Exploratory Data Analysis Report: Example KPI Report

KeyStone Predictive

Produced by KeyStone Predictive
www.keystonepredictive.com

1 Introduction

This report shows exploratory data analysis for Example Business using simulated transaction data. We looked at revenue, churn risk, customer behavior, and sales trends to identify areas for growth. Data confidentiality and security are addressed in client contracts.

Table 1: Variables Studied and Definitions

Variable	Definition
`customer_id`	Unique identifier for each customer in the dataset.
`purchase_date`	The date on which a transaction occurred.
`amount_spent`	Total dollar amount spent in a single transaction.
`recency_days`	Number of days since a customer's last purchase.
`frequency`	Total number of purchases made by a customer.
`monetary_value`	Total dollar amount spent by a customer across all transactions.
`churn_risk`	Classification of customer as high or low churn risk based on inactivity over 90 days.
`transactions`	Daily count of total transactions across all customers.

2 Sales Performance

2.1 Monthly Sales Trend

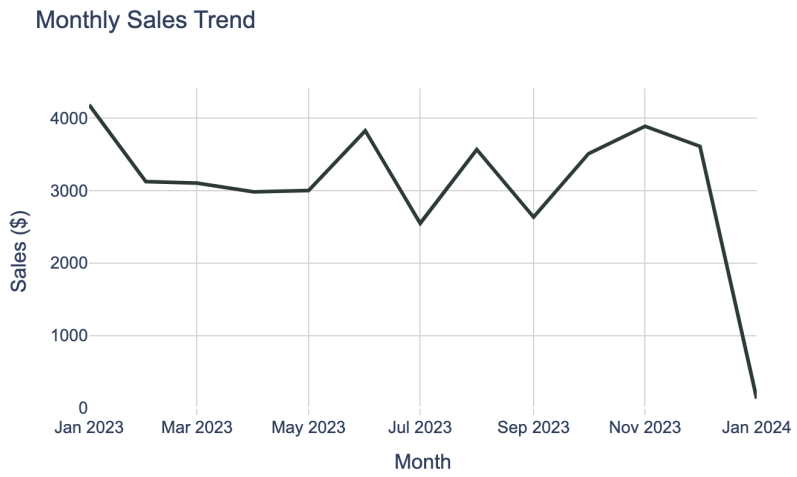

Figure 1: Monthly Sales Trend

This line chart shows the total dollars of sales processed each month over the year. By plotting month-over-month revenue, we can spot seasonal swings, spikes or dips in customer demand, and overall business momentum. **Key takeaway:** Sales remain relatively stable, with a noticeable peak in June and a dip in July—insights you can use to plan inventory and marketing pushes around those months.

2.2 Quarterly Sales Trend

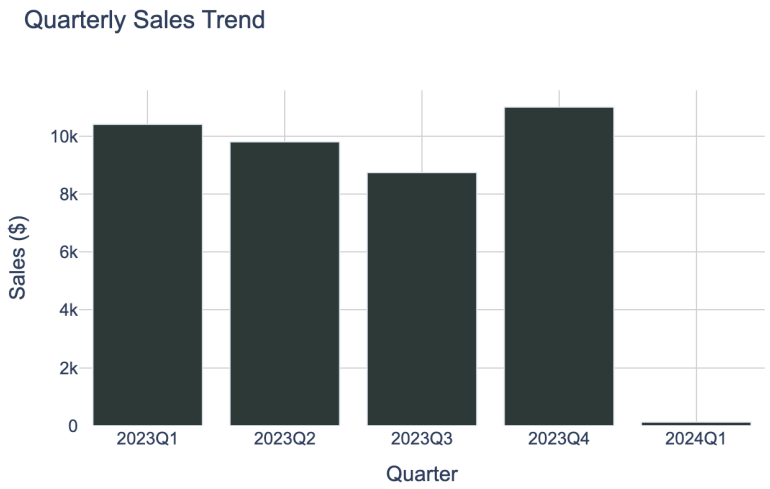

Figure 2: Quarterly Sales Trend

This chart shows total sales by quarter. It evens out monthly swings and shows overall trends. **Key takeaway:** Q4 had the highest sales, likely from holiday and year-end promotions. Focusing marketing in Q4 could boost revenue.

