

# Customer Shopping Behavior Analysis - Comprehensive Report

## Executive Summary

This comprehensive report presents findings from an in-depth analysis of customer shopping behavior for a leading retail company. The analysis examines 3,900 customer transactions to identify purchasing patterns, customer segmentation, and key factors influencing consumer decisions. The insights derived from this study will enable data-driven decision-making to improve sales, enhance customer satisfaction, and build long-term loyalty.

## Key Highlights

- **Total Revenue Analyzed:** \$233,081
- **Customer Base:** 3,900 unique transactions
- **Gender Distribution:** Male customers contribute 67.7% of total revenue
- **Customer Loyalty:** 79.8% of customers are classified as "Loyal"
- **Top Performing Age Group:** Young Adults contribute \$62,143 in revenue

## Table of Contents

1. [Introduction](#)
2. [Methodology](#)
3. [Data Preparation & Quality](#)
4. [Exploratory Data Analysis](#)
5. [Business Questions & Analysis](#)
6. [Key Findings](#)
7. [Business Recommendations](#)
8. [Conclusion](#)
9. [Appendix](#)

## 1. Introduction

### 1.1 Business Context

The retail industry is experiencing rapid transformation with changing consumer behaviors, increased competition, and evolving market dynamics. Understanding customer shopping patterns has become crucial for maintaining competitive advantage and driving sustainable growth.

### 1.2 Business Problem

Management has observed significant changes in purchasing patterns across:

- Demographics (age, gender, location)
- Product categories (clothing, footwear, accessories, outerwear)
- Sales channels and customer engagement metrics

### 1.3 Project Objectives

The primary objectives of this analysis are to:

1. **Identify Revenue Drivers:** Understand which customer segments generate the most revenue
2. **Optimize Discount Strategies:** Analyze the effectiveness of discount campaigns
3. **Improve Product Quality:** Identify top-performing products based on customer reviews
4. **Enhance Customer Experience:** Compare shipping methods and payment preferences
5. **Increase Customer Loyalty:** Analyze subscription benefits and repeat purchase behavior
6. **Segment Customers:** Classify customers into actionable segments (New, Returning, Loyal)
7. **Optimize Inventory:** Identify top-selling products within each category
8. **Target Marketing:** Understand age group preferences and revenue contribution

## 2. Methodology

### 2.1 Data Collection

- **Source:** Customer transaction database
- **Time Period:** Historical transaction data
- **Sample Size:** 3,900 records
- **Data Format:** CSV (Comma-Separated Values)

### 2.2 Tools & Technologies

| Tool             | Purpose  |
|------------------|--|
| Python 3.x       | Data cleaning, transformation, and feature engineering |
| Pandas           | Data manipulation and analysis                         |
| NumPy            | Numerical computations                                 |
| PostgreSQL       | Relational database for structured storage             |
| SQLAlchemy       | Python-PostgreSQL connectivity                         |
| Jupyter Notebook | Interactive development environment                    |
| SQL              | Business intelligence queries and aggregations         |

### 2.3 Analytical Approach

The analysis followed a structured approach:

1. **Data Ingestion:** Import raw data into Python environment
2. **Data Quality Assessment:** Identify missing values, duplicates, and anomalies
3. **Data Cleaning:** Handle missing values using statistical imputation

4. **Feature Engineering:** Create derived features for enhanced analysis
5. **Data Standardization:** Normalize column names and data formats
6. **Database Loading:** Transfer cleaned data to PostgreSQL
7. **SQL Analysis:** Execute business queries for insights extraction
8. **Visualization:** Create charts and dashboards for stakeholder communication

## 3. Data Preparation & Quality

### 3.1 Initial Data Assessment

#### Dataset Characteristics:

- Total Records: 3,900
- Total Features: 18
- Data Types: Mixed (numeric, categorical, text)
- Missing Values: Detected in 'Review Rating' column

