DATA ANALYSIS PROJECT – YYYY-GAME

1) Introduction

Case Study: An analysis on the gameplay data of the prototype "YYYY" by XXXX.

The purpose of this study is to gain insights into the prototype and give business suggestions based on these results.

Available data sets:

- Start
- Endgame
- Installs
- Logins
- Tutorial

Metrics Calculated:

- Daily number of sessions
- Daily number of unique users
- Average number of rounds per game per user
- Number of installs per day
- Retention based on Session Starts
- Retention based on Logins
- Average game duration per user
- Users in Tutorial with Step Count
- Average game metrics per user
- Max game metrics per user
- Average blast count w total number of sessions per user

This report will only reflect the results of these calculations. Calculation outputs are provided as an attachment (Attachment 1 – Outputs Combined) to this file.

Platforms used: SQLite3 Server, SQLite Studio, MS Excel, MS Word

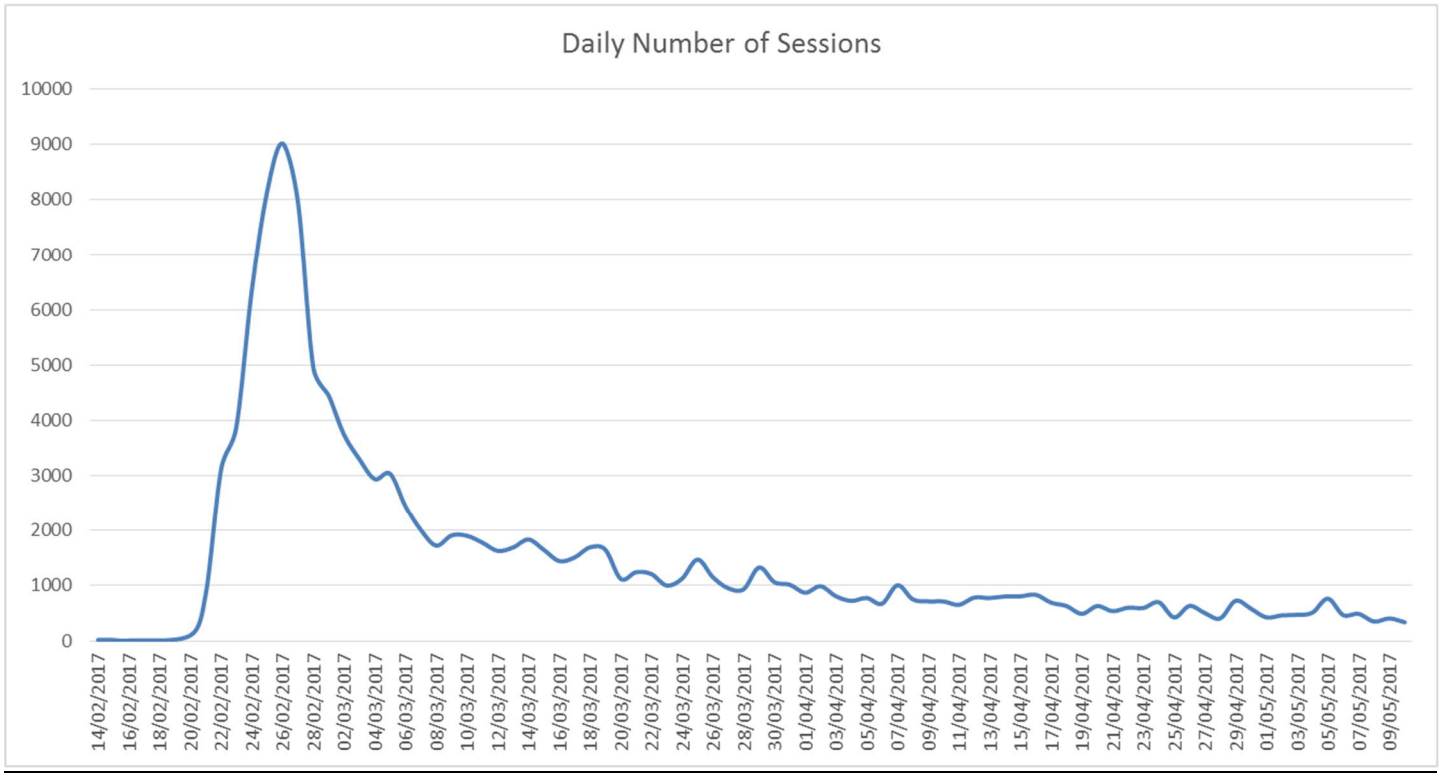
SQL queries are provided in the word file attached: "Attachment 2 – Queries"

