

Analisis Eksploratif Proporsi dan Perbandingan Harga Produk Farmasi yang Dihentikan versus Produk Aktif

Exploratory Analysis of the Proportion and Price Comparison of Discontinued versus Active Pharmaceutical Products

Muhammad Rijal^{1,*}; Nurkhalik Wahdania Asbara²; Pertiwi Ishak³; Firman Aziz⁴

^{1,2} Institut Teknologi dan Bisnis Nobel Indonesia 90221, Indonesia

^{3,4} Universitas Pancasakti, Makassar 90121, Indonesia

¹ rijal2303@gmail.com; ² khalikwahdania@nobel.ac.id; ³ pertiwi.ishak@unpacti.ac.id; ⁴ firman.aziz@unpacti.ac.id

* Corresponding author

Abstrak

Penghentian pemasaran produk obat merupakan fenomena global yang semakin mengkhawatirkan dalam dinamika rantai pasok farmasi. Penelitian terkait pengelolaan obat di fasilitas pelayanan kesehatan selama ini lebih banyak terbatas pada aspek manajemen internal, sementara studi di tingkat makro lebih banyak langsung mengembangkan model prediktif tanpa pemetaan eksploratif terlebih dahulu. Oleh karena itu, masih terdapat kesenjangan pengetahuan mengenai karakteristik dasar produk yang dihentikan dari perspektif analisis data pasar secara makro, yang justru menjadi fondasi penting sebelum pengembangan model prediksi. Penelitian ini bertujuan untuk (1) mengetahui proporsi produk obat yang dihentikan secara keseluruhan di pasar farmasi, dan (2) menganalisis perbedaan harga antara produk yang dihentikan dan produk yang masih aktif di pasaran. Penelitian menggunakan rancangan deskriptif kuantitatif dengan pendekatan berbasis data pada 253.973 produk farmasi dari basis data publik. Analisis dilakukan melalui tahap praproses data, analisis deskriptif, uji normalitas Shapiro-Wilk, uji komparatif Mann-Whitney, serta analisis kelompok berdasarkan bentuk sediaan, kelas terapi, pabrikasi, dan jumlah bahan aktif. Hasil penelitian menunjukkan proporsi produk yang dihentikan sebesar 3,11% (7.905 produk). Terdapat perbedaan harga yang signifikan secara statistik antara produk dihentikan dan produk aktif ($p\text{-value} = 0,0005$). Analisis kelompok mengidentifikasi bahwa bentuk sediaan inhaler (20,97%), spray (9,56%), dan respules (9,02%) memiliki proporsi penghentian tertinggi, sementara kelas terapi diuretik (7,79%) menempati peringkat pertama. Variasi antar pabrikasi sangat signifikan dengan salah satu pabrikasi mencapai tingkat penghentian 100%. Kebaruan penelitian ini terletak pada kerangka analisis data terstruktur yang menggabungkan proporsi, uji komparatif, dan analisis kelompok secara simultan pada skala data besar, yang sebelumnya belum diterapkan secara sistematis untuk memetakan risiko penghentian produk farmasi. Penelitian ini menyimpulkan bahwa penghentian produk farmasi terjadi pada sebagian kecil produk namun tetap signifikan, dengan karakteristik harga dan atribut produk tertentu yang berisiko lebih tinggi. Temuan ini memberikan dasar empiris untuk kebijakan ketahanan obat dan landasan awal bagi pengembangan model prediktif di masa mendatang.

Kata Kunci: penghentian produk obat, analisis data farmasi, perbedaan harga, data mining, ketahanan obat

Abstract

Drug product discontinuation is a growing global concern in pharmaceutical supply chain dynamics. Previous research on drug management at healthcare facilities has largely been confined to internal management aspects, leaving a knowledge gap regarding the characteristics of discontinued products from a macro market data analytics perspective. This study aims to (1) determine the overall proportion of discontinued pharmaceutical products in the market, and (2) analyze price differences between discontinued and active products. A descriptive quantitative design with a data-driven approach was employed, analyzing 253,973 pharmaceutical products from a publicly available database. The analysis involved data preprocessing, descriptive statistics, Shapiro-Wilk normality test, Mann-Whitney comparative test, and group analysis based on dosage form, therapeutic class, manufacturer, and number of active ingredients. Results showed that discontinued products accounted for 3.11% (7,905 products). A statistically significant price difference was found between discontinued and active products ($p\text{-value} = 0.0005$). Group analysis identified that inhalers (20.97%), sprays (9.56%), and respules (9.02%) had the highest discontinuation rates among dosage forms, while diuretics (7.79%) ranked first among therapeutic classes. Substantial variation was observed across manufacturers, with one manufacturer reaching a 100% discontinuation rate. This study concludes that pharmaceutical product discontinuation affects a small but significant proportion of products, with specific price characteristics and product attributes carrying higher risks. These findings provide an empirical foundation for drug security policies and the development of future predictive models.

Keywords: drug product discontinuation, pharmaceutical data analysis, price difference, data mining, drug security

Pendahuluan

Penghentian pemasaran produk obat (drug discontinuation) merupakan fenomena global yang semakin mengkhawatirkan dalam dinamika rantai pasok farmasi. Berdasarkan data dari berbagai negara, jumlah produk obat yang dihentikan pemasarannya menunjukkan tren peningkatan yang signifikan [1] melaporkan bahwa jumlah obat yang dihentikan di Jepang meningkat 1,4 kali lipat pada tahun 2021 dibandingkan tahun 2017, dengan 78,6% di antaranya merupakan obat generik. Fenomena ini tidak hanya berdampak pada ketersediaan produk di pasar, tetapi juga mencerminkan pola persaingan industri, tekanan ekonomi, serta dinamika rantai pasok yang kompleks.