### 3.2 Data Quality Issues

| Issue               | Column                | Action Taken                                     |
|---------------------|-----------------------|--|
| Missing Values      | Review Rating         | Median imputation by category                    |
| Inconsistent Naming | All columns           | Lowercase with underscores                       |
| Redundant Column    | Promo Code Used       | Removed (100% correlation with Discount Applied) |
| Long Column Name    | Purchase Amount (USD) | Renamed to purchase_amount                       |

### 3.3 Feature Engineering

#### 3.3.1 Age Group Segmentation

Created categorical age groups using quartile-based binning:

- **Young Adult:** Ages 18-30
- **Adult:** Ages 31-43
- **Middle Aged:** Ages 44-56
- **Senior:** Ages 57-70

**Rationale:** Enables targeted marketing strategies for different age demographics.

#### 3.3.2 Purchase Frequency (Days)

Converted categorical frequency to numerical days:

| Frequency Label | Days |
|-----------------|------|
| Never           | 0    |
| Rarely          | 1-2  |
| Occasionally    | 3-4  |
| Frequently      | 5-6  |
| Always          | 7+   |

|                |     |
|----------------|-----|
| Weekly         | 7   |
| Bi-Weekly      | 14  |
| Fortnightly    | 14  |
| Monthly        | 30  |
| Quarterly      | 90  |
| Every 3 Months | 90  |
| Annually       | 365 |

**Rationale:** Facilitates quantitative analysis of purchase patterns.

### 3.4 Data Standardization

#### Transformations Applied:

- Column names converted to lowercase
- Spaces replaced with underscores
- Simplified column name: Purchase Amount (USD) → purchase\_amount

### 3.5 Final Dataset

#### Clean Dataset Characteristics:

- Records: 3,900
- Features: 18 (17 original + 1 derived)
- Missing Values: 0
- Duplicates: 0
- Data Quality Score: 100%

## 4. Exploratory Data Analysis

### 4.1 Demographic Distribution

#### Gender Distribution

- **Female:** 1,943 customers (49.8%)
- **Male:** 1,957 customers (50.2%)

**Insight:** Nearly balanced gender distribution provides representative insights for both segments.

#### Age Distribution

- **Minimum Age:** 18 years
- **Maximum Age:** 70 years
- **Mean Age:** 44.07 years

- **Median Age:** 44 years

#### Location Distribution

- **Total States Covered:** 50 states
- **Geographic Coverage:** National presence

#### 4.2 Purchase Behavior

##### Purchase Amount Statistics

- **Mean Purchase:** \$59.76
- **Median Purchase:** \$60.00
- **Minimum Purchase:** \$20.00
- **Maximum Purchase:** \$100.00
- **Standard Deviation:** \$23.67

**Insight:** Moderate variability in purchase amounts with most transactions between \$36 and \$83.

##### Previous Purchases

- **Mean Previous Purchases:** 25.35
- **Median Previous Purchases:** 25.00
- **Range:** 1 to 50 previous purchases

**Insight:** High customer retention with significant repeat purchase behavior.

#### 4.3 Product Categories

##### Category Distribution

1. **Clothing:** 1,734 transactions (44.5%)
2. **Footwear:** 976 transactions (25.0%)
3. **Accessories:** 976 transactions (25.0%)
4. **Outerwear:** 214 transactions (5.5%)

#### 4.4 Customer Engagement

##### Subscription Status

- **Subscribed:** 1,053 customers (27.0%)
- **Not Subscribed:** 2,847 customers (73.0%)

##### Review Rating Distribution

- **Mean Rating:** 3.75/5.00
- **Median Rating:** 3.75/5.00
- **Range:** 2.5 to 5.0

**Insight:** Generally positive customer satisfaction with products.

## 5. Business Questions & Analysis

### 5.1 Revenue Analysis by Gender

**SQL Query:**

```
SELECT gender, SUM(purchase_amount) as revenue
FROM customer
GROUP BY gender;
```

**Results:**

| Gender | Revenue   | Percentage |
|--------|-----------|------------|
| Male   | \$157,890 | 67.7%      |
| Female | \$75,191  | 32.3%      |

