

Customer Shopping Behavior Analysis

This project analyzes customer shopping behavior using transactional data from 4000 purchases. We aim to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to guide strategic business decisions.



Project Overview & Dataset Summary

Project Goal

Uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to guide strategic business decisions.

Dataset Size

4000 rows, 18 columns of transactional data.

Key Features

Demographics, purchase details, shopping behavior (e.g., discounts, reviews, frequency).

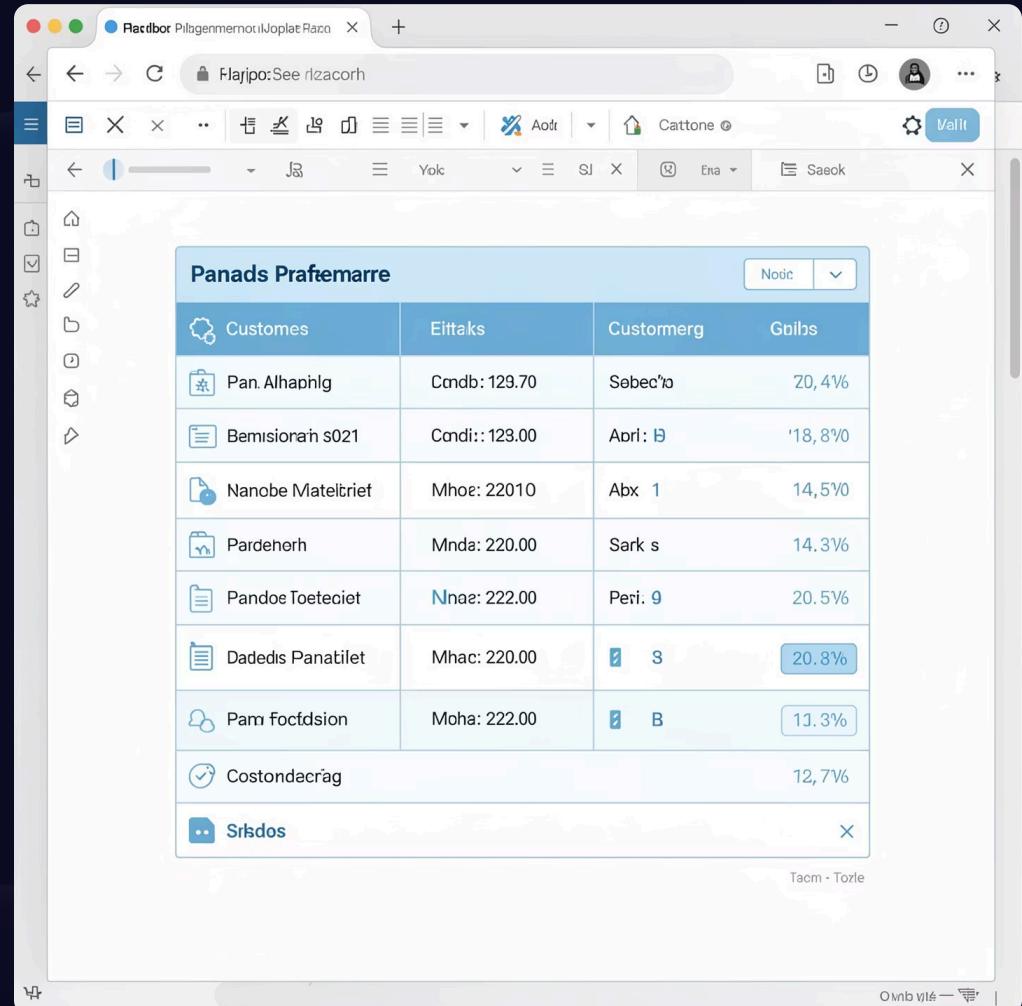
Missing Data

37 values in the Review Rating column.

