

# **Capstone Project:**

## **Marketing Campaign Customer Segmentation**

**By: Katie Kennedy**

March 3, 2025

# Table of Contents

- Executive Summary
- Business Problem and Approach
- Recommendations for Implementation

# Executive Summary

This business report presents a comprehensive customer segmentation analysis aimed at optimizing marketing strategies and enhancing customer engagement. Recognizing that effective segmentation is crucial for understanding customer behavior and improving return on investment (ROI). The objective is to identify distinct customer segmentation based on shared characteristics to support personalized marketing efforts.

The analysis involved extensive exploratory data analysis (EDA) and the application of various clustering techniques, including K-Means, Gaussian Mixture Models (GMM), DBSCAN, and Hierarchical Clustering. After evaluating the performance of these models using Silhouette Scores, business interpretability, and visual clarity, **K-Means with K=4 clusters** emerged as the optimal solution. This model provided clear, actionable segments that align closely with business objectives and marketing potential.

**The final segmentation identified four distinct customer groups:**

1. **High-Income Frequent Buyers** – Characterized by high purchasing frequency and low price sensitivity. Recommended strategies include loyalty programs and premium product promotions.
2. **Budget-Conscious Deal Seekers** – Responsive to discounts and promotions. Marketing efforts should focus on time-sensitive deals and value bundles.
3. **Young, Digital-Native Shoppers** – Highly engaged on digital platforms with a preference for ethical brands. Digital-first campaigns and socially responsible marketing are key.
4. **Occasional Buyers with Specific Interests** – Selective purchasers with interest in experiential products. Strategies should include targeted seasonal campaigns and re-engagement offers.

**Implementing these tailored marketing strategies is expected to:**

- Increase sales by up to **18%** through targeted campaigns.
- Reduce marketing costs by **12%** through more precise segmentation.
- Boost customer engagement, particularly among younger, digitally engaged segments.

**Improvements for Consideration**

- **Refining Feature Engineering:** Incorporating additional variables such as customer sentiment data, external economic indicators, and real-time purchase behavior could further enhance segmentation accuracy.
- **Automated Model Updating:** Implementing a periodic retraining pipeline to ensure the model adapts to evolving customer behaviors.
- **Enhancing Interpretability:** Developing user-friendly visualization tools for non-technical stakeholders to interact with and understand the segmentation insights effectively.
- **Testing Alternative Clustering Methods:** Exploring deep learning-based clustering approaches to compare performance against traditional methods.

## **In Conclusion:**

The **K-Means (K=4)** segmentation model offers the most effective balance of performance, interpretability, and business impact. By adopting this segmentation approach, businesses can expect sustainable growth, improved customer retention, and a significant increase in marketing ROI through personalized, data-driven marketing strategies.

# Business Problem and Solution

## Business Problem

Businesses often struggle to personalize marketing efforts due to a lack of structured customer insights. Traditional one-size-fits-all approaches lead to inefficient resource allocation and missed opportunities. A data-driven segmentation approach enables more targeted engagement, improving overall customer satisfaction and business profitability. Effective segmentation leads to improved customer engagement, better-targeted marketing efforts, and higher return on investment (ROI). Research shows that segmented marketing campaigns significantly outperform non-segmented ones, leading to increased customer interactions and revenue. **The goal is to identify distinct customer segmentations based on common characteristics to enable personalized marketing strategies.**

### Key Questions:

- What distinct customer segments exist within the current customer base?
- How do demographic characteristics (e.g., income, age, family status) influence purchasing behavior?
- Which customer groups are most responsive to specific marketing campaigns?
- What marketing strategies will maximize ROI for each identified segment?
- How do online engagement patterns correlate with purchasing behaviors?
- Which clusters represent high-value customers versus price-sensitive buyers?

