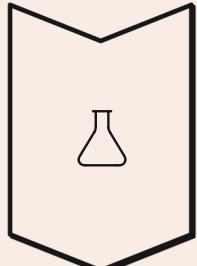




# Customer Sales Behaviour Analysis

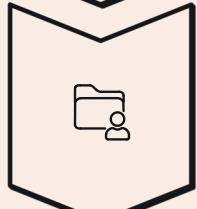
A comprehensive data analysis project exploring customer purchasing patterns through Python, SQL, and Power BI visualization.

# Project Workflow



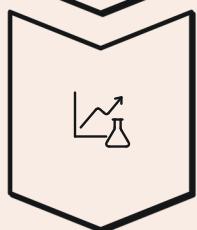
## Python Analysis

Data loading, EDA, feature engineering, and PostgreSQL export



## SQL Queries

Advanced queries to extract meaningful business insights



## Power BI

Interactive dashboard creation for visual storytelling



# Dataset Overview

**3,900**

**Total Records**

Customer transactions analyzed

**18**

**Attributes**

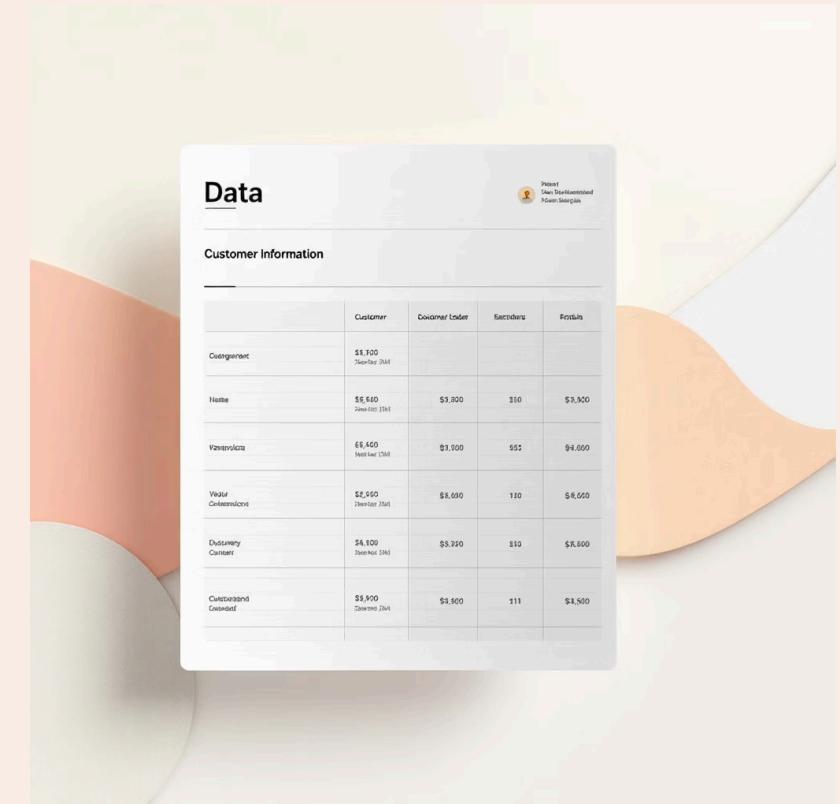
Data columns captured

**37**

**Missing Values**

In Review Rating column only

Dataset includes customer demographics, purchase details, shipping preferences, and behavioral patterns across multiple product categories.



Data				
	Customer	Customer Leader	Seconds	Product
Customer	\$8,100 200 rows 30M			
Name	\$6,640 200 rows 25M	\$3,300	110	\$3,300
Venue/Category	\$6,400 200 rows 25M	\$3,300	111	\$3,300
Visitor Demographics	\$2,450 200 rows 25M	\$8,050	110	\$8,050
Delivery Content	\$4,100 200 rows 30M	\$9,320	110	\$9,320
Categorize Content	\$5,900 200 rows 25M	\$9,500	111	\$9,500

