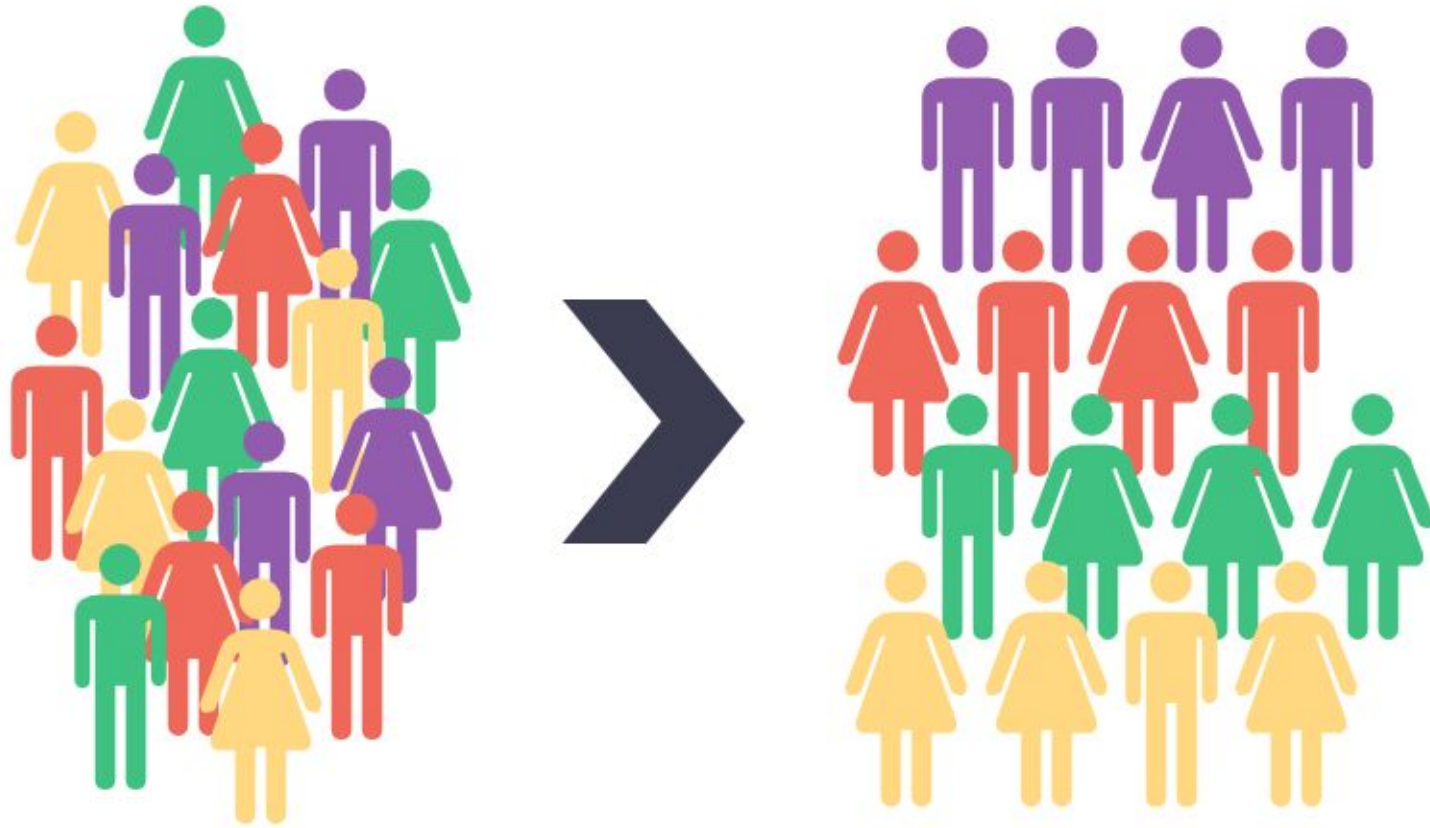
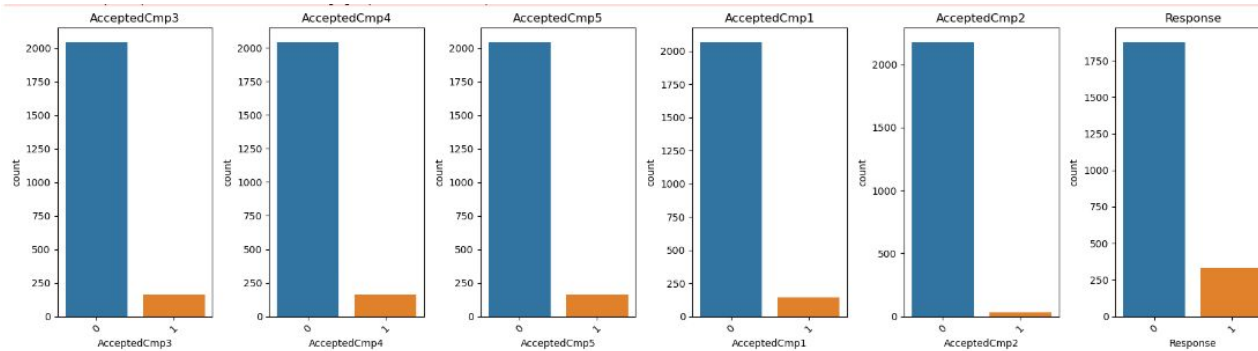


