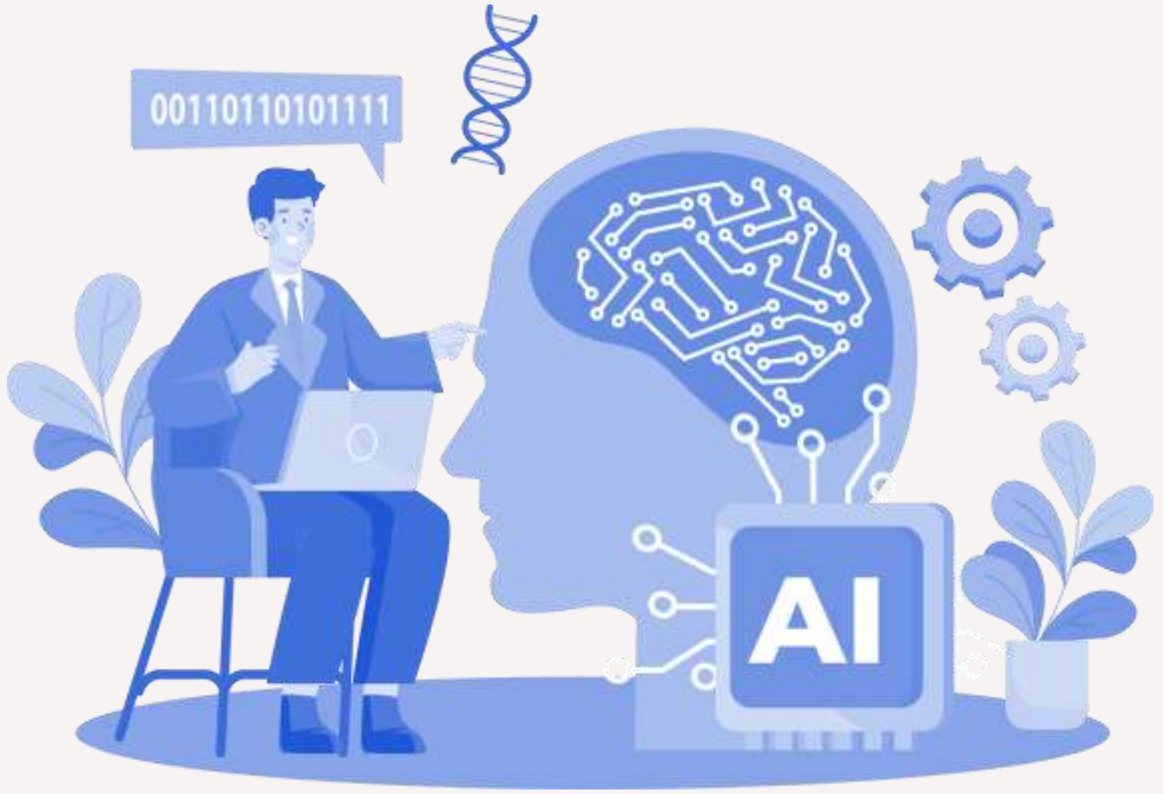
TEAM 2025103

Introduction

Problem Statement: Utilising optimisation techniques and data analytics, recommend a modified budget allocation for different marketing levers for upcoming months



Objectives

- Performance Driver Analysis
- Impact Analysis on Marketing ROI
- Optimizing Marketing Spending

Key Focus

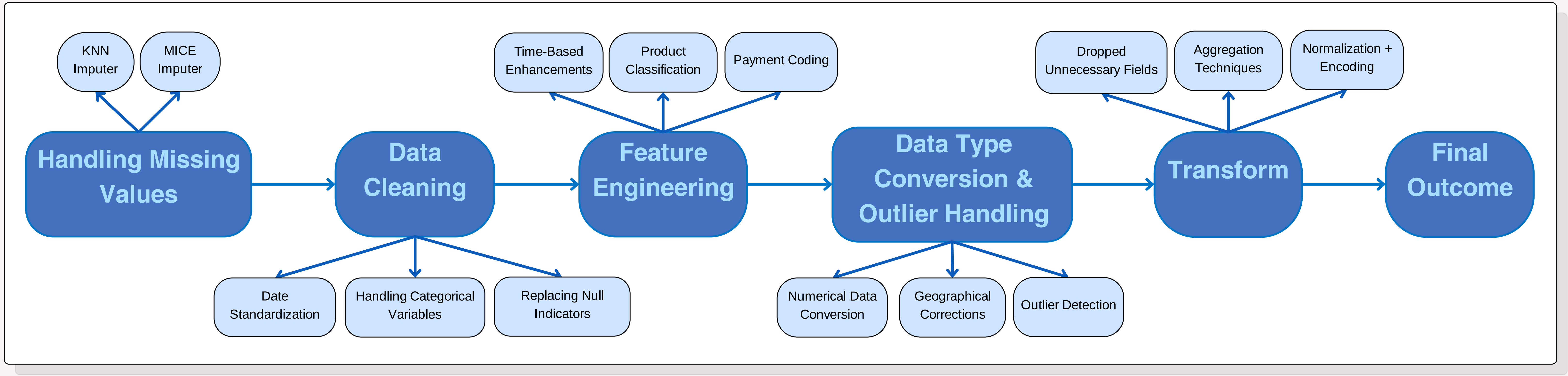
Identifying Key Performance Indicators (KPIs)

Measuring Effectiveness of Marketing Channels

Optimizing the Marketing Budget

Data Preprocessing

Given the nature of ElectroMart's business problem — evaluating marketing effectiveness, sales trends, and operational efficiencies — data preprocessing played a crucial role in preparing a structured and reliable dataset for further analysis.



