

Zepto Inventory Analysis

1. Project Overview

In the **Quick-Commerce (Q-Commerce)** ecosystem, success is measured by the ability to balance ultra-fast delivery with optimized **Unit Economics**. This Inventory Health project targets Zepto's product catalogue to identify operational inefficiencies. The objective is to audit the **Digital Shelf**, optimize **SKU Rationalization**, and minimize **Revenue Leakage** caused by inventory stockouts and pricing misalignments.

2. Dataset Summary

The analysis is backed by a robust dataset capturing a 360-degree view of the transactional ecosystem:

- **Volume:** 3,700+ high-frequency transaction records across 9 specialized attributes.
- **Key Feature Sets:**
 - **Customer Demographics:** Granular data on Age, Gender, Location, and Subscription Status to drive personalized marketing.
 - **Purchase Architecture:** Item-level details including Category, Purchase Amount, Seasonality, Size, and Color.
 - **Behavioral Metrics:** Tracking of Discount Application, and Purchase Frequency to measure customer lifetime value (CLV).

Table: `zepto`

Columns:

Category	text
name	text
mrp	int
discountPercent	int
availableQuantity	int
discountedSellingPrice	int
weightInGms	int
outOfStock	text
quantity	int

3. Exploratory Data Analysis (EDA)

The initial phase focused on auditing the structural integrity of the catalogue to ensure it is "Machine-Ready" for Zepto's automated systems.

- **Catalogue Nullity Audit:** Conducted a comprehensive scan of critical attributes including category, MRP, and stock status to ensure 100% data density for the search and recommendation engines.

- **In-Stock vs. OOS Mix:** Evaluated the **Fill Rate** by calculating the ratio of available products to out-of-stock items, a baseline KPI for customer reliability.
- **SKU Breadth Analysis:** Identified unique product counts across the taxonomy to understand category dominance and variety.
- **Duplicate SKU Detection:** Audited product naming conventions to identify potential SKU proliferation or redundant entries that could skew inventory reporting.

4. Data Analysis (Business Transactions & Revenue Intelligence)

These 8 queries were architected to simulate real-world e-commerce transactions and drive strategic decision-making:

1. **Value-Tier Identification:** Isolating the top 10 products by **Markdown Depth** to identify "Hero Products" that drive app traffic and customer acquisition.

```
90      -- Q1. Find the top 10 best-value products based on the discount percentage.
91  •  SELECT DISTINCT name, mrp, discountPercent
92    FROM zepto_1
93    ORDER BY discountPercent DESC
94   LIMIT 10;
95
```

Result Grid | Filter Rows: Export: Wrap Cell Content: Fetch rows:

name	mrp	discountPercent
Dukes Waffy Chocolate Wafers	45	51
Dukes Waffy Orange Wafers	45	51
Dukes Waffy Strawberry Wafers	45	51
Ceres Foods Fish Mustard Instant Liquid Masala	220	50
Ceres Foods Laal Maas Instant Liquid Masala	220	50
Ceres Foods Nalli Nihari Instant Liquid Masala	220	50
Chef's Basket Durum Wheat Elbow Pasta	160	50
Chef's Basket Durum Wheat Fusilli Pasta	160	50
Chef's Basket Durum Wheat Penne Pasta	160	50

2. **Revenue Opportunity Gap:** Identifying high-ticket items ($MRP > ₹300$) currently marked as **Out-of-Stock**. This highlights immediate lost revenue and procurement failure points.

```

97      -- Q2.What are the Products with High MRP but Out of Stock.
98 •  SELECT DISTINCT name, mrp
99   FROM zepto_1
100  WHERE outofstock = "true" AND mrp > 300
101  ORDER BY mrp DESC ;
102

```

name	mrp
Patanjali Cow's Ghee	565
MamyPoko Pants Standard Diapers, Extra Large...	399
Aashirvaad Atta With Mutigrains	315
Everest Kashmiri Lal Chilli Powder	310

3. **Category-Wise GMV Projections:** Calculating the total estimated revenue of current on-hand inventory to assess capital allocation across categories.

```

104      -- Q3.Calculate Estimated Revenue for each category.
105 •  SELECT category, SUM(discountedSellingPrice*availableQuantity) AS revenue
106   FROM zepto_1
107  GROUP BY category
108  ORDER BY revenue DESC;
109
110

