Retail Sales Intelligence

SectionA_G10 | Retail Analytics & Business Intelligence

Transforming Retail Data into Strategic Intelligence

PROJECT DETAILS

SECTOR

Retail Analytics / BI

INSTITUTE

Newton School of Technology

FACULTY

Satyaki Sir

DATE

February 2026

TEAM MEMBERS

- Suhaani Garg
- Vetriselvan R
- Aryan Vibhuti
- Arun Giri
- Sanath Waraikar
- Divyansh Rathore

2 Context & Problem Statement

BUSINESS CONTEXT

Retail chains rely on data-driven insights to optimize inventory planning, pricing strategy, and store-level performance. Sales performance is shaped by multiple factors — product category, pricing, outlet size, location tier, and product visibility.

Without centralized analytical dashboards, retail managers cannot reliably identify which variables drive revenue generation and where investment should be directed.

👤 **Store Managers**

📊 **Category Heads**

🏢 **C-Suite Exec**

Key Decision-Makers

PROBLEM STATEMENT

Retail management lacks a unified analytical framework to evaluate outlet performance, product contribution, and geographic distribution at a granularity sufficient for evidence-based capital allocation, format strategy, and category management decisions.

PROJECT OBJECTIVE

- Analyze item-level & outlet-level sales data across ₹18.59M revenue
- Identify patterns in product categories, outlet formats, and location tiers
- Build an interactive Excel dashboard supporting retail strategy decisions
- Deliver evidence-based recommendations for format expansion & optimization

3 Data Engineering — Source to Insight

🗃️ DATA SOURCE

- BigMart Sales Dataset
- Kaggle (Public)
- bigmart-sales-data

📏 DATASET SIZE

- 8,523 Rows
- 12 Key Columns
- Cross-sectional period

🏷️ COVERAGE

- 10 Outlet Stores
- 16 Product Categories
- 3 City Tiers

DATA CLEANING STEPS

Null Handling
Imputed Item_Weight nulls with category mean; filled missing Outlet_Size with 'Unknown'

Label Standardization
Normalized Item_Fat_Content: merged 'LF' → 'Low Fat', 'reg' → 'Regular'

Numeric Validation
Validated Item_Visibility, Item_MRP and Item_Outlet_Sales — no negative values found

DATA DICTIONARY (KEY COLUMNS)

Column	Type	Description
Item_Type	Cat.	Product category (16 types)
Item_Fat_Content	Cat.	Low Fat / Regular
Item_MRP	Float	Maximum Retail Price (₹)
Outlet_Type	Cat.	Store format (4 types)
Outlet_Location_Type	Cat.	City tier: Tier 1/2/3
Item_Outlet_Sales	Float	Target — sales revenue (₹)

4 KPI & Metrics Framework

₹18.59M
Total Gross Sales
Across 10 outlets

₹2,181
Avg Sales per Item
Overall efficiency

₹140.99
Avg Price Point
Average MRP

12.86 kg
Avg Item Weight
Product metric

KPI DEFINITIONS & BUSINESS LINKAGE

KPI	Formula	Business Question Answered
Total Gross Sales	SUM(Item_Outlet_Sales)	What is the overall revenue baseline?
Revenue Share %	Format Sales / Total × 100	Which formats/tiers drive the most value?
Avg Sales per Item	Format Sales / Item Count	Which format is most operationally efficient?
Location Tier Share	Tier Sales / Total × 100	Where should we expand geographically?
Fat Content Mix	Segment Sales / Total × 100	What product health trend patterns exist?

WHY THESE KPIs

01 Directly map to revenue strategy

02 Answer where to invest capital

03 Identify format efficiency gaps

04 Guide geographic expansion

05 Reveal consumer preferences

5 Key Insights from Exploratory Data Analysis

01

~70%

Supermarket Type1 Revenue Share

Dominates total sales but not the most efficient format per item sold

02

41.1%

Tier 3 City Revenue Contribution

Emerging markets outperform metro Tier 1 cities (24.1%) — counter-intuitive

03

₹2,820K

Fruits & Veg — #1 Category

Snack Foods close 2nd at ₹2,733K; essential goods anchor revenue base

04

64%

Low-Fat Product Revenue Share

Significant consumer preference signal or category composition effect

05

₹3,694

Type3 Avg Sales per Item

Highest efficiency vs Type1 (₹2,316) — scale ≠ efficiency paradox

06

₹340

Grocery Store Avg per Item

Dramatically below estate mean — format requires strategic review

6 Advanced Analysis

OUTLET EFFICIENCY ANALYSIS

Revenue per item sold disaggregated from scale:

- Type3: ₹3,694 avg (HIGHEST)
- Type1: ₹2,316 avg
- Type2: ₹1,995 avg
- Grocery: ₹340 avg (LOWEST)

Scale paradox confirmed — Type1 dominates total revenue but it ranks 2nd in per-item efficiency.