Benefits of Pre-Processing of Data

- Formatted and cleaned dataset
- Optimized for marketing effectiveness
- Enhanced delivery performance analysis
- Identified key business insights

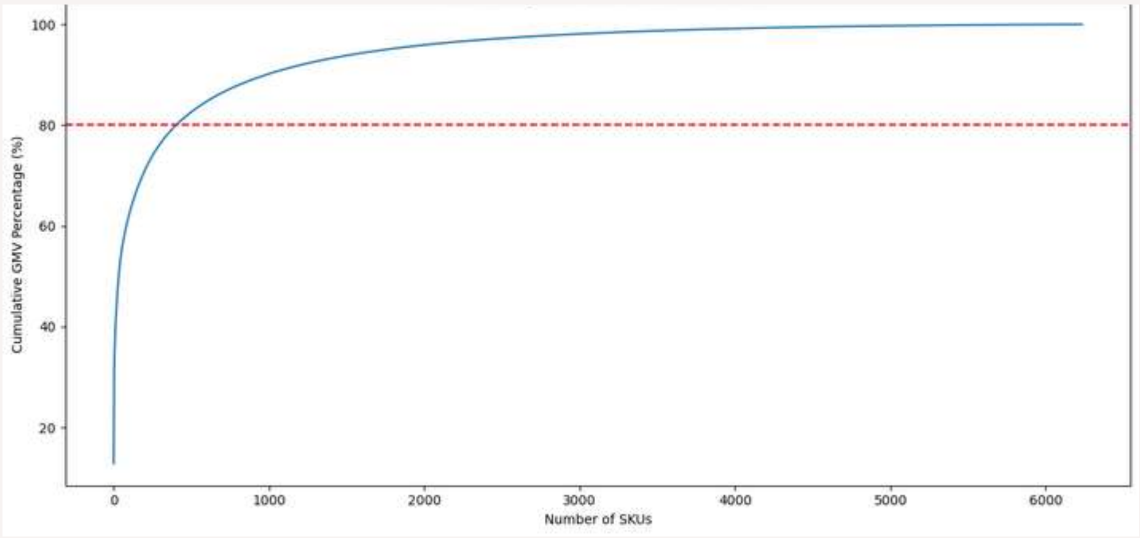
- Combined multiple datasets
- Utilized sophisticated imputation methods
- Applied advanced transformation techniques
- Supported data-driven decision-making

Exploratory Data Analysis

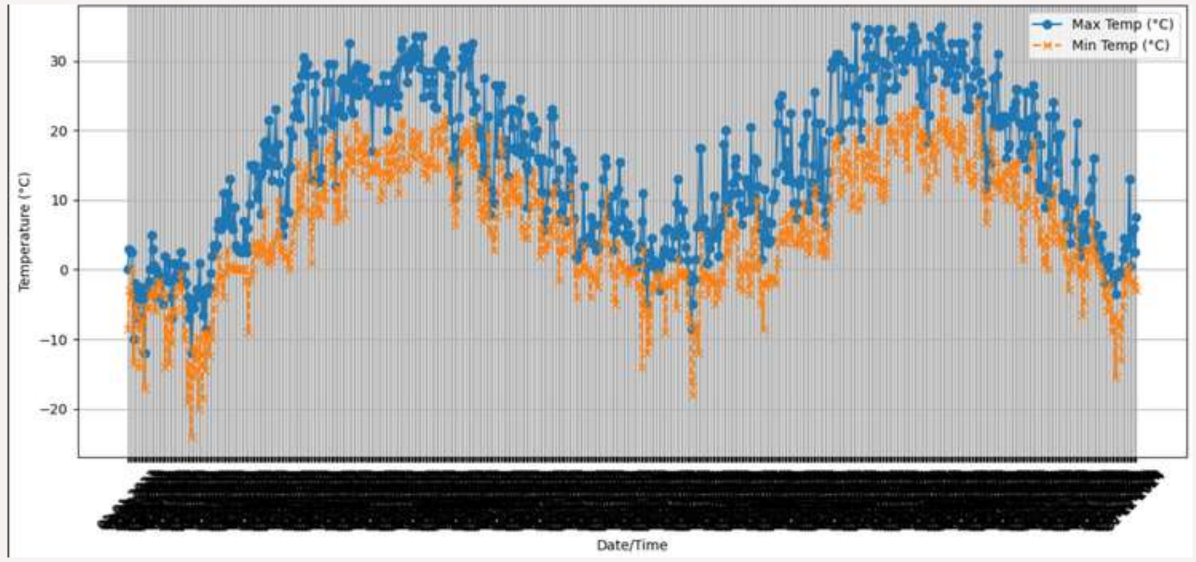
Exploratory Data Analysis (EDA) assists ElectroMart in discovering patterns and relationships for informed decision-making. It compares **revenue (total GMV)** with **ad expenses** to determine **marketing budget optimization** and analyzes effect of promotional discounts on revenues. **Customer satisfaction (NPS)** and **payment methods** are also compared with effects on revenue and order fulfillment.