# Data Preparation & Cleaning

01

## Data Loading

Imported customer shopping behavior CSV using pandas

02

## Exploratory Analysis

Examined data structure, identified 37 null values in  
Review Rating

03

## Null Value Treatment

Filled missing ratings with category-wise median values

04

## Column Standardization

Converted to lowercase, replaced spaces with underscores  
for SQL compatibility

# Feature Engineering

## New Attributes Created

### Age Group Segmentation

- Young Adult
- Adult
- Middle Aged
- Senior

Quartile-based customer age categorization

### Purchase Frequency Days

Mapped frequency labels to numeric day intervals  
(weekly=7, monthly=30, quarterly=90, annually=365)



# Key SQL Insights

## Revenue by Gender

Male customers: \$157,890

Female customers: \$75,191

Males generated 2.1x more revenue

## Discount Behavior

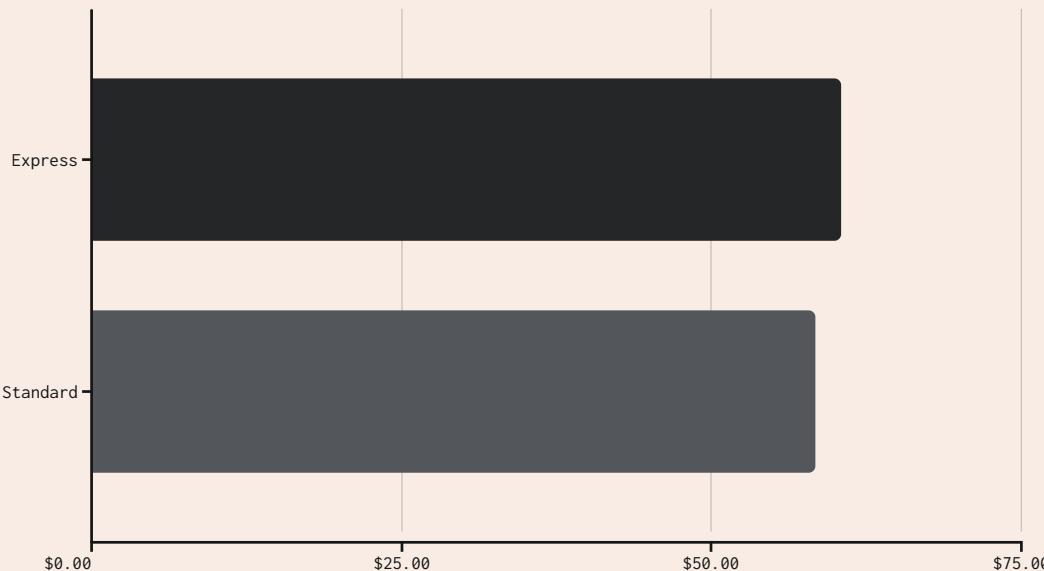
839 customers used discounts but still spent above average purchase amount