Kemajuan ilmu pengetahuan dan teknologi informasi menawarkan pendekatan baru dalam menganalisis fenomena penghentian produk obat. Salah satu pendekatan yang dapat digunakan adalah data mining, yang dapat menjadi referensi yang bermanfaat untuk menggali pola dan karakteristik dari produk-produk yang dihentikan. Dalam konteks analisis produk obat, teknik analisis data dapat digunakan untuk mengidentifikasi karakteristik produk yang berisiko dihentikan, seperti harga, bentuk sediaan, kelas terapi, dan jumlah pemasok [1]. Penelitian [1] berhasil mengungkap profil obat generik yang dihentikan di Jepang melalui analisis data kuantitatif, dan menemukan bahwa obat generik yang dihentikan memiliki karakteristik spesifik, yaitu harga yang lebih rendah dan jumlah pemasok yang lebih banyak dibandingkan produk yang masih aktif. Produk dengan harga di bawah 10 yen menyumbang 29,2% dari total produk yang dihentikan, sementara pada produk aktif hanya 15,0%. Demikian pula, produk yang dipasok oleh empat atau lebih perusahaan menyumbang 67,6% dari produk dihentikan, jauh lebih tinggi dibandingkan produk aktif yang hanya 38,4%. Temuan ini mengindikasikan bahwa tekanan harga dan persaingan pasar yang ketat berperan penting dalam keputusan penghentian produk.

Penelitian lain [2] menyoroti pemanfaatan big data dan teknik data mining dalam farmakovigilans untuk mengidentifikasi berbagai sinyal efek samping obat dari sistem pelaporan spontan (SRS), rekam medis elektronik (EHR), dan media sosial. Sementara itu, [3] menekankan bahwa pemanfaatan big data dan kecerdasan buatan tidak hanya berperan dalam deteksi efek samping, tetapi juga dapat digunakan untuk memantau pola penggunaan obat dan mengidentifikasi potensi gangguan pasokan sejak dini. Kedua studi ini menunjukkan bahwa pendekatan berbasis data dapat memberikan wawasan yang lebih dalam memetakan risiko dan karakteristik produk farmasi, termasuk produk yang berpotensi dihentikan. Pemanfaatan pendekatan berbasis data dan kecerdasan buatan dalam bidang kesehatan telah berkembang tidak hanya untuk analisis data farmasi, tetapi juga untuk mendukung diagnosis dan prediksi kondisi klinis. Misalnya [4] menerapkan metode machine learning untuk memprediksi depresi sebagai salah satu inovasi terkini dalam kesehatan mental, yang menunjukkan bahwa algoritma klasifikasi dapat dimanfaatkan untuk mendeteksi risiko gangguan psikologis secara dini. Sementara itu [5] mengoptimalkan pemanfaatan Internet of Medical Things (IoMT) untuk pelayanan kesehatan, dengan mengintegrasikan sensor-sensor akurat seperti suhu, denyut jantung, dan tekanan guna meningkatkan efisiensi sistem informasi kesehatan. Kedua studi ini memperkuat argumen bahwa pendekatan berbasis data dan teknologi informasi baik melalui machine learning maupun IoMT telah terbukti memberikan kontribusi nyata dalam meningkatkan kualitas layanan kesehatan, sekaligus membuka peluang untuk penerapan serupa dalam analisis rantai pasok farmasi, termasuk pemetaan risiko penghentian produk obat.

Di Indonesia, penelitian terkait pengelolaan obat selama ini lebih banyak berfokus pada evaluasi di tingkat fasilitas pelayanan kesehatan. [6] Mengevaluasi pengelolaan sediaan farmasi di Puskesmas Bara-Baraya Kota Makassar dan menemukan ketidaksesuaian pada aspek pengendalian persediaan, pendistribusian, serta pemusnahan dan penarikan obat. [7] Dalam tinjauan naratifnya mengidentifikasi bahwa pengelolaan obat di puskesmas di Indonesia masih menghadapi berbagai kendala, mulai dari perencanaan yang belum berbasis data epidemiologi hingga distribusi yang tidak merata. [8] Mengevaluasi pengelolaan obat rusak atau kedaluwarsa di Rumah Sakit Tk. II R. W. Mongisidi Manado dan menemukan kesesuaian sebesar 76% dengan pedoman yang ditetapkan, dengan faktor-faktor penyebab seperti perubahan regulasi BPJS, perubahan kebijakan pimpinan rumah sakit, pengadaan berlebih, perubahan pola konsumsi, dan penurunan kunjungan pasien. [9] Mengevaluasi pengelolaan obat yang mendekati kedaluwarsa di Instalasi Farmasi RS X Ngawi dan menemukan bahwa dari 60 jenis obat yang mendekati kedaluwarsa, 66,7% dapat diretur, 28,3% masih dapat digunakan, dan hanya 5,0% yang dimusnahkan. [10] Menganalisis pengelolaan obat di RSUD Budi Setia Kabupaten Minahasa menggunakan indikator efisiensi dan metode Hanlon, serta menemukan ketidaksesuaian pada beberapa indikator seperti kesesuaian item obat dengan formularium nasional (65,28%) dan alokasi dana pengadaan (14,74%).

Studi-studi tersebut memberikan gambaran yang komprehensif mengenai tantangan pengelolaan obat di fasilitas pelayanan kesehatan. Namun, kelima penelitian tersebut masih terbatas pada aspek manajemen internal fasilitas pelayanan dan belum secara spesifik menganalisis karakteristik produk obat yang dihentikan pemasarannya dari perspektif analisis data pasar secara makro dengan menggunakan kerangka data mining yang sistematis. Di sisi lain, penelitian internasional tentang penghentian produk farmasi cenderung langsung mengembangkan model prediktif tanpa terlebih dahulu melakukan pemetaan eksploratif yang mendalam terhadap karakteristik dasar data. Padahal, analisis eksploratif merupakan tahap krusial untuk memahami distribusi alami data, mendeteksi pola anomali, serta merumuskan hipotesis awal mengenai variabel-variabel yang paling berpengaruh, sebelum memasuki tahap pemodelan yang lebih kompleks. Tanpa fondasi eksploratif yang kuat, model prediktif berisiko mengalami bias dalam pemilihan fitur dan kesulitan dalam interpretasi hasil. Akibatnya, masih terdapat kesenjangan pengetahuan mengenai proporsi dan karakteristik harga produk obat yang dihentikan jika ditinjau dari perspektif data analitik, baik di tingkat nasional

maupun sebagai landasan bagi pengembangan prediksi di masa depan. Padahal, pemahaman mengenai pola dan karakteristik produk yang dihentikan sangat penting untuk merumuskan kebijakan yang efektif dalam menjaga ketahanan obat nasional.

