



CUSTOMER PROFILING

A Business Case for Targeted Marketing Strategy - A data-driven understanding of who our customers are, what they buy, and how we should engage them.

Presented by: Group 7
55-Pragya Gupta
71-Tejal Makwana
67-Anuja Shende

BUSINESS CONTEXT & OBJECTIVE

A

To identify which customers respond to marketing efforts and why

B

To understand customer spending patterns, loyalty, and engagement

C

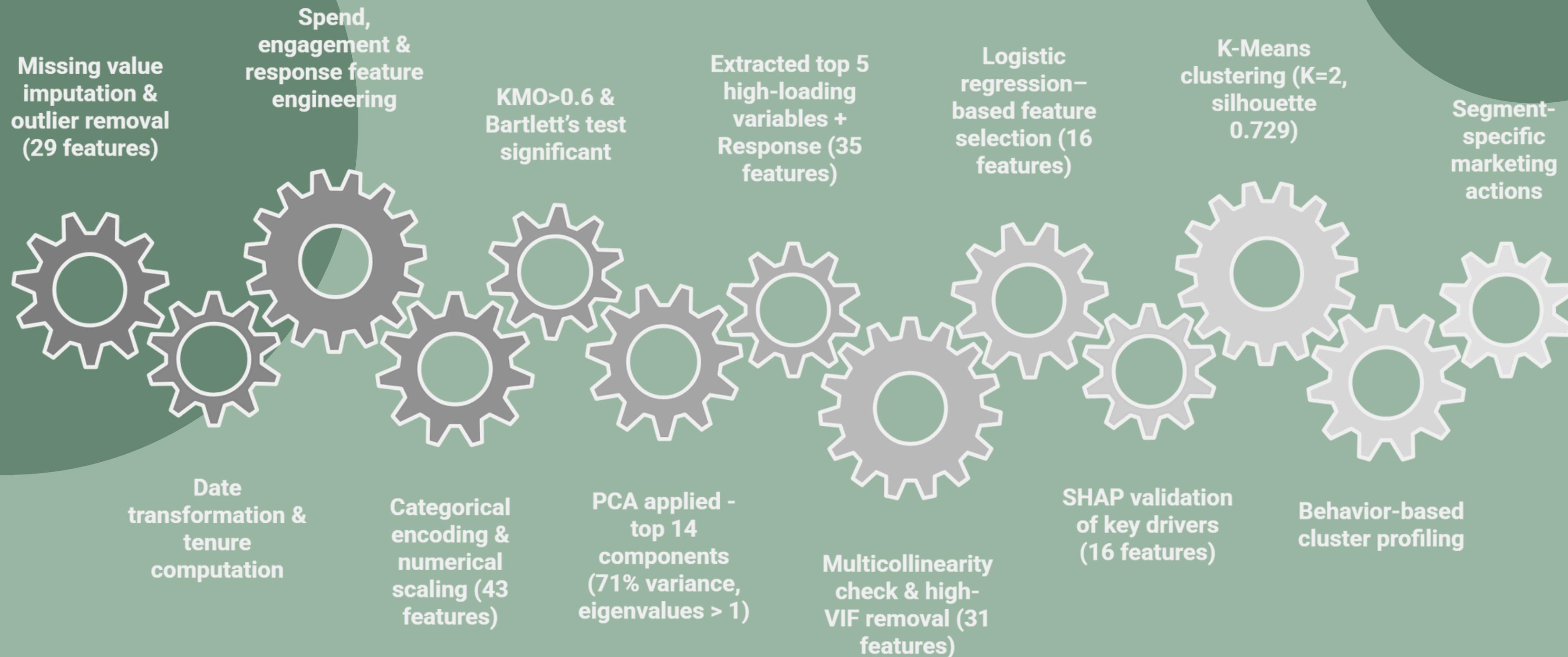
To categorize customers into clear, actionable segments

D

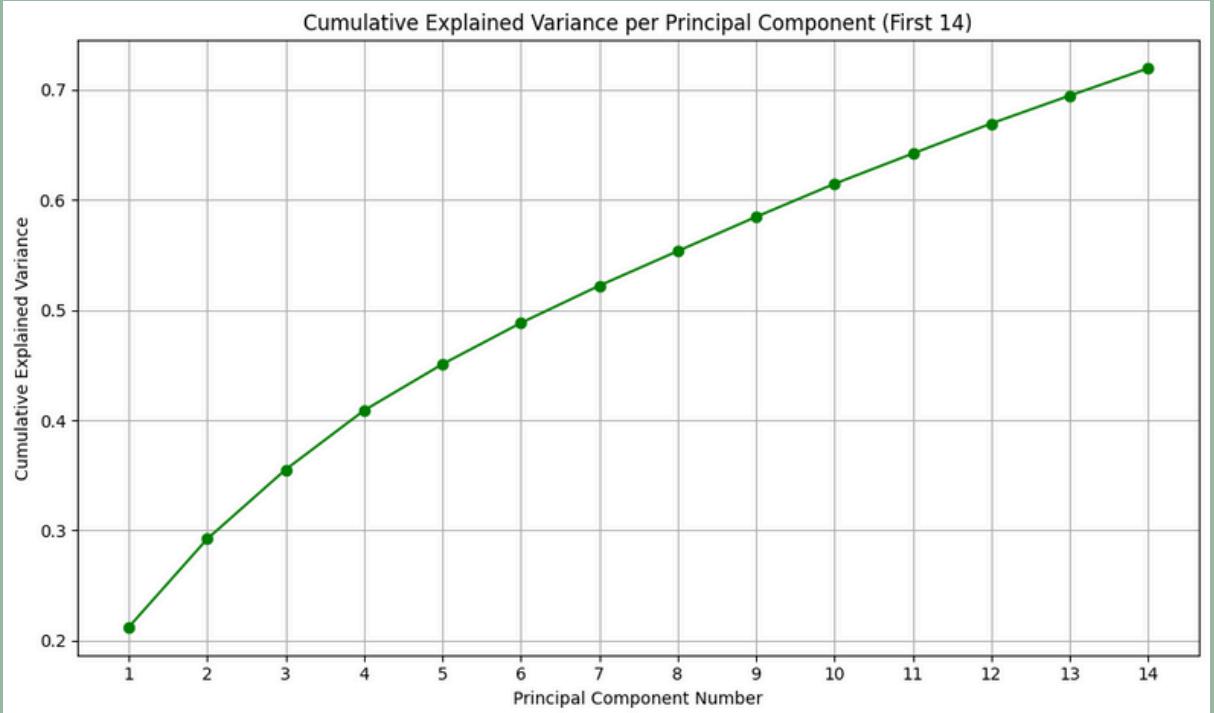
To design strategies that increase conversions, revenue, and long-term loyalty



