

# Retail Customer Segmentation Analysis



Anand Umrani



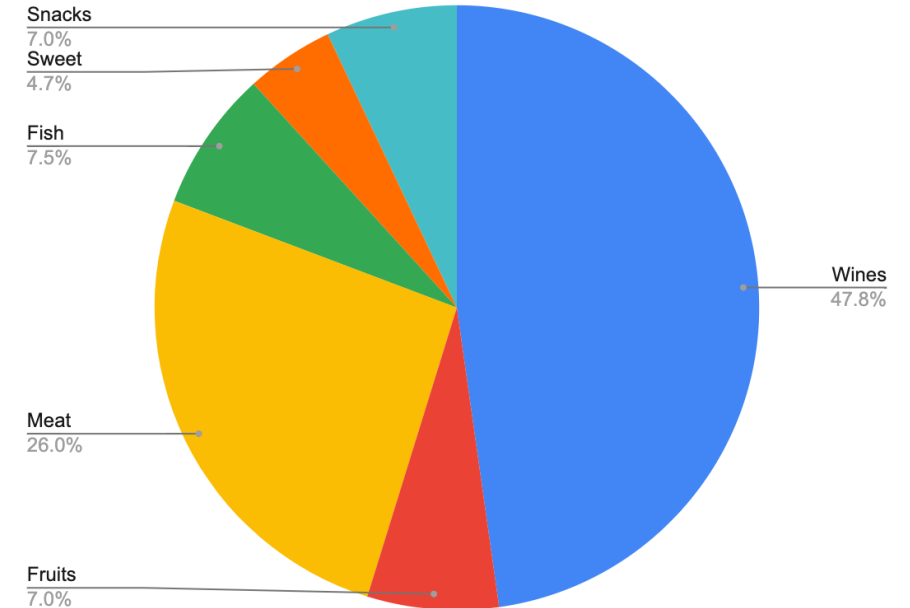
# Agenda

---

- Background, Problem Statement & Objectives
- Approach
- Insights from Analysis
- Strategy Recommendations

# Background, Problem Statement & Objective

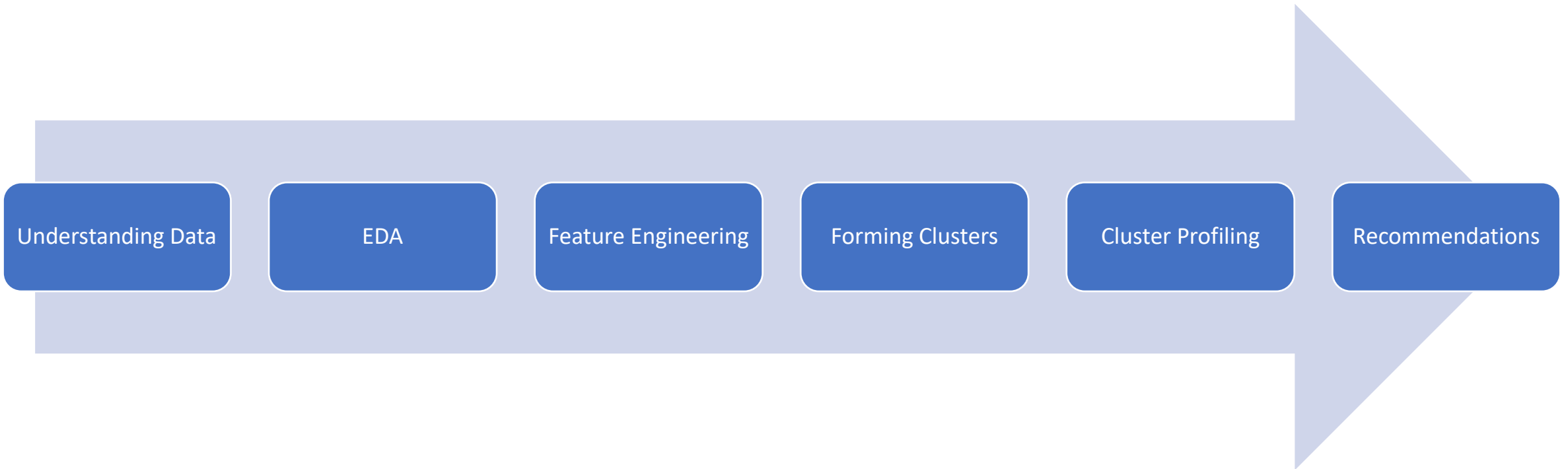
- Retailkart.com is a small and medium-scale organization that majorly deals in wine, fruit and meat products, having held around 35% market share.
- The company has been leading the domain offline for a long time.
- The company main challenge is to stay competitive against the rapidly expanding companies offering quick and free deliveries while still accepting online purchases at affordable costs
- The company's near-term objective is to maintain its overall market share, with the increasing competition offering 10-minute delivery, among other benefits.