## Clustering-Based Customer Segmentation for Marketing Success



# Business context



- **Marketing Challenges**
- **Campaign Overview:**
  - Six marketing campaigns
  - Minimal success and no revenue improvement
- **Campaign Results:**
  - < 15% customer acceptance for the first 5 campaigns
  - < 20% acceptance in the last campaign

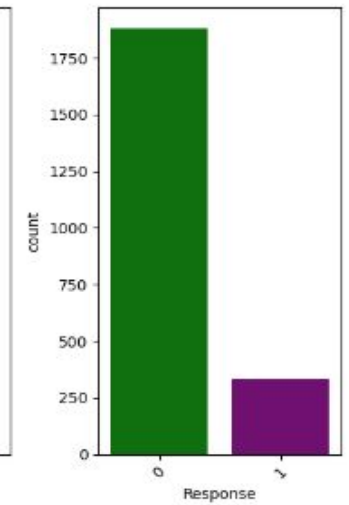
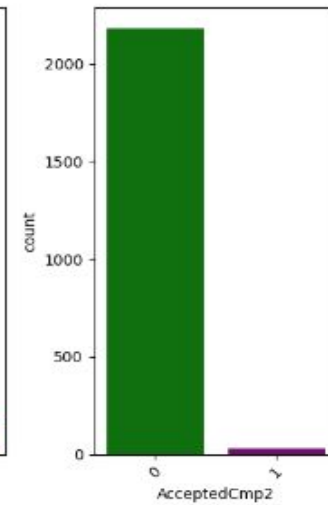
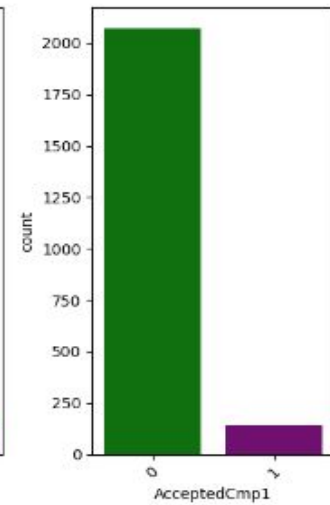
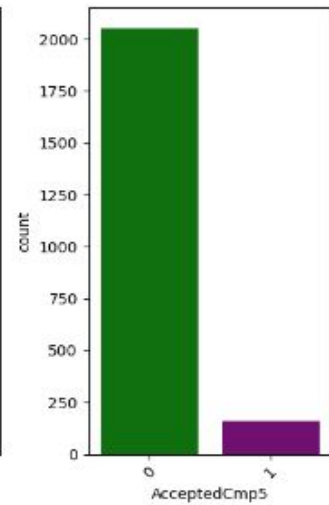
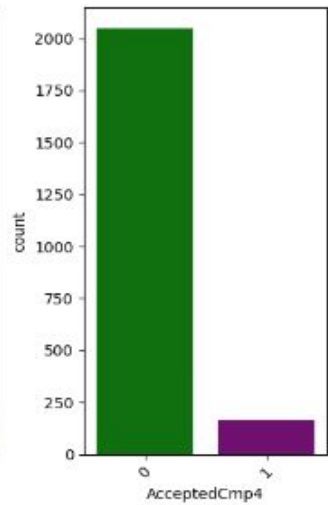
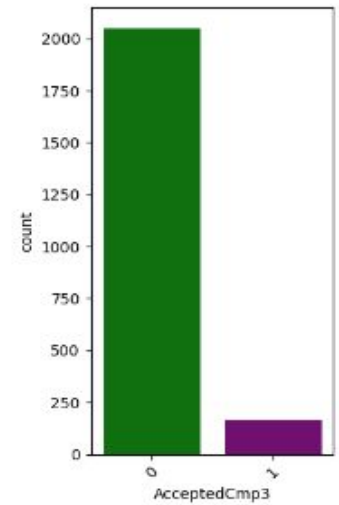
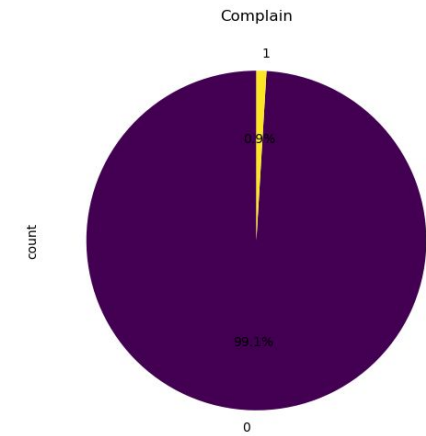
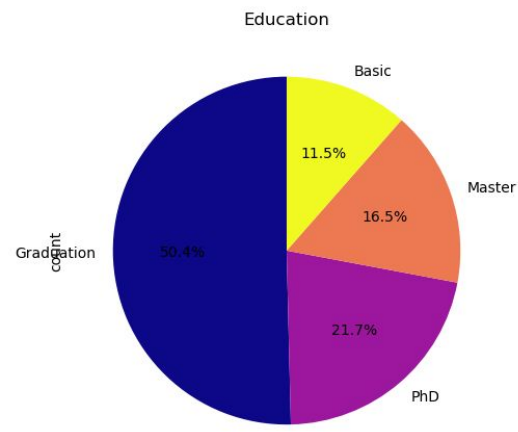
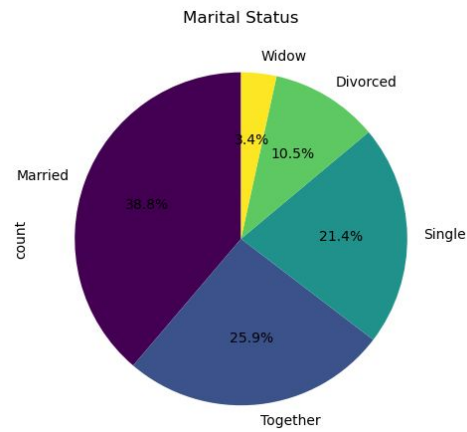
# Data source