3 Customer Churn Analysis

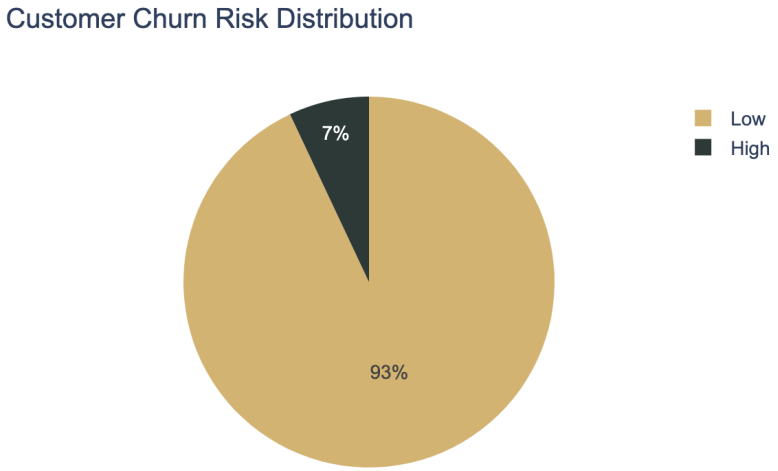

Figure 3: Customer Churn Risk Distribution

This pie chart groups customers into high or low churn risk using a 90-day cutoff. **Key takeaway:** About 7% are high risk. Improving retention for them can boost revenue and lower acquisition costs.

4 Transaction Patterns

4.1 Transaction Amounts

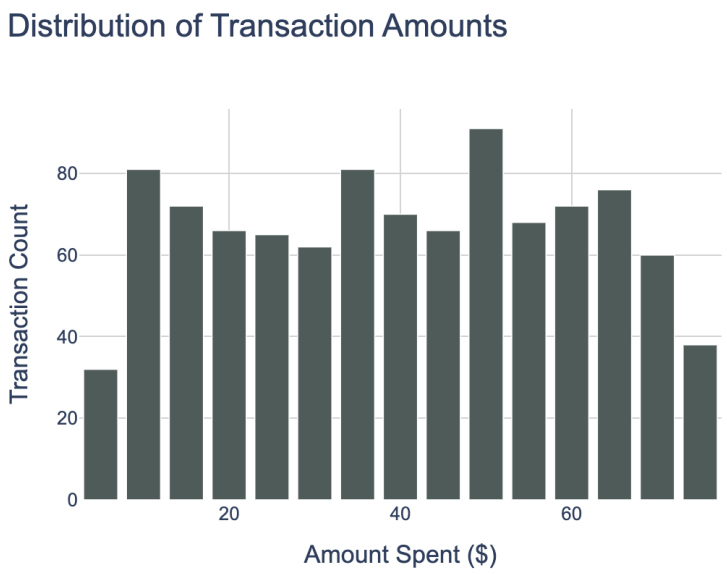

Figure 4: Distribution of Transaction Amounts

This histogram shows purchase amounts from $5 to $75. **Key takeaway:** Most purchases fall between $10 and $60, with the $45–$50 range seeing the highest volume and a smaller peak around $15–$20. Pricing products in these tiers may boost sales.

