Marketing Campaign Customer Segmentation

MIT Applied Data Science Program

April 2024

Executive Summary

OBJECTIVE	KEY INSIGHTS	NEXT STEPS
❖ Create distinct Customer Segments to help refine future marketing strategies ❖ Available customer data includes demographic info, past purchasing behavior, and past campaign responses	❖ K-Means clustering methods were used to identify 5 Customer Segments: ➢ **Budget Navigators** ➢ **Family Savers** ➢ **Moderate Spenders** ➢ **Gourmet Enthusiasts** ➢ **Wine Lovers**	❖ Target customers with Segment specific campaigns and marketing initiatives ❖ Continue to improve segmentation model by collecting new data and analyzing future campaigns

Problem Statement

- Businesses can optimize marketing strategies by **understanding customer behavior and preferences**
- Unsupervised learning techniques, such as **Dimensionality Reduction** and **Clustering**, can be used to divide an existing customer base into segments
- We can train various Clustering models on past customer data, including purchasing and campaign data, to determine the most effective **customer segments**
- Customer segmentation allows **tailored marketing approaches** to reach different segments, helping to increase ROI

Solution Approach

Step 1: Exploratory Data Analysis & Feature Engineering

Step 2: Scaling & Dimensionality Reduction

Step 3: Clustering & Segmentation

Step 4: Segment Profiling

Step 5: Solution Implementation

Data Overview

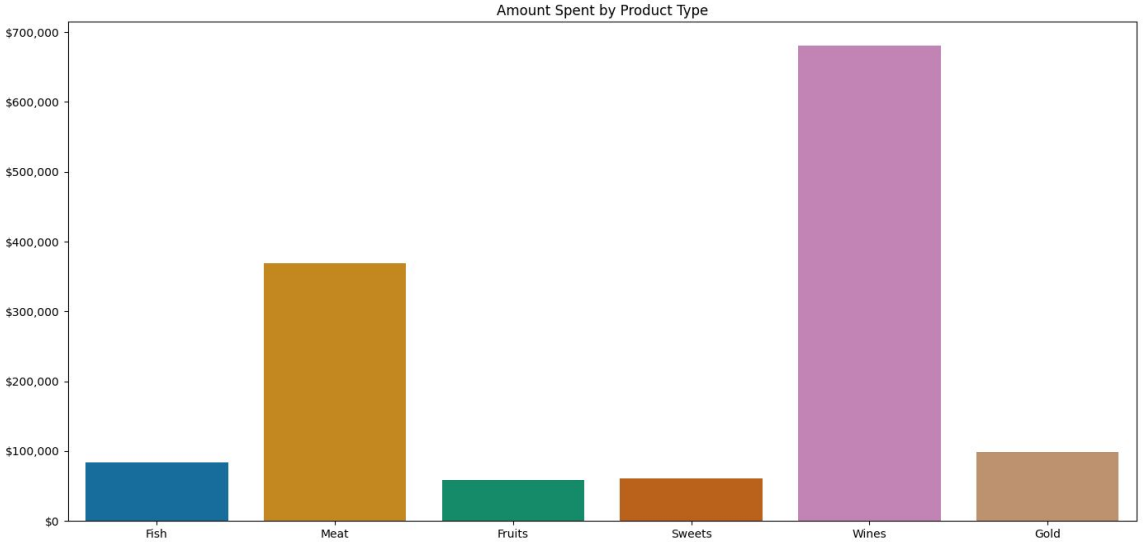
- Customer demographics
- Past purchasing behavior
- Past campaign responses