Dataset Collection & Initial Review

The data was sourced from Kaggle and loaded into Python for exploratory data analysis. This initial step ensures data accessibility and provides a preliminary view of its structure and content.

```
[1]: import pandas as pd  
import numpy as np  
  
[2]: df=pd.read_csv('shopping_trends .csv')  
  
[3]: df
```



A screenshot of a web browser window displaying a table titled "Panads Prafermarre". The table has four columns: "Customes", "Eittaks", "Customermerg", and "Gøths". The data rows are:

Customes	Eittaks	Customermerg	Gøths
Pan. Alhaphilg	Cndb: 129.70	Søbec"ø	20, 4%
Bemisiorah s021	Cnddi: 123.00	Apri: B	'18, 8%
Nanobe Mateltriet	Mhoe: 22010	Abx 1	14, 5%
Pardeherh	Mnda: 220.00	Sark s	14, 3%
Pandoe Toeteciet	Nnaæ: 222.00	Peri. 9	20, 5%
Dadedis Panatilet	Mhac: 220.00	2 S	20.8%
Pam Focfdision	Moha: 222.00	2 B	11.3%
Costondacrag			12, 7%
Srkdos			X

Data Cleaning & Preparation

01

Missing Data Handling

Checked data types, null values, and imputed missing values in 'Review Rating' using the median rating per product category.

02

Column Standardization

Renamed columns to SNAKE CASE for improved coding readability and consistency.

03

Feature Creation

Created 'age_group' by binning ages and 'purchase_frequency_days' from purchase data.

04

Data Consistency

Verified and dropped redundant 'promo_code_used' column.

Data Export & SQL Analysis

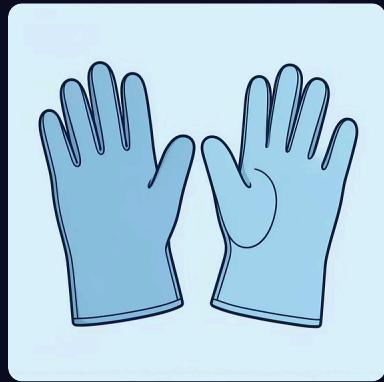
After cleaning, the data was saved to Excel/CSV and exported to MySQL for structured analysis. This allowed us to answer key business questions through targeted queries.

Example: Comparing total revenue generated by male vs. female customers.

Male	\$112,428
Female	\$52,365



Top Products & Shipping Insights



Top 5 Revenue Products

Gloves (3.89 avg rating), Sandals (3.85), Boots (3.82), Hat (3.81), Shorts (3.80).



Average Shipping Cost

Express shipping: \$61.34, Standard shipping: \$58.21.

Subscription & Discount Analysis

Subscribers vs. Non-Subscribers

Non-subscribers have higher total revenue (\$119,924) and average spend (\$59.66) compared to subscribers (\$44,869 total revenue, \$59.12 avg spend).

Customers making >5 purchases are more likely to subscribe (689 Yes vs. 1787 No).

Discounted Items

Sneakers (51.40% discount rate) have the highest number of purchases with discounts applied, followed by Jewelry (50.00%), Hat (49.57%), Sweater (49.56%), and Coat (47.97%).

Top Purchased Products by Category

1

Accessories

- Sunglasses (121)
- Hat (117)
- Scarf (116)

