



Customer Retail Analysis Project

Comprehensive Analysis of 20,000 Customer Transactions



PYTHON



SQL



POWER BI

Executive Summary

20,000 Records

Customer transaction records cleaned and processed

45,259 Missing Values

Handled across 22 data columns

13 SQL Queries

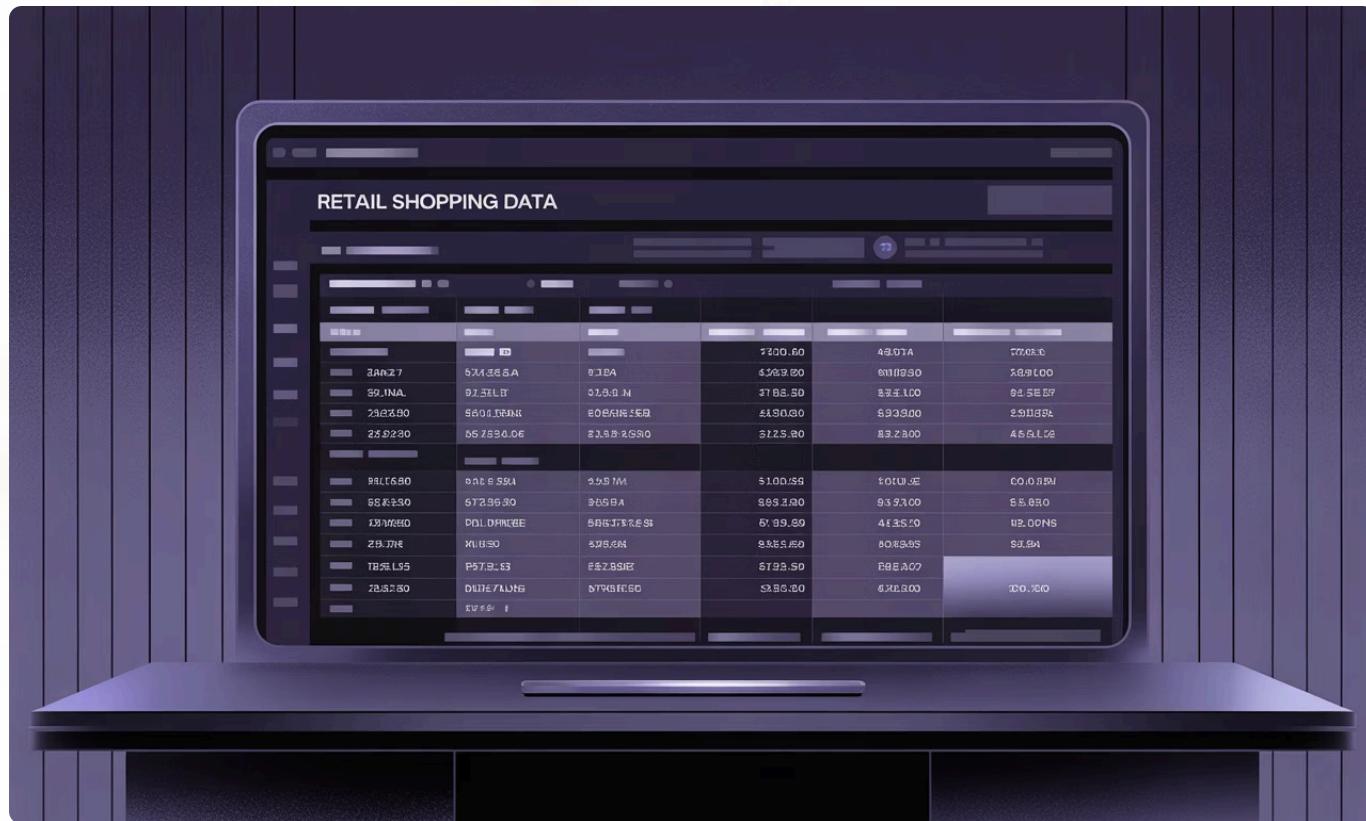
Executed for customer behavior analysis

Interactive Dashboard

Built with custom DAX measures in Power BI

This project demonstrates a complete data analytics workflow, transforming messy retail data into actionable business insights through systematic Python data cleaning, strategic MySQL analysis, and compelling Power BI visualization.

