Comprehensive Report: Player Segmentation and Persona Creation

---

Objective

The primary objective of this project was to segment players into distinct groups based on their behavior and preferences, define personas for each group, and provide actionable insights for targeted marketing and engagement strategies. Using a dataset containing game titles, discount levels, and review metrics, we analyzed player preferences and behaviors to create meaningful personas.

Data Preparation

1. Cleaning and Preprocessing

To ensure the dataset was ready for analysis:

- **Missing Values:**
  - Numerical columns: Missing values were replaced with the column mean to avoid skewing the data.
  - Categorical columns: Missing values were replaced with the mode.
- **Duplicates:**
  - Duplicate rows were removed to maintain data integrity.
- **Outlier Handling:**
  - Outliers in numerical columns were capped using the Interquartile Range (IQR) method to ensure fair clustering.

2. Feature Selection

Relevant features were selected to highlight key aspects of player behavior:

- **Sale Percentages:** Discount levels, such as `salepercentage_-75%` and `salepercentage_-50%`, to capture price sensitivity.

- **Review Sentiments:** Recent and overall review scores, such as `recentreviews_Overwhelmingly Positive` and `allreviews_Mostly Positive`, to gauge how players respond to game quality and feedback.

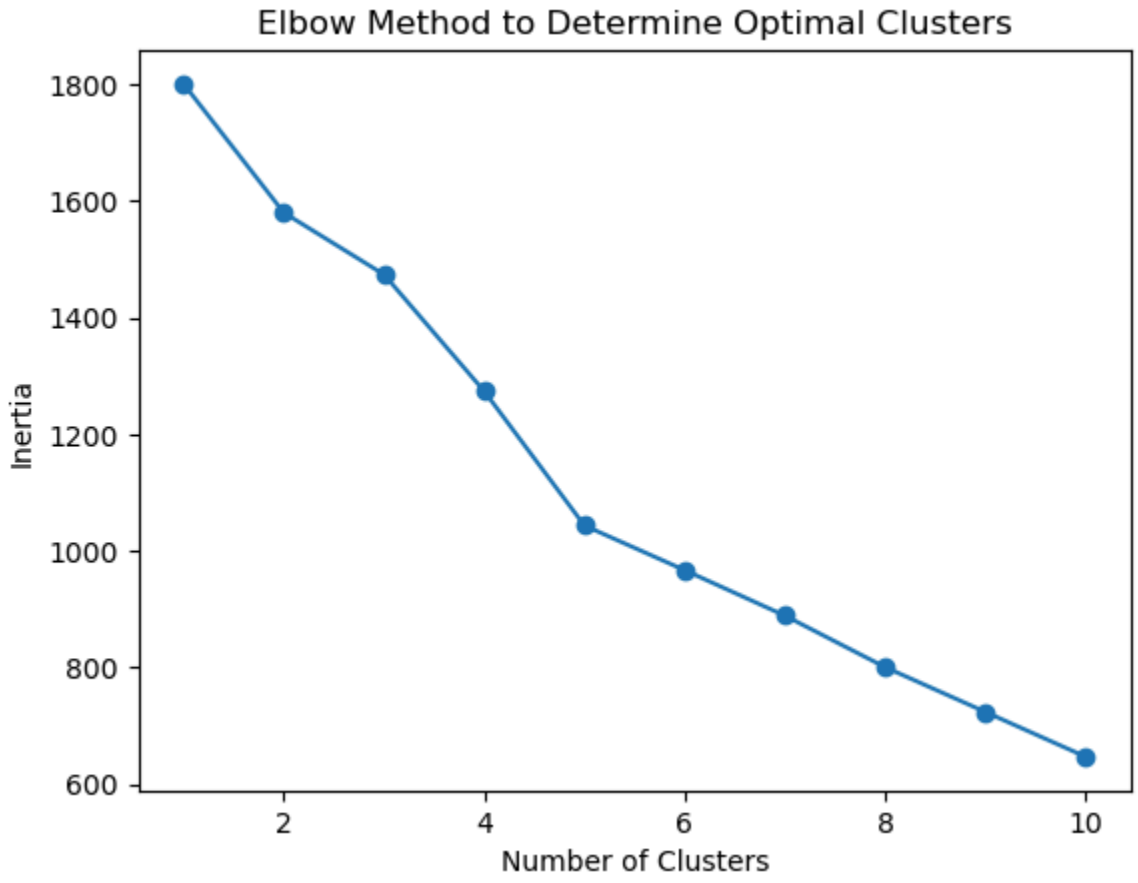
3. Feature Scaling

- Numerical features were normalized using the **MinMaxScaler** to bring all values to a comparable range, which is critical for clustering algorithms.

Clustering Methodology

1. Optimal Clusters Selection

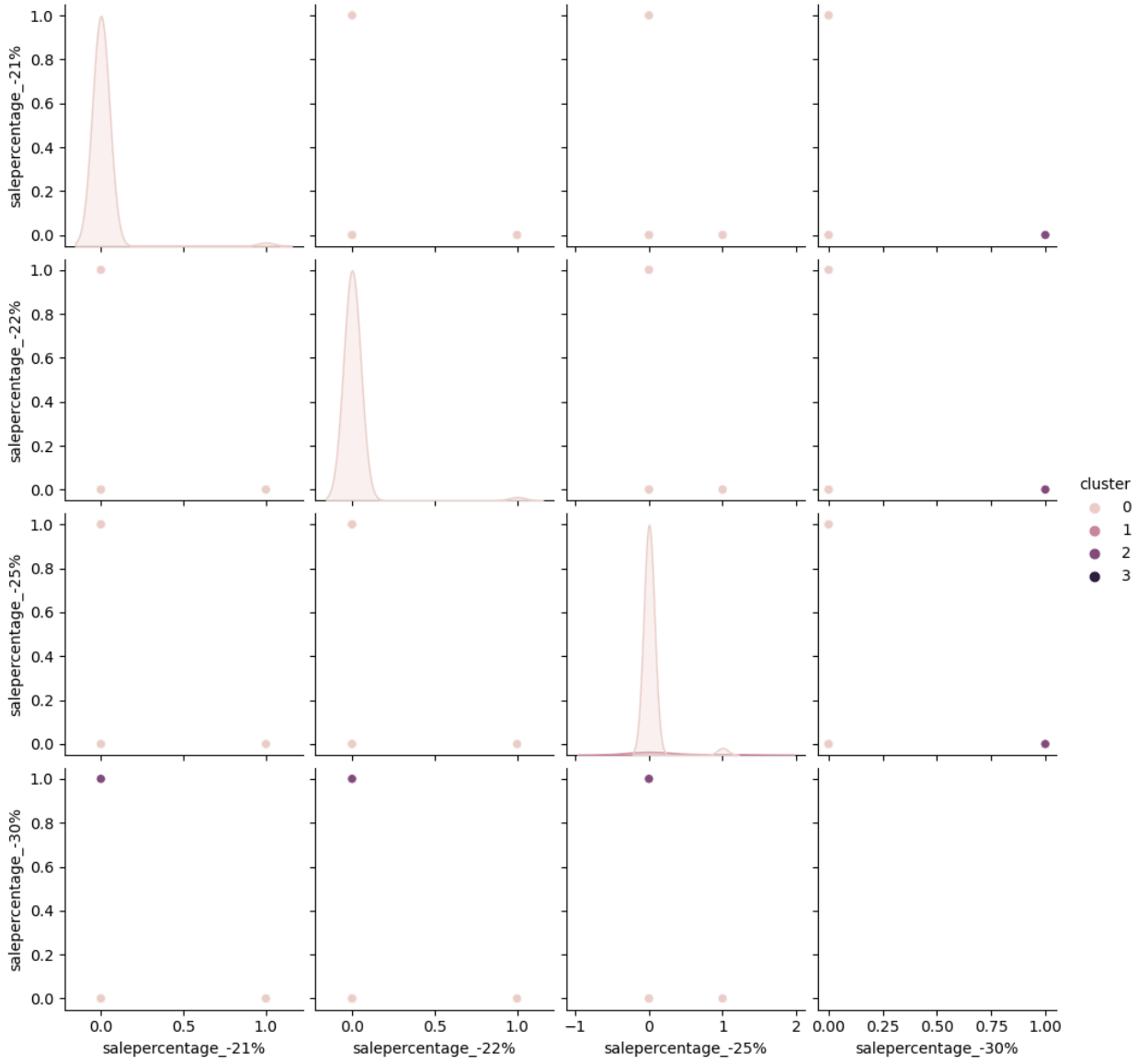
- The **Elbow Method** was used to determine the optimal number of clusters by plotting the inertia (sum of squared distances to cluster centroids) against the number of clusters.



- The optimal number of clusters was identified as **4**, providing a balance between interpretability and meaningful segmentation.

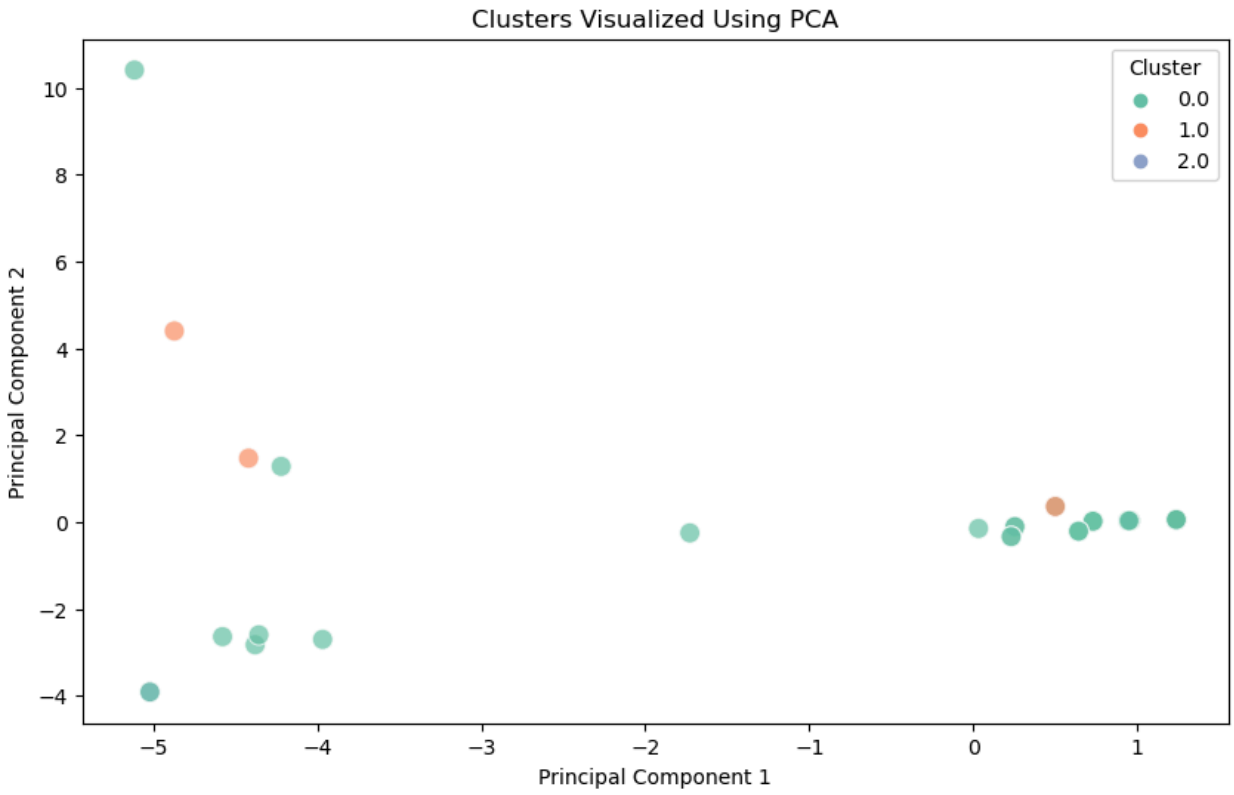
2. K-Means Clustering

- The **K-Means algorithm** was applied to segment players into 4 clusters based on selected features.



3. Principal Component Analysis (PCA)

- PCA was used to reduce the dimensionality of the dataset for visualization purposes.
- Two principal components captured approximately 30% of the variance, allowing for clear visualization of clusters in a 2D plot.



Cluster Analysis and Personas

Each cluster was analyzed to identify its unique characteristics and corresponding persona:

Cluster 0: Discount Seekers

- **Characteristics:**
  - Highest engagement with discounts, with high values for `salepercentage_-75%` and other discount-related features.
  - Moderate engagement with reviews, often indifferent to review sentiments.
- **Persona:**
  - Players primarily motivated by price reductions.
- **Strategic Recommendations:**
  - Offer frequent sales and highlight steep discounts in marketing campaigns.
  - Introduce "flash sales" to create urgency.
  - Promote game bundles with significant discounts.

Cluster 1: Positive Review Enthusiasts

- **Characteristics:**

    - Strong preference for highly rated games, with high values for `recentreviews_Overwhelmingly Positive`and `allreviews_Overwhelmingly Positive`.
    - Less sensitive to discounts.
- **Persona:**
    - Players who prioritize quality and rely on positive reviews.
- **Strategic Recommendations:**
    - Highlight critically acclaimed games and award-winning titles.
    - Use testimonials, review highlights, and quality-focused messaging in campaigns.
    - Emphasize community recognition and developer accolades.

Cluster 2: Mixed Sentiment Gamers

- **Characteristics:**
    - Engage with games that have polarizing reviews, such as `recentreviews_Mostly Negative` and `recentreviews_Mostly Positive`.
    - Likely driven by curiosity about niche or controversial games.
- **Persona:**
    - Players interested in exploring unique or polarizing games.
- **Strategic Recommendations:**
    - Market polarizing and niche games with curiosity-driven messaging, such as "Discover Hidden Gems" or "Games That Spark Debate."
    - Offer limited-time promotions to reduce perceived risk.
    - Encourage community discussions and content creation around these games.

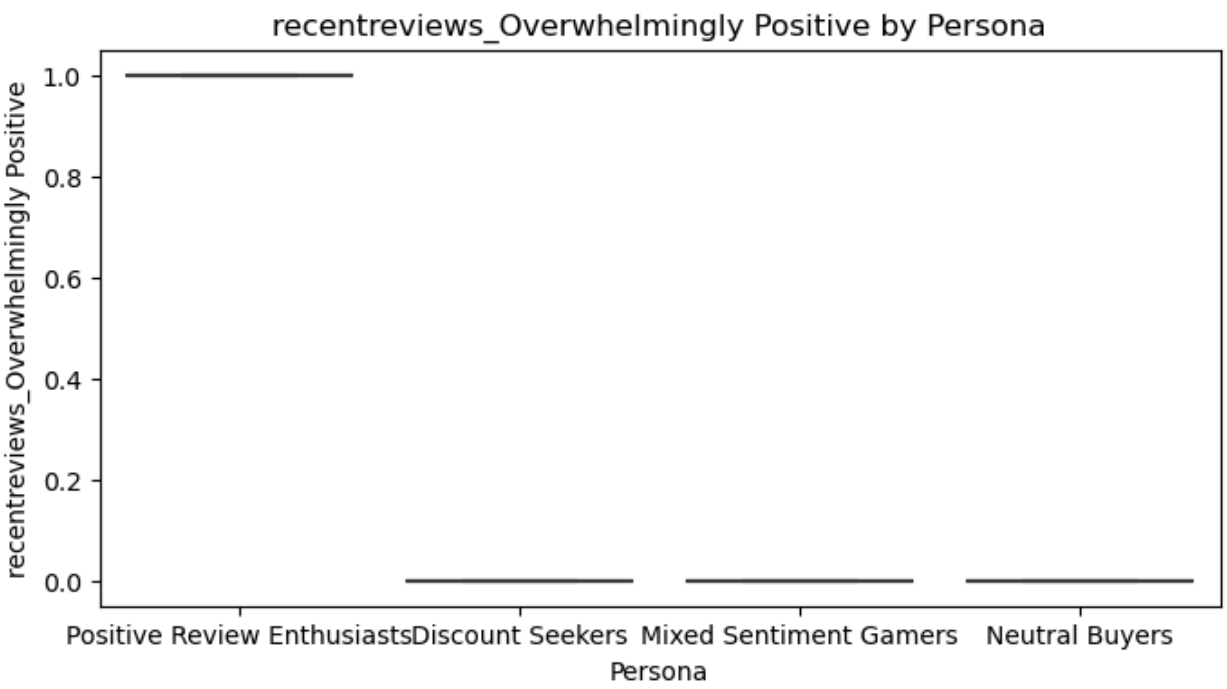
Cluster 3: Neutral Buyers

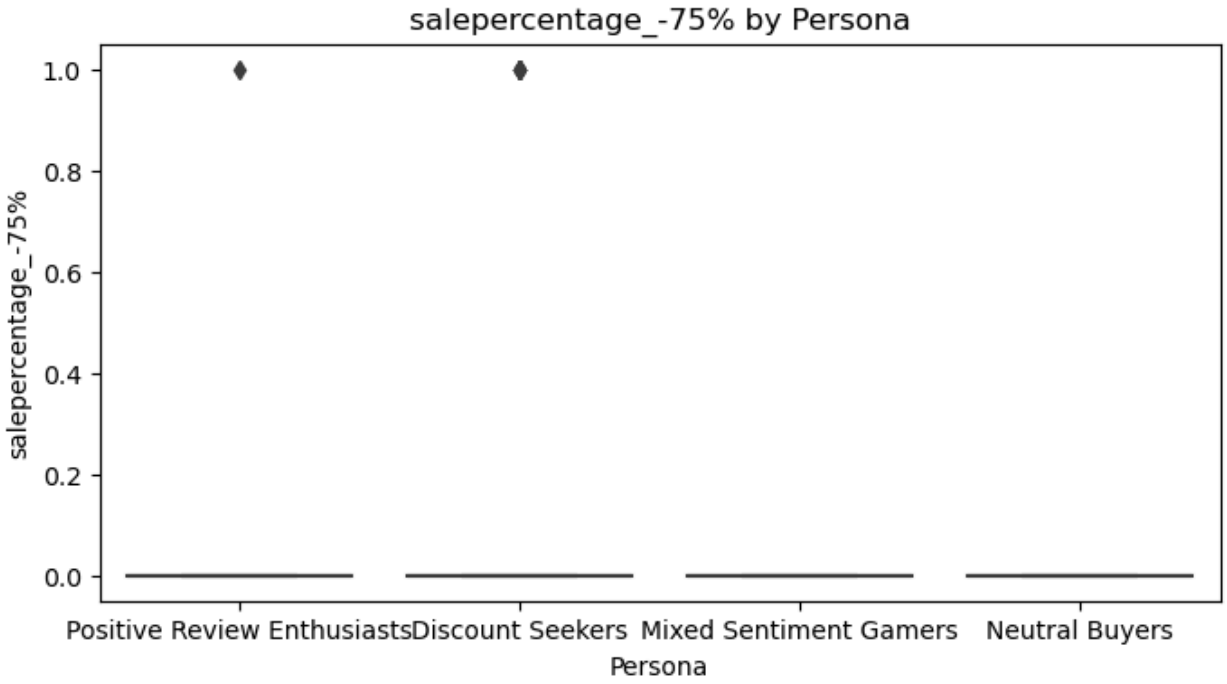
- **Characteristics:**
    - Balanced engagement across discounts and reviews.
    - Represent a general audience with no strong biases.
- **Persona:**
    - General audience with varied preferences.
- **Strategic Recommendations:**
    - Focus on broad marketing strategies, such as "Top Trending Games" or "Seasonal Favorites."
    - Introduce loyalty programs to encourage repeat purchases.
    - Pair new releases with related content, such as DLCs or merchandise.

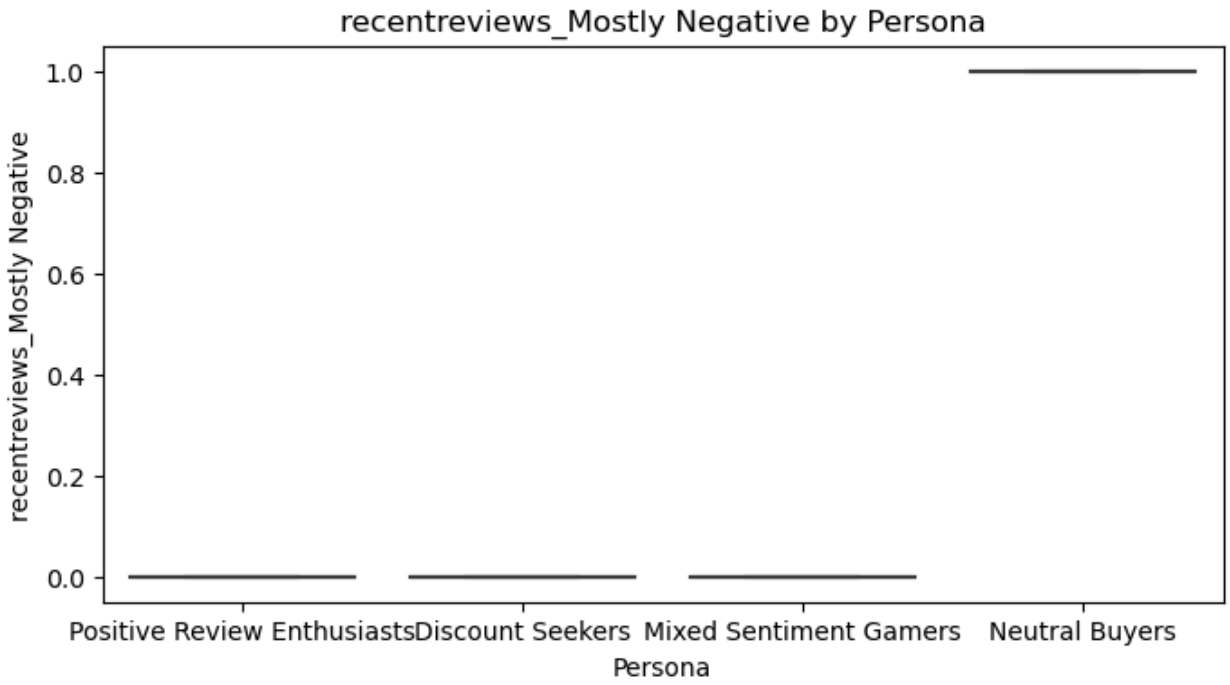
Visualizations

1. Cluster Feature Comparison

- A bar chart compared the mean values of key features (e.g., discounts and reviews) across clusters, highlighting distinct differences in behavior.
- **Insights:**
    - Discount Seekers had the highest sensitivity to price reductions.
    - Positive Review Enthusiasts engaged most with overwhelmingly positive reviews.


recentreviews_Overwhelmingly Positive by Persona
recentreviews_Overwhelmingly Positive
1.0
0.8
0.6
0.4
0.2
0.0
Positive Review Enthusiasts
Discount Seekers
Mixed Sentiment Gamers
Neutral Buyers
Persona


salepercentage_-75% by Persona
salepercentage_-75%
1.0
0.8
0.6
0.4
0.2
0.0
Positive Review Enthusiasts
Discount Seekers
Mixed Sentiment Gamers
Neutral Buyers
Persona



2. Persona Distribution

- A pie chart displayed the proportion of players in each persona:
    - **Discount Seekers** constituted the largest segment.
    - **Mixed Sentiment Gamers** formed the smallest but most unique group.

3. PCA Visualization

- A scatter plot visualized clusters in 2D space, showing clear separation between groups.

Targeted Strategy Recommendations

General Strategies for All Personas

- **Personalized Recommendations:**
    - Use machine learning to dynamically recommend games based on player segmentation.
- **Social Proof:**
    - Highlight reviews and community feedback to enhance trust and interest.
- **Seasonal Campaigns:**
    - Leverage key shopping periods (e.g., Black Friday, holiday sales) to boost engagement across personas.

Cluster-Specific Strategies

1. **Discount Seekers:**
    - Regularly offer significant discounts and bundle deals.
    - Promote loyalty-based incentives, such as points for purchases.
2. **Positive Review Enthusiasts:**
    - Create quality-driven campaigns featuring top-rated games.
    - Showcase developer achievements and awards.
3. **Mixed Sentiment Gamers:**
    - Highlight niche and polarizing games with curiosity-driven messaging.
    - Use limited-time offers to encourage exploration.
4. **Neutral Buyers:**
    - Focus on general trends and broad recommendations.
    - Offer subscription-based access to a variety of games.

Conclusion

This player segmentation analysis provided actionable insights into player behavior and preferences. By leveraging these insights, businesses can:

- Design tailored marketing strategies.
- Enhance user engagement.
- Maximize revenue through targeted promotions.

Further refinements, such as real-time personalization and advanced predictive modeling, can enhance the effectiveness of these strategies.