



Supermarket Customer Analysis and Insights

Data Analysis for Enhanced Customer Understanding and Strategic Recommendations

Vina Fasya Kartamanah

DTIDS-0206

Nov 2024



Contents

- Overview
- Data Understanding and Cleaning
- Customer Demographics and Spending Overview
- Customer Segmentation & Preferences Analysis
- Purchase Channel Analysis
- Summary
- Recommendation

Goal:

Analyze how customer demographics influence shopping behavior and spending across product categories like wine, fruits, meat, fish, sweets, and gold.

Purpose:

Provide actionable insights to improve supermarket operations, customer engagement, and sales strategies.



Data Overview

The supermarket customers dataset provides a detailed look at customer demographics, how they spend on products, their responses to promotions, and where they shop. It includes 2,240 rows and 29 columns, organized into four main categories:

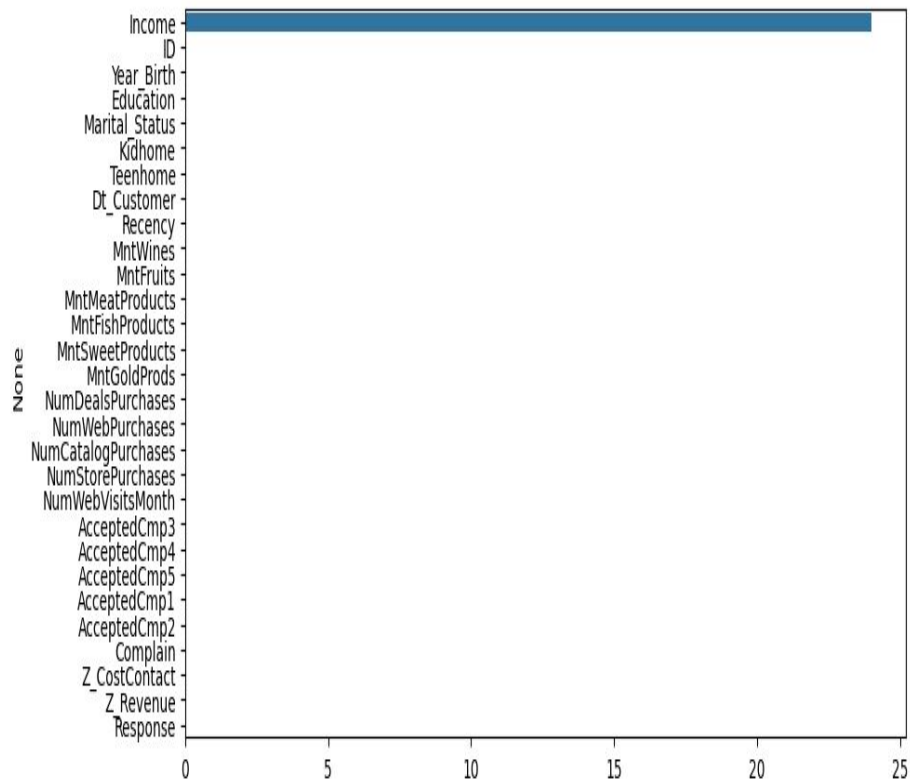
- Customer Demographics
- Product Spending
- Promotional Responses
- Purchase Channels



Data Understanding and Cleaning

- Outlier removal: Reduced skewness in spending and income.
- Handling missing values: Appropriate techniques to fill gaps (median) in 'Income' column (1.07% missing)
- Converting data types: from object to datetime for 'Dt_Customer' column
- Generating new features: Created age group, income group, and household composition columns.
- Merging semantically same value: Standardized redundant categorical values in marital and educational status

