

## Project Background

The success for every company is not just tied to offering competitive products and services, but also to have good insights into their customer profile and behaviour.

Knowing your existing and target customers is essential for effective marketing campaigns and advertisements.

Targeting the right customer segments is key. Aiming too broadly will increase cost and might attract only a small percentage of customers, but by keeping it too narrow the company could miss out on selling opportunities.

Customer segmentation can be a valuable tool to target the right people and provide them with the right buying incentives. With customer segmentation we divide the customers into groups based on common characteristics, like preferences, historic consumer behaviour or demographics. This information can then be used to tailor marketing initiatives to the individual groups and thus maximizing the potential response. In this context it is also important to analyze which marketing channels are effective for each customer. Older generations might respond better to classic channels like TV or email, while younger customers are more likely more engaged in social media.

In general, there are two major approaches for customer segmentation:

- **Segmentation based on demographics**

Here the focus is on analyzing who the customers are, by looking at aspects like age, gender, or income.

- **Segmentation based on behaviour**

In this view the analysis uses historic data and centers around the question of “What do customers do?”. This can include things like historic buying behavior, responses to past coupons, or, how long they have been a customer, etc.

## Aim of the project

The goal for this assignment is to analyze historic transactional purchase data and evaluate which groups of customers responded best to previous marketing campaigns. The data available for the analysis includes:

- Data about past 10 discount promotions
- Customer details for 17000 existing customers
- Transactional data with 306534 records including offer responses like (offer received, offer viewed, and offer completed) and purchase records for all customers with a focus on transactions that happened on the same time as offer completed.

The outcome of the analysis will outline the characteristics of one or more customer segments which have responded well to the promotional incentives. Here are the specific goals for the initiative:

- Identify clusters of customers that completed more transactions and/or show a significant increase in spending behaviour over the promotional period.
- Identify customers response / preferences to offer by age, gender, income and distribution channels.
- Provide an interpretation of clusters with visual representations of each cluster.
- Provide a clear description of customer characteristics that the company should focus their attention to and explain why.

## Data Exploration

The transactions table includes a mixture of offer events (offer received, viewed, completed) and purchase transaction records.

Our assumption is that a purchase transaction for customer that happens at the same time as an “offer completed” event ties these two events together.

Variables of interest per customer for transactions (customer purchase behaviour):

- What percentage of a customer’s total purchase amount did he spend on offers (offer acceptance rate)?  
 $\% \text{ amount spent on offers} = \text{amount spent on offers} / \text{total amount for all purchases}$
- What percentage of offers received did a customer act on with a purchase?  
 $\% \text{ offer completed} = \text{number of offers completed} / \text{number of offers received}$

### Who are our customers (demographics)?

The following exploratory plots give a high-level view on the overall customer demographics:

#### Customer age

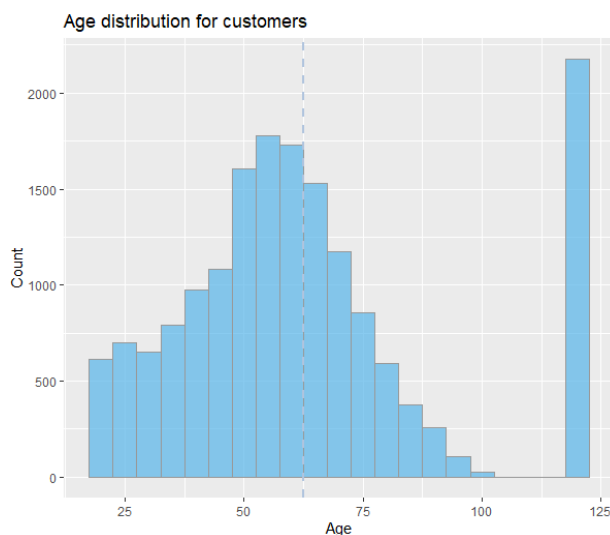


Figure 1 - Age distribution of customers

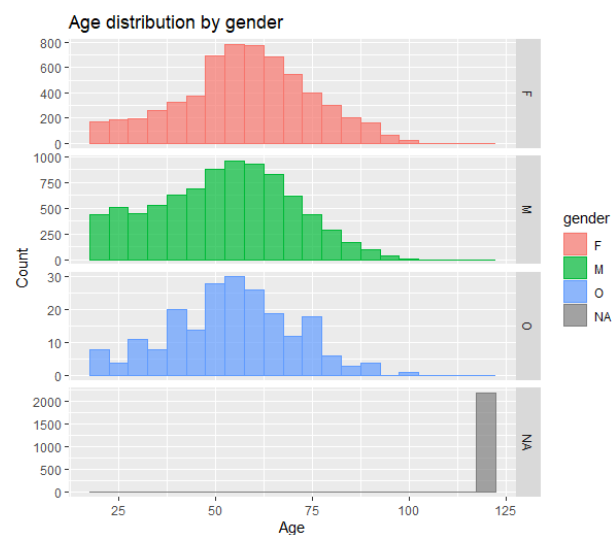


Figure 2 - Customer age distribution by gender

- Most customers are between 40 and 75 year with a mean of 62 years.
- The distribution for each gender category looks very similar with the male tail dropping off faster compared to the female distribution. This matches our expectations due to the shorter life expectancy for men.
- The age of 125 seems to be a default value for customers who have not been willing to provide their data. This will be treated as a separate group as we don't have any demographic details for these customers (except for their membership length).

#### Customer gender and income distribution

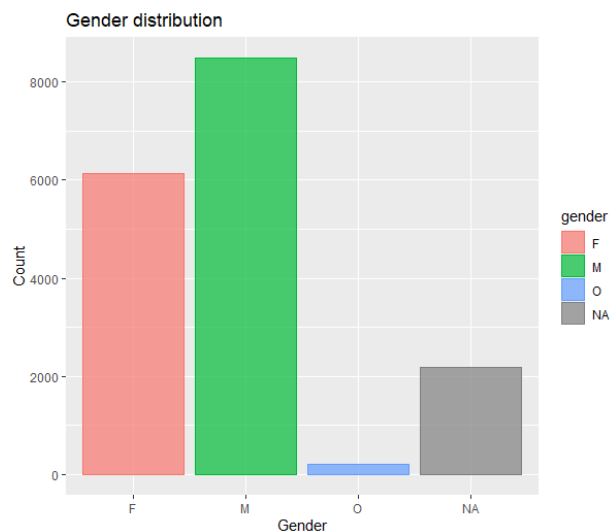


Figure 3 – Gender Distribution

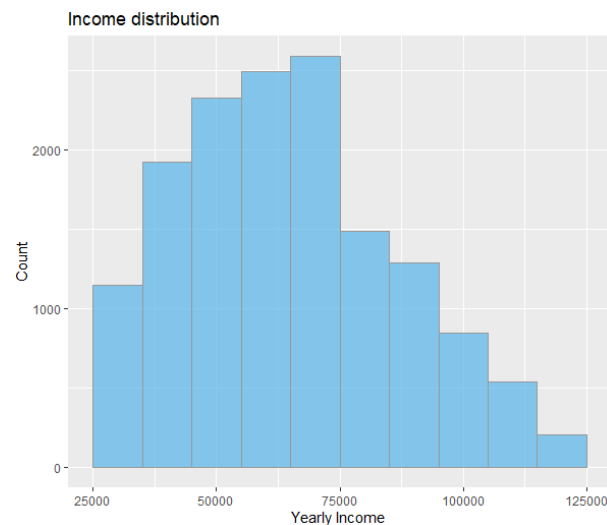


Figure 4 – Customer Income distribution

- We have more male than female customers
- The group for other gender identities is fairly small
- We have about 15% of customers with no information about their gender
- The income distribution is slightly right tailed with a mean of around 65K and a median of 64K pounds

### Membership (customer relationship) length

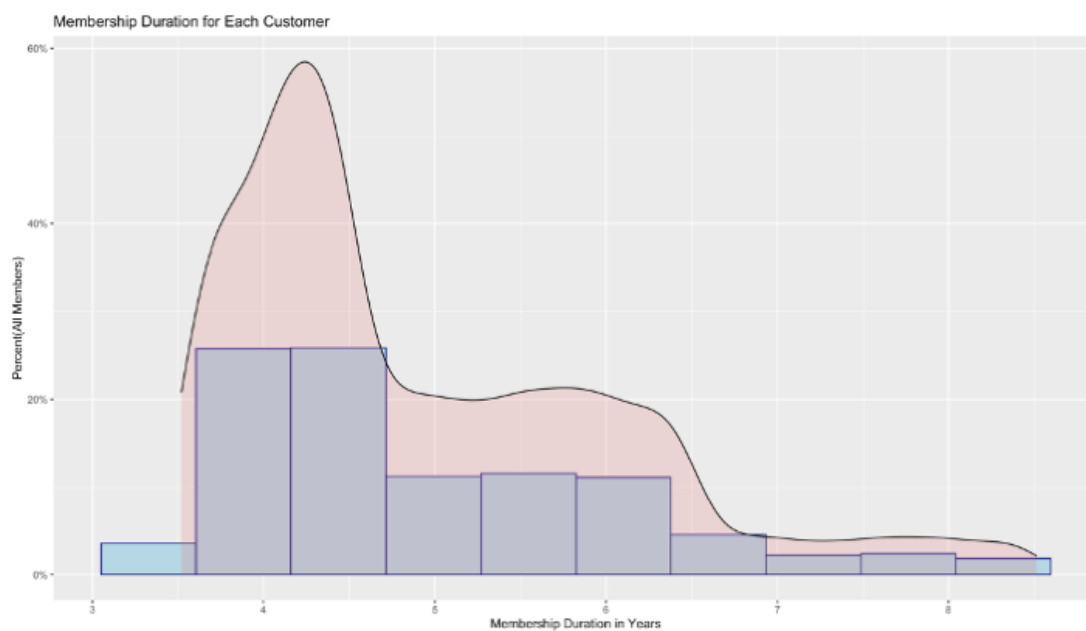


Figure 5 - Distribution of membership length in years since Feb. 1 2022

**Membership Length for each Customer:** Median of 4.5 years

	<4 yrs	4-5 yrs	5-6yrs	6-7yrs	7-8yrs	>8yrs
Counts	3389	6984	3514	2068	695	350
Percentage	19.94%	41.08%	20.67%	12.16%	4.09%	2.06%

- We can see that most individuals are customers for less than 5 years
- Only 6% of the customers are long term customers with more than 7 years

### How are the customers behaving?

The following plots show the offer acceptance rate by different characteristics:

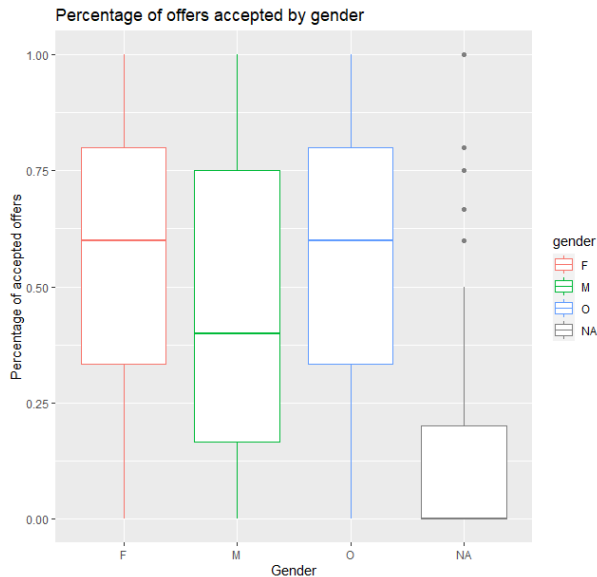


Figure 6 – Offer Acceptance rate by gender

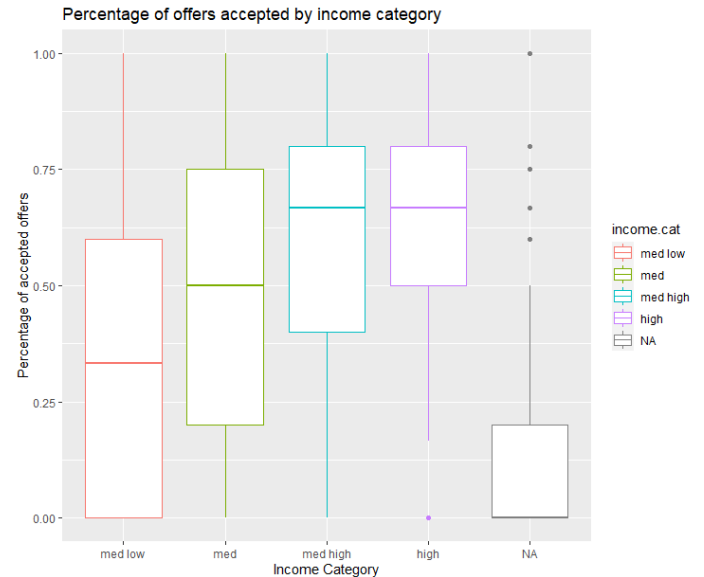


Figure 7 – Offer acceptance rate by income category:  
0 – 25000 (low), 25000-50000 (med low), 50000-75000 (med), 75000-100000 (med-high), 100000-125000 (high)

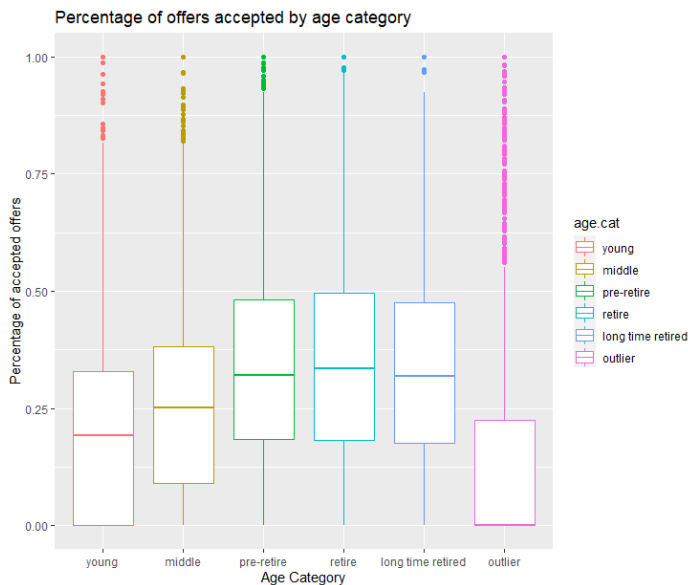


Figure 8 - Offer acceptance rate by age category  
<30 (young), 30-45 (middle), 45-60(pre-retirement), 60-75(retired), 75-100 (long term retired), >100 (outlier)

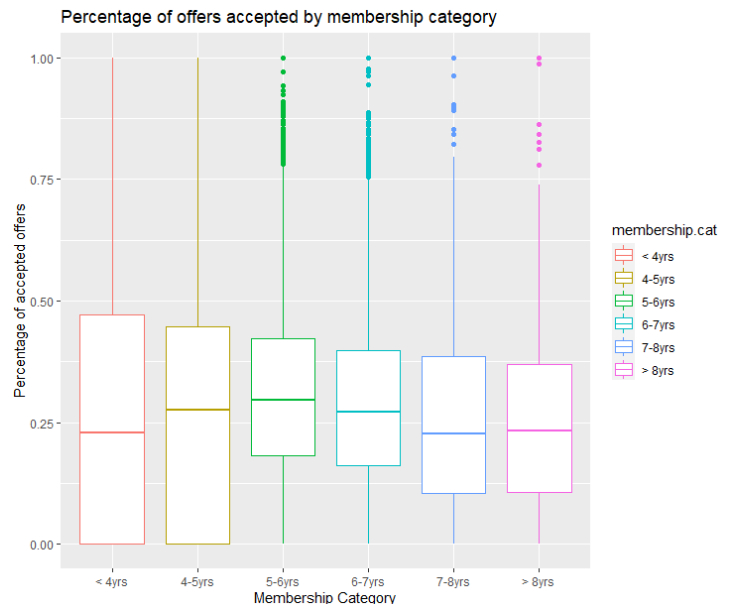


Figure 9- Offer acceptance rate by membership duration

- We can see that men seem to have a lower acceptance rate compared to women and other genders
- The acceptance rate is highest for income ranges medium-high and high (75000 annual income and above)
- It looks like the lower the income range is the lower the acceptance rate

- Age also seems to have an influence on offer acceptance. The highest acceptance rates can be seen for age groups above 45 years. Younger customers have a slightly lower rate. But we can also see that for the younger age groups there seem to be more customers with a rate closer to 1.
- The time of membership does not seem to have a major impact on the acceptance rate.

Does offer category have an influence on acceptance or transactions?

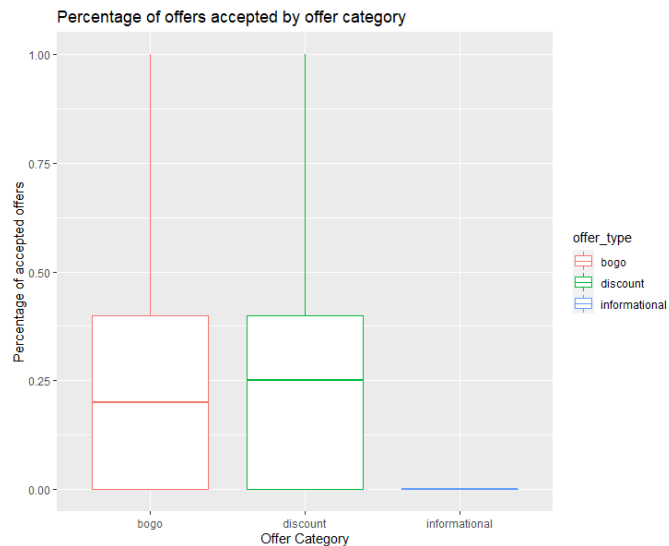


Figure 10- Acceptance rate by offer category

- It looks like informational offers have little to no acceptance
- BOGO and discounts are fairly close in their acceptance rate

#### Among all Transactions Happened Due to Completed Offer:

	<i><b>BOGO</b></i>	<i><b>Discount</b></i>
Counts	15,669	17,910
Percentage	46.66%	53.34%

Finally, here are a few plots exploring potential influences of customer spending behaviour.

#### **Amount Spent on transactions VS Membership duration:**

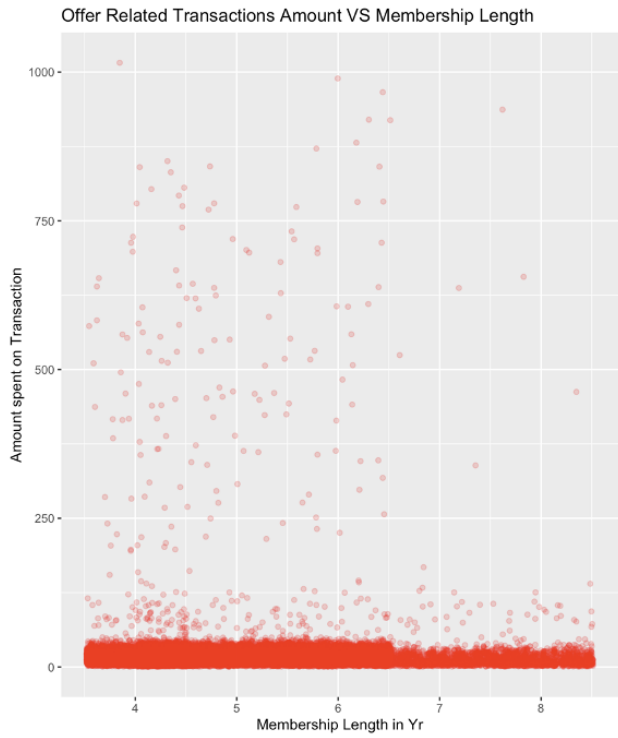


Figure 11 – Total amount spend by membership length

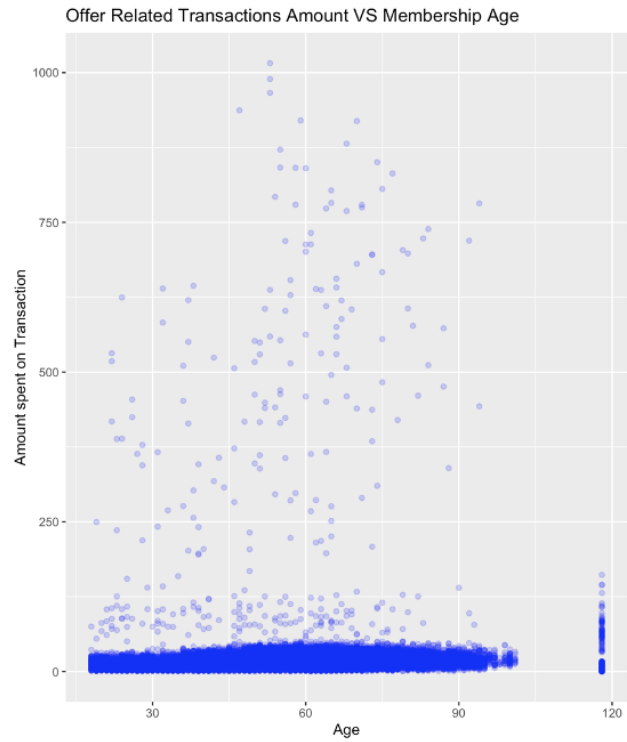


Figure 12 – Total amount spend by customer age

- Customers with four to six years of memberships contributed to more on transactions in terms of dollar value.
- Customers with age 45-75 contribute to more transactions in terms of dollar value