Important notes before looking at the results:

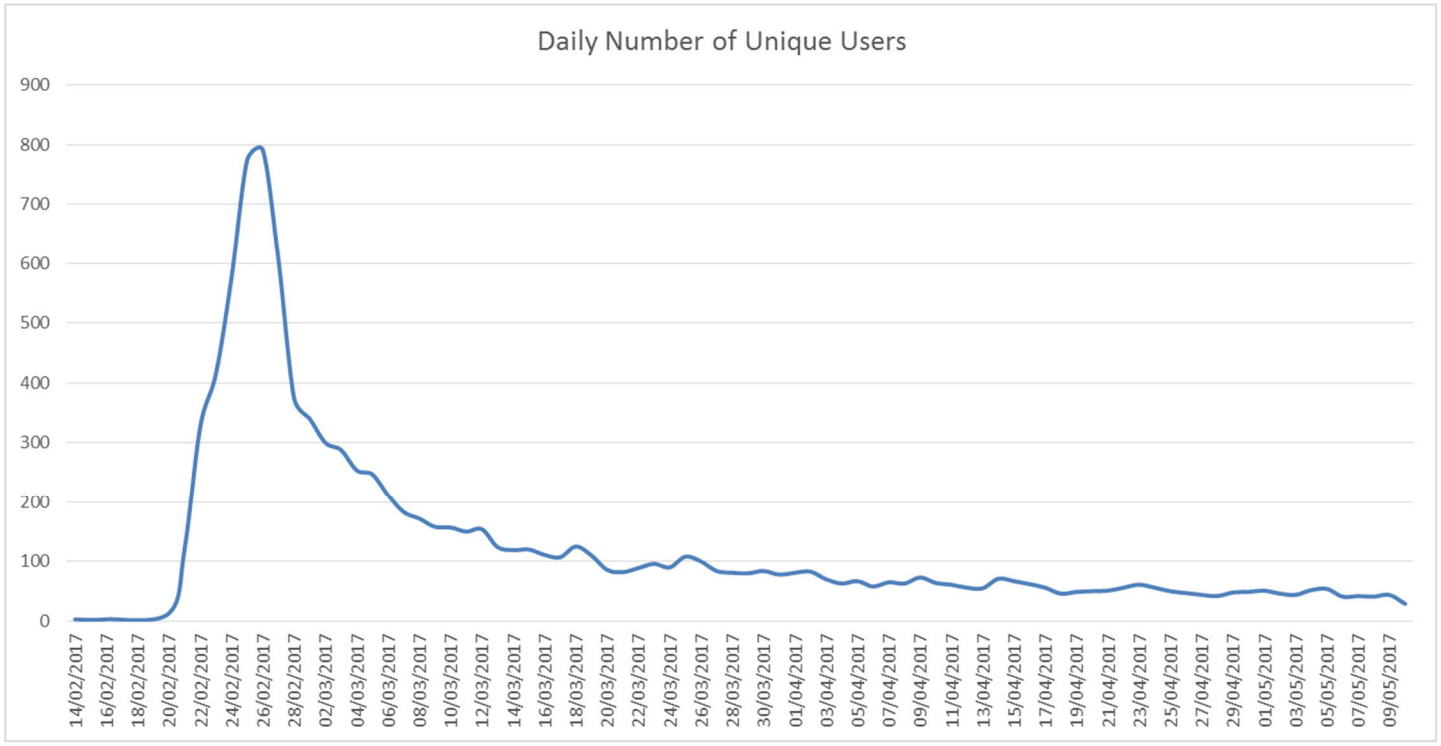
- It is noted that there are negative "s_time" values in the "Start" table (5 rows), these records are deleted.
- It is noted that there are duplicate rows in several tables, these duplicates are deleted.

