



BDM CAPSTONE PROJECT (SEP'25)

Optimizing Inventory and Service Offerings for a Multi-Service Retail Shop

SNEHA SARKAR (23F2004880)

About the Business

- ▶ **Cafe.com** is a multi-service retail shop established by **Mr. Biplab Ghosh** in 2018.
- ▶ The shop has been operating for seven years near **Suri Bus Stand, Birbhum, West Bengal**, a high-traffic and convenient location. It serves both local residents and daily commuters, offering a wide range of essential services and goods.
- ▶ The business started as a basic Xerox and printing service. Over time, it has expanded into: Photo printing, ID card support (Aadhaar, voter ID, license), Stationery items.
- ▶ The shop has been bootstrapped organically, based on the immediate needs of the local community. It maintains a steady and loyal customer base.
- ▶ The current goal is to streamline inventory and services using data-driven strategies to improve: Profitability & Operational efficiency.



Problem Statement & its Reasons

Problem Statements

1. Overstocking and Capital Blockage

- ▶ An excess stock of Low Demand items
- ▶ Due to this capital gets tied up, reducing liquidity.
- ▶ Leads to lower profitability due to slow-moving inventory.

2. Lack of Insights on Product Marketing

- ▶ There is no clear understanding of which items/services generate the most revenue.
- ▶ Decisions are based on assumptions rather than data.
- ▶ This leads to inefficient product marketing and sales.

3. Absence of Data Driven Demand Forecasting

- ▶ There is no system to predict demand or optimize inventory.
- ▶ Leads to missed market sales trends.
- ▶ This causes both overstocking and stock-outs.

Reasons

1. Internal Reasons

- ▶ **No systematic inventory management** or tracking of stock movement.
- ▶ **Sales not digitized**, making it difficult to monitor trends.
- ▶ **Owner relies on intuition**, not data, for purchase decisions.
- ▶ **No analytics tools** to identify best-selling or low-performing items/services.
- ▶ **Inefficient stock replenishment** due to limited visibility of demand.

2. External Reasons

- ▶ **Unpredictable customer preferences** in a walk-in retail environment.
- ▶ **Seasonal fluctuations** in demand for services and products.
- ▶ **Competition** from nearby shops offering similar services.
- ▶ **Rising supplier costs** affecting margins.
- ▶ **Limited shop space**, increasing the impact of overstocking.

Data Collection, Cleaning and Preprocessing

Column	Data Type	Description	Example	Missing Value (%)
date	String	Date of sell	14 oct 24	0
time	String	Time of sell	6:10 pm	0
customer	String	Customer category	student	14.83%
item	String	Name of SKU	kamlinnote book+apsar apencil+plastic-cover	0
quantity	String	Count of SKU	2+4+1	0
price/unit	String	Cost price of each SKU	50+5+5	0
total sell	String	Total Cost price	100+20+5	0
payment	String	Mode of payment	PayTM	8.91%

Uncleaned Data

1. Data Cleaning

- ▶ Imputed missing values with most frequent.
- ▶ Ensured all columns had valid values for analysis.

2. Standardization

- ▶ Expanded data values which were separated by symbols (eg +,.)
- ▶ Converted dates to **datetime** and prices/quantities to **numeric**.
- ▶ Normalized similar item names to a single value (e.g., all pencils → **"Pencil"**).

3. Time Formatting

- ▶ Time in AM/PM converted into 3 analysis-friendly bins: **Morning, Afternoon, Evening**.

4. Item Categorization

- ▶ Grouped items into broader service categories: **Document Services, Printing, Photo, Digital Services, Academic Supplies, Stationery, Office Supplies**.

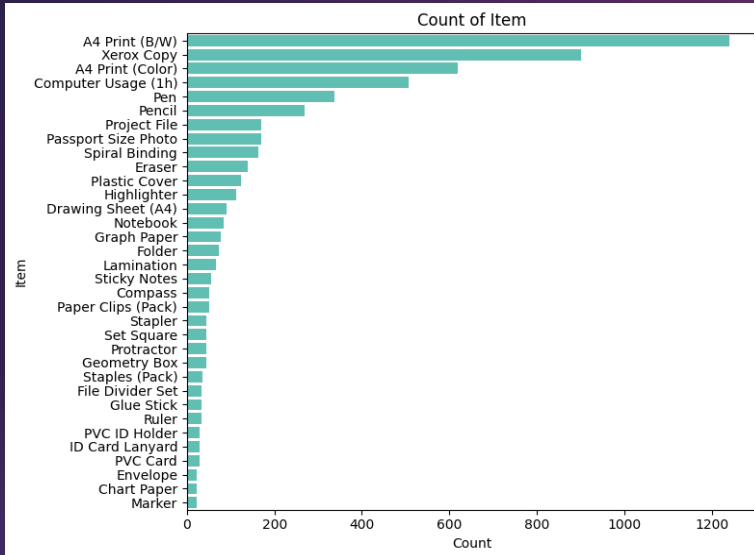
5. Payment Method Consolidation

- ▶ Payment variants were merged (UPI, PayTM, GooglePay → **UPI**).
- ▶ Final methods: **Cash, UPI, Card**.

