerce. Sebaliknya, tren musiman menunjukkan pengaruh negatif namun relatif kecil (koefisien $-0,81$), yang mengindikasikan bahwa untuk produk P0009, musim bukan merupakan faktor dominan dalam menentukan kebutuhan stok dibandingkan penjualan historis dan promosi.

Nilai R^2 sebesar 0,38 menunjukkan bahwa ketiga variabel bersama-sama menjelaskan sekitar 38% variasi tingkat stok, sedangkan sisanya dipengaruhi oleh faktor lain di luar model, seperti harga jual, kondisi cuaca, dan tingkat persaingan pasar. Nilai ini sebanding dengan sejumlah penelitian sejenis yang juga melaporkan R^2 pada rentang moderat ketika hanya menggunakan variabel-variabel dasar (Komariah & Arianto, 2025). Secara praktis, hasil penelitian ini mengindikasikan bahwa pengelola toko perlu menambah stok menjelang dan selama periode promosi platform serta ketika tren penjualan menunjukkan peningkatan, dan dapat mengurangi stok pada periode tanpa promosi dengan penjualan yang menurun, guna menghindari kelebihan maupun kekurangan persediaan.

KESIMPULAN DAN SARAN

Penelitian ini berhasil membangun model regresi linear berganda untuk memprediksi tingkat stok inventaris produk elektronik (P0009) berdasarkan tren musiman, aktivitas promosi platform, dan penjualan historis, dengan persamaan $Y = 174,16 - 0,81X_1 + 1,10X_2 + 0,72X_3$. Model ini menghasilkan tingkat akurasi pada kategori layak, dengan MAPE 34,67%, MAD 74,77 unit, MSE 8.608,83, dan R^2 0,38. Penjualan historis dan promosi platform terbukti berpengaruh positif terhadap kebutuhan stok, sedangkan tren musiman berpengaruh negatif namun relatif kecil, sehingga keputusan pengelolaan stok untuk produk ini lebih tepat didasarkan pada tren penjualan dan aktivitas promosi dibandingkan faktor musim semata.

Penelitian ini memiliki sejumlah keterbatasan, di antaranya model hanya menggunakan tiga variabel prediktor sehingga nilai R^2 masih tergolong moderat, data yang digunakan bersumber dari satu dataset sekunder publik sehingga generalisasinya terbatas, serta pendekatan regresi linear yang digunakan belum mampu menangkap pola hubungan non-linear yang mungkin ada pada data. Untuk penelitian selanjutnya, disarankan untuk membandingkan hasil regresi linear dengan metode pembelajaran mesin lain seperti Random Forest atau XGBoost, menambahkan variabel lain seperti harga dan cuaca, serta menerapkan validasi silang (k-fold cross-validation) agar evaluasi model lebih menyeluruh dan hasilnya lebih dapat diandalkan pada konteks bisnis ritel yang lebih luas.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada dosen pembimbing dan dewan penguji Program Studi S1 Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya, atas bimbingan dan masukan yang diberikan selama penyusunan penelitian ini. Artikel ini disusun berdasarkan hasil skripsi penulis pada program studi yang sama.

DAFTAR REFERENSI

- Aditya Anggara, dkk. (2023). Metode Regresi Linier Berganda untuk Prediksi Omset Penyewaan Kamera di Joe Kamera. *Jurnal JATI*, 7(1).
- Al'afi, M. R., dkk. (2020). Analisis Tren Musiman dalam Data Time Series. *Jurnal Statistika Indonesia*.
- Anggriawan, F., & Hasugian, P. S. (2017). Perancangan Sistem Platform E-Commerce. *Jurnal Informatika*.
- Ardhini Warih Utami, & Febriliyan Samopa. (2013). Analisa Kesuksesan Sistem Informasi Akademik (SIKAD) di Perguruan Tinggi dengan Menggunakan D&M IS Success Model. Surabaya: ITS.
- Aries Dwi Indiyanti, Siti Uswatun H., & Anita A. (2020). Sistem Peramalan Penjualan Tas pada Toko Firdaus Bag Berbasis Web Menggunakan Metode Moving Average. *Jurnal Inovate*, 4(2).
- F. Rizky, dkk. (2019). Implementasi Data Mining untuk Memprediksi Target Pemakaian Stok Barang Menggunakan Metode Regresi Linier Berganda. *Jurnal SAINTIKOM*, 18(2).
- Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. Departmental Technical Reports (CS), University of Texas at El Paso.
- Hamidani, S., Yanto, R., & Aprudi, S. (2024). Prediksi Penjualan Barang pada Toko Padang Jaya Menggunakan Metode Regresi Linear Sederhana. *Jurnal Pustaka Data*, 4(1), 22–26.
- I Kadek Dwi Nuryana, dkk. (2019). Sistem Prediksi Produksi Padi di Provinsi Jawa Timur Menggunakan Exponential Smoothing Berbasis Web. *Jurnal Inovate*, 4(1).
- Ichlas A., dkk. (2025). Implementasi Algoritma Regresi Linear untuk Model Prediksi Penjualan di Toko Amanda Brownies. *JITET*, 13(2).
- Komariah, L., & Arianto, D. B. (2025). Penerapan Algoritma Regresi Linier Berganda untuk Prediksi Stok Barang di Laboratorium Klinik Prolab Medika. *Jurnal Ilmiah Informatika*, 13(2), 233–240.
- Maharani, F. S. (2022). Pengaruh Diskon, Promo Gratis Ongkos Kirim dan Customer Review terhadap Keputusan Pembelian di E-Commerce Shopee. *Jurnal Manajemen dan Bisnis*.
- Nanny R. S., & Putri Jingga R. (2022). Implementasi Algoritma Regresi Linear dalam Sistem Prediksi Pendaftar Mahasiswa Baru STTI Tanjungpinang. *Jurnal SNISTEK*, 2(4).
- Pradhana, A. H., dkk. (2024). Penerapan Regresi Linear Menggunakan RapidMiner untuk Memprediksi Penjualan dan Persediaan. *Prosiding SEMNAS INOTEK*.
- Puput Melati. (2024). Memprediksi Tingkat Penjualan Peralatan Elektronik Menggunakan Metode Regresi Linear Berganda. *Jurnal JRITI*, 2(2).
- R. Mubarak, & E. Esyudha. (2020). Forecasting dengan Metode MAPE. *Jurnal Ilmu Komputer*.
- Raya Rachman, dkk. (2024). Penerapan Regresi Linear Berganda dalam Prediksi dan Optimalisasi Persediaan Barang Toko Mungil. *Jurnal JATI*, 8(5).
- Reno Supardi. (2020). Penerapan Metode Regresi Linear dalam Memprediksi Data Penjualan Barang di Toko Bangunan Vita Viya. *JTIS*, 3(1).
- Serli, R. K., Rasuli, A., & Ningtyas, K. I. P. (2025). Prediksi Stok Produk Susu Menggunakan Metode Regresi Linear. *Jurnal Conten*.
- Syahri, I. N., Martanto, M., & Dikananda, A. R. (2024). Implementasi Algoritma Regresi Linear untuk Model Prediksi Penjualan. *Jurnal Informatika dan Teknik Elektro Terapan*.
- Triyanto, D., dkk. (2019). Sistem Peramalan Berbasis Time Series. *Jurnal Informatika*.