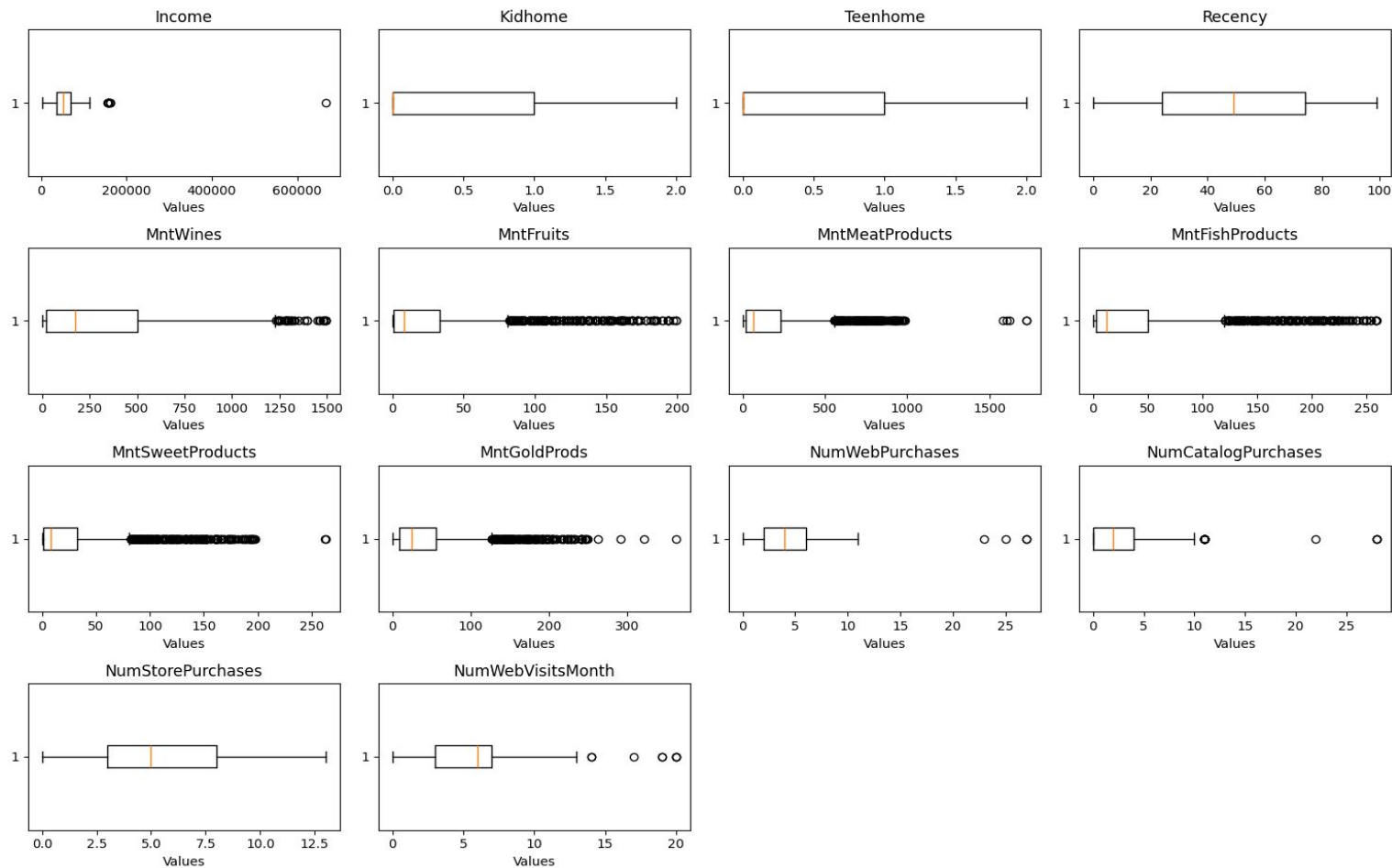
Data Cleaning and Understanding Process



#	Column	Non-Null Count		Dtype
0	ID	2240	non-null	int64
1	Year_Birth	2240	non-null	int64
2	Education	2240	non-null	object
3	Marital_Status	2240	non-null	object
4	Income	2216	non-null	float64
5	Kidhome	2240	non-null	int64
6	Teenhome	2240	non-null	int64
7	Dt_Customer	2240	non-null	object
8	Recency	2240	non-null	int64
9	MntWines	2240	non-null	int64
10	MntFruits	2240	non-null	int64
11	MntMeatProducts	2240	non-null	int64
12	MntFishProducts	2240	non-null	int64
13	MntSweetProducts	2240	non-null	int64
14	MntGoldProds	2240	non-null	int64
15	NumDealsPurchases	2240	non-null	int64
16	NumWebPurchases	2240	non-null	int64
17	NumCatalogPurchases	2240	non-null	int64
18	NumStorePurchases	2240	non-null	int64
...				
27	Z_Revenue	2240	non-null	int64
28	Response	2240	non-null	int64

dtypes: float64(1), int64(25), object(3)
memory usage: 507.6+ KB

Outliers in Numerical Data



Removing outliers from the key columns (Income, MntWines, MntGoldProds, and MntMeatProducts)

Filtered out about 8% of the data (180 rows out of 2240). this is to reduce the skew in those columns and provide a cleaner dataset



Customer Demographics and Spending Overview

Age Group Distribution

Majority of Customers: 31-45 and 46-60 age groups make up over 77% of the customer base.

Opportunity: Tailored promotions and offers targeting these groups can maximize engagement and sales.

Smaller Segments: 18-30 and 60+ groups offer potential for growth with targeted strategies.

Income Group Spending

High-Income Customers: Contribute most to total spending—ideal for premium offers and exclusive promotions.