2

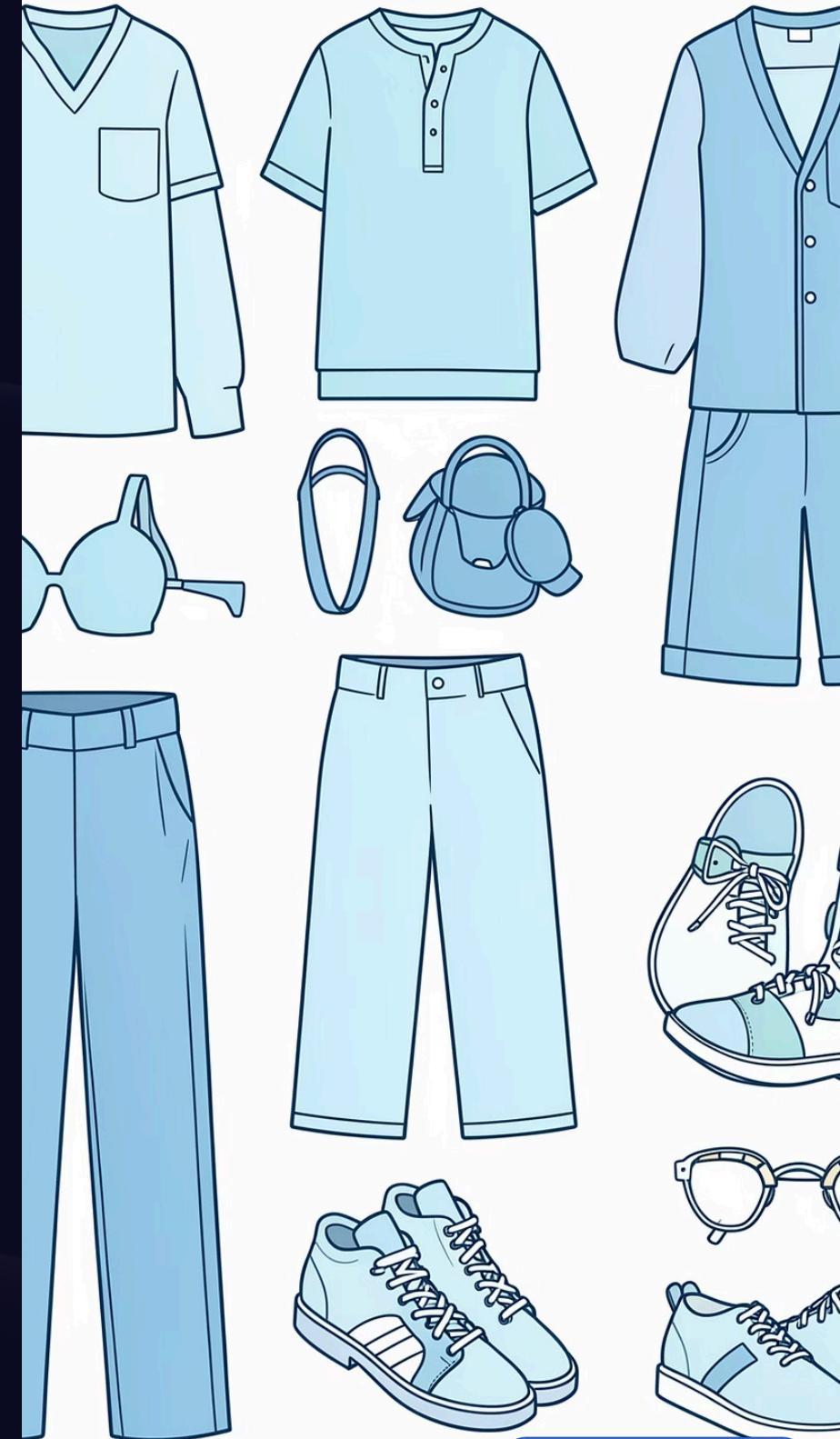
Clothing

- Pants (127)
- Shirt (126)
- Dress (118)

3

Footwear

- Sandals (112)
- Sneakers (105)
- Boots (98)



Customer Behavior Dashboard

Subscription_status

No

Yes

627

Number of customers

60.73

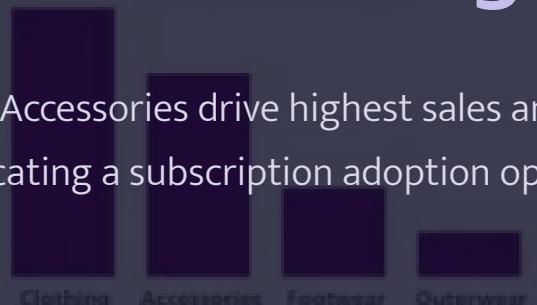
Average purchase number

3.76

Average review rating

Power BI Dashboard Insights

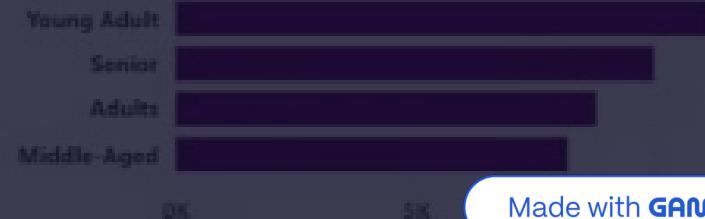
The dashboard reveals Clothing and Accessories drive highest sales and revenue. Young Adults contribute most to sales. Most customers are non-subscribers, indicating a subscription adoption opportunity. Actionable insights support data-driven marketing and sales strategies.



% of Customers by Subscribers



Sales by Category



Made with GAMMA

Business Recommendations

→ **Boost Subscriptions**

Promote exclusive benefits to increase subscriber adoption and retention.

→ **Customer Loyalty Programs**

Implement rewards for repeat buyers to foster loyalty.

→ **Review Discount Policy**

Optimize discount strategies to balance sales growth with profit margins.

→ **Product Positioning**

Highlight top-rated and best-selling products in marketing campaigns.

→ **Targeted Marketing**

Focus efforts on high-revenue age groups and users preferring express shipping.