**Analysis:**

- Male customers contribute more than double the revenue of female customers
- This disparity could indicate:
  - Higher average purchase amounts among males
  - Greater frequency of purchases by males
  - Product mix more appealing to male customers

**Business Implication:**

- Consider expanding product lines that appeal to female customers
- Investigate if marketing efforts are gender-balanced
- Explore opportunities to increase female customer engagement

### 5.2 High-Value Discount Users

**SQL Query:**

```
SELECT customer_id, purchase_amount
FROM customer
WHERE discount_applied = 'Yes'
AND purchase_amount >= (SELECT AVG(purchase_amount) FROM customer);
```

**Results:**

- **Total Customers:** 1,956 customers
- **Characteristic:** Used discounts but spent above average (\$59.76)

**Sample High-Value Discount Users:**

| Customer ID | Purchase Amount |
|-------------|-----------------|
|-------------|-----------------|

|   |  |      |
|---|--|------|
| 2 |  | \$64 |
| 3 |  | \$73 |
| 4 |  | \$90 |
| 7 |  | \$85 |
| 9 |  | \$97 |

#### **Analysis:**

- 50% of customers who used discounts still made above-average purchases
- Discounts are not necessarily associated with low-value transactions
- These customers represent optimal discount targets: price-sensitive but high-spending

#### **Business Implication:**

- Discounts can drive higher-value purchases when targeted correctly
- Consider personalized discount strategies for high-spending customers
- Monitor discount ROI to ensure profitability

### **5.3 Top-Rated Products**

#### **SQL Query:**

```
SELECT item_purchased,
       ROUND(AVG(review_rating::numeric), 2) as avg_rating
  FROM customer
 GROUP BY item_purchased
 ORDER BY avg_rating DESC
 LIMIT 5;
```

#### **Results:**

| Rank | Product | Average Rating |
|------|---------|----------------|
| 1    | Gloves  | 3.86           |
| 2    | Sandals | 3.84           |
| 3    | Boots   | 3.82           |
| 4    | Hat     | 3.80           |
| 5    | Skirt   | 3.78           |

#### **Analysis:**

- Top products span multiple categories (accessories, footwear, clothing)
- All top products exceed the overall average rating (3.75)

- Rating differences are modest (0.08 range), indicating consistent quality

#### **Business Implication:**

- Feature top-rated products prominently in marketing materials
- Study these products to identify quality factors to replicate
- Use customer reviews as testimonials for these items
- Consider bundling top-rated products

#### **5.4 Shipping Method Comparison**

##### **SQL Query:**

```
SELECT shipping_type,
       ROUND(AVG(purchase_amount), 2) as avg_purchase
  FROM customer
 WHERE shipping_type IN ('Standard', 'Express')
 GROUP BY shipping_type;
```

##### **Results:**

| Shipping Type | Average Purchase | Difference |
|---------------|------------------|------------|
| Express       | \$60.48          | +\$2.02    |
| Standard      | \$58.46          | Baseline   |

##### **Analysis:**

- Express shipping customers spend 3.5% more on average
- This could indicate:
  - Higher urgency/value placed on items
  - Wealthier or less price-sensitive customer segment
  - Premium product purchases requiring faster delivery

##### **Business Implication:**

- Consider offering express shipping incentives to increase order values
- Test dynamic pricing strategies for shipping
- Analyze if expedited shipping improves customer satisfaction and retention

#### **5.5 Subscription Impact Analysis**

##### **SQL Query:**

```
SELECT subscription_status,
       COUNT(customer_id) AS total_customers,
       ROUND(AVG(purchase_amount), 2) AS avg_spend,
       ROUND(SUM(purchase_amount), 2) AS total_revenue
  FROM customer
```

```
GROUP BY subscription_status;
```

### Results:

| Status              | Customers | Avg Spend | Total Revenue | % of Total |
|---------------------|-----------|-----------|---------------|------------|
| Yes (Subscribed)    | 1,053     | \$59.49   | \$62,645      | 26.9%      |
| No (Not Subscribed) | 2,847     | \$59.87   | \$170,436     | 73.1%      |

### Analysis:

- Only 27% of customers are subscribed
- Average spend is nearly identical between groups (\$0.38 difference)
- Non-subscribers generate significantly more total revenue due to larger base

**Surprising Finding:** Subscribers don't spend more per transaction, contrary to typical subscription model expectations.