## Top Rated Products

1. Gloves (3.86)
2. Sandals (3.84)
3. Boots (3.82)
4. Hat (3.80)
5. Skirt (3.78)

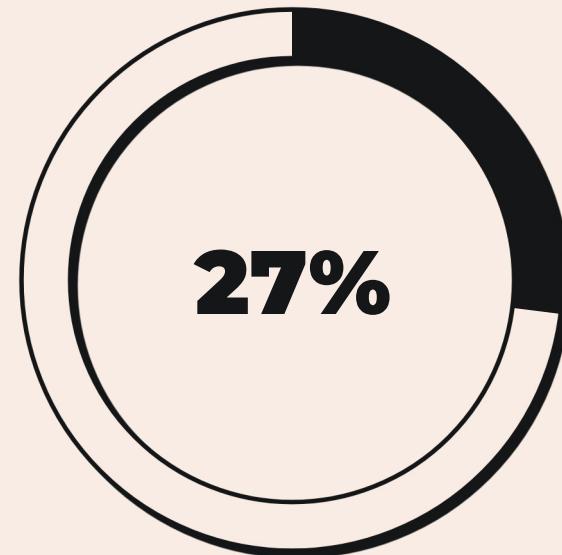
# Shipping & Subscription Analysis

## Shipping Type Comparison



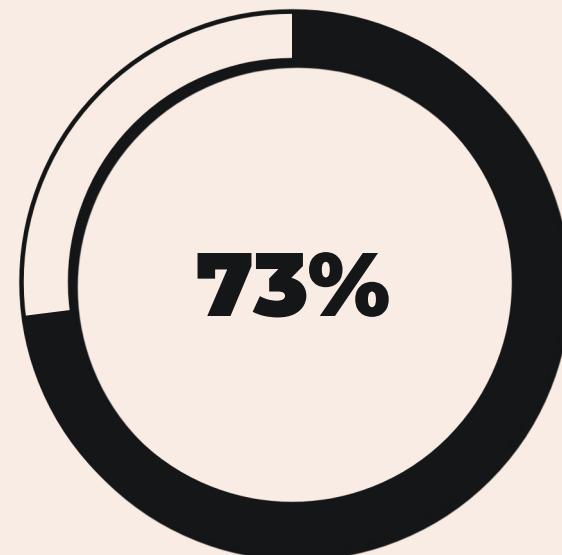
Express shipping customers spend slightly more on average

## Subscription Impact



## Subscribers

1,053 customers with \$62,645 revenue

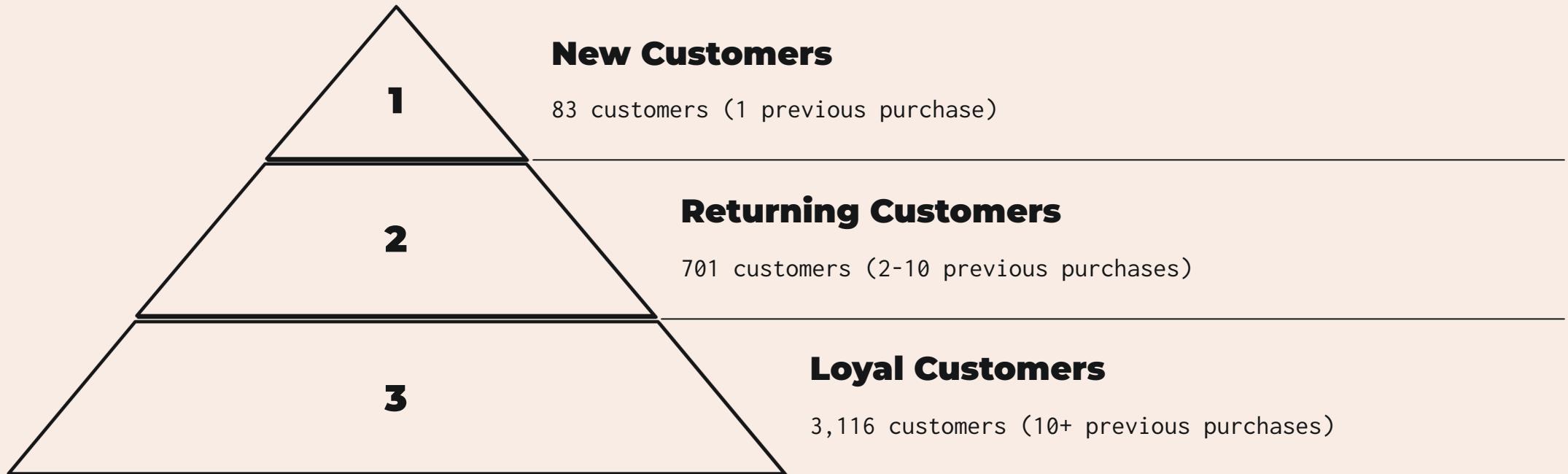


## Non-Subscribers

2,847 customers with \$170,436 revenue

Non-subscribers generate higher total revenue despite similar average spend (\$59.87 vs \$59.49)

# Customer Segmentation



The majority (80%) are loyal customers, indicating strong retention and repeat purchase behavior.

# Power BI Dashboard Highlights

## Interactive Visualizations

Revenue by subscription status, category breakdown, and gender analysis

## Key Metrics

3.9K customers, 3.75 avg rating,  
\$3.08K total revenue tracked

## Dynamic Filtering

Filter by gender, shipping type, and review ratings for deeper insights



# Key Takeaways



## Complete Data Pipeline

Successfully transformed raw data into actionable insights through Python cleaning, SQL analysis, and Power BI visualization



## Customer Behavior Patterns

Identified revenue differences by gender, discount usage trends, and strong customer loyalty (80% repeat buyers)



## Data-Driven Decisions

Interactive dashboard enables businesses to understand purchasing patterns and optimize marketing strategies