find insight

Shop Market **Analysis**

By Nadia Al-Salem

Project **Overview**

The Shop Market dataset comprises comprehensive sales records spanning multiple years, covering global markets including Africa, APAC, EMEA, EU, and the Americas. The dataset provides a detailed view of order transactions, customer segments, product categories, and financial metrics essential for business intelligence and predictive analytics.

Key Attributes

- Order & Shipping Details
- Customer Name & Segment
- Geographic Information
- Product Category & Sub-Category
- Sales Amount & Quantity
- Discount & Profit Margins
- Shipping Cost & Priority
- Temporal Data (Year)

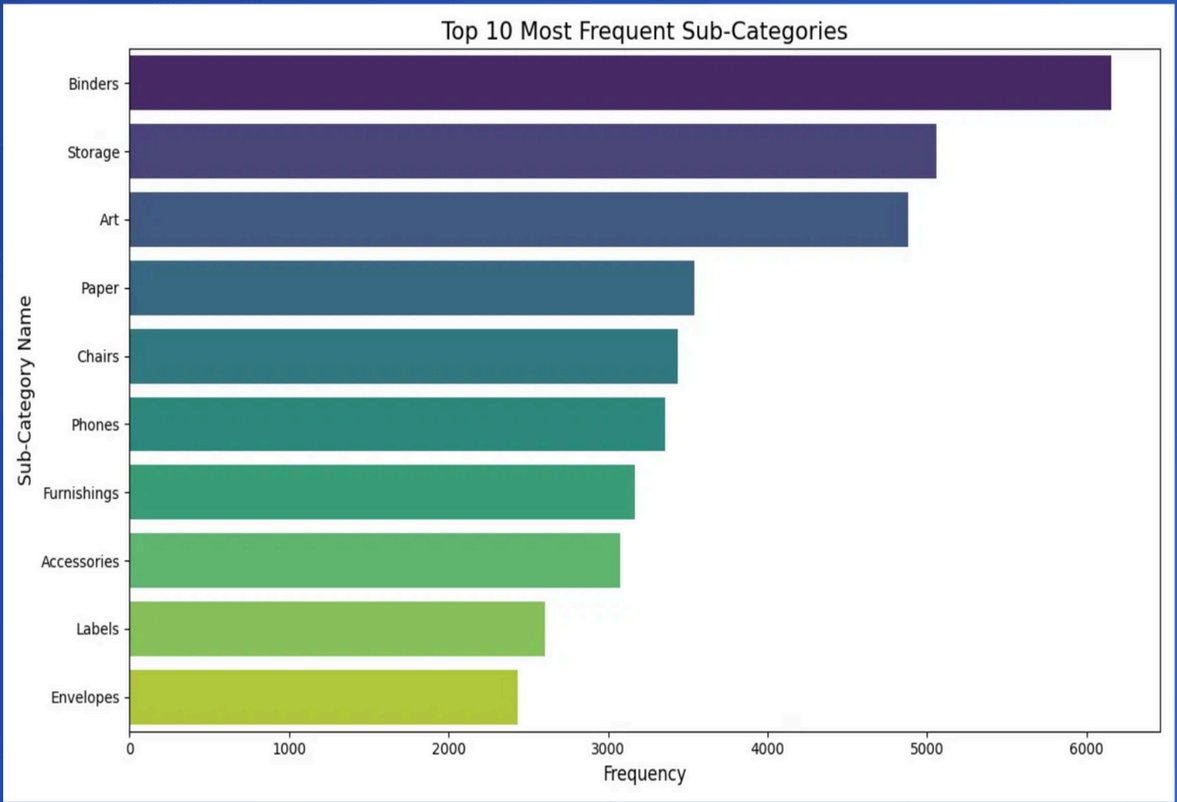
51,290
Total Records

21
Attributes

0
Missing Values

100%
Data Quality

Exploratory Data Analysis



Top Sub-Categories

Binders, Storage, and Art dominate the product mix, representing the highest transaction frequencies. These categories are critical for inventory planning and marketing focus.

Sales Distribution

Sales values are concentrated in the lower to mid-range, with occasional high-value outliers. This pattern suggests a broad customer base with varying purchasing power.

Profitability Patterns

While most transactions yield positive profit, a notable subset shows losses, particularly when discounts exceed optimal thresholds. Discount strategies require careful calibration.

Key Takeaway: Focus on optimizing inventory for top-selling sub-categories and refining discount policies to protect profit margins.

Predictive Modeling

The primary objective was to develop a predictive model capable of forecasting profit based on various features including product category, customer segment, geographic location, ship mode , and temporal factors. This enables proactive inventory management and revenue optimization.

Models Evaluated

- Cat Boost
- XGBoost
- Random Forest

Model Performance

Multiple regression models were trained and evaluated using standard metrics including Mean Absolute Error (MAE), Root Mean Squared Errpr and R Squared (R2). The LightGBM demonstrated superior performance by capturing complex non-linear relationships in the data.

Best Performing Model

LightGBM achieved the lowest error rates and highest predictive accuracy, making it the recommended model for sales forecasting in production environments.

Evaluation Metric	Best Model
MAE & RMSE	LightGBM

Conclusion & Recommendations

Summary of Findings

The Shop Market analysis revealed critical insights into sales performance, product category dynamics, and profitability drivers. Binders, Storage, and Art emerged as the most frequently purchased sub-categories. The analysis highlighted the significant impact of discount strategies on profit margins. Predictive modeling using Random Forest Regression demonstrated strong potential for sales forecasting, enabling data-driven decisions.

Strategic Recommendations

01

Optimize Inventory Management: Prioritize stock levels for top-selling sub-categories to meet customer demand and reduce stockouts.

02

Refine Discount Strategies: Implement data-driven discount thresholds to balance sales volume with profitability.

03

Leverage Predictive Analytics: Deploy the LightGBM model for sales forecasting to improve demand planning.

04

Investigate Customer Segments: Conduct deeper analysis of high-value segments to tailor marketing campaigns.

Thynk Unlimited

Thank You!

Thank you for joining us in exploring key shop market insights. Wishing you success in your investment journey!

www.nadiaalsalem.com