2) Metrics

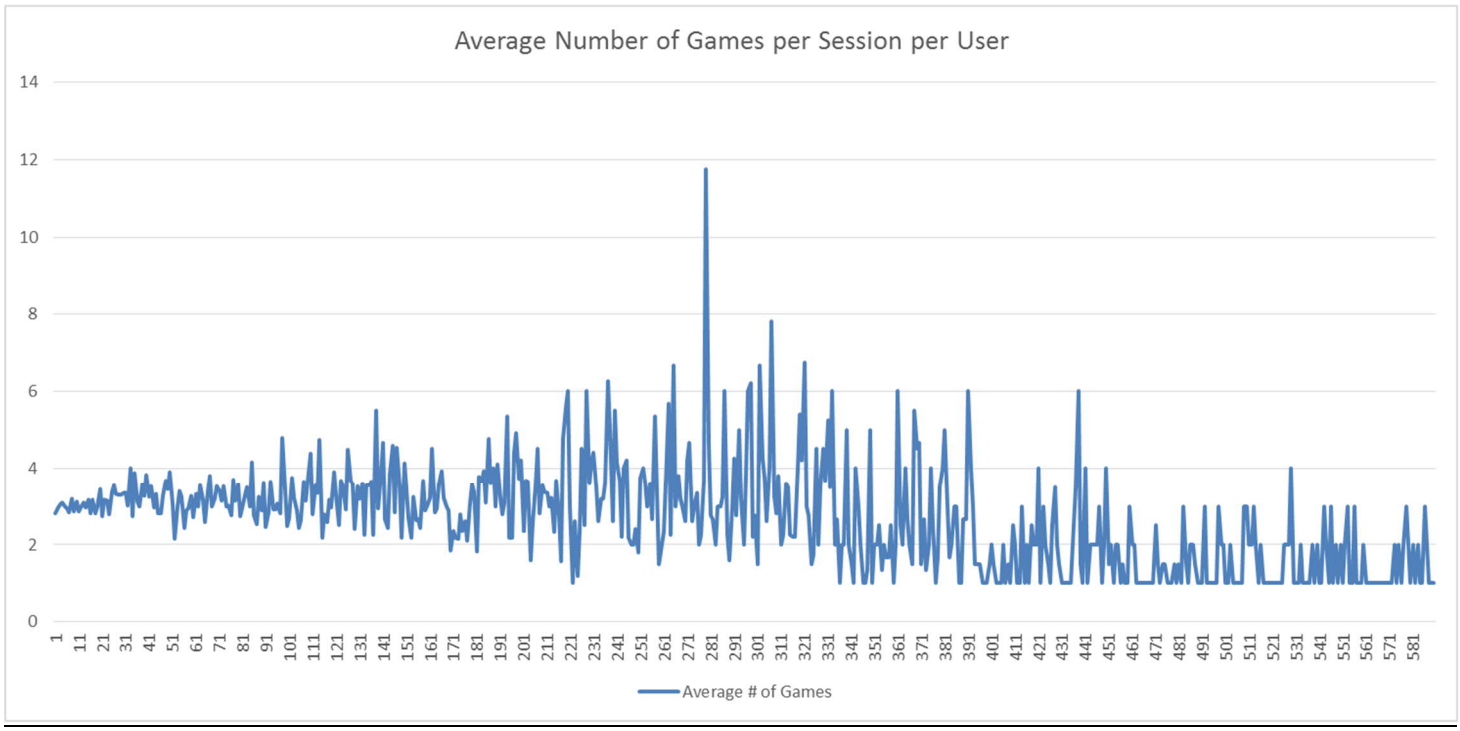
Daily number of sessions