- Customer Personality Analysis data frame from Kaggle.
- 2240 rows and 29 columns
- Data consist of categorical and numerical data.

People	Products
ID: Customer's unique identifier	MntWines: Amount spent on wine in last 2 years
<b>Year_Birth: Customer's birth year (NEW FEATURE-AGE)</b>	MntFruits: Amount spent on fruits in last 2 years
Education: Customer's education level	MntMeatProducts: Amount spent on meat in last 2 years
Marital_Status: Customer's marital status	MntFishProducts: Amount spent on fish in last 2 years
Income: Customer's yearly household income	MntSweetProducts: Amount spent on sweets in last 2 years
Kidhome: Number of children in customer's household	MntGoldProds: Amount spent on gold in last 2 years
Teenhome: Number of teenagers in customer's household	<b>NEW FEATURE- TOTAL</b>
<b>Dt_Customer: Date of customer's enrollment with the company(NEW FEATURE- MNTHS CUSTOMER)</b>	
Recency: Number of days since customer's last purchase	
Complain: 1 if the customer complained in the last 2 years, 0 otherwise	
Promotion	Place
NumDealsPurchases: Number of purchases made with a discount	NumWebPurchases: Number of purchases made through the company's website
AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise	NumCatalogPurchases: Number of purchases made using a catalogue
AcceptedCmp2: 1 if customer accepted the offer in the 2nd campaign, 0 otherwise	NumStorePurchases: Number of purchases made directly in stores
AcceptedCmp3: 1 if customer accepted the offer in the 3rd campaign, 0 otherwise	NumWebVisitsMonth: Number of visits to company's website in the last month
AcceptedCmp4: 1 if customer accepted the offer in the 4th campaign, 0 otherwise	
AcceptedCmp5: 1 if customer accepted the offer in the 5th campaign, 0 otherwise	
Response: 1 if customer accepted the offer in the last campaign, 0 otherwise	