Penelitian terkait prediksi penghentian produk farmasi telah berkembang dengan pesat seiring kemajuan teknologi kecerdasan buatan dan akses terhadap data berskala besar. [11] Mengembangkan model machine learning menggunakan Decision Tree dan Random Forest untuk memprediksi penghentian produk farmasi berdasarkan atribut produk seperti pabrikan, komposisi, dan harga, dengan analisis pentingnya fitur (feature importance) yang mengidentifikasi pabrikan dan komposisi sebagai faktor paling berpengaruh. [12] Menggunakan model bahasa besar (pretrained large language models/LLMs) untuk memprediksi kemungkinan suatu obat ditarik dari pasar karena masalah keamanan, mencapai AUC di atas 0,75 dan mengungguli model machine learning klasik serta model berbasis graf. Sementara itu, [13] Mengembangkan pendekatan explainable machine learning untuk memprediksi penghentian obat biologis pada pasien psoriasis menggunakan data real-world, yang menunjukkan potensi kecerdasan buatan dalam personalisasi pengobatan. Penelitian lain [14] mendemonstrasikan efektivitas LLM open-source dan closed-source (seperti GPT-4o dan BART) dalam mendeteksi penghentian obat dari data publik, dengan BART mencapai F1-score 0,86 tanpa fine-tuning. Berbagai pendekatan berbasis data ini menunjukkan bahwa analisis prediktif terhadap penghentian produk farmasi semakin matang dan beragam. Namun, studi-studi tersebut umumnya mengabaikan tahap eksplorasi mendalam terhadap karakteristik data dasar, sehingga temuan-temuan prediktif yang dihasilkan belum didukung oleh pemetaan awal yang komprehensif mengenai distribusi harga, bentuk sediaan, dan profil pabrikan dari produk yang dihentikan.

Kebaruan (state of the art) penelitian ini tidak terletak pada pengembangan metode statistik baru, melainkan pada penerapan kerangka analisis data yang terstruktur dan komprehensif meliputi preprocessing data (pemilihan fitur dan pembersihan data), analisis deskriptif untuk menghitung proporsi produk yang dihentikan, uji statistik komparatif (uji normalitas Shapiro-Wilk, diikuti dengan uji t-test atau Mann-Whitney), serta analisis kelompok berdasarkan bentuk sediaan, kelas terapi, jumlah bahan aktif, dan pabrikan yang secara khusus diterapkan pada data berskala besar (253.973 produk) produk obat yang dihentikan di pasar farmasi. Dengan demikian, penelitian ini memberikan kontribusi berupa pemetaan dasar (baseline mapping) yang selama ini belum tersedia, khususnya di konteks pasar farmasi skala luas, yang dapat menjadi fondasi bagi pengembangan model prediktif di masa mendatang. Berbeda dengan penelitian sebelumnya yang lebih berfokus pada aspek manajemen internal fasilitas pelayanan [6], [7], [8], [9], [10], atau langsung pada pemodelan prediktif tanpa eksplorasi awal [11], [12], [13], [14], penelitian ini mengisi celah tersebut dengan menyajikan profil eksploratif yang sistematis mengenai dinamika pasar obat dan faktor ekonomi yang memengaruhi keputusan penghentian produk dengan menggunakan pendekatan berbasis data.

Berdasarkan uraian di atas, penelitian ini mengangkat permasalahan mengenai analisis produk obat yang dihentikan menggunakan pendekatan data mining dan analisis statistik, serta menginvestigasi apakah terdapat perbedaan karakteristik harga yang signifikan antara produk yang dihentikan dan produk yang masih aktif. Penelitian sebelumnya telah mendemonstrasikan bahwa teknik analisis data dapat diterapkan secara luas di bidang kesehatan dan farmasi. Oleh karena itu, penelitian ini bertujuan untuk menganalisis produk obat yang dihentikan guna mendukung identifikasi pola produk-produk yang berisiko dihentikan pemasarannya. Secara spesifik, penelitian ini bertujuan untuk: (1) mengetahui proporsi produk obat yang dihentikan secara keseluruhan di pasar farmasi, dan (2) menganalisis perbedaan harga antara produk yang dihentikan dan produk yang masih aktif di pasaran.

Metode

Penelitian ini menggunakan rancangan penelitian deskriptif kuantitatif dengan pendekatan analisis data (data-driven approach) untuk menganalisis produk obat yang dihentikan pemasarannya. Metode yang digunakan terdiri dari beberapa tahapan, mulai dari pengumpulan data, praproses data, analisis deskriptif, uji statistik komparatif, hingga analisis kelompok.

A. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari platform Kaggle dengan judul Indian Pharmaceutical Products yang dipublikasikan oleh Rishabh Sharma [15]. Dataset utama yang digunakan bernama indian_pharmaceutical_products_clean.csv yang merupakan kumpulan data produk farmasi di India yang telah melalui proses pembersihan, standarisasi, dan pengayaan. Data ini terdiri dari lebih dari 250.000 produk farmasi dengan 15 kolom atribut, antara lain: nama obat (name), harga eceran dalam Rupee India (price_inr), status ketersediaan produk (is_discontinued), nama pabrikan (manufacturer), informasi kemasan (pack_size), bentuk sediaan (dosage_form), serta komposisi bahan aktif (num_active_ingredients).

B. Prosedur Analisis Data

Analisis data dalam penelitian ini dilakukan melalui serangkaian tahapan yang sistematis sebagai berikut:

1. Alat dan Perangkat Lunak

Seluruh proses analisis data dalam penelitian ini dilakukan menggunakan bahasa pemrograman Python dengan bantuan library pendukung, yaitu:

- Pandas: Untuk manipulasi dan analisis data
 - NumPy: Untuk perhitungan numerik
 - SciPy: Untuk uji statistik (Shapiro-Wilk, t-test, Mann-Whitney)
 - Matplotlib dan Seaborn: Untuk visualisasi data
 - Warnings: Untuk mengabaikan peringatan sistem yang tidak relevan
2. Praproses Data (Data Preprocessing)
- Sebelum dilakukan analisis lebih lanjut, data terlebih dahulu melalui tahap praproses untuk memastikan kualitas dan integritas data. Tahapan ini meliputi:
- Konversi tipe data: Kolom `is_discontinued` yang menunjukkan status penghentian produk dikonversi menjadi tipe data boolean (True untuk produk dihentikan, False untuk produk aktif) guna memudahkan perhitungan statistik.
 - Pemeriksaan nilai hilang (missing values): Dilakukan identifikasi terhadap nilai-nilai yang hilang pada kolom-kolom kunci yang akan dianalisis, yaitu `is_discontinued`, `price_inr`, `dosage_form`, `pack_size`, `num_active_ingredients`, `therapeutic_class`, dan `manufacturer`. Hasil pemeriksaan menunjukkan bahwa keenam kolom kunci tersebut tidak memiliki nilai hilang (0% missing values), sehingga seluruh data siap digunakan tanpa perlu imputasi. Untuk analisis harga, hanya data dengan harga yang tidak kosong (not null) yang digunakan namun karena tidak ada nilai hilang, seluruh 253.973 data harga digunakan dalam analisis.
 - Penanganan data duplikat: Dilakukan pemeriksaan terhadap kemungkinan adanya data duplikat pada seluruh fitur. Data duplikat diidentifikasi menggunakan fungsi `uplicated()` dan dihapus menggunakan `drop_duplicates()` dari library Pandas untuk memastikan setiap produk hanya terhitung satu kali dalam analisis.
 - Validasi kualitas data: Untuk memastikan data layak analisis, dilakukan validasi dengan langkah-langkah berikut: (1) pemeriksaan rentang harga (`price_inr`) untuk memastikan tidak ada nilai negatif atau nol yang tidak wajar, (2) pengecekan konsistensi kategori pada variabel kategorikal (`dosage_form`, `therapeutic_class`, `manufacturer`) dengan memastikan tidak ada label yang tidak sesuai atau typo, dan (3) perhitungan statistik ringkasan (summary statistics) untuk mendeteksi adanya nilai ekstrem (outlier) yang berpotensi mengganggu interpretasi. Seluruh hasil validasi menunjukkan bahwa data telah bersih dan konsisten, sehingga siap untuk tahap analisis berikutnya.
3. Analisis Deskriptif (Descriptive Analysis)
- Untuk menjawab rumusan masalah pertama mengenai proporsi produk yang dihentikan secara keseluruhan, dilakukan analisis deskriptif dengan langkah-langkah berikut:
- Menghitung total seluruh produk dalam dataset.
 - Menghitung jumlah produk dengan status dihentikan (`is_discontinued = True`) dan produk aktif (`is_discontinued = False`).
 - Menghitung persentase produk dihentikan terhadap total produk dengan rumus:
$$\text{Persentase Dihentikan} = \frac{\text{Jumlah Produk Dihentikan}}{\text{Total Produk}} * 100\%$$
 - Menghitung statistik ringkasan harga (mean, median, standar deviasi, kuartil) untuk masing-masing kelompok produk (aktif dan dihentikan) guna memberikan gambaran awal mengenai distribusi harga antar kelompok.
4. Uji Normalitas Shapiro-Wilk
- Sebelum melakukan uji komparatif harga, dilakukan uji normalitas untuk menentukan apakah data harga pada kedua kelompok terdistribusi normal. Uji normalitas yang digunakan adalah uji Shapiro-Wilk karena metode ini sesuai untuk data kontinu dan memiliki daya uji yang baik. Hipotesis yang diuji adalah:
- H_0 : Data terdistribusi normal ($p \geq 0,05$)
 - H^1 : Data tidak terdistribusi normal ($p < 0,05$)
- Karena ukuran data yang sangat besar, uji normalitas dilakukan terhadap sampel acak sebanyak 5.000 data dari masing-masing kelompok untuk efisiensi komputasi dengan tetap mempertahankan representativitas statistik.
5. Uji Perbedaan Harga
- Untuk menjawab rumusan masalah kedua mengenai perbedaan harga antara produk dihentikan dan produk aktif, dilakukan uji komparatif dengan ketentuan sebagai berikut:

- Jika data harga pada kedua kelompok terdistribusi normal ($p \geq 0,05$), maka digunakan uji independent t-test dengan asumsi varians tidak sama (equal variances not assumed)
- Jika data harga pada salah satu atau kedua kelompok tidak terdistribusi normal ($p < 0,05$), maka digunakan uji non-parametrik Mann-Whitney U test sebagai alternatif

Kedua uji tersebut digunakan untuk menguji hipotesis:

- H_0 : Tidak terdapat perbedaan harga yang signifikan antara produk dihentikan dan produk aktif
- H^1 : Terdapat perbedaan harga yang signifikan antara produk dihentikan dan produk aktif

Tingkat signifikansi (α) ditetapkan sebesar 0,05. Jika nilai p-value $< 0,05$, maka H_0 ditolak yang berarti terdapat perbedaan harga yang signifikan secara statistik.

6. Analisis Kelompok (Group Analysis)

Untuk memperoleh wawasan yang lebih mendalam mengenai karakteristik produk yang dihentikan, dilakukan analisis kelompok berdasarkan beberapa variabel kategorikal, yaitu:

- Bentuk sediaan (dosage_form): Mengidentifikasi bentuk sediaan apa yang memiliki proporsi penghentian tertinggi.
- Kelas terapi (therapeutic_class): Mengidentifikasi kelas terapi dengan proporsi penghentian tertinggi (10 besar)
- Pabrikan (manufacturer): Mengidentifikasi pabrikan dengan proporsi penghentian tertinggi, dengan batas minimum 5 produk per pabrikan untuk menghindari bias akibat ukuran sampel yang terlalu kecil
- Jumlah bahan aktif (num_active_ingredients): Menganalisis hubungan antara jumlah bahan aktif dalam suatu produk dengan kecenderungan penghentian

Untuk setiap analisis kelompok, dihitung tiga metrik: jumlah total produk dalam kelompok (count), jumlah produk yang dihentikan dalam kelompok (sum), dan proporsi penghentian (mean), yang kemudian dikonversi menjadi persentase.

7. Evaluasi dan Interpretasi Hasil

Hasil analisis dievaluasi berdasarkan nilai statistik yang dihasilkan pada setiap tahapan. Untuk analisis proporsi, interpretasi dilakukan dengan membandingkan persentase antar kategori. Untuk uji perbedaan harga, interpretasi didasarkan pada nilai p-value dengan ambang batas signifikansi 0,05. Hasil analisis kelompok disajikan dalam bentuk tabel ringkasan dan visualisasi untuk memudahkan pemahaman dan pengambilan kesimpulan. Perlu dicatat bahwa analisis dalam penelitian ini adalah membandingkan proporsi dan harga antar kelompok secara terpisah tanpa mengontrol variabel pengacau atau interaksi antar variabel secara simultan. Pendekatan ini dipilih karena tujuan utama penelitian adalah eksplorasi dan pemetaan dasar (baseline mapping), bukan identifikasi determinan secara simultan.

