

Customer Loyalty and Segmentation Insights for Turtle Games, *July 2024*

1. Background:

Turtle Games, a global retailer of board games, video games, toys, and books, seeks to leverage customer data to drive sales growth. The marketing team aims to understand how and why customers earn loyalty points, identify customer segments based on spending behavior and income, and use customer reviews to guide future campaigns.

This project analyzes transactional and review data using Python and R to uncover spending–loyalty relationships, segment customers, and assess review sentiment. Statistical checks such as skewness and kurtosis are applied to evaluate the reliability of loyalty points for predictive modeling. The results provide actionable insights for targeted marketing, customer retention, and product improvement.

2. Analytical approach:

To support Turtle Games’ goal of optimizing sales and marketing performance, I used a combined Python and R workflow for data ingestion, cleaning, exploration, modeling, and sentiment analysis.

Data Preparation

The dataset (nine columns, 1,999 entries) was imported and cleaned to ensure quality and relevance. Irrelevant columns, such as language and platform, were removed, missing values were handled with `dropna()`, and duplicate entries (especially in review and summary fields) were eliminated to ensure reliable sentiment analysis. Column names were standardized, and categorical variables (e.g., gender, education) were numerically encoded to enable decision tree modeling.

Exploratory Analysis

Python and R were used to explore data structure, distributional properties, and key trends. A pairplot of remuneration and spending score revealed multi-modal distributions and clear group separations, supporting the use of K-means clustering for customer segmentation. In R, the `moments` package and the `Jarque` package. The `test()` function was applied to assess the skewness and kurtosis of loyalty points, ensuring their suitability for predictive modeling.

Visualization & Statistical Insights

Using R’s `ggplot2`, I created boxplots of loyalty points across binned ranges of remuneration and spending scores, helping identify loyalty patterns and behavioral drivers. These visual findings directly informed segmentation and targeting strategies.

Sentiment Analysis

Python’s `TextBlob` library generated polarity scores for customer reviews and summaries. This highlighted common customer pain points and satisfaction drivers, guiding potential product improvements and marketing messaging.

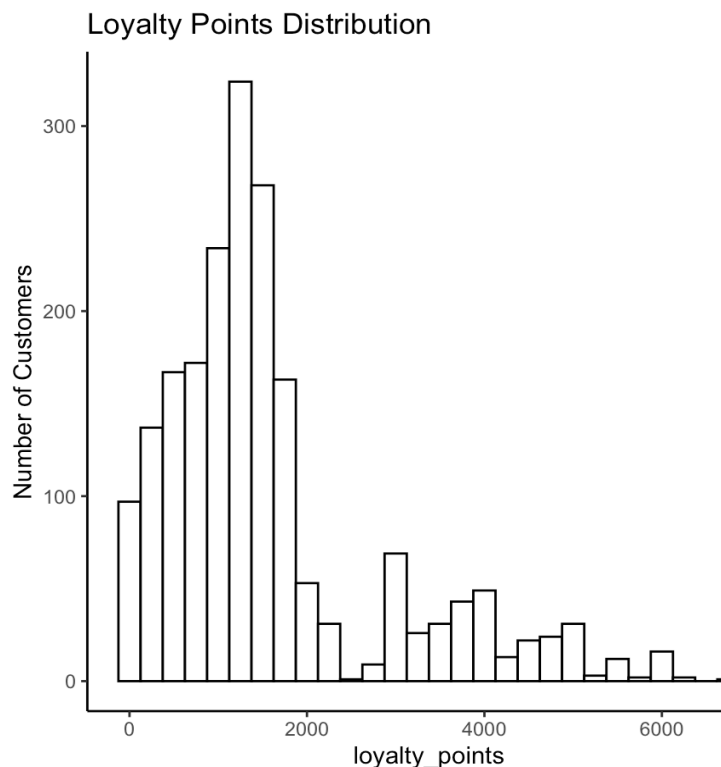
Code Structure

Custom Python functions were developed for key operations, ensuring a cohesive, maintainable, and scalable codebase — supporting future updates and collaboration.

Visualizations and Insights:

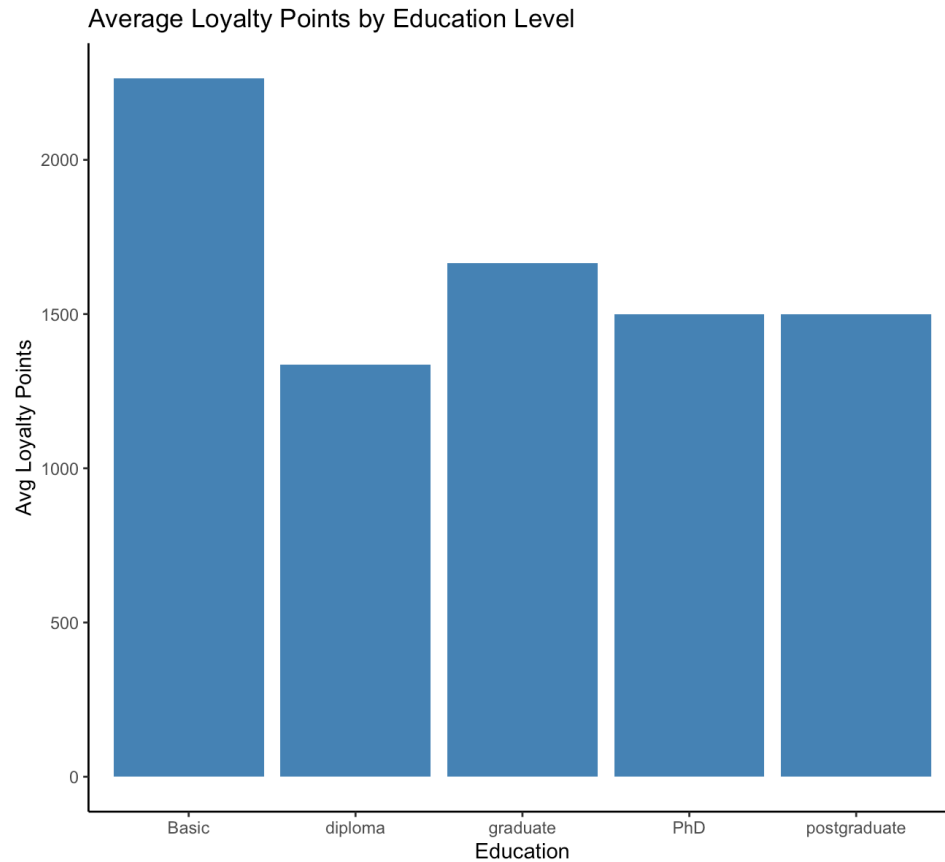
To derive actionable insights aligned with Turtle Games’ objectives of optimizing sales and marketing performance, a series of targeted visualizations was developed using Seaborn and Matplotlib in Python, along with ggplot2 and scatterplot3d in R. These visualizations explored relationships between customer demographics, spending behavior, and loyalty points, revealing key patterns and potential customer segments to inform data-driven marketing strategies..

- Customer Loyalty Points: Frequency Distribution:



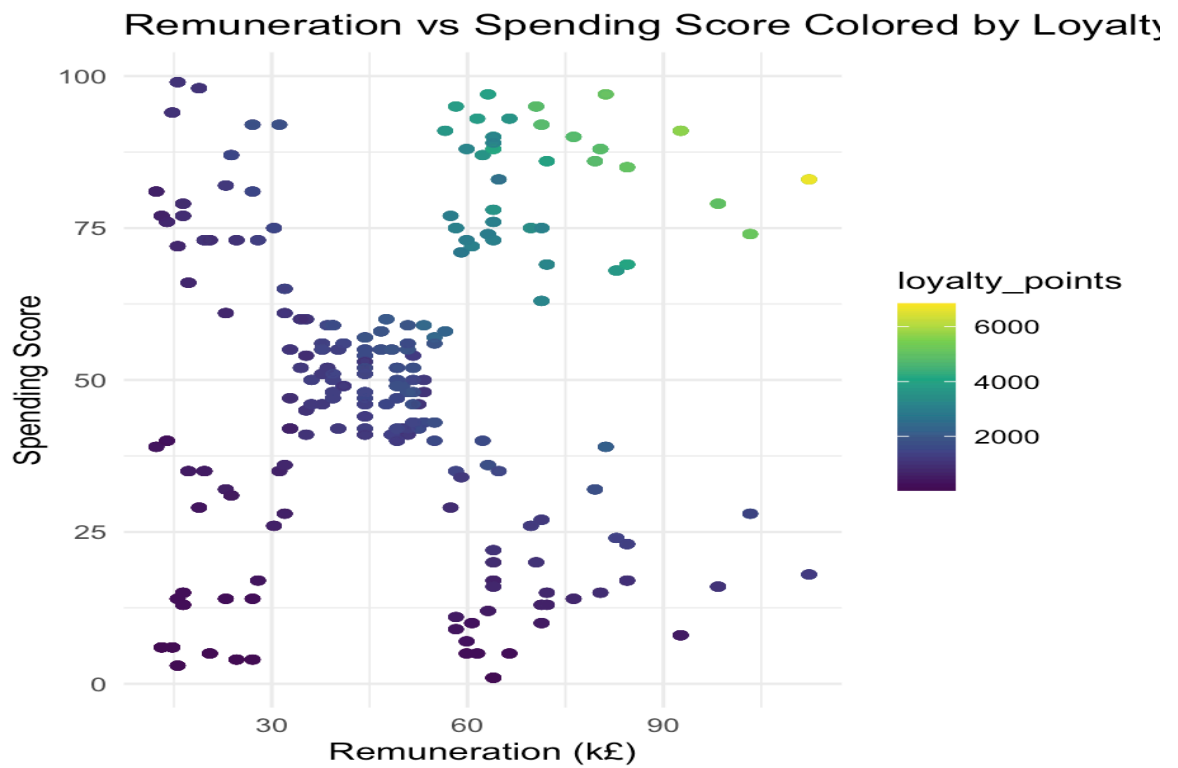
The distribution shows that most customers earn relatively low loyalty points, while a smaller segment accumulates significantly higher points. This suggests Turtle Games has a broad under-engaged customer base alongside a high-value segment of frequent or high-spending customers. Targeted campaigns could focus on increasing engagement within the larger, low-loyalty group while implementing retention strategies to maintain the loyalty of high-value customers.

- Loyalty points by education level



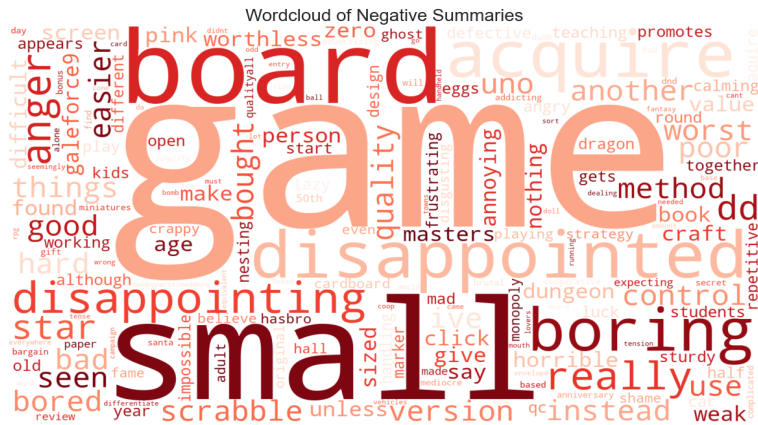
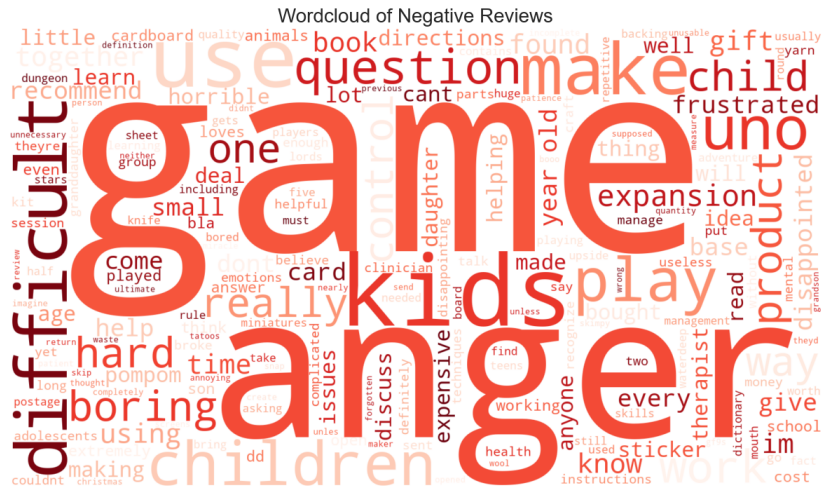
Customers with only a basic education record the highest average loyalty points, outperforming more educated groups. This challenges common assumptions about engagement drivers, indicating that loyalty is influenced more by factors such as income or purchasing preferences. Turtle Games could test tailored communication and rewards strategies across education levels to maximize engagement.

- Spending Score Colored by Loyalty Points



The scatter plot indicates that higher remuneration and spending scores tend to be associated with higher loyalty points. However, a notable segment of high-income customers remains under-engaged—representing untapped potential. Targeted campaigns and exclusive offers could convert these customers into high-value loyalists, maximizing retention and revenue.

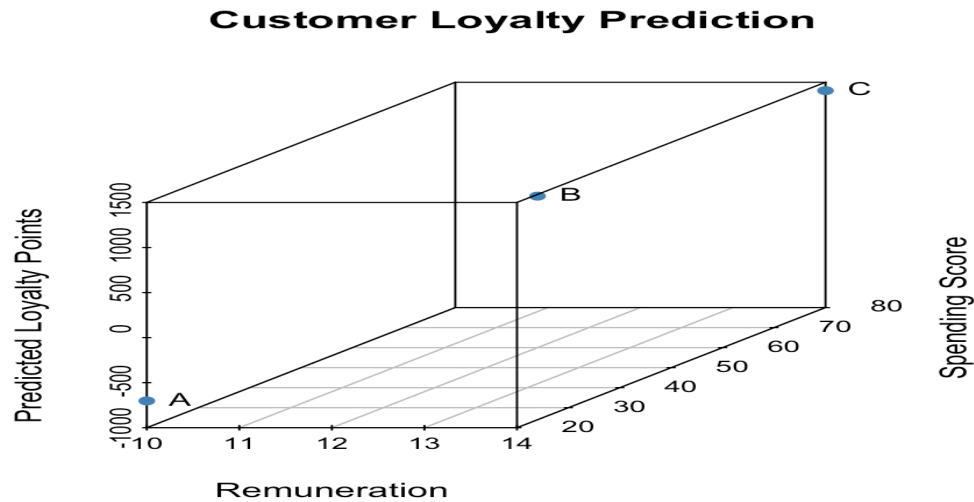
- Sentiment Analysis – Negative Review Insights



The sentiment analysis of negative reviews and summaries highlights recurring themes of customer dissatisfaction, with frequent mentions of terms such as “small,” “disappointed,” “boring,” and “confusing.” These patterns suggest that product size, clarity of gameplay instructions, and the overall customer experience may be falling short of expectations. For Turtle Games, these insights present an opportunity to address potential weaknesses in product design, packaging communication, and instructional materials. By proactively improving clarity, ensuring that size and quality meet customer expectations, and enhancing gameplay engagement, the company cannot only reduce the likelihood of negative feedback but also strengthen customer loyalty and brand reputation. Addressing these issues may help convert underwhelmed buyers into advocates, ultimately supporting higher retention rates and positive word-of-mouth promotion.

Patterns and predictions

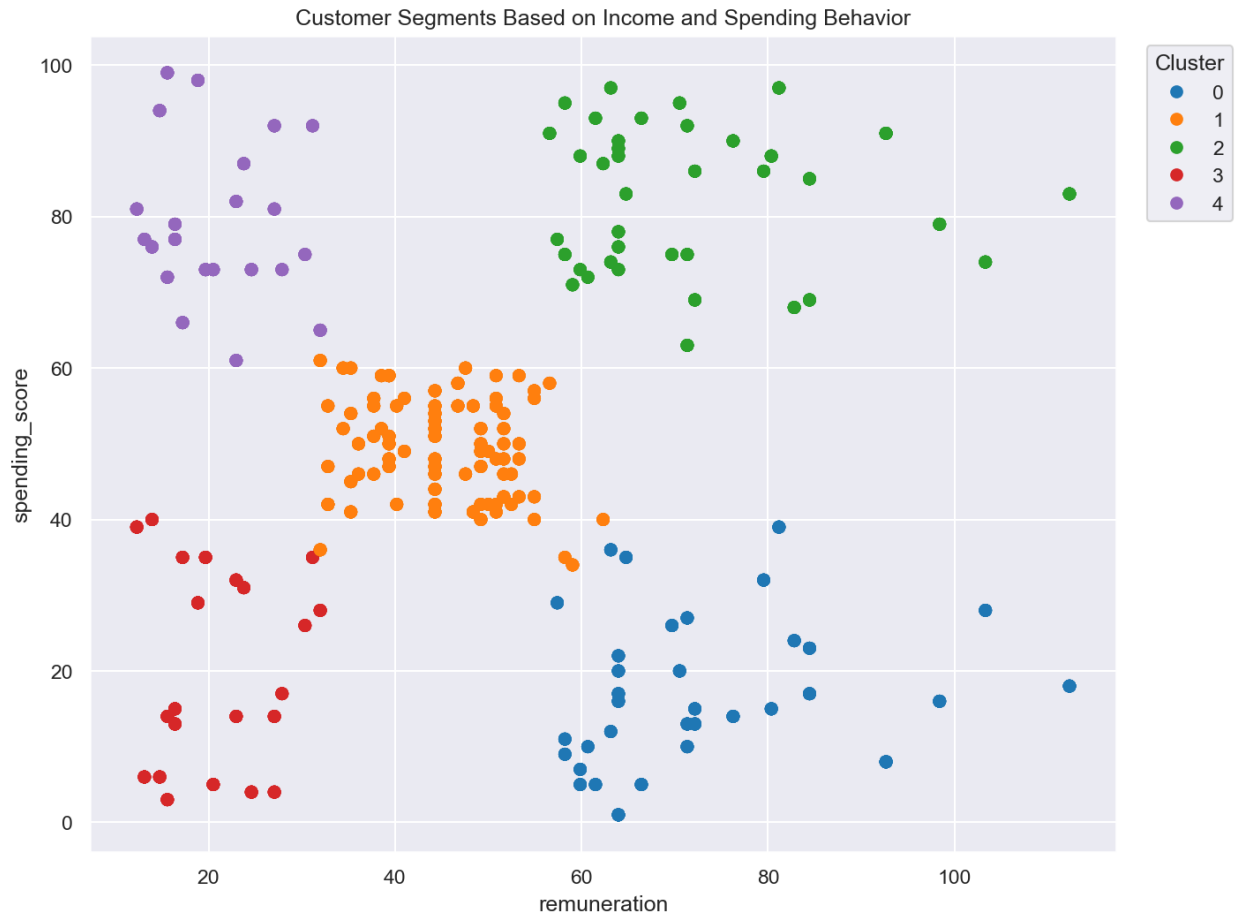
- Predicting Loyalty for Targeted Customer Strategies



Predicting Loyalty for Targeted Customer Strategies

Using a multiple linear regression model, loyalty points were predicted based on customers' remuneration and spending scores, revealing a strong positive relationship—higher income and higher spending are consistently linked to greater loyalty. A 3D visualization using three fictional profiles (A, B, and C) demonstrated how predicted loyalty shifts across different customer types, providing a clear visual link between spending capacity and engagement. This predictive capability gives Turtle Games a strategic advantage by allowing the identification of high-potential customers before loyalty declines. By leveraging these insights, the company can focus retention efforts on those most likely to generate long-term value, optimize marketing resources, and design tailored loyalty campaigns that deliver measurable business impact.

- **Customer Segmentation for Targeted Strategies (K-Means Analysis)**



K-means clustering was applied to segment customers into five distinct groups based on their remuneration and spending scores, uncovering clear behavioral patterns that extend beyond individual predictions. The analysis revealed diverse customer profiles, such as high-income, low-engagement customers who represent untapped potential, and lower-income customers who demonstrate strong loyalty despite limited spending power. Unlike regression, which predicts outcomes for individuals, clustering delivers group-level insights that enable Turtle Games to design targeted loyalty programs for underperforming segments while maintaining strong engagement with high-value groups. This segmentation empowers the company to align marketing strategies with actual behavioral patterns, driving more innovative retention efforts, maximizing campaign ROI, and fostering sustainable growth.

Conclusion & Recommendations

This analysis of Turtle Games' customer data combined Python and R techniques to deliver actionable insights across behavioral patterns, spending habits, and customer sentiment.

Key findings include:

- **Customer loyalty distribution** shows a broad base of under-engaged customers and a smaller group of high-value, highly loyal customers.
- **Demographic trends** reveal that loyalty is not solely tied to education level, challenging traditional assumptions about targeting.
- **Spending behavior analysis** highlights opportunities to engage high-income, low-loyalty customers with tailored campaigns.
- **Sentiment analysis** uncovers recurring product and experience concerns, such as size, clarity, and gameplay engagement, which can inform design improvements.
- **Predictive modeling** enables proactive targeting of high-potential customers based on income and spending profiles.
- **Customer segmentation** identifies distinct groups for precision targeting, ensuring loyalty strategies are matched to behavioral realities.

Strategic Recommendations:

- Launch retention campaigns for high-income, low-loyalty customers, using exclusive rewards or experiences.
- Address recurring product pain points by integrating sentiment findings into product development cycles.
- Tailor communication styles and offers to different education and demographic groups to maximize engagement.
- Maintain momentum with top-spending, high-loyalty customers by reinforcing value through VIP programs.
- Integrate clustering and predictive models into ongoing marketing analytics for continuous optimization.