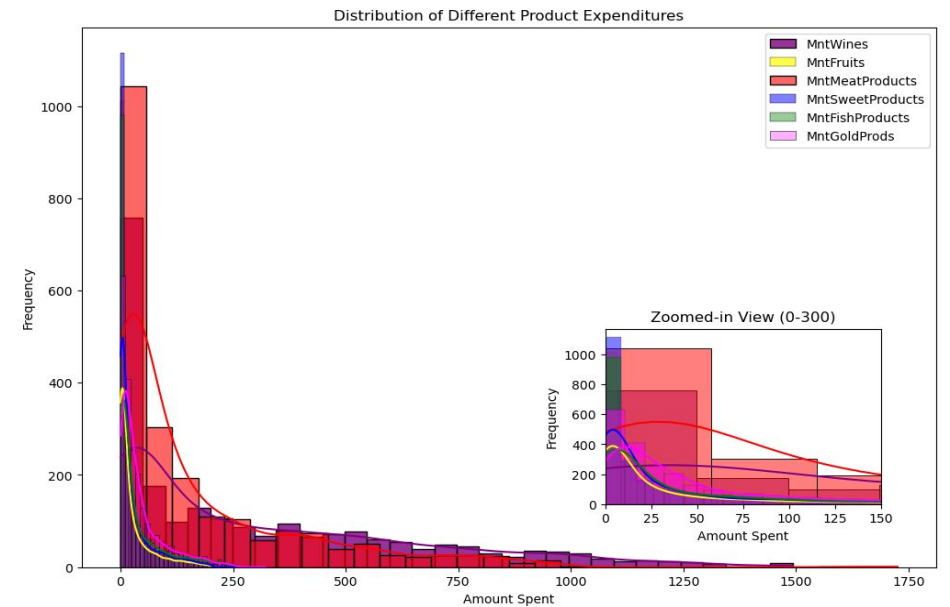
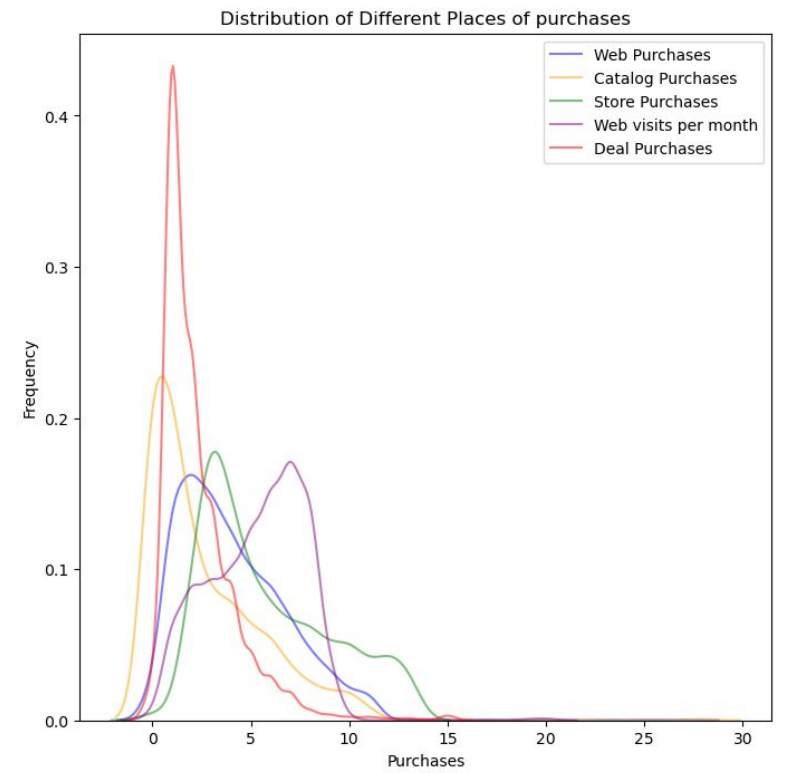
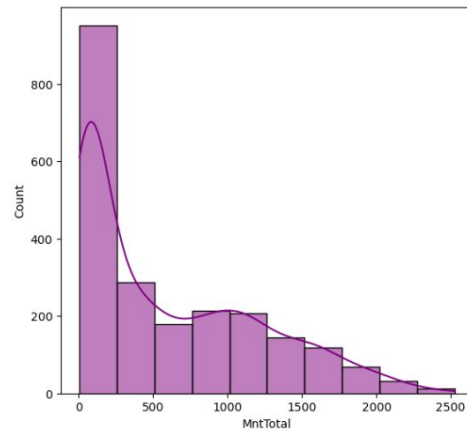