Hasil dan Diskusi

Dataset yang digunakan dalam penelitian ini terdiri dari 253.973 produk farmasi. Pemeriksaan nilai hilang pada kolom-kolom kunci yang dianalisis menunjukkan bahwa kolom is discontinued, price_inr, dosage_form, num_active_ingredients, therapeutic_class, dan manufacturer tidak memiliki nilai hilang, sehingga data siap untuk dianalisis lebih lanjut.

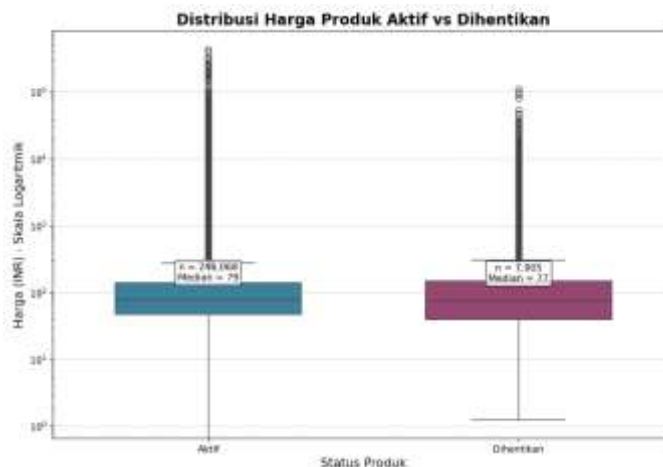
Berdasarkan hasil analisis, dari total 253.973 produk, sebanyak 7.905 produk dinyatakan dihentikan (discontinued), sedangkan 246.068 produk masih aktif di pasaran. Proporsi produk yang dihentikan secara keseluruhan adalah sebesar 3,11%.

Tabel 1. Proporsi Produk yang Dihentikan

<i>Kategori</i>	<i>Jumlah Produk</i>	<i>Persentase</i>
Aktif	246.068	96,89%
Dihentikan	7.905	3,11%
Total	253.973	100%

A. Analisis Perbedaan Harga

Distribusi harga berdasarkan status produk disajikan pada Gambar 1.

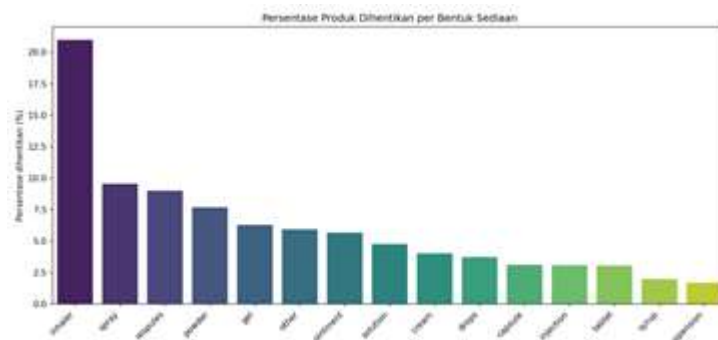


Gambar 1. Boxplot Distribusi Harga Produk Aktif vs Dihentikan

Berdasarkan Gambar 1, terlihat bahwa produk yang dihentikan dan produk aktif memiliki distribusi harga yang berbeda secara visual. Produk yang dihentikan (True) cenderung memiliki variasi harga yang lebih luas pada rentang tertentu, sementara produk aktif (False) menunjukkan distribusi yang lebih terkonsentrasi pada harga rendah hingga menengah. Skala logaritmik digunakan karena data harga sangat condong (skewed), sehingga pola perbedaan antar kelompok dapat terlihat lebih jelas.

Hasil uji normalitas Shapiro-Wilk menunjukkan bahwa data harga pada kedua kelompok tidak terdistribusi normal dengan nilai p-value < 0,001 (aktif: 0,0000; dihentikan: 0,0000). Oleh karena itu, uji perbedaan harga dilakukan menggunakan uji non-parametrik Mann-Whitney U test. Hasil uji menunjukkan nilai p-value = 0,0005, yang berarti terdapat perbedaan harga yang signifikan secara statistik antara produk yang dihentikan dan produk yang aktif di pasaran.

B. Analisis Berdasarkan Bentuk Sediaan

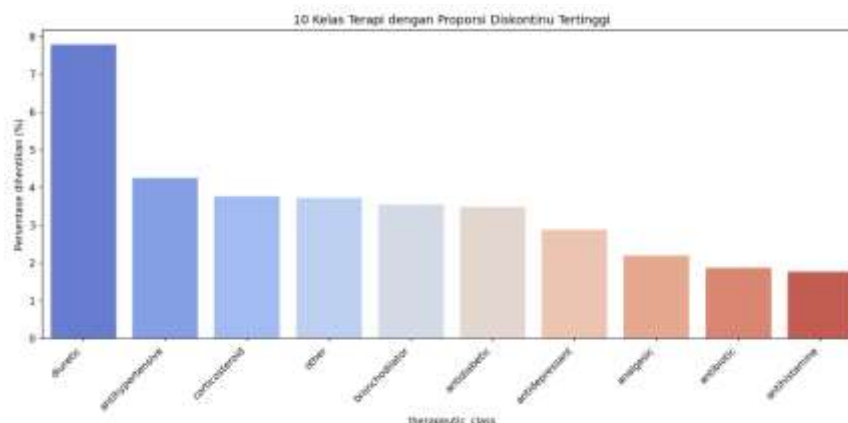


Gambar 2. Persentase Produk Dihentikan per Bentuk Sediaan

Berdasarkan Gambar 2, bentuk sediaan dengan proporsi penghentian tertinggi adalah inhaler (20,97%), spray (9,56%), dan respules (9,02%). Sebaliknya, suspension (1,70%), sirup (1,98%), dan tablet (3,04%) memiliki proporsi penghentian terendah.