Identify customer groups which have different purchasing preferences based on their affluence and other attributes and devise an effective marketing plan & preference-based user experience at a cohort/group level.

# Approach

---



Tools used: Python



# Data

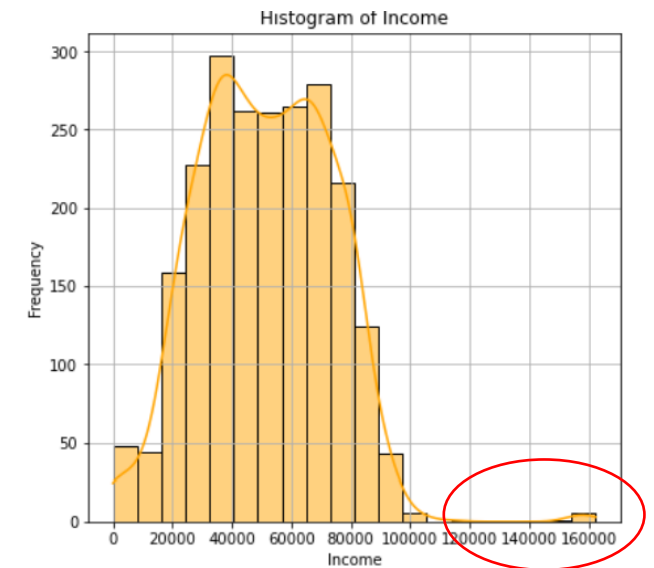
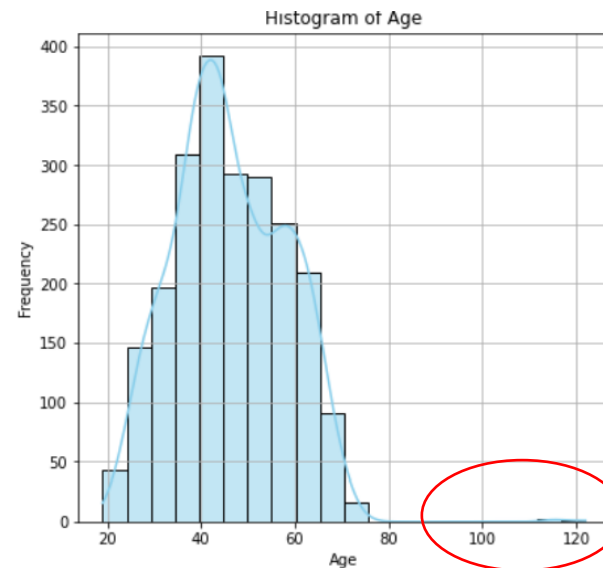
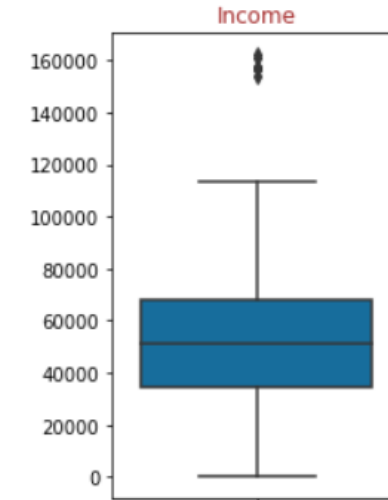
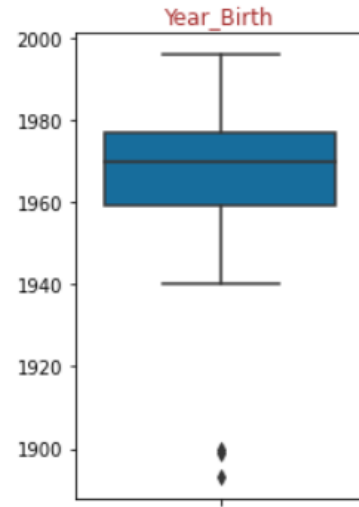
- The data is available at an aggregated level for each customer on the platform from appx from 2012 to 2014
- The data includes their historical transactions in each category, demographics, and reactions to promotions.
- Shape of Data
  - 2239 rows & 29 columns features
  - 26 numerical and 3 categorical features
- Data Quality
  - Very few missing values\* (Income 1 record & Responses 23 records)