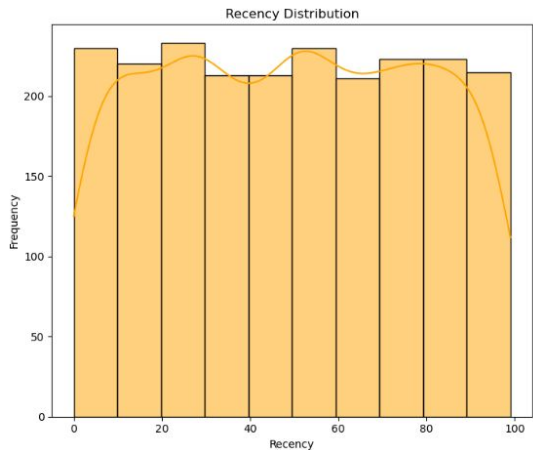
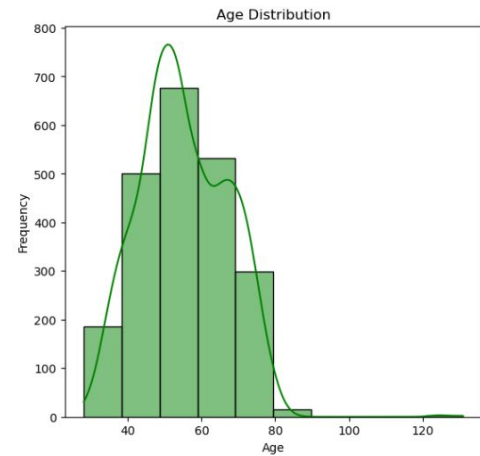
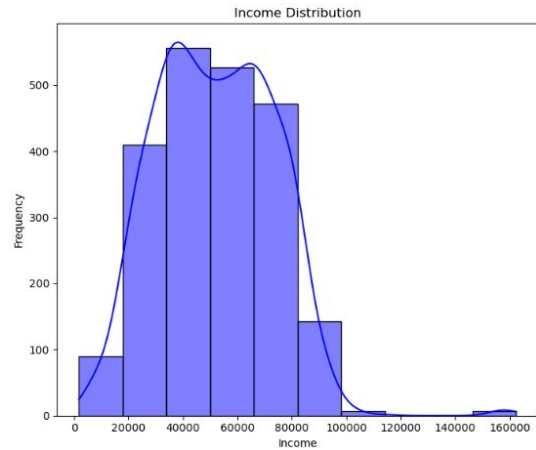
# EDA