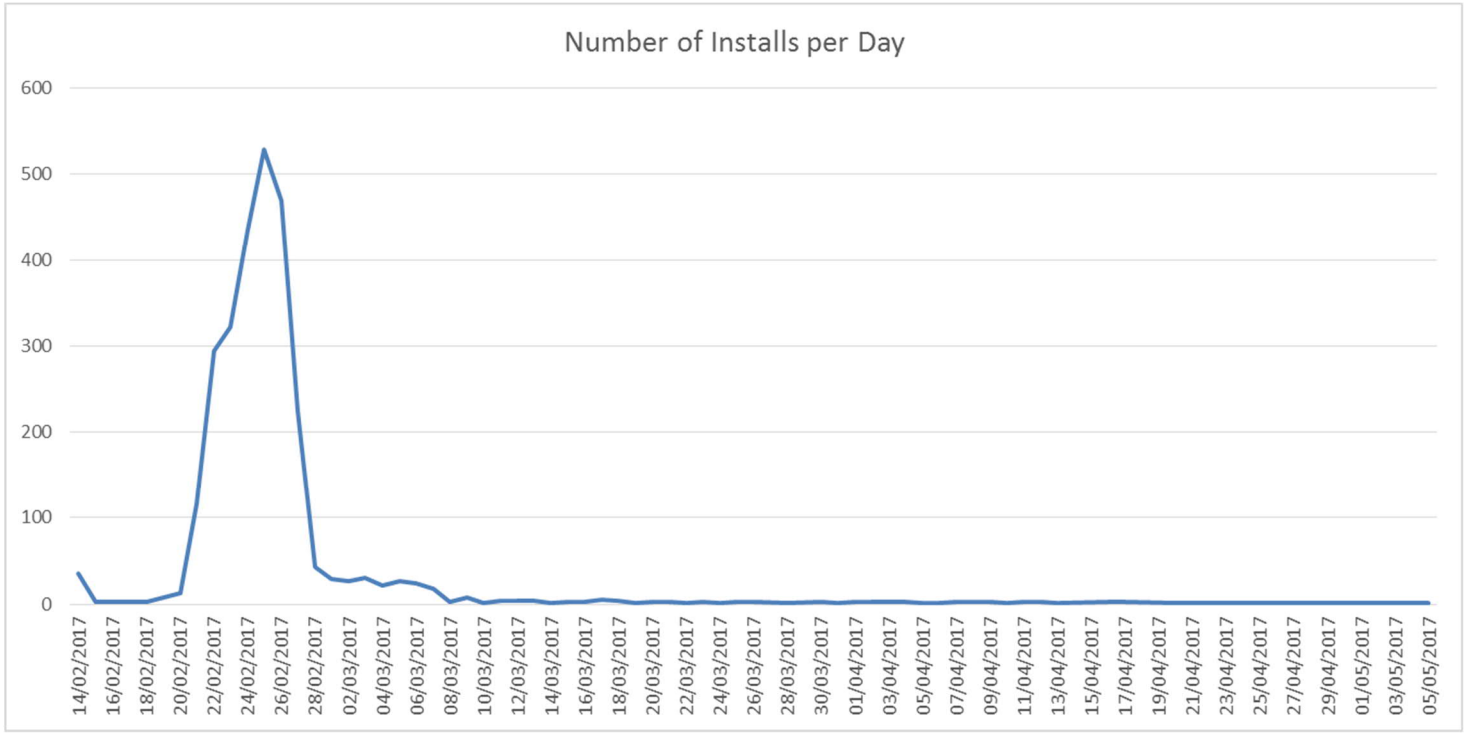
Daily number of unique users