\* Missing value for Income was imputed by 0. For Responses the null values were replaced with an approximation of whether the customer has accepted any campaigns at a later stage or not.

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

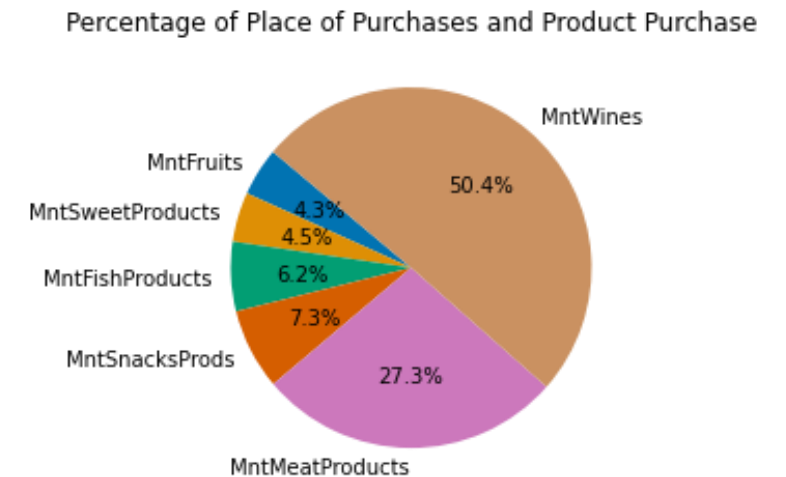
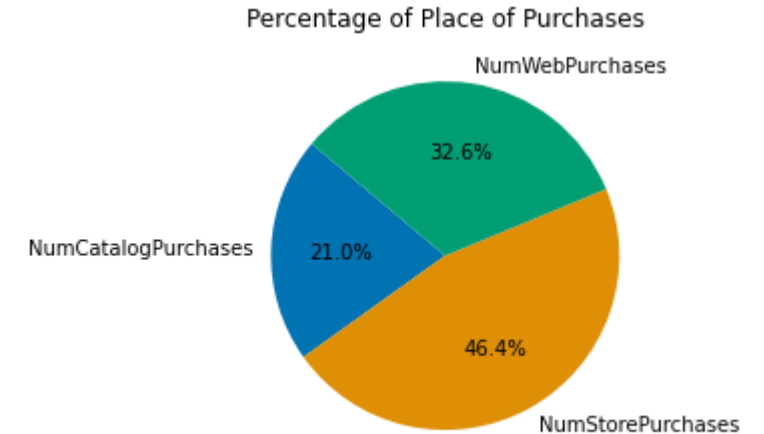
# EDA

- Outlier Analysis was done using Box Plot for all features
- Outliers were observed for Income and Year of Birth
- Outliers were removed for Income and after converting Year\_Birth to Age