4.2 Transactions Per Customer

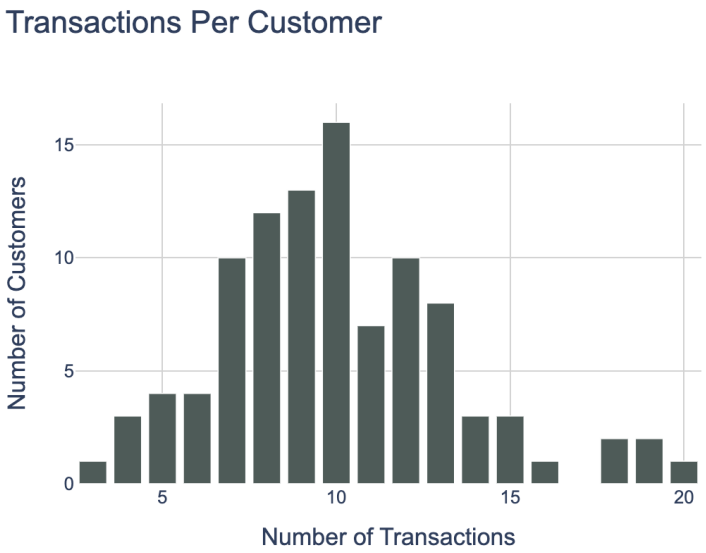

Figure 5: Transactions Per Customer

This histogram shows how often each customer bought over the year. Most made 8–12 purchases, with a small handful buying more than 15 times. **Key takeaway:** Some type of a loyalty or subscription program for the 8–12-purchase segment, and consider a VIP tier or special perks for the highest-frequency buyers.

4.3 Daily Transaction Volume

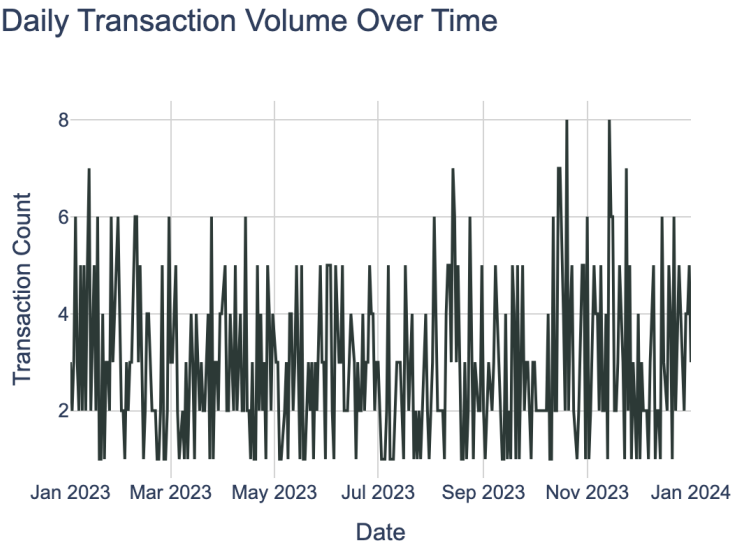

Figure 6: Daily Transaction Volume Over Time

This line chart shows daily sales counts over the year. **Key takeaway:** Volume usually ranges between 2 and 6 sales per day, with occasional peaks of 7–8 sales. We also see a lift in late Q4 around the holidays and higher counts on weekends during promotions. Consider testing flash deals on slower weekdays and using a 7-day average to plan staffing and inventory.