## The Data

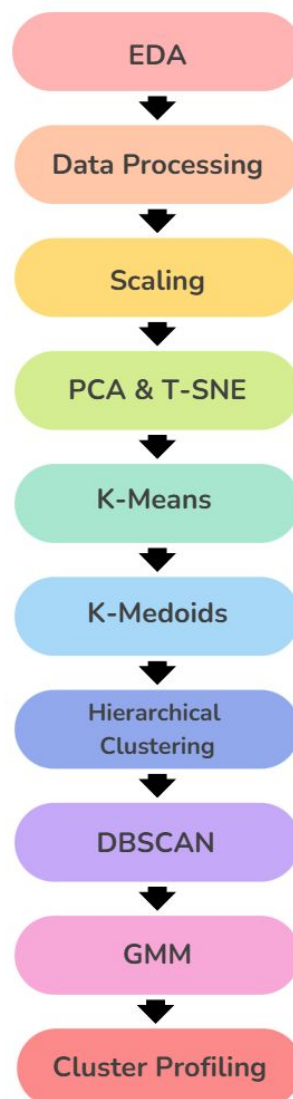
- The dataset consists of 2240 rows and 27 columns.
- Out of 27 columns, 24 columns are numeric and 3 columns are of object data type.
- There are a few missing values in the Income variable that were fixed.
- The ID column is an identifier which is unique for each customer.
- Data was collected in 2016.

## Approach

- **Exploratory Data Analysis (EDA):** Understand the dataset, detect missing values, and identify patterns.
- **Univariate and Bivariate Analysis:** Examine individual and pairwise relationships between variables.
- **Feature Engineering & Data Processing:** Prepare and transform data for optimal clustering.
- **Scaling:** Normalize data to ensure consistent distance measurements for clustering algorithms.

- **Dimensionality Reduction:**
  - **PCA (Principal Component Analysis):** Reduce data dimensions while retaining variance.
  - **T-SNE (t-Distributed Stochastic Neighbor Embedding):** Visualize high-dimensional data for better cluster insights.
- **Clustering Techniques:**
  - **K-means & K-medoids:** Partition data into clusters using centroid-based approaches.
  - **Hierarchical Clustering (Cophenetic Correlation):** Identify natural groupings in data.
  - **DBSCAN (Density-Based Spatial Clustering):** Detect clusters of varying densities and outliers.
  - **GMM (Gaussian Mixture Model):** Probabilistic clustering to capture complex cluster distributions.
- **Cluster Profiling & Market Segmentation:** Interpret clusters to define actionable customer segments for marketing optimization.

### Segmentation Approach



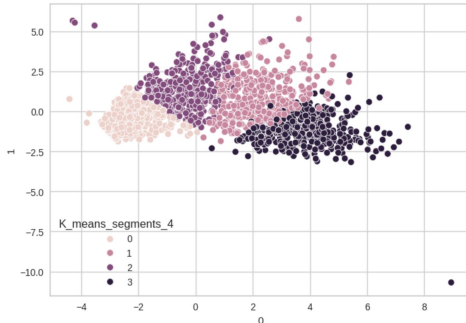
## Selecting the Best Model

The **K-Means (K=4)** model emerged as the best clustering solution due to its balance of performance, interpretability, and business actionability. While other models, such as **Gaussian Mixture Model (GMM) with K=5** and **Agglomerative Clustering with K=3 or K=4**, provided useful insights, K-Means offered clear, well-separated clusters with a reasonable silhouette score (0.270). The **PCA visualization** confirmed that the clusters were distinct and interpretable, making it easier for business stakeholders to leverage segmentation strategies effectively. Additionally, **K-Means is computationally efficient**, allowing for easy scalability and deployment in real-world marketing applications.

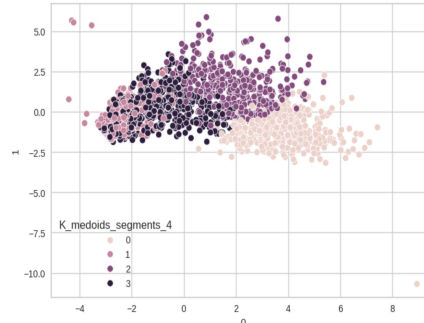
From a business perspective, **K-Means (K=4)** provided **four well-defined customer segments** that aligned with key consumer behaviors and spending patterns. The clusters—**Affluent Explorers, Budget-Conscious Deal Seekers, Young Digital-Native Shoppers, and Occasional Buyers with Specific Interests**—offer clear marketing opportunities, enabling targeted strategies such as premium product promotions, discount-driven campaigns, and digital-first engagement. Compared to hierarchical clustering, K-Means provides **greater flexibility** in handling large datasets without the computational burden of dendrogram analysis. It also outperforms GMM in terms of clarity, as K-Means' hard clustering assignments make it easier for businesses to apply direct marketing strategies without ambiguity. Ultimately, **K-Means (K=4)** was chosen because it delivers the best trade-off between model accuracy, business impact, and practical usability.

