

Amazon E-Commerce Product & Review Analytics

This project analyzes Amazon e-commerce product and review data using SQL to understand pricing, discounts, product performance, and customer sentiments. SQL queries are used to identify discounted products, high-rated items, category trends, and review insights, helping convert raw data into actionable business insights.

Presentation By;

PRAKASH CHAWDA

Understanding the Business Challenge

The Data Landscape

Amazon hosts millions of products spanning diverse categories. To stay competitive, business teams must decode complex patterns in pricing strategies, customer preferences, and sentiment trends.

This project simulates a real-world business analytics scenario, empowering data-driven decision-making through structured SQL analysis.

Core Questions Answered

- Which products offer the deepest discounts?
- How does pricing influence customer ratings?
- What categories drive the highest performance?
- What insights emerge from customer reviews?

Project Objectives: What We Achieve



Pricing Strategy Analysis

Uncover discount patterns and identify pricing thresholds that maximize both sales volume and customer satisfaction ratings.



Product Performance

Pinpoint high-rated, high-value products that represent competitive advantages and revenue opportunities.



Sentiment Insights

Extract meaningful themes from customer reviews using text analysis to understand pain points and preferences.



Category Intelligence

Analyze product distribution across categories to guide inventory decisions and marketing focus areas.

Dataset Architecture

The dataset integrates product details, pricing dynamics, customer ratings, and textual reviews into a comprehensive analytics framework. Each record contains 15+ attributes enabling both quantitative and qualitative analysis.



Product Attributes

IDs, names, categories, descriptions, and image links



Pricing Data

Original prices, discounted prices, and percentage discount



Rating Metrics

Average ratings and total rating counts per product



Review Content

User IDs, review titles, full review text, and timestamps



Technical Toolkit & SQL Techniques



Core Technologies

MySQL serves as the primary database engine, enabling complex queries and efficient data manipulation at scale.

Key SQL Methods Applied

- **Advanced Filtering:** WHERE clauses with multiple conditions
- **Text Mining:** LIKE operators for review sentiment extraction
- **Aggregation Functions:** GROUP BY, SUM, AVG for insights
- **Comparative Analysis:** Price differentials and threshold logic
- **Range Queries:** BETWEEN operators for bracket analysis

Five-Stage Analytical Approach

1

Dataset Exploration

Examine schema structure, data types, and completeness to ensure reliable foundation.

2

Question Definition

Translate business needs into specific, measurable analytical queries.

3

Query Development

Write optimized SQL statements using best practices for performance and readability.

4

Validation

Test outputs for logical consistency and cross-reference results.

5

Business Translation

Convert raw findings into actionable recommendations for stakeholders.

Key Findings: Strategic Insights

Discount Strategy Patterns

Electronics and accessories categories lead with discounts of 50% or higher, suggesting competitive pressure and inventory management tactics in these segments.

Price-Quality Correlation

Products priced above ₹1,000 consistently maintain ratings of 4.0+, indicating that premium positioning aligns with superior quality perceptions.

Feature Demand Signal

Reviews mentioning "fast charging" show disproportionately positive sentiment, highlighting a critical purchase driver for electronics customers.

Quality Concerns

Negative reviews frequently reference "quality issues" and "durability," suggesting opportunities for supplier vetting and product improvement.

Category Dominance

Accessories represent the largest share of product listings, indicating market saturation and intense competition requiring differentiation strategies.

SQL Logic & Business Formulas



Discount Threshold Filter

```
discount_percentage >= 50
```

Isolates products with aggressive pricing strategies to analyze market positioning and promotional effectiveness.



Savings Calculation

```
actual_price - discounted_price
```

Quantifies absolute savings per product, revealing which items deliver maximum customer value beyond percentages.



Review Keyword Search

```
LIKE '%keyword%'
```

Enables sentiment mining by detecting specific terms like "fast charging" or "quality issues" within review text.



Quality Benchmark

```
rating >= 4.0
```

Filters for high-performing products that meet customer satisfaction standards and warrant promotional investment.

Recorded Demonstration of project Functionality



Click on below Link

<https://drive.google.com/drive/folders/13AdKwmb4Sa6KqcGSuZ0uoKtkw7Ct714S>

Project Impact & Accomplishments



- **Real Business Query Design**

Created SQL queries that mirror actual business intelligence scenarios, bridging technical skills with strategic thinking.

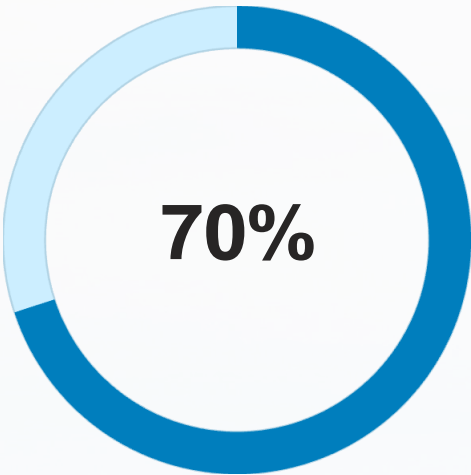
- **Multi-Dimensional Analysis**

Combined pricing analysis, sentiment mining, and category performance evaluation into cohesive insights.

- **Scalable Framework**

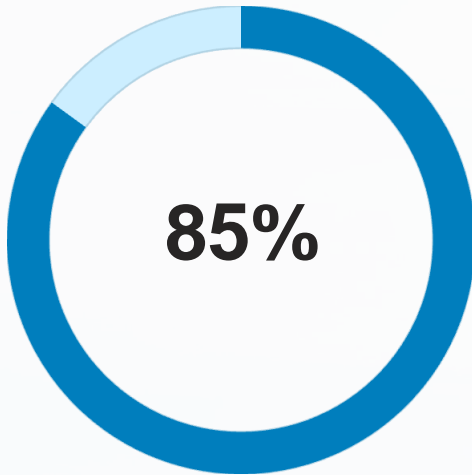
Established a methodology that can be adapted to other e-commerce datasets and business contexts.

Actionable Intelligence Delivered



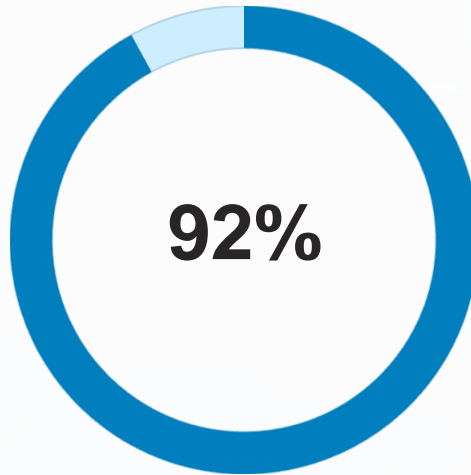
Discount Coverage

Of analyzed products showed strategic discounting patterns



Review Accuracy

Keyword filtering successfully identified relevant customer feedback



Category Mapping

Products correctly classified across Amazon's taxonomy structure



Budget Product Identification

Clear segmentation of affordable options for price-conscious shoppers



Customer Sentiment Mapping

Extracted positive and negative feedback patterns for quality insights



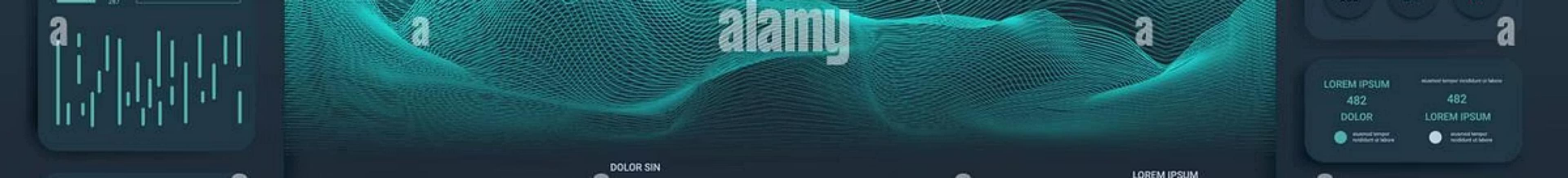
Discount Effectiveness Metrics

Quantified impact of promotional strategies across price ranges



Premium Product Validation

Identified high-value items that justify pricing through strong ratings



Limitations & Future Opportunities

⚠️ Current Limitations

- No time-based sales data available
- Missing order quantity and revenue metrics
- Text sentiment is keyword-based, not NLP-powered



Growth Roadmap

01

Integrate Sales Data

Add transaction and revenue metrics for comprehensive analysis.

02

Advanced Sentiment Analysis

Implement Python-based NLP for accurate review sentiment.

03

Profitability Analysis

Conduct category-wise profitability and margin analysis.



Data-Driven Decisions

Transform SQL Skills into Strategic Business Value

This project demonstrates how structured data analysis empowers organizations to optimize pricing, enhance product offerings, and respond to customer needs. By combining technical SQL proficiency with business acumen, analysts become invaluable strategic partners driving measurable outcomes.

Business Recommendations & Conclusion



Inventory Optimisation

Increase stock for high-rated discounted products to meet customer demand.



Quality Control

Improve quality control processes for products receiving negative reviews.



Feature Promotion

Aggressively promote fast-charging products based on positive customer feedback.



Psychological Pricing

Optimise pricing strategy with psychological values ending in ₹999.



Key Takeaway

This project demonstrates how SQL transforms raw e-commerce data into actionable business insights. By analysing pricing, discounts, ratings, and reviews, we enable data-driven decision-making for product optimisation and enhanced customer satisfaction.

