# Customer Purchasing Analysis: Patterns, Preferences, and Trends in Retail

Prepared by

**Kingsley Agu**

## Introduction

This report offers a detailed analysis of customers purchasing habits to figure out patterns, preferences and trends. The dataset used for this analysis gives us a complete view of how people shop, including details about what they buy, who they are, and how they pay. It includes information like:

- When and where transactions happened.
- What products were bought and for how much.
- Customer details like age, location, and preferred payment methods.

## Objectives

To uncover meaningful insights and actionable recommendations, the following key questions were addressed during analysis:

1. What are the most popular and underperforming products?
2. Which age groups and locations generate the highest revenue?
3. What are the preferred payment methods, and how do they vary by season?
4. Which seasonal trends impact revenue and purchasing revenue?
5. How can these insights inform targeted marketing and sales strategies?

## Data Analysis Steps

### Step 1: Data Cleaning with SQL

To ensure the data was ready for analysis, I used SQL to clean and organize the dataset:

- **Database Creation**: Imported the dataset into a SQL database for efficient querying and cleaning.
- **Renaming Columns**: Updated column names for better readability and consistency using the snake_case format.
- **Data Validation**: Verified the cleaned dataset to ensure consistency and accuracy.

### Step 2: Exploratory Data Analysis (EDA)

- Performed exploratory analysis on transactional trends, product categories, and customer demographics using SQL queries.
- Identified revenue trends, customer preferences, and seasonal peaks by grouping and aggregating the data.

### Step 3: Key Metrics Calculation

Used SQL functions to calculate key metrics such as :

- **Total Revenue**: SUM(purchase_amount_usd)
- **Average Purchase Value (APV)**: AVG(purchase_amount_usd)
- **Location Analysis**: SUM(purchase_amount_usd) grouped by location.

### Step 4: Visualization Preparation

- Designed clear, professional visualizations from exported csv files of SQL queries for the key metrics calculations to convey insights effectively.

