Challenge Description

A leading gaming company wants to understand which players are likely to spend money in-game and how much they might spend.

Your task is to:

1. Clean and prepare player session data

2. Engineer useful features

3. Build a simple model to predict player spending in USD

You can use Excel, Python (Pandas, Scikit-learn), or any other tool to create their model.

Stage 1: Data Cleaning & Exploration

✅ Task:

Handle missing values

Remove duplicates and incorrect data

Explore trends in session duration, engagement, and spending.

Curveball: "The marketing team believes 'time spent in-game' is a key factor. Validate this assumption."

Stage 2: Feature Engineering

✅ Task:

- Create new features (e.g., average session duration, engagement score)

- Select which features are most predictive of spending

- Normalize and prepare data for modeling.

Curveball: The company wants to predict spending for new players with limited data—suggest an approach.

Stage 3: Build & Submit the Prediction Model

✅ Task:

- Build a simple regression or classification model

- Evaluate the model (R², MAE, or Accuracy)

- Submit a CSV file with predicted player spending