Low-Income Customers: Largest group by percentage but lower spending—respond well to value-driven promotions.

Household Composition

With Children (75.49%): Dominant segment; target with family-oriented and bulk purchase promotions.

Without Children (24.51%): Higher average spending; suitable for premium and personalized offers.

Customer Demographics and Insights for Supermarket Spending Patterns

Acquisition Year: (All) Age Group: (All) Income Group: (All)

Total Customer

2,060

Total Spending \$

1,085,382

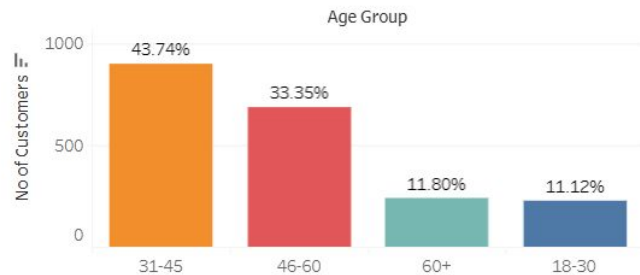
Average Spending

526.9

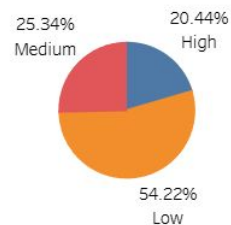
Customers by Income Group

Low	Medium	High
54.22%	25.34%	20.44%

Customer Demo - Age



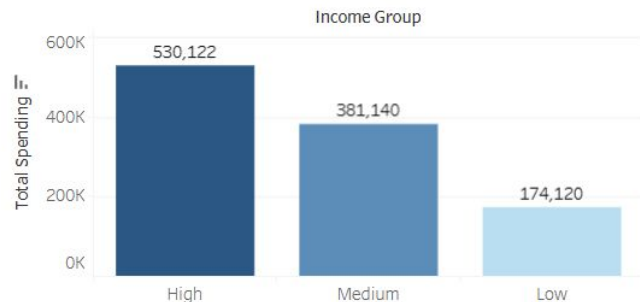
Customer Demo - Income



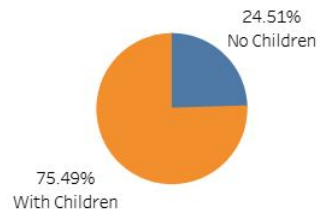
Objectives:

1. Understand how customer demographic characteristics influence shopping behavior
2. Identify the most profitable customer segments and determine effective marketing strategies for each segment
3. Optimize inventory management and merchandising based on customer preferences and shopping behavior.