PROJECT FLOW



ANALYTICAL VALIDATION & MODEL SELECTION RESULTS



PC 14: 0.7196

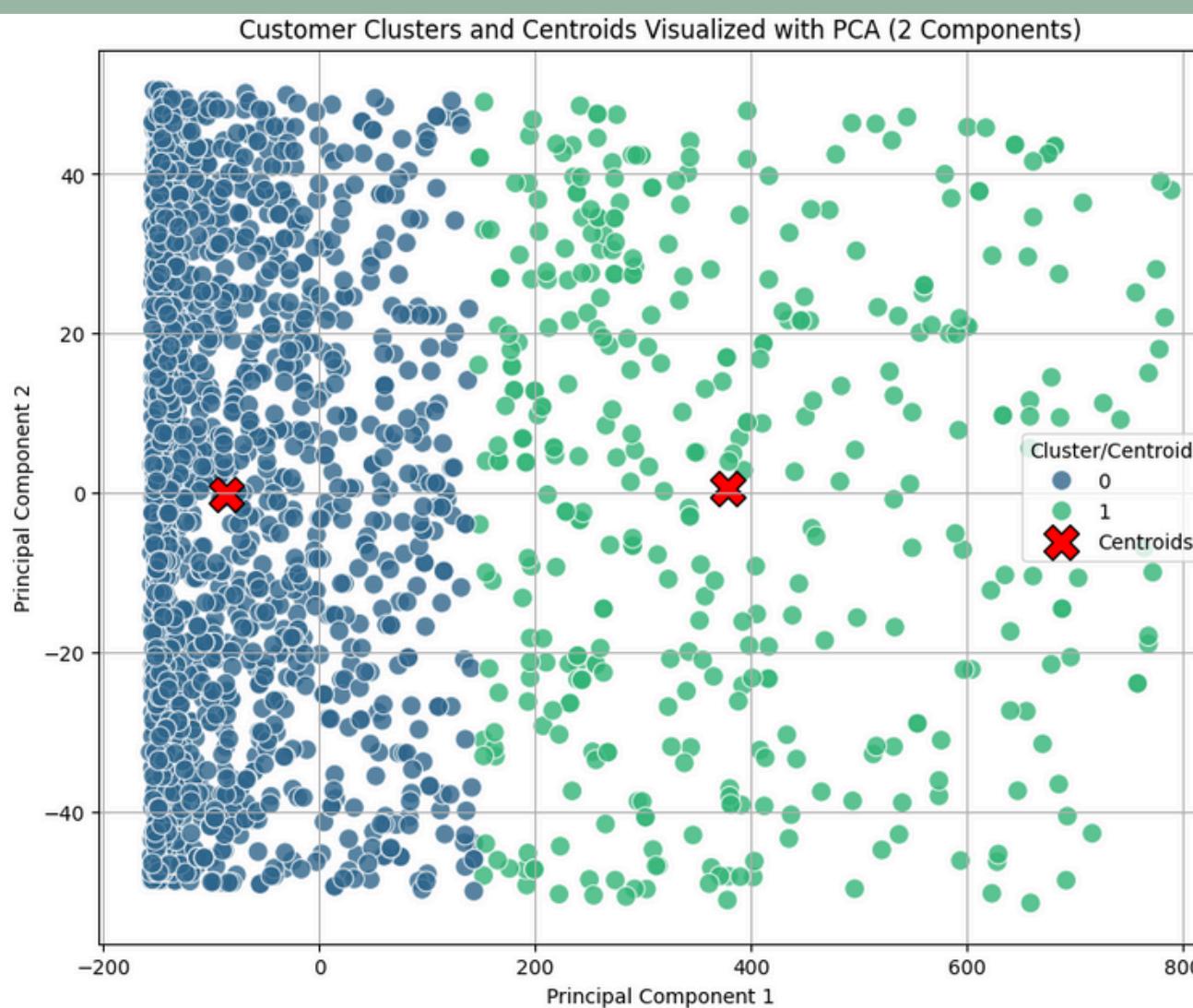
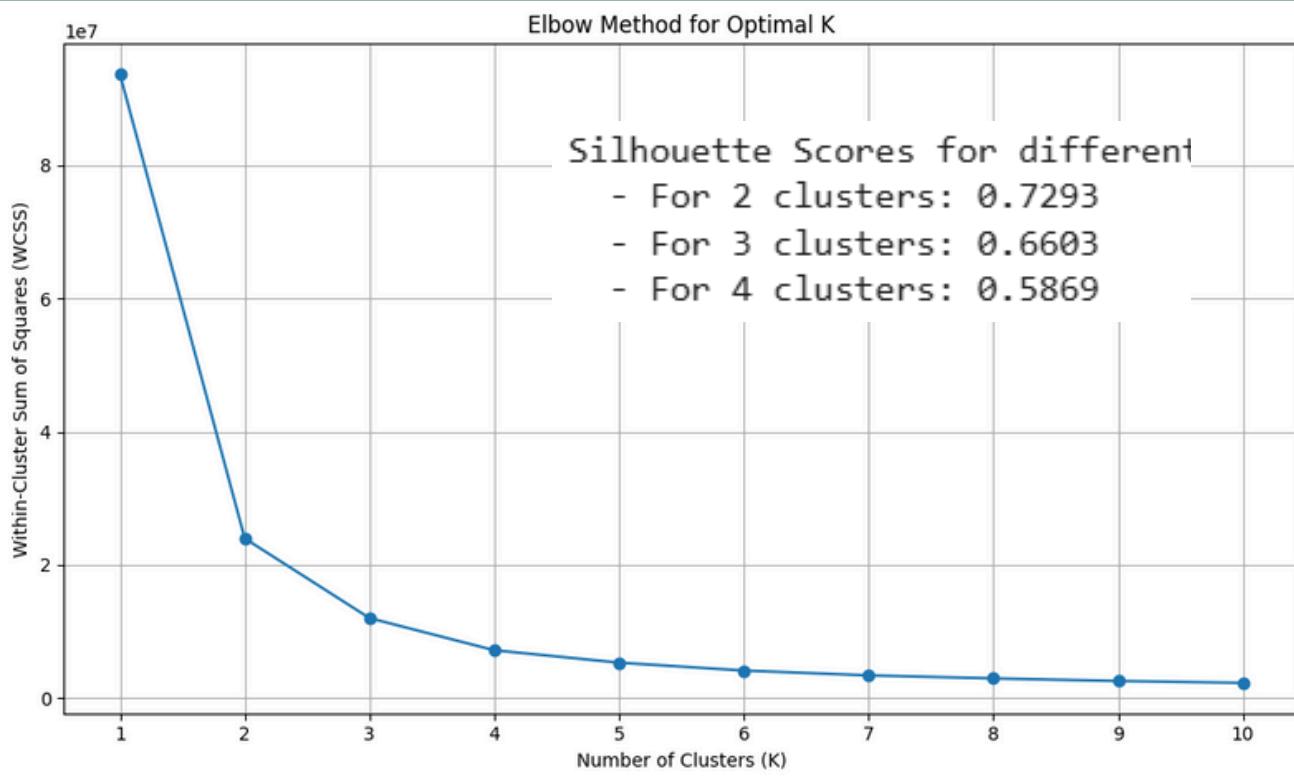
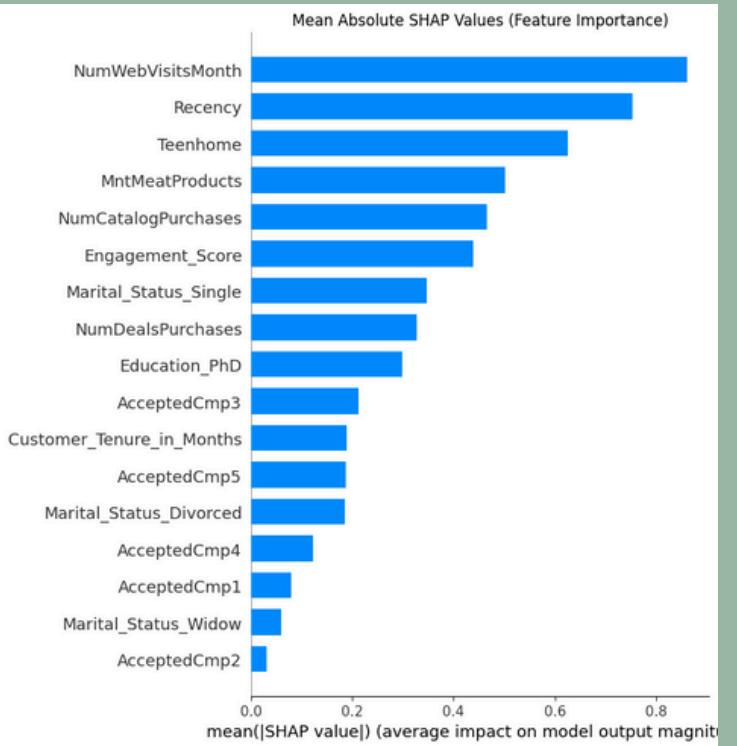
Eigenvalues for each principal component:

- Principal Component 1: 8.7053
- Principal Component 2: 3.2818
- Principal Component 3: 2.5843
- Principal Component 4: 2.2016
- Principal Component 5: 1.7248
- Principal Component 6: 1.5349
- Principal Component 7: 1.3874
- Principal Component 8: 1.2876
- Principal Component 9: 1.2754
- Principal Component 10: 1.2317
- Principal Component 11: 1.1278
- Principal Component 12: 1.1190
- Principal Component 13: 1.0374
- Principal Component 14: 1.0165

	feature	VIF
20	Customer_Tenure	inf
21	Kidhome	inf
16	Teenhome	inf
11	Customer_Tenure_in_Months	inf
34	Dependency_Ratio	inf
0	const	4.224987e+04
29	Marital_Status_Married	5.158129e+02
26	Marital_Status_Together	4.165237e+02
30	Marital_Status_Single	3.634961e+02
9	Marital_Status_Divorced	2.059031e+02
1	Marital_Status_Widow	7.145624e+01

	feature	VIF
0	const	40435.386432
10	Customer_Tenure_in_Months	6.388268
28	Enrollment_Year	6.331907
2	Campaign_Response	6.082372
6	Income	5.210510
22	MntWines	4.008452
4	Education_Graduation	3.379631
7	NumCatalogPurchases	3.250198
25	MntMeatProducts	3.173184
12	AcceptedCmp3	2.988894
23	Education_PhD	2.890076

