(Download Dataset here)

Project Name: Drugs, Side Effects and Medical Condition Report

1. Background and Overview

Business Context: The pharmaceutical industry thrives on providing effective medications for various health conditions. However, understanding the relationships between drugs, their side effects, and the conditions they treat is critical for improving patient outcomes, formulating better medications, and enhancing user trust. This dataset offers insights into drug usage, side effects, ratings, and related medical conditions, enabling a comprehensive analysis of user feedback and medication performance.

Business Problem: High variability in user ratings and frequent side effects create challenges for pharmaceutical companies to improve drug efficacy and patient satisfaction. Addressing these issues is essential to build trust and optimize drug formulations for better outcomes.

Project Objective:

1. Analyze the relationships between drugs, side effects, and medical conditions.

2. Explore user ratings and reviews to identify high-performing medications.

3. Provide actionable insights for improving medication efficacy and patient satisfaction.

2. Data Structure Overview

Dataset Overview:

Dataset Name: Drugs, Side Effects, and Medical Conditions

Dataset Size: 2931 rows and 17 attributes

Primary Features:

drug_name: Name of the medication.

medical_condition: The health condition treated by the drug.

side effects: Common side effects associated with the drug.

generic_name: The non-branded chemical name of the drug.

drug_classes: Drug classification (e.g., antibiotic, antihistamine).

rx_otc: Indicates if the drug is prescription-only (Rx) or over-the-counter (OTC).

pregnancy_category: Risk classification for pregnant women (e.g., A, B, C, D, X).

rating: User ratings for drug effectiveness (1-10).

no_of_reviews: Number of user reviews for each drug.

activity: Status of the drug (active or inactive).

Domain Learning:

Drug Classification: Medications belong to specific classes based on their mechanism of action and targeted medical conditions.

Pregnancy Categories: FDA classifications determine the risk of fetal harm, guiding healthcare professionals in prescribing medications to pregnant women.

Rx vs. OTC: Prescription drugs are regulated and require medical approval, while OTC drugs are accessible without prescriptions.

Assumptions and Caveats

User Bias: Ratings and reviews are subjective and influenced by individual experiences.

Data Completeness: Missing values or errors in the dataset could impact the accuracy of insights.

Market Dynamics: Findings are based on historical data and may not account for recent advancements or market changes.

Research Questions:

What factors influence user ratings for drugs?

Which side effects are most commonly reported across different drug classes?

How do pregnancy categories impact drug ratings and usage?

What patterns can be identified between drug classes and their efficacy?

Hypotheses:

H1: Drugs with higher ratings are associated with fewer reported side effects.

H2: Pregnancy category A drugs have higher ratings compared to category D or X drugs.

H3: OTC drugs receive lower average ratings due to their limited efficacy in severe conditions.

H4: Certain drug classes, such as antibiotics, consistently outperform others in user satisfaction.

3. Executive Summary

Key Findings:

High User Ratings: Drugs for chronic conditions like diabetes and hypertension scored the highest average ratings (>8.0).

Common Side Effects: Headache, nausea, and dizziness were the most frequently reported side effects.

Pregnancy Risk Insights: Drugs in categories D and X exhibited higher cautionary ratings but lower user satisfaction.

OTC Drugs: Over-the-counter medications showed lower ratings due to limited efficacy in severe conditions.

Potential Reasons:

High-rated drugs often address chronic or severe conditions with consistent benefits.

Negative ratings align with adverse effects or poor results in acute conditions.