Total Spending Overview by Income Group

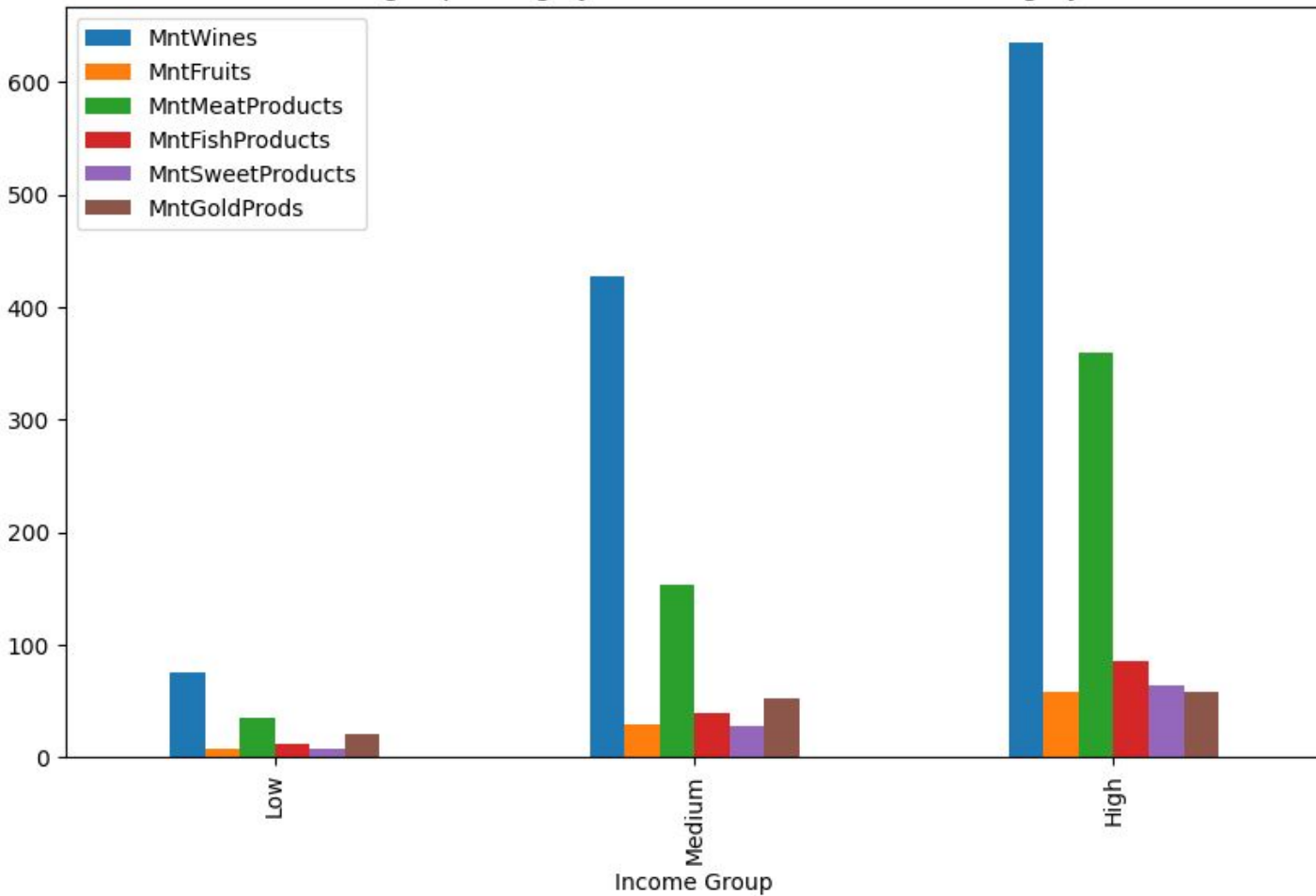


Customers With vs. Without Children

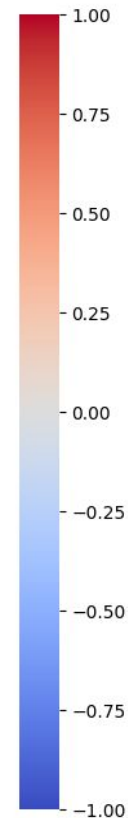


Income groups were created using min, median, and max values. High-income group contributes significantly to total spending.

Average Spending by Income Level and Product Category



High-income households lead in wine and meat spending, presenting an opportunity for premium offerings.

[illegible]

Income	Total_Spending	MntWines	MntFruits	MntMeatProducts	MntFishProducts	MntSweetProducts	MntGoldProds	NumWebPurchases	NumCatalogPurchases	NumStorePurchases	Recency
--------	----------------	----------	-----------	-----------------	-----------------	------------------	--------------	-----------------	---------------------	-------------------	---------



Customer Segmentation & Preferences

Product Spending Trends

Top Categories: Wine and meat are the highest spending categories, with wine leading significantly.

Opportunity: Focus on premium offers and loyalty programs around wine to capitalize on its popularity.

Spending by Income Group

High-Income Customers: Spend the most on both wine and meat, indicating their potential for premium and exclusive offers.

Medium and Low-Income Groups: Moderate to lower spending on wine and meat, suggesting value-driven promotions may increase engagement.

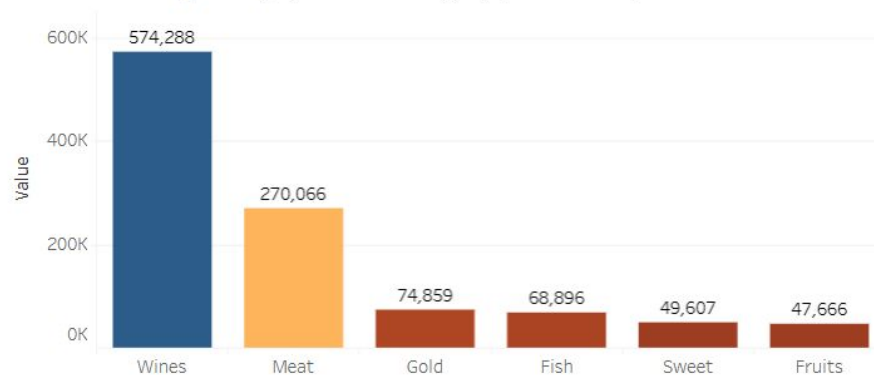
Campaign Acceptance

Successful Campaigns: Campaigns 3, 4, and 5 show strong customer acceptance rates.

Improvement Opportunity: Campaigns 1 and 2 have lower engagement and may benefit from refinement or budget reallocation to successful campaigns.