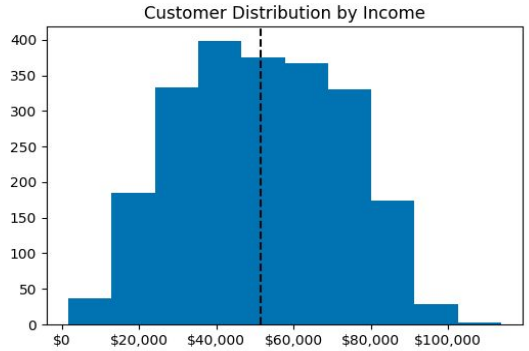
EDA Challenges

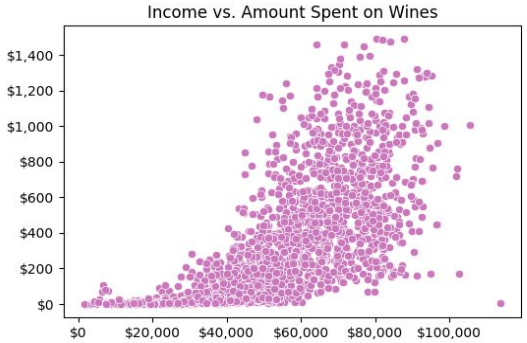
- Missing values
- Outliers

Key Trends

- Wine products generate the most revenue
- Strong positive correlation (0.73) between Income and amount spent on Wines





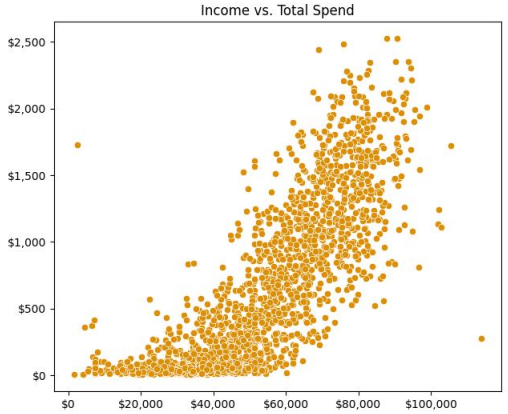


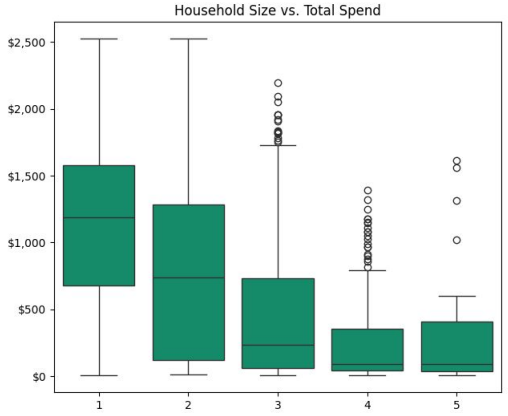
Key Features Created

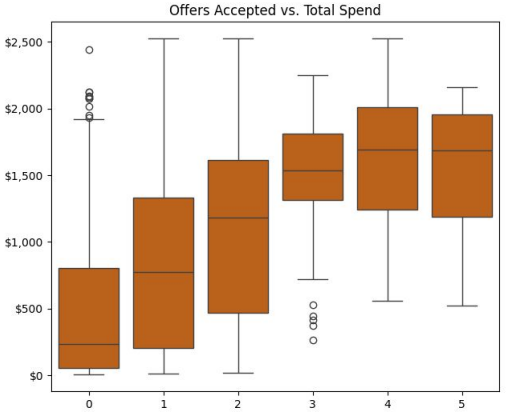
- Total Spend & Total Purchases
- Total Offers Accepted
- Household Size

Key Trends

- Income positively correlated with Total Spend (0.82) and Total Purchases (0.70)
- Negative correlation between Total Spend and Household Size (-0.43)
- Positive correlation between Total Spend and Offers Accepted (0.46)
- 27% of customers have accepted a past offer