5 RFM Correlation Insights

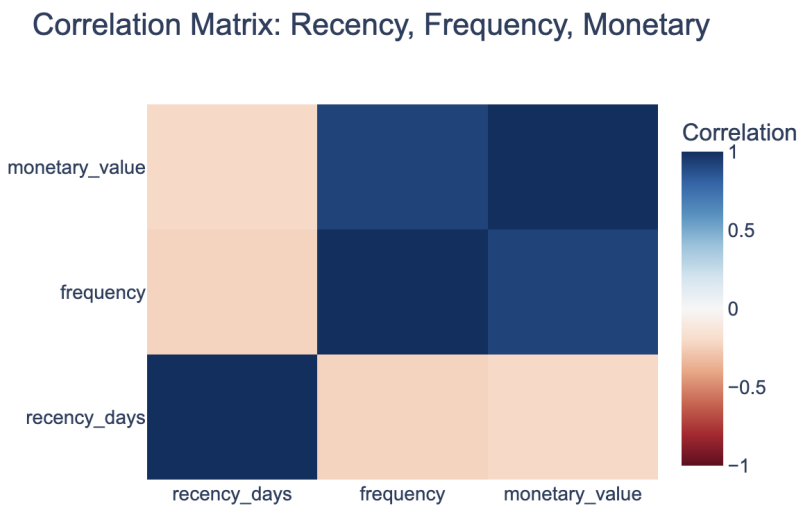

Figure 7: Correlation Matrix: Recency, Frequency, Monetary

This heatmap shows how recent, frequent, and high-spending customers relate to each other. **Key takeaway:** Frequency and monetary value are strongly correlated—customers who spend more tend to buy more often. Recency is negatively related to both, meaning the longer it's been since someone last purchased, the less they tend to buy or spend overall. This pattern can help prioritize re-engagement: those with low recency but past high frequency or spend may still be worth targeting.

6 Summary and Next Steps

Summary Insights

- **Total Revenue:** $40,081.60 — shows total income generated over the year.
- **Average Transaction Value:** 1000 — helps identify spending behavior per visit.
- **Churn Risk Distribution:** About 7% high risk — shows where customer retention efforts are most needed.
- **Most Active Buyers:** 8–12 purchases per year — indicates a strong base of loyal repeat customers.
- **Sales Volume:** 2–6 per day, higher in Q4 — reflects healthy day-to-day performance with holiday upside.
- **Correlation Trends:** Recency negatively related to frequency/spend — shows long-inactive customers are less likely to return without re-engagement.

Key Observations

- Sales are stable month-to-month with a summer peak and dip.
  Helps businesses plan inventory and marketing around seasonal trends.
- Q4 is the strongest quarter.
  Suggests focusing promotions and staffing for the holiday season.
- Most purchases are in the $10–$60 range.
  Suggests pricing sweet spots for bundling or new offers.
- Customers buying 8–12 times per year represent a core segment.
  Great candidates for loyalty programs and personalized follow-ups.
- Weekend and holiday periods see spikes in daily sales.
  Useful for timing campaigns, staffing, and flash sales.

Possible Improvements

- Enrich the dataset with product categories and customer demographics.
  Helps personalize marketing and understand customer needs.
- Capture data over multiple years.
  Enables better trend analysis and more accurate forecasting.

- Track promotional activity and response rates.
  Helps assess ROI and optimize marketing spend.
- Add weekly or hourly time resolution.
  Useful for peak-time staffing or flash promotion targeting.

Potential Machine Learning Applications

- **Customer Churn Prediction** (e.g., Random Forest, XGBoost)
  Helps identify at-risk customers and prompt retention outreach.
- **Sales Forecasting** (e.g., Prophet, ARIMA, LSTM)
  Supports smarter inventory, cash flow planning, and seasonal prep.
- **Customer Segmentation** (e.g., K-Means, DBSCAN)
  Enables targeted messaging, loyalty offers, and product recommendations.
- **Anomaly Detection**
  Flags unexpected spikes or drops in sales that may indicate issues or opportunities.
- **Recommendation Systems**
  Boosts revenue by surfacing relevant products based on past buying behavior.

Interactive Dashboards (Direct Links)

The following sample dashboards are hosted on Dropbox and will open directly in your browser. This setup is for demonstration only and is not used for client-facing deliverables or data hosting.

- Monthly Sales Trend
- Quarterly Sales Trend
- Transaction Amount Distribution
- Transactions Per Customer
- Daily Transaction Volume
- RFM Correlation Matrix

These interactive dashboards are best viewed in Chrome, Firefox, or Edge.

Produced by KeyStone Predictive
www.keystonepredictive.com