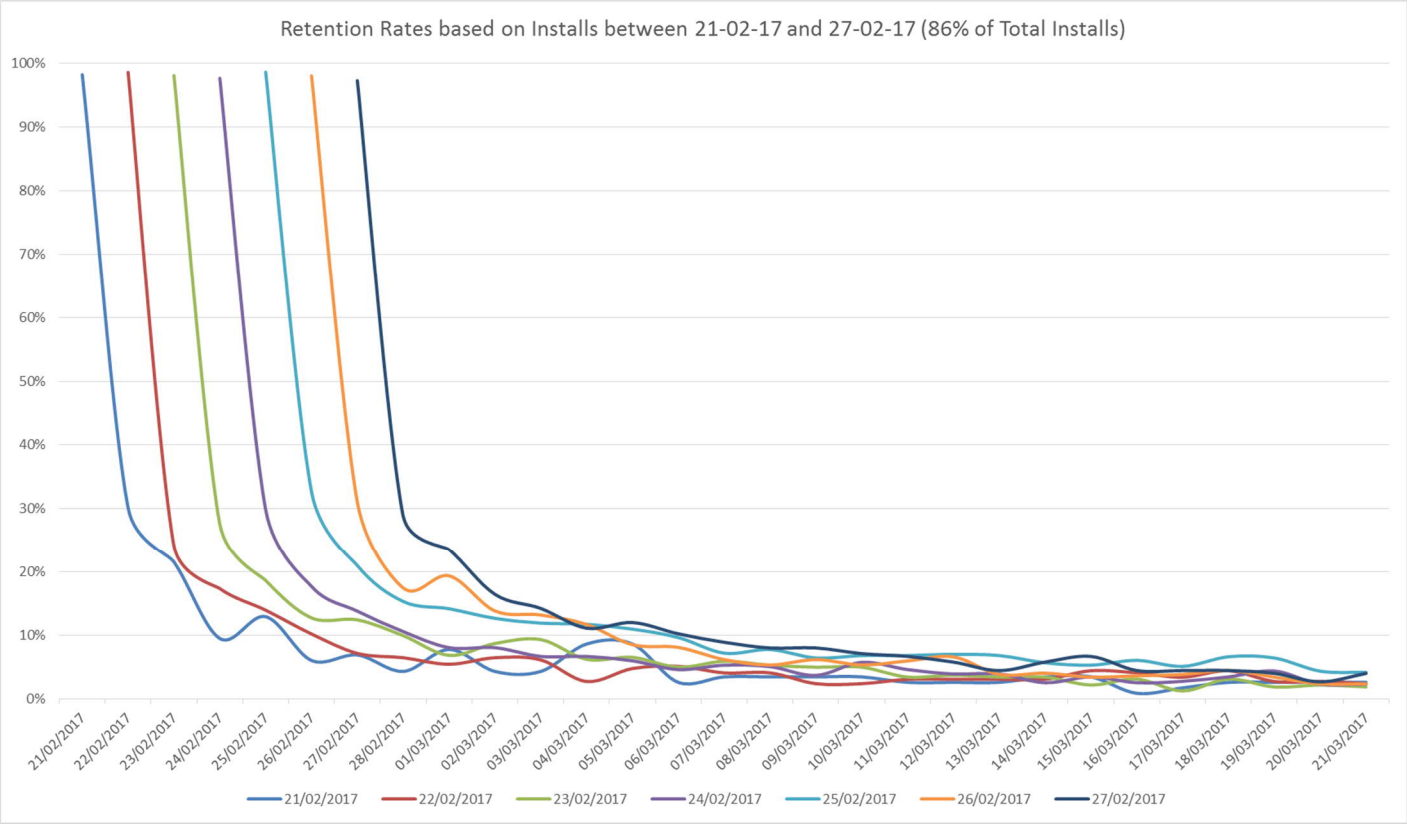
Average number of games per session per user