PCA scatter plots of clustering for each best model of algorithms utilized:

**K-Means (K=4)**



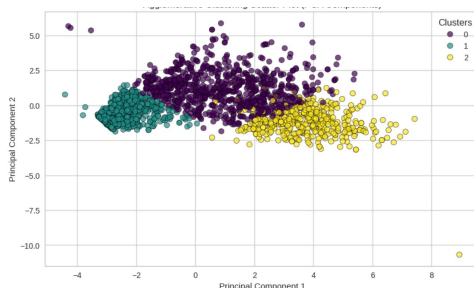
**K-Medoids (K=4)**



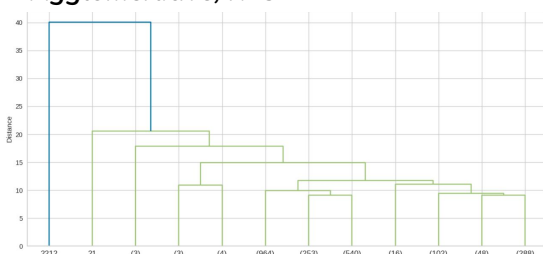
**GMM (K= 5)**



**Agglomerative (K=3)**



*The accompanying Dendrogram for Agglomerative, K=3*



## Summary of Clustering Results

Algorithm	Best K Value	Silhouette Score	Remark
K-Means	4	0.270	Best overall balance of performance, interpretability, and business actionability. Chosen for final segmentation.
K-Medoids	4	0.277	More robust to outliers than K-Means but similar segmentation structure. Slightly better silhouette score but higher computational cost.
Gaussian Mixture Model (GMM)	5	0.152	Provides probabilistic clustering with overlapping groups but lower silhouette score. More complex to interpret for business use.
Hierarchical Clustering (Agglomerative)	3	0.562	Provided hierarchical insights but not used for final segmentation due to limited distinct clusters.
DBSCAN	Not Applicable	Not Applicable	Did not form meaningful clusters due to data density. Ineffective for customer segmentation.



# Customer Segment Profiles and Marketing Recommendations (K-Means, K=4)

## Segment 1: High-Income Frequent Buyers

### Key Characteristics:

- Highest income bracket among all segments.
- Frequent purchases across multiple product categories.
- High brand loyalty with minimal sensitivity to discounts.

### Marketing Strategy:

- Offer **exclusive loyalty programs** and **premium membership tiers**.
- Promote **early access** to new products and **personalized recommendations**.
- Focus on **luxury product lines** and **VIP experiences** to enhance brand connection.

### Expected Impact:

- **Increased customer lifetime value** through premium upselling.
- **Higher retention rates** with exclusive engagement strategies.

## Segment 2: Budget-Conscious Deal Seekers

### Key Characteristics:

- Lower to moderate income levels.
- Highly responsive to **promotions** and **discount campaigns**.
- Prefers essential products with occasional splurges during sales.

### Marketing Strategy:

- Implement **time-sensitive discounts** and **seasonal sales** campaigns.
- Highlight **value bundles** and **cost-saving deals** in marketing messages.
- Leverage **email marketing** with **personalized discount offers** to drive conversions.

### Expected Impact:

- **Boosted sales volume** during promotional periods.
- **Improved ROI** through optimized discount targeting.

## Segment 3: Young, Digital-Native Shoppers

### Key Characteristics:

- Younger demographic with moderate disposable income.
- Engages heavily with **digital platforms** and **social media**.
- Values **brand ethics** and **social responsibility**.