# EDA

- Appropriate exploratory analysis was conducted to extract useful insights (whether directly useful for business or for eventual modeling/feature engineering).
- Insights from EDA:
  - Most of the customers are in the age group of appx 35 to 50 yrs with appx 50k as income
  - Majority of the customers are Graduates followed by PhD and Masters Degree.
  - Customers who are married and staying together are more than divorced, single, etc.
  - Maximum customers dont have kids or teens at home.
  - Majority of the customers are purchasing from the store followed by web. Which is not that bad.
  - Majority of the customers are buying Wines and Meat Products.
  - Very few customers are purchasing fruits and sweet products



# Feature Engineering

Features Engineered	Method/Approach
Age	Taking 2015 as-of date since data is till end of 2014 and using Year_Birth
Total_Spent	Sum of amount spent on all products purchased by each customer
Member_for	Difference between 01-01-2015 and Dt_Customer
Total_Kids	Sum of Kidhome and Teenhome
Total_Purchases	Sum of purchases made from places – Web, Catalogue, Store
Total_AcceptedCmp	Sum of all accepted campaign columns.
Clubbing Product Categories <ul style="list-style-type: none"><li>- Meat &amp; Fish</li><li>- Snacks &amp; Fruits</li><li>- Alcohol</li></ul>	Sum of purchase amount for Meant and Fish Sum of purchase amount for Snacks, Fruits and Sweets Renames Wines
Education: <ul style="list-style-type: none"><li>- High_Edu</li><li>- Basic_Edu</li></ul>	Customers with education - PhD, Cycle, Graduation, Master Customers with education – Basic
Marital_Status <ul style="list-style-type: none"><li>- Single</li><li>- Not-Single</li></ul>	Divorced, Widow, Alone, YOLO, Absurd Married, Together



# Correlation Insights

## High Income Customers

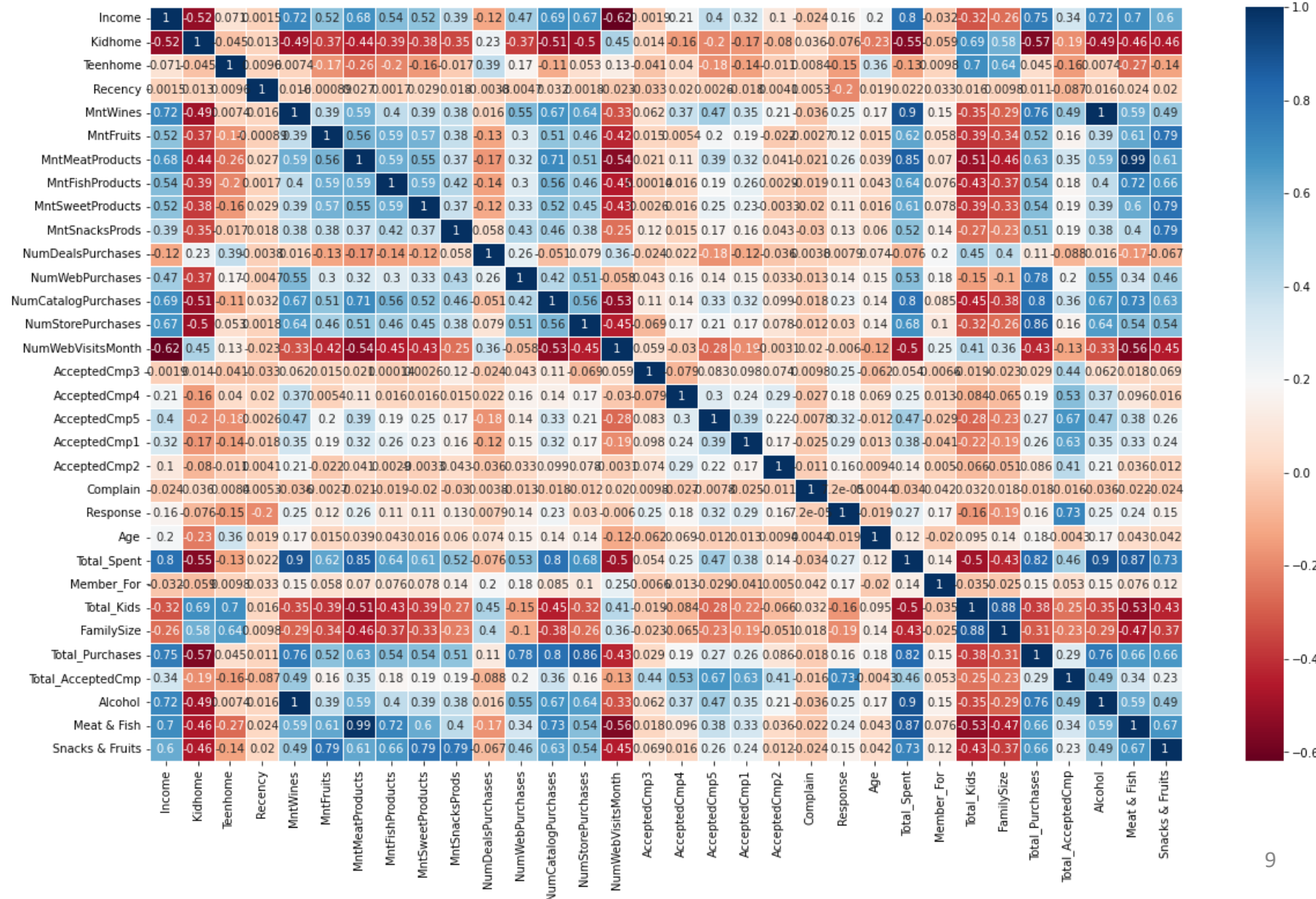
- spend more and purchase more.
- visit the company's website less frequently.
- Have few numbers of purchases made with a discount