GEOGRAPHIC DEMAND ANALYSIS

Counter-intuitive tier hierarchy:

- Tier 3: 41.1% (₹7.64M) — HIGHEST
- Tier 2: 34.8% (₹6.47M)
- Tier 1: 24.1% (₹4.48M) — LOWEST

Explanation: Lower competition, higher outlet dependency, and favorable cost economics in Tier 3 markets.

PRODUCT PORTFOLIO SEGMENTATION

Category revenue contribution (top 5):

- Fruits & Vegetables: ₹2.82M (15.2%)
- Snack Foods: ₹2.73M (14.7%)
- Household: ₹2.06M (11.1%)
- Frozen Foods: ₹1.83M (9.8%)
- Dairy: ₹1.52M (8.2%)

Conclusion: Essential goods dominate, reducing upside but ensuring stability.

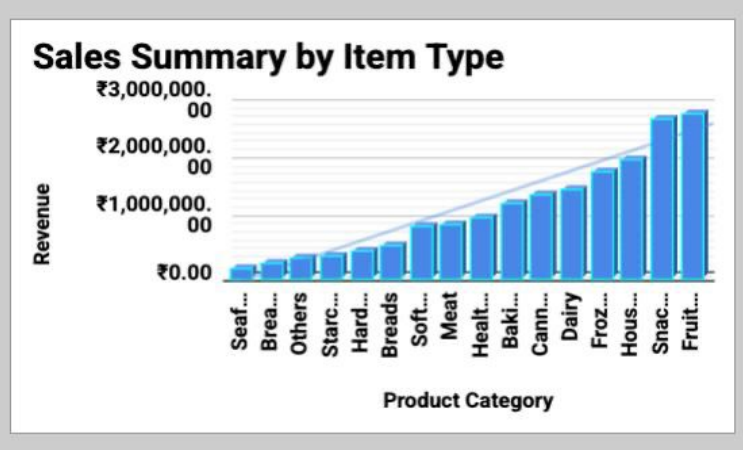
FAT CONTENT SEGMENTATION

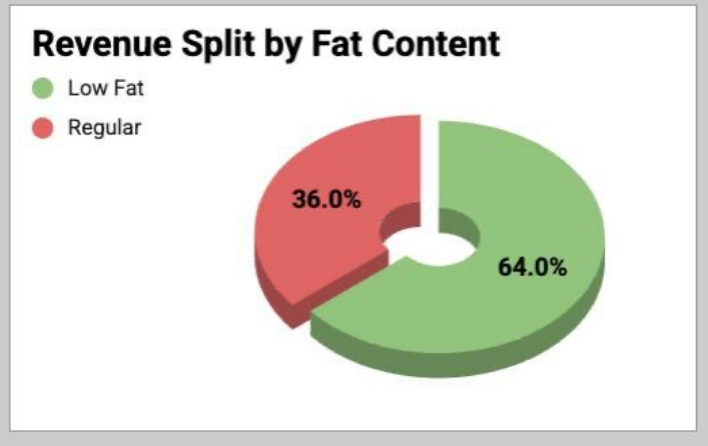
Revenue split by fat classification:

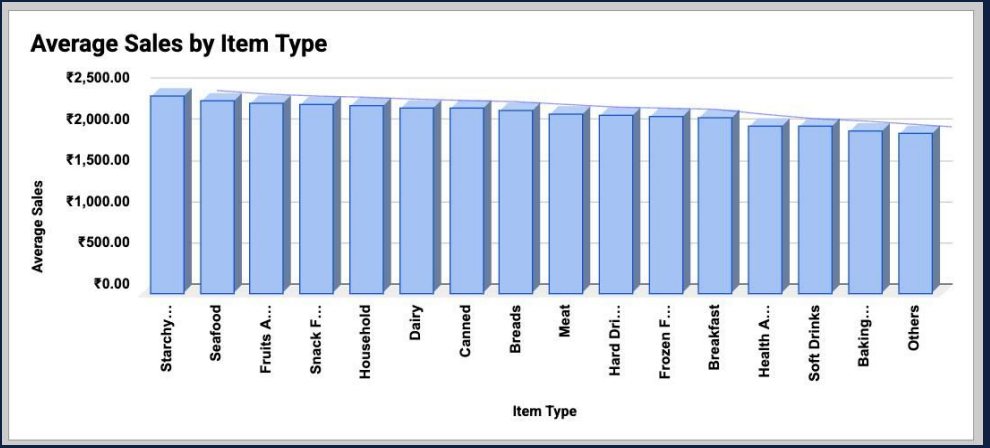
- Low Fat: 64.0% (₹11.90M)
- Regular: 36.0% (₹6.69M)

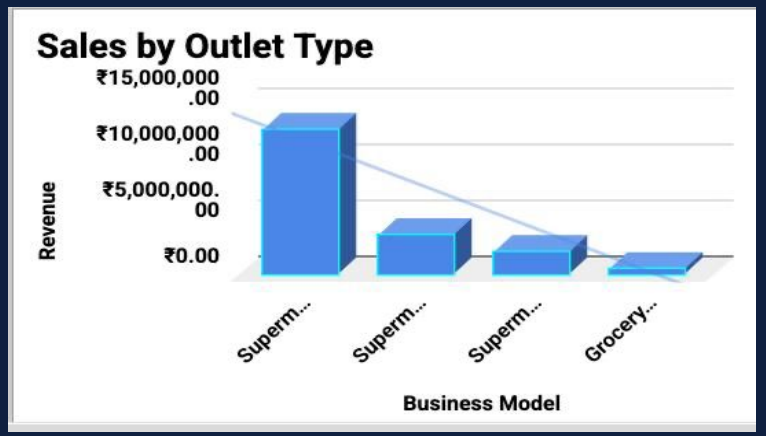
Two explanations tested: (1) Genuine consumer health preference, (2) Category composition (fruits, veg naturally classified as low-fat). Both factors contribute.

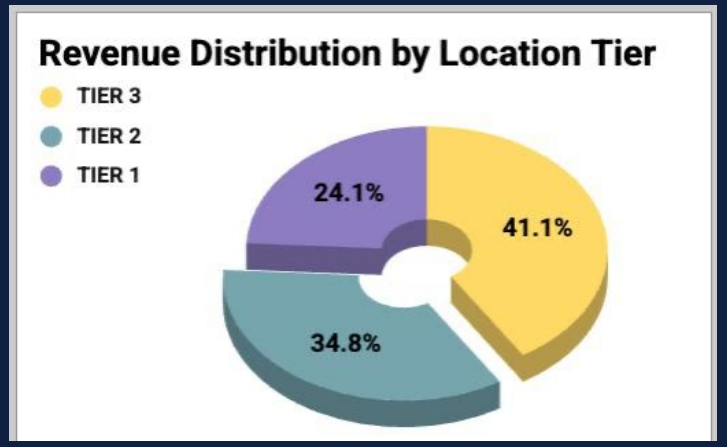
7 Dashboard Walkthrough

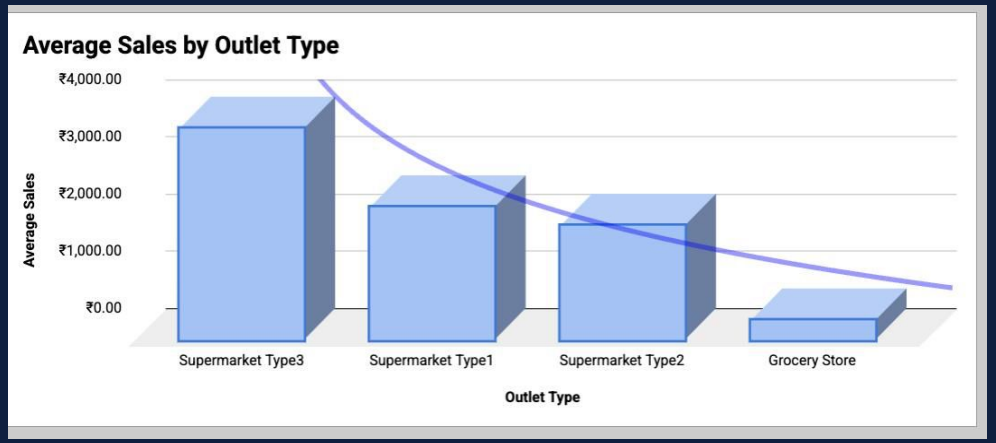












Interactive slicers available: Outlet Type | Location Tier | Item Type | Fat Content