Number of Installs per Day

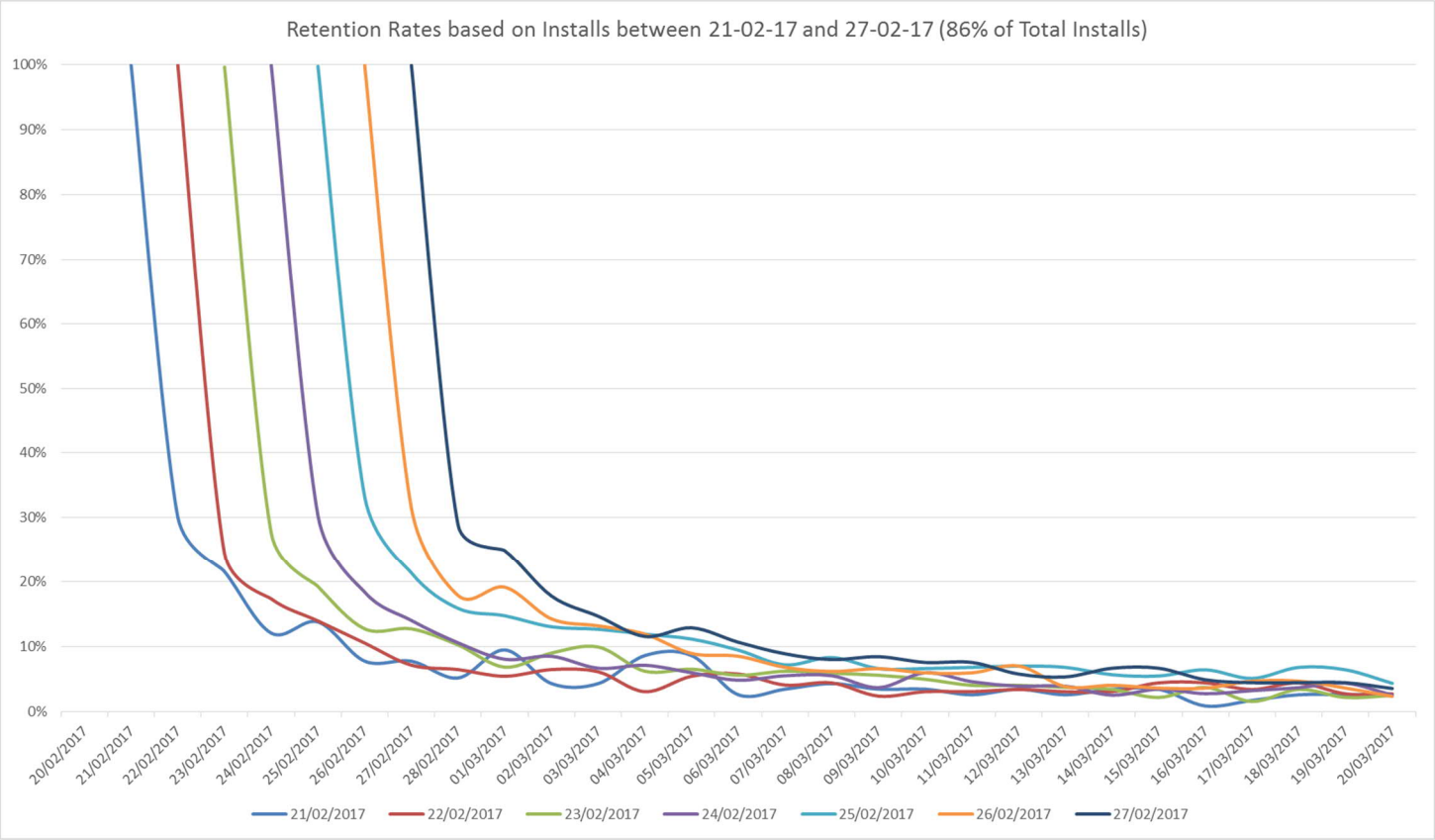


Retention – Based on session starts



# of Inst.	116	294	322	435	529	470	225	43	29	26	31	22	26	24	18
Date	21/02/2017	22/02/2017	23/02/2017	24/02/2017	25/02/2017	26/02/2017	27/02/2017	28/02/2017	01/03/2017	02/03/2017	03/03/2017	04/03/2017	05/03/2017	06/03/2017	07/03/2017
21/02/2017	98%														
22/02/2017	30%	99%													
23/02/2017	22%	24%	98%												
24/02/2017	9%	17%	28%	98%											
25/02/2017	13%	14%	19%	30%	99%										
26/02/2017	6%	10%	13%	18%	33%	98%									
27/02/2017	7%	7%	12%	14%	21%	31%	97%								
28/02/2017	4%	6%	10%	11%	15%	17%	29%	98%							
01/03/2017	8%	5%	7%	8%	14%	19%	24%	26%	97%						
02/03/2017	4%	6%	9%	8%	13%	14%	16%	23%	28%	96%					
03/03/2017	4%	6%	9%	7%	12%	13%	14%	16%	21%	31%	94%				
04/03/2017	9%	3%	6%	7%	12%	12%	11%	12%	14%	19%	29%	95%			
05/03/2017	9%	5%	7%	6%	11%	9%	12%	5%	17%	19%	19%	27%	100%		
