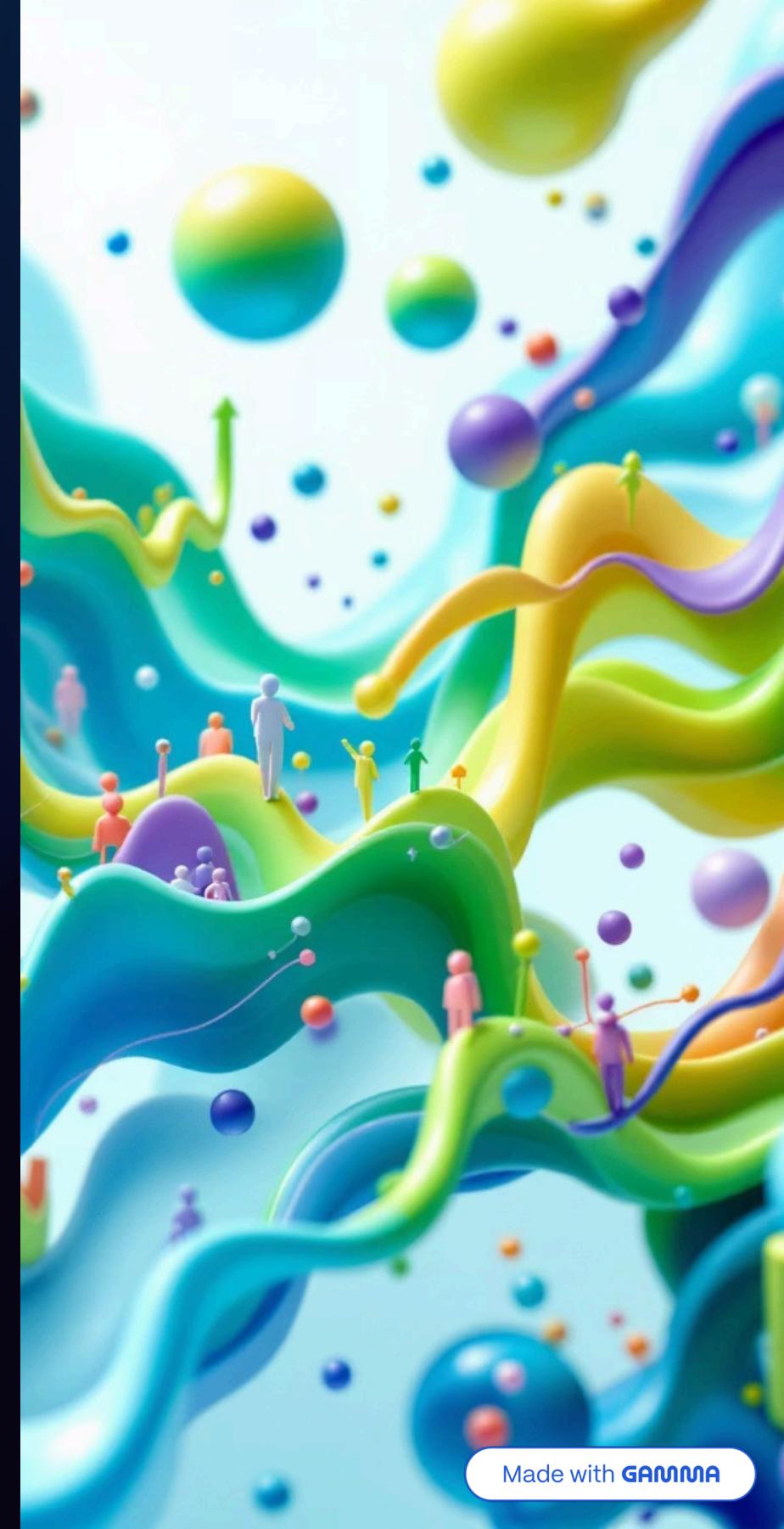