```

category	revenue
Cooking Essentials	337131
Munchies	337131
Personal Care	270849
Paan Corner	270849
Packaged Food	224385
Ice Cream & Desserts	224385
Chocolates & Candies	224385
Home & Cleaning	122661
Health & Hygiene	64180

4. **Premium Price Sensitivity Audit:** Filtering high-MRP items (>₹500) with low discounts (<10%) to identify products that may be experiencing **Velocity Stagnation**.

```

111      -- Q4. Find all products where MRP is greater than ₹500 and discount is less than 10%.
112 •   SELECT DISTINCT name, mrp, discountpercent
113   FROM zepto_1
114 WHERE mrp > 500 AND discountpercent < 10
115 ORDER BY mrp;
116
117

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	name	mrp	discountpercent
▶	Amul Pure Ghee (Pouch)	505	1
	Fresh Mutton Mince - Kheema (Halal)	515	3
	Amul Ghee (Tin)	520	2
	Surf Excel Matic Front Load Detergent Powder	530	4
	Kellogg's Real Almond & Honey Corn Flakes	550	5
	L'Oreal Casting Creme Ebony Black Hair Color	550	0
	L'Oreal Paris Casting Creme Gloss Hair Color (Da...)	550	0
	L'Oreal Paris Casting Creme Gloss Hair Color, M...	550	0
	L'Oreal Paris Excellence Creme Hair Color, 3.16 ...	550	0

5. **Markdown Efficiency Ranking:** Identifying the top 5 categories by average discount to evaluate which departments are leading the company's "Value" perception.

```

118      -- Q5. Identify the top 5 categories offering the highest average discount percentage.
119 •   SELECT category, ROUND(AVG(discountPercent), 2) AS avg_discount
120   FROM zepto_1
121 GROUP BY category
122 ORDER BY avg_discount DESC
123 LIMIT 5;
124
125

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows: |

	category	avg_discount
▶	Fruits & Vegetables	15.46
	Meats, Fish & Eggs	11.03
	Packaged Food	8.32
	Ice Cream & Desserts	8.32
	Chocolates & Candies	8.32

6. **Unit-Value Engineering:** Analyzing the price-per-gram for products over 100g to identify the most cost-effective options for bulk-purchasing consumers.

```

126      -- Q6. Find the price per gram for products above 100g and sort by best value.
127 •  SELECT DISTINCT name, discountedSellingPrice, weightInGms,
128     ROUND(discountedSellingPrice/weightInGms, 2) AS price_per_gram
129   FROM zepto_1
130   WHERE weightInGms > 100
131   ORDER BY price_per_gram;
132

```

	name	discountedSellingPrice	weightInGms	price_per_gram
▶	Aashirvaad Iodised Salt	19	1000	0.02
	Onion	21	1000	0.02
	Onion	57	3000	0.02
	Shubh kart - Nirmal sugandhi mogra wet dhoop ...	28	1160	0.02
	Tata Salt	24	1000	0.02
	Vicks Cough Drops Menthol	20	1160	0.02
	Baby Potato	16	500	0.03
	Beetroot	13	500	0.03
	Carrot	15	500	0.03

7. **Logistics Payload Tiering:** Segmenting the catalog into **Low, Medium, and Bulk** weight classes to optimize the allocation of the delivery fleet (Bikes vs. Electric Three-wheelers).

```

134      -- Q7. Group the products into categories like Low, Medium, Bulk.
135 •  SELECT DISTINCT name, weightInGms,
136   CASE WHEN weightInGms < 1000 THEN "Low"
137     WHEN weightInGms < 5000 THEN "Medium"
138     ELSE "Bulk"
139   END AS weight_category
140   FROM zepto_1;
141
142

```

	name	weightInGms	weight_category
▶	Onion	1000	Medium
	Tomato Hybrid	1000	Medium
	Sponge Gourd	300	Low
	Tender Coconut	58	Low
	Coriander Leaves	100	Low
	Ladies Finger	250	Low
	Zama Organics Carrot	250	Low
	Safal Frozen - Mixed Vegetables	1000	Medium
	Zama Organics Sweet Potato	250	Low
	Potato	1000	Medium
	Lemon	200	Low