Customer Segmentation & Preferences - Spending and Campaign Engagement

Total Customer Spending by Product Category (Last 2 Years)



Year of Customer Acquisition Date

(All)

Household Composition

(All)

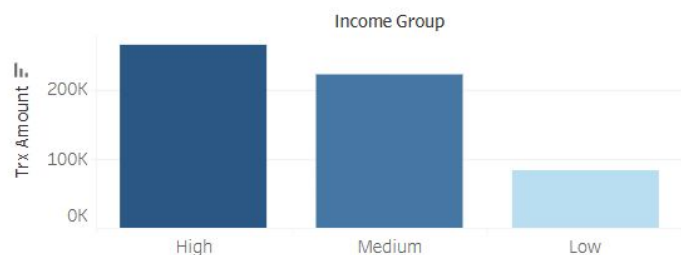
Age Group

(All)

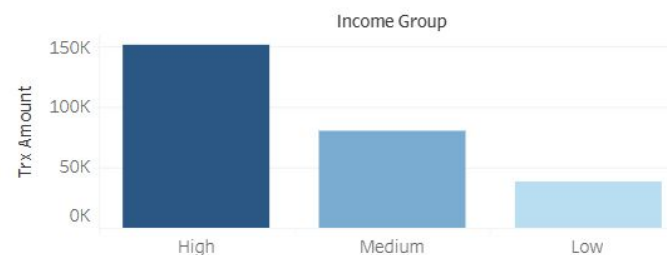
Income Group

(All)

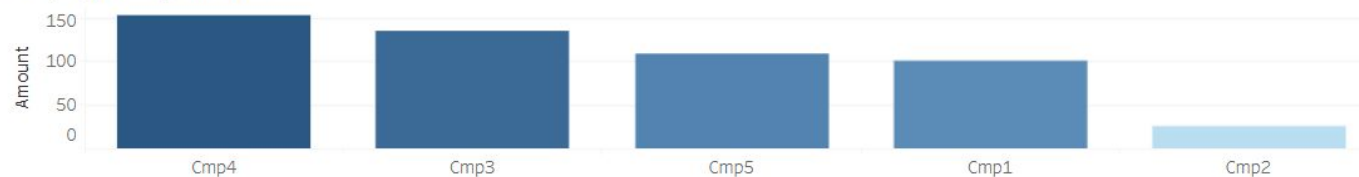
Wine Spending



Meat Spending



Campaign Acceptance Rate





Purchase Channel Analysis

Purchase Channel Preferences by Household Composition

Households with Children: High engagement across all channels (store, web, and catalog), particularly in-store.

Households Without Children: Lower overall engagement but consistent use of all channels.

Channel-Specific Behavior

Store Purchases: Dominant for all income groups, indicating the continued importance of in-person shopping experiences.

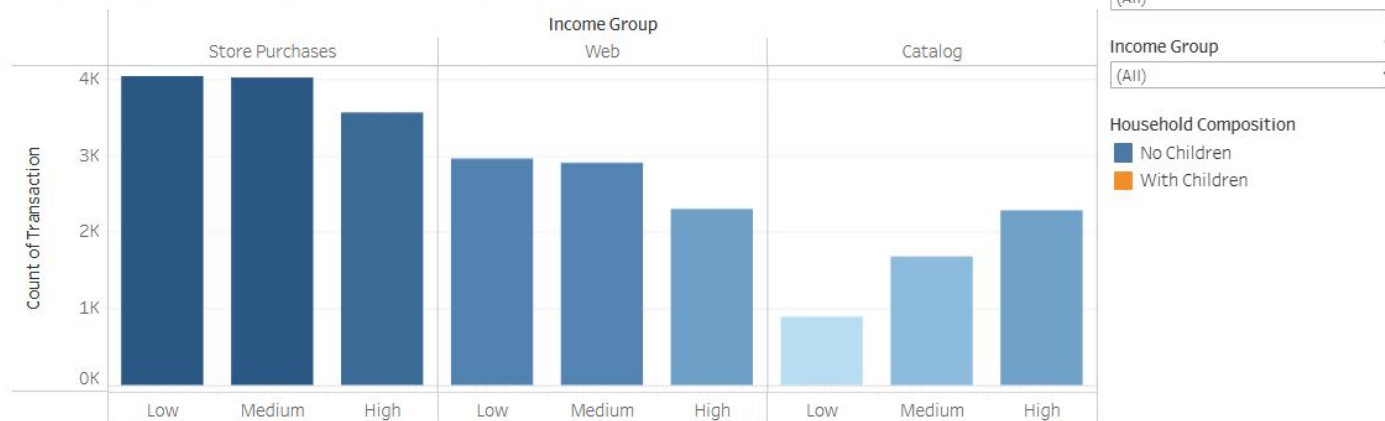
Web and Catalog: Moderate usage across income groups, suggesting opportunities to enhance online and catalog offers.

