Customer Shopping Behavior Analysis

Analyzing 3,900 purchases to uncover actionable insights through Python, SQL, and Power BI

Project Scope

01

Data Preparation

Python for cleaning and transformation

02

Structured Analysis

SQL for business questions

03

Visualization

Power BI for interactive insights

Dataset Overview

- 3,900 transactions analyzed
- 18 data columns
- Demographics, purchases, behavior

Key Data Points

- Customer demographics
- Purchase patterns
- Shopping preferences

Python Data Preparation

Data Loading & Exploration

Imported with Pandas, checked structure with df.info() and df.describe()

Missing Data Handling

Category-level median imputation for review ratings
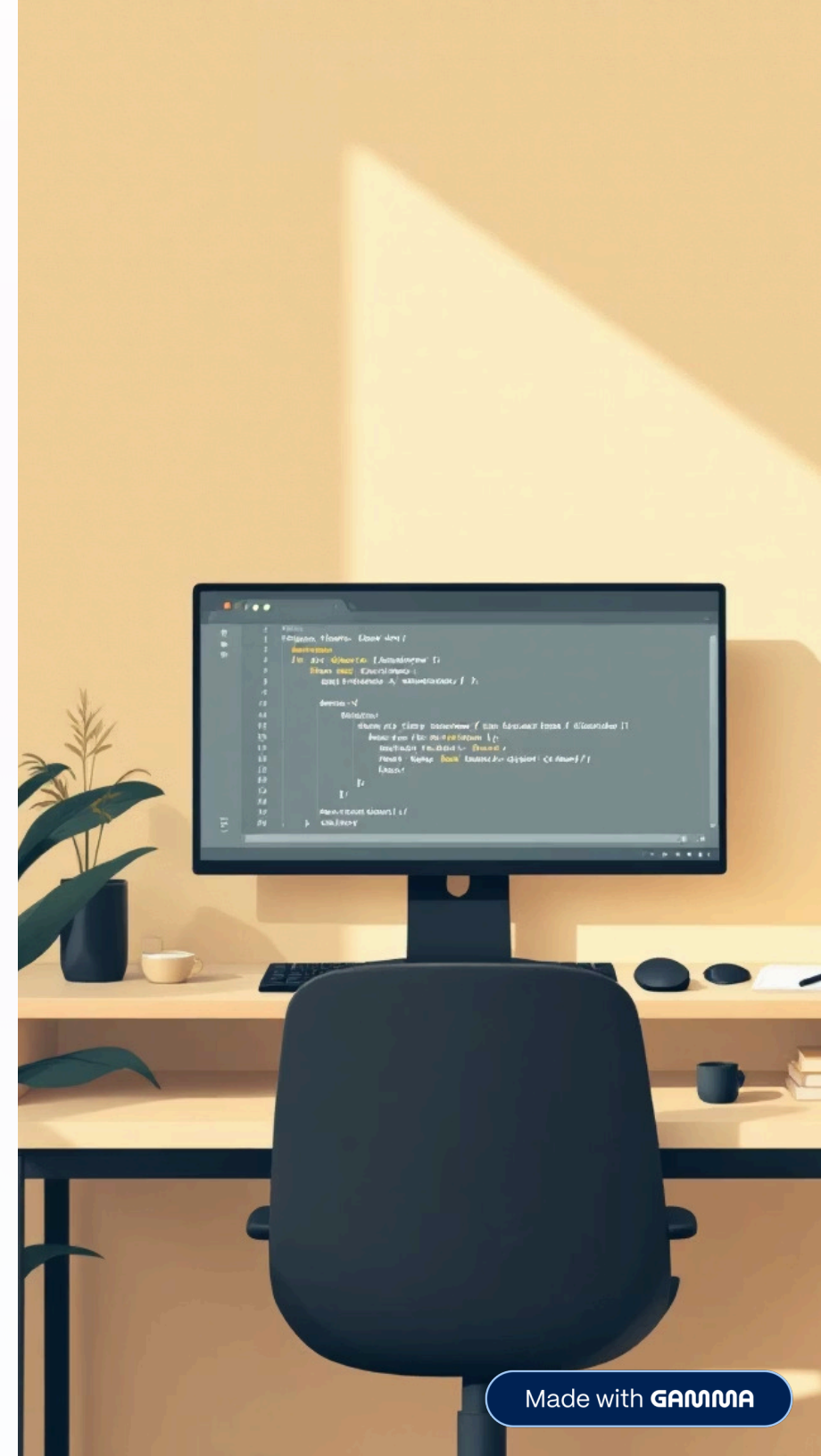
Standardization