### Business Implication:

- Subscription model may need reevaluation
- Consider adding exclusive benefits to increase subscription value proposition
- Investigate subscription cancellation reasons
- Test tiered subscription models with varying benefits

## 5.6 Products with Highest Discount Usage

### SQL Query:

```
SELECT item_purchased,
       ROUND(100.0 * SUM(CASE WHEN discount_applied = 'Yes'
                               THEN 1 ELSE 0 END) / COUNT(*), 2) AS discount_rate
  FROM customer
 GROUP BY item_purchased
 ORDER BY discount_rate DESC
 LIMIT 5;
```

### Results:

| Rank | Product  | Discount Rate |
|------|----------|---------------|
| 1    | Hat      | 50.00%        |
| 2    | Sneakers | 49.66%        |
| 3    | Coat     | 49.07%        |
| 4    | Sweater  | 48.17%        |

|   |       |        |
|---|-------|--------|
| 5 | Pants | 47.37% |
|---|-------|--------|

### Analysis:

- Approximately 50% of purchases for these products used discounts
- These items may require discounts to move inventory
- Could indicate:
  - Overpriced base prices
  - High competition in these product segments
  - Seasonal items requiring promotional support

### Business Implication:

- Review base pricing for these high-discount products
- Assess if discounts are eroding profit margins
- Consider alternative promotional strategies (bundles, loyalty rewards)
- Evaluate if these products truly need heavy discounting

## 5.7 Customer Segmentation

### SQL Query:

```

WITH customer_type AS (
  SELECT customer_id, previous_purchases,
  CASE
    WHEN previous_purchases = 1 THEN 'New'
    WHEN previous_purchases BETWEEN 2 AND 10 THEN 'Returning'
    ELSE 'Loyal'
  END AS customer_segment
  FROM customer
)
SELECT customer_segment,
  COUNT(*) AS customer_count
FROM customer_type
GROUP BY customer_segment;
  
```

### Results:

| Segment   | Count | Percentage | Definition              |
|-----------|-------|------------|-------------------------|
| Loyal     | 3,116 | 79.9%      | >10 previous purchases  |
| Returning | 701   | 18.0%      | 2-10 previous purchases |
| New       | 83    | 2.1%       | 1 previous purchase     |

### Analysis:

- Exceptionally high loyalty rate (79.9%)
- Very small new customer acquisition (2.1%)

- Strong customer retention once acquired

#### **Customer Lifetime Value Implications:**

- Loyal customers represent the most valuable segment
- New customer acquisition appears to be a challenge
- Focus should be on both retention AND acquisition

#### **Business Implication:**

- Develop VIP programs for loyal customers (3,116 customers)
- Create onboarding experiences for new customers to accelerate loyalty
- Implement retention campaigns for returning customers
- Investigate marketing channels for new customer acquisition

#### **5.8 Top Products by Category**

##### **SQL Query:**

```

WITH item_counts AS (
    SELECT category,
           item_purchased,
           COUNT(customer_id) AS total_orders,
           ROW_NUMBER() OVER (PARTITION BY category
                               ORDER BY COUNT(customer_id) DESC) AS item_rank
    FROM customer
   GROUP BY category, item_purchased
)
SELECT item_rank, category, item_purchased, total_orders
  FROM item_counts
 WHERE item_rank <= 3;

