

# **Project Proposal: Market Basket Analysis of Game Purchases**

**Problem:** Game publishers and retailers often struggle to optimize their sales strategies and improve revenue by identifying effective cross-promotion and bundling opportunities. By performing a market basket analysis of game purchases, we aim to uncover associations between games that are frequently purchased together, enabling more strategic promotional efforts and improved bundling strategies.

## **Context**

The gaming industry is highly competitive with a vast array of games available for purchase. Publishers and retailers need to understand consumer purchasing patterns to effectively market and bundle games. By analyzing transaction data to find associations between games, businesses can enhance their promotional strategies, increase cross-selling opportunities, and potentially boost overall sales.

## **Criteria for Success**

- **Accurate Association Rules:** Successfully identify and validate strong association rules between games purchased together.
- **Actionable Insights:** Provide actionable recommendations for cross-promotion and bundling based on the identified associations.
- **End-to-End Pipeline:** Develop and demonstrate a complete data science pipeline from data acquisition to analysis and reporting.

## **Scope of Solution Space**

- **Data Acquisition:** Obtain and preprocess transaction data related to game purchases.
- **Exploratory Data Analysis (EDA):** Analyze the frequency and patterns of game purchases.
- **Market Basket Analysis:** Implement association rule mining techniques to identify frequent item sets and strong association rules.
- **Feature Engineering:** Extract features and transform data to enhance the analysis.
- **Visualization:** Create visualizations to present the findings in a comprehensible manner.
- **Reporting:** Prepare a detailed report or presentation summarizing the methodology, findings, and recommendations.

## **Constraints**

- **Data Quality:** Ensure the data is clean, complete, and representative of the overall purchase behavior.
- **Computational Resources:** Efficiently manage resources for processing and analyzing potentially large datasets.
- **Time:** Complete the project within the given timeframe, ensuring thorough analysis and accurate results.

## **Stakeholders**

- **Game Publishers:** Interested in optimizing their marketing and bundling strategies based on purchase patterns.
- **Retailers:** Looking to enhance cross-selling opportunities and increase sales through effective promotions.
- **Consumers:** Benefiting from relevant game recommendations and attractive bundles.
- **Data Science Community:** Gaining insights into the application of market basket analysis in the gaming industry.

## **Data Sources**

**Primary Data Source:** The Steam Store Games dataset from Kaggle, which includes transaction data related to game purchases. This dataset provides a comprehensive view of game transactions and allows for detailed market basket analysis.

## **Approach**

1. **Data Acquisition and Preprocessing:**
  - Load the dataset from Kaggle and clean the data (e.g., handle missing values, remove duplicates).
  - Transform the data into a format suitable for market basket analysis, such as transaction item sets.
2. **Exploratory Data Analysis (EDA):**
  - Analyze the frequency of game purchases and co-purchases.
  - Visualize common game pairs and purchase patterns using bar charts and heatmaps.
3. **Market Basket Analysis:**
  - Apply association rule mining techniques such as Apriori or FP-Growth to identify frequent item sets.
  - Generate association rules and evaluate them using metrics like support, confidence, and lift to find strong associations.
4. **Feature Engineering:**
  - Extract features from the transaction data that might enhance the analysis, such as purchase frequency and game attributes.

- Combine these features with the association rules to provide more nuanced insights.
- 5. Visualization:
  - Create visualizations to present the identified associations and recommendations, such as network graphs of game associations and heatmaps of rule metrics.
  - Provide actionable insights for cross-promotion and bundling strategies.
- 6. Reporting:
  - Compile the findings into a comprehensive report or presentation.
  - Include code, methodology, results, and actionable recommendations for stakeholders.

## **Deliverables**

- **Code:** A complete set of scripts for data preprocessing, market basket analysis, and visualization.
- **Report:** A detailed report outlining the problem, methodology, findings, and recommendations.
- **Presentation:** A slide deck summarizing the project, key insights, and actionable recommendations for stakeholders.