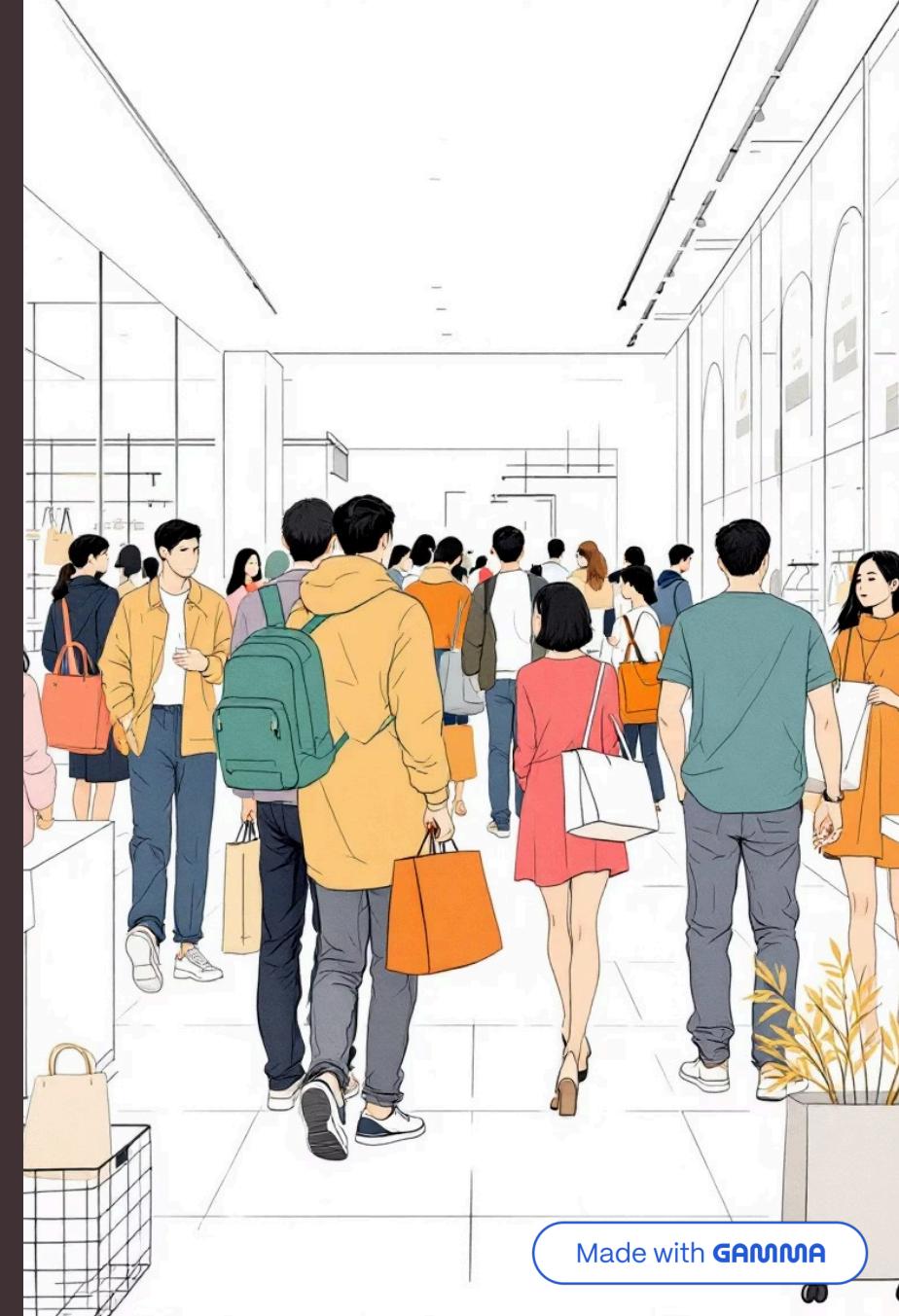


# Customer Shopping Behavior Analysis

Uncovering insights from transactional data to guide strategic business decisions.