```

##### **Results:**

| Category    | Rank | Product    | Total Orders |
|-------------|------|------------|--------------|
| Accessories | 1    | Jewelry    | 171          |
|             | 2    | Sunglasses | 161          |
|             | 3    | Belt       | 161          |
| Clothing    | 1    | Blouse     | 171          |
|             | 2    | Pants      | 171          |
|             | 3    | Shirt      | 169          |
| Footwear    | 1    | Sandals    | 160          |
|             | 2    | Shoes      | 150          |

|           |   |          |     |
|-----------|---|----------|-----|
|           | 3 | Sneakers | 145 |
| Outerwear | 1 | Jacket   | 163 |
|           | 2 | Coat     | 161 |
|           | 3 | Dress    | 149 |

#### Analysis:

- Order volumes are remarkably consistent across categories
- Top products differ by only 10-26 orders within categories
- Suggests balanced inventory distribution

#### Category Insights:

1. **Accessories:** Jewelry leads, followed by sunglasses and belts
2. **Clothing:** Blouses and pants are equally popular
3. **Footwear:** Sandals slightly outperform shoes and sneakers
4. **Outerwear:** Jackets are the clear category leader

#### Business Implication:

- Ensure adequate stock levels for top 3 items in each category
- Create category-specific promotional bundles
- Use top products as "hero items" in category displays
- Monitor seasonal variations for outerwear and footwear

## 5.9 Repeat Buyers and Subscription Correlation

#### SQL Query:

```
SELECT subscription_status,
       COUNT(customer_id) AS repeat_buyers
  FROM customer
 WHERE previous_purchases > 5
 GROUP BY subscription_status;
```

#### Results:

| Subscription Status | Repeat Buyers (>5 purchases) |
|---------------------|------------------------------|
| No                  | 2,518                        |
| Yes                 | 958                          |

#### Detailed Analysis:

- **Total Repeat Buyers:** 3,476 (89.1% of all customers)
- **Subscribed Repeat Buyers:** 958 (27.6% of repeat buyers)

- **Non-Subscribed Repeat Buyers:** 2,518 (72.4% of repeat buyers)

**Key Finding:** The majority of repeat buyers (72.4%) do NOT subscribe, despite demonstrating loyalty through multiple purchases.

**Analysis:** This represents a significant opportunity gap:

- Customers are loyal but not converting to subscriptions
- Subscription benefits may not be compelling enough
- Price sensitivity or perceived lack of value in subscription

**Business Implication:**

- **Critical Action Item:** Investigate why loyal customers aren't subscribing
- Survey repeat buyers to understand barriers to subscription
- Redesign subscription benefits to appeal to repeat buyers
- Test subscription incentives (discounts, free shipping, exclusive products)
- Consider a loyalty program as an alternative to subscriptions

## 5.10 Revenue Contribution by Age Group

**SQL Query:**

```
SELECT age_group,
       SUM(purchase_amount) AS total_revenue
  FROM customer
 GROUP BY age_group
 ORDER BY total_revenue DESC;
```

**Results:**

| Age Group           | Revenue  | Percentage | Customer Count |
|---------------------|----------|------------|----------------|
| Young Adult (18-30) | \$62,143 | 26.7%      | ~975           |
| Middle Aged (44-56) | \$59,197 | 25.4%      | ~975           |
| Adult (31-43)       | \$55,978 | 24.0%      | ~975           |
| Senior (57-70)      | \$55,763 | 23.9%      | ~975           |

**Analysis:**

- Revenue is relatively evenly distributed across age groups
- Young Adults contribute the most (\$62,143)
- Seniors contribute the least (\$55,763)
- Difference between highest and lowest is only \$6,380 (11.4%)

**Insights by Age Group:**

1. **Young Adults (18-30)**

- Highest revenue contributors
- Likely tech-savvy, active shoppers
- Potential for digital marketing strategies

## 2. Middle Aged (44-56)

- Second-highest revenue
- Likely established careers, higher disposable income
- May prefer quality over price

## 3. Adults (31-43)

- Moderate revenue contribution
- Possibly balancing family and career expenses
- May be more price-conscious

