Customer Segmentation using K-Means Clustering

# Introduction / Problem Statement

Customer segmentation is the process of dividing customers into groups based on similar characteristics. This helps companies design targeted marketing strategies, improve customer satisfaction, and increase revenue.  
  
Objective: To use K-Means clustering to identify different customer groups based on demographic and behavioral data.

# Dataset Description

- Dataset: Mall Customers Dataset  
- Rows: 200, Columns: 5  
- Features:  
 - CustomerID – Unique ID  
 - Gender – Male/Female  
 - Age – Customer’s age  
 - Annual Income (k$) – Income level  
 - Spending Score (1–100) – Shopping behavior

# Exploratory Data Analysis (EDA)

The EDA phase focused on understanding customer demographics and spending trends.

* **Age Distribution:** Customers range from young adults (~20s) to older groups (~60s), with the majority in the **30–45 age range**.
* **Annual Income:** Spans from ~15k to ~140k, with most customers concentrated between **40k and 80k**.
* **Spending Score:** Shows wide variation; some customers with high incomes spend little, while some with lower incomes spend heavily.

Key Observations:

* Younger customers tend to have higher spending scores.
* Higher income does not always mean higher spending.
* Distinct groups of cautious vs. luxury vs. value-focused spenders were visible in scatter plots.

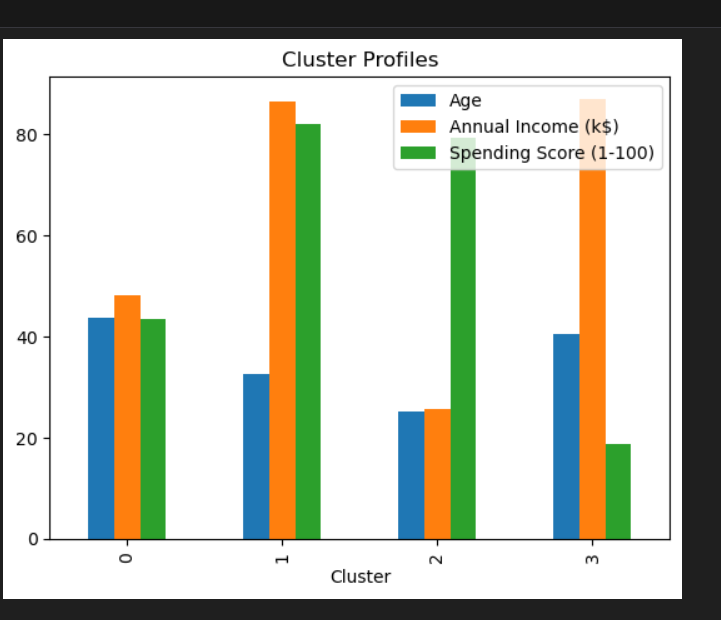
# Methodology

- Preprocessing: Data cleaning, scaling (StandardScaler)  
- Clustering Algorithm: K-Means clustering  
- Optimal Cluster Selection: Elbow Method & Silhouette Score

# Results

Key outputs:  
- Elbow Method plot → Chosen K = 4  
- PCA scatter plot of clusters  
- Cluster profile table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cluster** | **Age (Avg)** | **Annual Income (Avg)** | **Spending Score (Avg)** | **Interpretation** |
| **0** | ~44 (middle-aged) | ~48k (middle-income) | ~43 (moderate) | Mainstream / Mid-Market — largest group, average spenders. Focus on loyalty, mid-range bundles, seasonal offers. |
| **1** | ~33 (young) | ~87k (very high) | ~82 (very high) | High-Value / Luxury Shoppers — young, affluent, spend heavily. Target with exclusivity, VIP service, premium upsell. |
| **2** | ~25 (very young) | ~26k (low) | ~79 (high) | Young Value Seekers — low income but high spending. Use promotions, social-first campaigns, gamified rewards. |
| **3** | ~40 (middle-aged) | ~87k (high) | ~19 (very low) | Cautious Affluents — wealthy but reluctant to spend. Overcome barriers with guarantees, education, risk-free trials. |



**Segment-wise Marketing Recommendations**

**Cluster 0 — Mainstream / Mid-Market**

**Profile:**

* Largest group (101 customers).
* Age: ~44 years, middle-aged.
* Income: ~48k (middle-income).
* Spending: Moderate (~43).