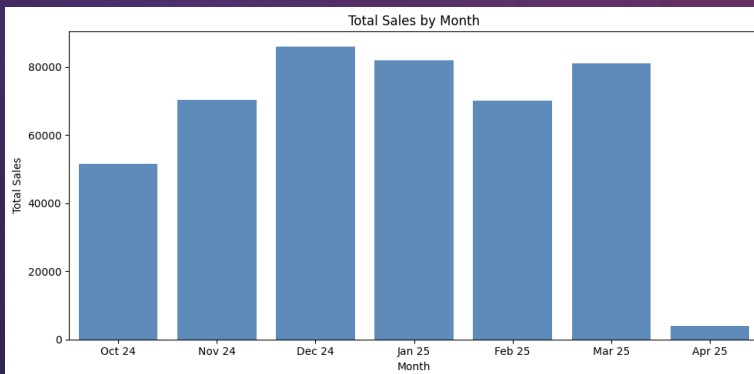
6. Final Structuring (Pre-processed Dataset)

- ▶ Organized columns in a logical order:
Date → Time → Item → Category → Quantity → Unit Price → Total Price → Payment Method.

Descriptive Statistics



SKU Count Distribution



Sales by Month Distribution

Overall Dataset Summary

- ▶ **Time Period:** Oct 2024 – Apr 2025
- ▶ **Total Transactions:** 1,740
- ▶ **Number of SKUs:** 34
- ▶ **Total Revenue:** ₹4,42,780.00
- ▶ **Average Transaction Value (ATV):** ₹76.53

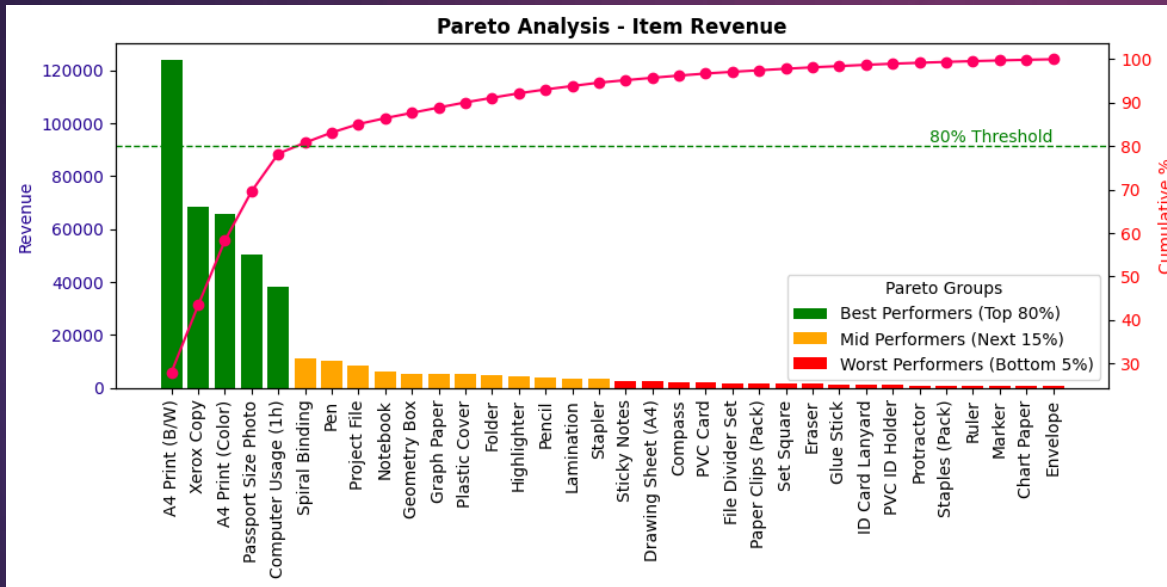
Central Tendency of Numerical Columns

- ▶ Generally, **1–2 units** per transaction.
- ▶ Maximum demand was observed in **Printing/Xerox** services.
- ▶ **Low-value (₹):** Pens, pencils, photocopies, basic stationery (₹2–₹20).
- ▶ **Mid-value (₹):** Printing, color prints, digital services (₹10–₹50).
- ▶ **High-value (₹):** Lamination, PVC cards, accessories (₹50–₹150+).