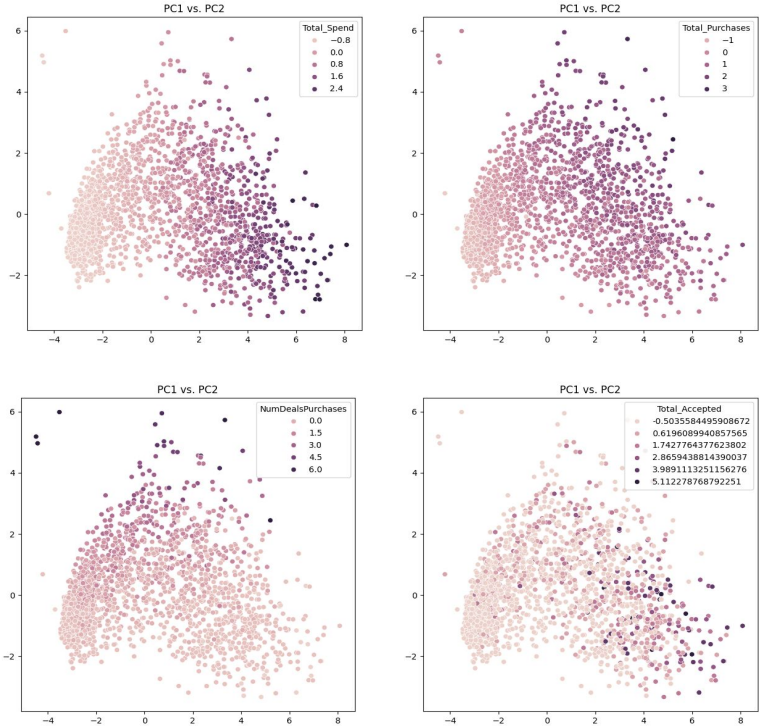


Dimensionality Reduction

List of features included in PCA and final training data:

- Total Spend
- Spend by category (Fish, Meats, Fruits, Sweets, Wines, Gold)
- Total Purchases
- Amount Per Purchase
- Purchases by channel (Store, Web, Catalog, Discount)
- Number of Web Visits
- Recency (of last purchase)
- Customer Tenure
- Total Offers Accepted

(Demographic features were eliminated from Clustering model, but later used for segment profiling)



Principal Component Analysis (PCA)

- PCA was initialized with 17 principal components
- PC1 captures variations related to overall spending behavior
- PC2 captures differences in channel engagement, including discount purchases

(Data was transformed with StandardScaler prior to applying PCA)

Clustering & Segmentation

Model Testing

The following models were tested for their ability to identify effective Customer Segments:

- K-Means
- K-Medoids
- Hierarchical (Agglomerative) Clustering
- DBSCAN
- Gaussian Mixture Model

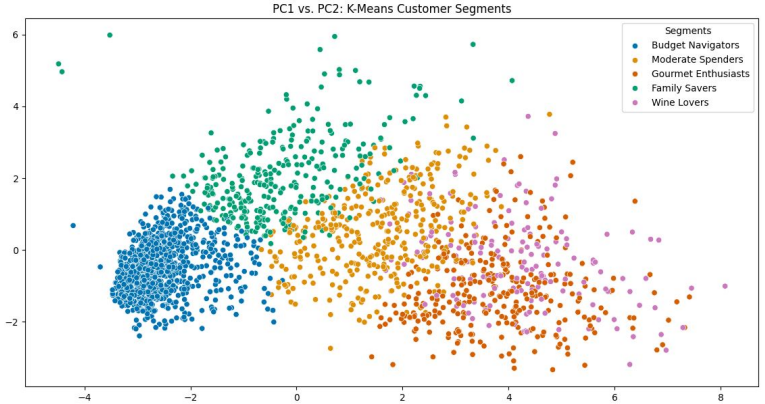
Criteria for model success included:

- Silhouette Score
- Cluster Interpretability
- Relevant Insights
- Potential Repeatability

Recommended Model

K-Means with 5 Clusters, trained on 10 PCs

- Silhouette Score = 0.26
- 10 PCs capture 90% of the variability in the data



Segment Profiling

Budget Navigators	Family Savers	Moderate Spenders	Gourmet Enthusiasts	Wine Lovers
Lowest-income customers Lowest spending across all categories Fewest purchases across all channels Largest segment, with 988 customers Youngest age group Frequent web visits Least likely to accept campaign offers	Middle-income customers with families Highest preference for discount purchases Relatively low spending in all product categories Household size of 2-5 Higher than average web visits and web purchases	Mid-to-high income customers with average spending Mid-range spending across all categories Oldest age group Highest preference for in-store purchases Households tend to be smaller Least likely to accept campaign offers	High-income customers buying food products Mid-to-high spending across all categories Highest-spending on Sweets, Fruits, Meat, and Fish Higher preference for catalog purchases Households tend to be small with no kids	High-income customers buying wines Mid-to-high spending across all categories Highest-spending on Wines, and overall Smallest segment, with 172 customers Higher preference for catalog purchases Most likely to accept campaign offers