Summary of the Final Logistic Regression Model: Logit Regression Results					
Dep. Variable:	Response	No. Observations:	1712		
Model:	Logit	Df Residuals:	1694		
Method:	MLE	Df Model:	17		
Date:	Thu, 04 Dec 2025	Pseudo R-squ.:	0.3874		
Time:	05:37:49	Log-Likelihood:	-430.62		
converged:	True	LL-Null:	-702.89		
Covariance Type:	nonrobust	LLR p-value:	7.628e-105		
coef	std err	z	P> z	[0.025]	
const	-5.2364	0.528	-9.908	0.000	-6.272
Marital_Status_Widow	1.5999	0.409	3.907	0.000	0.797
NumDealsPurchases	0.2282	0.055	4.117	0.000	0.120
NumCatalogPurchases	0.2081	0.047	4.409	0.000	0.116
Marital_Status_Divorced	1.2217	0.272	4.499	0.000	0.690
Customer_Tenure_in_Months	0.0569	0.016	3.512	0.000	0.025
Recency	-0.0317	0.003	-9.172	0.000	-0.038
AcceptedCmp3	1.7334	0.259	6.681	0.000	1.225
AcceptedCmp5	1.6306	0.329	4.950	0.000	0.985
Teenhome	-1.2381	0.222	-5.584	0.000	-1.673
AcceptedCmp1	1.2525	0.323	3.876	0.000	0.619
AcceptedCmp2	1.5072	0.638	2.364	0.018	0.257
AcceptedCmp4	0.9531	0.338	2.816	0.005	0.290
Education_PhD	0.8526	0.200	4.260	0.000	0.460
MntMeatProducts	0.0034	0.001	5.468	0.000	0.002
Marital_Status_Single	1.2489	0.210	5.951	0.000	0.838
NumWebVisitsMonth	0.4883	0.065	7.546	0.000	0.361
Engagement_Score	-0.3009	0.065	-4.598	0.000	-0.429