Distribution Insights

- ▶ Majority of sales come from **low-price, high-frequency** services.
- ▶ High-value items selling in **lower volume** generate **higher per-unit revenue**.
- ▶ Sales show **seasonal spikes**, with certain months showing increased demand.
- ▶ Most transactions occur during the **5 PM – 10 PM** time window.
- ▶ **Students are the largest customer segment.**
- ▶ **Cash & UPI** payment methods are most frequent.

1. 80/20 Revenue Impact Assessment (Pareto Analysis)



80/20 Pareto Analysis

Explanation

Pareto analysis uncovers which SKUs contribute most to the revenue. This helps prioritize the high-revenue generating ones for marketing and resource allocation.

Results

- ▶ A small group of **5 service SKUs** (A4 B/W print, color print, Xerox, passport photos, computer usage) generates **~80% of total revenue**.
- ▶ These top items are all **services**, not products.
- ▶ Mid performers (12 SKUs) contribute **~15%** which are mostly stationery/items.
- ▶ Low performers add **very little**, taking up space but not much revenue.

Findings

- ▶ **Services are the main strengths**, not retail products.
- ▶ Focus should be on **improving services quality, promotion, and pricing**.
- ▶ Sales of **Mid-performers** can be boosted by trying to **bunch them along with high performers**
- ▶ Optimizing resources around top SKUs can **boost profitability**.

2. Market Basket Analysis (Cross-Selling Insights)



Cross-Selling Insights

Explanation

This helps to spot services that are frequently bought together. We use Association rule mining from **mlxtend python library**. Metrics such as Support, Confidence, Lift point out chances for cross-selling. That guides bundled promotions, service placement ideas.

Results

- ▶ Black & white printing acts as a **central hub**, mid performers like folders, staplers, graph paper, are frequently bought together.
- ▶ Color prints show **very few cross-sells**, means these are usually not purchased together with other items.
- ▶ Network graph shows **strong frequent pairings** around A4 B/W print.

Findings

- ▶ This might mean people tend to **impulse-purchase** when getting **B/W prints**.
- ▶ **Cross-selling opportunities** exist: place folders, clips, staplers near print area.
- ▶ Color print customers are **purpose-driven**, not impulse buyers.
- ▶ Bundles or counterside accessories can **increase basket value**.

3. Cyclic Demand Decomposition (Seasonality Analysis)

Explanation

Preprocessed time-series data breaks demand into Trend, Seasonal, Residual parts. This reveals recurring cycles, and improves forecasting accuracy.

Results

- ▶ Demand shows **regular repeating patterns** (seasonal cycles).
- ▶ We observe a strong, stable demand from **Oct–Feb**.
- ▶ There is a noticeable drop in **Mar–Apr**, possibly seasonal or external.
- ▶ Residuals are small → model explains demand well.

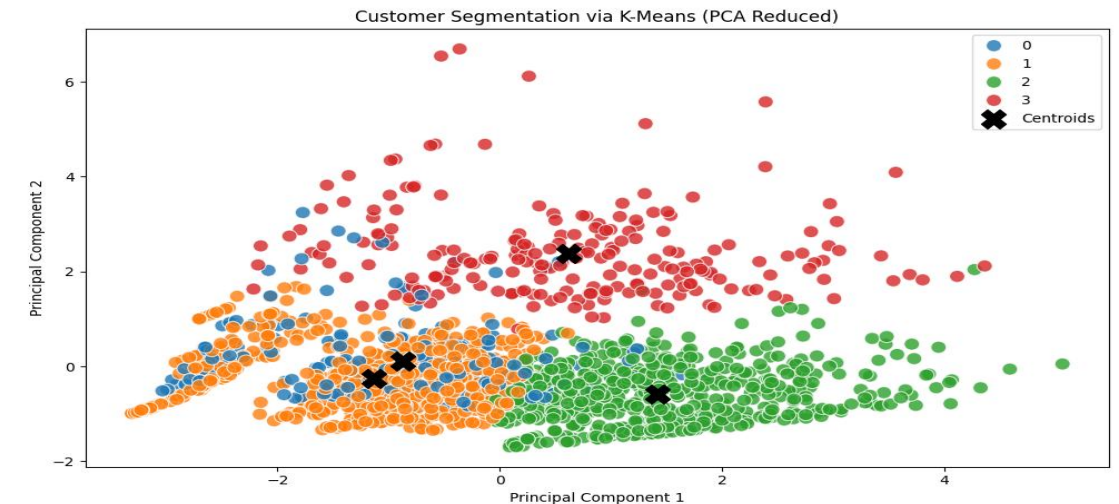
Findings

- ▶ Demand is mostly **predictable with clear cycles**.
- ▶ Seasonal variations can guide **stocking and promotions**.
- ▶ The **Mar–Apr decline** requires investigation (season, exams, events).
- ▶ Planning should target **peak months more heavily**.