## Customers with High Spent

- do not to visit the company's website.
- buy more wines and meat products
- make a high number of purchases made using a catalog

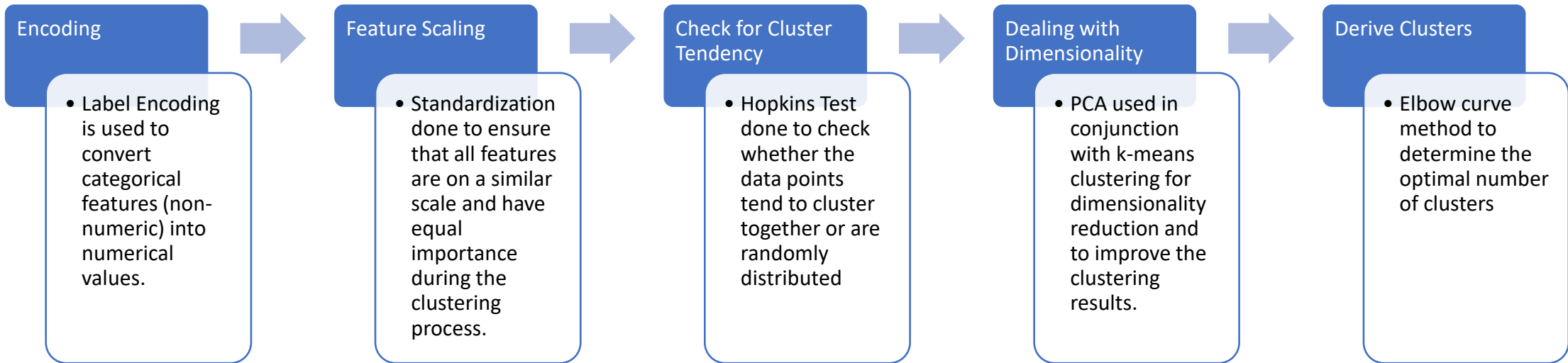
## Customers with kids

- spend less and purchase less.
- have a high number of purchases made with a discount