Spending Overview by Household Composition

Total Spending: Households with children contribute more to total spending despite having lower average spending per person, reflecting their dominance in customer numbers.

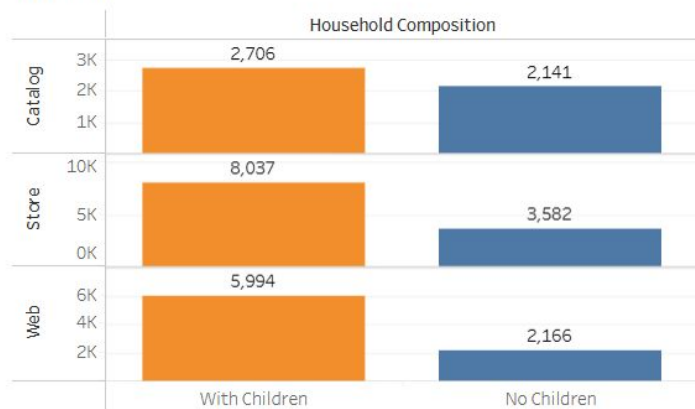
Opportunity: Tailor promotions for families to increase per-person spending while maintaining engagement.

Purchase Channel Breakdown by Household Composition

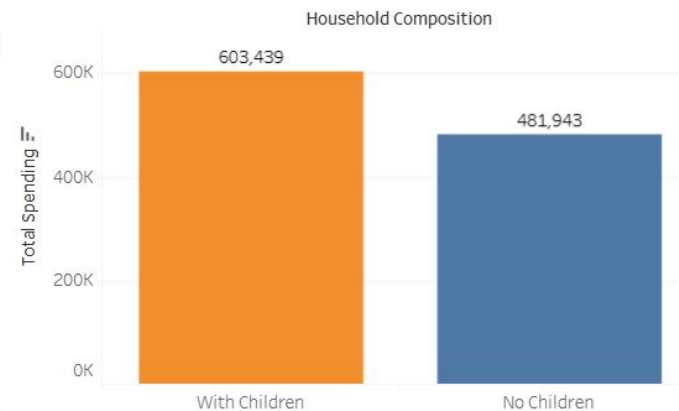


This is to reveal how family dynamics influence choices, enabling targeted marketing and promo tailored to families vs individuals for max engagement and sales impact.

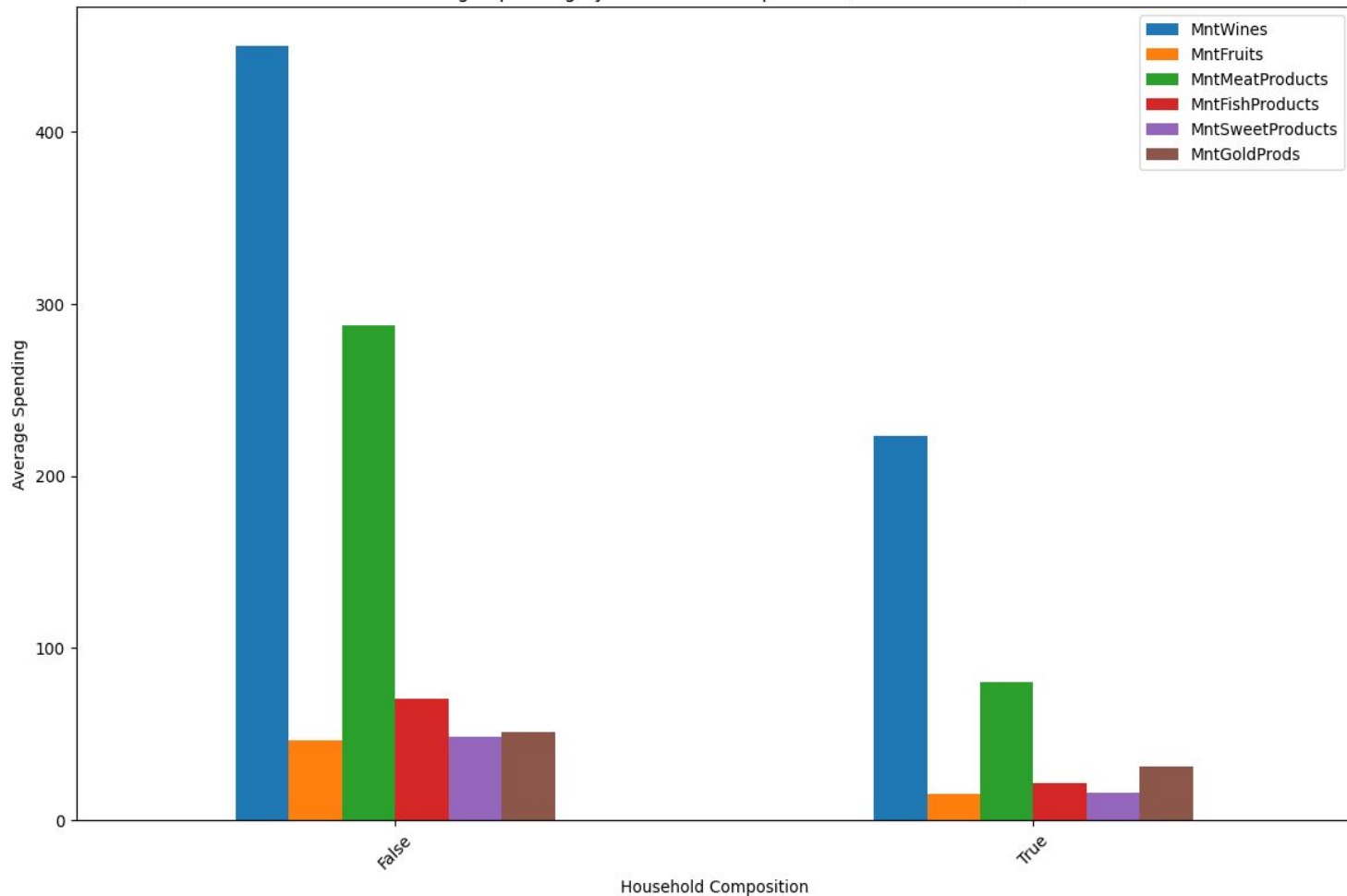
Channel Transaction Between Households With and Without Children



