GreedyGame: Session Calculation

Introduction

A gaming session happens when a user plays a game on a device. The session typically starts with a GGSTART event and ends with a GGSTOP event. This document describes a detailed approach on how to calculate number of sessions (valid and total) of a game and the average session time on the dataset which has been provided by GreedyGame. A detailed insights about users is also provided.

Data Preparation

The given dataset is in json format, it has to be converted into more usable form to be used by python. Steps for preparation of data are as follows:

1. The data from the file is loaded and is converted into a line list name dic_file.
2. An empty set of game_id named 'gme' is created in which game_id is stored.
3. Corresponding to each game_id (gme) three empty dictionaries are created named game_dic, game_dev, time_data. game_dic contains list for each game_id, in each list information like device id (ai5), SDK version, event (ggstart or ggstop), ts and timetamp is stored. game_dev contains data corresponding to particular game_id and device_id. time_data contais data about timestamp.
4. game_dev is sorted on timestamp.
5. The data thus prepared and stored in game_dev is used for session calculation.

Algorithm for Session Calculation

Case 1: if event entry is ggstop and previous event entry is ggstart then time between ggstop and ggstart is calculated and stored in sess_time.

Case 2: if event entry is ggstop and previous entry is ggstop then there discrepancy in data i.e. multiple ggstop.

Case 3: if event entry is ggstart and previous event entry is ggstart then there is discrepancy in data (multiple ggstart).

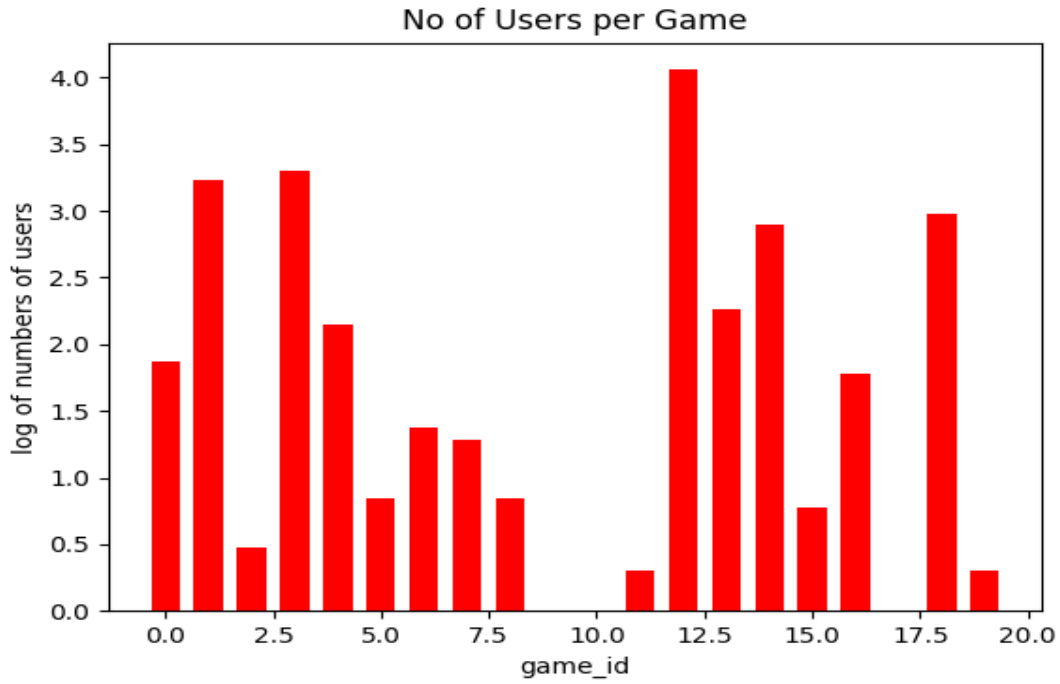
Case 4: if event entry is ggstat and previous event entry is ggatop then time difference between ggstart and ggstop is calculated and condition is checked if it is greater than 30 seconds or not. If this is greater than 30 sec and sess_time calculated earlier is greater than 60 sec then a valid session is counted. If sess_time is greater than 1 sec and less than 60 sec, then invalid session is counted.

In previous loop total time of valid session is also calculated which is used for calculation of average time of valid session.

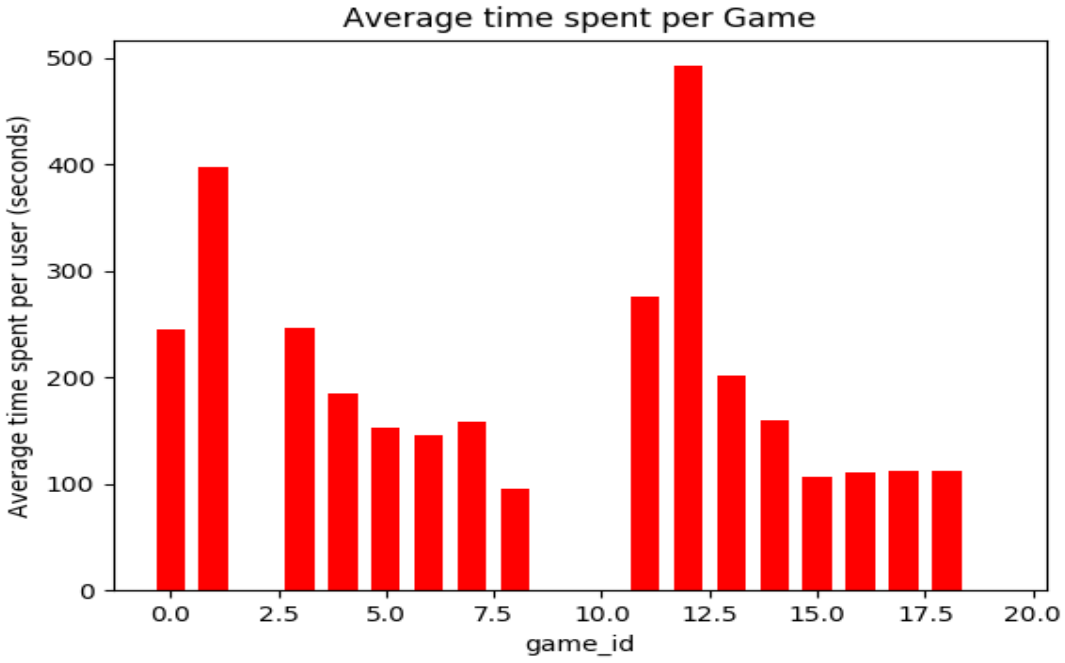
Insights about Users

To show the relation of users with games three bar graphs are plotted.

1. Number of users per game: In this graph log of number of users vs game_id is plotted. Each game_id is given a number from 1 to 20. It shows that number of users of game_id corresponding to number 12 is maximum.



2. Average time spent per user per game: In this gaph average time spent by user per game is plotted against game_id. this follows more or less similar attern as above. Here we see that among these games users spent their time mostly on game corresponding to number 12.



3. User vs Time Spent: This graph shows the relationship between fraction of users and corresponding time spent on game. This shows most of the users spent very less time on game (about 125 seconds). There are very few users who spent 2500-3000 seconds on game.