06/03/2017	3%	5%	5%	5%	10%	8%	10%	7%	7%	12%	13%	27%	15%	96%	
07/03/2017	3%	4%	6%	5%	7%	6%	9%	5%	0%	12%	10%	14%	19%	17%	89%
08/03/2017	3%	4%	5%	5%	8%	5%	8%	7%	14%	8%	10%	14%	23%	8%	39%
09/03/2017	3%	2%	5%	4%	6%	6%	8%	7%	3%	8%	10%	14%	19%	8%	28%
10/03/2017	3%	2%	5%	6%	7%	5%	7%	9%	7%	8%	10%	9%	19%	13%	22%
11/03/2017	3%	3%	3%	5%	7%	6%	7%	5%	3%	8%	10%	14%	12%	13%	22%
12/03/2017	3%	3%	4%	4%	7%	7%	6%	9%	7%	15%	6%	9%	12%	17%	22%
13/03/2017	3%	3%	3%	4%	7%	4%	4%	12%	0%	0%	10%	5%	12%	4%	6%
14/03/2017	3%	3%	3%	3%	6%	4%	6%	7%	0%	4%	16%	5%	15%	8%	11%
15/03/2017	3%	4%	2%	3%	5%	3%	7%	9%	3%	0%	13%	5%	19%	8%	11%
16/03/2017	1%	4%	3%	3%	6%	4%	4%	5%	3%	0%	10%	14%	8%	8%	17%
17/03/2017	2%	3%	1%	3%	5%	4%	4%	12%	3%	4%	6%	14%	8%	8%	6%
18/03/2017	3%	4%	3%	3%	7%	4%	4%	9%	0%	4%	10%	9%	8%	4%	17%
19/03/2017	3%	3%	2%	4%	6%	3%	4%	5%	3%	4%	13%	9%	8%	4%	11%
20/03/2017	3%	3%	2%	2%	4%	2%	3%	5%	0%	4%	10%	9%	8%	8%	11%
21/03/2017	3%	2%	2%	2%	4%	2%	4%	2%	0%	0%	6%	5%	8%	4%	17%

