

# PSYCHOGRAPHIC SEGMENTATION OF ONLINE SHOPPERS USING HIERARCHICAL CLUSTERING IN R

## 1. Background

A mid-sized U.S. e-commerce brand dealing in lifestyle products sought to deepen customer understanding beyond surface-level demographics. While sales and browsing data were available, the brand lacked segmentation based on psychological traits, values, and purchase motivations. The goal was to create psychographic profiles using survey and behavioral data, to drive hyper-targeted email and ad campaigns.

## 2. Objective

- To cluster U.S. online shoppers based on psychographic attributes using R
- To identify actionable consumer personas for brand messaging and digital marketing
- To deliver a reusable segmentation model with integrated reporting

## 3. Data Summary

- **Source:** Brand-led Qualtrics survey + transactional dataset
- **Volume:** 5,487 matched users (survey and purchase history)
- **Variables:**
  - Psychographics: Attitude to pricing, value-seeking behavior, environmental concern, brand loyalty, self-perception (Likert scale)
  - Purchase history: Recency, frequency, monetary value (RFM)
  - Engagement: Newsletter opens, product review submissions, social follows
  - Demographics: Age group, Gender, State

## 4. Methodology

### 4.1 Data Preparation

- Cleaned and standardized Likert scale data (1 to 5) using `scale()`
- Normalized all continuous variables using z-scores

- Applied reverse scoring for negative-framed survey items
- Conducted correlation matrix inspection to reduce multicollinearity among psychographic attributes

## 4.2 Hierarchical Clustering

- Chose **Gower's distance** (via daisy() from cluster) to handle mixed data types
- Dendrogram generated using Ward's method
- Determined **5 cluster solution** using dendrogram cuts and NbClust::NbClust()
- Validated clusters with silhouette width scores and cluster stability indices

## 5. Segment Profiles

Cluster	Share	Psychographic Insight	Engagement Strategy
C1	28%	Eco-conscious, brand-loyal, moderate spenders	Use sustainability badges and loyalty coupons
C2	24%	Trend-driven, high value seekers, fast decision-makers	Push new arrivals and influencer-style emails
C3	19%	Price sensitive, compares frequently, waits for sales	Schedule flash deals and price-drop alerts
C4	17%	Indifferent to brand, buys only by function or need	Retarget through utility-first messaging
C5	12%	Highly engaged reviewers, refer friends, low purchase	Build referral campaigns and feedback loops

## 6. Results

- Segment-specific marketing lifted email CTR by **14.6%** on average
- Returning customer rate improved by **10.1%** in C1 after personalized eco-branding
- Conversion from ads increased by **11%** in C2 via urgency-based creatives
- Survey responses aligned with transaction patterns with **82% predictive consistency**

## 7. Deliverables

- Psychographic segmentation model (saved as R object for reusability)
- R Markdown based report with visual analysis and profiling summary
- Dashboard prototype using Shiny with filters by age, state, or purchase segment
- Ready-to-use CSVs for CRM upload including cluster tags
- Recommendation deck for marketing team with message framework by cluster

## 8. Tools and Packages Used

- cluster, dendextend, NbClust, ggplot2, dplyr, forcats, shiny

## 9. Future Scope

- Automate quarterly re-segmentation pipeline using Rscript via cron jobs
- Integrate psychographic clusters into Facebook Ad custom audience tags
- Extend the model to mobile app engagement data and in-app quiz responses