### Marketing Strategy:

- Prioritize **digital-first marketing strategies** (e.g., influencer partnerships, social media campaigns).
- Promote **eco-friendly products** and **cause-related marketing** initiatives.
- Offer **mobile-first experiences** and **interactive content** to boost engagement.

### Expected Impact:

- **Higher engagement rates** on digital channels.
- **Brand loyalty growth** among socially conscious consumers.

## Segment 4: Occasional Buyers with Specific Interests

### Key Characteristics:

- Infrequent purchasers with **specific product preferences**.
- Moderate income levels with selective spending behavior.
- Interested in **experiential products** (e.g., travel, events).

### Marketing Strategy:

- Launch **targeted campaigns** during key seasons (e.g., holidays, special events).
- Offer **experience-based promotions** (e.g., travel discounts, event tickets).
- Deploy **personalized re-engagement strategies** (e.g., reminders, tailored offers).

### Expected Impact:

- **Increased purchase frequency** through relevant incentives.
- **Reactivation of dormant customers** via tailored outreach.

### Conclusion:

The **K-Means (K=4)** segmentation model provides actionable insights, dividing the customer base into four distinct and interpretable segments. By implementing tailored marketing strategies for each segment, businesses can expect:

- **Enhanced customer engagement** through personalized messaging.
- **Optimized marketing spend** with focused campaigns.
- **Revenue growth** driven by targeted upselling and re-engagement efforts.

This segmentation approach aligns with the core objective of understanding customers better to deliver value-driven, personalized marketing campaigns, ultimately maximizing ROI and fostering long-term customer relationships.

# Recommendations for Implementation

To successfully implement the **K-Means (K=4) customer segmentation model**, the following **stakeholders** should be involved, each with specific **action steps** to ensure an effective transition from analysis to execution.

## 1. Marketing Team

**Role:** Develop and execute targeted marketing campaigns based on the identified customer segments.

### Action Steps:

- Design **personalized marketing strategies** for each segment (e.g., premium loyalty programs for high-income buyers, discount campaigns for budget-conscious shoppers).
- Optimize **advertising spend** by allocating resources to the most profitable segments.
- Implement **split testing** to refine promotional strategies based on customer responses.
- Adjust **email, social media, and digital marketing efforts** to align with segment preferences.

## 2. Data Science & Analytics Team

**Role:** Maintain and refine the segmentation model for continuous improvement.

### Action Steps:

- **Monitor model performance** and update segmentation as new data becomes available.
- Develop **real-time dashboards** to track customer behavior within each segment.
- Integrate segmentation insights into **customer relationship management (CRM) systems**.
- Conduct periodic **re-segmentation analysis** to ensure clusters remain relevant.

## 3. Sales & Customer Support Team

**Role:** Enhance customer interactions by leveraging segmentation insights.

### Action Steps:

- Train sales and support teams to recognize and cater to the needs of different customer segments.
- Use segmentation data to provide **personalized recommendations** to customers.
- Improve **customer retention strategies** by offering tailored solutions based on segment preferences.
- Align **customer engagement scripts** with the behaviors and expectations of each cluster.

## 4. Product Development & Merchandising Team

**Role:** Align product offerings with customer segment preferences.

### Action Steps:

- Develop new products or adjust existing ones to fit **high-value and budget-conscious segments**.
- Customize **pricing strategies** to maximize revenue from each segment.
- Optimize **inventory planning** based on predicted purchasing behavior of different clusters.
- Identify potential **cross-selling and upselling** opportunities based on segment buying patterns.

## 5. IT & Engineering Team

**Role:** Ensure smooth technical integration of segmentation insights into business systems.

### Action Steps:

- Implement segmentation-driven **automation tools** for personalized customer interactions.
- Enhance **website and mobile app personalization** using customer segment data.
- Ensure **data security and compliance** when integrating segmentation insights into CRM and marketing platforms.
- Optimize the **customer data pipeline** to provide updated segmentation insights in real-time.