**Detailed Recommendations:**

* **Product Strategy:** Offer **value-for-money bundles** (e.g., family packs, subscription boxes).
* **Pricing Strategy:** Introduce **tiered pricing** — mid-range SKUs + premium upsell.
* **Engagement:** Launch **seasonal promotions** (festive offers, back-to-school campaigns).
* **Loyalty Programs:** Introduce a **points-based system** with birthday rewards.
* **Channel Strategy:** Email + WhatsApp updates; in-app personalized offers.
* **Retention Focus:** Offer post-purchase service (extended warranty, loyalty vouchers).

**Cluster 1 — High-Value / Luxury Shoppers**

**Profile:**

* 39 customers.
* Younger (avg 33).
* Very high income (~87k).
* High spend score (~82).

**Detailed Recommendations:**

* **Product Strategy:** Push **exclusive / limited edition products**.
* **Pricing Strategy:** Premium positioning, not discounts — highlight exclusivity.
* **Engagement:** Launch a **VIP club** (priority access, early launches, concierge service).
* **Personalization:** Use **AI recommendations** (past high-ticket purchases).
* **Branding:** Collaborate with **influencers & luxury partners**.
* **Upsell Strategy:** Push high-margin add-ons (warranties, accessories).
* **Retention:** Invite to **private events** (online/offline).

**Cluster 2 — Young Value Seekers**

**Profile:**

* 22 customers.
* Very young (avg 25).
* Low income (~26k).
* Surprisingly high spending score (~79).

**Detailed Recommendations:**

* **Product Strategy:** Introduce **affordable, trendy items**.
* **Promotions:** Flash sales, discount coupons, “Buy 1 Get 1” offers.
* **Engagement:** Use **social-first campaigns** (TikTok, Instagram Reels, gamified contests).
* **Loyalty:** Introduce **gamified rewards** (spin-the-wheel, streak points).
* **Urgency Marketing:** SMS/Push: “Hurry! 2 hours left”.
* **Retention:** Offer **student discounts** or subscription-lite options.
* **Cross-Sell:** Focus on **frequent low-cost add-ons**.

**Cluster 3 — Cautious Affluents**

**Profile:**

* 38 customers.
* Middle-aged (~40).
* High income (~87k).
* Low spend score (~19).

**Detailed Recommendations:**

* **Barriers:** They have money but hesitate to spend (trust issues, risk aversion).
* **Product Strategy:** Highlight **durability & quality guarantees**.
* **Pricing Strategy:** Value-based positioning (justify premium pricing with long-term savings).
* **Engagement:** Detailed product comparisons, video explainers, customer testimonials.
* **Retention:** **Risk-reversal offers**: Free trials, easy returns, warranty extensions.
* **Nurturing:** Email sequences educating on features, safety, and value-for-money.
* **Cross-Sell:** Push “must-have essentials” first → gradually introduce higher-value products.

**Strategic Summary**

* **Cluster 0 (Mainstream):** Drive **repeat & cross-sell** via loyalty.
* **Cluster 1 (Luxury):** Focus on **exclusivity, VIP service, and premium upsell**.
* **Cluster 2 (Young Value Seekers):** Capture **volume with promotions & social-first engagement**.
* **Cluster 3 (Cautious Affluents):** Overcome **trust barriers with guarantees, education, and trials**.

# Conclusion

The clustering analysis successfully segmented customers into meaningful groups. These insights can be used to improve marketing strategies, optimize customer engagement, and increase profitability.