8. Cumulative Inventory Payload: Calculating the total weight per category to assist in Dark Store Floor-loading and warehouse storage planning.

```

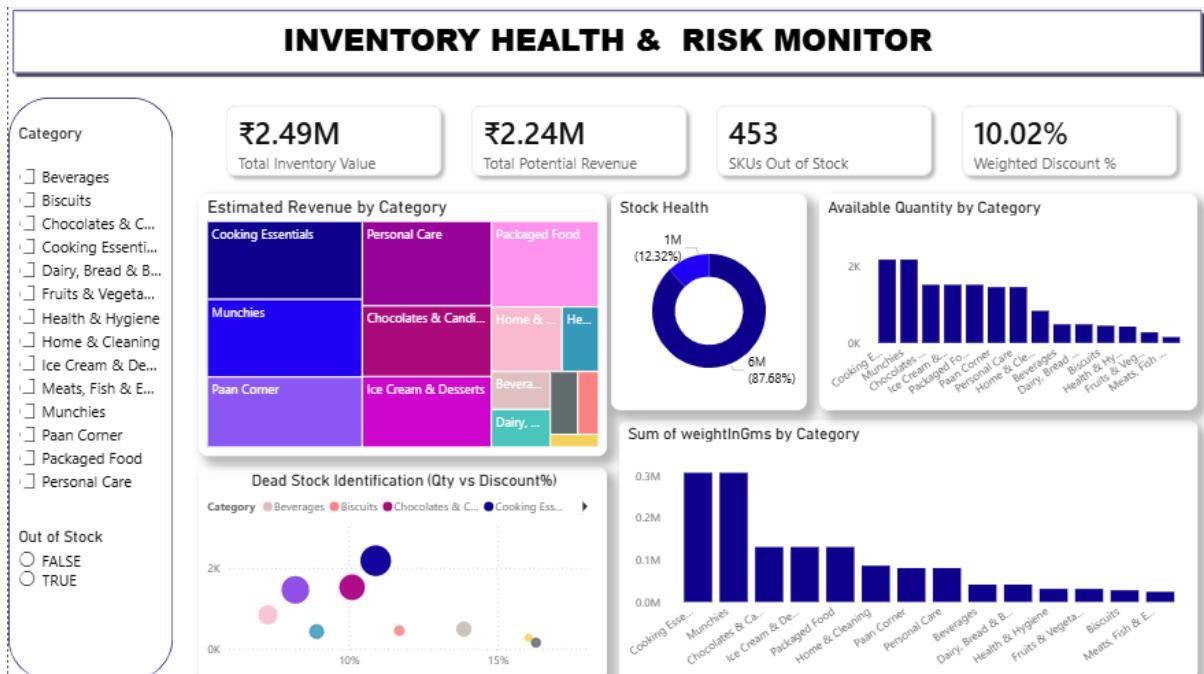
143    -- Q8.What is the Total Inventory Weight Per Category.
144 •   SELECT category, SUM(weightInGms*availableQuantity) AS total_weight
145     FROM zepto_1
146     GROUP BY category
147     ORDER BY total_weight DESC;
148

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	category	total_weight		
▶	Cooking Essentials	1404326		
	Munchies	1404326		
	Packaged Food	490797		
	Ice Cream & Desserts	490797		
	Chocolates & Candies	490797		
	Home & Cleaning	373161		
	Personal Care	348187		
	Paan Corner	348187		
	Beverages	143735		
	Dairy, Bread & Batter	143735		
	Health & Hygiene	142904		
	Fruits & Vegetables	01704		

5. Dashboard in Power BI

Finally, we built an interactive dashboard in Power BI to present insights visually.



6. Business Recommendations

- **Recover Lost GMV:** Prioritize the immediate procurement of "High-Ticket OOS" items identified in the analysis. Restocking items above ₹300 can significantly increase **Average Order Value (AOV)** with minimal marketing spend.
- **Aggressive Markdown Realignment:** Categories currently offering the highest average discounts should be evaluated for **Margin Contribution**. If these are not "Loss Leaders," the markdown strategy should be moderated to improve profitability.
- **Last-Mile Fleet Optimization:** Use the **Payload Tiering** data to dynamic-route deliveries. "Bulk" category orders should be automatically flagged for vehicles with higher payload capacities to reduce rider fatigue and delivery delays.
- **Inventory Velocity Re-engagement:** For high-MRP items with low discounts that are not moving, implement a tiered discount strategy (e.g., 5-10% increase in markdown) to unlock liquidity from slow-moving stock.
- **Assortment Rationalization:** Consolidate redundant SKUs identified in the duplication audit to simplify the user interface and reduce picking errors at the fulfillment centers.