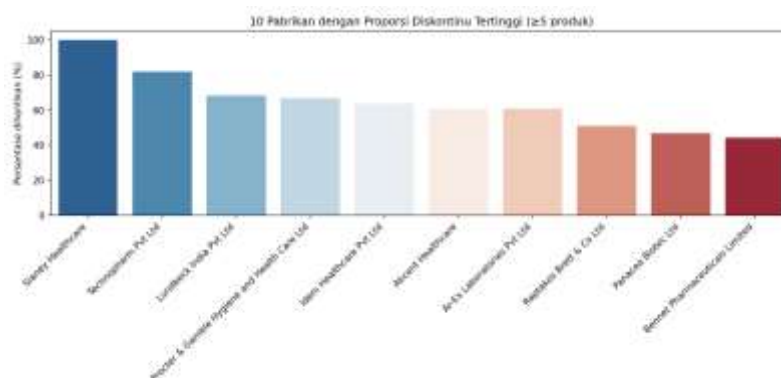
C. Analisis Berdasarkan Kelas Terapi

Berdasarkan Gambar 3, kelas diuretic memiliki proporsi penghentian tertinggi (7,79%), diikuti oleh antihypertensive (4,26%) dan corticosteroid (3,77%). Kelas dengan proporsi terendah di antara sepuluh besar adalah antihistamine (1,77%) dan antibiotic (1,87%).



Gambar 3. Sepuluh Kelas Terapi dengan Proporsi Diskontinu Tertinggi

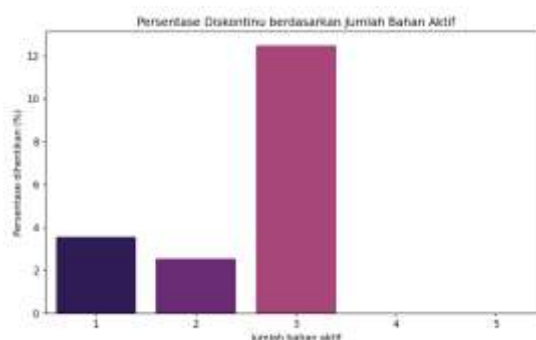
D. Analisis Berdasarkan Pabrik



Gambar 4. Sepuluh Pabrik dengan Proporsi Diskontinu Tertinggi (≥5 Produk)

Berdasarkan Gambar 4, Slaney Healthcare memiliki tingkat penghentian tertinggi (100%), diikuti Technopharm Pvt Ltd (81,82%) dan Lundbeck India Pvt Ltd (68,18%).

E. Analisis Berdasarkan Jumlah Bahan Aktif



Gambar 5. Persentase Diskontinu berdasarkan Jumlah Bahan Aktif

Berdasarkan Gambar 5, produk dengan 3 bahan aktif memiliki proporsi penghentian tertinggi (12,50%). Namun, jumlah produk dalam kategori ini sangat kecil (8 produk), sehingga hasil ini perlu diinterpretasi dengan hati-hati. Sebagian besar produk memiliki 1 atau 2 bahan aktif dengan proporsi masing-masing 3,56% dan 2,55%.

F. Diskusi

Berdasarkan Tabel 1, proporsi produk yang dihentikan adalah sebesar 3,11% dari total 253.973 produk. Angka ini menunjukkan bahwa mayoritas produk farmasi dalam dataset (96,89%) masih aktif di pasaran. Temuan ini sejalan dengan [1] yang melaporkan bahwa jumlah obat yang dihentikan di Jepang menunjukkan peningkatan dari tahun ke tahun.

Berdasarkan Gambar 1, rata-rata harga produk dihentikan (442,63 INR) lebih tinggi dibandingkan produk aktif (265,00 INR). Namun, nilai median menunjukkan pola berbeda, di mana produk aktif memiliki median 79,0 INR dan produk dihentikan 77,0 INR. Hasil uji Mann-Whitney menunjukkan $p\text{-value} = 0,0005$, yang berarti terdapat perbedaan harga yang signifikan secara statistik antara produk dihentikan dan produk aktif.

Perbedaan antara rata-rata dan median harga yang sangat kontras ini mengindikasikan bahwa data harga pada kedua kelompok sangat condong (skewed), dengan keberadaan produk-produk berharga ekstrem yang memengaruhi nilai rata-rata. Temuan ini sejalan dengan penelitian [11] yang mengidentifikasi bahwa harga merupakan salah satu atribut penting dalam prediksi penghentian produk, namun tidak berdiri sendiri harga perlu dianalisis bersama dengan bentuk sediaan dan strategi pabrikan untuk memperoleh gambaran yang lebih utuh. Dengan kata lain, produk yang dihentikan cenderung memiliki variasi harga yang lebih lebar, yang mungkin mencerminkan strategi penetapan harga yang berbeda antar pabrikan atau segmen pasar yang dilayani.

Berdasarkan Gambar 2, bentuk sediaan inhaler, spray, dan respules memiliki proporsi penghentian tertinggi. Ketiga bentuk sediaan ini merupakan produk dengan rute administrasi inhalasi yang memerlukan teknologi formulasi khusus. Sebaliknya, tablet dan kapsul yang merupakan bentuk sediaan paling umum memiliki proporsi penghentian yang relatif rendah. Hal ini sejalan dengan [2] bahwa produk dengan biaya produksi tinggi dan permintaan terbatas memiliki risiko penghentian lebih besar. Berdasarkan Gambar 3, kelas diuretic memiliki proporsi penghentian tertinggi. Diuretik dan antihipertensi merupakan kelas terapi dengan jumlah produk yang besar dan persaingan yang ketat.

Tingginya proporsi penghentian pada kelas diuretik (7,79%) dan antihipertensi (4,26%) perlu dipahami dalam konteks pasar obat kardiovaskular yang sangat kompetitif dengan banyaknya produk generik. Dalam pasar yang padat, produk-produk dengan harga kurang kompetitif atau pangsa pasar kecil cenderung dihentikan untuk mengoptimalkan portofolio pabrikan. Temuan ini konsisten dengan Baba et al.[1] yang melaporkan bahwa produk generik dengan tekanan harga tinggi lebih berisiko dihentikan. Di sisi lain, kelas dengan proporsi rendah seperti antibiotik (1,87%) dan antihistamin (1,77%) mungkin mencerminkan permintaan yang relatif stabil dan kebutuhan klinis yang berkelanjutan, sehingga pabrikan cenderung mempertahankan produk-produk tersebut meskipun menghadapi persaingan.

