Retail Sales Analysis - Project Report

Executive Summary

This analysis examines retail sales data from October 2022 to March 2023, covering 75 transactions across 4 product categories. The dataset includes customer demographics, purchase patterns, and sales performance metrics, revealing key insights into customer behavior, category performance, and sales trends.

Dataset Overview

Data Summary

- **Total Records**: 75 transactions
- **Unique Customers**: 38 customers
- **Product Categories**: 4 (Clothing, Electronics, Beauty, Home)
- **Date Range**: October 2022 - March 2023
- **Gender Distribution**: 63.15% Male, 36.84% Female
- **Age Range**: 22-68 years

Data Quality Assessment

- **Initial Records**: 75 transactions
- **Records with NULL values**: 20 (26.66% of total)
- **Clean Records After Cleaning**: 57 (76.00% of total)
- **Common NULL Issues**: Missing prices, customer IDs, and transaction details

Key Findings & Insights

1. Sales Performance by Category

```
-- Electronics leads in total sales despite fewer transactions
SELECT category, COUNT(*) as transactions, SUM(total_sale) as total_revenue
FROM retail_sales
GROUP BY category
ORDER BY total_revenue DESC;
```

Results:

Category	Transactions	Total Revenue	Average Transaction Value
Electronics	15	$15,848.83	$1,056.59
Clothing	28	$2,962.75	$105.81
Beauty	17	$2,105.27	$123.84
Home	15	$2,063.57	$137.57

Insight: Electronics dominates revenue despite fewer transactions, indicating high-value purchases in this category.

2. Customer Segmentation Analysis

Top 5 Customers by Spending:

```
-- Top 5 customers contribute significantly to revenue
SELECT customer_id, COUNT(*) as purchases, SUM(total_sale) as total_spent
FROM retail_sales
GROUP BY customer_id
ORDER BY total_spent DESC
LIMIT 5;
```

Results:

Customer ID	Purchases	Total Spent	Category Preference
1007	5	$4,468.93	Electronics
1020	4	$3,659.89	Electronics/Beauty
1002	3	$1,499.96	Electronics
1001	4	$507.45	Clothing/Beauty
1003	3	$819.94	Home

Insight: Customer 1007 is the highest spender with $4,469 across 5 purchases, primarily in Electronics.

3. Temporal Analysis

Best Performing Months:

```
-- November 2022 was the best month overall
SELECT
    Year,
    Month,
    AverageSale
FROM (
        SELECT
            YEAR(sale_date) AS Year,
            MONTH(sale_date) AS Month,
            ROUND(AVG(total_sale),2) AS AverageSale,
            DENSE_RANK() OVER(PARTITION BY YEAR(sale_date) ORDER BY AVG(total_sale) DESC) AS AverageRank
        FROM retail_sales
        GROUP BY 1,2
    )t
WHERE AverageRank = 1;
```

Key Months:

1. **December 2022**: Peak performance with highest transactions and revenue
2. **February 2023**: Consistent sales with high-value transactions

Shift Analysis:

```
-- Afternoon is the busiest shopping time
WITH hourly_sale
AS
(
    SELECT
        *,
        CASE
            WHEN HOUR(sale_time) < 12 THEN "Morning"
            WHEN HOUR(sale_time) BETWEEN 12 AND 17 THEN "Afternoon"
            ELSE "Evening"
        END AS Shift
    FROM retail_sales
)
SELECT
    Shift,
    COUNT(*) AS TotalOrders
FROM hourly_sale
GROUP BY Shift;
```

Results:

Shift	Orders	Percentage
Afternoon	32	56.14%

Shift	Orders	Percentage
Morning	24	42.10%
Evening	1	1.75%

Insight: 56.14% of purchases occur in the afternoon (12 PM - 5 PM).

4. Demographic Insights

Gender Distribution by Category:

```
-- Gender preferences vary by category
SELECT
    gender,
    category,
    COUNT(*) AS TotalTransactions
FROM retail_sales
GROUP BY 1,2
ORDER BY 2;
```

Key Observations:

- **Electronics**: Male-dominated (100% of purchases)
- **Beauty**: Female-dominated (92.85% of purchases)
- **Clothing**: Balanced gender distribution (54.54% F & 45.45% M)
- **Home**: Female-dominated (80%)

Age Analysis by Category:

```
-- Beauty category attracts younger customers
SELECT category,
       ROUND(AVG(age), 1) as avg_age,
       MIN(age) as min_age,
       MAX(age) as max_age
FROM retail_sales
GROUP BY category
ORDER BY avg_age;
```

Results:

Category	Average Age	Age Range
Beauty	33.8 years	22-45
Clothing	37.4 years	22-68
Electronics	40.6 years	31-68
Home	41.3 years	27-65

Insight: Beauty products attract the youngest customers (avg 33.8 years), while Home products attract the oldest (avg 41.3 years).

🎯 Business Recommendations

1. Inventory Optimization

- **Increase Electronics Stock**: High-value category with premium pricing
- **Optimize Beauty Inventory**: Focus on younger demographic preferences
- **Seasonal Planning**: Stock up before November (peak month)

2. Marketing Strategies

- **Targeted Campaigns**:
  - **Electronics**: Target males aged 35-40
  - **Beauty**: Target females aged 25-35
  - **Afternoon Promotions**: Capitalize on peak shopping hours

- Loyalty Programs: Focus on top 5 customers who contribute 45% of revenue

3. Customer Experience Improvements

- **Personalized Recommendations**: Based on purchase history and demographics
- **Extended Afternoon Hours**: Align with peak shopping times
- **Age-Appropriate Marketing**: Different strategies for different age groups

4. Revenue Growth Opportunities

- **Cross-selling**: Electronics customers → accessories/insurance
- **Upselling**: Beauty customers → premium product lines
- **Bundling**: Create category-based bundles (e.g., Home decor sets)

Conclusion

The retail sales analysis reveals a healthy business with strong performance in Electronics, clear demographic patterns, and identifiable peak periods. Key success factors include:

- Strong Electronics category performance
- Loyal high-value customers
- Clear shopping time preferences
- Balanced category portfolio

Strategic Priority:

Focus on retaining top customers while expanding the customer base through targeted marketing in underperforming time slots and demographic segments.