 Overall Response Rate
14%

 Cluster 1 Size
397

 Total Customers
2140

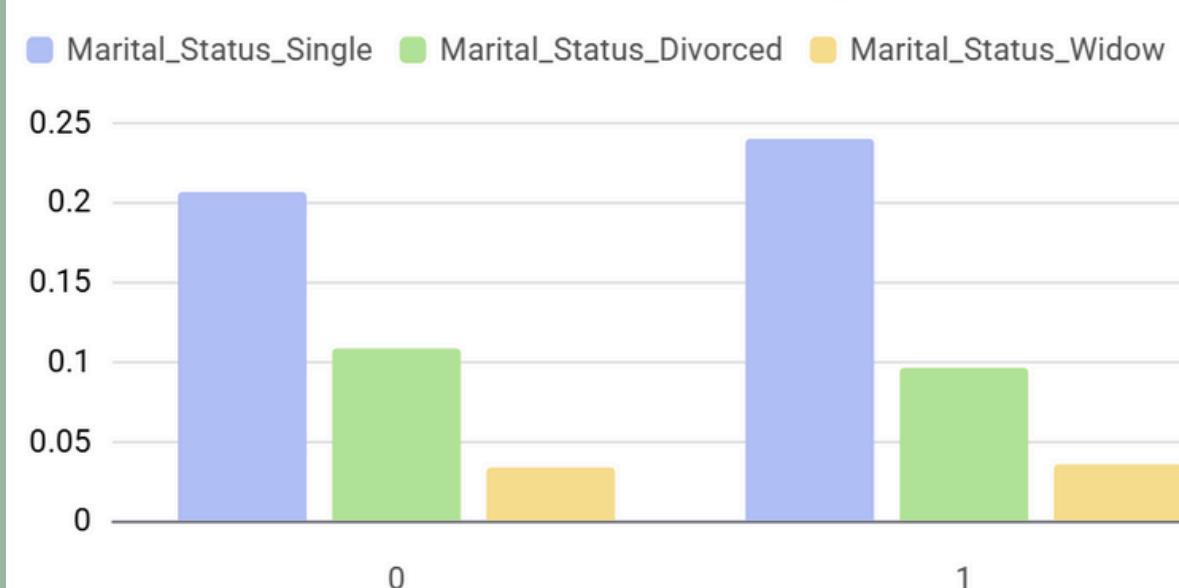
 Cluster 0 Size
1743

Cluster 0: Value Conscious Buyers
Cluster 1: High Value Engaged Shoppers

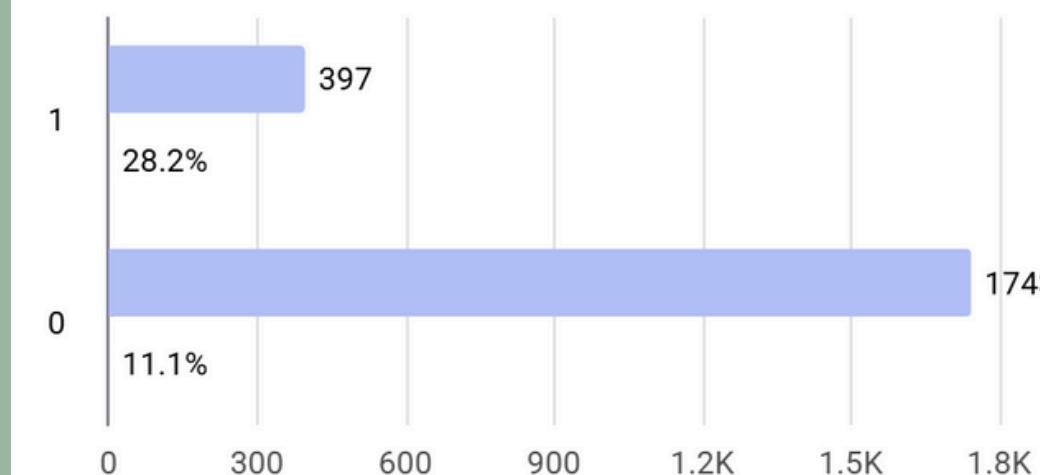
Average Accepted Campaigns by Cluster



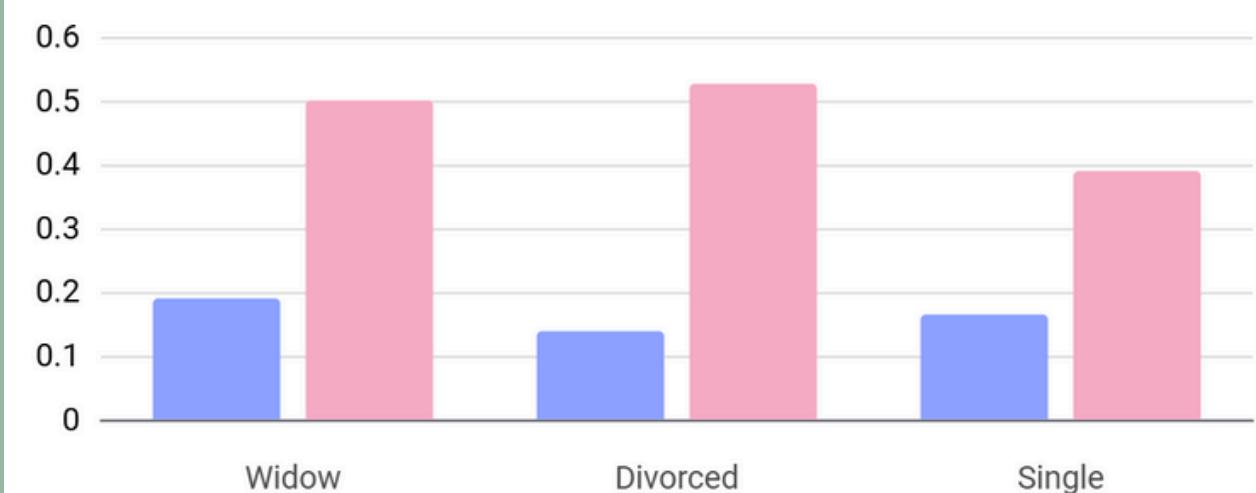
Marital Status Distribution by Cluster



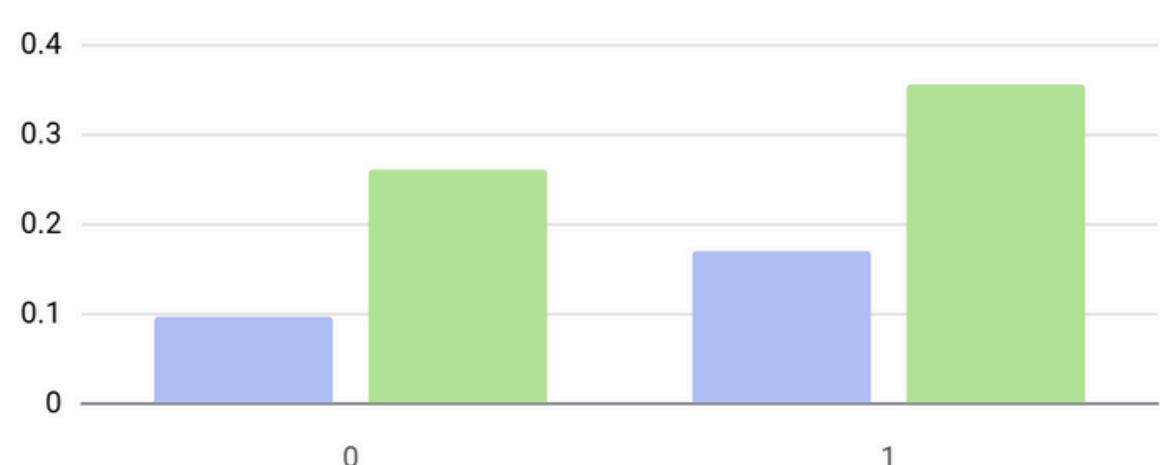
Customer Count and Response Rate by Cluster



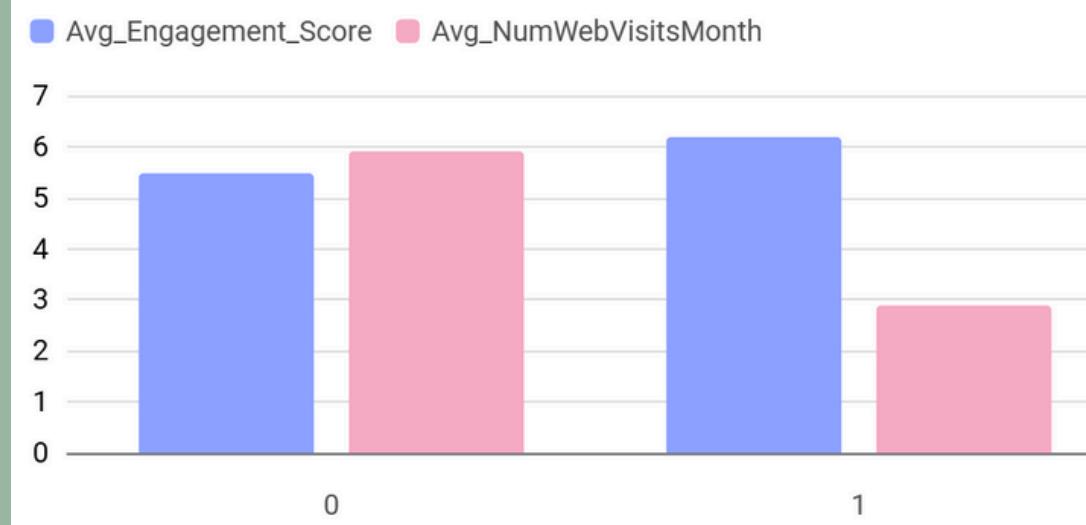
Average Response Rate by Marital Status and Cluster