Measures of EDA

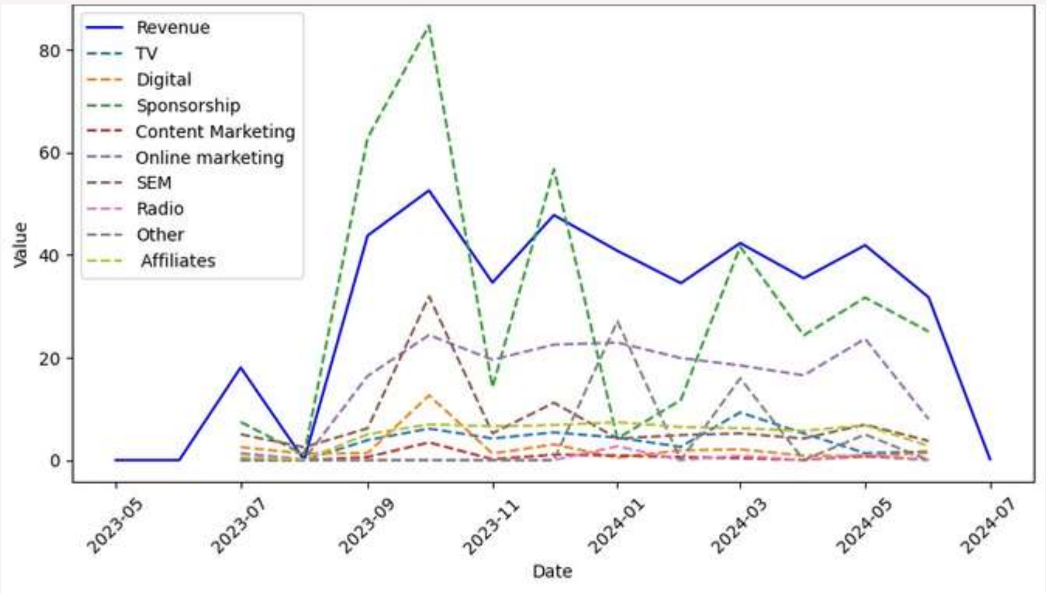
- Delivery Performance
- Seasonal Patterns
- Category Contribution
- Strategic Insights



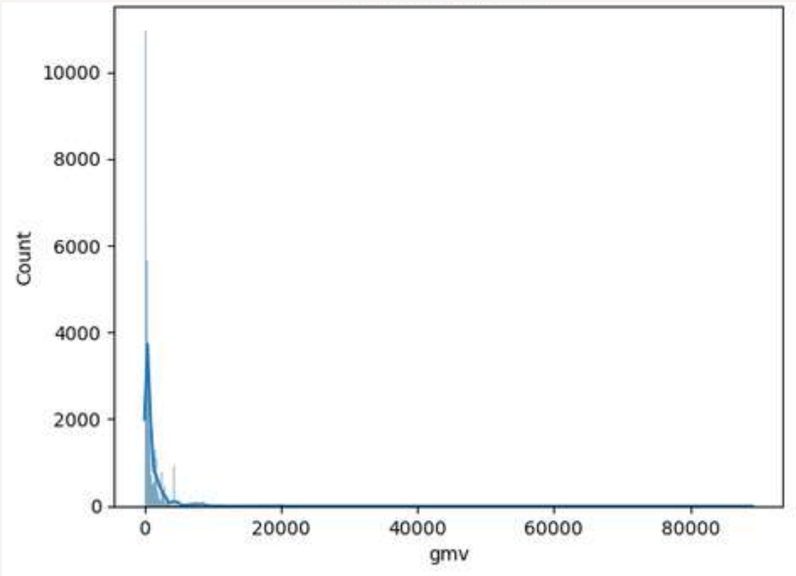
Pareto Analysis of SKU-total GMV Contribution