Recommendations & Next Steps

Solution Implementation

- Work with Marketing team to reevaluate past campaign messaging
- Create new segment-specific marketing strategies
- Conduct A/B tests for new messaging and outreach channels

Potential Benefits

- Increased marketing efficiency and ROI
- Increased customer loyalty
- Ability to target prospects more effectively, creating a competitive advantage

Potential Risks

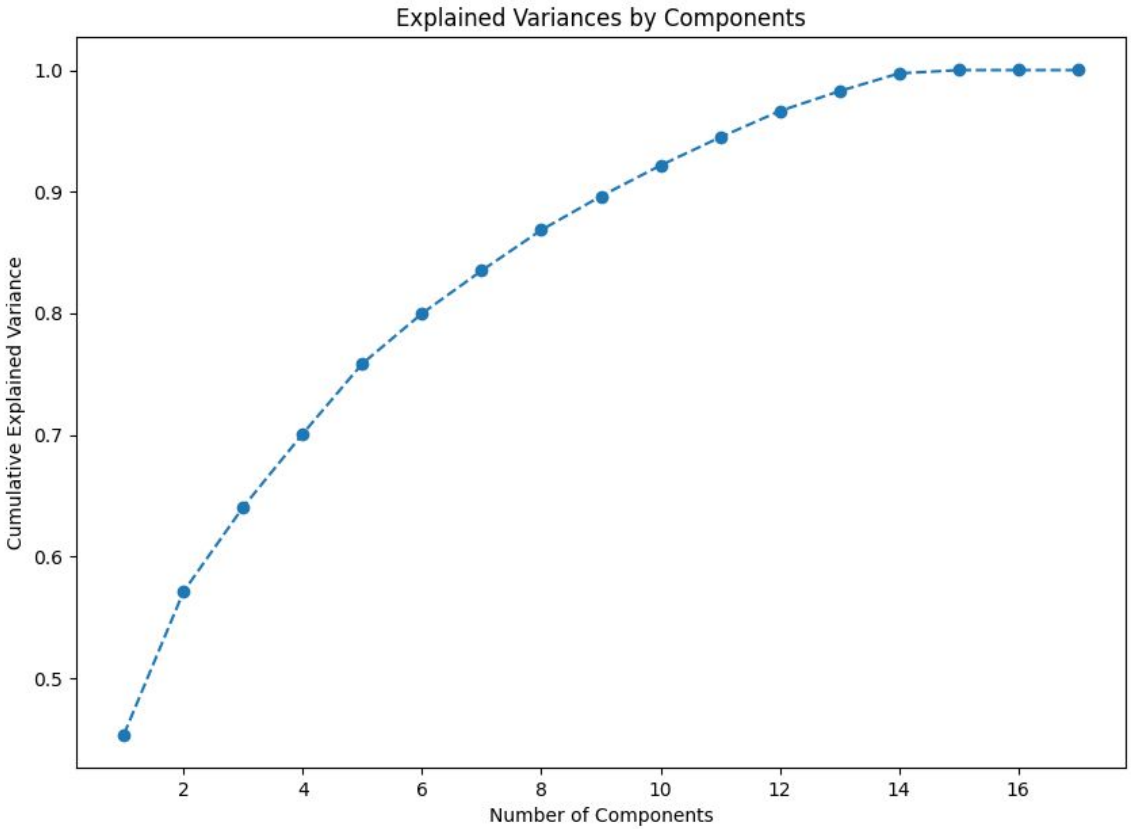
- Model overfitting or overgeneralization
- Operational challenges

Next Steps for Analysis

- Validate and adjust K-Means model based on A/B test results
- Incorporate additional data to refine K-Means model and segment profiles