## Exploring categorical variables



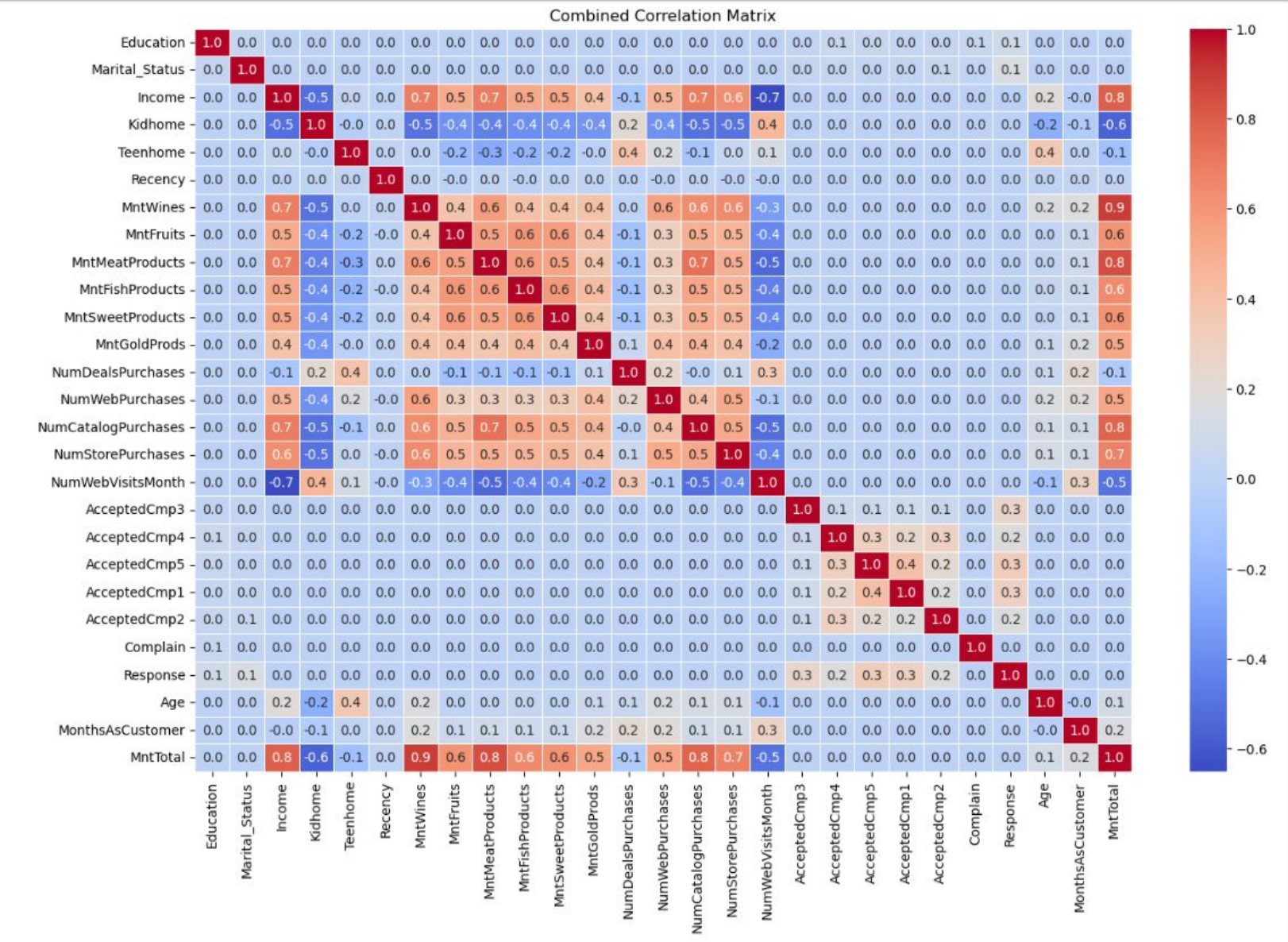
# EDA

## Exploring Numerical variables



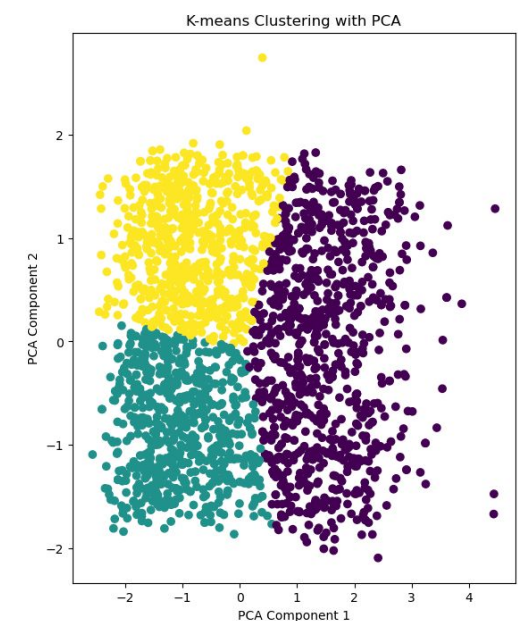
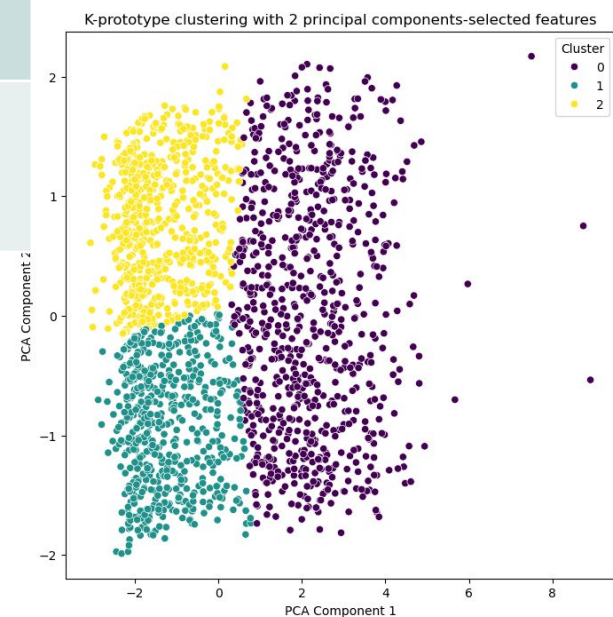
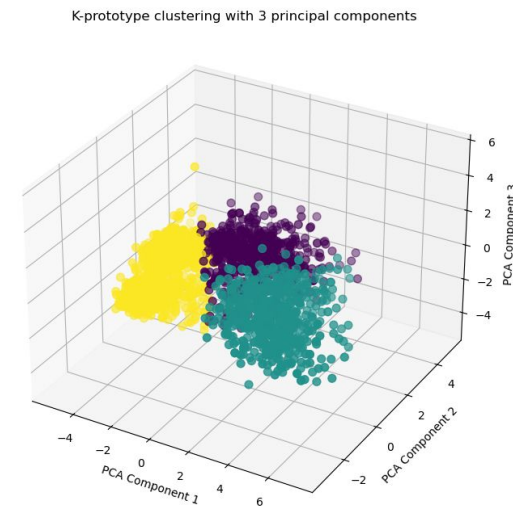
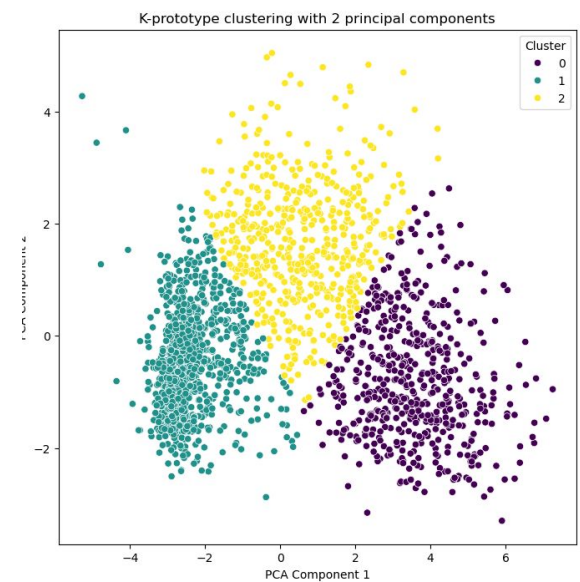


Correlation Matrix



# Clustering

Cluster	Evaluation Metrics
K-Prototype clustering with 2 principle components	Silhouette Score: 0.352 Davies-Bouldin Index: 1.118 Calinski-Harabasz Index: 1780.5
K-Prototype clustering with 3 principle components	Silhouette Score: 0.304 Davies-Bouldin Index: 1.271 Calinski-Harabasz Index: 1353.7
K-prototype clustering with 2 principal components-selected features	Silhouette Score: 0.182 Davies-Bouldin Index: 1.532 Calinski-Harabasz Index: 559.7
K-means Clustering with PCA	Silhouette Score: 0.127 Davies-Bouldin Index: 3.215 Calinski-Harabasz Index: 658.156