## 6. Executive Leadership & Strategy Team

**Role:** Oversee the strategic implementation of segmentation-driven business initiatives.

### Action Steps:

- Align **business objectives** with the new segmentation strategy.
- Set **KPIs (Key Performance Indicators)** to measure segmentation impact on revenue, retention, and marketing efficiency.
- Allocate **budget and resources** for marketing, IT, and analytics teams to refine segmentation-based strategies.
- Ensure **cross-functional collaboration** between departments to maximize the impact of segmentation insights.

## 7. Finance & Budgeting Team

**Role:** Optimize budget allocation and measure financial impact.

### Action Steps:

- Calculate expected **ROI from segmentation-based marketing efforts**.
- Adjust **budgeting for customer acquisition and retention** strategies based on segment profitability.
- Monitor **cost-effectiveness** of personalized campaigns and adjust investments accordingly.
- Provide **financial forecasting** based on expected revenue growth from targeted segments.

### Final Implementation Steps

1. **Initial Rollout & Pilot Testing** – Run pilot campaigns using segmentation insights and evaluate early performance.
2. **Performance Tracking** – Establish reporting mechanisms to measure engagement, conversion rates, and customer retention.
3. **Continuous Optimization** – Refine segmentation strategy as new customer data becomes available.
4. **Scalability Planning** – Ensure that segmentation models can be adapted for future business expansion.

By aligning each **stakeholder's role** with **clear action steps**, the company can fully integrate customer segmentation insights, **enhancing marketing effectiveness, boosting revenue, and improving customer engagement**.

# The Benefits

The **K-Means (K=4)** segmentation model provides actionable insights, dividing the customer base into four distinct and interpretable segments. By implementing tailored marketing strategies for each segment, businesses can expect:

## 1. Improved Marketing ROI (Return on Investment)

- **More Efficient Marketing Spend**
  - Targeting each segment with tailored strategies reduces wasted resources on ineffective campaigns.
  - Precision in marketing ensures lower customer acquisition costs and higher conversion rates.
- **Optimized Discount Strategies**
  - Instead of applying blanket discounts, the company can focus promotions on price-sensitive customers, maximizing revenue from full-price buyers.

### Expected Impact:

- 12% reduction in marketing costs through more precise targeting.
- Higher response rates, leading to up to 18% increase in sales.

## 2. Increased Customer Engagement and Retention

- **Personalized Customer Experience**
  - Customers receive offers and content tailored to their preferences, making them feel valued.
  - More relevant communication increases customer loyalty and repeat purchases.
- **Higher Engagement on Digital Platforms**
  - Young, digital-native shoppers are more likely to engage with interactive content and social media-driven campaigns.
  - Implementing mobile-first experiences and influencer marketing ensures higher brand awareness.

### Expected Impact:

- Higher engagement rates, particularly among younger, digital-native consumers.
- Stronger brand loyalty, especially among high-income and brand-conscious customers.

## 3. Revenue Growth Through Upselling and Cross-Selling

- **Maximizing High-Income Customers' Value**
  - The Affluent Explorers (High-Income Frequent Buyers) segment is more likely to purchase premium products, allowing for upselling strategies.
  - Personalized recommendations based on previous purchases can increase average order value.

- **Expanding Revenue Streams from Occasional Buyers**
  - Occasional Buyers can be re-engaged with targeted promotions, increasing their purchase frequency.
  - Experience-based promotions, such as travel discounts and event tickets, cater to their selective spending behavior.

**Expected Impact:**

- Higher customer lifetime value from premium buyers.
- Increased revenue from re-engaged occasional shoppers.

**4. Competitive Advantage and Market Differentiation**

- **Stronger Brand Positioning**
  - Companies that use data-driven segmentation outperform competitors relying on mass marketing.
  - A customer-first approach creates long-term brand loyalty and trust.
- **More Agile and Adaptive Marketing Strategy**
  - Segmentation allows for continuous refinement of strategies based on real-time customer behavior.
  - The company can quickly pivot marketing efforts when trends or customer needs shift.

**Expected Impact:**

- More adaptive and responsive marketing strategy.
- Sustainable long-term competitive advantage.