Max and Min Temperature Over Time (2023-2024) after Imputation

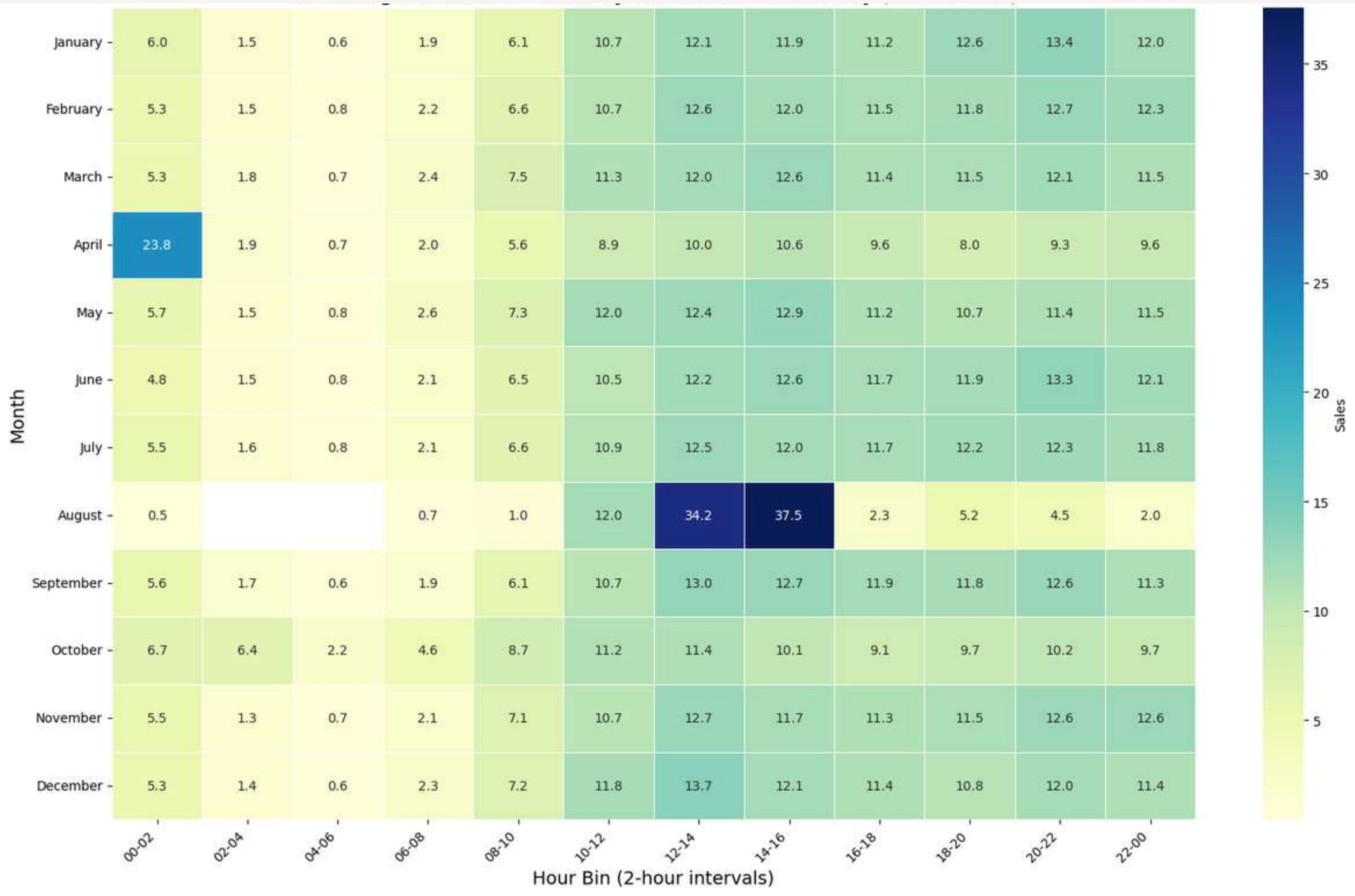


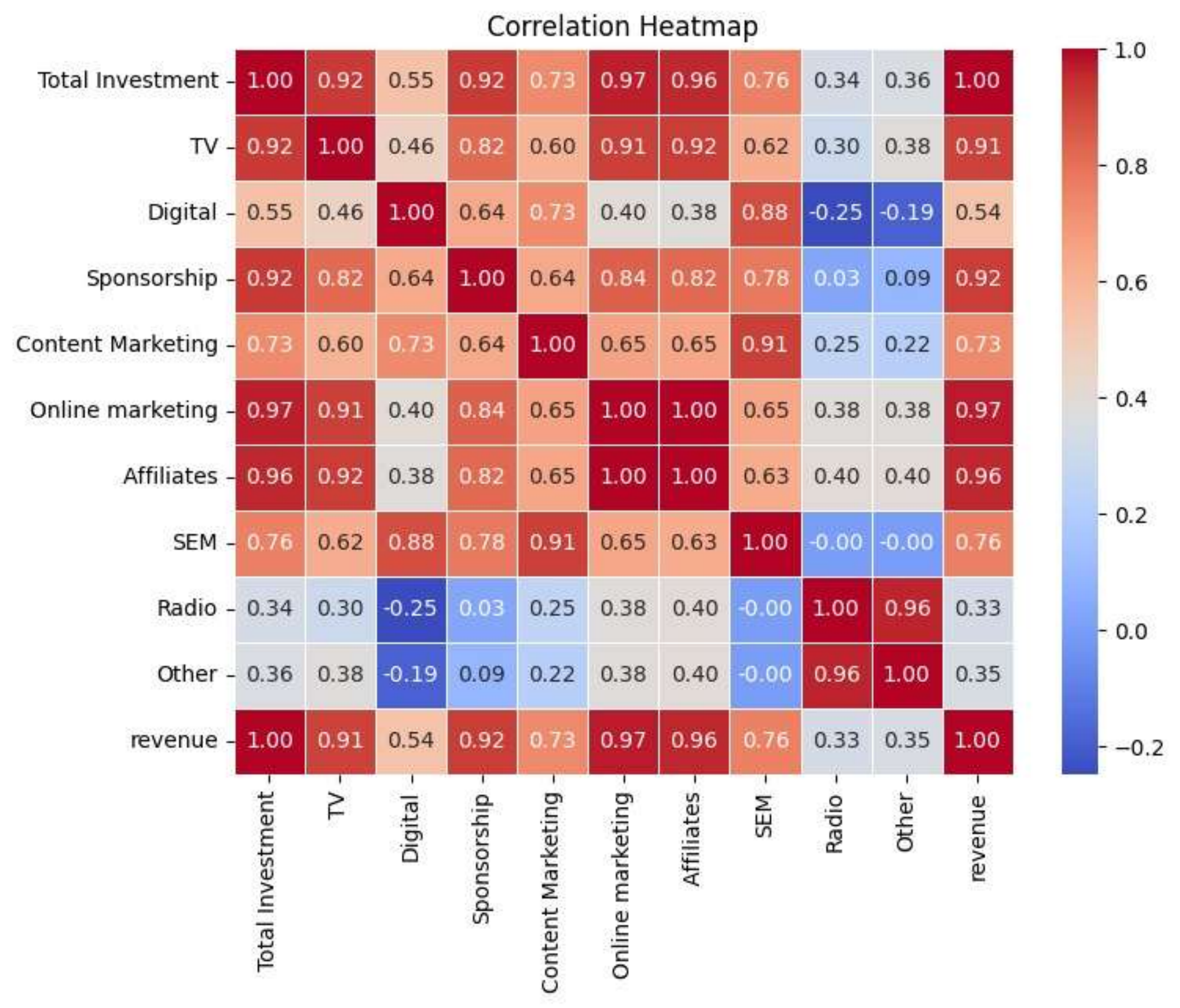
Monthly Revenue and Media Investments



Frequency Distribution of Gross Merchandise Value (GMV)

Exploratory Data Analysis





Percentage sales Distribution by Month and Time of Day (2-hour Bins)

Correlation Heatmap of Revenue with various marketing channels

Key Performance Indicators (KPIs)

KPIs/KRAs /KRIs

Customer Lifetime Value (CLTV) (12378 CAD)

Avg Order Value × No. of Orders × Retention Period

Retention Rate (9.4%)

number of customers at end of period-number of new customers/number of customers at start of period *100

Discount Seeker (9126)

(number of customers when discount>50%)/Total customers *100)

Order Fulfillment Efficiency (81.6%)

((order with deliverybdays<p.sla)/total orders *100)

Discount Rate Threshold (46%)

cumulative_gmv ≥ (0.90 × total_gmv)

Monthly Warehouse Inefficiency

(orders with (b_days-procurement_sla)>0)/no. of total orders

Rain/Snow Delay (21%)

((avg Sales on precipitation days - avg Sales on non-precipitation days) / avg Sales on non-precipitation days) × 100%

Change in Weekend Sales (31%)

((Weekend avg. sales - Weekday avg. sales) / Weekday avg. sales) × 100%

Key Performance Indicators (KPIs)

Sales and Revenue KPIs

- **Delivery Timeliness:** (Number of on-time deliveries / Total number of deliveries) × 100%.
- First Order Rate : (Number of first-time orders / Total number of orders) × 100%.
- Repeat Order Rate: (Number of repeat orders / Total number of orders) × 100%.
- Total Spends per Customer (CLTV - Customer Lifetime Value): Formula: Average Order Value × Number of Orders × Retention Period.
- High Spender: Customers with spending >75th percentile of all customer spending.
- Low Spender: Customers with spending <=75th percentile of all customer spending

Environmental Factors