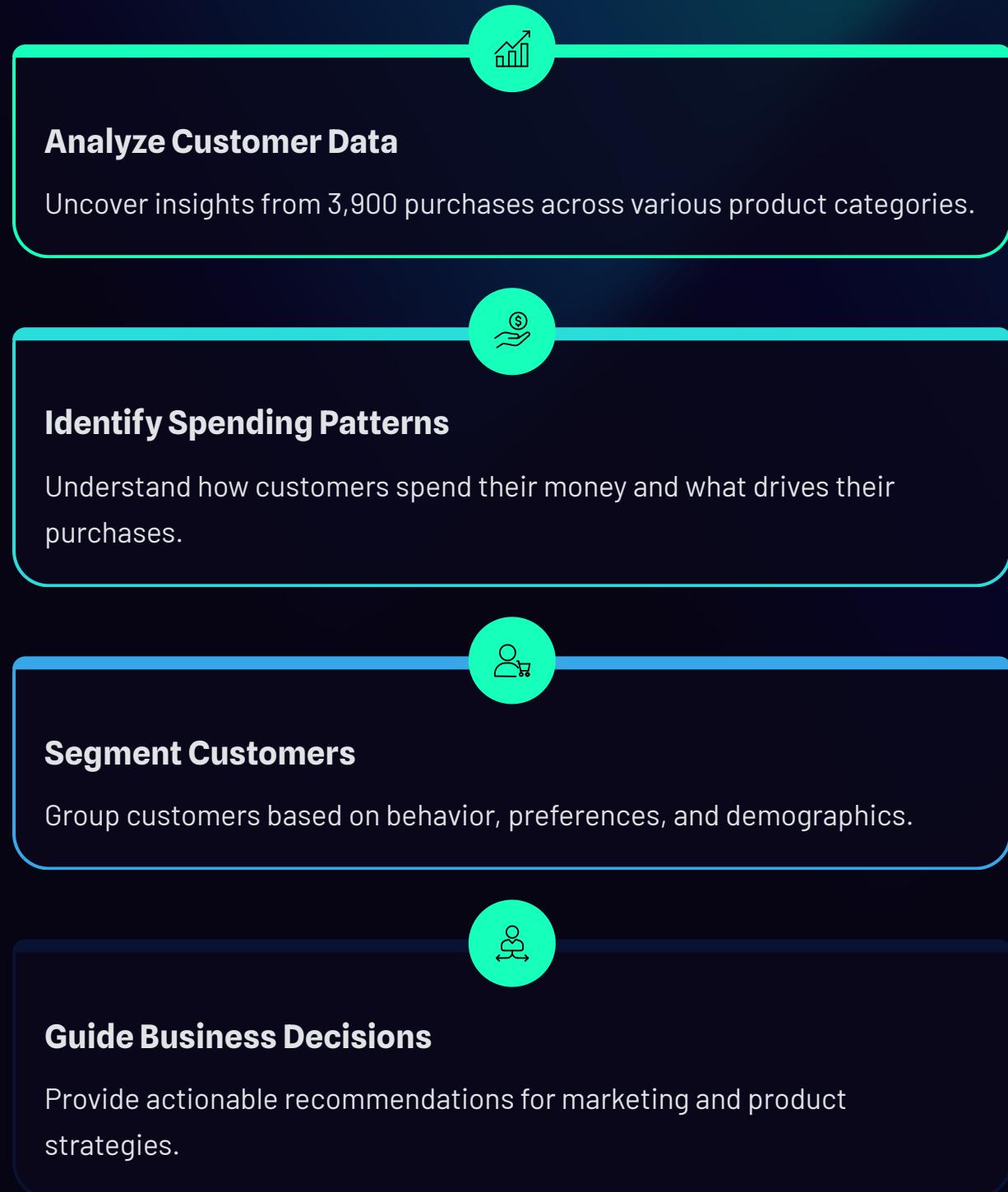