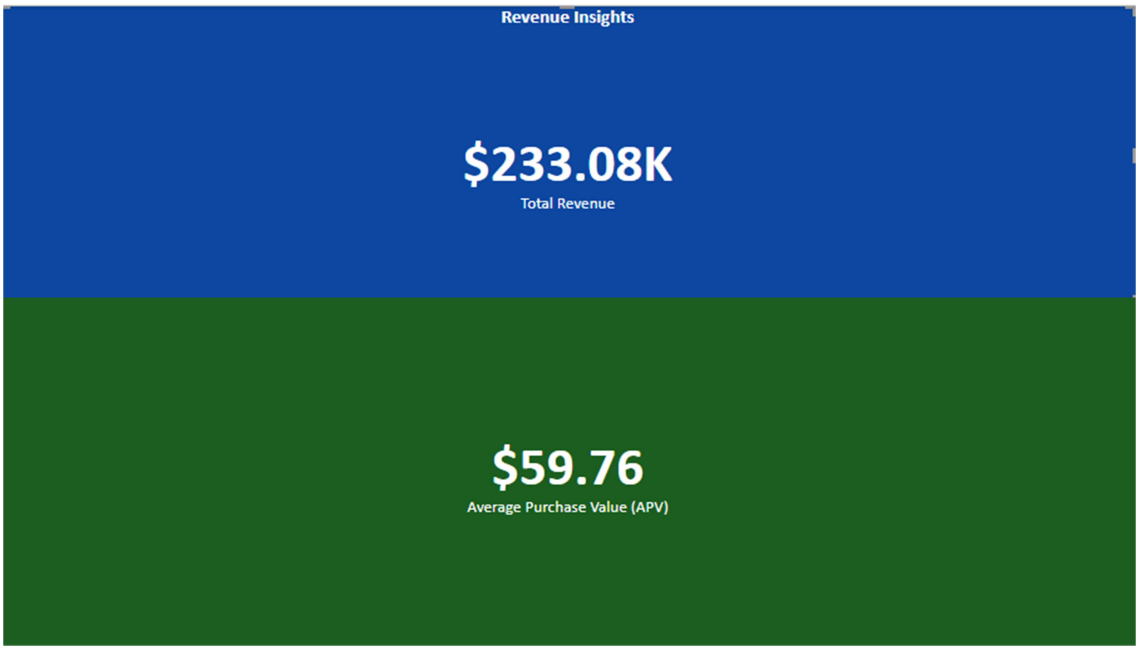
## Key Metrics Calculation with SQL

### Metric 1: Total Revenue and Average Purchase value (APV)

The **total revenue** measures the overall sales performance while the **average purchase value** measures the average amount spent per transaction.

```
#Total Revenue and Average Purchase Value
SELECT SUM(purchase_amount_usd) AS total_revenue,
AVG(purchase_amount_usd) AS average_purchase_value
FROM shopping_trends;
```

The total revenue generated is $233,081.00 and the average purchase value is $59.76.

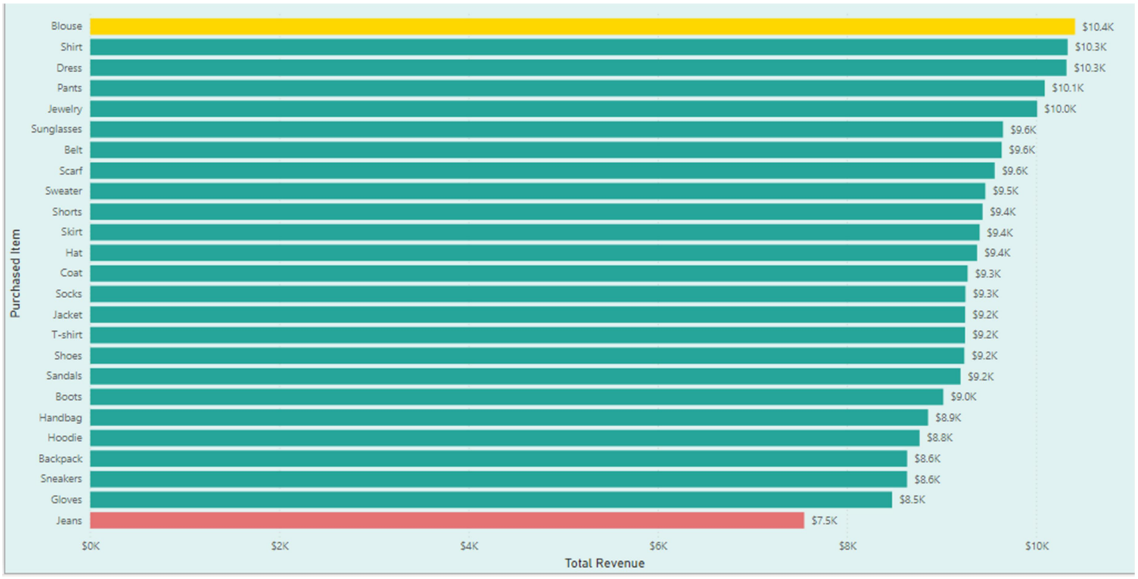


## Metric 2: Most Popular and Underperforming Products

This calculation measures the products with the highest and lowest total revenue.

```
#Most Popular Products(Items and Categories)
SELECT
item_purchased,
COUNT(*) AS items_sold,
SUM(purchase_amount_usd) AS total_revenue
FROM shopping_trends
GROUP BY item_purchased
ORDER BY items_sold DESC;
```

The **most popular** product is Blouse with total revenue of $10,410.00 and the **underperforming** product is Jean with total revenue of $7,548.00.