Project Overview & Dataset



Dataset Characteristics

- 20,000 customer transactions across 22 attributes
- 45,259 missing values (11.27% of total cells)
- Historical retail data covering multiple product categories
- Seven major geographic markets analyzed

Technology Stack

Python for data cleaning and exploration, MySQL for database queries and analysis, Power BI with DAX for interactive dashboards, and Excel for documentation.

Data Cleaning Process

01

Column Standardization

Cleaned column names by removing spaces and special characters, replacing spaces with underscores for consistency

03

Duplicate Removal

Identified and removed duplicate transactions to ensure data integrity and accuracy

Python libraries used: Pandas for data manipulation, NumPy for numerical operations, Seaborn and Matplotlib for statistical visualization.

02

Missing Value Treatment

Applied category-based median imputation for numerical data across 16 columns with missing values

04

Feature Engineering

Created Age_Group segments and calculated purchase frequency metrics for enhanced analysis

Key Statistical Findings

\$1,001

Average Purchase

Mean transaction value
across all customers

42

Median Age

Middle point of customer age
distribution

24

Previous Purchases

Average number of past
transactions per customer

3.0

Average Rating

Mean customer review score
out of 5 stars

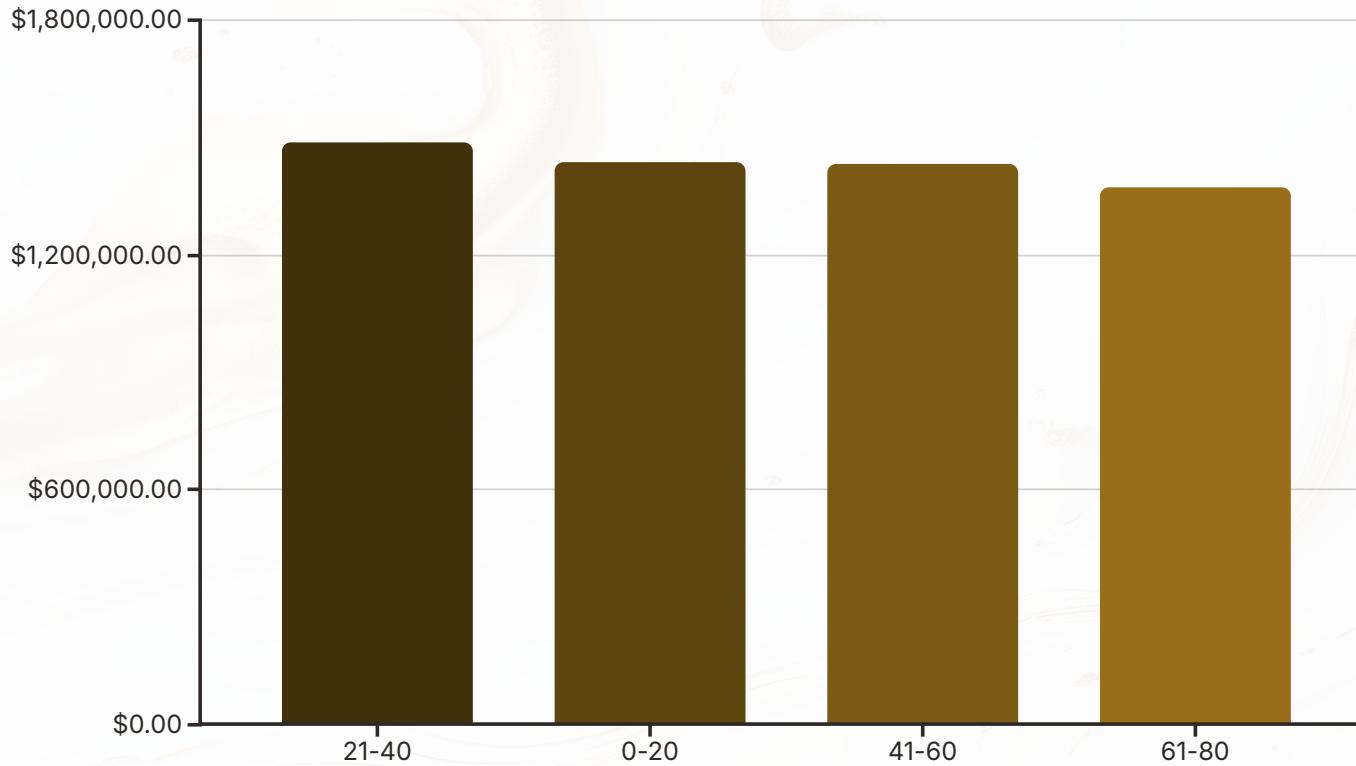
Customer Demographics

Age range spans 16-69 years with a median of 42. Geographic presence covers seven major markets including USA, UK, Germany, China, India, France, and Nigeria.

Purchase Patterns

Transaction values range from \$5 to \$1,999, with an average of \$1,001. Customers show strong repeat behavior with an average of 24 previous purchases.

SQL Analysis: Revenue Insights



Age Group Spending Analysis

The 21-40 age demographic emerges as the highest-spending segment at \$1.49M total revenue, making them the key target market for marketing initiatives.

Discount Impact Discovery

SQL analysis revealed that discounts slightly reduce average spend: discounted purchases average \$816.20 versus \$824.99 for non-discounted, suggesting discounts attract budget-conscious shoppers rather than driving higher spending.

Product Performance & Geographic Revenue



Shirts: \$1.15M

Top revenue generator,
nearly double the next
highest category



Watches: \$608K

Second-highest revenue
product with strong demand



Headsets: \$607K

Electronics category showing
consistent performance



Shoes: \$590K

Strong footwear category
with growth potential

UK leads geographic revenue at \$691,548, followed by USA (\$657,476) and Germany (\$655,034). India shows highest discount dependency with 155 transactions.

Power BI Dashboard & DAX Measures