A complex network graph composed of numerous small, semi-transparent nodes and a dense web of thin, dark gray lines. The nodes are scattered across the left side of the slide, with a higher concentration in the lower-left quadrant. Some nodes are highlighted with larger, solid black circles, particularly in the center-right area. The overall effect is one of data connectivity and complexity.

# Project Overview

## Goal

Analyze 3,900 purchases to understand spending patterns, customer segments, product preferences, and subscription behavior.

## Data Source

Transactional data across various product categories.

## Outcome

Provide actionable insights for strategic business decisions.

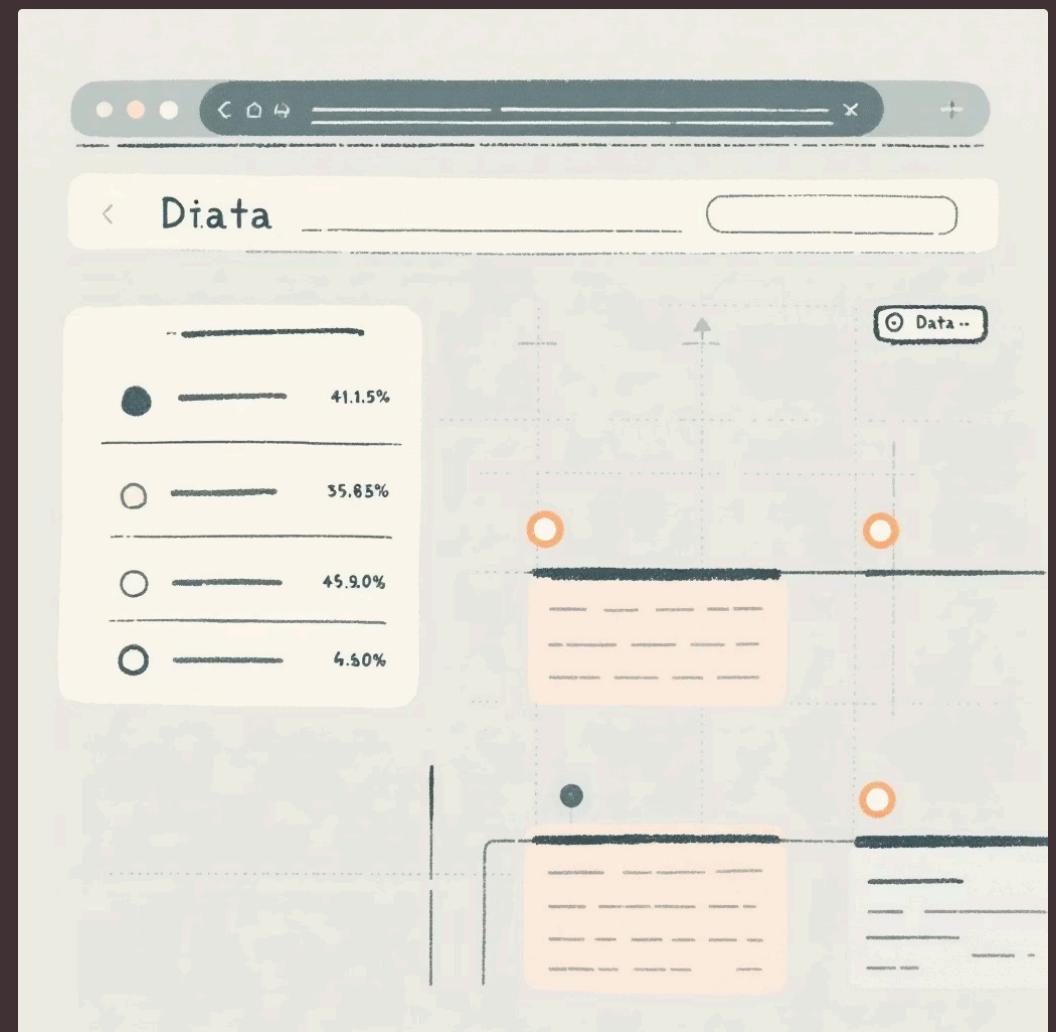
# Dataset Summary

## 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)

## Data Snapshot

- 3,900 Rows
- 18 Columns
- 37 missing values in "Review Rating"



# Exploratory Data Analysis (Python)

01

---

## Data Preparation

Loading, initial exploration, and handling missing values (imputed median rating).