Appendix

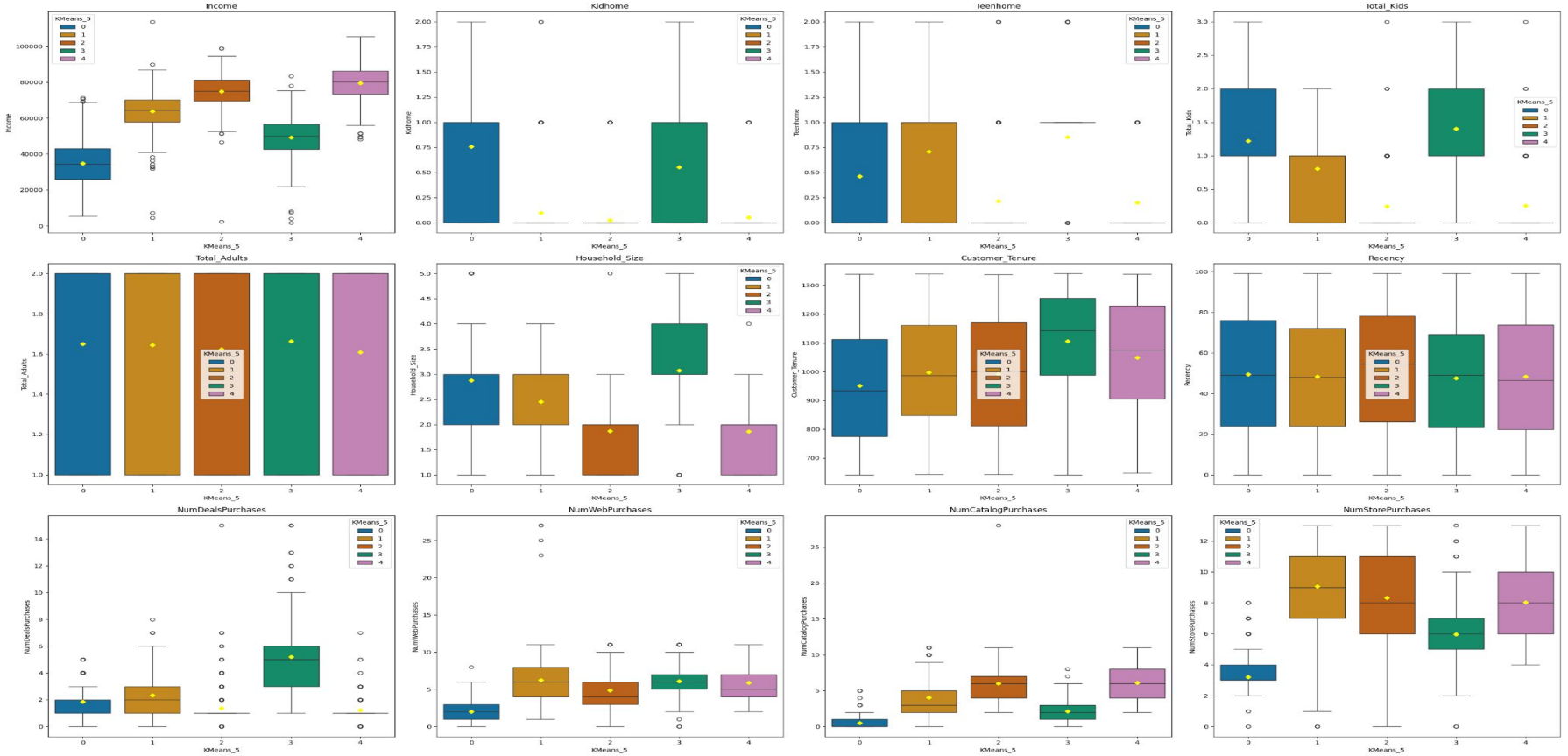
PCA Explained Variance



Model Comparison

Model	Clusters	Silhouette Score
K-Means	5	0.26
K-Medoids	5	0.13
GMM	5	0.12
Agglomerative	3	0.26
Agglomerative	5	0.23
DBSCAN	2	0.54

K-Means Segment Comparison



K-Means Segment Comparison