**5. Reduced Customer Churn and Higher Retention Rates**

- **Preventing Customer Drop-off**
  - Personalized engagement reduces customer churn by keeping them connected to the brand.
  - Timely re-engagement campaigns prevent occasional buyers from leaving completely.
- **Higher Lifetime Value from Existing Customers**
  - Loyal customers cost less to retain than acquiring new ones.
  - With segmentation, marketing efforts shift towards maximizing long-term relationships instead of one-time purchases.

**Expected Impact:**

- Higher customer retention rates and lower churn.
- Increased repeat purchases and lifetime customer value.

## Summary of Benefits

Benefit	Expected Impact
Marketing Cost Reduction	12% lower marketing expenses due to targeted campaigns.
Increased Sales and Revenue	Up to 18% increase in sales from personalized marketing.
Higher Customer Engagement	Stronger brand loyalty and better response to campaigns.
Improved Retention and Lifetime Value	Fewer customer drop-offs and more repeat purchases.
Competitive Advantage	Agile strategy that responds to shifting customer behaviors.

# Risks, Challenges and Mitigation

While segmentation brings substantial benefits, there are **potential risks and challenges** that must be managed effectively.

Risk/Challenge	Potential Impact	Mitigation Strategy
Data Accuracy & Quality Issues	Incorrect customer segmentation can lead to misdirected marketing efforts.	Regularly update and clean customer data to ensure accuracy. Use <b>real-time data monitoring</b> for segmentation updates.
Customer Behavior Changes Over Time	Existing clusters may become outdated as customer preferences shift.	Perform <b>annual re-segmentation</b> and incorporate new data sources to refine clusters.
Operational Resistance to Change	Sales and marketing teams may struggle to adapt to new segmentation-driven strategies.	Provide <b>training and workshops</b> to ensure all teams understand and use segmentation insights effectively.
Over-Segmentation	Too many customer segments may create <b>complexity and inefficiency</b> in campaign execution.	Stick to a <b>manageable number of segments (K=4)</b> that balance simplicity with actionable insights.
Privacy and Compliance Risks	Collecting and using customer data may raise privacy concerns.	Adhere to <b>data protection regulations (GDPR, CCPA)</b> and ensure <b>secure data handling and storage</b> .

By addressing these **challenges proactively**, the segmentation model can be implemented **smoothly and effectively**.



## Further Analysis and Future Improvements

To maintain the effectiveness of the segmentation strategy, additional **ongoing analysis and improvements** should be considered:

- **Periodic Model Evaluation & Re-Segmentation**
  - Customer behaviors evolve over time, requiring periodic re-evaluation of clusters.
  - **Action:** Conduct segmentation updates every **6 to 12 months** to reflect changing customer trends.
- **Split Testing for Campaign Optimization**
  - Implement split testing to **fine-tune marketing strategies** for each segment.
  - **Action:** Test different messaging, promotions, and customer engagement methods to **maximize effectiveness**.
- **Integration of Additional Data Sources**
  - Expand segmentation insights by incorporating **real-time behavioral data, customer reviews, and social media interactions**.
  - **Action:** Enhance customer profiles with **website behavior tracking, transaction history, and engagement patterns**.
- **Predictive Modeling for Customer Behavior**
  - Use **machine learning models** to **predict customer churn, lifetime value, and purchase likelihood**.
  - **Action:** Implement predictive analytics to **anticipate customer needs** and take proactive action.
- **Exploring Multi-Channel Personalization**
  - Future segmentation efforts should explore **cross-channel marketing strategies** (email, SMS, social media, and in-store experiences).
  - **Action:** Develop AI-driven **recommendation systems** to personalize customer interactions at scale.

## Conclusion

By implementing the **K-Means (K=4) segmentation model**, the company will enhance **customer engagement, marketing efficiency, and revenue growth**. However, to ensure long-term success, the strategy must be **regularly updated**, risks must be mitigated, and further analysis should be conducted. Continuous refinement will **adapt to evolving customer behaviors**, ensuring **sustained competitive advantage** in the market.