Total Spending Overview by Household Composition

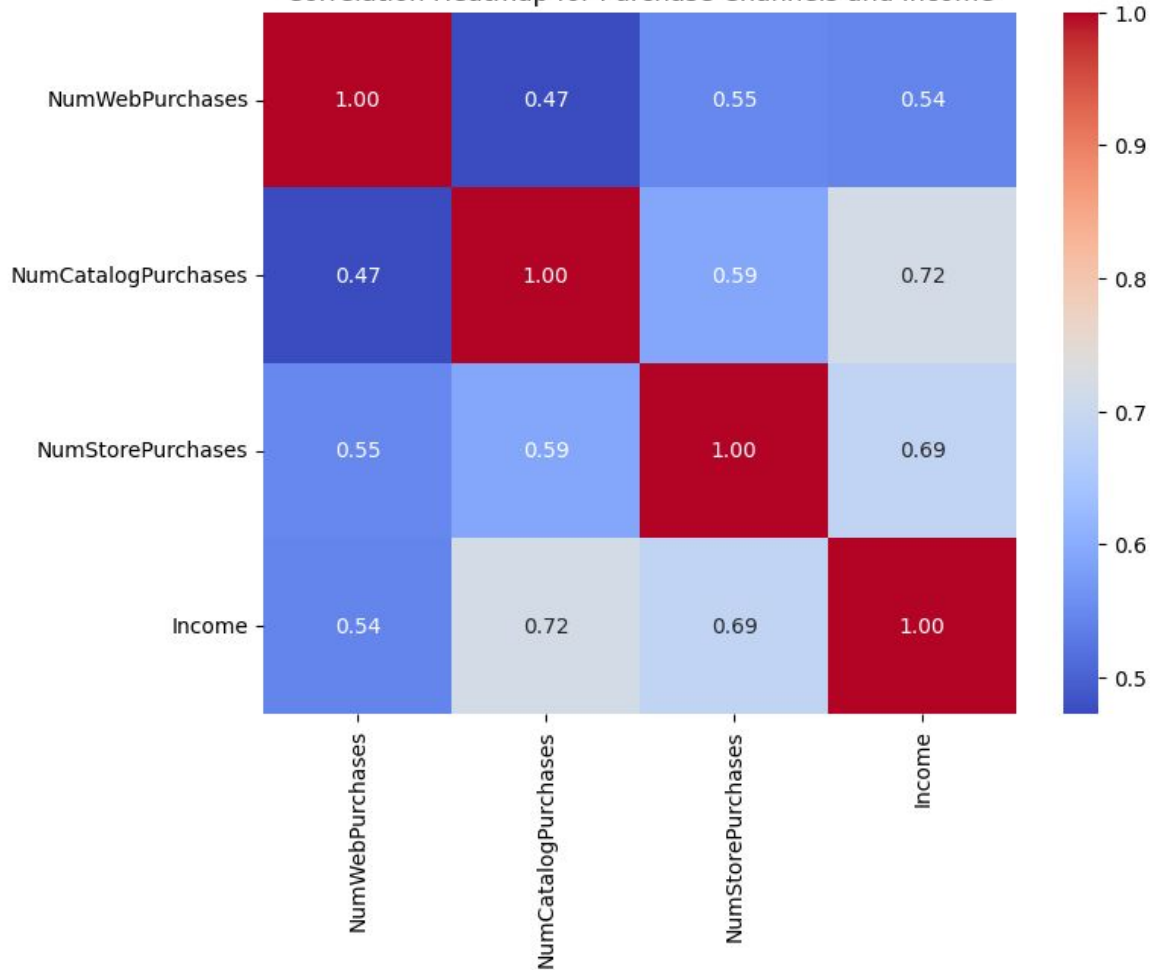


Average Spending by Household Composition (Has Kids or Teens)



- Households with children, despite being more numerous, have lower average spending per category compared to those without children.
- Opportunity: Targeted promotions and value-driven offers can help boost spending among families.

Correlation Heatmap for Purchase Channels and Income



Income and Total Spending

- Strong positive correlation (0.81). Indicates that higher-income customers spend more, making income a strong predictor of spending behavior.

Product Spending Patterns

- **MntWines:** High correlation with total spending (0.91) and with other product categories like meat (0.62). Suggests wine spending significantly drives overall customer spending.
- **MntMeatProducts:** Strong contribution to total spending (0.86 correlation). Shows importance in customer preferences across multiple segments.



Summary

Customer Insights: High-income customers and families with children are key drivers of spending, with a strong preference for wine and meat.

Campaign Performance: Campaigns 3, 4, and 5 are highly effective, while Campaign 2 has the lowest engagement, offering room for strategic adjustment or reallocation.

Spending Trends: Households with children contribute more to total spending, while those without children have higher average spending per person.

Channel Engagement: Multi-channel shoppers show strong engagement, with store purchases being the most common.

Growth Potential: Lower-demand categories like fruits, sweets, and fish offer opportunities for increased sales through targeted promotions.



Focused Marketing

High-Income Households

These customers spend the most, especially on wine and meat. Focus on promoting premium products to this group. Average spending on wine by high-income households is \$600 per year, and on meat is \$270 per year.

Households with Children

Families with children tend to spend more overall. Target these households with promotions for popular products like meat and sweets. Households with children spend on average 30% more on meat and 25% more on sweets compared to households without children.

Age Group 31-45

This age group makes up the largest part of our customer base, accounting for 43.7% of customers. Cater to their needs by offering family-oriented promotions and loyalty programs.