Berdasarkan Gambar 4, terdapat variasi yang sangat signifikan dalam proporsi penghentian antar pabrikan. Variasi ini menunjukkan bahwa keputusan penghentian produk sangat dipengaruhi oleh strategi bisnis dan manajemen portofolio masing-masing pabrikan. Temuan ini sejalan dengan [11] yang mengidentifikasi bahwa pabrikan merupakan salah satu faktor paling berpengaruh dalam memprediksi penghentian produk farmasi.

Kasus Slaney Healthcare dengan tingkat penghentian 100% dari total 5 produknya memerlukan interpretasi yang sangat hati-hati. Angka 100% tidak serta-merta dapat diartikan bahwa pabrikan tersebut memiliki kualitas produk yang buruk atau gagal dalam memenuhi regulasi. Sebaliknya, proporsi 100% pada portofolio yang sangat kecil (hanya 5 produk) lebih mungkin mencerminkan keputusan strategis bisnis, seperti penghentian seluruh lini produk di pasar tertentu, peralihan fokus ke produk lain, atau bahkan akuisisi/rebranding oleh perusahaan lain. Dengan demikian, angka ekstrem seperti ini sebaiknya tidak digunakan sebagai indikator tunggal kinerja atau kualitas pabrikan, melainkan sebagai sinyal awal yang memerlukan investigasi lebih mendalam dengan melibatkan variabel tambahan seperti riwayat izin edar, volume penjualan, dan perubahan kepemilikan perusahaan. Penelitian lanjutan dengan analisis survival atau regresi logistik dapat membantu mengidentifikasi apakah faktor-faktor seperti ukuran portofolio dan diversifikasi produk secara signifikan memengaruhi risiko penghentian.

Berdasarkan Gambar 5, produk dengan 3 bahan aktif memiliki proporsi penghentian tertinggi. Namun, jumlah produk dalam kategori ini sangat kecil. Hal ini sejalan dengan [13] bahwa kompleksitas produk dapat memengaruhi keputusan penghentian.

Penelitian ini memiliki beberapa keterbatasan. Pertama, data yang digunakan hanya mencakup satu wilayah. Kedua, penelitian ini hanya menganalisis faktor-faktor tertentu tanpa mempertimbangkan variabel lain yang tidak ada dalam data seperti masa izin edar atau pangsa pasar. Meskipun demikian, hasil ini memberikan dasar empiris untuk analisis lebih lanjut.

Kesimpulan

Berdasarkan hasil analisis data terhadap 253.973 produk farmasi, penelitian ini berhasil menjawab kedua rumusan masalah yang telah ditetapkan. Pertama, proporsi produk yang dihentikan secara keseluruhan adalah sebesar 3,11% atau sebanyak 7.905 produk dari total 253.973 produk, sedangkan sisanya sebanyak 246.068 produk (96,89%) masih aktif di pasaran. Temuan ini menunjukkan bahwa penghentian produk farmasi merupakan fenomena yang terjadi pada sebagian kecil produk, namun tetap signifikan untuk diperhatikan dalam konteks ketersediaan dan keberlanjutan pasokan obat. Kedua, terdapat perbedaan harga yang signifikan secara statistik antara produk yang dihentikan dan produk yang aktif di pasaran ($p\text{-value} = 0,0005$). Rata-rata harga produk dihentikan (442,63 INR) lebih tinggi dibandingkan produk aktif (265,00 INR), namun nilai median produk aktif (79,0 INR) sedikit lebih tinggi dari produk dihentikan (77,0 INR). Perbedaan antara rata-rata dan median ini menunjukkan bahwa distribusi data harga pada kedua kelompok sangat condong dengan adanya produk-produk berharga ekstrem. Selain itu, analisis kelompok

menunjukkan bahwa produk dengan bentuk sediaan inhaler (20,97%), spray (9,56%), dan respules (9,02%) memiliki proporsi penghentian tertinggi, sedangkan kelas terapi diuretic (7,79%) merupakan kelas dengan proporsi penghentian tertinggi. Terdapat variasi yang sangat signifikan antar pabrikan, di mana Slaney Healthcare memiliki tingkat penghentian 100% untuk seluruh produknya. Produk dengan 3 bahan aktif memiliki proporsi penghentian tertinggi (12,50%), namun jumlah produk dalam kategori ini sangat kecil sehingga memerlukan interpretasi.

Deklarasi Kepentingan

Penulis menyatakan tidak terdapat konflik kepentingan dalam penelitian ini. Penelitian ini merupakan bagian dari tugas dan tanggung jawab penulis sebagai dosen dalam melaksanakan penelitian di bidang data science dan farmasi. Tidak ada afiliasi, hubungan keuangan, atau kepentingan pribadi yang dapat mempengaruhi objektivitas, analisis, atau interpretasi hasil penelitian ini. Penelitian ini juga merupakan kontribusi ilmiah penulis untuk pengembangan ilmu pengetahuan dan mendukung kegiatan tridharma perguruan tinggi

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada PT. Irmex Digital Akademika yang telah memberikan dukungan fasilitas dalam pelaksanaan penelitian ini. Ucapan terima kasih juga disampaikan kepada rekan satu tim, Firman Aziz, serta seluruh pihak yang telah memberikan bantuan berharga dalam penelitian ini, baik secara langsung maupun tidak langsung, mulai dari proses awal, analisis, hingga penyusunan naskah artikel. Penelitian ini tidak menerima pendanaan khusus dari lembaga pendanaan mana pun di sektor publik, komersial, atau nirlaba.

Ketersediaan Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang tersedia untuk umum dan dapat diakses secara bebas melalui platform Kaggle pada tautan: <https://www.kaggle.com/datasets/rishgeeky/indian-pharmaceutical-products>. Dataset tersebut merupakan data publik yang telah melalui proses pembersihan dan standarisasi oleh kontributor. Seluruh kode analisis dan hasil olahan data yang digunakan dalam penelitian ini tersedia dan dapat direproduksi. Penulis menyatakan bahwa tidak ada data privat atau rahasia yang digunakan dalam penelitian ini. Untuk reproduksi hasil penelitian, pembaca dapat mengakses dataset melalui tautan yang telah disebutkan dan menjalankan kode analisis yang tercantum dalam metode penelitian.