# Clustering/Segmentation Approach

---

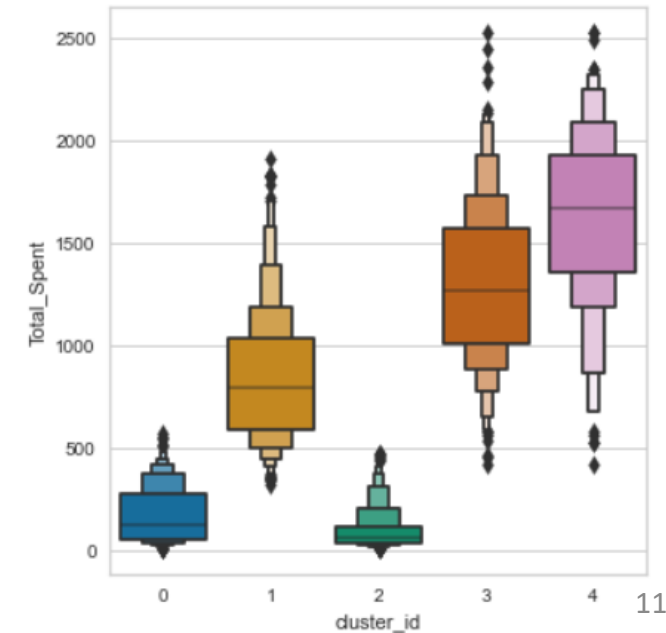
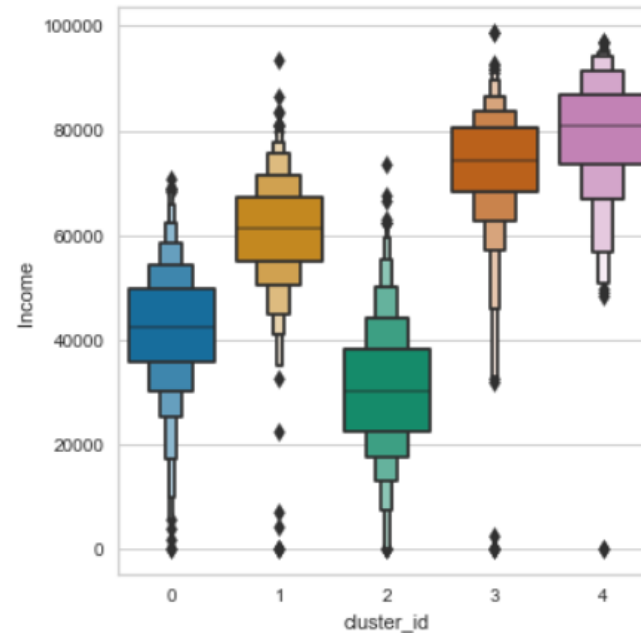
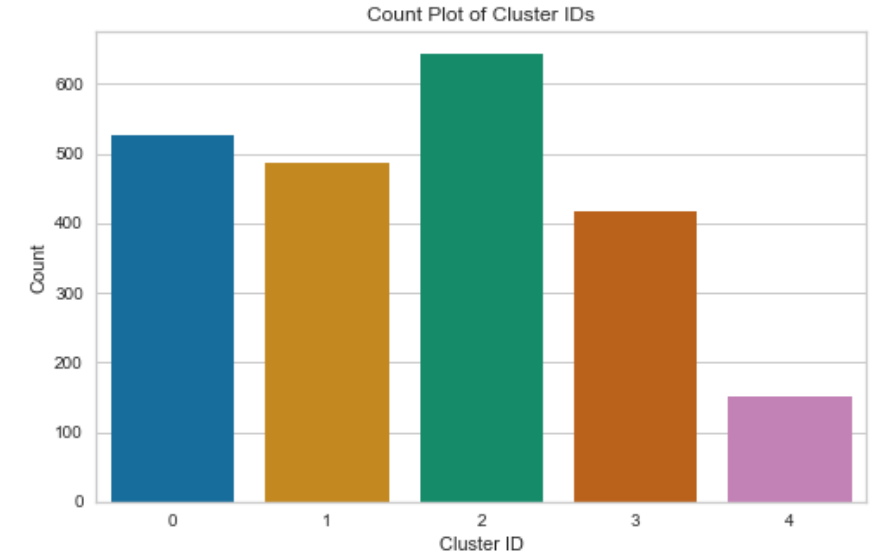


