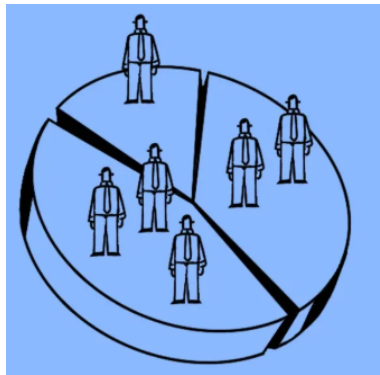# Data-Driven Customer Segmentation for Marketing Optimization
Uncovering Profitable Segments to Maximize Campaign ROI

Data Science Institute - University of Toronto
Cohort 8 Capstone Project
Machine Learning Team - 8
March 7, 2026

Table of contents

Executive Summary

We've analyzed customer data to identify four distinct segments that naturally exist within our customer base. These segments differ significantly in who they are (demographics), how they shop (channel preference), and what they buy (product categories). Each segment requires a different marketing approach to maximize response rates and ROI.

PART 1: MEET YOUR CUSTOMERS – THE FOUR SEGMENTS

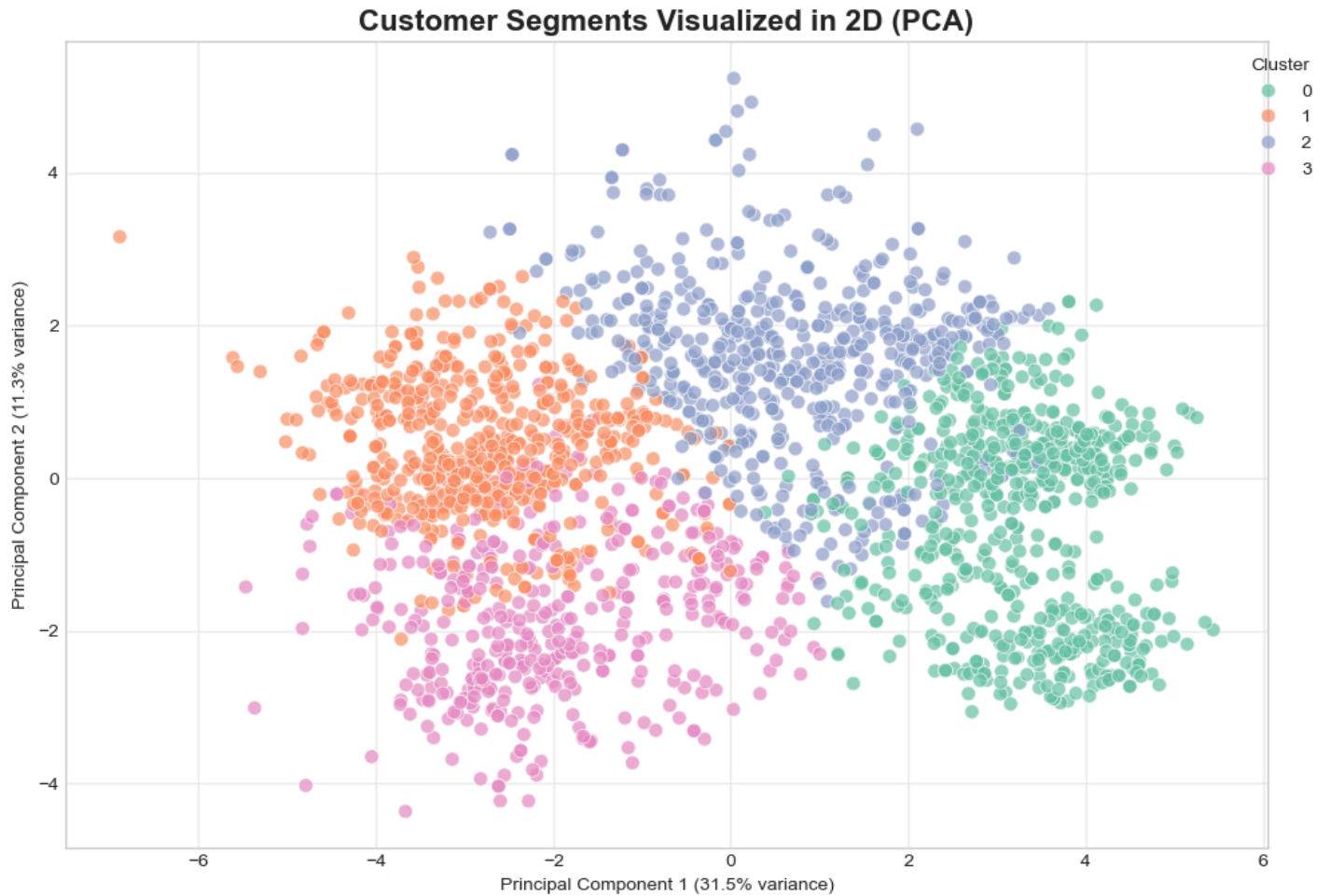| Segment	Size	Who They Are	How They Shop	What They Buy	Marketing Opportunity
0: The Affluent Traditionalists	642 (28.8%)	Highest income ($73.6K), older (58 yrs), no kids at home	Love catalogs & in-store, rarely visit website	Heavy spenders across all categories, especially wine & meat	Catalog mailers, premium product offers, in-store events
1: The Budget-Conscious Parents	614 (27.5%)	Lowest income ($33.6K), younger (54 yrs), largest families (3.2 members)	Web-only shoppers, frequent site visits	Lowest spending across all products	Family bundles, value deals, online-exclusive coupons
2: The Established Families	562 (25.2%)	Middle-high income ($57.3K), oldest (60.5 yrs), have children at home	Multi-channel: store, web, deals	High spending, balanced across categories	Cross-channel campaigns, loyalty rewards, family-focused promotions
3: The Emerging Singles	411 (18.4%)	Lower-middle income ($36.5K), younger (55 yrs), small households	Web visitors, some store purchases	Light spending, basic product mix	Build loyalty early, email campaigns, entry-level products