- Correlation of Weather and Sales: Pearson correlation coefficient between weather metrics and sales figures.
- Impact of Rain/Snow on Delay: ((Sales on precipitation days - Sales on non-precipitation days) / Sales on non-precipitation days) × 100%.
- Change in Sales on Weekends : Sales on weekends are 19% higher campared to weekdays = ((Weekend avg. sales - Weekday avg. sales) / Weekday avg. sales) × 100%.
- Average Mean Temperature Impact: The relationship between mean temperature and purchasing behavior. Measured through correlation coefficients and regression analysis. This supports seasonal merchandising strategies.

Discount and Pricing KPIs

- Discount Seeker: (Number of discounted purchases / Total purchases by customer) × 100%.
- Discount Rate Threshold: where cumulative_gmv ≥ (0.90 × total_gmv)
- Customer Retention Rate: 7.7% of customers continue to make purchases: Repeat Customers/Total Unique Customers × 100%. This measures loyalty and satisfaction.
- Warehouse Efficiency: (Total orders processed / Labor hours) or (Items picked per hour). This indicates operational productivity.

Risk Metrics (KRIs)

- Delivery Risk: (Delayed deliveries / Total deliveries) × 100%.
- Low NPS Risk: (Detractors / Total respondents) × 100%.
- High Churn Risk: Historical churn rate or predictive model based on behavior patterns. This helps identify at-risk customers for intervention.

Operational KRAs

- Customer Order Frequency: (Total Orders in a Month)/(Total Unique Customers in that Month)
- SLA Compliance: (Number of deliveries meeting SLA / Total deliveries) × 100%. This measures operational reliability.
- Order Cancellation Rate: Number of canceled orders / Total number of orders
- Variance of Procurement SLA : Standard deviation of 53.46.

Optimizing Marketing Budget

Using a logarithmic model, we found the importance of each marketing channel. And based on that, we solved an optimization problem to find optimal budget of each channel.