Key Performance Indicators

- Total Revenue: \$20.02M across all transactions
- Total Customers: 20,000 unique customer records
- Average Transaction: \$1,001.13 per purchase
- Discount Usage Rate: 26.7% of all transactions

Custom DAX Calculations

Eight custom measures created including Total Revenue, Average Transaction Value, Discount Revenue Impact, Revenue Growth %, Customer Retention Rate, and VIP Customer Revenue for dynamic analysis.

Strategic Business Recommendations



Target High-Value Segments

Focus marketing on 21-40 age group (\$1.49M revenue) and develop exclusive VIP programs since VIP customers spend 6.6% more per transaction



Optimize Product Mix

Increase inventory for Shirts (\$1.15M revenue) and bundle complementary products like watches and headsets to boost transaction value



Refine Discount Strategy

Implement tiered discounts to encourage higher spending and limit deep discounts to slow-moving inventory since current discounts reduce average spend



Geographic Expansion

Maintain premium positioning in UK (\$691K revenue), expand range in USA, and develop price-sensitive strategies for India and France markets

Project Impact & Future Enhancements

Key Achievements

- Successfully cleaned 22 data attributes with 100% integrity
- Identified \$1.15M revenue opportunity in Shirts category
- Discovered 6.6% spending premium among VIP customers
- Created actionable insights across customer segments and markets

Technical Proficiency Demonstrated

Advanced Python data manipulation, complex SQL querying with subqueries and aggregations, Power BI dashboard design with custom DAX measures, and comprehensive statistical analysis.

Next Steps

- Implement predictive modeling for customer lifetime value
- Develop real-time data streaming for live updates
- Build recommendation engine using purchase history
- Conduct sentiment analysis on customer feedback

This end-to-end analysis transformed raw data into strategic intelligence, providing evidence-based recommendations for sustainable business growth and competitive advantage in retail markets.

Business Recommendation

This recommendation is the result of extensive research and analysis into emerging market trends and opportunities.

- The key highlights of this recommendation are...
- First, it will allow us to diversify our product portfolio and tap into a rapidly growing consumer segment.
- Second, the financial projections show strong potential for increased revenue and profitability over the next 3-5 years.
- Third, the implementation plan outlines a phased approach that minimizes risk and disruption to our core business.
- In summary, I believe this recommendation represents a compelling strategic opportunity that is well-aligned with our company's vision and values.