PART 2: WHAT MAKES EACH SEGMENT UNIQUE
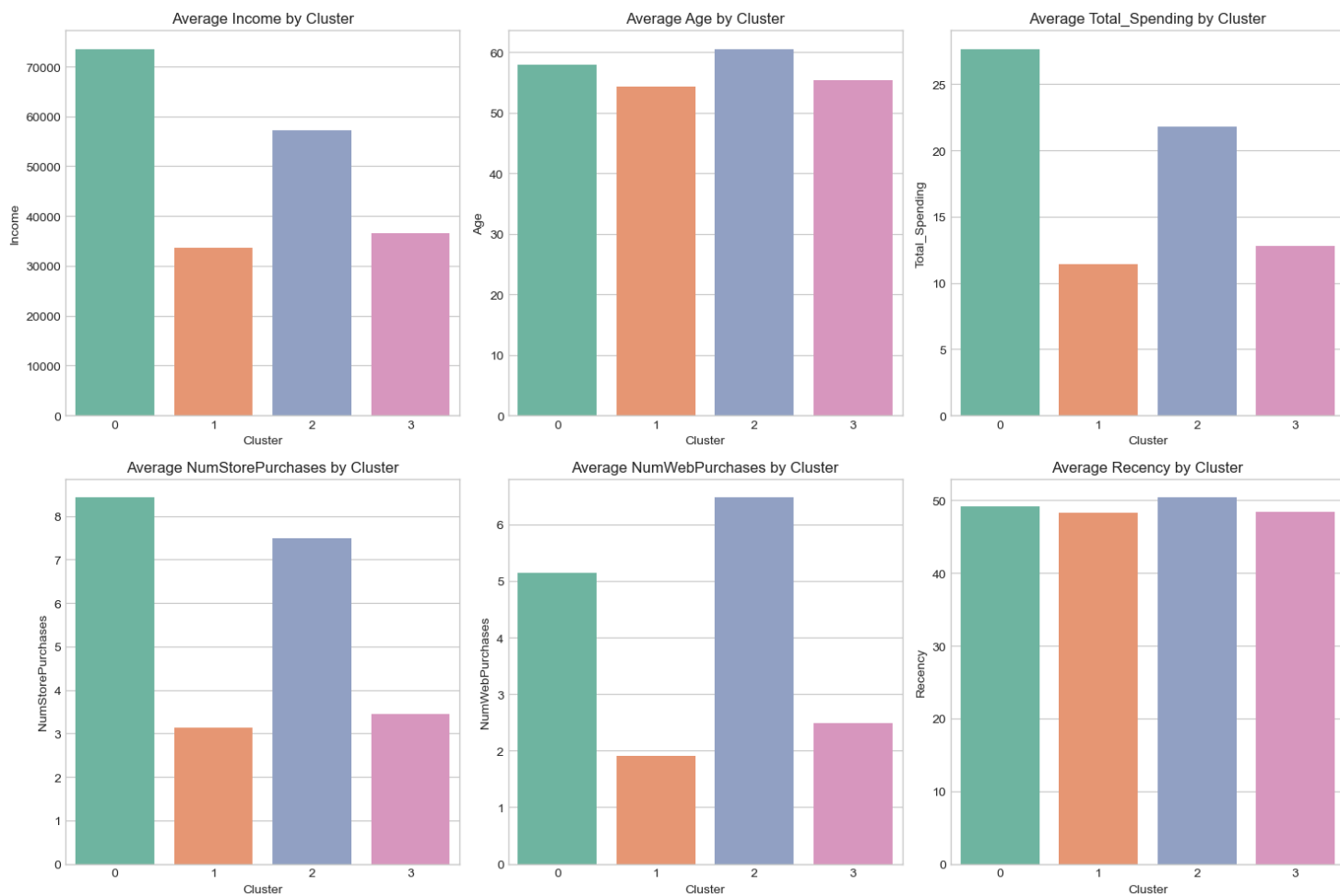

Fig 1: Customer Segments (PCA)

Fig 2 : Cluster Profiles (mean values)

Income & Spending Comparison

Segment	Income	Total Spending	Primary Channel
0 (Affluent Traditionalists)	↑↑↑ High	↑↑↑ High	Store + Catalog
1 (Budget-Conscious Parents)	↓↓ Low	↓↓ Low	Web
2 (Established Families)	↑↑ Medium-High	↑↑Medium-High	All Channels
3 (Emerging Singles)	↓ Medium-Low	↓ Medium-Low	Web + Store

Channel Preference at a Glance

Segment 0:	STORE (8.4) + CATALOG (5.6) → Traditional shoppers
Segment 1:	WEB (6.5 visits/month) → Digital-first families
Segment 2:	STORE (7.5) + WEB (6.5) → Omnichannel families
Segment 3:	WEB (6.5) + STORE (3.4) → Occasional shoppers

PART 3: ACTIONABLE MARKETING STRATEGIES

Segment 0 – The Affluent Traditionalists (28% of customers)

Strategy: **"Premium Treatment"**
- Send high-quality catalog mailers featuring premium wines and meats.
- Invite to exclusive in-store tasting events.
- Cross-sell gold products to this high-spending group.
- Avoid: Heavy digital advertising (they don't visit websites often).

Segment 1 – The Budget-Conscious Parents (27% of customers)

Strategy: **"Digital Value"**
- Email campaigns with family-focused bundles.
- Website-exclusive discount codes.
- Promote deals and value-sized purchases.
- Avoid: Catalog mailings (they don't use catalogs).

Segment 2 – The Established Families (25% of customers)

Strategy: **"Omnichannel Engagement"**
- Consistent presence across all channels.
- Loyalty program with family rewards.