Campaign Strategy

Revise Campaign 2

Campaign 2 has been the least successful, with an acceptance rate of only 5%. Revise its message or target audience, or reallocate some of its budget to successful campaigns like Campaigns 3, 4, and 5.

Remove Campaign 2

Given its low acceptance rate of only 5%, consider discontinuing Campaign 2 altogether. Reallocate the entire budget to successful campaigns such as Campaigns 3, 4, and 5, which have consistently higher engagement rates. Focus on strategies that have proven to work, such as targeted messaging, better timing, and customer segmentation used in the successful campaigns.

Boost Successful Campaigns

Campaigns 3, 4, and 5 performed well, each with an acceptance rate of 15% or higher. Allocate an additional 20% of the marketing budget to these campaigns to expand their success.



Product Offerings and Inventory

Keep Popular Products in Stock

Wine and meat are the top sellers, with total sales of \$574,288 and \$270,066 respectively over the last two years. Ensure these products are always in stock, especially during promotions.

Expand Family-Friendly Products

Since many of our customers are families, offer more family-sized products, bundles, or meal kits. Families with children account for 75.5% of the customer base, and their average spending is 35% higher compared to households without children.

Tailor Products by Income

Offer discounts or value packs for lower-income customers, who represent 54.2% of the customer base and spend an average of \$174,120 annually. Introduce premium versions of popular products for high-income customers, whose annual spending is \$530,122 on average.



Customer Engagement

Loyalty Programs for High-Value Customers

Create or enhance loyalty programs for high-income customers and families. Offer exclusive deals or early access to promotions. High-income households represent 20.4% of customers and contribute 45% of total revenue.

Personalized Email Campaigns

Send targeted email offers based on income, household type, and age to increase relevance and engagement. Households with children and high-income earners are prime candidates, as they spend 30-50% more across product categories.

Monitor Results:

Track the success of the updated marketing strategies.



Links

<https://public.tableau.com/app/profile/vina.k2719/viz/shared/RG3WY4GP9>

<https://github.com/vinafasya/Capstone2-Data-Analysis>