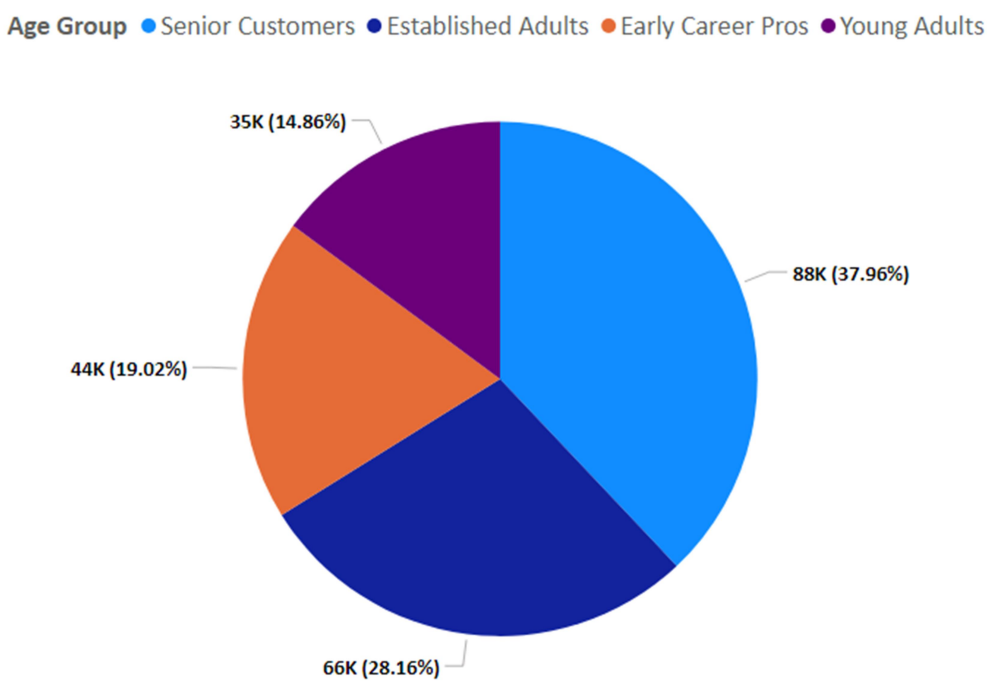



## Metric 3: Customer Demographics by Age Group

This calculation aims to analyze the revenue contribution across different age groups.

```
-- By Age Group
SELECT
    CASE
    WHEN age < 18 THEN 'Under 18'
    WHEN age BETWEEN 18 AND 25 THEN 'Young Adults'
    WHEN age BETWEEN 26 AND 35 THEN 'Early Career Pros'
    WHEN age BETWEEN 36 and 50 THEN 'Established Adults'
    ELSE 'Senior Customers'
    END AS age_group,
    COUNT(*) AS customers,
    SUM(purchase_amount_usd) AS total_revenue
    FROM shopping_trends
```

The result shows that the Senior Customers and Established Adults combined generate more than 50% of the overall total revenue with $88,480.00 and $65,629.00 respectively.

