



Customer Shopping Behavior Analysis

Uncovering insights from 3,900 purchases to drive strategic business decisions

Dataset at a Glance

3,900

Total Purchases

Transactions analyzed

18

Data Points

Features per customer

4

Categories

Product segments

50

Locations

Geographic spread

Key Features Tracked

- Customer demographics & subscription status
- Purchase details & seasonal patterns
- Shopping behavior & review ratings

Data Quality

37 missing values in Review Rating column,
imputed using median by category

Python Data Preparation

01

Data Loading & Exploration

Imported dataset, checked structure with `df.info()` and summary statistics

02

Cleaning & Standardization

Handled missing values, renamed columns to snake case

03

Feature Engineering

Created `age_group` and `purchase_frequency_days` columns

04

Database Integration

Connected to MySQL and loaded cleaned data for SQL analysis

Revenue Insights

Gender Revenue Comparison

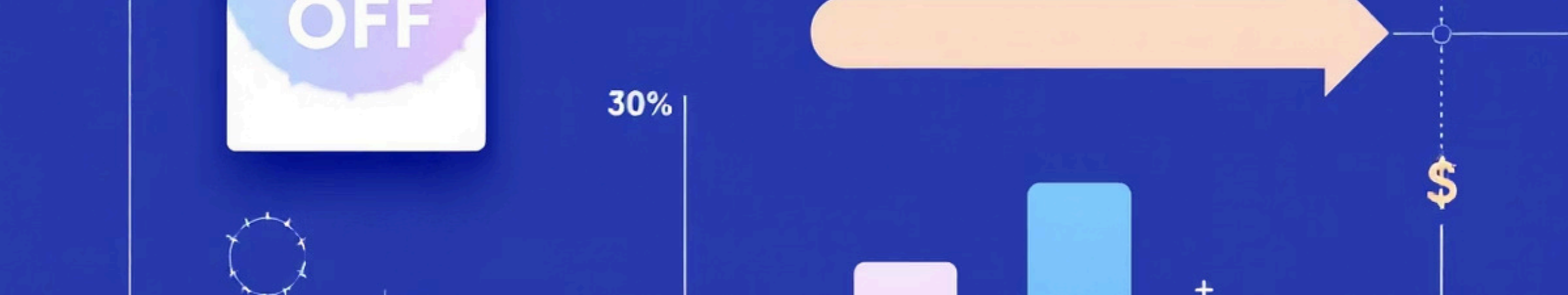
Male customers generate higher total revenue across all categories

Result Grid			Filter Rows:
	gender	total_revenue	
▶	Male	157890	
	Female	75191	

Age Group Contribution

Revenue distribution reveals key demographic segments driving sales

Result Grid			Filter Rows:
	age_group	total_revenue	
▶	Senior	55763	
	Adult	55978	
	Middle-aged	59197	
	Young Adult	62143	




Discount & Spending Patterns

High-Spending Discount Users


Customers using discounts still spending above average purchase amount

Result Grid



Filter Rows:

Export:



	customer_id	gender	discount_applied	purchase_amount
▶	2	Male	Yes	64
	3	Male	Yes	73
	4	Male	Yes	90
	7	Male	Yes	85
	9	Male	Yes	97
	12	Male	Yes	68
	13	Male	Yes	72
	16	Male	Yes	81
	20	Male	Yes	90

Discount-Dependent Products

Top 5 products with highest percentage of discounted purchases identified

Result Grid	Filter Rows:
item_purchased	discount_percentage
Hat	50.00
Sneakers	49.66
Coat	49.07
Sweater	48.17
Pants	47.37



Product Performance

Top-Rated Products

5 products with highest average review ratings drive customer satisfaction

Result Grid			Filter Rows:	
	item_purchased	avg_product_rating		
▶	Gloves	3.86		
	Sandals	3.84		
	Boots	3.82		
	Hat	3.8		
	Handbag	3.78		

Category Leaders

Top 3 most purchased products identified within each category

Result Grid			Filter Rows:		Export:		W
	item_rank	category	item_purchased	total_orders			
	1	Accessories	Jewelry	171			
	1	Clothing	Blouse	171			
	1	Footwear	Sandals	160			
	1	Outerwear	Jacket	163			
▶	2	Accessories	Sunglasses	161			
	2	Clothing	Pants	171			
	2	Footwear	Shoes	150			
	2	Outerwear	Coat	161			
	3	Accessories	Belt	161			
	3	Clothing	Shirt	169			
	3	Footwear	Sneakers	145			



Shipping & Delivery Analysis

Standard vs Express

Average purchase amounts compared across shipping types reveal customer preferences

Result Grid			Filter Rows:
	shipping_type	avg_purchase_amt	
▶	Express	60.48	
	Standard	58.46	

Customer Segmentation



Customers classified by purchase history: New, Returning, and Loyal segments

Result Grid			Filter Rows:
	customer_segment	total_counts	
▶	Loyal	3116	
	Returning	701	
	New	83	

Subscription Impact

Subscribers vs non-subscribers: average spend and total revenue comparison

Result Grid					Filter Rows:	Export:	Wrap
	subscription_status	total_customers	avg_spend	total_revenue			
▶	Yes	1053	59.49	62645			
	No	2847	59.87	170436			

Repeat Buyer Behavior

Customers with >5 purchases show higher subscription likelihood

Result Grid			Filter Rows:
	subscription_status	repeat_buyers	
▶	Yes	958	
	No	2518	

Power BI Dashboard

Interactive visualization bringing all insights together for data-driven decisions

Key KPIs

Revenue, customer growth, category performance

Visual Storytelling

Clear, engaging insights from raw data

Business Intelligence

Empowering smarter strategic decisions



Strategic Recommendations



Boost Subscriptions

Promote exclusive benefits for subscribers



Loyalty Programs

Reward repeat buyers to build loyal segment



Review Discount Policy

Balance sales boosts with margin control



Product Positioning

Highlight top-rated items in campaigns



Targeted Marketing

Focus on high-revenue age groups and express shipping users