Penggunaan Ai Dan Deklarasi Penggunaan Ai Generatif

Penulis menyatakan bahwa dalam penelitian ini, alat kecerdasan buatan (AI) generatif digunakan sebagai alat bantu untuk mendukung pekerjaan, namun tidak sebagai pengganti pemikiran kritis manusia dan tetap berada di bawah pengawasan serta kendali penulis. Penggunaan AI dalam penelitian ini meliputi:

1. *Parafrase dan Penyempurnaan Naskah:* Alat AI generatif (DeepSeek) digunakan untuk membantu parafrase dan penyempurnaan tata bahasa dalam penulisan naskah artikel, termasuk bagian pendahuluan, metode, hasil, dan diskusi. Seluruh konten yang dihasilkan oleh AI telah ditinjau, diedit, dan disesuaikan oleh penulis untuk memastikan keakuratan, koherensi, dan kesesuaian dengan temuan penelitian.
2. *Debugging dan Optimasi Kode Python:* Alat AI generatif (DeepSeek) digunakan untuk membantu mengidentifikasi dan memperbaiki kesalahan dalam kode Python yang digunakan untuk analisis data. Seluruh kode dijalankan dalam lingkungan PyCharm dan hasil analisis yang diperoleh langsung ditulis tanpa modifikasi dari AI. Penulis bertanggung jawab penuh atas kebenaran kode, interpretasi hasil, dan kesimpulan yang diambil.
3. *Penggunaan AI untuk Modifikasi Gambar:* Tidak ada penggunaan AI untuk memodifikasi gambar dalam penelitian ini. Seluruh visualisasi (boxplot, bar chart) dihasilkan langsung dari kode Python menggunakan library Matplotlib dan Seaborn tanpa modifikasi atau manipulasi gambar oleh AI.

Penulis menegaskan bahwa seluruh konten yang dihasilkan oleh AI telah melalui proses verifikasi dan penyuntingan oleh penulis, dan penulis bertanggung jawab penuh atas isi artikel yang dipublikasikan. Penelitian ini merupakan hasil kerja intelektual penulis yang didukung oleh alat AI sebagai sarana bantu teknis dan editorial, bukan sebagai kontributor substansial dalam penelitian.

Daftar Pustaka

- [1] Baba A. *et al.*, "Profile of Generic Drugs Associated with the Discontinuation of the Supply," *YAKUGAKU ZASSHI*, vol. 143, no. 12, pp. 1069–1073, Dec. 2023, doi: 10.1248/yakushi.23-00098.
- [2] C. L. Ventola, "Big Data and Pharmacovigilance: Data Mining for Adverse Drug Events and Interactions".
- [3] R. Hussain, "Big data, medicines safety and pharmacovigilance," *J of Pharm Policy and Pract*, vol. 14, no. 1, p. 48, Dec. 2021, doi: 10.1186/s40545-021-00329-4.

- [4] M. Rijal, F. Aziz, and S. Abasa, "PREDIKSI DEPRESI: INOVASI TERKINI DALAM KESEHATAN MENTAL MELALUI METODE MACHINE LEARNING," *Journal Pharmacy and Application of Computer Sciences*, vol. 2, no. 1, pp. 9–14, Feb. 2024, doi: 10.59823/jopacs.v2i1.47.
- [5] A. Z. Ifani, V. Alviani, D. Suryanti, and N. W. Asbara, "Pengoptimalan Internet of Medical Things (IoMT) untuk Pelayanan Kesehatan," *Jurnal Fokus Elektroda : Energi Listrik, Telekomunikasi, Komputer, Elektronika dan Kendali*, vol. 9, no. 4, pp. 234–238, Nov. 2024, doi: 10.33772/jfe.v9i4.949.
- [6] K. Fauziah, Nurmiati Muchlis, and W. Hamzah, "Evaluasi Pengelolaan Sediaan Farmasi Di Puskesmas Bara-Baraya Kota Makassar," *woph*, vol. 5, no. 5, pp. 717–728, Oct. 2024, doi: 10.33096/woph.v5i5.1953.
- [7] D. Wahyono and I. Hapsari, "Tinjauan Naratif : Evaluasi Pengelolaan Obat Puskesmas di Indonesia," vol. 9, no. 1, 2025.
- [8] G. Citraningtyas, G. E. Rundengan, F. Manawan, and M. Rumondor, "EVALUASI PENGELOLAAN OBAT RUSAK ATAU KEDALUWARSA DI RUMAH SAKIT TK. II R. W. MONGISIDI MANADO," vol. 7, no. 2, 2024.
- [9] D. Nurcahyani and A. D. Saputri, "Evaluasi Pengelolaan Obat Yang Mendekati Waktu Kedaluwarsa Tahun 2024 Di Instalasi Farmasi RS X Ngawi," vol. 14, 2025.
- [10] H. M. Rumagit, T. Sunarni, and I. Purwidyaningrum, "Analisis Pengelolaan Obat Dan Strategi Perbaikan Dengan Metode Hanlon Di Instalasi Farmasi RSU Budi Setia Kabupaten Minahasa Provinsi Sulawesi Utara," *jfi*, vol. 21, no. 2, pp. 59–69, Nov. 2024, doi: 10.31001/jfi.v21i2.825.
- [11] S. Suhartono, "Predicting Pharmaceutical Product Discontinuation Using Decision Tree and Random Forest Algorithms Based on Product Attributes," *Int. J. Appl. Inf. Manag.*, vol. 5, no. 2, pp. 86–97, Jul. 2025, doi: 10.47738/ijaim.v5i2.101.
- [12] E. Mazuz, G. Shtar, N. Kutsky, L. Rokach, and B. Shapira, "Pretrained transformer models for predicting the withdrawal of drugs from the market," *Bioinformatics*, vol. 39, no. 8, p. btad519, Aug. 2023, doi: 10.1093/bioinformatics/btad519.
- [13] A. B. Hussain *et al.*, "P002 Predicting psoriasis biologic drug discontinuation: an explainable machine learning approach applying British Association of Dermatologists Biologics and Immunomodulators Register data," *Br J Dermatol*, vol. 193, no. Supplement_1, p. ljaf085.030, Jul. 2025, doi: 10.1093/bjd/ljaf085.030.
- [14] W. Trevena *et al.*, "Using Large Language Models to Detect and Understand Drug Discontinuation Events in Web-Based Forums: Development and Validation Study," *J Med Internet Res*, vol. 27, p. e54601, Jan. 2025, doi: 10.2196/54601.
- [15] "Indian Pharmaceutical Products." Accessed: Jun. 29, 2026. [Online]. Available: <https://www.kaggle.com/datasets/rishgeeky/indian-pharmaceutical-products>