Title : Customer Segmentation Analysis
Author : Pratik Ganguli

Executive summary

Understanding customer behaviour is a cornerstone for developing effective marketing strategies in any retail business. By segmenting the customer base, a supermarket can better tailor its offerings, personalize experiences, and create more efficient marketing campaigns. Market segmentation allows businesses to target specific groups with unique needs and preferences, resulting in improved customer satisfaction, loyalty, and increased sales. This report focuses on identifying key customer segments for Marketplace, a supermarket chain, to optimize its marketing efforts and drive growth.

The dataset used in this analysis contains various demographic and transactional information, including customer income, shopping habits, and product preferences. Univariate and bivariate analyses were employed to explore the relationship between key variables. Univariate analysis focused on individual variables to summarize their distributions, while bivariate analysis was used to investigate correlations, such as the relationship between customer income and spending. Outliers were identified and addressed to ensure the integrity of the data. These analyses help identify meaningful patterns in customer behaviour and set the stage for more advanced segmentation.

To deepen the understanding of customer differences, cluster analysis was conducted, revealing four distinct customer segments. Each cluster was characterized by unique demographic, age and shopping behaviour traits. The first cluster consists of moderate-income families with young children, focused on budget-friendly purchases. The second cluster includes low-income, price-sensitive families primarily seeking essentials. The third cluster represents high-income individuals, drawn to premium products and in-store experiences. Finally, the fourth cluster consists of convenience-driven shoppers who prioritize digital engagement and seamless online and offline shopping experiences. These segments highlight the diversity within Marketplace's customer base, each requiring distinct marketing strategies.

Based on these findings, several targeted strategies are recommended. For moderate-income families, Marketplace should focus on in-store promotions, family-oriented bundles, and discounts. Budget-conscious families should be engaged with loyalty programs and regular promotions on essential products. High-income customers would benefit from exclusive product offerings, personalized services, and a premium shopping experience. For convenience-driven shoppers, an enhanced omnichannel strategy that integrates online and offline shopping experiences, along with personalized digital promotions, is key to fostering loyalty and driving sales growth.

By implementing these strategies, Marketplace can effectively engage its diverse customer base, improve retention, and achieve its sales objectives
Introduction

Understanding customer behaviour is at the heart of any successful retail business, and customer segmentation plays a pivotal role in achieving this. For supermarkets like Marketplace, segmenting customers based on shared characteristics allows for targeted strategies that enhance sales, improve customer satisfaction, and foster long-term loyalty. This report applies statistical methods to identify distinct customer segments and propose actionable strategies tailored to the needs of Marketplace's diverse customer base.

The analysis examines key variables such as consumer composition (age, marital status, income), purchasing behaviour (monthly spending, product preferences), and engagement with campaigns (promotion responses). The dataset includes information from 2,150 customers, capturing income levels (ranging from $1,730 to $162,397) and average monthly spending (around $231.02). These insights will help Marketplace understand how different customer groups interact with its offerings and inform strategies for better engagement.
Customer segmentation divides a broad customer base into smaller, more manageable groups based on similar behaviours or attributes. By identifying patterns within these segments, Marketplace can tailor marketing efforts, optimize product offerings, and design loyalty programs that meet the unique needs of each group.
In a competitive retail environment, these insights are invaluable. This report provides practical recommendations that will help Marketplace address its challenges, capitalize on emerging customer trends, and strengthen its position as a customer-focused supermarket.

Methodology

The analysis utilizes univariate, bivariate, and cluster analysis techniques to segment the customer base and examine relationships between key variables, ensuring a comprehensive understanding of customer behaviour and preferences. Univariate analysis is employed to summarize and explore individual variables, detecting outliers and identifying distribution patterns through histograms and boxplots. This process helps clean the data and ensures that only representative values are considered in further analysis. Bivariate analysis is used to investigate relationships between pairs of variables. Techniques such as the Chi-Square test (to assess relationships between categorical variables), T-test (to compare means between two groups), ANOVA (to evaluate mean differences across multiple groups), and correlation analysis (to measure the strength and direction of relationships between continuous variables) allow for a deeper understanding of variable interactions and statistically significant patterns.

Cluster analysis, specifically two-step clustering, is the core technique of this report. Two-step clustering was chosen for its ability to handle both numerical and categorical data, making it ideal for identifying meaningful customer segments based on demographic and purchasing behaviours. This unsupervised machine learning technique allows for grouping customers without predefined categories, making it suitable for discovering new patterns within complex datasets.

To ensure the integrity of the analysis, data preprocessing will involve detecting and addressing outliers in key variables, such as age, to prevent distortion in segmentation results. Duplicate entries, such as IDs, will be excluded if they do not contribute to meaningful insights. Correlation checks will identify and address redundancy among variables by combining or removing highly correlated ones. A variance check will ensure that selected variables exhibit sufficient variability to support distinct and actionable clustering outcomes. These steps are critical for achieving robust and insightful segmentation results.

For detailed figures and visualizations, refer to *Appendix A*.

Variables Used for Cluster Analysis

The variables chosen for cluster analysis were selected for their significant relevance in capturing customer behaviour and their potential to yield actionable insights. Age was included to uncover generational differences in purchasing preferences. Income serves as a key indicator of purchasing power, allowing for the categorization of customers based on their spending capacity. Family composition, including *KidHousehold* and *TeenHousehold,* provides insights into the influence of household structure on shopping habits, especially for products related to

children or teenagers. Purchasing behaviour, represented by *NStorePurchases* and *NWebPurchases,* helps identify whether customers prefer shopping online or in-store, providing essential insights for crafting effective omnichannel strategies. Finally, *SpendProducts* was included to measure overall customer value by assessing their spending habits. Since this data represents the most recent month, it also provides valuable insights into current spending trends and patterns, reflecting the latest customer behaviour and preferences. This makes it an essential variable for identifying recent shifts in consumer spending, which can inform timely and targeted marketing strategies.
Together, these variables provide a well-rounded understanding of customer preferences and behaviours, ensuring that the clustering process generates meaningful segments that can be targeted with tailored marketing strategies.

Analysis Results

The analysis of Marketplace's customer base provided valuable insights into customer demographics, shopping habits, and spending behaviour. The univariate analysis revealed that the average customer is 51.07 years old, with an average annual income of $52,043, and spends about $230.95 each month on products. Customers made, on average, 4.08 online purchases and 5.81 in-store purchases. This indicates that while many customers enjoy the convenience of shopping online, a significant portion still prefers to shop in physical stores, reflecting a mixed shopping behaviour.

In terms of household composition, 42% of customers have young children, and 48% have teenagers. This suggests that family-related factors, such as buying products for children or teenagers, could influence shopping decisions and preferences, making family-oriented marketing campaigns potentially more effective for this group. The bivariate analysis explored relationships between key variables, revealing significant patterns. A T-test comparing customers who accepted marketing campaigns with those who did not showed a strong link between campaign acceptance and higher spending. Customers who responded to campaigns had a tendency to spend more, with a p-value of less than 0.001, indicating that targeted marketing strategies can drive higher customer spending.

Additionally, a correlation analysis between income and spending showed a moderate but significant relationship ($r = 0.71$, $p < 0.001$), suggesting that customers with higher incomes are more likely to make larger purchases. This finding highlights the importance of focusing on affluent customer segments for targeted marketing and promotions, as they are more likely to spend more overall.

For detailed figures and visualizations, refer to *Appendix B*.

The analysis of Marketplace's customer base, utilizing a two-step clustering methodology, identified four distinct customer segments, each characterized by a unique blend of demographic factors, income levels, and purchasing behaviours. With a silhouette score of 0.5, this segmentation provides valuable insights into how family structure and income influence consumer spending patterns. Active profiling was employed to interpret and label these clusters based on their defining characteristics, enabling a deeper understanding of the diverse customer groups. The following section provides a detailed breakdown of each cluster, outlining their key traits and behaviours, which are essential for developing targeted strategies tailored to each group's specific needs and preferences.

For detailed figures and visualizations, refer to *Appendix D*.

Moderate-Income Families with Young Children (Cluster 1)

This segment comprises families with a moderate average income of $44,429.57, all of whom have young children and teenagers in their households. These families demonstrate steady spending habits, with an average monthly expenditure of $63.33. Their shopping behaviour shows a slight preference for in-store purchases (4.09) compared to online shopping (3.74), possibly reflecting a preference for physically evaluating products and seeking cost-saving opportunities in-store. Despite their moderate income, they have the lowest spending on products among all segments in the past month. The name of this cluster emphasizes their stable financial situation, with the presence of both young children and teenagers significantly influencing their purchasing decisions as they focus on meeting the essential needs of their family.

Low-Income Budget-Conscious Families (Cluster 2)

Customers in this group are characterized by their lower average income of $32,309.59. The majority (82.1%) have small kids, which likely adds financial pressure, leading to cautious and frugal spending habits. Their average monthly expenditure of $70.01 is the second-lowest across the clusters, and they have the fewest store purchases (3.58) and web purchases (2.58). The name highlights their focus on stretching their budgets and prioritizing essential purchases for their families.

High-Income Product Enthusiasts (Cluster 3)

This cluster consists of affluent customers with a high average income of $74,436.40, none of whom have children. They exhibit significantly higher spending behaviour, averaging $594.40 per month, and have the highest frequency of in-store purchases (8.15). This group likely values premium products and experiences, seeking quality over quantity. The name underscores their strong purchasing power and keen interest in indulging in diverse or high-end products.

Digitally-Driven Convenience Shoppers (Cluster 4)

This segment consists of customers with a moderate average income of $58,072.59, typically living in households with teenagers rather than young children. They exhibit a strong preference for digital shopping, with the highest average monthly web purchases at 5.53, indicating a high level of engagement with online platforms. At the same time, they maintain a balanced shopping approach, making 7.21 in-store visits on average. Their monthly spending is substantial, at $201.82, highlighting a tendency to seek convenience and variety in their shopping habits. The name "Digitally-Driven Convenience Shoppers" effectively reflects this group's tech-savvy nature and focus on balancing the ease of online shopping with the benefits of in-store experiences.

For detailed figures and visualizations, refer to *Appendix C*.

Passive analysis was conducted to examine the characteristics of the clusters in relation to variables not included in the clustering process, providing an additional layer of validation for the identified segments. In this case, the variable *AcceptedCampaigns* was analyzed using a Chi-Square test, as outlined in the Appendix. The results indicate that Cluster 3, labeled as High-Income Product Enthusiasts, demonstrated a significantly higher acceptance rate of marketing campaigns compared to the other segments. Furthermore, this group also exhibited higher monthly expenditure, reinforcing the correlation between campaign acceptance and increased spending. This finding aligns with insights from the bivariate analysis, where a strong relationship between *SpendProduct* and campaign acceptance was observed. These results

suggest that customers who engage with campaigns are likely to spend more, highlighting the importance of targeted marketing strategies in driving higher customer spend.

For detailed figures and visualizations, refer to *Appendix E*.

Segment-specific strategies

The segmentation analysis has identified four distinct customer groups, each defined by unique characteristics, preferences, and shopping behaviours. These insights enable Marketplace to design tailored strategies that resonate with the specific needs of each segment. By implementing strategies such as family-focused promotions, loyalty programs, premium memberships, and digital-first initiatives, Marketplace can enhance customer engagement, improve satisfaction, and drive revenue growth. This data-driven approach ensures optimal resource allocation while addressing the diverse preferences of the customer base, minimizing churn, and fostering long-term loyalty. Through these targeted efforts, Marketplace positions itself as a leader in customer-centric retail innovation.

Moderate-Income Families with Young Children (Cluster 1)

For this segment, consisting of families managing moderate incomes and a balanced preference for both in-store and online shopping, the recommended strategy centers on Budget-Conscious Family Bundles and Promotions. This approach emphasizes value by offering discounts on essential items such as groceries, household goods, and baby care products, addressing their practical and family-oriented needs. Exclusive bundles that combine frequently purchased items into cost-effective packages will further attract this group. Both digital and in-store campaigns should be utilized to maximize reach—digital platforms like email newsletters and social media ads can promote convenience and accessibility, while in-store promotions can immediately capture attention with visible discounts and dedicated family shopping areas. These strategies resonate with this group's focus on affordability and practicality, helping to enhance loyalty, increase engagement across shopping channels, and build a strong, lasting connection with these customers.

Low-Income Budget-Conscious Families (Cluster 2)

For this segment, characterized by limited incomes and a strong focus on essential goods, the recommended approach is centered on Discounted Essentials and Loyalty Rewards. By launching targeted discount campaigns on essential items such as groceries, cleaning supplies, and children's products, Marketplace can alleviate the financial strain on these families. Additionally, introducing a simple and easy-to-use loyalty program will provide rewards for frequent purchases of everyday items, incentivizing repeat shopping and enhancing customer loyalty. Digital marketing strategies, such as email and SMS campaigns, can be leveraged to notify customers of limited-time discounts and loyalty rewards, ensuring maximum reach and engagement. These initiatives address their budget-conscious mindset while fostering long-term relationships through accessible and practical solutions.

High-Income Product Enthusiasts (Cluster 3)

To appeal to this affluent segment, the recommended strategy revolves around Premium Product Offerings and Exclusive Experiences. Marketplace can introduce curated premium product collections, limited-edition releases, and VIP experiences, such as private shopping

events or early access to sales. Personalization will also be key—tailored recommendations based on purchase history and exclusive online platforms for high-end goods will enhance the shopping experience. These approaches cater to their preference for quality and exclusivity, aligning with their high spending and frequent in-store visits. By offering premium memberships with perks like concierge shopping services or personalized consultations, Marketplace can strengthen brand loyalty and position itself as a preferred destination for upscale customers.

Digitally-Driven Convenience Shoppers (Cluster 4)

For this tech-savvy group, the best approach is to focus on Seamless Omni-Channel Shopping with Digital Incentives. Marketplace should create an integrated shopping experience that transitions smoothly between online and in-store environments. Initiatives such as exclusive online discounts, loyalty points redeemable across channels, and mobile app features like in-store pickup options can drive engagement. Personalized digital marketing, such as push notifications for flash sales or tailored recommendations, will further appeal to their convenience-driven shopping preferences. These strategies enhance their ability to shop flexibly while reinforcing Marketplace's commitment to ease and efficiency. The dual emphasis on digital platforms and physical stores ensures comprehensive engagement with this segment.

Marketplace's cluster-based segmentation offers valuable insights that enable the development of highly targeted strategies, each designed to meet the distinct needs of its diverse customer segments. Whether through budget-friendly family bundles, premium offerings for high-income shoppers, or seamless omni-channel experiences for digitally engaged customers, each strategy is crafted to maximize relevance and engagement. By aligning these tailored approaches with the preferences of specific segments, Marketplace can drive greater customer satisfaction, foster brand loyalty, and achieve long-term growth. This customer-centric approach will strengthen the company's position as an industry leader in delivering exceptional shopping experiences.

Conclusion

The two-step cluster analysis conducted on Marketplace's customer base has successfully segmented customers into four distinct groups: Moderate-Income Families with Young Children, Low-Income Budget-Conscious Families, High-Income Product Enthusiasts, and Digitally-Driven Convenience Shoppers. Each segment reveals specific characteristics in terms of income, family structure, shopping behaviour, and product preferences, offering valuable insights into how different groups interact with the supermarket's offerings. By identifying these segments, the analysis highlights the importance of understanding diverse customer needs to create targeted, effective strategies.

The insights from the segmentation enable Marketplace to develop highly focused, personalized marketing strategies for each cluster. For Moderate-Income Families with Young Children, emphasizing budget-friendly family bundles and in-store promotions would address their need for affordability while enhancing customer loyalty. The Low-Income Budget-Conscious Families segment can be engaged through discounted essentials and simple loyalty rewards, offering them cost-effective solutions that encourage repeat business. High-Income Product Enthusiasts can be attracted with premium product lines, exclusive shopping experiences, and personalized services, aligning with their preference for quality and exclusivity. Finally, Digitally-Driven Convenience Shoppers will benefit from a seamless omnichannel shopping experience, supported by digital incentives and personalized online offers.

By implementing these segment-specific strategies, Marketplace can significantly enhance customer satisfaction, increase spending, and foster long-term loyalty. This customer-centric approach not only optimizes marketing efforts and product offerings but also strengthens Marketplace's competitive edge in the retail market, positioning it as a brand that deeply understands and meets the needs of its diverse customer base.

Appendix

Appendix A:

This appendix contains all the univariant analysis evidence and graphs.

Statistics

		Age	Income	NWebPurchases	NStorePurchases	SpendProducts
N	Valid	2150	2150	2150	2150	2150
	Missing	0	0	0	0	0
Mean		51.17	52054.79	4.08	5.81	231.02
Minimum		24	1730	0	0	1
Maximum		127	162397	27	13	1727

Figure 1.1: Univariate analysis for metric variables

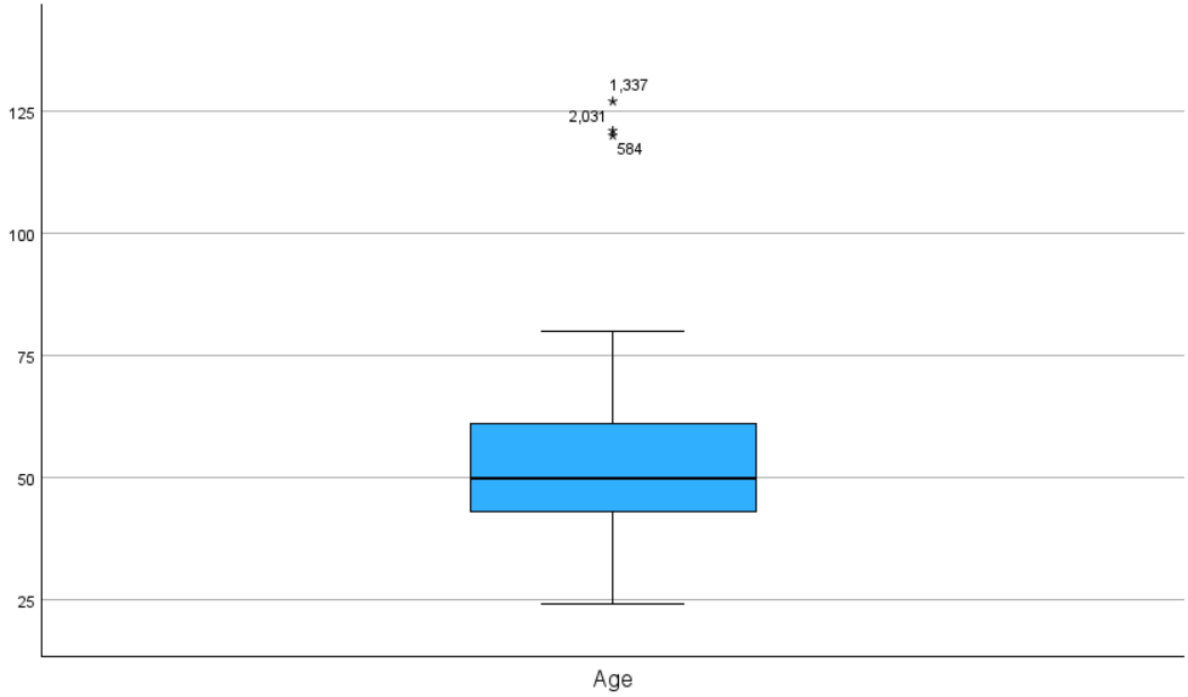

Figure 1.2 Boxplot for Age outliers

Statistics

		KidsHousehold	TeensHousehold	AcceptedCampaign	Complaint
N	Valid	2150	2150	2150	2150
	Missing	0	0	0	0
Mean		.42	.48	.21	.01
Median		.00	.00	.00	.00
Mode		0	0	0	0

*Figure **Error! No text of specified style in document.**3 Univariate analysis for non-metric data*

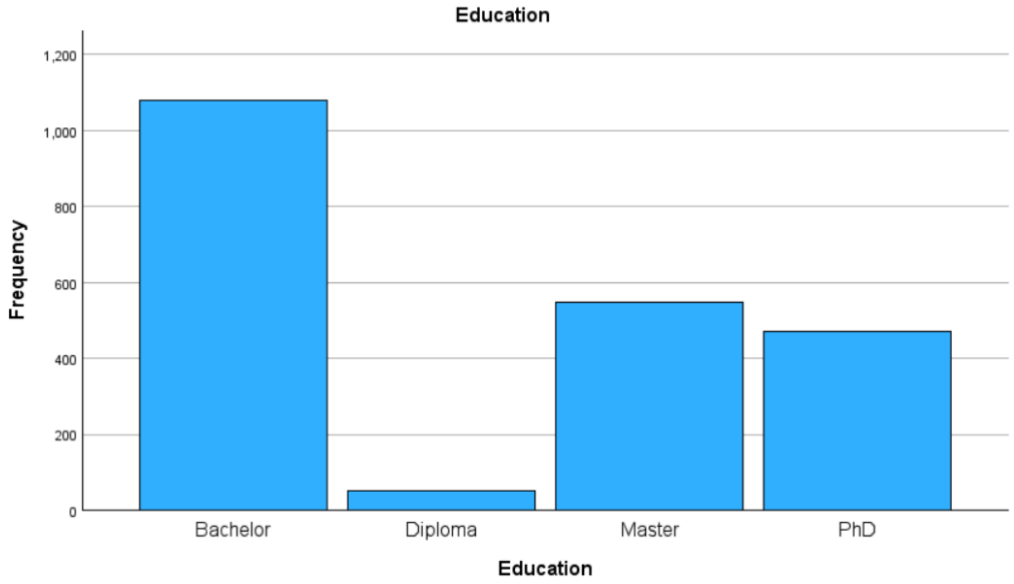

Figure 1.4 Bar Graph for Education

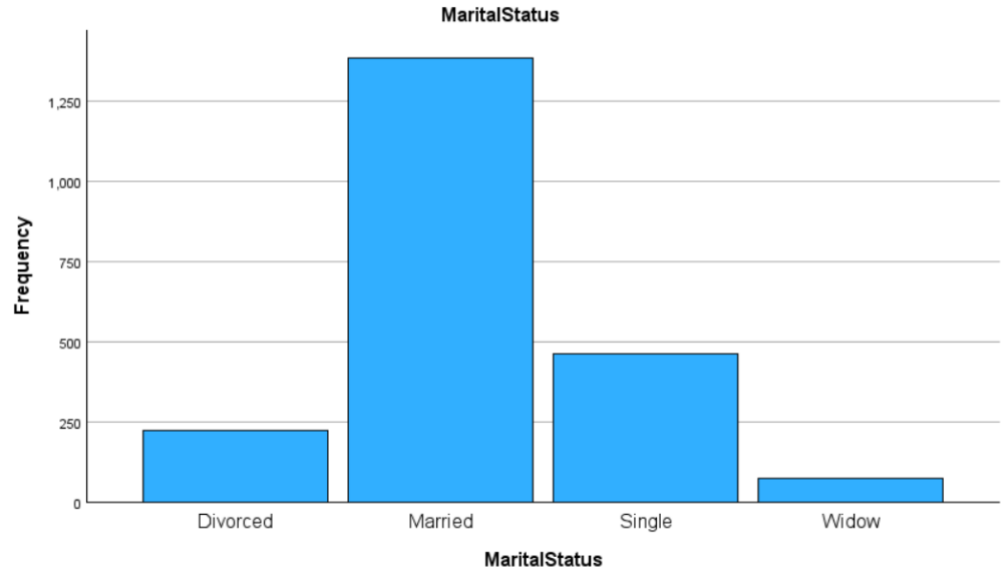

Figure 1.5 Bar Graph for Marital Status

Appendix B:

This appendix contains all the Bivariate analysis evidence and graphs

T-Test

Group Statistics

	AcceptedCampaign	N	Mean	Std. Deviation	Std. Error Mean
SpendProducts	0	1706	192.50	254.422	6.160
	1	441	379.69	340.284	16.204

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
SpendProducts	Equal variances assumed	117.837	<.001	-12.777	2145	<.001	<.001	-187.184	14.650	-215.913	-158.454
	Equal variances not assumed			-10.798	573.264	<.001	<.001	-187.184	17.335	-221.232	-153.135

Figure 2.1 T-test for SpendProduct and AcceptedCampaign

Correlations

Correlations

		Income	SpendProducts
Income	Pearson Correlation	1	.713**
	Sig. (2-tailed)		<.001
	N	2147	2147
SpendProducts	Pearson Correlation	.713**	1
	Sig. (2-tailed)	<.001	
	N	2147	2147

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 2.2 Correlation test for SpendProduct and Income

Graph

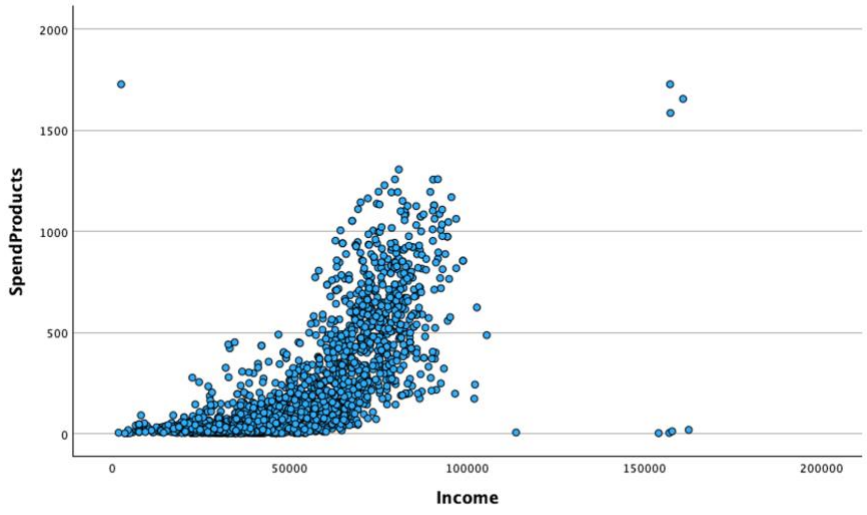

Figure 2.3 Scatter plot for spendproduct and Income

Appendix C:

This appendix contains all the Two-Step cluster analysis evidence

TwoStep Cluster

Model Summary

Algorithm	TwoStep
Inputs	7
Clusters	4

Cluster Quality

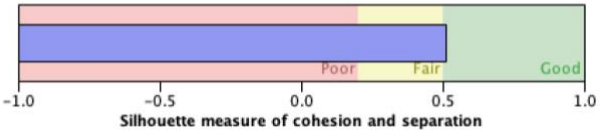

Figure 3.1 Silhouette score for two-step cluster

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	1	2	3	4
Label				
Description				
Size	18.9% (405)	28.4% (610)	23.6% (507)	29.1% (625)
Inputs	Income 44,429.57	Income 32,309.59	Income 74,436.40	Income 58,072.59
	KidsHousehold 1 (100.0%)	KidsHousehold 1 (82.1%)	KidsHousehold 0 (100.0%)	KidsHousehold 0 (100.0%)
	SpendProducts 63.33	SpendProducts 70.01	SpendProducts 594.40	SpendProducts 201.82
	TeensHousehold 1 (100.0%)	TeensHousehold 0 (100.0%)	TeensHousehold 0 (100.0%)	TeensHousehold 1 (100.0%)
	NStorePurchases 4.09	NStorePurchases 3.58	NStorePurchases 8.15	NStorePurchases 7.21
	Age 54.57	Age 42.38	Age 52.47	Age 56.14
	NWebPurchases 2.99	NWebPurchases 2.58	NWebPurchases 4.96	NWebPurchases 5.53

Figure 3.2 Cluster information

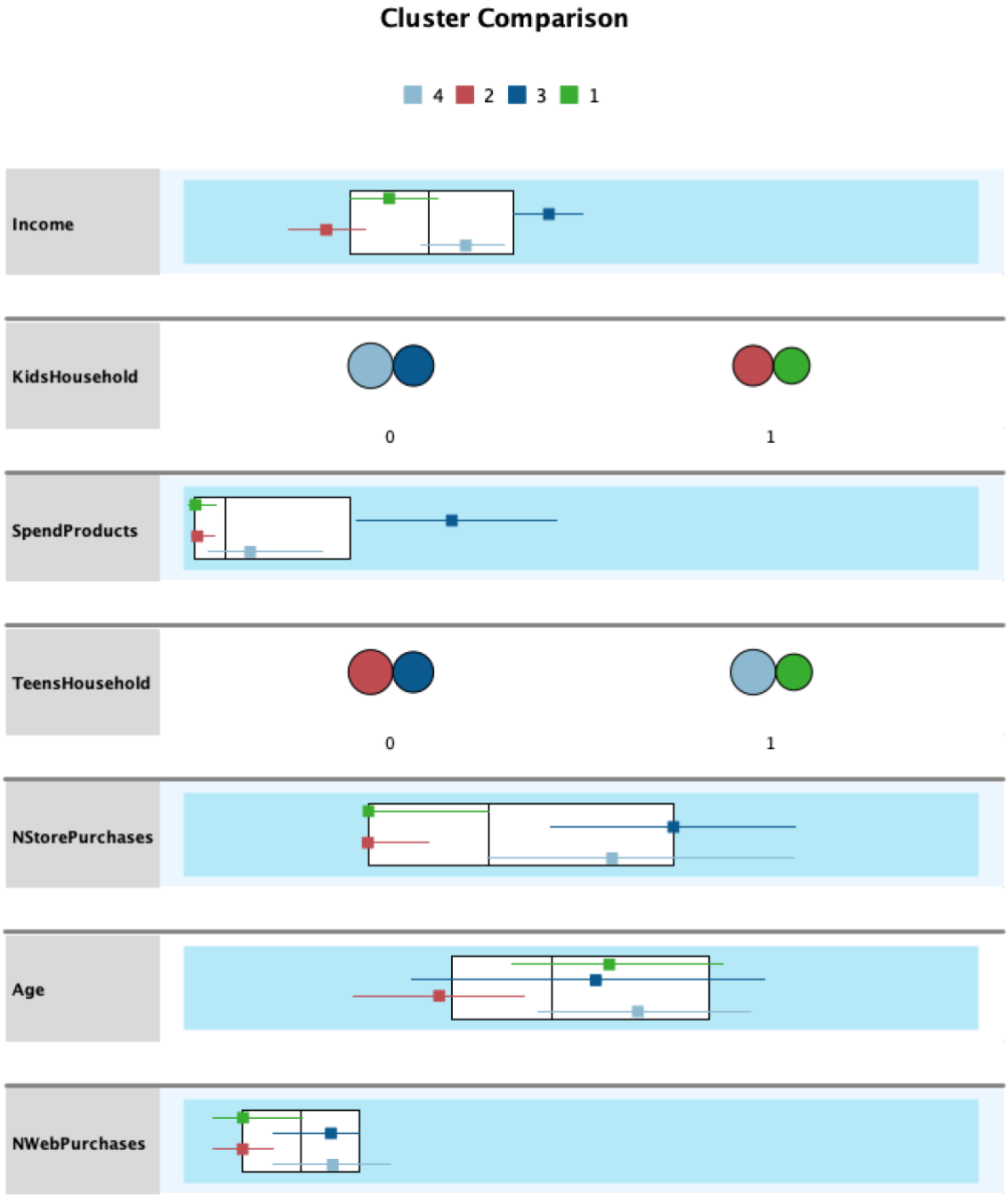

Figure 3.3 Comparison of Clusters

Appendix D:

This appendix contains all the Active profiling cluster analysis evidence

Descriptives

						95% Confidence Interval for Mean			
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Income	1	505	74623.44	14793.087	658.284	73330.13	75916.76	22507	160803
	2	612	32292.92	14815.019	598.862	31116.85	33469.00	1730	157733
	3	625	58072.59	13526.345	541.054	57010.09	59135.10	4428	157243
	4	405	44429.57	15070.063	748.837	42957.47	45901.68	4023	162397
	Total	2147	52043.53	21507.963	464.177	51133.25	52953.81	1730	162397
Age	1	505	52.41	14.257	.634	51.16	53.66	25	80
	2	612	42.46	8.414	.340	41.80	43.13	24	74
	3	625	56.14	8.931	.357	55.44	56.84	38	74
	4	405	54.57	8.758	.435	53.72	55.43	39	74
	Total	2147	51.07	11.700	.252	50.57	51.56	24	80
NWebPurchases	1	505	4.96	2.340	.104	4.76	5.17	0	27
	2	612	2.59	1.949	.079	2.43	2.74	0	11
	3	625	5.53	2.904	.116	5.30	5.76	0	25
	4	405	2.99	2.328	.116	2.76	3.21	0	11
	Total	2147	4.08	2.738	.059	3.96	4.19	0	27
NStorePurchases	1	505	8.16	2.884	.128	7.91	8.41	0	13
	2	612	3.58	1.989	.080	3.42	3.74	0	13
	3	625	7.21	3.086	.123	6.97	7.46	0	13
	4	405	4.09	2.154	.107	3.88	4.30	0	12
	Total	2147	5.81	3.251	.070	5.67	5.95	0	13

Figure 4.1 ANOVA test: Descriptive values for all the metric variables used for clustering

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Income	Based on Mean	.564	3	2143	.639
	Based on Median	.515	3	2143	.672
	Based on Median and with adjusted df	.515	3	2094.369	.672
	Based on trimmed mean	.513	3	2143	.673
Age	Based on Mean	114.625	3	2143	<.001
	Based on Median	113.921	3	2143	<.001
	Based on Median and with adjusted df	113.921	3	1840.343	<.001
	Based on trimmed mean	114.631		43	<.001
NWebPurchases	Based on Mean	31.181		43	<.001
	Based on Median	32.966		143	<.001
	Based on Median and with adjusted df	32.966	3	2074.649	<.001
	Based on trimmed mean	32.782	3	2143	<.001
NStorePurchases	Based on Mean	106.202	3	2143	<.001
	Based on Median	108.000	3	2143	<.001
	Based on Median and with adjusted df	108.000	3	2092.019	<.001
	Based on trimmed mean	113.978	3	2143	<.001

Double-click to activate

Figure 4.2 ANOVA test: Homogeneity test values for all the metric variables used for clusterin

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Income	Between Groups	5.424E+11	3	1.808E+11	860.411	<.001
	Within Groups	4.503E+11	2143	210134251.10		
	Total	9.927E+11	2146			
Age	Between Groups	67272.727	3	22424.242	212.191	<.001
	Within Groups	226471.070	2143	105.679		
	Total	293743.798	2146			
NWebPurchases	Between Groups	3555.647	3	1185.216	202.687	<.001
	Within Groups	12531.207	2143	5.848		
	Total	16086.854	2146			
NStorePurchases	Between Groups	8257.542	3	2752.514	408.895	<.001
	Within Groups	14425.813	2143	6.732		
	Total	22683.355	2146			

Figure 4.3 ANOVA test: Anova test values for all the metric variables used for clustering

Income

			Subset for alpha = 0.05			
	TwoStep Cluster Number	N	1	2	3	4
Tukey B[a,b]	2	612	32292.92			
	4	405		44429.57		
	3	625			58072.59	
	1	505				74623.44

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 520.608.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Figure 4.4 Tukeys B PostHoc test for Income

Multiple Comparisons

Dependent Variable		(I) TwoStep Cluster Number	(J) TwoStep Cluster Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Income	Tamhane	1	2	42330.520*	889.929	<.001	39984.77	44676.27
			3	16550.852*	852.101	<.001	14304.64	18797.07
			4	30193.871*	997.043	<.001	27564.55	32823.19
		2	1	-42330.520*	889.929	<.001	-44676.27	-39984.77
			3	-25779.669*	807.078	<.001	-27906.57	-23652.76
			4	-12136.650*	958.850	<.001	-14665.28	-9608.02
		3	1	-16550.852*	852.101	<.001	-18797.07	-14304.64
			2	25779.669*	807.078	<.001	23652.76	27906.57
			4	13643.019*	923.849	<.001	11206.27	16079.77
		4	1	-30193.871*	997.043	<.001	-32823.19	-27564.55
			2	12136.650*	958.850	<.001	9608.02	14665.28
			3	-13643.019*	923.849	<.001	-16079.77	-11206.27
	Dunnett T3	1	2	42330.520*	889.929	<.001	39984.95	44676.09
			3	16550.852*	852.101	<.001	14304.82	18796.88
			4	30193.871*	997.043	<.001	27564.81	32822.94
		2	1	-42330.520*	889.929	<.001	-44676.09	-39984.95
			3	-25779.669*	807.078	<.001	-27906.43	-23652.91
			4	-12136.650*	958.850	<.001	-14665.03	-9608.27
		3	1	-16550.852*	852.101	<.001	-18796.88	-14304.82
			2	25779.669*	807.078	<.001	23652.91	27906.43
			4	13643.019*	923.849	<.001	11206.53	16079.51
		4	1	-30193.871*	997.043	<.001	-32822.94	-27564.81
			2	12136.650*	958.850	<.001	9608.27	14665.03
			3	-13643.019*	923.849	<.001	-16079.51	-11206.53
Age	Tamhane	1	2	9.946*	.720	<.001	8.05	11.84
			3	-3.731*	.728	<.001	-5.65	-1.81
			4	-2.163*	.769	.030	-4.19	-.13
		2	1	-9.946*	.720	<.001	-11.84	-8.05
			3	-13.677*	.493	<.001	-14.98	-12.38
			4	-12.109*	.552	<.001	-13.57	-10.65
		3	1	3.731*	.728	<.001	1.81	5.65
			2	13.677*	.493	<.001	12.38	14.98
			4	1.568*	.563	.032	.08	3.05
		4	1	2.163*	.769	.030	.13	4.19
			2	12.109*	.552	<.001	10.65	13.57
			3	-1.568*	.563	.032	-3.05	-.08
	Dunnett T3	1	2	9.946*	.720	<.001	8.05	11.84
			3	-3.731*	.728	<.001	-5.65	-1.81
			4	-2.163*	.769	.030	-4.19	-.13
		2	1	-9.946*	.720	<.001	-11.84	-8.05
			3	-13.677*	.493	<.001	-14.98	-12.38
			4	-12.109*	.552	<.001	-13.57	-10.65
		3	1	3.731*	.728	<.001	1.81	5.65
			2	13.677*	.493	<.001	12.38	14.98
			4	1.568*	.563	.032	.08	3.05
		4	1	2.163*	.769	.030	.13	4.19
			2	12.109*	.552	<.001	10.65	13.57
			3	-1.568*	.563	.032	-3.05	-.08
NWebPurchases	Tamhane	1	2	2.376*	.131	<.001	2.03	2.72
			3	-.565*	.156	.002	-.98	-.15
			4	1.979*	.156	<.001	1.57	2.39
		2	1	-2.376*	.131	<.001	-2.72	-2.03
			3	-2.941*	.140	<.001	-3.31	-2.57
			4	-.397*	.140	.028	-.77	-.03
		3	1	.565*	.156	.002	.15	.98
			2	2.941*	.140	<.001	2.57	3.31
			4	2.544*	.164	<.001	2.11	2.98
		4	1	-1.979*	.156	<.001	-2.39	-1.57
			2	.397*	.140	.028	.03	.77
			3	-2.544*	.164	<.001	-2.98	-2.11
	Dunnett T3	1	2	2.376*	.131	<.001	2.03	2.72
			3	-.565*	.156	.002	-.98	-.15
			4	1.979*	.156	<.001	1.57	2.39
		2	1	-2.376*	.131	<.001	-2.72	-2.03
			3	-2.941*	.140	<.001	-3.31	-2.57
			4	-.397*	.140	.028	-.77	-.03
		3	1	.565*	.156	.002	.15	.98
			2	2.941*	.140	<.001	2.57	3.31
			4	2.544*	.164	<.001	2.11	2.98
		4	1	-1.979*	.156	<.001	-2.39	-1.57
			2	.397*	.140	.028	.03	.77
			3	-2.544*	.164	<.001	-2.98	-2.11
NStorePurchases	Tamhane	1	2	4.579*	.151	<.001	4.18	4.98
			3	.946*	.178	<.001	.48	1.42
			4	4.069*	.167	<.001	3.63	4.51
		2	1	-4.579*	.151	<.001	-4.98	-4.18
			3	-3.633*	.147	<.001	-4.02	-3.24
			4	-.510*	.134	<.001	-.86	-.16
		3	1	-.946*	.178	<.001	-1.42	-.48
			2	3.633*	.147	<.001	3.24	4.02
			4	3.123*	.163	<.001	2.69	3.55
		4	1	-4.069*	.167	<.001	-4.51	-3.63
			2	.510*	.134	<.001	.16	.86
			3	-3.123*	.163	<.001	-3.55	-2.69
	Dunnett T3	1	2	4.579*	.151	<.001	4.18	4.98
			3	.946*	.178	<.001	.48	1.42
			4	4.069*	.167	<.001	3.63	4.51
		2	1	-4.579*	.151	<.001	-4.98	-4.18
			3	-3.633*	.147	<.001	-4.02	-3.24
			4	-.510*	.134	<.001	-.86	-.16
		3	1	-.946*	.178	<.001	-1.42	-.48
			2	3.633*	.147	<.001	3.24	4.02
			4	3.123*	.163	<.001	2.69	3.55
		4	1	-4.069*	.167	<.001	-4.51	-3.63
			2	.510*	.134	<.001	.16	.86
			3	-3.123*	.163	<.001	-3.55	-2.69

*. The mean difference is significant at the 0.05 level.

Figure 4.5 Tamhane PostHoc test for all metric variables

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
KidsHousehold * TeensHousehold	2147	100.0%	0	0.0%	2147	100.0%

KidsHousehold * TeensHousehold Crosstabulation

Count

		TeensHousehold		
		0	1	Total
KidsHousehold	0	616	625	1241
	1	501	405	906
Total		1117	1030	2147

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6.723[a]	1	.010		
Continuity Correction[b]	6.498	1	.011		
Likelihood Ratio	6.730	1	.009		
Fisher's Exact Test				.010	.005
Linear-by-Linear Association	6.720	1	.010		
N of Valid Cases	2147				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 434.64.

b. Computed only for a 2x2 table

Figure 4.6 Chi-Square test for KidHousehold and TeenHousehold

Appendix E:

This appendix contains all the Passive profiling cluster analysis evidence and graphs

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
AcceptedCampaign * TwoStep Cluster Number	2147	100.0%	0	0.0%	2147	100.0%

AcceptedCampaign * TwoStep Cluster Number Crosstabulation

Count

		TwoStep Cluster Number				
		1	2	3	4	Total
AcceptedCampaign	0	291	547	504	364	1706
	1	214	65	121	41	441
Total		505	612	625	405	2147

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	211.885[a]	3	<.001
Likelihood Ratio	198.107	3	<.001
Linear-by-Linear Association	106.490	1	<.001
N of Valid Cases	2147		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 83.19.

Figure 5.1 Chi-Square test for AcceptedCampaign

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Complaint * TwoStep Cluster Number	2147	100.0%	0	0.0%	2147	100.0%

Complaint * TwoStep Cluster Number Crosstabulation

Count

		TwoStep Cluster Number				
		1	2	3	4	Total
Complaint	0	502	605	621	399	2127
	1	3	7	4	6	20
Total		505	612	625	405	2147

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	2.825[a]	3	.419
Likelihood Ratio	2.766	3	.429
Linear-by-Linear Association	.868	1	.352
N of Valid Cases	2147		

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 3.77.

Figure 5.2 Chi-Square test for Complaints

Reference

Grammarly: Free AI Writing Assistance. (2024). Grammarly.com. https://app.grammarly.com/ddocs/2663250977

IBM. (2023). SPSS software. IBM. https://www.ibm.com/spss

Google Scholar. (n.d.). Scholar.google.com. https://scholar.google.com/schhp?hl=en&as_sdt=0

ChatGPT. (2024). ChatGPT. ChatGPT; OpenAI. https://chatgpt.com/

Su-Yeon Kim; Tae-Soo Jung; Eui-Ho Suh; Hyun-Seok Hwang. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. , 31(1), 101–107.