4. K-Means Customer Segmentation (Behavioral Segmentation)

Explanation

Customers were classified into classes based on their purchase amount, spending habits. It uncovers different segments for targeted marketing and retention efforts.



Findings

- ▶ Enables personalized offers for **high-value and regular buyers**.
- ▶ Helps improve **inventory planning** for each group.
- ▶ Targeted marketing can **boost retention and revenue**.

5. Service Utilization Dynamics (Trend Mining)

Explanation

Used to understand the trends that follow the sales in the market. It spots peak demand periods, changing customer preferences. That aids operational planning.

Results

- ▶ **Printing services** have the highest and most stable revenue.
- ▶ **Photo services** show fluctuating but have meaningful contributions to revenue.
- ▶ **Digital services** bring lower but steady income.

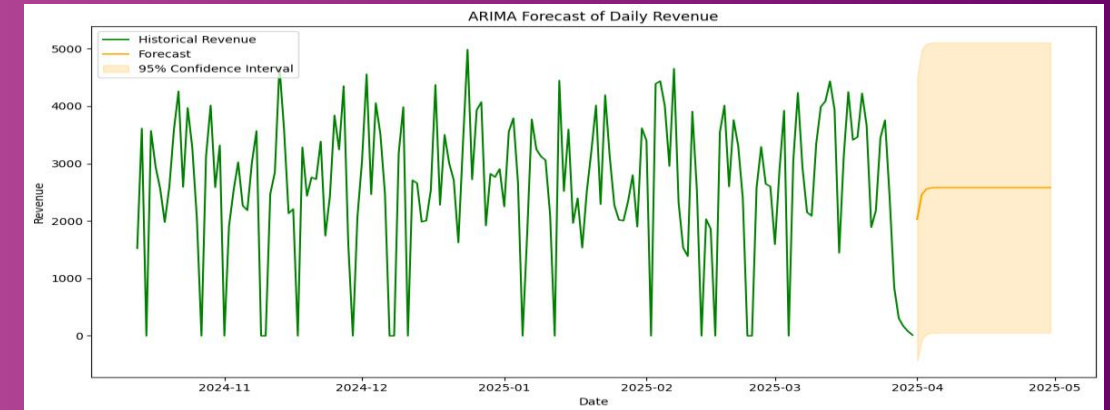
Findings

- ▶ Demand is mostly **predictable with clear cycles**.
- ▶ Printing is the **core revenue driver**, with consistent demand.
- ▶ Photo services are **seasonal/occasion-based** but profitable.
- ▶ Digital services are **niche** with potential for growth.
- ▶ Prioritize printing efficiency; promote photo & digital strategically.

6. ARIMA Time Series Forecasting (Demand Forecasting)

Explanation

Using the time-based data the model predicts future service needs. It picks up trends, seasonality. All that supports better inventory, resource planning.



Findings

- ▶ Revenue is **stable overall but unpredictable short-term**.
- ▶ Operations should focus on **managing variability**, not expecting growth.
- ▶ Emphasize **core stable services** for steady income.
- ▶ Forecasting helps with **moderate planning**, not exact predictions.

Recommendations and Conclusions

1. Focus on High-Performing Services

- ▶ Prioritize top revenue drivers (printing, Xerox, passport photos, computer usage). Ensure stock availability, improve counters, and offer bulk deals.

2. Reduce Low-Performing Inventory

- ▶ Phase out slow-moving stationery items. Use shelf space for fast-selling, service-linked products.

3. Use Cross-Selling Opportunities

- ▶ Place folders, clips, staplers near print counters. Offer small bundles; train staff for gentle add-on suggestions.

4. Align with Seasonal Demand

- ▶ Stock up before high-demand months; reduce during slow periods. Adjust staffing and run promotions in low months.

5. Targeted Customer Segmentation

- ▶ Create purchase perks for high-value customers. Offer student-friendly combos and digital payment rewards.

6. Improve Forecasting & Planning

- ▶ Maintain a simple 30-day forecast for restocking and staffing. Use Excel/Python dashboards updated with daily sales.

7. Shift to Digital Operations

- ▶ Adopt low-cost POS systems; digitize transactions. Use automated sheets/dashboards for real-time insights.

8. Aesthetic redesign of the shop

- ▶ The business should focus on improving the look and feel of the shop, like adding modern designs and posters, for customer retention.