02

---

## Column Standardization

Renamed columns to snake\_case for readability.

03

---

## Feature Engineering

Created 'age\_group' and 'purchase\_frequency\_days'.

04

---

## Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.



# Data Analysis using SQL

1

## Revenue by Gender

Female: \$75,191, Male: \$157,890

2

## High-Spending Discount Users

Identified customers using discounts but spending above average.

3

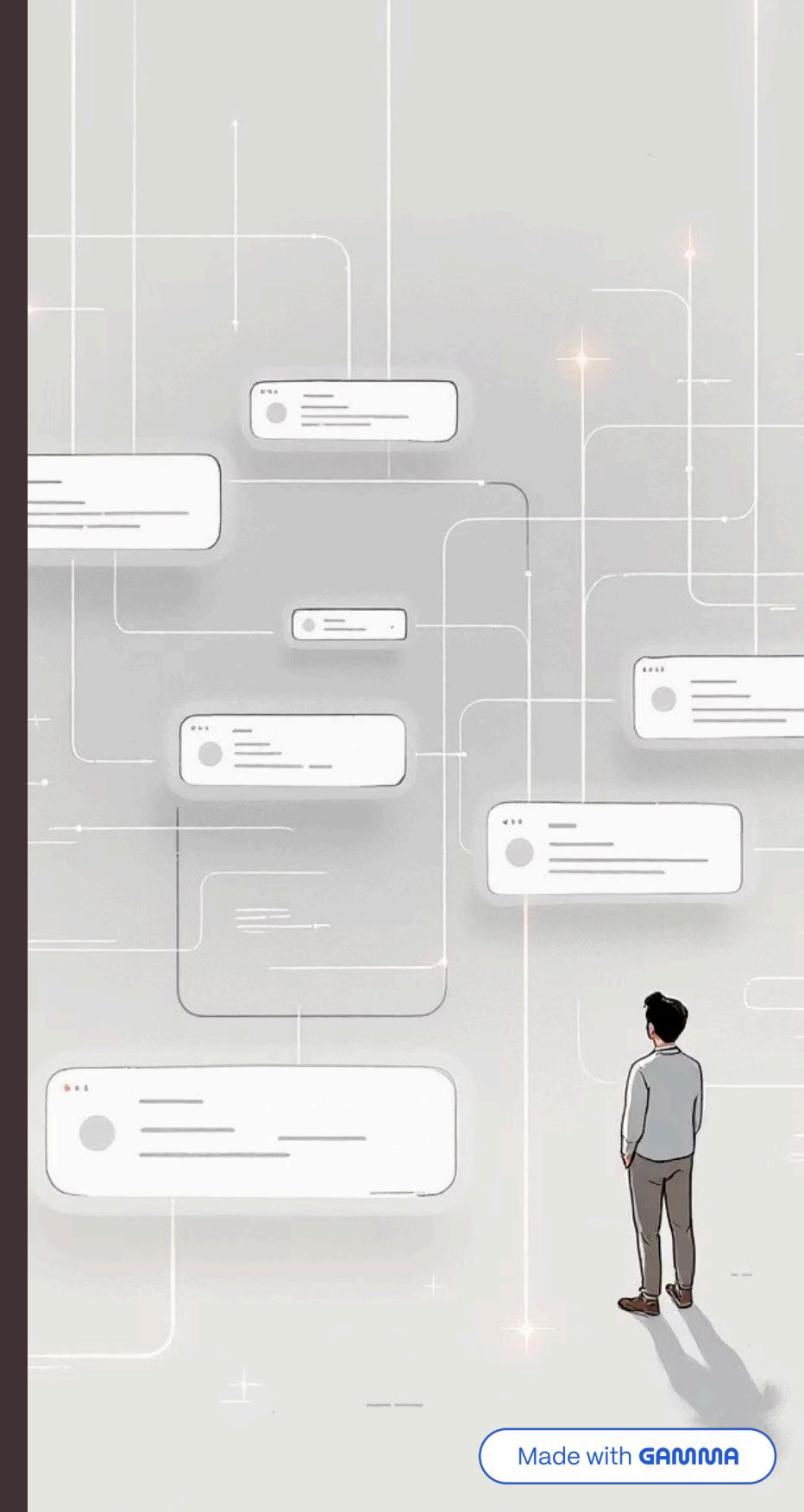
## Top 5 Products by Rating

Gloves (3.86), Sandals (3.84), Boots (3.82), Hat (3.80), Skirt (3.78)