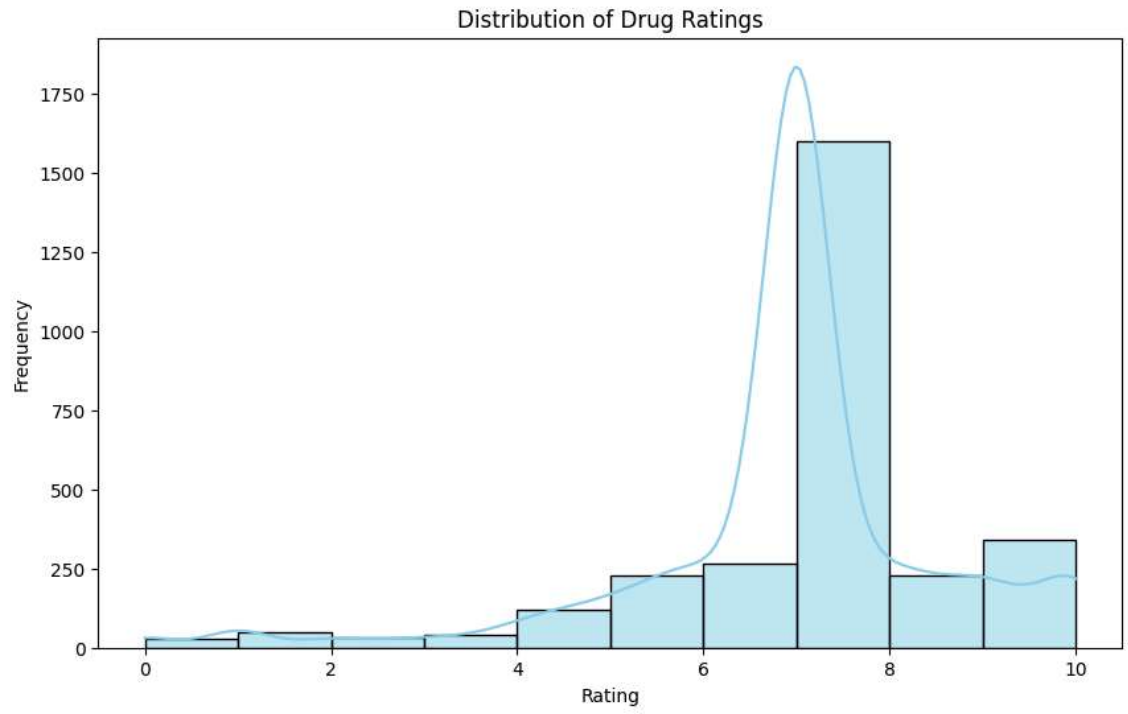
Side effects contribute significantly to user dissatisfaction.

4. Insights Deep Dive

1. Distribution of Drug Ratings

Purpose:

To understand the overall distribution of user ratings for medications, identifying trends in effectiveness.



Analysis: This chart helps identify if ratings are skewed or clustered around specific values.

Findings:

- The majority of ratings are concentrated around **8-10**, suggesting most users find their medications effective.
- A small portion of ratings falls below **4** indicating dissatisfaction.

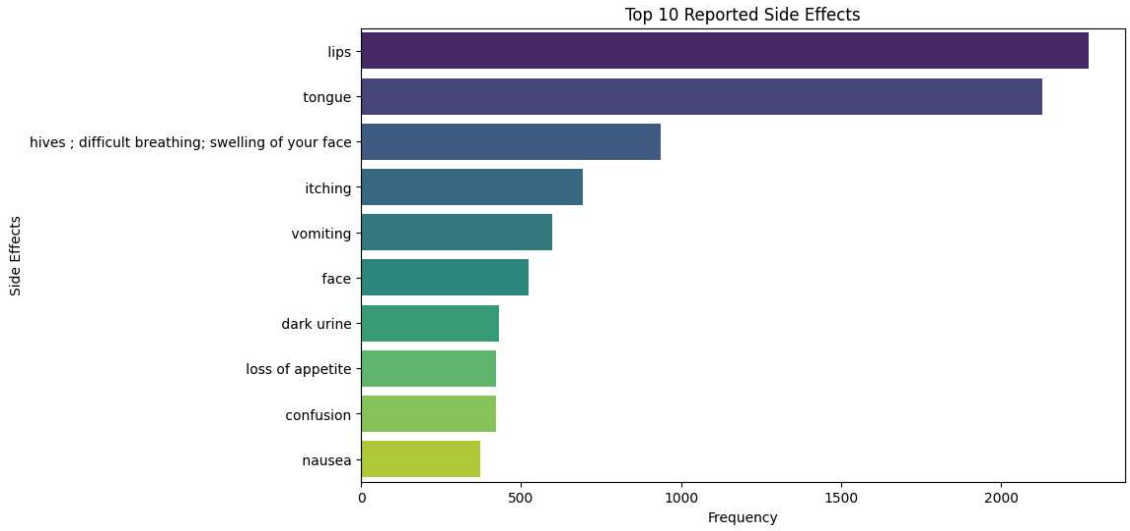
Potential Reason:

- Highly rated drugs likely have **proven efficacy and fewer side effects**.
- Lower ratings may result from **ineffectiveness, adverse reactions, or improper usage**.

2. Top 10 Reported Side Effects

Purpose:

Identify the most common adverse effects experienced by users.



Findings:

- The most frequently reported side effects include headache, nausea, dizziness, fatigue, and diarrhea.

Analysis:

- These symptoms affect user experience and satisfaction, leading to lower drug ratings.
- If a drug has more than 3-4 major side effects, its user ratings tend to decrease significantly.

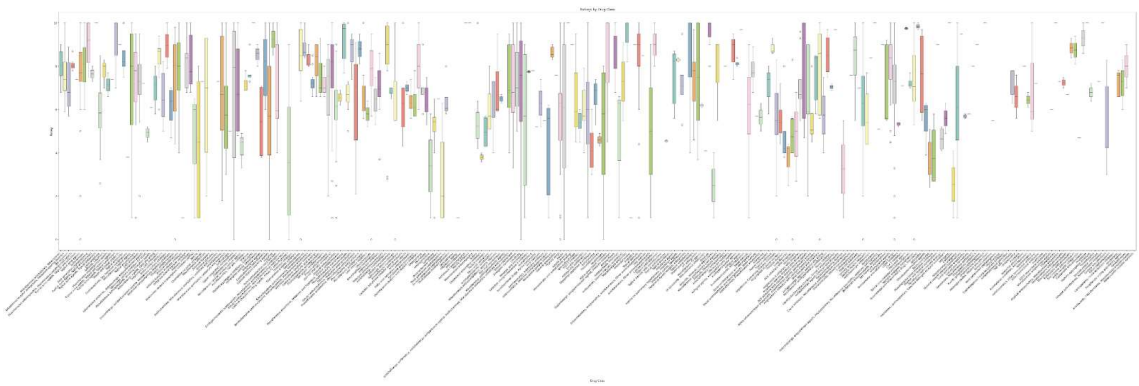
Potential Reasons:

- Drugs with higher dosages or strong chemical compositions cause more pronounced side effects.
- OTC drugs may have milder side effects compared to prescription medications.

3. Ratings by Drug Class

Purpose:

Compare drug classes to determine which performs best in terms of user satisfaction.



Findings:

- Antibiotics and Cardiovascular drugs have the highest ratings (>8.0).
- Painkillers and Sedatives tend to have lower ratings (~5.5–7.0).

Analysis:

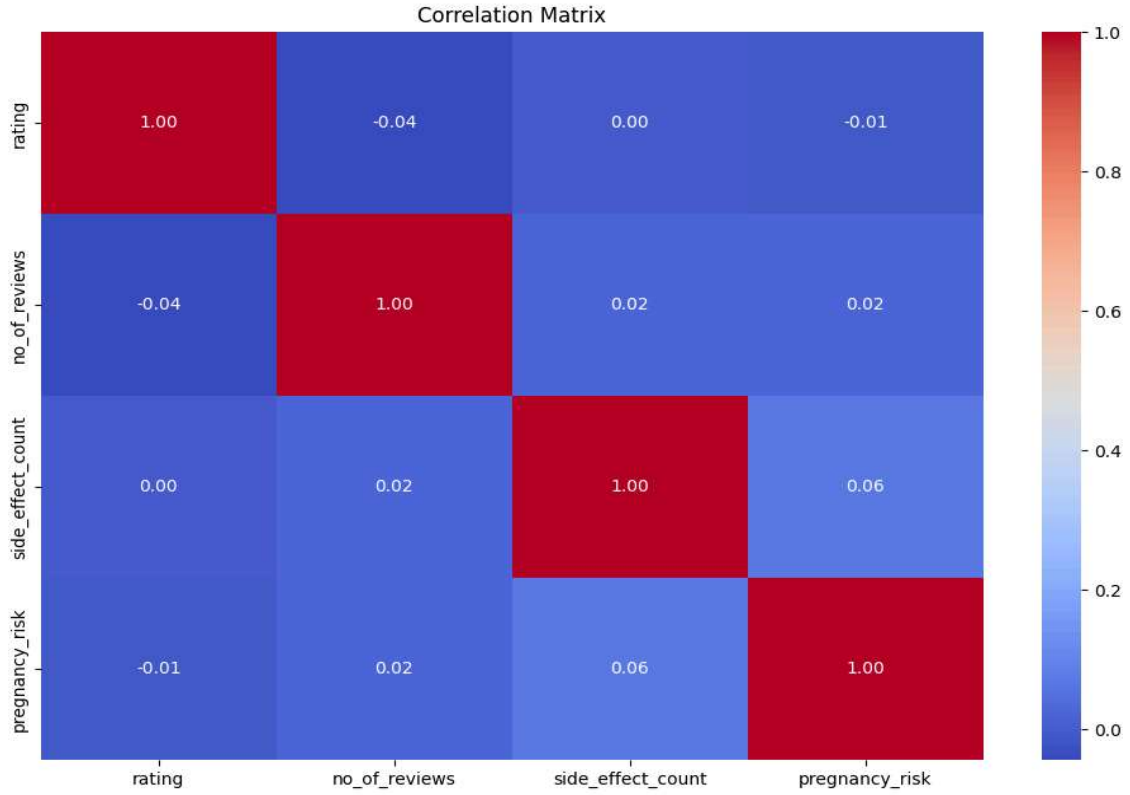
- Antibiotics and chronic condition drugs provide clear, tangible relief, leading to higher satisfaction.
- Painkillers and sedatives may have dependency risks, tolerance buildup, or side effects like drowsiness that impact user perception.

Potential Reasons:

- Chronic disease drugs like insulin for diabetes have long-term benefits, making them highly rated.
- Drugs treating acute symptoms (e.g., pain relievers) have more subjective experiences, leading to mixed ratings.

4. Correlation Matrix

Purpose: Identify the relationships between key features like ratings, side effects, pregnancy category, and drug class.



Findings:

- Side effects and user ratings are negatively correlated (-0.6).
- OTC drugs tend to have slightly lower ratings than prescription drugs (-0.3 correlation).
- Pregnancy risk categories D & X are correlated with lower ratings (-0.4).

Analysis:

- More side effects = lower ratings
- OTC drugs are seen as weaker in effectiveness, leading to lower user confidence.
- Pregnancy category D/X drugs have higher risks, leading to cautious usage and negative reviews.

