Research Report: Customer Segmentation for Rubies Mall

Problem Story

Rubies mall, our client, is planning a huge marketing campaign and we have been employed as analysts to ensure the campaign is a success. Our job is to effectively segment their customer base for targeted advertisement and a personalized customer experience.

Solution Roadmap

For our analysis, we employed a machine learning technique (K-means clustering) to effectively segment the customers by similarities in features (age, gender, spending score, and annual income).

Methodology

01

Data Exploration

We explored the dataset, checking for data type correctness, duplicates, and null values.

The dataset contained 5 features. Customer ID was dropped because it was irrelevant to the analysis.

02

Feature Selection for Clustering

We focused on Age, Annual Income and Spending Score for clustering, as their interaction is key for the marketing segmentation.
Reasoning: These features directly reflect purchasing power and shopping tendencies.

03

Optimal Cluster Number

We used the Elbow Method (Inertia vs. K) to determine the optimum number of clusters to reflect the features.

04

K-Means Clustering

We applied K-means with the chosen K (3 clusters) and visualized results via scatter plots.

Features Engineering

The customers have the following features represented in the dataset: ID, Age, Gender, Income (in thousand Dollars), Spending Score (spending on a scale of 1-100).

Age, Income and Spending Score are the relevant features for this analysis. These features have little to no relation with each other.

Gender was encoded using OHE (One Hot Encoding) for model input.

Data Quality

The Rubies mall data is clean, anomaly-free, and representative of the target population.

Pricing Uniformity

Rubies mall maintains uniform pricing; income and spending score are valid differentiators.

Building the model

The K-means clustering algorithm was implemented in this analysis due to its simplicity, and scalability over variate features and dataset.

After calculating inertia over a range, **4** was found to be the optimum number of clusters for this segmentation through the Elbow method. Instead, we used **3** clusters for our analysis, this is because we had a total of 4 features and would risk overfitting the model.

Results: Key Customer Segments

Our analysis revealed several distinct customer clusters based on Age, Annual Income and Spending Score:

Potential Spenders - Cluster 0

Low Spending: Customers in this cluster have an evenly distribution among ages and annual income but a low spending score between (0 to 60)

Customers in this cluster might entail a range from occasional visitors or infrequent spenders to customers with untapped financial potential.

Need to encourage engagement through targeted promotions, BOGO offers, or experiential marketing.

VIP Customers - Cluster 1

High Income – High Spending: with high Annual income (120 to 140) have high Spending Score(20 to 80) and an Age between 0 to 35 with cases of a few Outliers.

Highly engaged, generous spenders. Ideal for exclusive loyalty programs, premium product previews, and personalized concierge services.

Value Seekers - Cluster 2

Low Income – High Spending: Young customers with ages (Between 0 to 30) have an annual income between 30 to 125 and spend a lot with a score between 0 to 85.

We need to engage this cluster with discount campaigns, bundled deals, and early sales access.

Recommendations for Action

Based on the identified customer segments, we propose the following targeted marketing strategies:

- **Targeted Discounts:** Offer exclusive discounts to "Potential Spenders" (High Income – Low Spending) to boost engagement and wallet share.
- **Loyalty Rewards:** Enhance premium loyalty programs for "VIP Customers" (High Income – High Spending) with personalized services and exclusive access.
- **BOGO Promotions:** Engage "Value Seekers" (Low Income – High Spending) with BOGO deals, bundles, and seasonal sales.
- **Awareness Campaigns:** Launch broad campaigns and introductory offers for "Occasional Buyers" (Low Income – Low Spending) to increase visits and exploration.
- **Dynamic Segmentation:** Implement quarterly reviews of clustering analysis to ensure responsive segmentation to evolving customer behaviors and market trends.

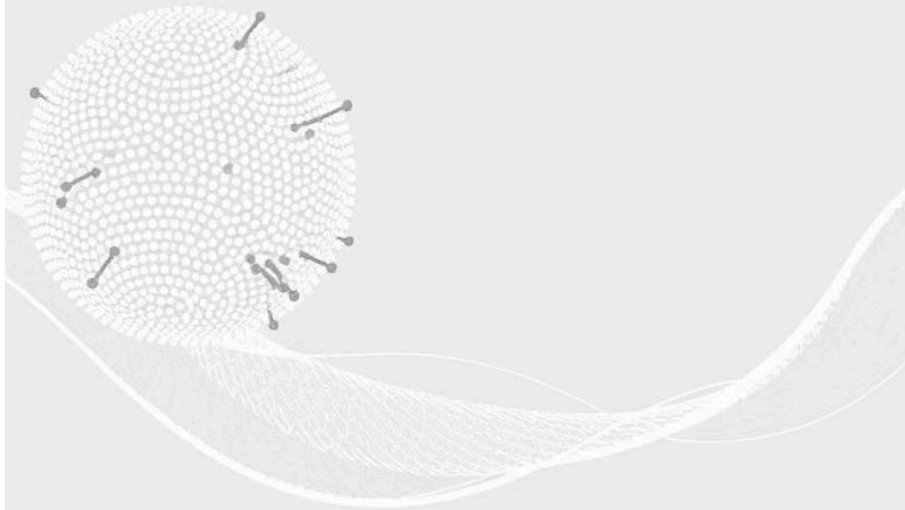
Limitations & Future Considerations

Methodological Limitations

- The "Spending Score" is a subjective metric. Incorporating granular purchase data (e.g., transaction history) would provide more robust insights.

Assumptions for Future Analysis

- This analysis assumes consistent customer behavior. Economic shifts or market trends could invalidate current segments.
- Future analyses could integrate demographic, psychographic, and online behavior data for richer segmentation.



Conclusion & Next Steps

K-means clustering effectively segmented mall customers by income and spending, enabling a shift from generic to personalized marketing. This provides a strong foundation for efficient resource allocation and enhanced customer experiences.

Targeted Promotions

Implement campaigns for each segment to maximize impact and ROI.

Personalized Experiences

Deliver experiences that resonate with individual customer needs.

Continuous Optimization

Periodically refine segments to adapt to evolving customer behaviors.

Deeper Insights

Incorporate transaction histories, demographics, and psychographics for richer segmentation.