# CAPSTONE PROJECT REPORT

**Title:** An Analytical Study of Retail Profitability Drivers Using Transaction-Level Data

**Sector:** Retail & Commerce

**Team Details:**


- Guru Manohar Gupta

- Dadi Dinesh

- Mouli Srivastava

- Ravleen Singh

- Aarsh Bhatnagar

- Vidit Sachan

**Institute:** Newton School Technology

# 1. EXECUTIVE SUMMARY

## Problem

Retail businesses often achieve strong sales volumes but struggle with consistent profitability. Analysis of 9,994 transactions (2014–2017) shows an overall profit margin of 12.47%, yet 18.72% of transactions are loss-making. A significant portion of these losses remain hidden within aggregated revenue figures, making it difficult for management to identify profit leakage drivers related to discounting, category performance, and regional variation.

## Approach

This project analyzes transaction-level retail data using Google Sheets as the primary analytics platform.

The workflow included:
- Data cleaning and standardization
- Creation of derived KPIs such as Profit Margin and Loss Flag
- Pivot table analysis across Category, Region, Segment, and Discount levels
- Development of an interactive dashboard for performance monitoring

The objective was to identify where profits are generated versus where margins are eroded and to build a structured decision-support framework.

## Key Insights

- 18.72% of transactions generate negative profit.
- 60.93% of loss-making transactions occur when discounts exceed 30%, identifying discounting as the primary profit leakage driver.
- Furniture operates at a low profit margin of 2.49%, significantly below Technology (17.40%) and Office Supplies (17.04%).
- Profit contribution is uneven across categories and regions, indicating performance concentration rather than balanced growth.

### Key Recommendations

• Implement discount control thresholds, particularly above the 30% range.
• Re-evaluate pricing and promotional strategies for the Furniture category.
• Prioritize investment and promotion in higher-margin categories.
• Use the dashboard for continuous monitoring of margin risk and category performance.

## 2. Sector and Business Context

### Sector Overview

The Retail & Commerce sector operates on high-volume transactional sales across multiple product categories and regions. Profitability depends not only on revenue generation but also on effective pricing, discount control, and category performance management.

### Current Challenges

Retail businesses often face margin pressure due to excessive discounting and uneven category contribution. In the analyzed dataset (2014–2017), 18.72% of transactions are loss-making, indicating hidden profit leakage despite overall positive sales performance.

### Why This Problem Was Chosen

The dataset shows clear variation in profitability across discount levels and product categories, with a significant concentration of losses in high-discount transactions. This highlights the need for KPI-driven analysis and dashboard monitoring to improve profitability stability and decision-making.

## 3. Problem Statement

Retail businesses often experience a mismatch between revenue growth and profitability. High sales volumes may coexist with declining margins due to excessive discounting, inefficient product mixes, or region-specific performance issues. Without transaction-level analysis, these profit leakages remain hidden.

**Core Problem Statement:**
**How can retail management improve profitability and ensure sustainable revenue growth by analyzing transaction-level sales data, discount behavior, and regional performance?**

The objective is to identify where profits are generated versus where they are eroded and to support strategic decision-making using data-driven insights.

## 4. Project Objectives

The key objectives of this capstone project are:

- Analyze overall sales and profit performance across products, regions, and customer segments
- Evaluate the impact of discounting strategies on profitability
- Identify loss-making transactions and underperforming categories
- Compare regional and category-level performance to detect inefficiencies
- Generate actionable insights to support managerial decision-making

## 5. Dataset Description

The analysis uses the **Superstore Sales Dataset**, which contains detailed transaction-level data from a U.S. retail chain.

**Dataset Characteristics:**

- **Nature:** Retail / Commerce transactional data
- **Granularity:** Row-level (each row represents a product sold within an order)
- **Size:** ~10,000 rows (approximately)
- **Time Period:** 2014–2017

**Key Columns:**

- Order details: Order ID, Order Date, Ship Date, Ship Mode
- Customer details: Customer ID, Customer Name, Segment
- Location details: City, State, Region
- Product details: Category, Sub-Category, Product Name
- Financial metrics: Sales, Quantity, Discount, Profit

## 6. Data Cleaning and Preparation

The raw dataset contained inconsistencies and required structured cleaning. All preparations were performed in **Google Sheets**.

**Cleaning Steps Performed:**

- Removed duplicate records using Order ID validation
- Standardized text fields (Category, Sub-Category, Product Names, Customer Names)
- Converted Order Date and Ship Date into consistent date formats
- Rounded Sales and Profit values to two decimal places
- Ensured consistent numeric formatting across financial fields

**Derived Columns Created:**

- Order Year and Order Month from Order Date
- Shipping Delay calculated as Ship Date minus Order Date
- Discount Percentage derived from discount values
- Loss Flag to identify negative-profit transactions

These steps ensure analytical accuracy and dashboard readiness.