METHODOLOGY

Logarithmic Modeling

Marketing expenditure often exhibits decreasing marginal effectiveness, making linear regression unsuitable for modeling returns. The logarithmic specification provides a tractable approximation of this effect

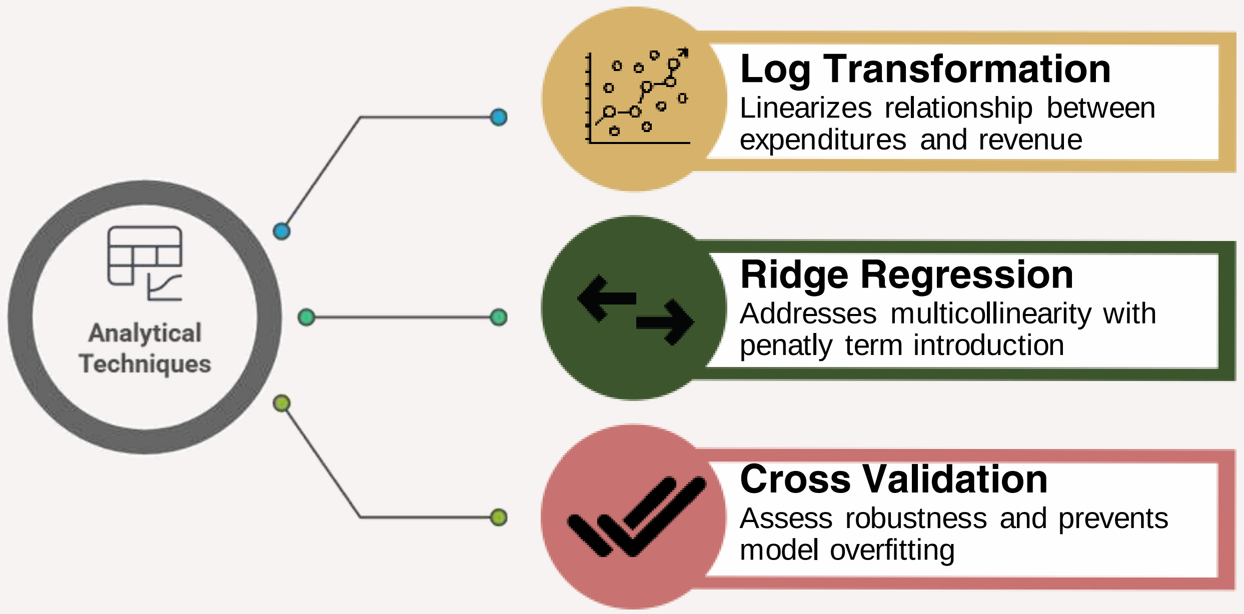
$$R_i = a_i + b_i \cdot \log(S_i)$$

Aggregate Revenue Representation:

$$R_{total} = \sum_{i=1}^{10} (a_i + b_i \cdot \log(S_i))$$

R_i : Revenue generated by marketing
S channel
a_i b_i : Spend allocated to channel
: Coefficients estimated via regression

Enhancing Model Accuracy and Interpretability



OPTIMIZATION FRAMEWORK

Maximize total revenue subject to fixed budget constraint

$$\text{Maximize} \sum_{i=1}^{10} (a_i + b_i \quad \log(S_i)) \quad \text{Subject to} \quad \sum_{i=1}^{10} S_i = \text{Total Budget}$$