Snake_case formatting for cross-tool compatibility

Feature Engineering

Age group segmentation and purchase frequency indicators created for deeper analysis

Database Integration

Connected to PostgreSQL for SQL analysis and Power BI linking

Revenue Insights

₹157K

Male Customers

Higher revenue contribution

₹75K

Female Customers

Secondary revenue segment

₹60.48

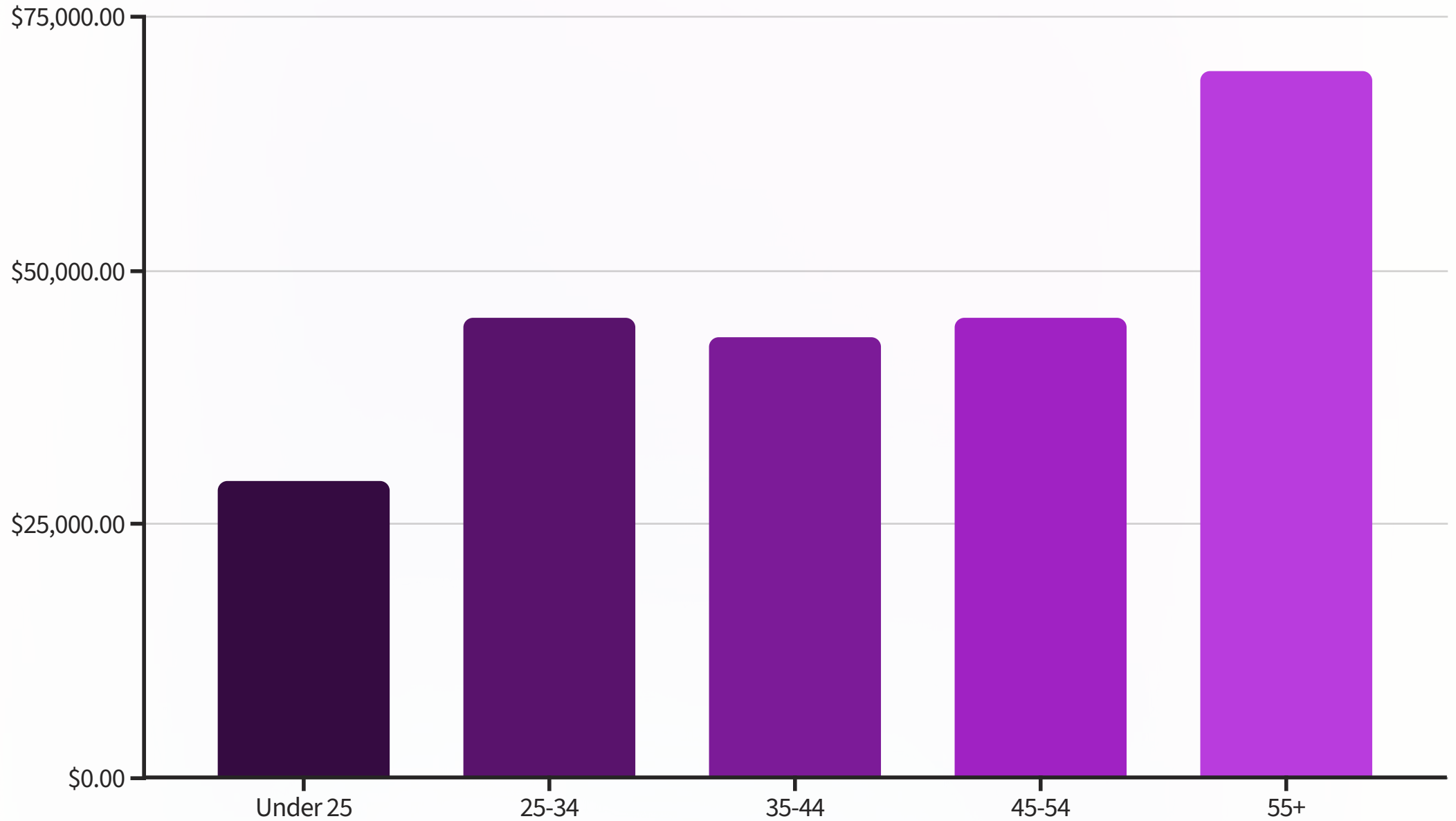**Express Shipping**

Average purchase amount

₹58.86

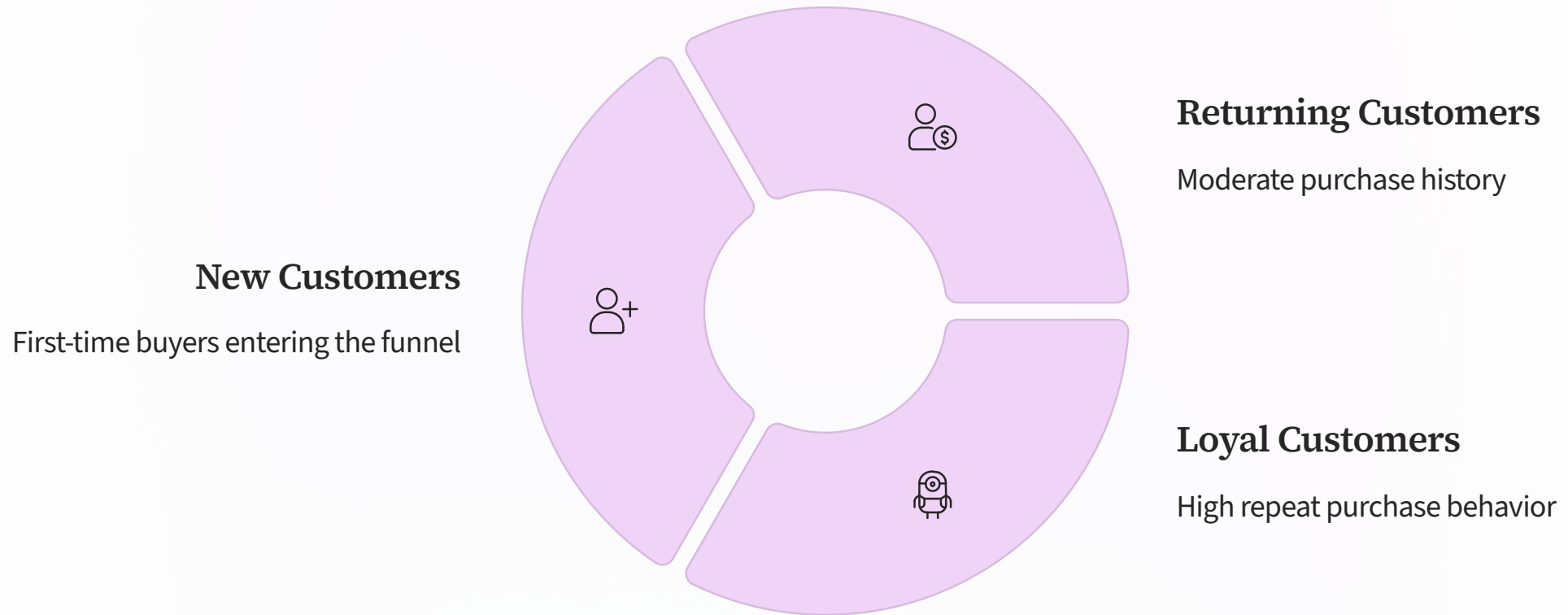**Standard Shipping**

Average purchase amount

Age Group Revenue Analysis

Older demographics (55+) contribute highest revenue at ₹69,590, representing the most valuable customer segment