## 7. Key Performance Indicators (KPIs)

The following KPIs were designed and maintained:

- **Total Sales** – Overall revenue generated
- **Total Profit** – Net profitability
- **Profit Margin** – Profit as a percentage of sales
- **Count of Orders** – Computed total orders in the dataset (cleaned)

## 8. Advanced Analysis

The analysis followed a structured exploratory workflow:

- Pivot tables summarize sales and profit by Category, Region, Segment, and Year, delay of orders, Profit.
- Trend analysis examined changes in sales and profit over time
- Comparative analysis evaluated category-wise and region-wise performance
- Discount-to-profit analysis assessed pricing effectiveness
- Filters and slicers enabled interactive drill-down analysis

## 9. Dashboard Design

A comprehensive dashboard was built in **Google Sheets**.

**Dashboard Components:**

- KPI summary cards (Sales, Profit, Profit Margin, Orders Count)
- Category-wise and Sub-Category wise comparison charts
- Sales and profit trend charts
- Discount vs Profit analysis visuals
- Most selling category per region
- Interactive filters for Region, Category, and Segment, Shipping speed, sub-category, loss-severity.


# 10. Insights Summary

- Profit margin erosion strongly linked to high discount bands (>30%)
- Furniture contributes high revenue but low profitability ratio
- Technology provides highest profit efficiency
- West region is the most stable profit contributor
- Central region shows higher concentration of loss transactions
- Consumer segment drives majority of total sales

- Corporate segment shows better margin stability
- Loss-making transactions cluster within specific sub-categories
- Shipping delay shows minimal correlation with profit decline
- Discount control could significantly improve overall margin performance

# 11. Recommendations

- Introduce discount caps to prevent margin erosion
- Focus investments on high-margin categories
- Reevaluate pricing strategies for Furniture
- Replicate best practices from high-performing regions
- Use dashboards for continuous performance monitoring

# 12. Impact Estimation

The analysis of 9,994 transactions shows an overall profit margin of 12.47%, with 18.72% of transactions generating negative profit.

## Cost Savings

60.93% of all loss-making transactions occur when discounts exceed 30%.

Controlling high-discount transactions could reduce nearly half of current loss exposure and improve overall profit margin by an estimated **3–5 percentage points**, depending on implementation discipline.

## Profitability Optimization

Furniture generates only 2.49% profit margin compared to:

• Office Supplies – 17.04%
• Technology – 17.40%

Improving Furniture pricing or discount structure could significantly increase overall profitability without increasing sales volume.

### Risk Reduction

With nearly 1 in 5 transactions being loss-making, implementing margin thresholds and discount monitoring can materially reduce profit volatility and protect earnings stability.

## 12. Limitations

Dataset includes only final Profit values; no detailed cost breakdown is available for deeper margin analysis.

• No return or refund data is provided, so post-sale profitability adjustments cannot be evaluated.

• Data is limited to the period 2014–2017 and may not reflect current market conditions.

• No inventory or supply chain variables are available; operational analysis is limited to Shipping Delay only.

• Customer data is limited to segments; no repeat purchase or lifetime value information is available.

• External factors such as competition, economic changes, or promotional campaigns are not captured in the dataset.

## 13. Future Scope

• Incorporate detailed cost components to perform deeper margin and cost-driver analysis.

• Integrate return and refund data to evaluate true net profitability.

• Develop predictive models to forecast sales and profit trends based on discount and category patterns.

• Perform customer-level analysis such as repeat purchase behavior and lifetime value estimation.

• Expand the dashboard into a real-time monitoring system for continuous profitability tracking.

## 14. Conclusion

This project analyzed 9,994 retail transactions (2014–2017) using structured data cleaning, KPI computation, pivot analysis, and a Google Sheets dashboard.

The findings show an overall profit margin of 12.47%, with 18.72% of transactions generating losses. Notably, 60.93% of loss-making transactions occur when discounts exceed 30%, highlighting discount control as a key profitability lever. Category analysis further revealed that Furniture operates at a significantly lower margin (2.49%) compared to Technology (17.40%) and Office Supplies (17.04%).

The dashboard provides a practical decision-support framework to monitor discount risk, optimize category performance, and improve profitability stability.

## 15. Team Contribution Details

| Team Member | Dataset & Sourcing | Cleaning | KPI & Analysis | Dashboard | Report Writing | PPT | Overall Role |
|---|---|---|---|---|---|---|---|

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Guru Manohar Gupta** | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | Team Lead |
| **Dadi Dinesh** | ☑ | ☑ | ☑ | ☑ | | | Data Analyst |
| **Mouli Srivasthava** | | ☑ | ☑ | ☑ | | ☑ | Data Analyst |
| **Ravleen Singh** | ☑ | ☑ | ☑ | | | ☑ | Data Analyst |
| **Aarsh Bhatnagar** | ☑ | ☑ | | | ☑ | ☑ | Data Analyst |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| **Vidit Sachan** | | | | | | ☑ | Data Analyst |

## Declaration

We declare that this project is our original work and has been completed in accordance with academic integrity guidelines.

**Signatures:**

Guru Manohar Gupta :

Dadi Dinesh :

Aarsh Bhatnagar :

Vidit Sachan :

Ravleen Singh :

Mouli Srivastava :

**Date: 18-02-2026**