Gradient Ascent

Iterative adjustment to maximize revenue

Apply Constraints

Ensure budget limits and channel bounds

Sensitivity Analysis

Evaluate robustness under scenarios

Key Benefits

Adaptability

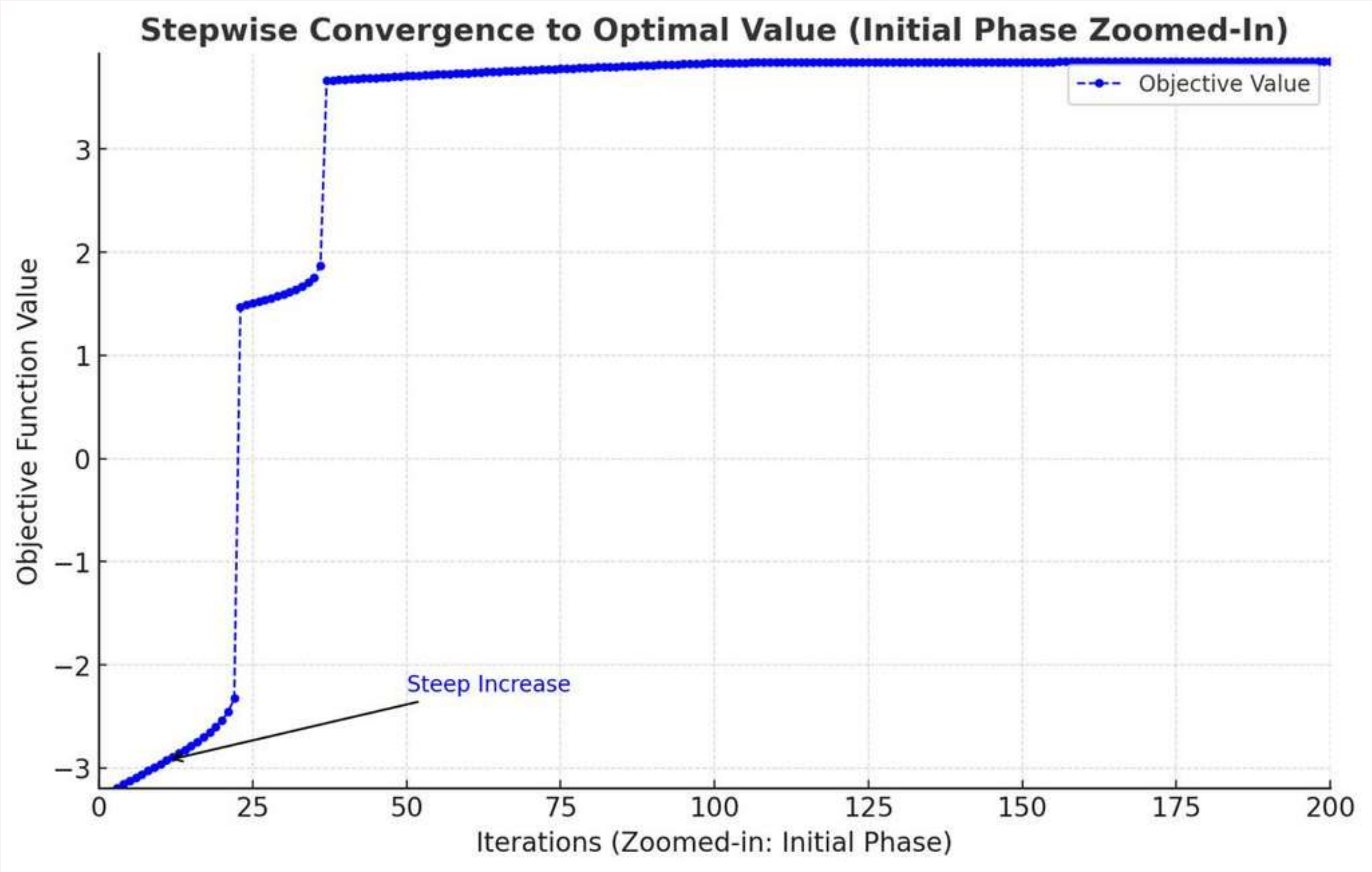
Accommodating business constraints and channel-specific limits

Computational efficiency

Rapid convergence with minimal iterations

Optimizing Marketing Budget

Objective Function Value vs Iterations



RESULTS AND INSIGHTS

Ridge Regression demonstrated strong predictive reliability, yielding high R^2 values and effectively capturing diminishing returns

The **log-linear relationship** identified key inflection points where additional spending became inefficient, enabling targeted budget reductions.

129% ROI | **29% Net Profit**

Analysis revealed that previous budgets were **overallocated** to saturated channels
Our model successfully redirected funds to **higher-elasticity channels**

CONCLUSIONS

The **logarithmic model** effectively captures diminishing returns in marketing expenditures

Ridge Regression proved valuable in mitigating overfitting and ensuring stable coefficient estimation despite collinearity

The **gradient-based optimization** approach improved marketing efficiency by reallocating resources to high-ROI channels

Inventory Optimization

Factors for Optimized Inventory Allocation:

- **Procurement Efficiency**
- **Sales Contribution**

Objective

Maximize sum of the weighted unit allocations across all subcategories. To achieve this, we employed a gradient descent optimization algorithm, ensuring an efficient and data-driven approach to inventory management.

$$Maximize \sum_{i=1}^{n} (WeightedAverageSLA_i \times PercentageSales_i \times Units_i)$$

Where n is the number of subcategories, the optimal percentage distribution of warehouse space units for each subcategory is represented in the pie chart below, with the weights provided as:

Process

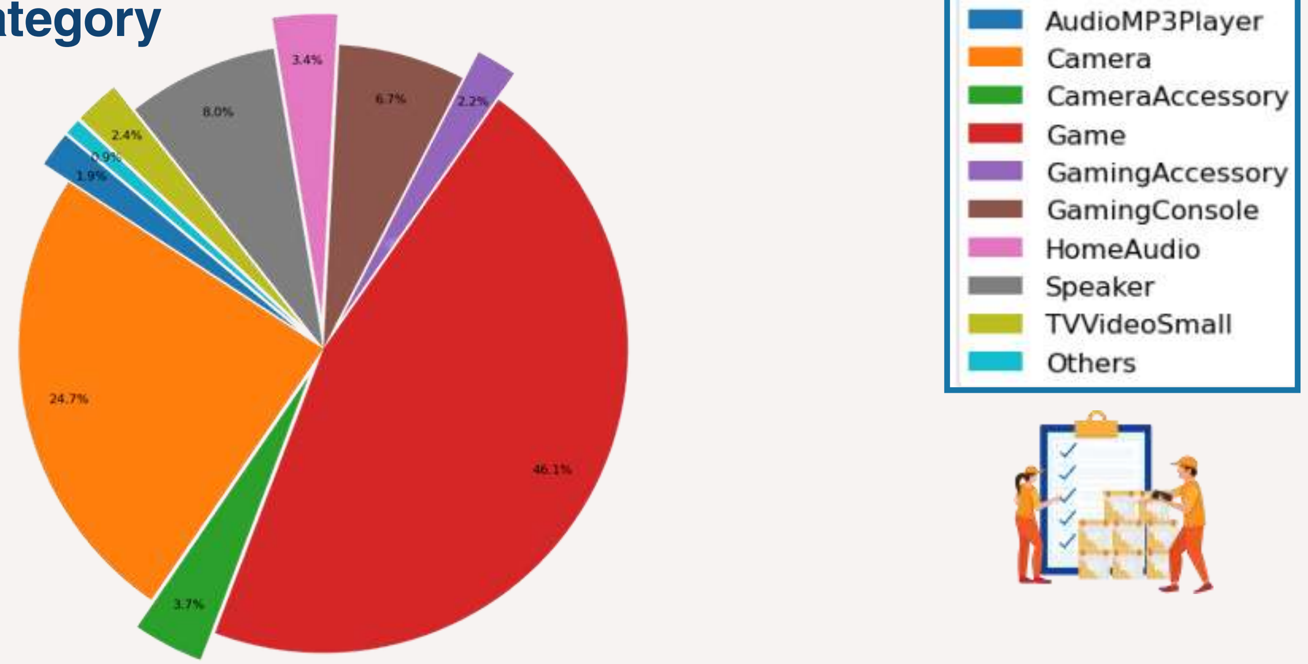
1) Calculate Service Level Agreements for Each Product Subcategory