# Cluster Characteristics

K-Prototype clustering with 2 principal components

## Cluster 0-

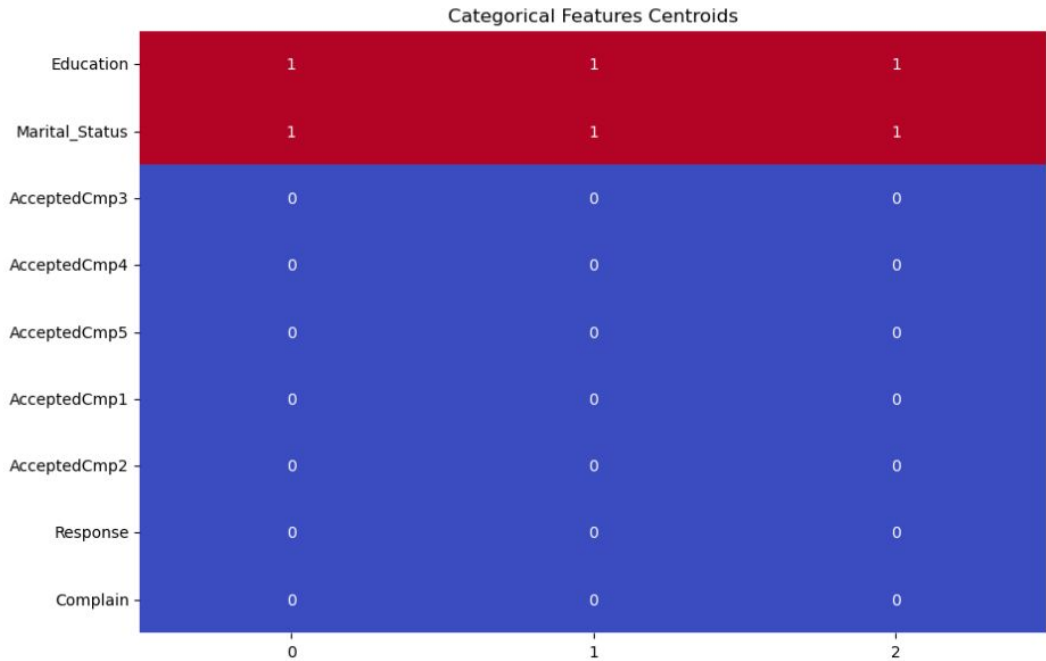
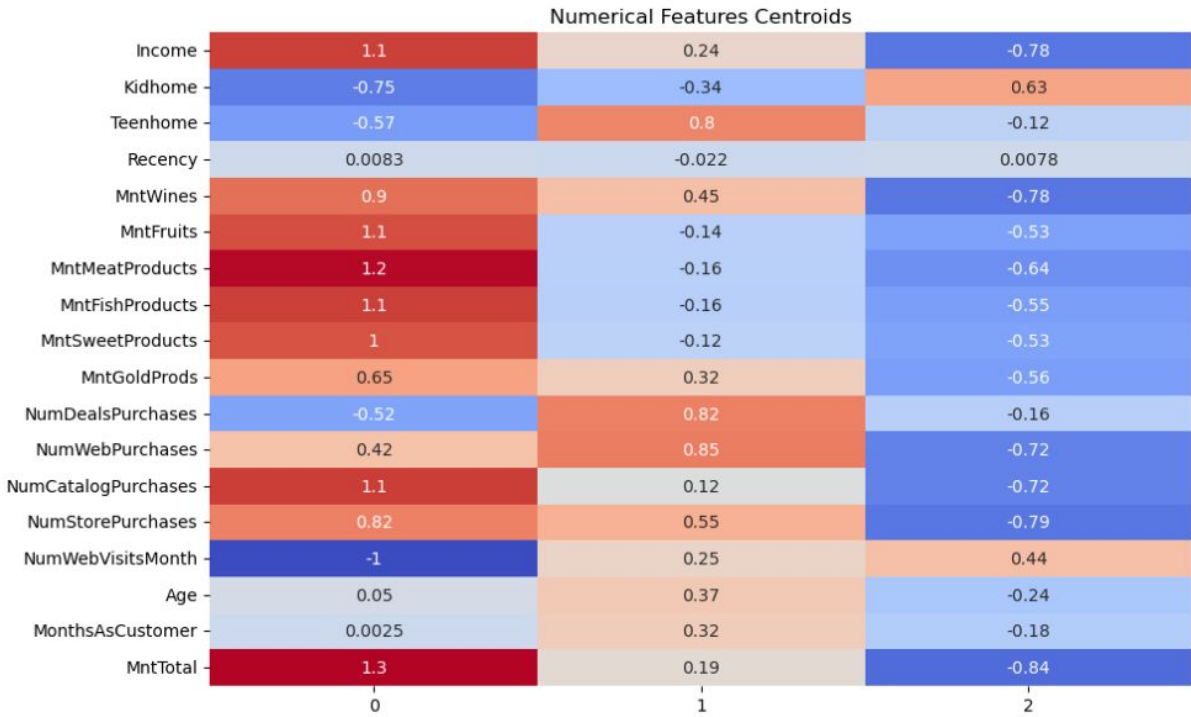
This cluster represents customers with high income and high spending across various product categories. They are likely well-educated and married, with fewer children and teenagers at home.

## Cluster 1-

This cluster represents customers with moderate income and spending. They are well-educated and married, with more teenagers at home.

## Cluster 2-

This cluster represents customers with lower income and spending across various product categories. They are well-educated and married, with more children at home.





# Future work, Feedback and suggestions

