1. Name of the project and team members

Project Name:
Analyzing Player Activity and Recommendation Efficiency of Top Steam Games

Group members:
Jiahang Liu (student id: 7071108622)

Renjie Huang(students id: 1626544075)

2. What problem are you trying to solve?

Background:

Steam is one of the largest platforms for PC games, and it provides valuable information about how many people are playing a game and how many recommendations it receives. At first glance, games with higher player activity often seem to receive more recommendations. However, a large player base does not always mean that players enjoy the game more. Even among popular games, some titles receive strong community support while others receive relatively fewer recommendations.

Problems:

Because raw numbers can be misleading, our project uses **recommendation efficiency**, which we defined as the number of recommendations per peak player (total recommendations divided by peak players). This helps normalize differences in game size, allowing fair comparisons among the most played games on Steam.

Using player activity data from **SteamCharts** and recommendation data from the **Steam Storefront API**, this project focuses on three questions:
(1) How efficiency differs among the most popular games,
(2) Whether certain genres tend to have higher or lower efficiency, and
(3) Which games show unusually high or low recommendation efficiency compared to their player base.

3. How will you collect data and from where?

Data will be collected using two reliable online sources:

• SteamCharts(Web Scraping with request + BeautifulSoup):

From the "Top Games" page, we will extract:

- Game name

- Current players

- Last 30 days peak players

• Steam Storefront API (requests + JSON)

Using different app ID which we can directly find in steam, we will query:

https://store.steampowered.com/api/appdetails?appids={appid}

This API provides:

- Total recommendations (recommendations.total)

- Genre

  These two sources combined provide all data needed for analysis.

4. What analysis will you do and what visualizations will you create?

As we mentioned above, we will compute a normalized measure of **Recommendation Efficiency**:

$$\text{Efficiency} = \frac{\text{recommendations in total}}{\text{peak players}}$$

This allows fair comparison across different kind of games.

Planned analysis includes:

- Correlation between player activity (current/30 days peak) and recommendation totals

- Identifying games with unusually high or low recommendation efficiency

- Check the rate: if games with higher recommendation efficiency tend to have more users

Planned visualizations:

- Scatter plot: peak players vs recommendations (log-scale)

- Bar chart: top and bottom games by efficiency

- Histogram: distribution of efficiency values to illustrate how recommendation behavior varies across games.

- Linear plot by genre: plot the relationship between user count and recommendation efficiency in different genre to see whether recommendation efficiency is positively related to the number of users. (FPS, RPG etc.)

All analysis and visualizations will be implemented in Python using requests, BeautifulSoup, pandas, matplotlib, and seaborn.