Customer Shopping Behavior Analysis

This project analyzes customer shopping behavior using transactional data from 3,900 purchases. Our goal is to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to guide strategic business decisions.

Project Overview



Dataset Summary

Key Features

- Customer demographics (Age, Gender, Location, Subscription Status)
- Purchase details (Item, Category, Amount, Season, Size, Color)
- Shopping behavior (Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type)

Dataset Snapshot

Rows: 3,900

Columns: 18

Missing Data: 37 values in Review Rating column



Exploratory Data Analysis (EDA) in Python

01

Data Loading & Exploration

Imported dataset using `pandas` and used `df.info()` and `.describe()` for initial structure and summary statistics.

02

Missing Data Handling

Imputed missing values in the **Review Rating** column using the median rating of each product category.

03

Column Standardization

Renamed columns to **snake_case** for improved readability and documentation.

04

Feature Engineering

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

05

Database Integration

Loaded the cleaned DataFrame into PostgreSQL for SQL analysis.

Data Analysis using SQL: Key Business Insights

Structured analysis in PostgreSQL answered key business questions, revealing critical spending patterns and customer behaviors.



Revenue by Gender

Male customers generated \$157,890, Female customers \$75,191.



Top 5 Products by Rating

Gloves, Sandals, Boots, Hat, and Skirt received the highest average review ratings.



Shipping Type Comparison

Express shipping had a slightly higher average purchase amount (\$60.48) than Standard (\$58.46).



Customer Segmentation & Discount Impact

Customer Segments

- **Loyal:** 3,116 customers
- **Returning:** 701 customers
- **New:** 83 customers

Subscribers vs. Non-Subscribers

Yes	1053	59.49	62645.00
No	2847	59.87	170436.00

Discount-Dependent Products

Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), and Pants (47.37%) had the highest percentage of discounted purchases.



