

DATA-DRIVEN MARKETING STRATEGIES FOR ENHANCED BUSINESS PERFORMANCE

Proposals for Customer Retention, Purchase Pattern Analysis, and Sentiment Analysis

STRUCTURE

- business questions
- data overview
- insights and analysis
- quantitative approach
- marketing campaign proposals
- expected outcomes



BUSINESS QUESTIONS

- Key Business Questions:
 - How can we retain high-value customers?
 - What unexpected relations can we find in our purchase patterns?
 - How can we automate the extraction of valuable insights from customer feedback?

DATA OVERVIEW

- Datasets Used:

- Transactions Data
- Customer Reviews
- Labelled Review Sentiments
- Customer and Product Information

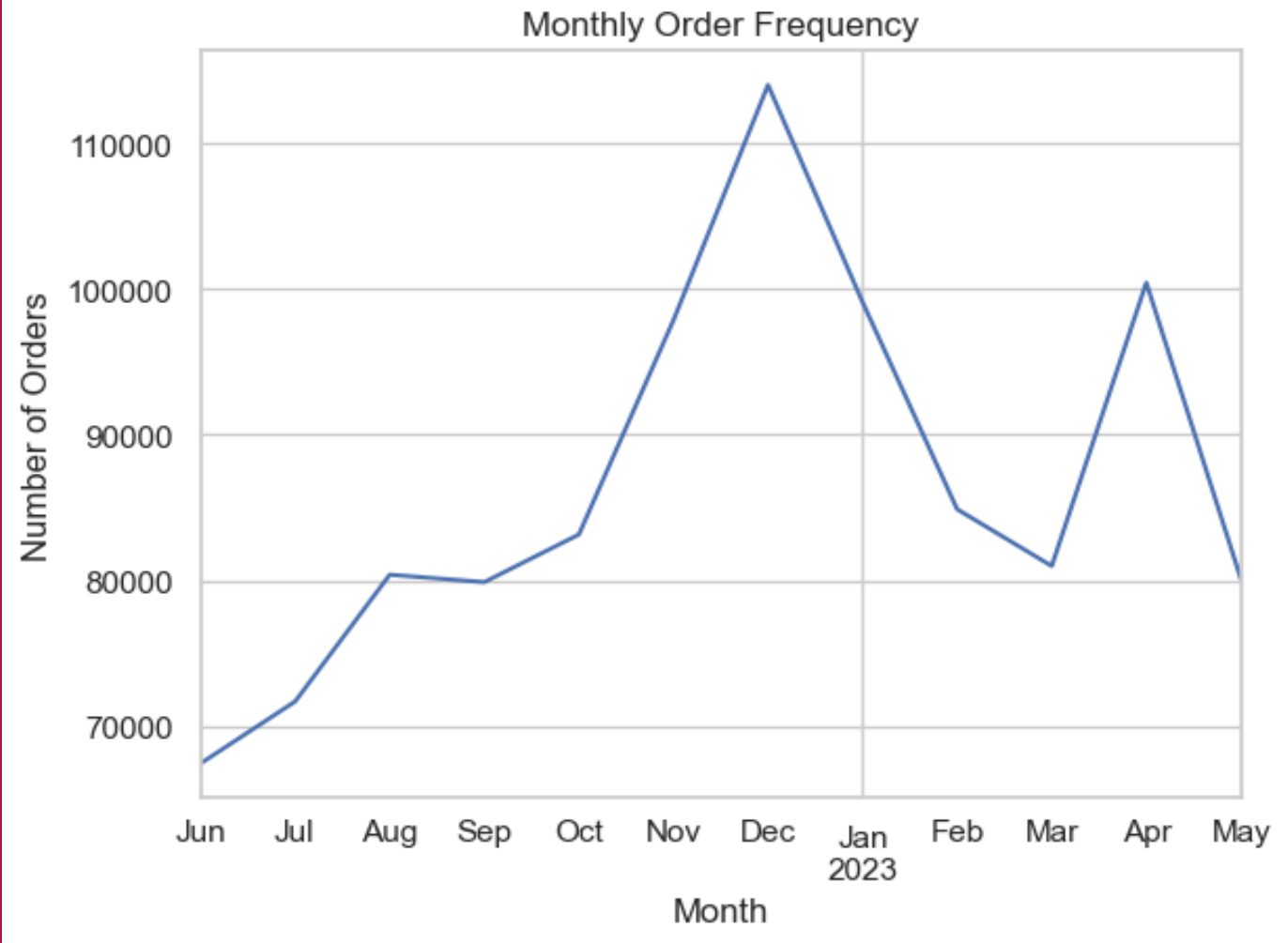
- Key Metrics:

- Customers: 140,946
- Total transactions: 371,804
- Unique products: 88,538
- Customer reviews: 105,710
- Labelled reviews: 462,744

DATA OVERVIEW

- Monthly order frequency

This line chart represents the number of orders placed each month. It helps to identify trends and patterns in purchasing behavior over time. The peak observed in December indicates a significant increase in orders during the holiday season, followed by a decline in January and subsequent fluctuations.

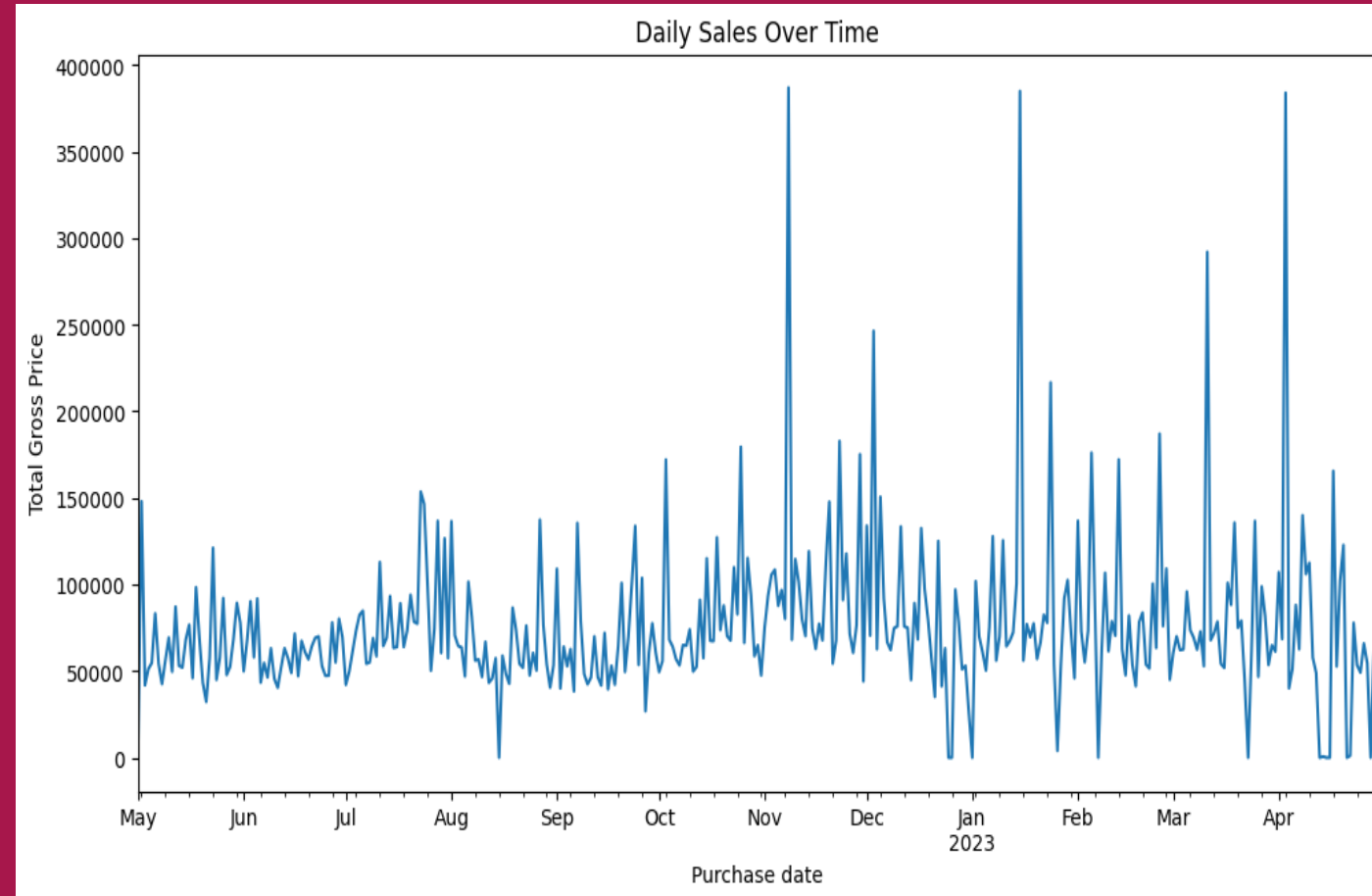


"Of all the months of the year there is not a month one half so welcome to the young, or so full of happy associations, as the last month of the year." - Charles Dickens

DATA OVERVIEW

Key Observations:

- **Seasonal Peaks:** There are notable spikes in sales around specific months. For instance, high peaks are observable in November, December, and subsequent months. These could correlate with holiday seasons or promotional periods like Black Friday or Christmas sales.
- **General Trends:** While the daily sales show variability, there are underlying trends where sales rise and fall in a pattern that may correspond to consumer behavior, marketing campaigns, or external economic factors.
- **Low and High Points:** The lowest sales days typically occur in periods outside of major shopping seasons, which might indicate off-peak times for the business.
- **End of Year Surge:** A significant surge in sales towards the end of the year suggests increased consumer purchasing activity, possibly influenced by the holiday season.



"Of all the months of the year there is not a month one half so welcome to the young, or so full of happy associations, as the last month of the year." - Charles Dickens

DATA OVERVIEW

- **Size of Words:** The size of each word in the cloud corresponds to its frequency in the reviews. Larger words appear more frequently, making them key points of focus.
- **Prominent Words:** "Love", "Good", and "Great" are prominently featured, suggesting positive feedback and high satisfaction among our customers.
- **"Buy" and "Tried"** indicate purchasing behavior and willingness to experiment with different products.
- **"Amazon" and "Price"** reflect the influence of online shopping platforms and cost considerations on purchasing decisions.
- Words like **"Tea"**, **"Coffee"**, and **"Chocolate"** might point to popular product categories or flavors that are frequently discussed by customers.

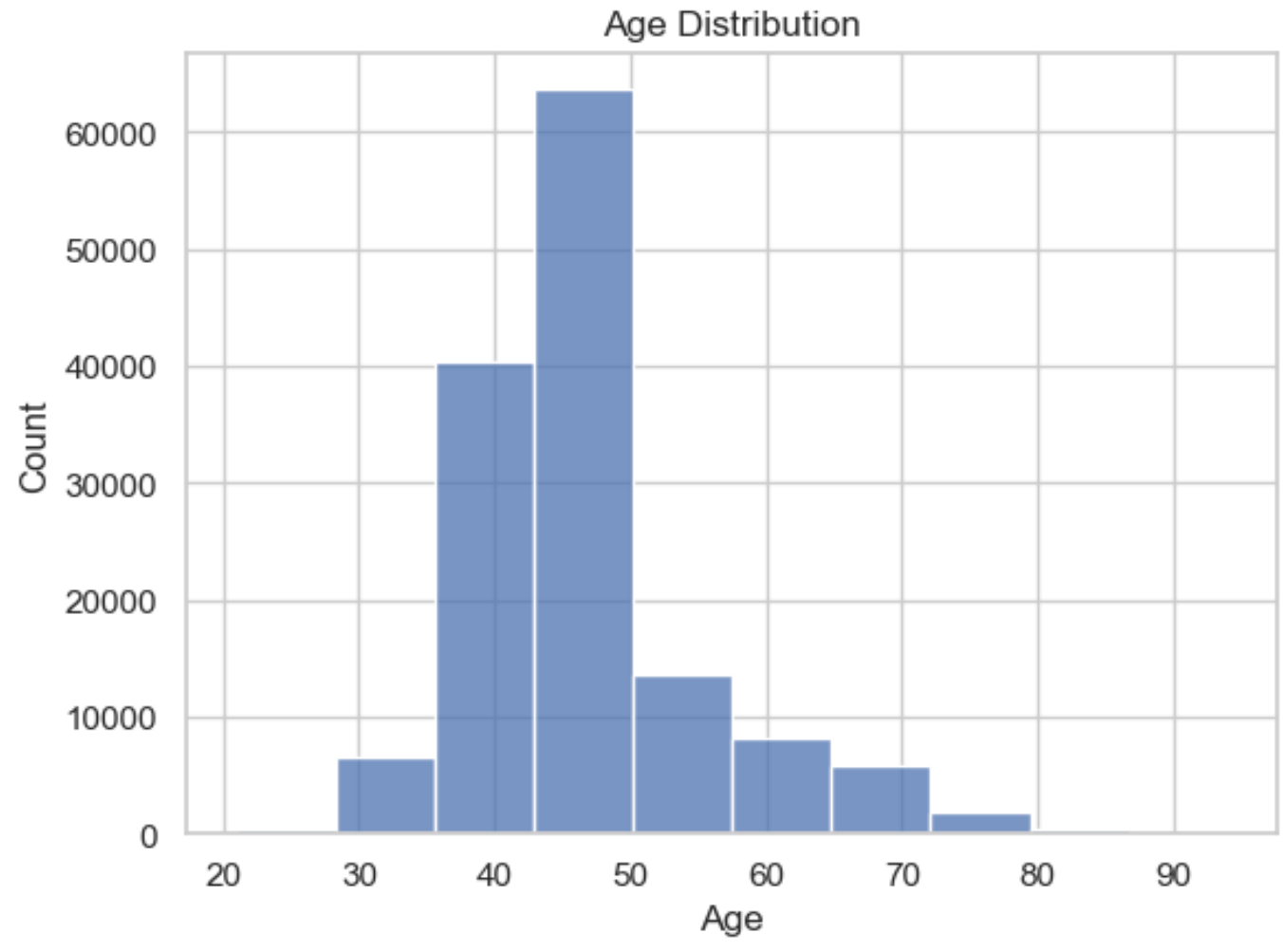


"Of all the months of the year there is not a month one half so welcome to the young, or so full of happy associations, as the last month of the year." – Charles Dickens

DATA OVERVIEW

- Age distribution

This histogram shows the distribution of customer ages. The majority of customers fall within the 40-50 age range, indicating that middle-aged adults are the primary customers.



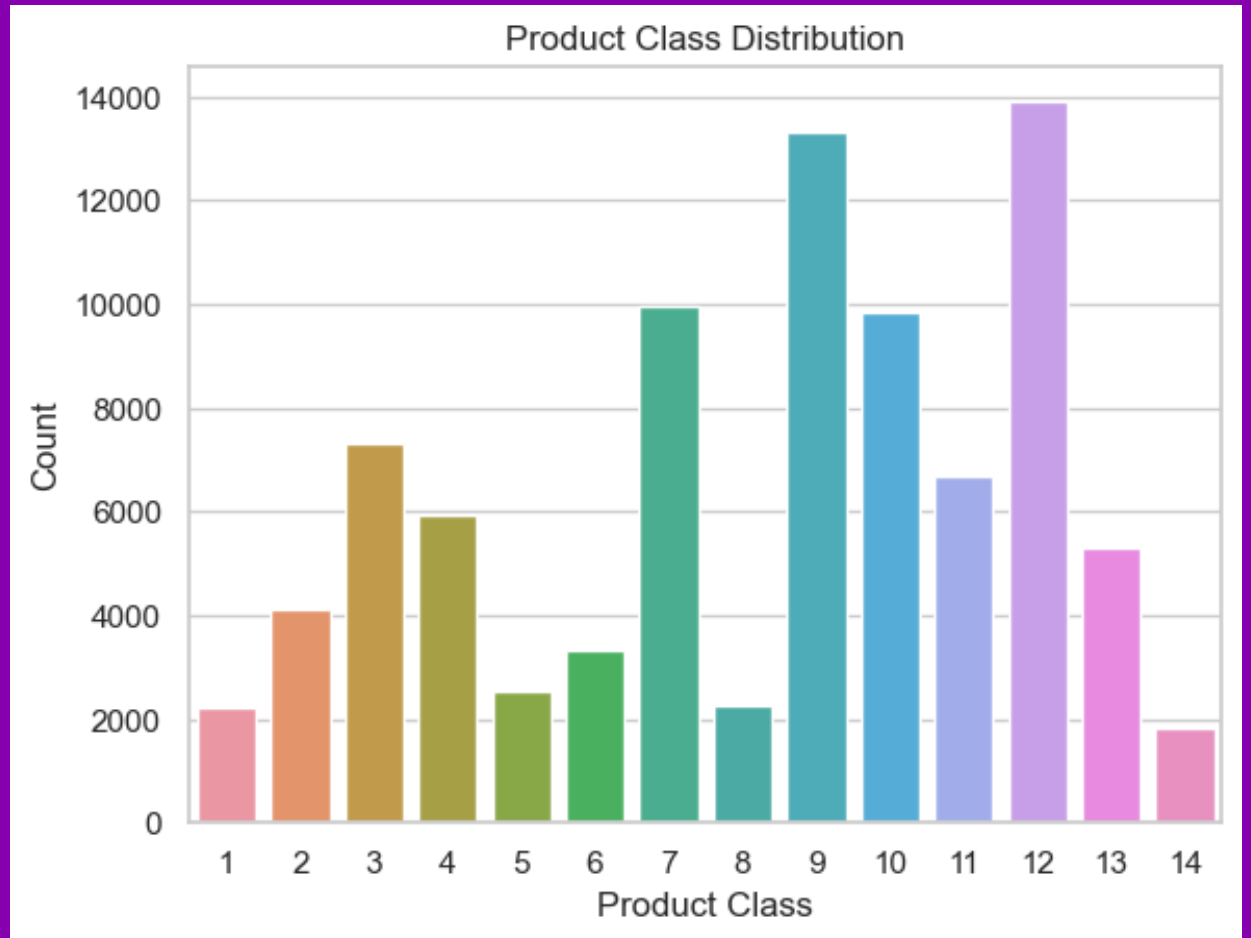
"Forty is the old age of youth; fifty the youth of old age."

– Victor Hugo

DATA OVERVIEW

- Product class distribution

This bar chart displays the distribution of products across different product classes. It highlights which product categories are most and least popular among customers. Class 9 and 13 have the highest counts, suggesting they are the most frequently purchased categories.



DATA OVERVIEW

- Gender distribution

This bar chart represents the gender distribution of customers. There is a slightly higher number of female customers compared to male customers.



DATA OVERVIEW

- Loyalty type distribution

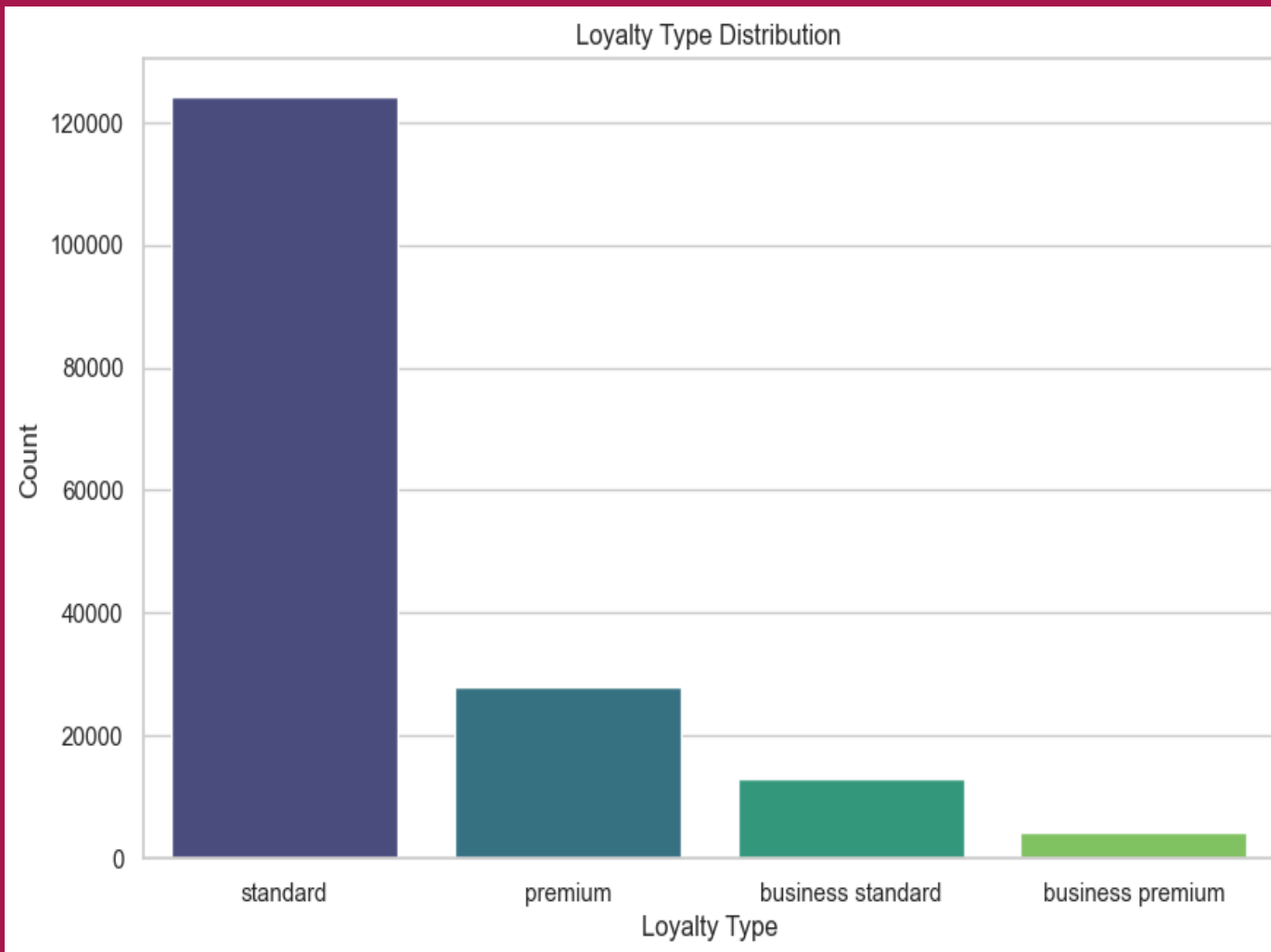
This bar chart illustrates the distribution of customers across different loyalty types. The categories include:

Standard: The majority of customers fall into this category, indicating they are part of the basic loyalty program.

Premium: A smaller but significant number of customers are in the premium loyalty program, which likely offers more benefits.

Business Standard: These customers belong to a business-focused standard loyalty program.

Business Premium: The smallest group, indicating business customers with premium loyalty status.



DATA OVERVIEW

- **Business Premium:** Customers in this category have the highest average number of orders, suggesting they are highly engaged and make frequent purchases.
- **Premium:** These customers also have a high average number of orders, indicating strong engagement.
- **Business Standard:** Customers in this category have a moderate average number of orders.
- **Standard:** Customers in the standard loyalty program have the lowest average number of orders, reflecting lower engagement compared to other loyalty types.

- **Average number of orders per loyalty type**

This bar chart shows the average number of orders made by customers in each loyalty type category. Key observations include: Business Premium, Premium, Business Standard, Standard.



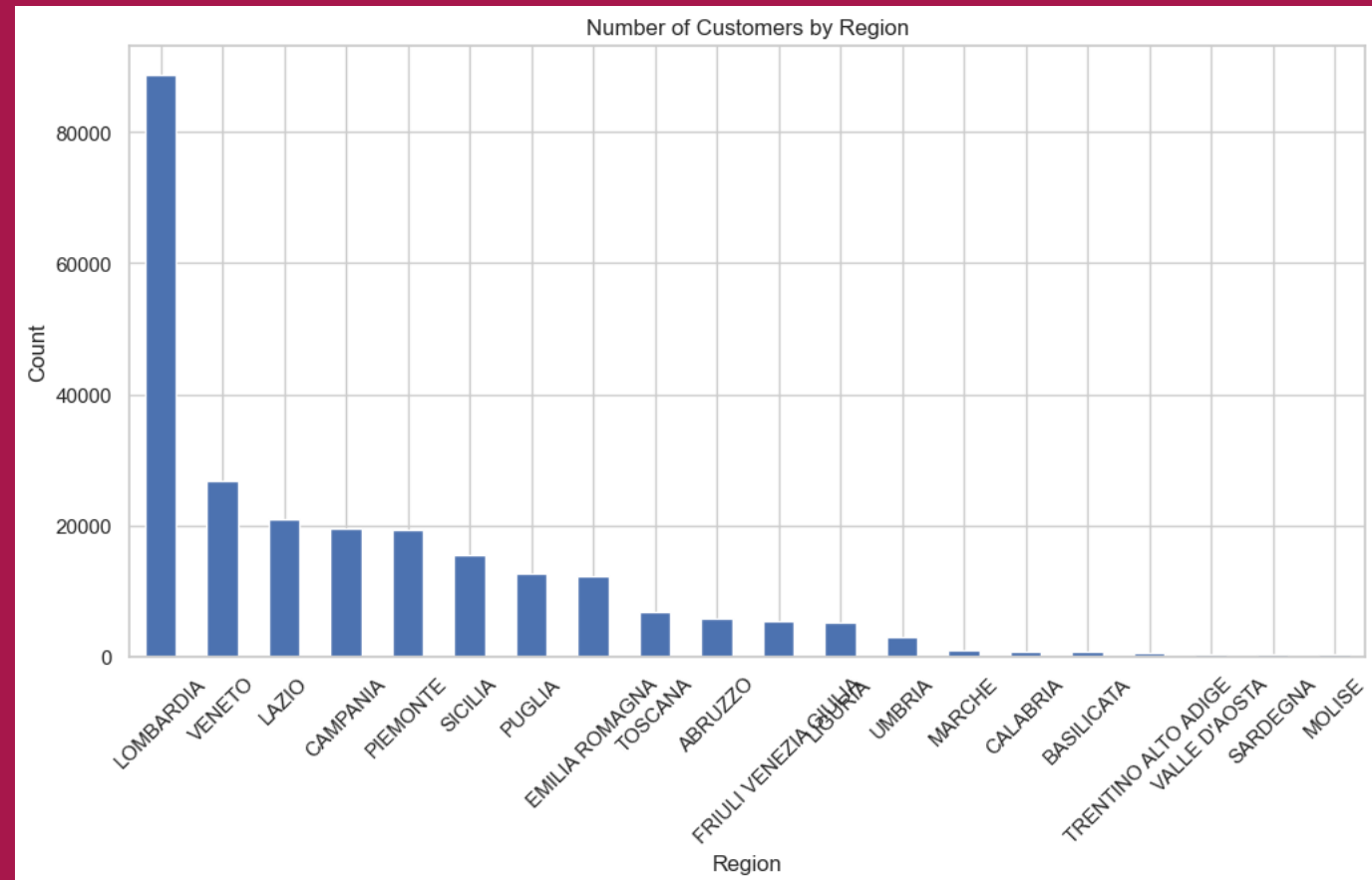
DATA OVERVIEW

Key Observations:

- **Lombardia** stands out with the highest customer count, significantly more than other regions, suggesting it as our primary market.
- Following Lombardia, regions like **Veneto** and **Lazio** also show substantial customer bases, indicating strong market presence.
- On the other end, **Molise**, along with **Valle d'Aosta** and **Basilicata**, has the smallest customer base, which might indicate regions with potential for market expansion or underpenetration.

Number of customers by region

This bar chart illustrates the distribution of customer base across various regions, highlighting where our market penetration is strongest and where there are opportunities for growth.



CUSTOMER FOCUS

- RFM Analysis

RFM (Recency, Frequency, Monetary) analysis is a customer segmentation technique that categorizes customers based on their purchasing behavior.

- Recency: How recently a customer made a purchase
- Frequency: How often a customer makes purchases
- Monetary: How much a customer spends

Each RFM component is divided into quartiles and assigned scores from 1 to 4. The RFM score is then concatenated to create a three-digit score (e.g., '444' for the highest value customers). Customers with RFM scores of '444', '434', '443', '433', '442', '432' are considered **high-value**.

CUSTOMER FOCUS

- **Churn Prediction Model**

A machine learning model to predict the likelihood of customer churn among high-value customers.

Target Variable: Churn is defined as customers who haven't made a purchase in the last 60 days (Recency > 60).

Algorithm Chosen: Random Forest Classifier was selected for performance benchmarking. This choice is justified by:

- Ability to handle non-linear relationships
- Robustness to outliers and noise
- Provides feature importance rankings

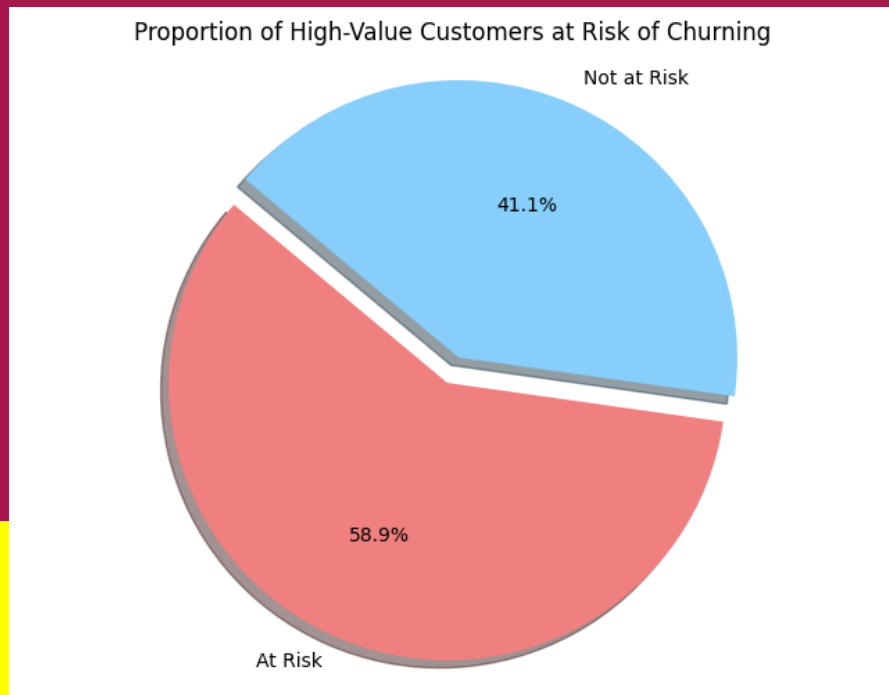
Performance Metrics:

- **Accuracy:** Measures overall correctness of predictions.
- **Precision:** Measures the accuracy of positive predictions.
- **Recall:** Measures the ability to find all positive instances.
- **F1-score:** Harmonic mean of precision and recall.
- **Confusion Matrix:** Provides a breakdown of correct and incorrect predictions by class

CUSTOMER RETENTION ANALYSIS

- **High-value customers and churned**

This pie chart provides a visual representation of the distribution of high-value customers, with a specific focus on those who have churned.



- **High-Value Customers Retained:**

The **blue** segment makes up about 41.1% of the pie, suggesting that a smaller yet substantial portion of the customer base is secure and not likely to churn.

- **High-Value Customers Churned:**

The **red** segment covers approximately 58.9% of the pie, indicating that a significant majority of high-value customers are considered at risk of churning.

MARKETING STRATEGY FOR CUSTOMER RETENTION

- **Retention Campaign:**

- **Target Audience:** High-value customers at risk of churn
- **Channels:** email/SMS campaigns, push notifications, social media engagement
- **Offers:** personalized discounts, referral programs, win-back offers

- **Expected Outcomes:**

- **Reduced Churn Rate:** by targeting high-value customers with personalized offers and engagement, we expect a significant reduction in churn rates.
- **Increased Customer Lifetime Value (CLV):** enhanced loyalty and engagement will lead to increased spending and higher CLV among high-value customers.
- **Improved Customer Satisfaction:** personalized and timely offers will improve overall customer satisfaction and foster a stronger relationship with the brand.
- **Higher Average Order Value (AOV):** cross-selling and up-selling strategies will lead to an increase in the average order value.

- **Selection Criteria:**

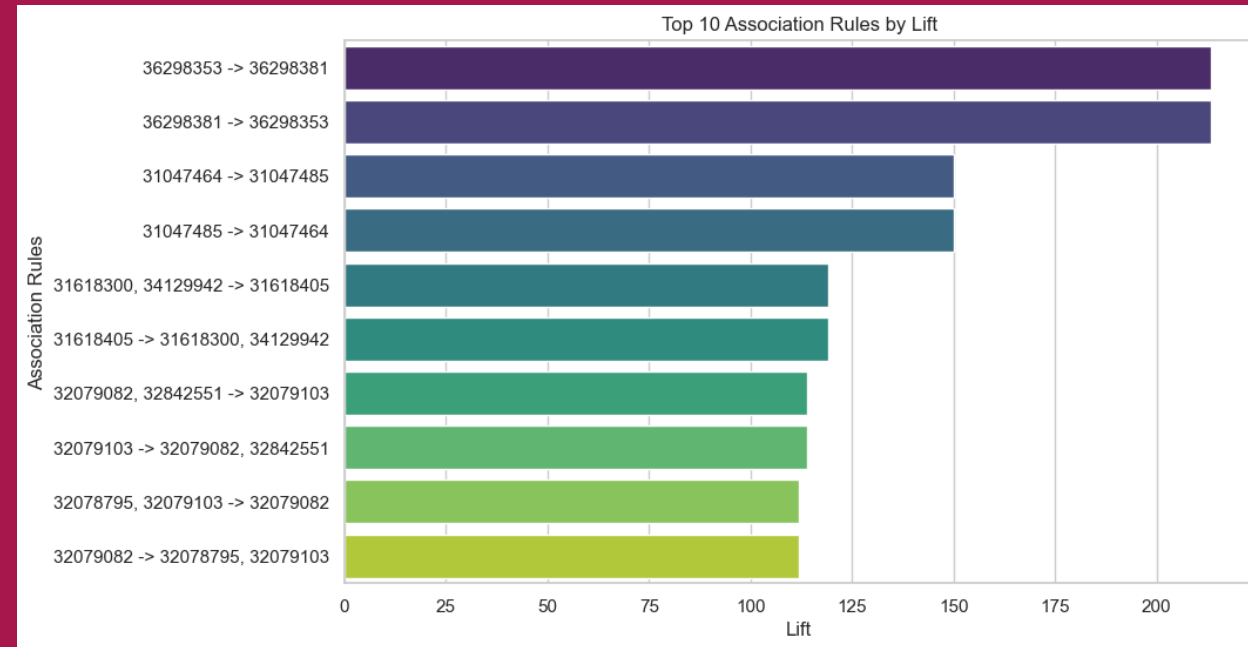
- Customers with an RFM (Recency, Frequency, Monetary) score indicating high value but decreased engagement or purchasing frequency in the last 60 days.

- **Discount strategy:**

- Offer a tiered discount system where the discount rate increases with the purchase amount: 10% off purchases over \$100, 15% off purchases over \$200, and 20% off purchases over \$300.
- Special discount of 25% on bundled products related to their previous purchases to encourage higher basket sizes.

PURCHASE PATTERN ANALYSIS

- **X-axis (Lift):** The x-axis represents the lift value. Higher lift values indicate stronger associations between items.
- **Y-axis (Association Rules):** The y-axis lists the association rules in the format "antecedents -> consequents". For example, "36298353 -> 36298381" means that if a customer buys product 36298353, they are highly likely to also buy product 36298381.



"If you're going through hell, keep going."

- Winston Churchill

PURCHASE PATTERN ANALYSIS

Key findings:

- **Strong Associations:**

The top association rules have very high lift values, indicating strong relationships between the products involved.

For example, the rule "36298353 -> 36298381" has the highest lift value, suggesting that customers who buy product 36298353 almost always buy product 36298381 as well.

- **Cross-Selling Opportunities:**

These high-lift rules reveal opportunities for cross-selling. By recommending related products based on these associations, the company can increase sales and enhance customer experience.

For instance, if a customer purchases product 31047464, suggesting product 31047485 as well could be a successful cross-selling strategy.

CROSS-SELLING CAMPAIGN

- **Channel**
 - Email
- **Target Audience**
 - Purchased product **36298353** in the last 60 days
 - Have not yet purchased product **36298381**
 - Opened at least one marketing email in the last 90 days (for email engagement)
- **Discount strategy**
 - Design bundle offers for the groups of products that show strong associations. For example, a special price when buying all three of 31618300, 34129942, and 31618405 together.

Profitability Rationale

- **High conversion potential:** The strong association (lift close to 200) suggests customers who bought 36298353 are highly likely to be interested in 36298381.
- **Increased Customer Lifetime Value (CLV):** Successfully cross-selling increases the customer's investment in our product ecosystem.
- **Efficient targeting:** By focusing on recent buyers, we're reaching out when the initial product is still fresh in their minds.

SENTIMENT ANALYSIS OVERVIEW

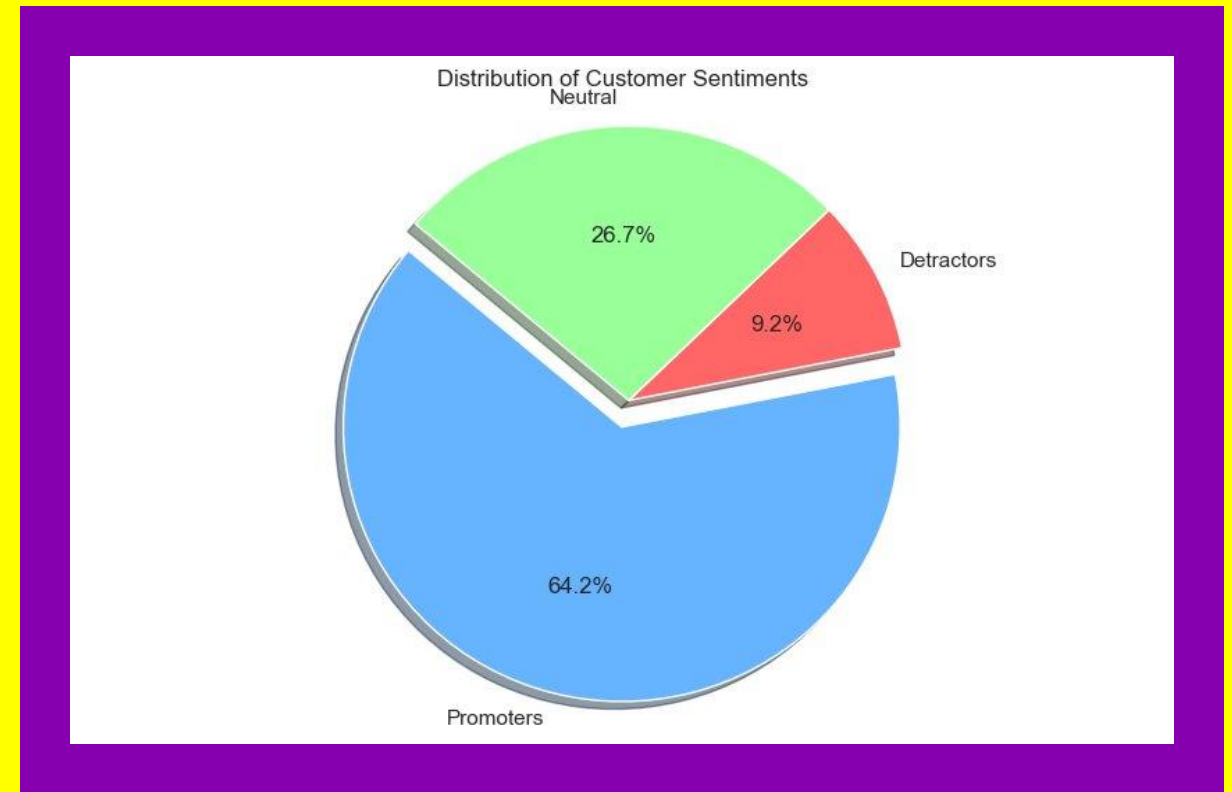
- **Sentiment Analysis:**

The dataset already included a 'sentiment_label' column, which categorized reviews into positive, neutral, and negative sentiments.

Utilized the provided sentiment_column to understand the overall sentiment distribution.

Categories:

- Positive(Promoters)
- Neutral
- Negative(Detractors)



SENTIMENT DISTRIBUTION INSIGHTS

Insights from Sentiment Analysis:

- **Promoters (Positive Sentiments):**

Majority: The largest segment (64.2%) consists of customers who have left positive reviews, indicating overall satisfaction with the products and services. These customers are likely to be loyal and can act as brand advocates.

- **Neutral Sentiments:**

Significant Portion: A notable portion (26.7%) of reviews are neutral. These reviews often contain factual feedback without strong emotions, which can be valuable for understanding product functionality and areas for improvement.

- **Detractors (Negative Sentiments):**

Smaller Segment: A smaller segment (9.2%) consists of negative reviews. While this is a minority, addressing the issues raised by these customers is crucial for improving satisfaction and reducing churn.

MARKETING STRATEGY FOR SENTIMENT-BASED ENGAGEMENT

- 1. Detractor Engagement Campaign

Target Audience: High-value customers (RFM scores: 444, 434, 443, 433, 442, 432) and churn probability greater than 70%

Channel: Email primary, followed by personalized phone call.

- Subject: "[Customer Name], We Value Your Feedback - Let's Make Things Right"
- Content:
 - Acknowledge their status as a valued customer
 - Express concern about their recent experiences
 - Offer to schedule a personal call with a dedicated customer success manager

Phone Call Follow-up

Loyalty Recovery Offer:

- 20% discount on next purchase

MARKETING STRATEGY FOR SENTIMENT-BASED ENGAGEMENT

- **2. Promoter Amplification Campaign**

Target Audience: High-value customers (RFM scores: 444, 434, 443, 433, 442, 432) and churn probability smaller than 30%

Channel: Email primary.

- Subject: "[Customer Name], Your Loyalty Deserves a Reward!"
- Content:
 - Thank them for their continued loyalty
 - Highlight their status as a top customer
 - Introduce the referral program and loyalty rewards

Social Media

- Platform: Instagram and Facebook
- Content: User-generated content featuring promoters' positive experiences
- Hashtag campaign: #[BrandName]LoyaltyRewards

Loyalty Amplification Offer:

- Double points on purchases for the next 30 days
- Exclusive early access to new products

CONCLUSION

Summary of Key Points:

- **Customer Retention Strategies:**

- RFM Analysis identified high-value customers at risk of churn.

- Churn Prediction Model leveraged Random Forest and SMOTE to predict churn risk.

- Marketing Campaigns designed to target high-value customers with personalized offers, reducing churn and increasing customer lifetime value (CLV).

- **Purchase Pattern Analysis:**

- Market Basket Analysis (MBA) revealed key association rules in product purchases.

- Cross-Selling Campaigns developed to recommend complementary products, aiming to increase average order value (AOV).

- **Sentiment Analysis:**

- Customer Feedback Analysis used sentiment labels to categorize reviews.

- Engagement Campaigns targeted detractors and promoters to address negative feedback and amplify positive word-of-mouth.