## 4. Seniors (57-70)

- Lowest revenue contribution
- May have different product preferences
- Could be more selective in purchases

### **Business Implication:**

- Develop age-specific marketing campaigns
- Young Adults: Focus on trends, social media, influencer marketing
- Middle Aged/Adults: Emphasize quality, value, convenience
- Seniors: Highlight comfort, reliability, customer service
- Consider age-appropriate product recommendations
- Test age-segmented email campaigns

## **6. Key Findings**

### **6.1 Revenue Insights**

1. **Gender Gap:** Male customers generate 67.7% of total revenue (\$157,890 vs \$75,191)
2. **Age Factor:** Young Adults (18-30) are the highest revenue contributors (\$62,143)
3. **Total Revenue:** \$233,081 from 3,900 transactions
4. **Average Transaction:** \$59.76 per purchase

### **6.2 Customer Loyalty Findings**

1. **Exceptional Retention:** 79.9% of customers are "Loyal" (>10 purchases)
2. **Acquisition Challenge:** Only 2.1% are new customers (1 previous purchase)
3. **Subscription Paradox:** Loyal customers (with many repeat purchases) are not subscribing
4. **Repeat Buyer Profile:** 89.1% of customers have made more than 5 purchases

### **6.3 Product Performance**

1. **Top-Rated Products:** Gloves (3.86), Sandals (3.84), Boots (3.82)
2. **High-Discount Products:** Hat (50%), Sneakers (49.66%), Coat (49.07%)
3. **Category Leaders:**
  - Accessories: Jewelry (171 orders)
  - Clothing: Blouse & Pants (171 orders each)
  - Footwear: Sandals (160 orders)
  - Outerwear: Jacket (163 orders)

#### 6.4 Marketing & Operations

1. **Discount Effectiveness:** 50% of discount users still spend above average
2. **Shipping Preference:** Express shipping customers spend 3.5% more (\$60.48 vs \$58.46)
3. **Subscription Status:** 73% of customers are not subscribed despite high loyalty
4. **Review Ratings:** Average 3.75/5.0 indicating good overall satisfaction

#### 6.5 Customer Segmentation

- **Loyal Customers:** 3,116 (79.9%) - Core revenue base
- **Returning Customers:** 701 (18.0%) - Growth opportunity
- **New Customers:** 83 (2.1%) - Acquisition challenge

### 7. Business Recommendations

#### 7.1 Immediate Actions (0-3 Months)

##### 1. Expand Female Customer Engagement

**Priority:** High **Issue:** Female customers contribute only 32.3% of revenue

##### Actions:

- Conduct focus groups with female customers to understand preferences
- Expand product lines appealing to female demographics
- Create gender-specific marketing campaigns
- Analyze competitor offerings for female customers
- Test targeted social media campaigns for women

**Expected Impact:** 15-20% increase in female customer revenue

##### 2. Redesign Subscription Program

**Priority:** High **Issue:** 72.4% of repeat buyers don't subscribe despite loyalty

##### Actions:

- Survey non-subscribing repeat buyers to identify barriers
- Redesign subscription benefits package:
  - Exclusive early access to new products

- Free express shipping
- Birthday/anniversary discounts
- VIP customer service line
- Points-based reward system
- Create tiered subscription levels (Silver, Gold, Platinum)
- Offer trial subscription periods

**Expected Impact:** Increase subscription rate from 27% to 45%

### 3. Optimize Discount Strategy

**Priority:** Medium **Issue:** 50% of certain products are sold with discounts

**Actions:**

- Review base pricing for high-discount products (Hat, Sneakers, Coat, Sweater, Pants)
- Test reduced discount frequencies to assess impact on sales
- Implement dynamic pricing based on inventory levels
- Focus discounts on new customer acquisition rather than all purchases
- Create bundle offers as alternative to item-level discounts