# Cluster/Segment Profiles

Total 5 Segments were identified as follows:

- Cluster 4: High Income High Spenders
- Cluster 3: High Income Avg Spenders
- Cluster 1: Avg Income Avg Spenders
- Cluster 0: Avg Income Low Spenders
- Cluster 2: Low Income Low Spenders

[Click here to Refer Annexure after Thank You slide for detailed visualization](#)



# Segment Behavior

---

## High Income High Spending

- Spend a lot of amount in purchasing meat & fish and alcohol
- They make very less Deal Purchases and make most of purchases from store
- Their visit to the website in last month is very low but their response to past campaigns is high
- Majority of them don't have kids at home

## High Income Avg Spenders

- This group has sizable number of customers
- They do spend on an average considerable amount of purchases on all products
- Their last month visit to the web is the lowest in all clusters
- Their place of purchase amount is almost similar to high income and high spend segment
- They have smaller family size

## Avg Income Avg Spenders

- This segment spends more amount on alcohol and fruits+snacks+sweets than meat+fish
- Their web purchases are comparatively higher than other segments
- Their web visits last month were high and so were the purchases made through deals
- They have a slightly higher family size

## Avg Income Low Spenders

- This segment spends low amount on purchases
- They do more deal purchases
- Their visit to web last month was higher than others
- They have the largest family size

## Low Income Low Spenders

- This segment is quite sizable in number of segmented customers
- As the name suggests they very less purchases and amount spent is also very low
- Whatever purchases they do is through stores
- However, their last month visit to website is highest among all segments
- They have large family size

# Customized Recommendations for Segments

---

- **High Income High Spending Segment:**

- Enhance the online shopping experience by improving the website's user interface, making it user-friendly and visually appealing.
- Provide online deals and promotions to encourage more web purchases from this segment.
- Implement a loyalty program for frequent online shoppers with discounts or wallet points for future purchases.
- Targeted email marketing and personalized product recommendations based on their past online purchases.

- **High Income Avg Spenders Segment:**

- Develop online-exclusive offers and limited-time promotions to attract this segment to make more web purchases.
- Increase online advertising and promote special deals available only on the website.
- Offer free shipping or discounts on bulk purchases.
- Implement campaigns to remind this segment of products they viewed but did not purchase during their previous online visits.

- **Avg Income Avg Spenders Segment:**

- Create loyalty rewards that encourage repeat purchases and offer special schemes for online shoppers.
- Implement a referral program to encourage satisfied customers to refer friends and family to make online purchases.
- Email marketing to inform this segment about exclusive online deals and upcoming promotions.

- **Avg Income Low Spenders Segment:**

- Create limited-time online flash sales to attract this segment.
- Offer free shipping for online orders to remove potential barriers to online shopping for this segment.
- Flash product descriptions and customer reviews on the landing page of website to build trust and confidence in online purchases.
- Implement a user-friendly checkout process with various payment options to modernize online purchases.
- Improve websites performance for loading and checkout.

# Summary of Recommendations

---

- Retailkart.com should use web analytics to understand customer/user behavior and identify potential points of improvement on the website.
- Regularly update and refresh Retailkart.com website content to keep it engaging and relevant to the target audience.
- Leverage social media and online advertising to direct traffic to the website during promotional periods.
- Encourage customer feedback and reviews to build trust and credibility among online shoppers.
- Monitor and optimize the website's loading speed to provide a seamless online shopping experience.