8 Strategic Recommendations

R1

HIGH

Expand in Tier 3 Markets

Linked to Insight 02

Prioritize new Supermarket openings in high-performing Tier 3 sub-markets. Lower competitive density and favorable cost economics make Tier 3 the highest-conviction expansion priority.

R2

HIGH

Replicate Type3 Efficiency Model

Linked to Insight 05

Commission an operational audit of Supermarket Type3 to codify its ₹3,694 per-item efficiency. Apply learnings to Type1 — a 5% improvement on Type1's base would yield ~₹645K additional revenue.

R3

MEDIUM

Protect Essential Goods Core

Linked to Insights 03, 04

Strengthen supply chain for Fruits & Vegetables and Snack Foods — the ₹5.55M combined revenue anchor. Any supply disruption would have outsized estate-level impact.

R4

MEDIUM

Improve Pricing Discipline in Type1

Linked to Insights 01, 05

Type1 generates 69.5% of revenue but has the lowest per-item efficiency. Investigate assortment mix, pricing strategy, and markdown frequency to close the efficiency gap with Type3.

R5

LOW-MED

Selective Premium Category Diversification

Linked to Insight 03

Test 2–3 premium or lifestyle categories in Type3 and Type1 formats. Current revenue is anchored in low-margin essentials — diversification can improve mix and reduce category concentration risk.

9 Impact & Value — The Business Case

Tier 3 Expansion

₹929K

Projected Uplift

5% revenue uplift

Confidence: Moderate

Type1 Efficiency

₹391K–651K

Projected Uplift

3–5% efficiency gain

Confidence: Mod-Low

Category Diversification

₹372K

Projected Uplift

2% premium mix uplift

Confidence: Low

Combined Scenario

₹1.7M–2.0M

Projected Uplift

All levers executed

Confidence: Moderate

🎯 **Combined scenario projects a 9–11% revenue uplift on ₹18.59M base = ₹1.7M–₹2.0M additional annual revenue**

📊 Data-Driven

All recommendations backed by ₹18.59M of actual transactional evidence

🎯 Clear Strategy

Tier 3 expansion + Type3 efficiency replication + portfolio diversification

✅ Low Execution Risk

Phased approach with performance gates before material capital commitment

10 Limitations & Next Steps

⚠ DATA LIMITATIONS

No Profit Margin Data

Revenue analysis cannot assess profitability. High-revenue formats may underperform on margin.

No Customer Demographics

Fat content skew and category preferences cannot be attributed to behavioral vs. compositional factors.

No Time-Series Data

Cross-sectional only — revenue trajectories and seasonal patterns cannot be assessed.

No Competitive Intelligence

Tier 3 outperformance hypotheses are plausible but unvalidated without market density data.

No Promotional Data

Revenue spikes from promotions cannot be isolated from underlying demand trends.

🚀 ANALYTICS ROADMAP

HORIZON 1 — 0 to 6 Months

- Outlet-level revenue tracking per KPI
- Basket-level transaction data capture
- Category-level margin tracking

HORIZON 2 — 6 to 18 Months

- Customer segmentation framework
- Tier 3 sub-market demand modeling
- Time-series revenue forecasting
- Migrate to Power BI dashboard

HORIZON 3 — 18 to 36 Months

- ML demand forecasting by category
- Real-time performance monitoring
- Price optimization engine

RETAIL SALES

Thank You

We're grateful for your time and attention

SectionA_G10 | Retail Analytics & Business Intelligence | Newton School of Technology | February 2026

Suhaani Garg | Vetriselvan R | Aryan Vibhuti | Arun Giri | Sanath Waraikar | Divyansh Rathore

₹18,591,125.41 | 8,523 Records | 10 Outlets | 16 Categories | 3 City Tiers