Average Response Rate by Education and Cluster



Engagement and Web Visits by Cluster



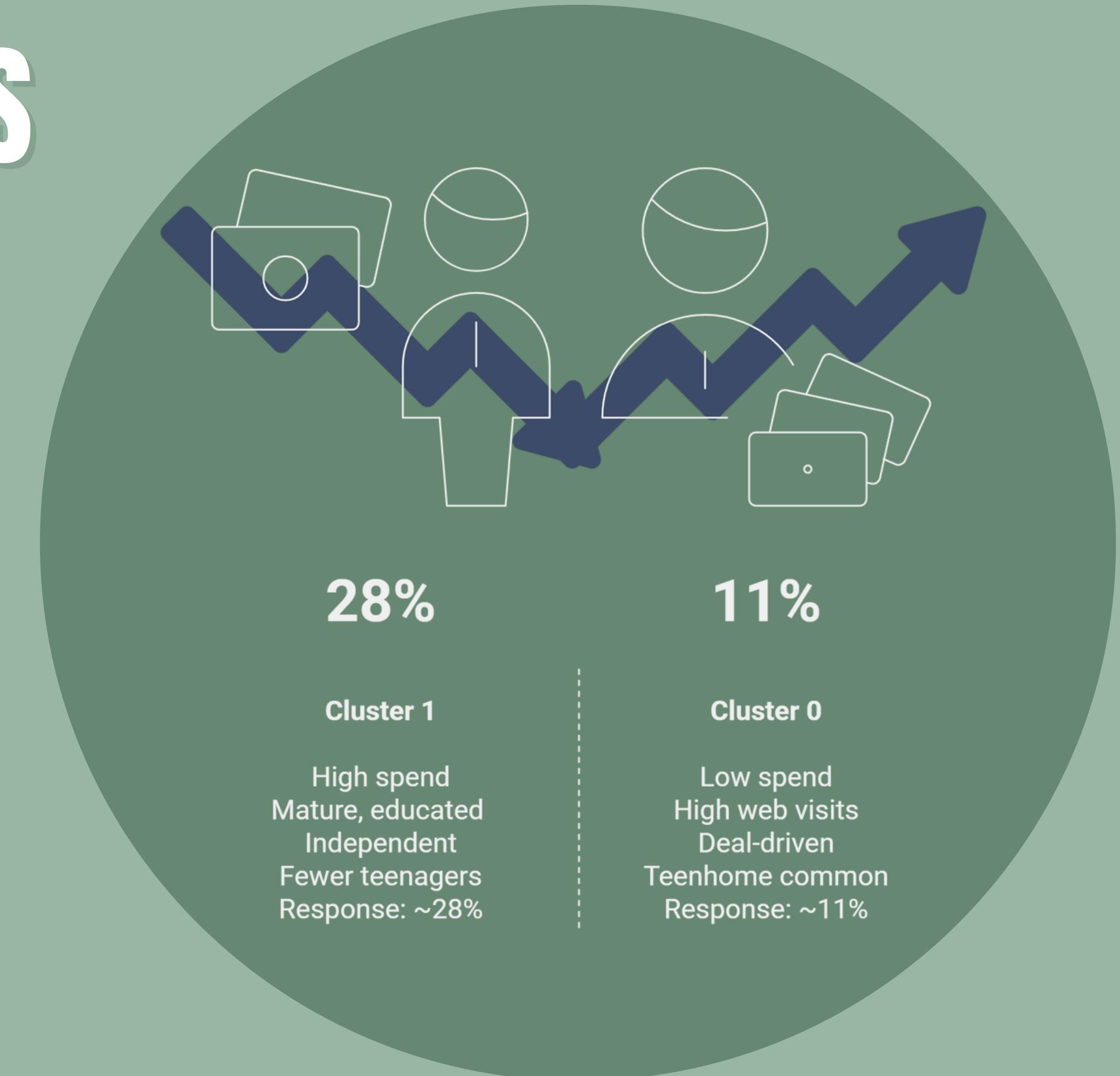
THE TWO SEGMENTS

Segment 1: Budget Browsers (Cluster 0)

- Largest segment ($\approx 81\%$) with lower overall spending, mainly browses online, and responds mostly to deals.
- More likely to have teenagers at home, moderate tenure, and low response/accepted campaign rates.

Segment 2: Elite-Engaged Buyers (Cluster 1)

- Smaller segment ($\approx 19\%$) with high spending, higher engagement, and strong response/accepted campaign rates.
- Typically fewer teenagers, more educated, mature individuals, and prefer direct purchases via catalog or store.



KEY BEHAVIOURAL DRIVERS



Recency of Purchase
Recent purchases ↑ response
“Lower Recency = Higher Response”



Catalog & Channel Engagement

Catalog shoppers respond more

“More Catalog Purchases = Higher Engagement”



High-Value Spending Behaviour

High meat spenders = premium buyers

“Higher Meat Spend = High-Value Segment”