**Expected Impact:** Improve profit margins by 5-8%

## 7.2 Short-Term Initiatives (3-6 Months)

### 4. New Customer Acquisition Campaign

**Priority:** High **Issue:** Only 2.1% of customers are new

**Actions:**

- Develop referral program: "Refer a friend, both get 20% off"
- Invest in targeted digital advertising (Google Ads, Facebook, Instagram)
- Partner with influencers to reach new audiences
- Offer first-purchase discounts for new customers
- Create compelling "new customer welcome kit"
- Optimize website for SEO and user experience

**Expected Impact:** Increase new customer acquisition by 50-75%

### 5. Age-Targeted Marketing

**Priority:** Medium **Issue:** Untapped potential in specific age segments

**Actions:**

- **Young Adults:** Social media campaigns, influencer partnerships, trend-focused content
- **Middle Aged:** Email marketing emphasizing quality and convenience
- **Adults:** Family-oriented promotions, work-from-home product lines

- **Seniors:** Traditional marketing channels, emphasis on comfort and reliability
- Create age-specific product recommendations on website

**Expected Impact:** 10-15% increase in conversion rates per age group

## 6. Product Quality Excellence Program

**Priority:** Medium **Issue:** Leverage top-rated products for brand building

**Actions:**

- Feature top-rated products (Gloves, Sandals, Boots, Hat, Skirt) prominently
- Use positive reviews in marketing materials
- Create "Customer Favorites" section on website
- Implement quality control benchmarks based on top products
- Train staff on features that make these products successful

**Expected Impact:** Overall review rating increase from 3.75 to 4.0

## 7.3 Long-Term Strategies (6-12 Months)

### 7. Comprehensive Loyalty Program

**Priority:** High **Issue:** Need to reward and retain 79.9% loyal customer base

**Actions:**

- Implement points-based loyalty program
- Create VIP tiers with increasing benefits
- Offer exclusive products for loyal customers
- Host special events for top-tier members
- Develop loyalty app with gamification elements
- Provide birthday and anniversary rewards

**Expected Impact:** Increase customer lifetime value by 25-30%

## 8. Shipping Optimization

**Priority:** Medium **Issue:** Express customers spend more but may face higher costs

**Actions:**

- Negotiate better rates with shipping partners
- Offer free express shipping threshold (\$75+)
- Test subscription benefit: free express shipping for members
- Implement "ship from store" for faster local delivery
- Provide real-time tracking for all shipments

**Expected Impact:** Increase average order value by 5-10%

## 9. Category-Specific Strategies

**Priority:** Medium **Issue:** Optimize performance for each product category

#### **Actions by Category:**

##### **Accessories (25% of sales):**

- Promote jewelry bundles (earrings + necklace)
- Create seasonal sunglasses collections
- Offer belt + wallet combinations

##### **Clothing (44.5% of sales - largest category):**

- Expand blouse and pants varieties
- Create outfit bundles (shirt + pants)
- Introduce "complete the look" recommendations

##### **Footwear (25% of sales):**

- Highlight comfort features of sandals
- Create seasonal shoe collections
- Offer shoe care products and bundles

##### **Outerwear (5.5% of sales - smallest category):**

- Expand jacket styles for different seasons
- Create coat + accessory bundles
- Market outerwear for various weather conditions

**Expected Impact:** 15% increase in category-specific sales

## **7.4 Technology & Analytics Investments**

### **10. Advanced Analytics Implementation**

**Priority:** High

#### **Initiatives:**

- Implement customer segmentation ML models
- Deploy recommendation engine for personalized shopping
- Create churn prediction model
- Build customer lifetime value (CLV) calculator
- Develop dynamic pricing algorithms
- Implement A/B testing framework for all initiatives

**Expected Impact:** Data-driven decisions increasing ROI by 20-30%

## **8. Conclusion**

### **8.1 Summary of Findings**

This comprehensive analysis of 3,900 customer transactions has revealed critical insights into shopping behavior, customer loyalty, and revenue drivers. The company demonstrates exceptional customer retention (79.9% loyal customers) but faces challenges in new customer acquisition (2.1%) and subscription conversion (27%).