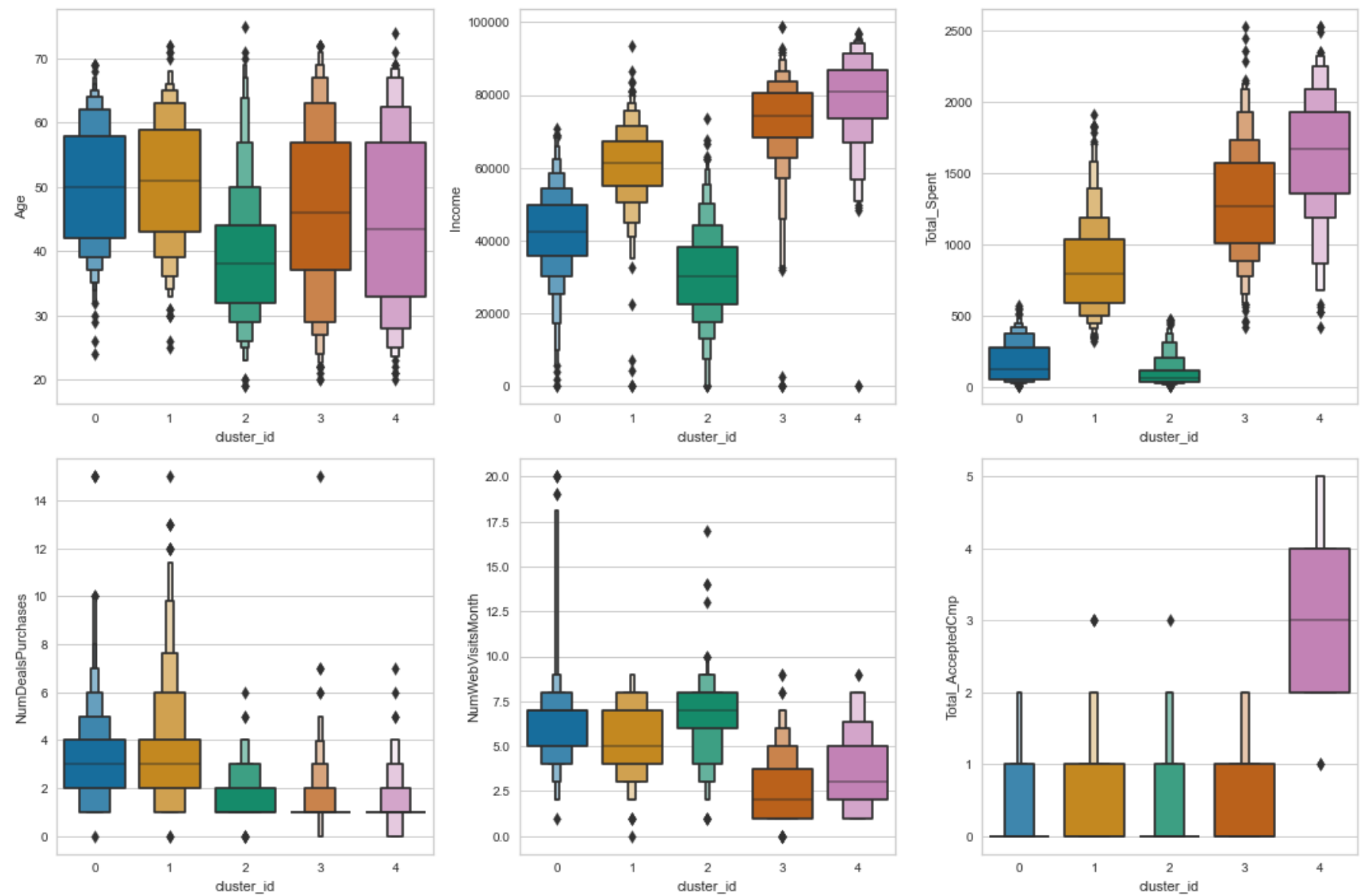




# Thank you

---

# Cluster Behavior Visualization (1/2)



# Cluster Behavior Visualization (2/2)