- Retargeting campaigns across web and in-store.
- Focus: Keep them engaged everywhere.

Segment 3 – The Emerging Singles (18% of customers)

Strategy: **"Growth & Retention"**
- Welcome email series to build relationships.
- Introduce higher-value products gradually.
- Monitor recency—they're at risk of churn.
- Focus: Increase basket size over time.

PART 4: MODEL VALIDATION & ROBUSTNESS

Silhouette Score: 0.14 – What does this mean?

Think of silhouette score as measuring "how different are our segments from each other?" A score of 0.14 indicates moderate separation—our segments are distinct enough to be useful, but customers don't fall into perfectly isolated boxes (which is normal for human behavior).

Why we're confident:
- Clear, logical patterns across all features (income aligns with spending, family size aligns with shopping behavior).
- Each segment has a distinct channel preference.
- Segment sizes are balanced (18-28% each), no single segment dominates.

Risks to monitor:
- Customer behavior evolves—segments should be refreshed annually.
- External factors (economy, competitors) may shift patterns.
- 18% of customers (Emerging Singles) are lighter spenders—they need attention to prevent churn.

PART 5: NEXT STEPS

Timeline	Action
First step	Share segments with marketing team; tag customers in CRM
Next Campaign	Design one test campaign per segment, measure response lift
Quarterly	Track segment migration (are customers moving between groups?)
Annual	Re-run clustering with new data to validate/update segments