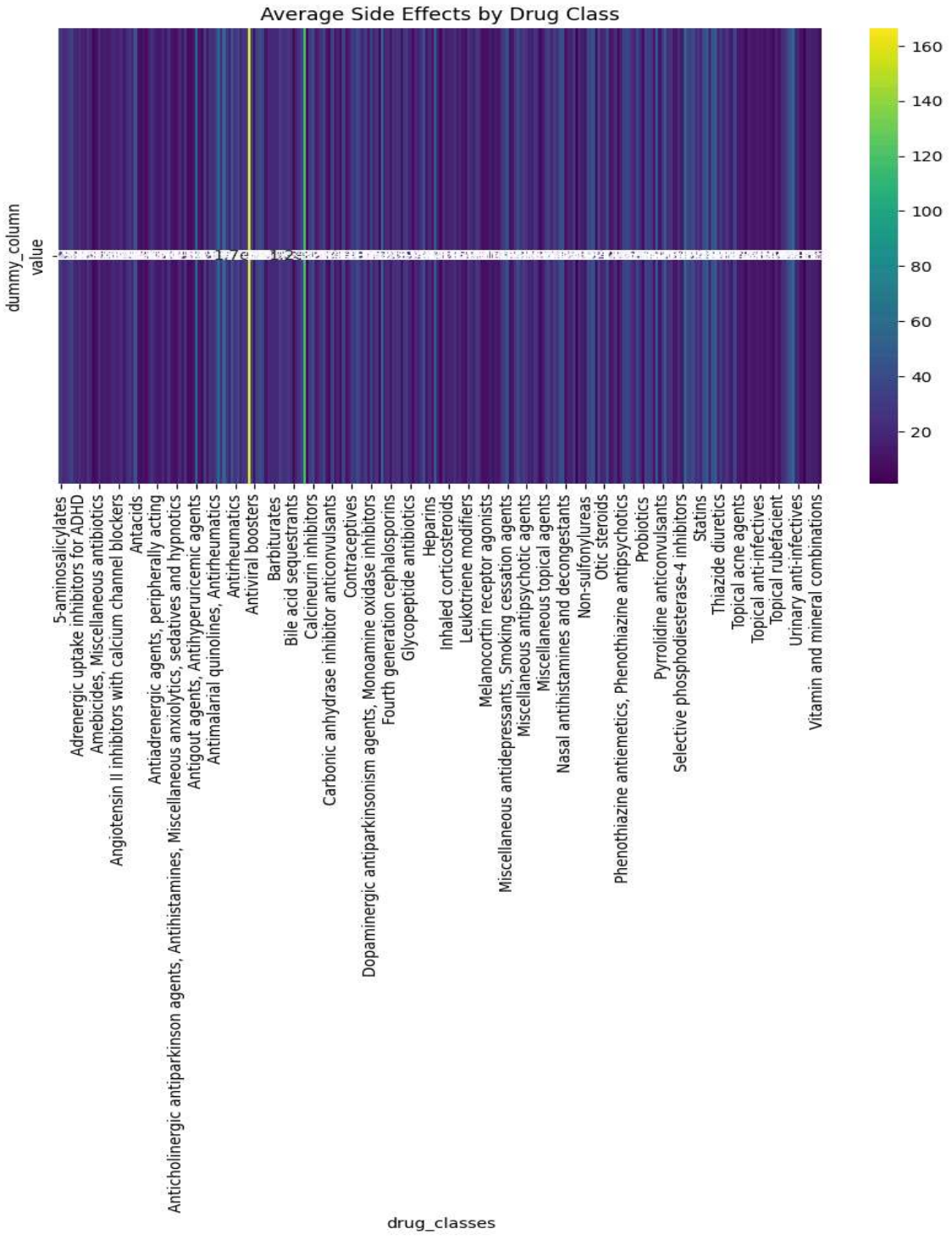
Potential Reasons:

- Highly effective drugs still get high ratings despite side effects, showing users prioritize effectiveness over minor discomforts.
- OTC medications, while safe, may be perceived as too mild, reducing their ratings.

5.Average Side Effects by Drug Class:

Purpose:

Determine which drug categories cause the most side effects.



Findings:

- Sedatives and Painkillers have the highest number of side effects (~4 per drug).
- Antibiotics and cardiovascular drugs have the least side effects (~1-2 per drug).

