Retail Inventory and Sales Analysis Using SQL – Zepto Case Study

# 1. Project Title

Retail Inventory and Sales Optimization Using SQL-Based Data Analysis – A Case Study on Zepto

# 2. Objective

The objective of this project is to perform an in-depth analysis of Zepto’s retail inventory data using SQL to uncover insights related to stock availability, discount strategies, pricing effectiveness, and category-wise performance. This aims to enhance decision-making in inventory management and revenue optimization.

# 3. Tools & Technologies

• Database: PostgreSQL  
• Language: SQL  
• Dataset: Zepto product catalog and inventory (CSV)

# 4. Dataset Description

The dataset consists of various product attributes, including:  
- Serial\_id – Unique identifier  
- Category – Product category  
- Name – Product name  
- MRP – Maximum retail price  
- Discount\_percent – Discount applied  
- Available\_Quantity – Units available  
- Discounted\_Selling\_Price – Price after discount  
- WeightInGms – Product weight in grams  
- OutOfStock – Boolean indicating stock status  
- Quantity – Listed quantity for display

# 5. Methodology

## 5.1 Table Creation

A SQL table `zepto` was created to store the dataset:  
  
CREATE TABLE zepto (  
 Serial\_id SERIAL PRIMARY KEY,  
 catagory VARCHAR(120),  
 name VARCHAR(150) NOT NULL,  
 mrp NUMERIC(8,2),  
 discount\_percent NUMERIC(5,2),  
 availabile\_Quantity INTEGER,  
 discounted\_Selling\_price NUMERIC(8,2),  
 weightInGms INTEGER,  
 OutOfStock BOOLEAN,  
 Quantity INTEGER  
);

# 6. Data Exploration

- Total Record Count: Verified with SELECT COUNT(\*) FROM zepto;  
- Null Check: Ensured no missing values exist in critical fields.  
- Category Listing: Extracted unique product categories.  
- Stock Analysis: Counted how many products are out of stock vs. in stock.  
- Duplicate Products: Identified repeated product names across multiple rows.

# 7. Data Cleaning

- Removed Products with Zero MRP: DELETE FROM zepto WHERE mrp = 0;  
- Normalized Price: Converted paise to rupees using SQL update queries.

# 8. Business Insights & Analytical Queries

## Q1: Top 10 Best-Value Products by Discount

SELECT DISTINCT name, mrp, discount\_percent FROM zepto ORDER BY discount\_percent DESC LIMIT 10;

## Q2: High MRP but Out of Stock

SELECT DISTINCT name, mrp FROM zepto WHERE OutOfStock = TRUE AND mrp > 300 ORDER BY mrp DESC;

## Q3: Estimated Revenue per Category

SELECT DISTINCT catagory, SUM(Discounted\_Selling\_Price \* availabile\_Quantity) AS total\_revenue FROM zepto GROUP BY catagory ORDER BY total\_revenue DESC;

## Q4: Products with High Price but Low Discounts

SELECT DISTINCT name, mrp, discount\_percent FROM zepto WHERE mrp > 500 AND discount\_percent < 10 ORDER BY mrp DESC, discount\_percent DESC;

## Q5: Top 5 Categories by Average Discount

SELECT DISTINCT catagory, ROUND(AVG(discount\_percent), 2) AS avg\_dis FROM zepto GROUP BY catagory ORDER BY avg\_dis DESC LIMIT 5;

## Q6: Price per Gram for Products >100g

SELECT DISTINCT name, weightInGms, Discounted\_Selling\_Price, ROUND(Discounted\_Selling\_Price / weightInGms, 2) AS price\_per\_gram FROM zepto WHERE weightInGms >= 100 ORDER BY price\_per\_gram;

## Q7: Weight Category Classification

SELECT DISTINCT name, weightInGms, CASE WHEN weightInGms <= 500 THEN 'Low' WHEN weightInGms <= 1000 THEN 'Medium' ELSE 'Bulk' END AS weightCategory FROM zepto;

## Q8: Total Inventory Weight per Category

SELECT catagory, SUM(weightInGms \* availabile\_Quantity) AS total\_weight FROM zepto GROUP BY catagory ORDER BY total\_weight;

# 9. Results & Conclusion

This SQL-based analysis provided actionable insights into Zepto's inventory. Key findings included:  
- Best-value products and most discounted categories.  
- High-revenue generating categories.  
- High-priced products going out of stock.  
- Product classifications based on weight and price-per-gram value.  
  
These insights can assist Zepto’s operations team in managing stock better, adjusting pricing strategies, and improving customer satisfaction through targeted discounts and efficient logistics.

# 10. Future Scope

- Integration with visualization tools like Tableau or Power BI for dashboards.  
- Predictive analytics on product demand using time series models.  
- Incorporating supplier-level data for supply chain optimization.