2) Analyze the Percentage Contribution of each Subcategory to Total Sales

3) Obtain a Coefficient by multiplying Weighted SLA and Sales Contribution

4) Use coefficients to Maximize Sum of Weighted Unit Allocations across subcategories

Optimal Percentage Distribution of Inventory space by Subcategory



ElectroMart Analytics

Dashboard

Aug 23

Data period: Jul 23 - Sep 23

Overview | Performance Drivers | Budget Optimization | Product Analysis

View Full Year

Revenue (Aug 23)

$5,201

Monthly revenue

Marketing Spend (Aug 23)

$633,038

Allocated for Aug 23

ROI (Aug 23)

0.01x

Return on marketing investment

NPS Score (Aug 23)

60

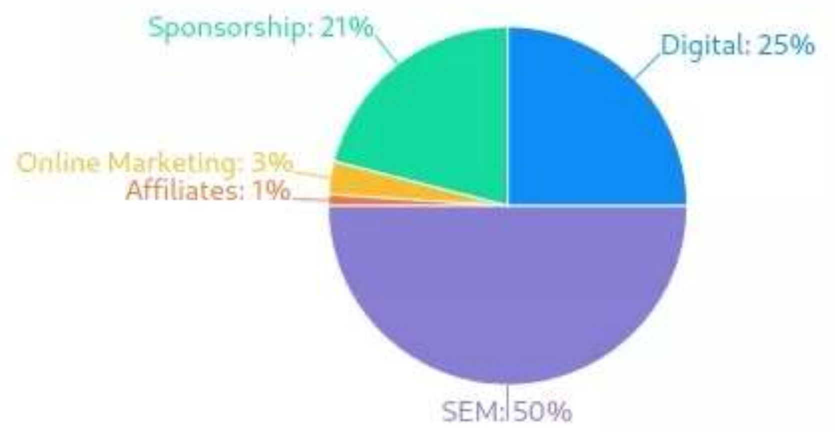
Customer satisfaction

Revenue Trend

Monthly revenue, orders, and average order value

Marketing Channel Distribution

Spend allocation by channel (Aug 23)



The ROI (0.0082) is extremely low, indicating poor marketing efficiency and an urgent need for optimization. Nearly half of the budget is spent on SEM (49.63%), but it is not driving strong sales, while Digital (25.24%) and Sponsorship (21.00%) are also underperforming. Content Marketing (0.0001%) and Radio (~0%) are almost neglected, missing long-term brand-building opportunities. The NPS Score (60.0) shows improved customer satisfaction, but sales remain weak. Urgent action is needed to reevaluate the SEM strategy, optimize budget allocation, and explore organic marketing efforts to boost ROI.

Additional Dataset Suggestions

Below are several supplementary dataset requirement suggestions, each designed to enhance the depth and accuracy of our analysis, ultimately enabling us to derive more comprehensive and insightful conclusions.

Customer Demographic and Behavioural Data

- Age
- Gender
- Geographical Location
- Income levels
- Web browsing history
- Session Duration
- Click Through Rates
- Cost Abandonment Rate
- Wishlist Addition
- Loyalty Program

Marketing Campaign Metrics

- Cost Per Click (CPC)
- Cost per acquisition
- Social Media Engagements
- Engagement data from social media (likes, comments, shares etc)
- Customer acquisition source data (e.g., organic, referral, paid ads)
- Return on Ad Spend (ROAS)

Product and Inventory Information

- Product ratings, reviews, customer feedback (via socials)
- Stock levels
- Supply chain lead times
- Inventory turnover rates
- Competitor pricing trends
- Warranty claims
- Return Rates per Product

Transactional and Financial Insights

- Refunds and cancellation rates
- Discount redemption behaviour
- Cost of Goods Sold (COGS),
- profit margins per SKU
- Threshold purchase patterns

External and Contextual Factors

- Competitor market share comparison data
- Customer service data i.e. response time
- Resolution Rates
- Issue Frequency
- Industry trends and seasonality effects (apart from sales)

THANK YOU