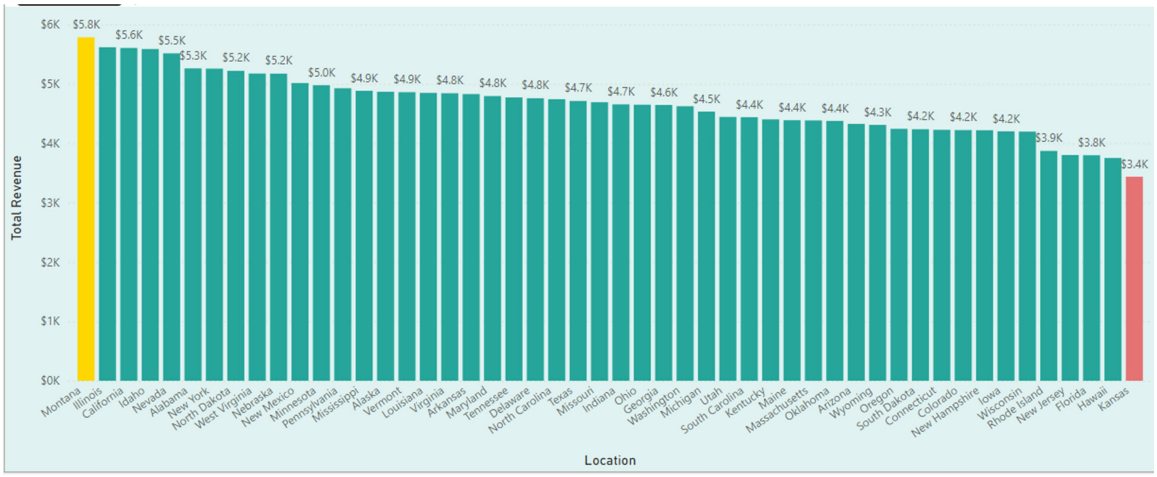



## Metric 4: Revenue Distribution by Location

This calculation examines which locations generate the highest and lowest total revenue.

```
#Location Analysis
SELECT
location,
COUNT(*) AS customers,
SUM(purchase_amount_usd) AS total_revenue
FROM shopping_trends
GROUP BY location
ORDER BY total_revenue DESC;
```

The result shows that Montana has the highest total revenue of $5,800.00 and Kansas has the lowest total revenue of $3,400.00.

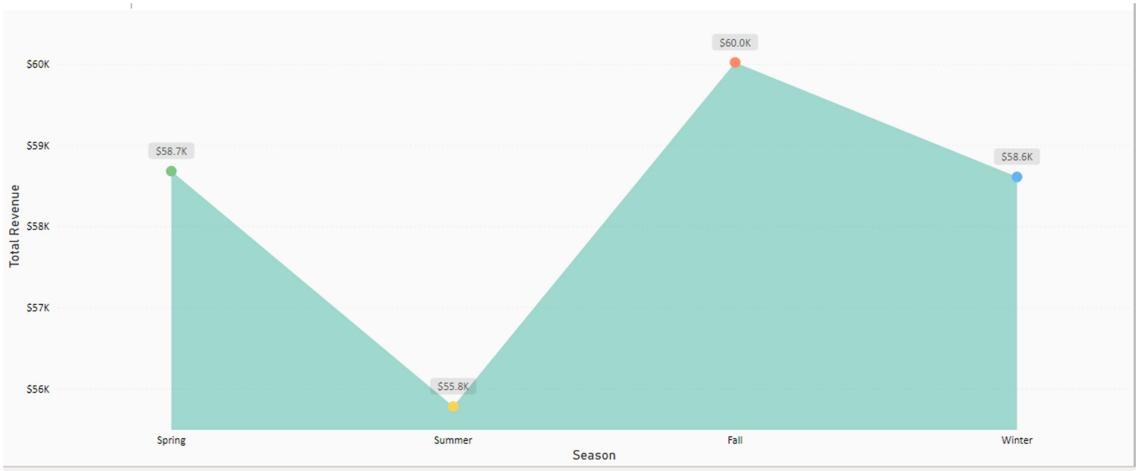



## Metric 5: Seasonal Trends

This calculation analyzes how revenue fluctuates across different seasons of the year.

```sql
#Seasonal Trends
SELECT
season,
COUNT(item_purchased) AS items_sold,
SUM(purchase_amount_usd) AS total_revenue
FROM shopping_trends
GROUP BY season
ORDER BY total_revenue DESC;
```

The result shows that revenue has not really been consistent with up and down movement across the different seasons of the year.

