Executive Summary

This project presents an SQL-based exploratory and analytical study of **Zepto's product inventory and pricing dataset**. The objective was to understand product availability, pricing structure, discount strategies, stock health, and category-wise sales contribution using structured SQL queries.

A relational database named **zepto** was created, and a single table was populated using CSV import. The dataset contains product-level information such as category, product name, MRP, discount percentage, available quantity, discounted selling price, product weight, stock status, and quantity metrics.

Through a series of descriptive, aggregate, and window-function-based SQL queries, the analysis highlights key business insights including high-value categories, aggressive discounting patterns, stock availability ratios, and pricing efficiency at a per-gram level. These insights can help Zepto optimize inventory planning, pricing, discount strategies, and revenue generation.

Project Objectives

The main objectives of this project were:

- To explore and validate Zepto's product dataset using SQL
- To analyze pricing, discounts, and stock availability across categories
- To identify top-performing and underperforming product categories
- To evaluate sales contribution using discounted selling prices
- To derive actionable business insights and recommendations

Data Overview

The dataset consists of the following key attributes:

- **Category** – Product category (e.g., grocery, essentials, etc.)
- **Product Name** – Name of the item
- **MRP** – Maximum Retail Price
- **Discount Percentage** – Discount applied on MRP
- **Discounted Selling Price** – Final selling price after discount
- **Available Quantity** – Units available in stock
- **Weight in Grams** – Product weight
- **Out of Stock Flag** – Stock status indicator
- **Quantity** – Pack quantity or unit count

Initial data checks confirmed successful data import and availability for analysis.

Key Analysis & Insights

1. Product Availability & Inventory Health

- The **total available product quantity** was calculated to assess overall inventory strength.
- Products marked as **out of stock** were identified, helping highlight potential demandsupply gaps.
- Category-wise **stock availability ratio** revealed which categories maintain better inventory health and which require replenishment planning.

2. Category-Level Product Distribution

- Distinct product categories were identified, giving a clear view of assortment breadth.
- Category-wise **total available quantity** showed where inventory investment is concentrated.

3. Pricing & MRP Analysis

- **Average MRP per category** highlighted premium versus budget-oriented categories.
- The **top 5 most expensive products** were identified using window functions, useful for premium product tracking.

4. Discount Strategy Evaluation

- Products with **discounts greater than 10%** were extracted to study aggressive discounting.
- Category-wise **average discount percentage** revealed how discounts vary across segments.
- One category emerged with the **highest average discount**, indicating either competitive pricing or inventory clearance strategy.

5. Sales Contribution Analysis

- Total sales value was calculated as:

discountedSellingPrice × availableQuantity

- The **highest revenue-generating category** was identified using CTEs and ranking functions.
- The **top 3 categories contributing to total discounted sales** accounted for a major portion of revenue, showing strong category dominance.

6. Value-for-Money (Price per Gram)

- Price efficiency was evaluated by calculating **MRP per gram**.
- Products with the **lowest price per gram** provide better value and can be promoted to price-sensitive customers.

7. Deep Discount Identification

- Products where the **discounted price was below 80% of MRP** were identified, signaling heavy markdowns that may affect margins but boost volume.

Business Impact

This analysis demonstrates how SQL can be effectively used to:

- Support **data-driven inventory decisions**
- Optimize **pricing and discount strategies**
- Identify **high-revenue and high-risk categories**
- Improve **stock availability and demand planning**

The project reflects strong foundational skills in SQL, including aggregation, filtering, window functions, CTEs, and analytical reasoning.

Recommendations

1. Inventory Optimization

- Categories with **low stock availability ratios** should be prioritized for replenishment to prevent lost sales.
- Frequently out-of-stock products indicate high demand and should be stocked more aggressively.

2. Smarter Discount Strategy

- Categories with **very high average discounts** should be reviewed to ensure profitability is not compromised.
- Instead of flat discounts, consider **targeted or time-based promotions**.

3. Revenue Growth Focus

- Invest more in the **top 3 revenue-generating categories**, as they contribute most to overall sales.
- Bundle low-performing products with high-performing ones to increase movement.

4. Pricing Efficiency

- Promote products with **low price per gram** as “best value” items.
- Review premium-priced products to ensure they justify higher MRPs through brand or quality.

5. Advanced Analytics (Future Scope)

- Add **time-based sales data** to analyze trends and seasonality.
- Include **customer purchase behavior** for demand forecasting.

- Build a **dashboard (Power BI / Tableau)** using these SQL outputs for real-time business monitoring.

Conclusion

This Zepto SQL analysis project successfully translates raw product data into meaningful business insights. It showcases practical SQL skills and demonstrates how structured querying can support strategic decisions in e-commerce and quick-commerce platforms. With further enhancements and visualization, this analysis can evolve into a comprehensive decision-support system.