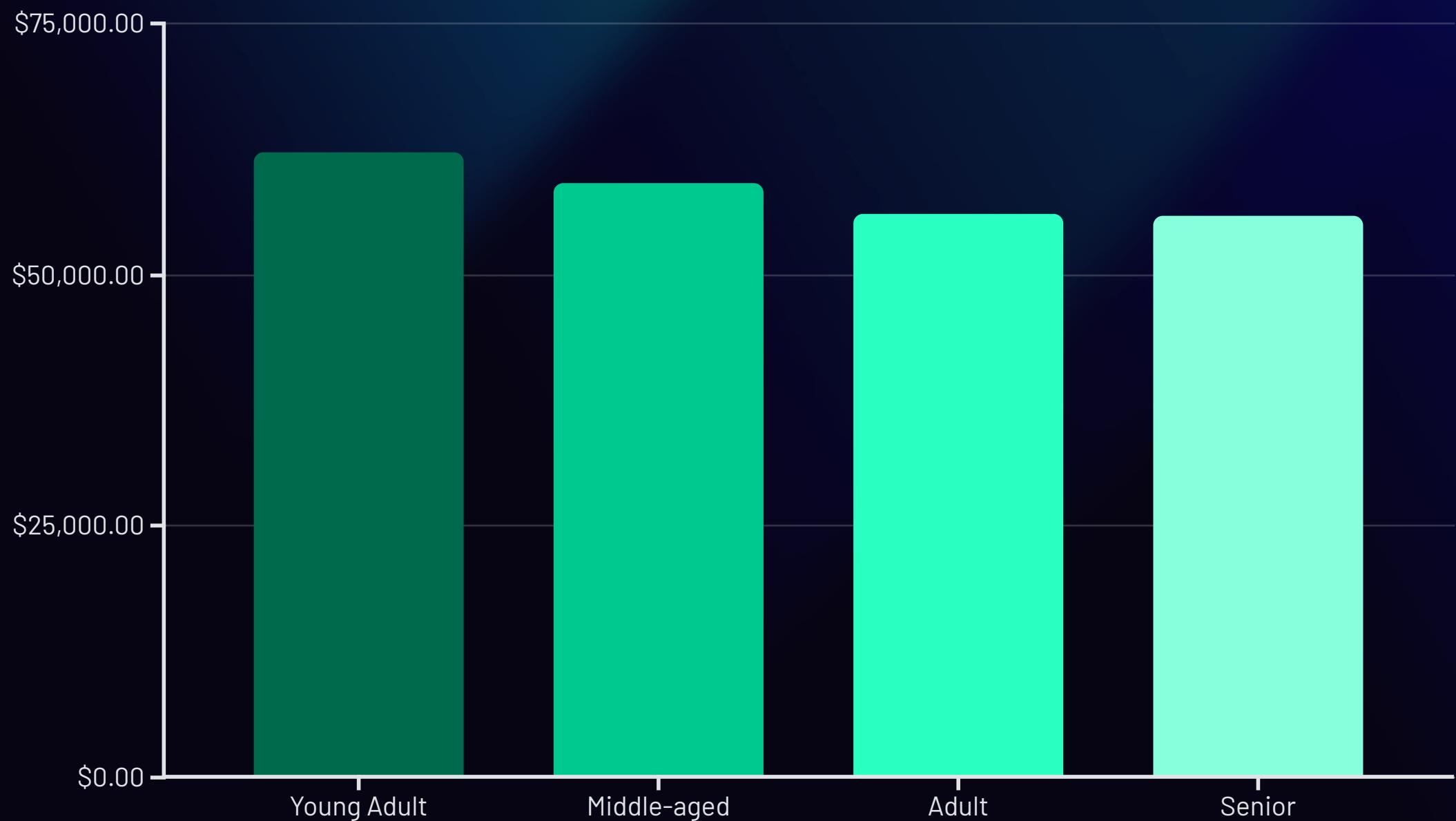
Top Products by Category

Understanding product popularity within each category helps optimize inventory and marketing efforts.

1	Accessories	Jewelry	171
2	Accessories	Sunglasses	161
3	Accessories	Belt	161
1	Clothing	Blouse	171
2	Clothing	Pants	171
3	Clothing	Shirt	169
1	Footwear	Sandals	160
2	Footwear	Shoes	150
3	Footwear	Sneakers	145
1	Outerwear	Jacket	163
2	Outerwear	Coat	161

Revenue by Age Group

Different age groups contribute varying amounts to total revenue, informing targeted marketing strategies.



Young Adults generate the highest revenue, followed closely by Middle-aged customers. This highlights opportunities for age-specific campaigns.

Interactive Dashboard in Power BI

An interactive Power BI dashboard was created to visualize key insights, providing a dynamic tool for business users.



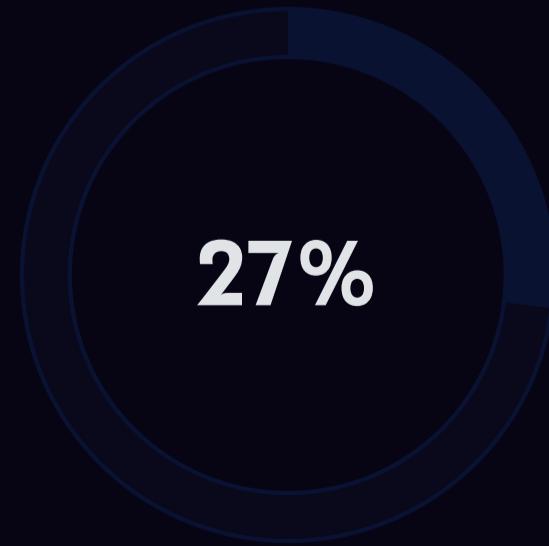
Customers



Avg. Purchase



Avg. Review



Subscribers



Business Recommendations



Boost Subscriptions

Promote exclusive benefits to increase subscriber base.



Customer Loyalty Programs

Reward repeat buyers to foster loyalty and move them into the "Loyal" segment.



Review Discount Policy

Balance sales boosts with margin control to optimize profitability.



Targeted Marketing

Focus efforts on high-revenue age groups and express-shipping users for maximum impact.