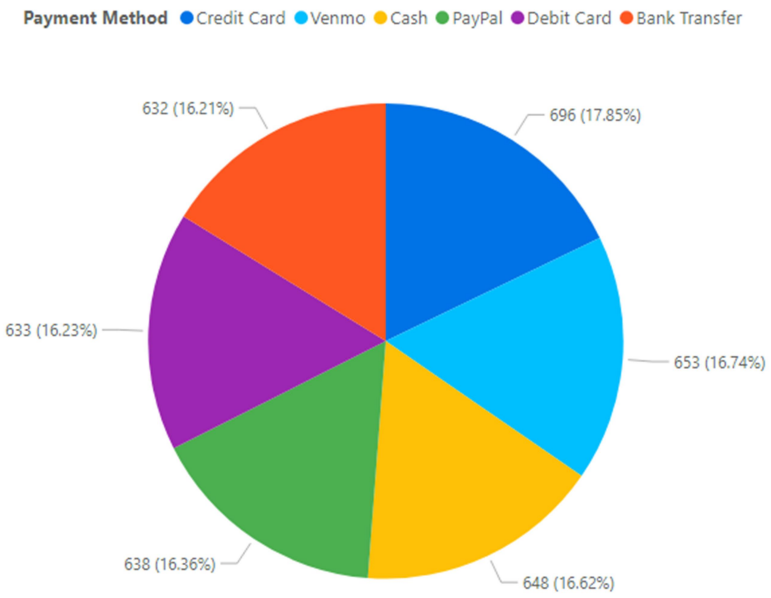


## Metric 6: Preferred Payment Methods Distribution

This calculation shows the distribution of transactions across different payment methods.

```
#Preferred Payment Methods
SELECT
payment_method,
COUNT(item_purchased) AS transactions,
SUM(purchase_amount_usd) AS total_revenue
FROM shopping_trends
GROUP BY payment_method
ORDER BY total_revenue DESC;
```

The result shows that the most used payment method is **Credit Card** while the least used payment method is **Bank Transfer**.




## Findings and Recommendations

### Product Performance

1. **Most Popular Product**:
   - Blouses are the best sellers, earning $10,410.00.
   - **Recommendation**: Stock up on blouses and promote them.
2. **Underperforming Product**:
   - Jeans make less money ($7,548.00) and don't sell as well.
   - **Recommendation**: Look into why Jeans aren't selling and either

## Customer Demographics

1. **Who Spends the Most:**
    - Senior customers (50+) are the biggest spenders, contributing 37.96% of the total revenue, and they love shoes.
    - **Recommendation**: Focus on marketing shoes and other products to senior customers.
2. **What Each Age Group Likes**:
    - **18 – 25 (Young Adults)**: Dresses($1917.00)
    - **26 – 35 (Early Career Pros)**: Shirts ($2861.00)
    - **36 – 50 (Established Adults)**: Scarves ($3,632.00)
    - **50+ (Senior Customers)**: Shoes ($4,217.00)
    - **Recommendation**: Advertise products that each age group prefers to them.

## Revenue by Location

1. **Top Location**:
    - Montana makes the most money ($5,784.00).
2. **Lowest Location**:
    - Kansas makes the least money ($3,437.00).
    - **Recommendation**: Find out why sales are low in Kansas and try to improve sales there.

## Seasonal Trends

1. **Best and Worst Season**:
    - Fall makes the most money ($60,018.00) while Summer makes the least ($55,777.00).
    - **Recommendation**: Focus sales and promotions in Fall when people spend the most.
2. **Seasonal Payment Methods**:
    - Credit cards are most used in Fall and Winter. Bank Transfers and Venmo are popular in Spring and Winter.
    - **Recommendation**: Offer deals based on how people like to pay during different seasons.

## Payment Methods

1. **Most and Least Popular Methods**:
    - Credit Cards are used the most, while Bank Transfers are the least used.
    - **Recommendation**: Make it easier for people to use Credit Cards and encourage other methods like Bank Transfer with promotions.
2. **How Different Age Groups Pay**:
    - **18 – 25 (Young Adults)**: Prefer PayPal.
    - **26 – 35 (Early Career Pros)**: Like Debit Cards and Credit Cards equally.
    - **36 – 50 (Established Adults)**: Prefer Credit Cards.
    - **50+ (Senior Customers)**: Mostly use Credit Cards.
    - **Recommendation**: Offer discounts or bonuses for their favorite payment methods.

## Conclusion

This analysis offers a data-driven view of customer shopping trends. By leveraging these insights, businesses can implement targeted strategies to maximize sales, improve customer satisfaction, and streamline operational efficiencies.