## 8.2 Strategic Priorities

The analysis identifies three strategic priorities:

1. **Customer Acquisition:** Dramatically increase new customer acquisition through referral programs and targeted marketing
2. **Subscription Transformation:** Redesign the subscription program to convert loyal repeat buyers
3. **Revenue Diversification:** Balance gender revenue contribution and optimize discount strategies

## 8.3 Expected Business Impact

Implementation of recommended strategies is projected to deliver:

- **Revenue Growth:** 20-25% increase in first year
- **Customer Acquisition:** 50-75% more new customers
- **Subscription Rate:** Increase from 27% to 45%
- **Profit Margins:** Improvement of 5-8% through discount optimization
- **Customer Lifetime Value:** 25-30% increase through loyalty programs

## 8.4 Next Steps

### 1. Immediate (Week 1-2):

- Present findings to executive leadership
- Prioritize recommendations based on resources and impact
- Assign ownership for each initiative

### 2. Short-term (Month 1-3):

- Launch female customer research initiative
- Begin subscription program redesign
- Implement discount strategy optimization

### 3. Ongoing (Month 3+):

- Execute new customer acquisition campaigns
- Roll out age-targeted marketing
- Build comprehensive loyalty program
- Invest in analytics and technology infrastructure

## 8.5 Measuring Success

Establish KPIs to track progress:

- Monthly revenue growth by gender
- New customer acquisition rate

- Subscription conversion rate
- Average discount percentage by product
- Customer retention rate
- Net Promoter Score (NPS)
- Customer lifetime value (CLV)
- Return on marketing investment (ROMI)

## 9. Appendix

### 9.1 Data Dictionary

| Column Name            | Data Type | Description                                |
|------------------------|-----------|--|
| customer_id            | Integer   | Unique customer identifier                 |
| age                    | Integer   | Customer age (18-70)                       |
| gender                 | Text      | Male or Female                             |
| item_purchased         | Text      | Product name                               |
| category               | Text      | Clothing, Footwear, Accessories, Outerwear |
| purchase_amount        | Numeric   | Purchase value in USD                      |
| location               | Text      | US State                                   |
| size                   | Text      | S, M, L, XL                                |
| color                  | Text      | Product color                              |
| season                 | Text      | Winter, Spring, Summer, Fall               |
| review_rating          | Numeric   | Rating 1-5                                 |
| subscription_status    | Text      | Yes or No                                  |
| shipping_type          | Text      | Standard, Express, Free Shipping, etc.     |
| discount_applied       | Text      | Yes or N                                   |
| previous_purchases     | Integer   | Count of past purchases                    |
| payment_method         | Text      | Credit Card, Debit Card, PayPal, etc.      |
| frequency_of_purchases | Text      | Weekly, Monthly, Quarterly, etc.           |

|                         |         |   |
|-------------------------|---------|---|
| age_group               | Text    | Young Adult, Adult, Middle Aged, Senior |
| purchase_frequency_days | Integer | Numeric representation of frequency     |

## 9.2 SQL Query Repository

All SQL queries used in this analysis are available in the `customer_behavior.sql` file in the project repository.

## 9.3 Technical Specifications

- **Python Version:** 3.8+
- **PostgreSQL Version:** 12+
- **Key Libraries:** pandas 1.3+, numpy 1.21+, sqlalchemy 1.4+
- **Database Name:** Customer\_Behavior
- **Table Name:** customer

## 9.4 References

- Dataset: Customer Shopping Behavior (retail transaction data)
- Analysis Period: 2025
- Report Author: Kaushal Divekar
- Project Repository: [https://github.com/CodeWithKaushal/Customer\\_Shopping\\_Behavior\\_Analysis](https://github.com/CodeWithKaushal/Customer_Shopping_Behavior_Analysis)

## Report End

*This report is confidential and intended for internal business use only.*

## Document Information: