Chapter 5

Latent Similarity Clustering of Video Games Based on Euclidean Distance and PCA

Diana Bratić[1]

Abstract

This paper presents a multi-criteria similarity analysis of video games using quantitative variables from an available dataset. The research includes the following variables: user rating, number of recommendations, average playing time (overall and in the last two weeks), and percentage of positive reviews. The research aims to develop a similarity model for games in a multidimensional space defined by these attributes and to identify patterns and groupings based on their quantitative profiles. The data was standardized to ensure comparability across variables with different scales. Euclidean distance was used to measure similarity between games, as it is intuitively interpretable in real space: the distance between two games is calculated as the square root of the sum of squared differences across all dimensions. This metric enables accurate positioning of games within the attribute space and forms the basis for hierarchical clustering. Principal component analysis (PCA) was applied to reduce dimensionality and facilitate visual interpretation of the results.

Preliminary findings indicate the existence of several stable clusters, including games with high ratings and recommendations but relatively short playing time, as well as a group of games played extensively but rated lower by users. These combinations suggest distinct usage patterns and perceived value, which are not directly aligned with traditional categories such as genre or publisher. The approach presented in this study can serve as a foundation for structuring large-scale game datasets and as a starting point for developing classification and recommendation algorithms based on objective rather than subjective product characteristics.

1 University of Zagreb, Faculty of Graphic Arts, Department of Printing Processes, 10 000 Zagreb, Croatia

https://doi.org/10.58830/ozgur.pub865.c3503

INTRODUCTION

This paper presents a multi-criteria similarity analysis of video games based on quantitative variables from a publicly available dataset. The analysis includes the following variables: user rating, number of recommendations, average playtime (overall and in the last two weeks), and percentage of positive reviews. The objective is to develop a similarity model in a multidimensional space defined by these attributes and to identify patterns and groupings based on their quantitative profiles.

The data was standardized to ensure comparability across variables with different scales. Euclidean distance was used to measure similarity between games, as it is directly interpretable in real space: the distance between two games is computed as the square root of the sum of squared differences across all dimensions. This metric provides precise positioning of games in the attribute space and serves as the basis for hierarchical clustering. Principal Component Analysis (PCA) was applied to reduce dimensionality and support visual interpretation of the results.

Preliminary results indicate the presence of several stable clusters, including games with high ratings and recommendations but relatively short playtime, as well as games with extensive playtime but lower user ratings. These configurations reflect distinct usage patterns and perceived value, which are not directly aligned with traditional categories such as genre or publisher. The proposed approach may serve as a foundation for structuring large-scale video game datasets and as a starting point for developing classification and recommendation algorithms based on objective behavioral attributes rather than subjective labels.

To contextualize this approach, the next section reviews recent studies that operationalize game similarity using behavioral data and explores methodological developments in this domain. In existing literature, similarity between video games is commonly defined through genre classification, thematic elements, game mechanics, or target user groups. Traditional models rely on genre labels and content descriptions as a basis for categorization, though these methods are increasingly criticized for their lack of objectivity and inconsistency across platforms. Due to the subjective nature of genre labeling, there is a growing demand for similarity frameworks grounded in measurable, quantitative indicators of user interaction.

Recent studies have sought to operationalize similarity using behavioral data such as total playtime, session counts, recommendations, or user ratings. These approaches bypass predefined categories and instead use actual user behavior to reveal latent structures not captured by standard classifications. Xu et al. (2023) introduced the DUIP model, which integrates LSTM and LLM techniques to predict preferences based on behavioral sequences. Their results on the "Games" dataset suggest that interaction-based dynamics outperform traditional content- and collaboration-based filters. Qiu et al. (2024) similarly focus on behavioral modeling by combining temporal and static features to analyze mobile MMO game users. Using methods such as Time Series K-means, DeepWalk, and LINE, they identify five distinct clusters and link network properties to customer retention. Their findings validate the effectiveness of multi-component behavior analysis for personalization and engagement prediction. In a related study, Zhou et al. (2021) propose the TCBS algorithm, which blends k-means clustering with a smoothed sequence of Euclidean distances to capture behavioral patterns in casual games. The study highlights that the choice of distance metric (e.g., ED, DTW, CORT) critically influences clustering outcomes, reinforcing the need for careful selection of dissimilarity functions in behavioral analysis.

Complementary evidence is provided by group of authors (Calvo-Morata et al., 2025), who use learning analytics to refine the design and evaluation of serious games. They analyze behavioral traces from 134 high school students, focusing on task duration, number of attempts, and error rates. Their findings reveal gender-based strategy differences and confirm the value of quantitative interaction data in content adaptation, performance assessment, and evidence-based development. Similarly, other group of authors (Lu et al., 2023) propose a framework for feature engineering based on telemetry data from serious games. By examining the gameplay of 373 students, they identify key behavioral patterns linked to knowledge levels, including task completion speed, interaction diversity, and key event frequency. The authors emphasize the importance of aligning game design with analytical goals to maximize the utility of behavioral data in modeling engagement and learning. Freire et al. (2016) extend this approach through the Game Learning Analytics (GLA) framework, which uses event-based logs to derive rich behavioral features for content personalization and design validation. Kang et al. (2017) further support this view, demonstrating that task sequences and activity duration can uncover learning strategies and inform serious game design through detailed behavioral profiling. Finally, Normoyle and Jensen (2015) apply a non-parametric Bayesian model in an FPS context to identify overlapping affiliations across latent behavioral styles. Their approach, which integrates sequential behavior with hierarchical inference, yields stable and interpretable classifications that move beyond traditional categorical models. Using real-world Battlefield 3 data, they validate the role of behavioral metrics in uncovering hidden dimensions of user interaction.

Furthermore, the analysis of behavioral patterns across datasets with mixed data types, such as numerical metrics and categorical labels, requires careful selection of both dissimilarity measures and clustering algorithms. One proposed solution combines the Gower coefficient, the Winsor-Huber loss function for numeric variables, and the entropy-based distance for categorical attributes, achieving improved efficiency in segmenting users and predicting churn in mobile games (Perišić and Pahor, 2022). Although their work primarily focuses on churn prediction, the underlying methodological framework exhibits strong transferability to behavior-driven clustering tasks, including latent similarity modeling in video games.

Modeling similarity between games also faces technical challenges due to the heterogeneity of development tools and production pipelines. A functional classification of game development tools comprising user interaction, content creation, and integration components has been proposed as a way to better understand this complexity (Toftedahl and Engström, 2019). As the term *game engine* often encapsulates a variety of unrelated systems, consistent categorization based on technical features alone proves insufficient, reinforcing the need for models grounded in usage data. A more recent contribution by Grelier and Kauflan (2023) demonstrates the potential of combining behavioral features with metadata to automatically cluster and semantically tag video game groups. By integrating optimized UMAP projection, hierarchical clustering, and semantic analysis of game titles, their model identifies meaningful groupings within a dataset of over 30,000 games. These groupings, named according to shared themes and functional properties, enable enhanced content discovery, recommendation, and market trend analysis without reliance on predefined genre labels.

Such examples illustrate how multi-criteria similarity frameworks can integrate diverse quantitative dimensions without depending on static descriptors like genre or publisher. Behavior-based models can reveal natural clusters that more accurately reflect actual engagement patterns and user preferences. Support for this approach also comes from research into esports consumption. A study on esports audiences identified four distinct user profiles: all-round gamers, conventional gamers, observers, and recreational gamers, based on both playing and viewing behaviors (Jang et al., 2021). The findings highlight how multiple forms of engagement, including hardware ownership and content consumption, jointly shape user segmentation, underscoring the value of multidimensional models in understanding user behavior in digital games.

Taken together, the reviewed literature supports the validity of multi-criteria similarity modeling grounded in user behavior, offering a pathway to uncover latent structures of similarity that transcend traditional genre classifications and technical taxonomies.

METHODOLOGY

This study employs a quantitative approach to multidimensional similarity analysis of video games based on user behavior. Design of data structuring and transfer aligns with standard networked system models (Kurose and Ross, 2016). The selected dataset consists of 1,000 video games sourced from open aggregation platforms, chosen based on data completeness and high distributional diversity. The sample size reflects a balance between representativeness and the computational demands of the applied methods. Specifically, the quadratic time complexity of hierarchical clustering, denoted as $O(n^2)$, was taken into consideration. For larger datasets, this complexity would necessitate a shift toward more scalable but less interpretable methods, such as *K*-means or DBSCAN.

Variables and Normalization

In order to analyze latent patterns of similarity between video games, five behavioral variables were defined to quantify key aspects of user interaction. The selection of variables is based on their relevance to evaluating engagement, user recommendations, and perceived game quality. All data were obtained from the publicly accessible Steam Games Dataset available on the Kaggle platform.

The data sample comprises $n = 1000$ unique video games, each described by five variables as follows:

- score rank (x_1): a numerical aggregation of user scores on a scale from 1 to 10, normalized according to preference distribution,
- recommendations (x_2): the total count of positive recommendations made by users within the community,
- average playing time forever (x_3): the cumulative average duration (in hours) users have spent in the game,
- average playing time two weeks (x_4): short-term engagement, measured using the same time unit,

- percentage of positive reviews (x_5): the proportion of positive ratings relative to the total number of user reviews.

Each variable represents a distinct behavioral dimension, enabling the development of a multidimensional similarity model grounded in actual usage patterns rather than static genre classifications. Together, these measures define distinct dimensions of behavior and support the construction of a multidimensional similarity model based on actual content usage rather than static genre labels.

To align the scales of all variables and enable fair comparison, *z*-score standardization was applied, converting each value into its standardized deviation from the mean:

$$z_{ik} = \frac{x_{ik} - \mu_k}{\sigma_k}$$

where x_{ik} is the value of the *k*-th variable for the *i*-th game, μ_k is the mean of that variable, and σ_k is its standard deviation. This transformation results in a matrix representation $\boldsymbol{Z} \in \mathrm{R}^{n \times 5}$, standardized with respect to central tendency and dispersion ($\mu = 0, \sigma = 1$), thereby ensuring metric neutrality in both distance calculations and projection analyses.

Similarity Measure

To quantify the dissimilarity between pairs of instances, the standard Euclidean distance was used, defined as the l_2 norm of the difference between feature vectors in the five-dimensional space R^5:

$$d(i, j) = \|\boldsymbol{z}_i - \boldsymbol{z}_j\|_2 = \sqrt{\sum_{k=1}^{5} (\boldsymbol{z}_{ik} - \boldsymbol{z}_{jk})^2}$$

where $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$ are five-dimensional vectors of standardized behavioral characteristics for games *i* and *j*. This metric assumes orthogonality of dimensions and exhibits high sensitivity to outliers, which can highlight deviant patterns in user behavior.

In addition, a dissimilarity matrix is generated as output:

$$\boldsymbol{D} \in \mathbb{R}^{n \times n}, \quad \boldsymbol{D}_{i,j} = d(i, j)$$

which fulfills all axioms of a metric space and serves as the input for all subsequent analytical procedures, including dimensionality reduction and hierarchical clustering.

The equivalent vectorized form is used to efficiently compute the complete distance matrix:

$$\boldsymbol{D} = \sqrt{(\mathrm{diag}(\boldsymbol{ZZ})^{\top} \cdot \boldsymbol{v}^{1\top} \cdot \mathrm{diag}(\boldsymbol{ZZ}^{\top})^{\top} - 2\boldsymbol{ZZ}^{\top})}$$

where $\boldsymbol{Z} \in \mathrm{R}^{n\times 5}$ is the matrix of standardized features and $\boldsymbol{v} \in \mathrm{R}^{n}$ is the unit vector. This enables numerically efficient computation of the matrix $\boldsymbol{D}$ without explicit pairwise iteration.

Dimensionality reduction with PCA

To identify latent patterns of variation and reduce redundant dimensions within the standardized behavioral space R^5, principal component analysis (PCA) was applied, which is based on the spectral decomposition of the empirical covariance matrix.

Let $\boldsymbol{Z} \in \mathrm{R}^{n\times 5}$ be the $\boldsymbol{Z}$-score matrix of the standardized traits, with each row representing one game and the columns representing five behavioral variables. The covariance matrix $\boldsymbol{C} \in \mathrm{R}^{5\times 5}$ is then calculated as follows:

$$\boldsymbol{C} = \frac{1}{n-1}\boldsymbol{Z}^{\top}\boldsymbol{Z}$$

Spectral decomposition of the matrix $\boldsymbol{C}$ is performed by solving the eigenvalue equation $\boldsymbol{C}_{vk} = \lambda_k \boldsymbol{v}_k$, where $\lambda_k \in \mathrm{R}$ are the eigenvalues (ordered in descending magnitude) and $\boldsymbol{v}_k \in \mathrm{R}^5$ are the corresponding orthonormal eigenvectors. This yields the following factorization:

$$\boldsymbol{C} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\top}$$

where $\boldsymbol{V}$ is the matrix of principal axes (eigenvectors), and $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_5)$ is the diagonal matrix of the variances of the principal components. The projection of the original data space onto a reduced-dimensional subspace R^q, where $q < 5$, is calculated as:

$$\boldsymbol{Z}^{(q)} = \boldsymbol{Z}\boldsymbol{V}_q$$

where $\boldsymbol{V}_q \in \mathrm{R}^{5\times q}$ is the matrix of the q most dominant eigenvectors. In this study, the first two components were selected ($q = 2$), which together explain more than 90% of the total variance, enabling a topologically faithful visualization of the data structure and the detection of clustering configurations.

PCA is used here not only as a means of dimensionality compression and noise filtering but also as a mapping technique into a latent space with maximum inertia, which facilitates the interpretation of geometric relationships between the games. This approach is consistent with established practices in multivariate data analysis (Hair et al., 2019). Dimensionality reduction was subsequently applied to the dissimilarity matrix to ensure consistency with the previously used metric structure.

Implementation in Python

All data processing phases were carried out using the Python programming language (version 3.11), along with the following libraries:

- pandas: for data import, filtering, and preparation of table structures (*pandas.read_csv*, *DataFrame.loc*, *dropna*),
- numpy: for performing numerical operations and matrix-based transformations (*numpy.array*, *numpy.mean*, *numpy.std*, *numpy.dot*),
- scikit-learn: for standardizing features using the z-score method (*StandardScaler*), applying principal component analysis (*PCA*), and performing hierarchical clustering based on distances (*AgglomerativeClustering*),
- scipy.cluster.hierarchy: for implementing Ward's clustering method and constructing dendrograms (*linkage*, *dendrogram*),
- matplotlib.pyplot: for creating scatter plots, dendrograms, and arranging visual elements (*scatter*, *subplots*, *tight_layout*),
- seaborn: for advanced statistical visualizations such as heatmaps and boxplots (*heatmap*, *boxplot*, *clustermap*).

The implementation strategy draws on practical principles outlined in recent literature on applied deep learning (Howard and Gugger, 2020; Elgendy, 2022).

The standardized variables were used to compute principal component scores (*pca_scores*) and clustering outcomes (*labels_here*, *linkage_matrix*). The code is implemented in the script *anomaly_detection.py* and further documented in the notebook *visual_mapping.ipynb*. This methodological structure establishes a coherent analytical pathway from behavioral variable selection to dimensionality reduction and cluster formation, thereby enabling the identification of latent similarity patterns which are further examined in the following section.

RESULTS AND DISCUSSION

This chapter presents the key results of the analysis, along with their interpretation in relation to the methodological framework and research objectives. The analysis begins with the examination of principal components obtained through principal component analysis (PCA), followed by the results of hierarchical clustering based on the reduced feature space. Visualizations are used to illustrate patterns observed in the dataset, while the discussion addresses the main findings, limitations, and potential implications of the results

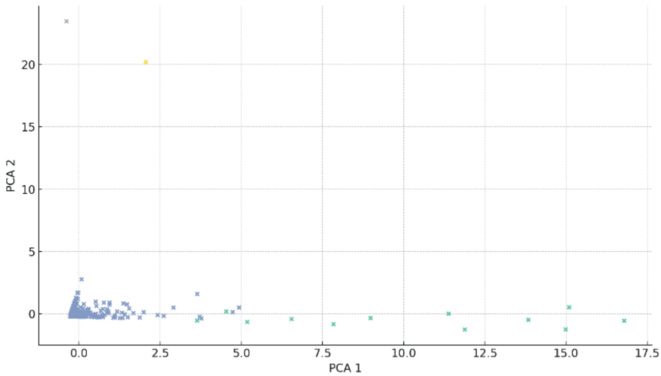

Figure 1. PCA Scatter Plot showing projection of games in two-dimensional space using principal components

Based on the results shown in Figure 1, the spatial distribution of video games in a two-dimensional space can be observed, derived through principal component analysis (PCA). Each point in the plot corresponds to a single game, and the distances between points reflect behavioral similarity based on the five standardized variables. Several distinct groups can be identified, indicating the presence of clusters with similar characteristics. In addition to the main groupings, several outlying points deviate from the overall structure, suggesting the presence of potentially anomalous games. These outliers, which do not align with prevailing behavioral patterns, are of particular interest for further analysis as they may represent atypical titles with unique properties.

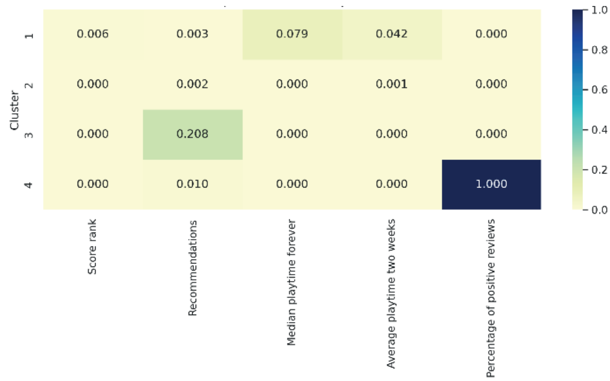

Figure 2. Heatmap of standardized mean values for each variable across identified clusters

Figure 2 presents a heat map displaying the standardized mean values of the five behavioral variables for each identified cluster. The rows correspond to individual clusters, while the columns represent the variables: *score rank, recommendations, mean playtime forever, mean playtime two weeks*, and *percentage of positive reviews*. Darker shades indicate higher average values within a cluster for a given variable. Cluster 3 stands out with a notably higher number of recommendations, while Cluster 4 is characterized by a dominant proportion of positive ratings. This visualization enables a concise comparison of behavioral profiles across clusters and supports their interpretation based on aggregated patterns.

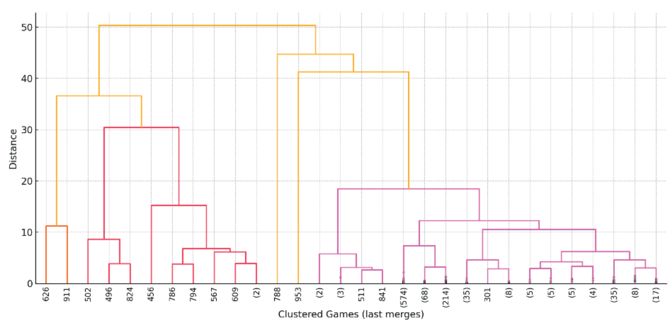

Figure 3. Dendrogram of hierarchical merging of games using Ward's method

The dendrogram in Figure 3 illustrates the hierarchical clustering of video games using Ward's method, where individual games and emerging clusters are iteratively merged based on similarity. Shorter branches at the lower levels reflect games with highly similar behavioral profiles, while longer branches at higher levels represent the union of more distinct groups. This visualization offers a detailed view of the clustering process underlying the PCA projection and reveals the relative proximity between clusters. The resulting structure supports the interpretation of cluster cohesion and separation within the reduced feature space.

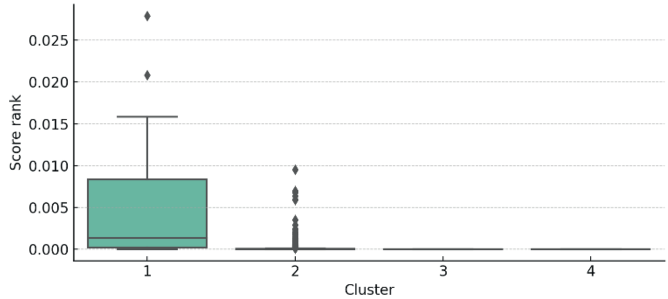

Figure 4. Boxplot of score rank distributions across the four clusters

Figure 4 displays the distribution of the score rank variable across the clusters using a box plot, which shows the median, interquartile range and extreme values. Cluster 1 has the widest range of scores, including several high-ranking outliers, which indicates substantial internal variability as some games in this group are highly rated while others are not. In contrast, clusters 2, 3 and 4 show uniformly low and tightly grouped scores with minimal variation. This pattern suggests that the games in these clusters are generally lower-ranking titles. The observed distribution confirms that score rank is a key differentiating factor, especially in defining cluster 1.

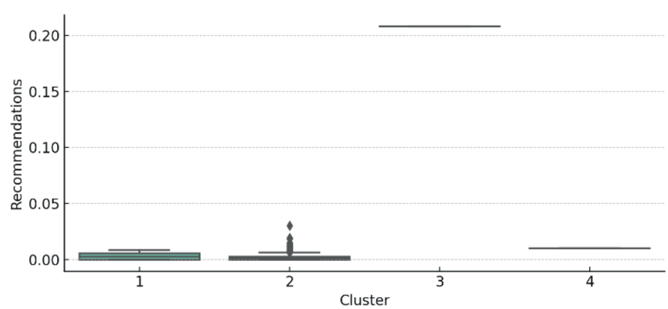

Figure 5. Boxplot of number of recommendations per game in each cluster

Figure 5 presents the distribution of the number of recommendations across clusters using a box plot. Cluster 3 stands out clearly, as it contains games with exceptionally high frequencies of recommendation compared to all other clusters. The values in this cluster are well above the overall average, pointing to strong visibility within recommendation systems. Clusters 1 and 2, apart from a few isolated cases, show consistently low recommendation counts. Cluster 4 is characterized by a moderate but steady level of recommendations without significant deviations. This distribution indicates that games in Cluster 3 are shared or promoted considerably more often than those in the remaining clusters.

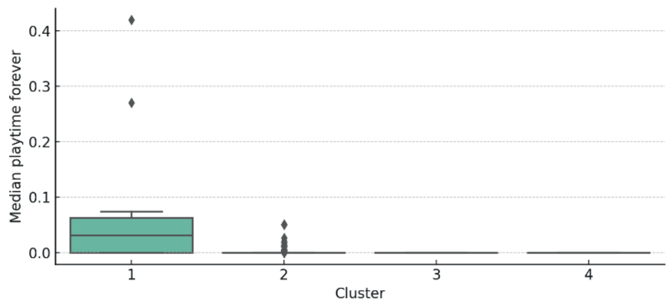

Figure 6. Boxplot of median total playtime forever by cluster

The box plot in Figure 6 shows the distribution of the median total playing time across clusters. Cluster 1 exhibits a notably higher median and a broader range of values compared to the other clusters, including several prominent outliers. This suggests that players associated with this group tend to spend more time in gameplay overall. In contrast, Clusters 2, 3, and 4 show consistently low and narrowly distributed values, indicating that games in these clusters generally experience limited long-term user engagement.

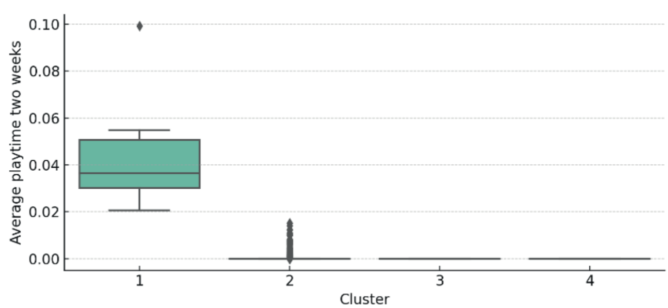

Figure 7. Boxplot of median total playtime two weeks by cluster

Figure 7 shows the average play time in the last two weeks for each cluster. Cluster 1 clearly stands out with higher values and greater variability, indicating active and ongoing user engagement within this group. The remaining three clusters exhibit consistently low values, suggesting that the games in these groups are currently played infrequently or lack a stable user base. This metric is useful for identifying titles with sustained short-term activity.

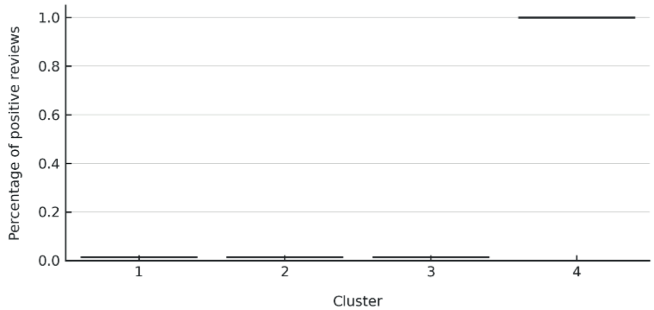

Figure 8. Boxplot of percentage of positive user reviews across clusters

The boxplot in Figure 8 shows the distribution of the variable *percentage of positive reviews* across the clusters. Cluster 4 stands out clearly, with all games in this group having exclusively positive ratings (value 1.0). In contrast, the other clusters display values near zero, with minimal variation. This suggests that the games in cluster 4 represent niche titles that are exceptionally well received, whereas the remaining groups consist of games with predominantly negative or mixed feedback. Including this variable adds a sentiment dimension to the interpretation of user behavior.

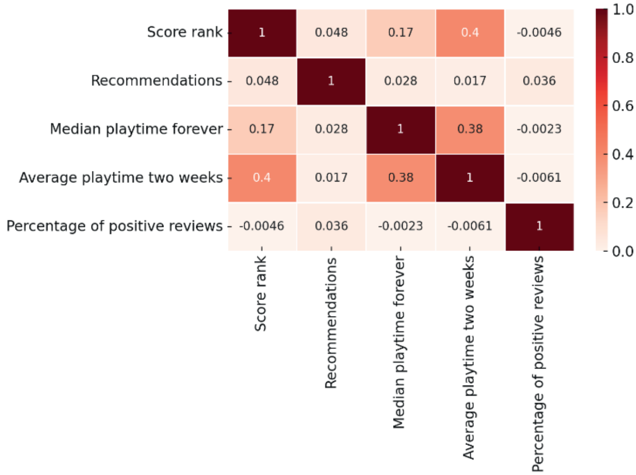

Figure 9. Correlation Matrix of All Five Standardized Input Variables

Figure 9 shows the correlation matrix in the form of a heat map between five key variables: *score rank, recommendations, average playtime forever, average playtime for two weeks, and percentage of positive reviews*. The visualization presents the strength of the linear Pearson correlation using a color gradient, where darker shades indicate stronger positive associations. The strongest correlation is observed between median *playtime forever* and *average playtime for two weeks* (r = 0.38), suggesting that users who have played a game extensively in the past continue to play it actively. A moderate correlation is found between *score rank* and *average playtime for two weeks* (r = 0.4), indicating that higher-ranked games may attract more recent user engagement. In contrast, the variable *percentage of positive reviews*, representing the share of positive user ratings, shows almost no correlation with the other variables, implying that user satisfaction is not directly tied to engagement metrics. These findings support the need for a multi-criteria analytical framework to adequately capture user behavior.

Taken together, the findings highlight distinct behavioral patterns across game clusters and reinforce the value of combining dimensionality reduction with hierarchical methods to uncover latent structures and interpret divergence within the dataset.

CONCLUSION

The results show that it is possible to group video games based on user behavior variables. Hierarchical clustering enabled the identification of stable and interpretable engagement patterns, independent of genre or publisher. The resulting clusters reflect latent structures derived from actual usage rather than declarative categories. The application of principal component analysis (PCA) allowed for a two-dimensional visualization of multidimensional similarities between games, making the relationships between clusters more apparent. It was observed that behaviors related to recommendations and playtime did not always align, indicating different forms of user engagement.

The detected patterns confirm the feasibility of objective, data-driven modeling of product similarity. The resulting similarity maps may serve as a basis for developing recommendation systems and content discovery tools. In addition, the clusters can be used to train supervised models or refine existing taxonomies of digital products. Since the approach is based solely on behavioral data, it can be applied across languages, markets, and game formats. Future work could include the integration of temporal behavioral patterns, session-based metrics, and genre classifications, along with the exploration of supervised models for improved anomaly classification. The method is repeatable and scalable, offering a robust framework for product classification based on real-world usage patterns. Expanding the dataset may further increase the model's robustness and generalizability.

The Python script used for the analysis is available on request.

References

Calvo-Morata A et al., 2025. Learning Analytics to Guide Serious Game Development: A Case Study Using Articoding, Computers, 14(4): 122.

Elgendy M, 2020. Deep Learning for Vision Systems. 1st ed., Manning Publications Co, USA.

Freire M et al, 2016. Game Learning Analytics: Learning Analytics for Serious Games. Learning, Design, and Technology. (Editors: Spector M, Lockee B, Childress M). Springer, pp.1-29, Germany.

Grelier N, Kaufmann S, 2023. Automated clustering of video games into groupswith distinctive names. (Editors: Figueroa P, Di Iorio A, Guzman del Rio D, Gonzalez Clua EW, Cuevas Rodriguez L) Entertainment Computing – ICEC 2024. Lecture Notes in Computer Science, Vol 15192. Springer, pp.223-231, Germany.

Hair JF et al., 2019. Multivariate Data Analysis: A Global Perspective. 7th ed., Pearson, USA.

Howard J, Gugger S, 2020. Deep Learning for Coders with fastai & PyTorch, 1st ed., O´Reilly, USA.

Jang WW et al., 2021. Clustering Esports Gameplay Consumers via Game Experiences, Frontiers in Sports and Active Living, 3(1): 1-12.

Kang J, Liu M, Qu W, 2017. Using gameplay data to examine learning behavior patterns in aserious game, Computers in Human Behavior, 72(1): 757-770.

Kurose JF, Ross KW, 2016. Computer Networking: A Top-Down Approach. 7th ed., Pearson, USA.

Lu W et al., 2023. Serious Game Analytics by Design: Feature Generation and Selection Using Game Telemetry and Game Metrics: Toward Predictive Model Construction, Journal of Learning Analytics, 10(1): 168-188.

Normoyle A, Jensen ST, 2021. Bayesian Clustering of Player Styles for Multiplayer Games. 11th Artificial Intelligence and Interactive Digital Entertainment Conference, November 14-18, 2015, 163-169, Santa Cruz, CA, USA.

Perišić A, Pahor M, 2022. Clustering mixed-type player behavior data for churn prediction in mobile games, Central European Journal of Operazions Research, 31(1): 165-190.

Qiu Y, Gong Y, Liu G, 2024. User Behavior Analysis and Clustering in a MMO Mobile Game: Insights and Recommendations, Preprint, 2024: 1-16.

Toftedahl M, Engström H, 2019. A Taxonomy of Game Engines and the Tools that Drive the Industry. DiGRA 2019 Conference: Game, Play and the Emerging Ludo-Mix, August 6-10, 2019, 1-17, Kyoto, Japan.

Xu X et al., 2025. Enhancing User Intent for Recommendation Systems via Large Language Models, Preprints, 2025: 1-9.

Zhou Y, Hu Z, Liu Y, 2021. Analyzing User Behavior Patterns in Casual Games Using Time Series Clustering. 2nd International Conference on Computing and Data Science (CDS), January 28-29, 2021, 372-382, Stanford, CA, USA.

Conflict of Interest

The author has declared that there is no conflict of interest”.

Author Contributions

All aspects of the study, including conceptualization, methodology, analysis, and writing, were carried out by the sole author.