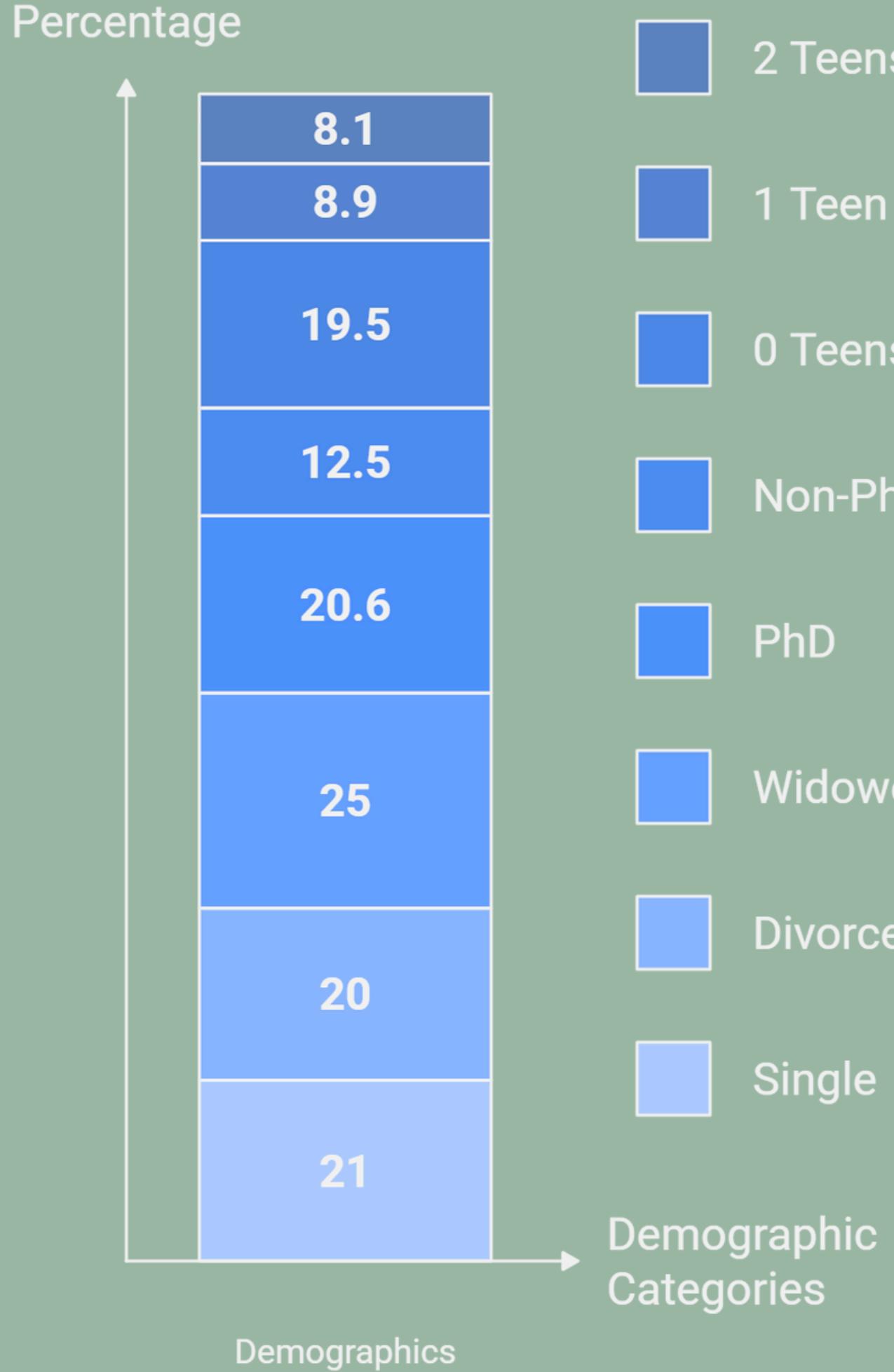


Campaign Acceptance History

Campaign acceptors = highly responsive

“Accepted Cmp → 47%–62.5% Response”





DEMOGRAPHICS THAT DRIVE RESPONSE

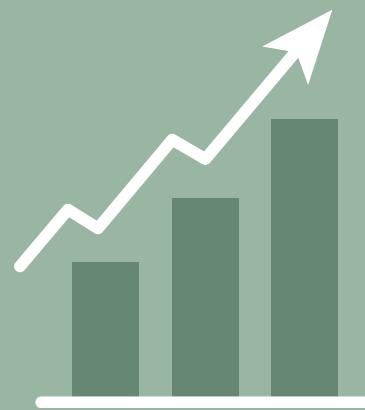
- Singles / divorced / widowed → higher response
- PhD holders respond more
- No teenagers → highest response



SEGMENT-SPECIFIC MARKETING STRATEGIES

Cluster 0: Value-led engagement

Cluster 1: Premium personalization



Value-Conscious
(Cluster 0)



Web Personalization



Discount Triggers



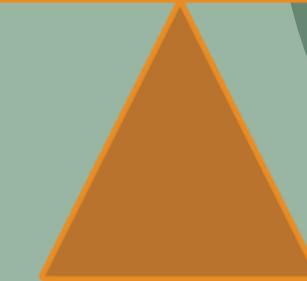
Re-engagement Nudges



Deal-Focused Bundles



Family-Oriented Messaging



High-Value
(Cluster 1)