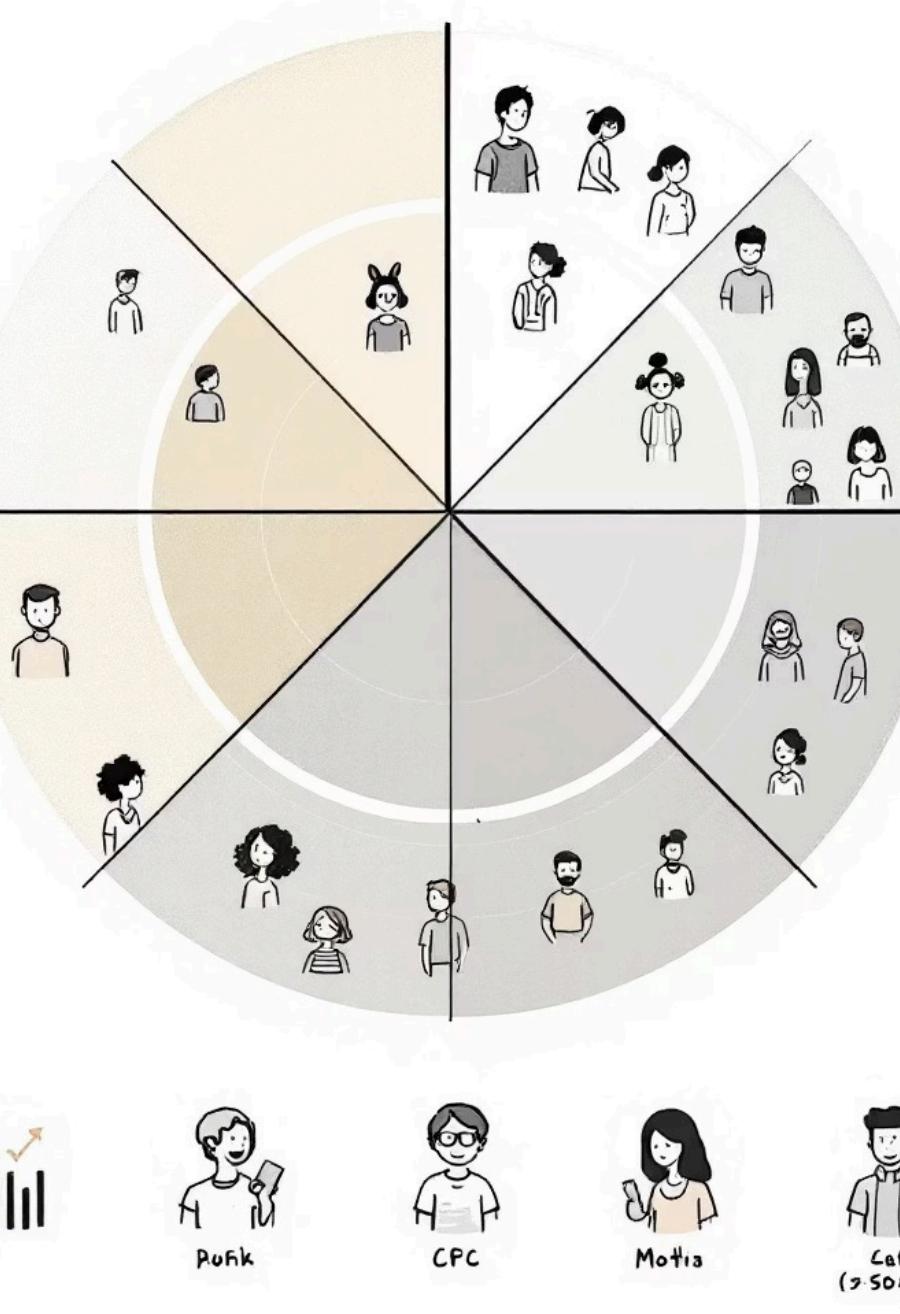
4

## Shipping Type Comparison

Standard: \$58.46, Express: \$60.48



# Customer Segmentation



# SQL Insights: Customer Behavior

1

## Subscribers vs. Non-Subscribers

Subscribers: 1,053 customers,  
Avg Spend: \$59.49, Total  
Revenue: \$62,645

Non-Subscribers: 2,847  
customers, Avg Spend:  
\$59.87, Total Revenue:  
\$170,436

2

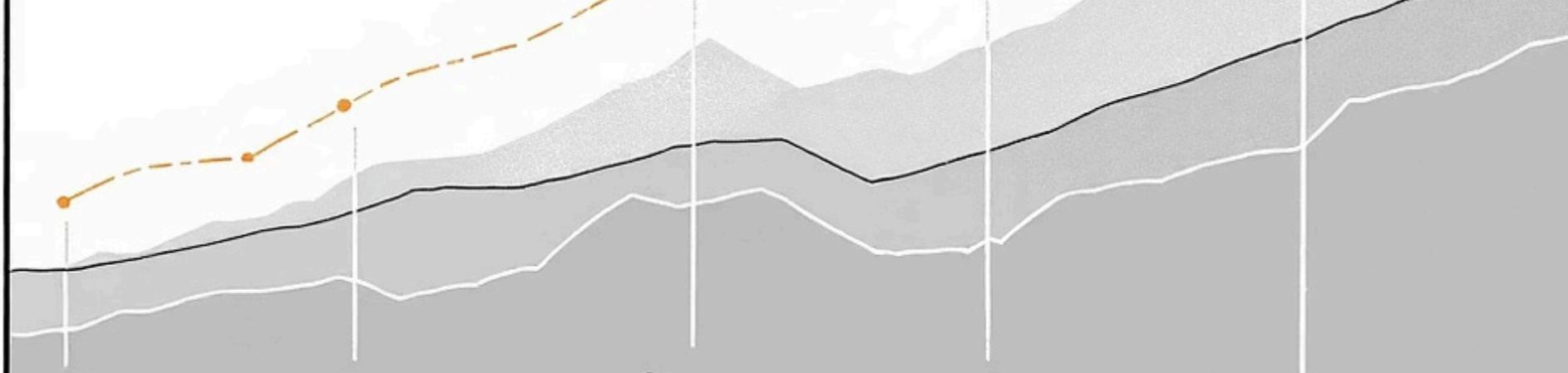
## Discount-Dependent Products

Hat (50%), Sneakers (49.66%),  
Coat (49.07%), Sweater  
(48.17%), Pants (47.37%)

3

## Customer Segmentation

Loyal (3,116), Returning (701), New (83)



# SQL Insights: Product & Age

1

## Top 3 Products per Category

Accessories: Jewelry, Sunglasses, Belt

Clothing: Blouse, Pants, Shirt

Footwear: Sandals, Shoes, Sneakers

Outerwear: Jacket, Coat

2

## Repeat Buyers & Subscriptions

Repeat buyers (>5 purchases): 958 subscribers vs. 2,518 non-subscribers.

3

## Revenue by Age Group

Young Adult: \$62,143

Middle-aged: \$59,197

Adult: \$55,978

Senior: \$55,763

# Power BI Dashboard

An interactive dashboard was built to present these insights visually.





# Business Recommendations

## → Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

## → Customer Loyalty Programs

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

## → Review Discount Policy

Optimize discount strategies to balance sales with profit margins.

# Strategic Actions

## → Product Positioning

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

## → Targeted Marketing

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