Customer Segmentation

Returning Customers

Moderate purchase history

New Customers

First-time buyers entering the funnel

Loyal Customers

High repeat purchase behavior

Repeat buyers with 5+ purchases show significantly higher subscription rates: 958 subscribers vs 2,518 non-subscribers

Top Performing Products

By Category

- **Clothing:** Blouse, Pants, Shirt
- **Footwear:** Sandals, Shoes, Sneakers
- **Outerwear:** Jacket, Coat
- **Accessories:** Sunglasses, Jewelry

Highest Rated Items

Top 5 products identified by average review ratings drive customer trust and conversions

Products with highest discount dependency require strategic pricing review

Subscription Impact

Subscribers

Higher average spend and total revenue contribution

Non-Subscribers

Lower lifetime value but larger customer base

High-spending discount users identified: customers using promotions while maintaining above-average purchase amounts represent valuable conversion targets
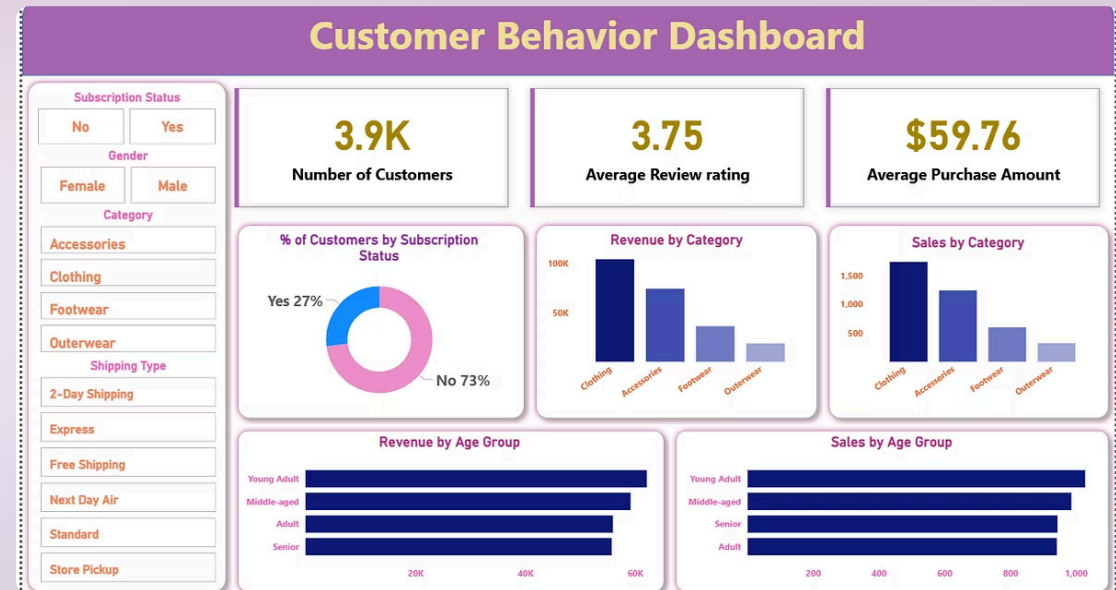
Power BI Dashboard

Key Metrics

Total customers, average rating, purchase value

Revenue Analysis

Sales distribution by category and subscription status

Interactive Filters

Slice by gender, location, shipping, age group

Customer Behavior Dashboard

Strategic Recommendations

Boost Subscriptions
Priority shipping, exclusive discounts, early product access

Loyalty Programs
Points, cashback, milestone rewards for repeat buyers

Strategic Discounting
Target promotions on responsive products, protect margins

Targeted Marketing
Focus on high-revenue age groups and express shipping users

Made with GAMMA