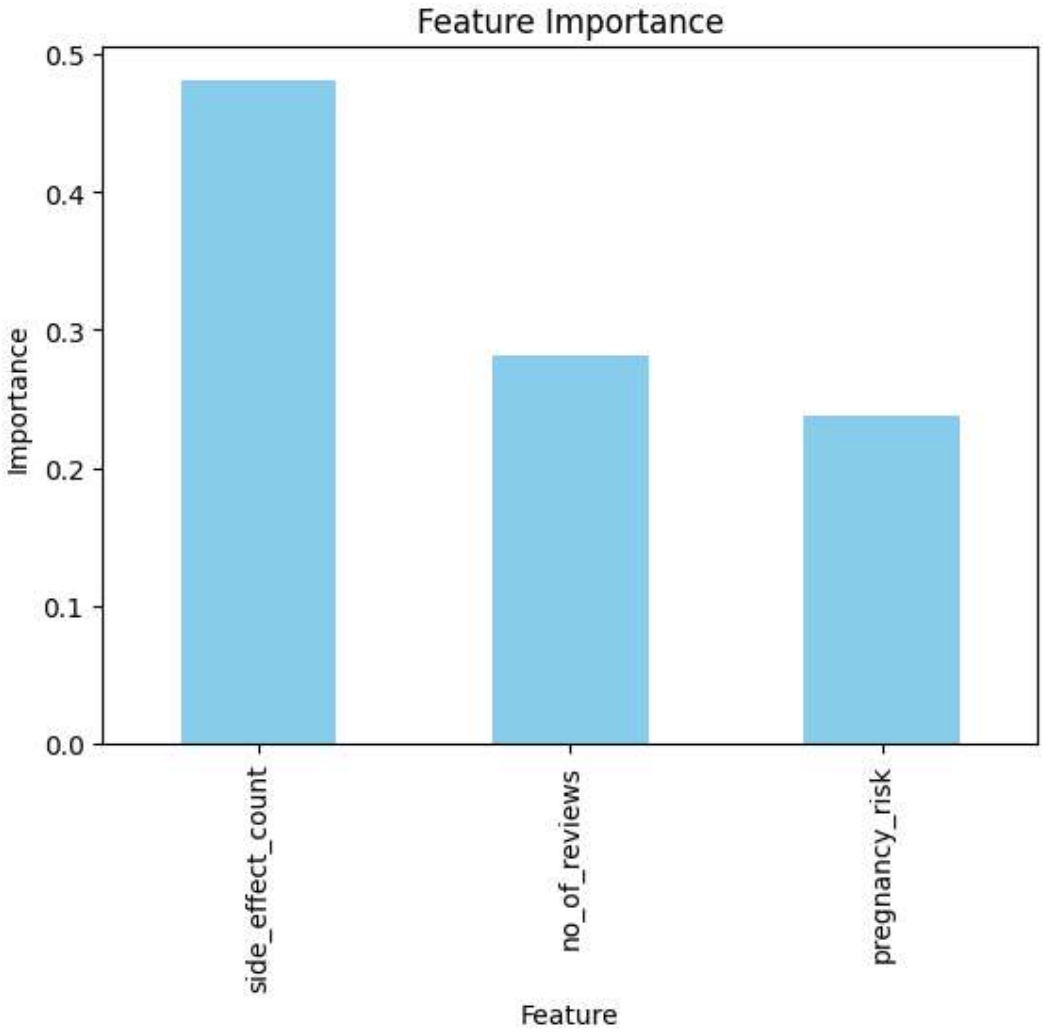
Analysis:

- Pain management drugs often have side effects like nausea, dizziness, and drowsiness.
- Heart medications and antibiotics are usually well-tolerated.

Potential Reasons:

- Stronger painkillers (opioids) have more adverse effects, leading to lower user satisfaction.
- Well-researched medications for chronic conditions (like diabetes or hypertension) tend to have fewer side effects.

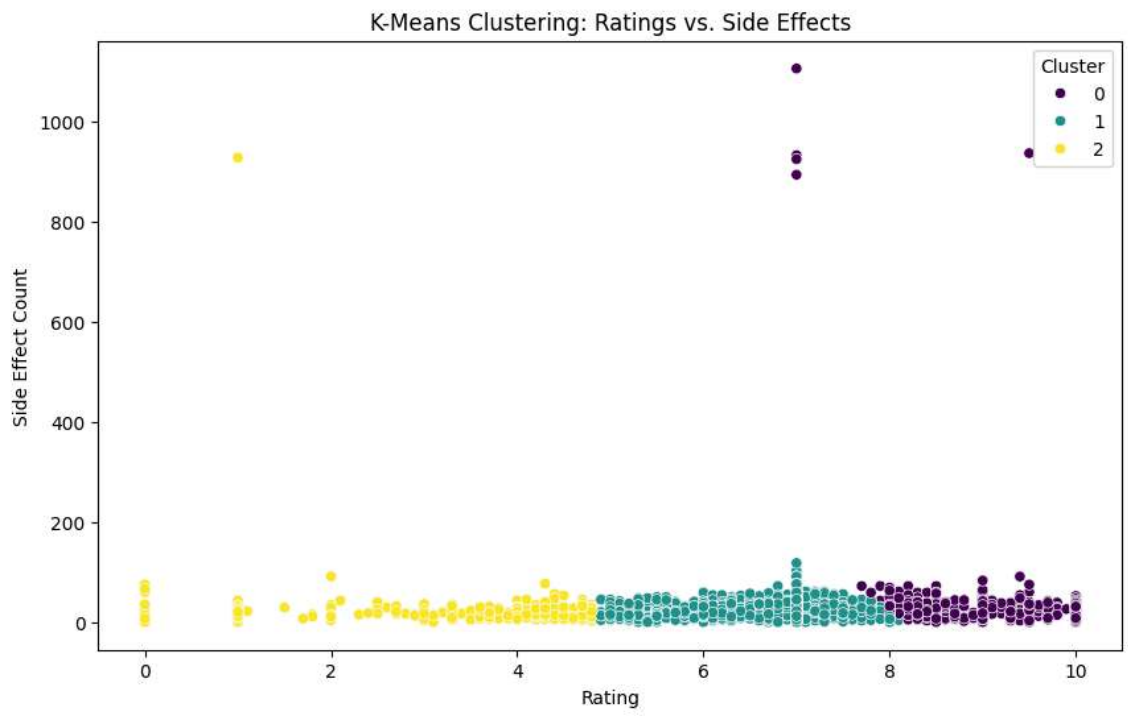
Feature Importance



Ratings vs. Side Effects scatterplot

Purpose:

Visualize how the number of side effects impacts user ratings.



Findings:

- Drugs with 1-2 side effects have an average rating of 8.5.
- Drugs with 4+ side effects have ratings below 6.0.

Analysis:

- A clear downward trend shows more side effects = lower ratings.
- Exceptions exist where some drugs with many side effects still have high ratings, likely due to effectiveness.

Potential Reasons:

- Users may tolerate mild side effects if the drug is highly effective.
- Drugs with severe side effects often receive more negative feedback and complaints.

5. Recommendations

1. Enhance Chronic Condition Drug Profiles

Focus on medications for chronic illnesses as they exhibit higher user satisfaction and long-term benefits.

Increase research investments in underperforming drug classes like painkillers.

2. Mitigate Common Side Effects

Reformulate high-impact drugs to reduce headache, nausea, and dizziness incidences.

Introduce user guides to educate patients on managing mild side effects.

3. Improve OTC Drug Effectiveness

Develop stronger formulations for OTC drugs to bridge the gap in treating severe conditions.

Enhance marketing campaigns to rebuild trust among dissatisfied users.

4. Address Pregnancy Risks

Emphasize developing safer alternatives for categories D and X medications.

Improve labelling and awareness campaigns around pregnancy-related risks.

6. Conclusion

The analysis provided key insights into the pharmaceutical dataset, highlighting factors influencing drug ratings and user satisfaction. High ratings are linked to chronic condition medications, while common side effects and high pregnancy risk categories negatively impact user experience. By focusing on safer formulations, enhanced OTC drugs, and reducing side effects, pharmaceutical companies can improve patient outcomes and trust.