Premium Bundles



Loyalty Exclusives



High-Quality Product Focus



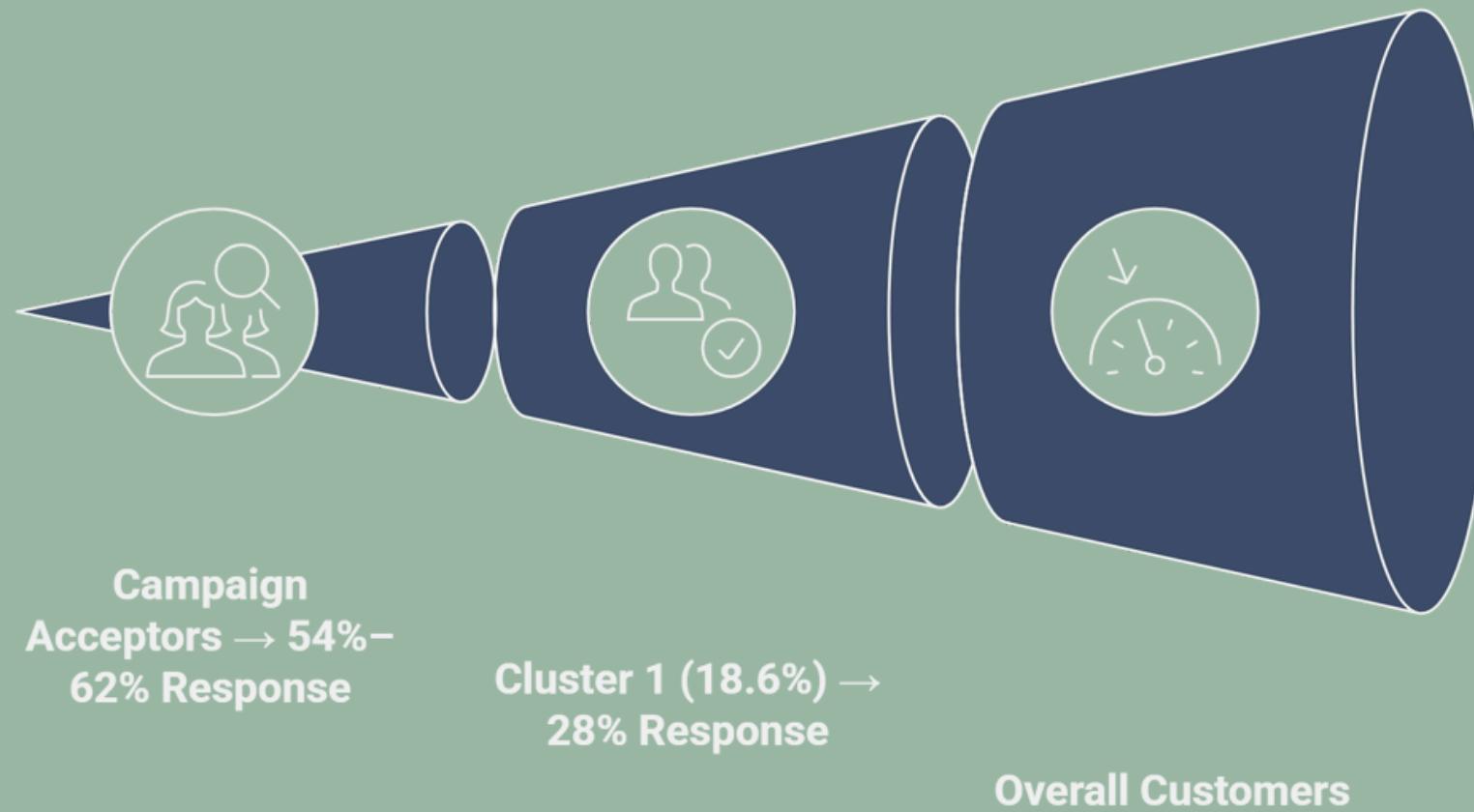
Curated Recommendations



Catalog-First Offers



Customer Response Funnel

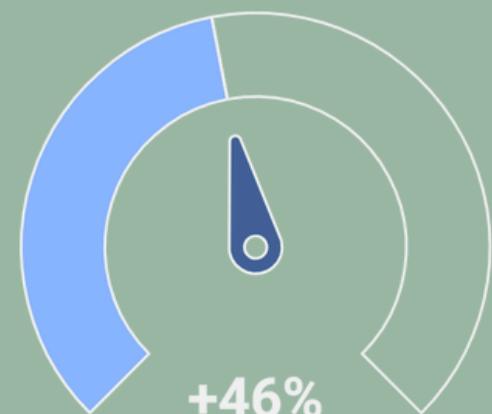


HIGH-IMPACT INSIGHTS FOR CAMPAIGN PLANNING

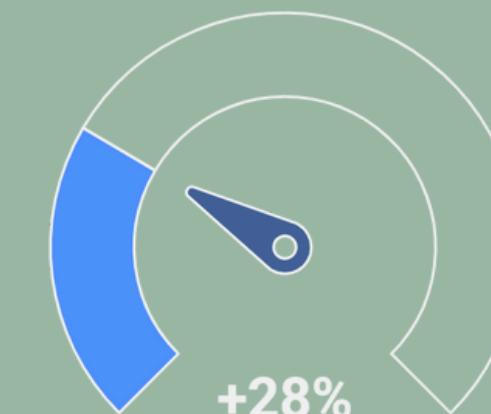
- Prioritize Cluster 1
- Use Cmp1, Cmp2, Cmp5 styles
- Target: recent + high spend
- Demographic micro-segments



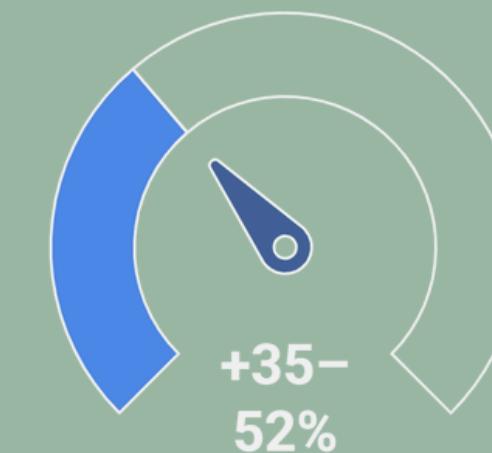
Response Uplift by Customer Segment



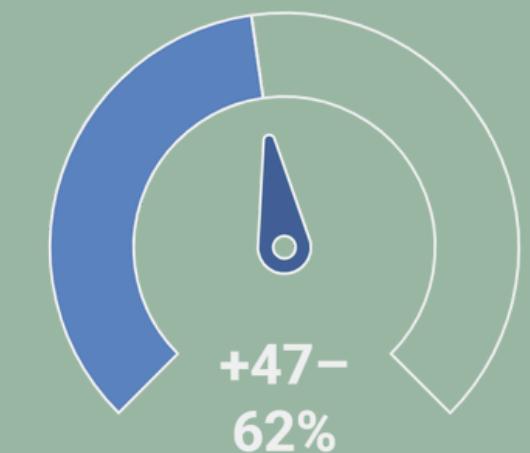
Recent buyers



High meat spend



PhD / Single



Cmp acceptors

THANK YOU



Streamlit App [🔗](#)