Retention – Based on Logins

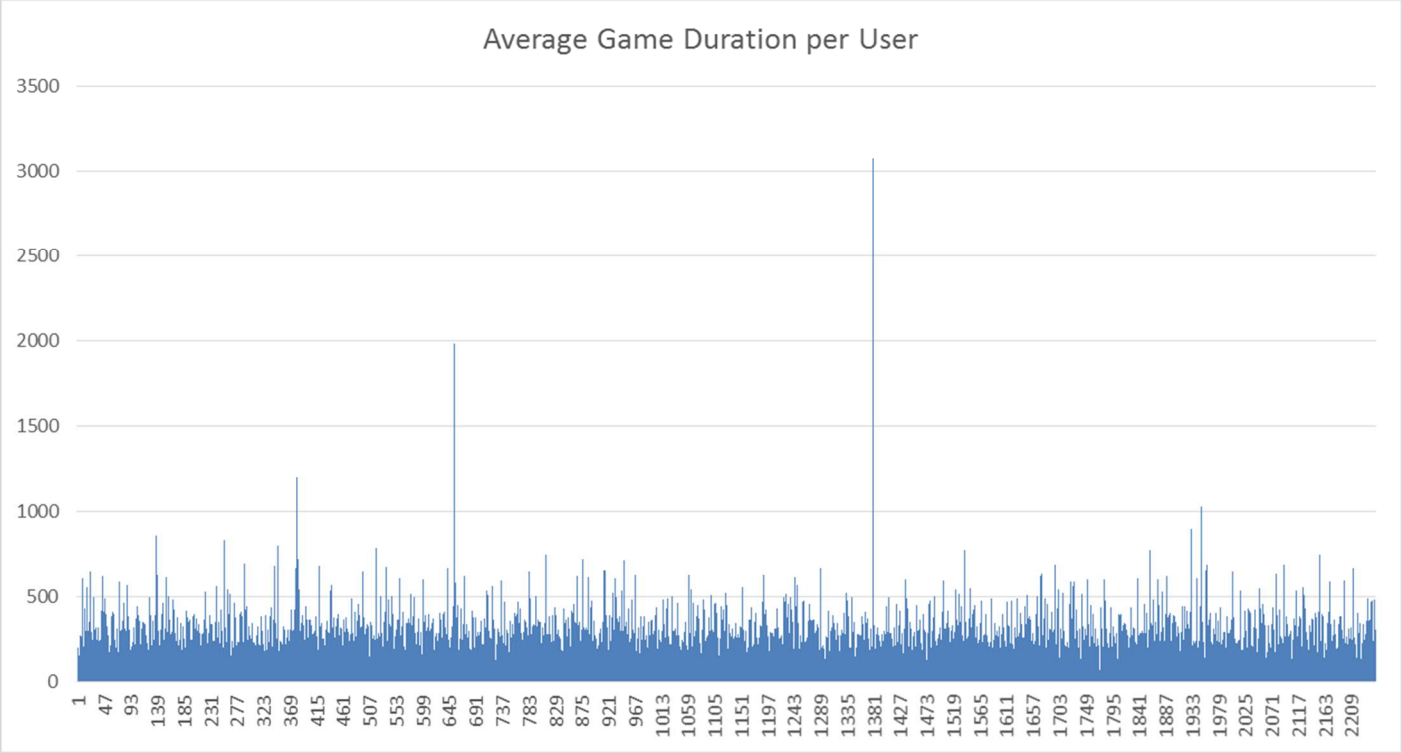


# of Inst.	116	294	322	435	529	470	225	43	29	26	31	22	26	24	18
Date	21/02/2017	22/02/2017	23/02/2017	24/02/2017	25/02/2017	26/02/2017	27/02/2017	28/02/2017	01/03/2017	02/03/2017	03/03/2017	04/03/2017	05/03/2017	06/03/2017	07/03/2017
21/02/2017	100%														
22/02/2017	30%	100%													
23/02/2017	22%	24%	100%												
24/02/2017	12%	17%	28%	100%											
25/02/2017	14%	14%	19%	30%	100%										
26/02/2017	8%	11%	13%	18%	33%	100%									
27/02/2017	8%	7%	13%	14%	21%	31%	100%								
28/02/2017	5%	6%	10%	11%	16%	18%	29%	100%							
01/03/2017	9%	5%	7%	8%	15%	19%	25%	26%	97%						
02/03/2017	4%	6%	9%	9%	13%	14%	18%	23%	28%	96%					
03/03/2017	4%	6%	10%	7%	13%	13%	15%	16%	21%	31%	100%				
04/03/2017	9%	3%	6%	7%	12%	12%	12%	12%	14%	19%	29%	100%			
05/03/2017	9%	5%	7%	6%	11%	9%	13%	5%	17%	19%	19%	27%	100%		
06/03/2017	3%	6%	6%	5%	9%	9%	11%	7%	7%	12%	13%	27%	15%	100%	
07/03/2017	3%	4%	6%	6%	7%	7%	9%	5%	0%	12%	10%	14%	19%	17%	100%
08/03/2017	4%	4%	6%	6%	8%	6%	8%	7%	17%	8%	10%	14%	23%	8%	39%
09/03/2017	3%	2%	6%	4%	7%	7%	8%	7%	3%	8%	10%	14%	19%	8%	28%
10/03/2017	3%	3%	5%	6%	7%	6%	8%	9%	7%	8%	13%	9%	19%	13%	22%
11/03/2017	3%	3%	4%	5%	7%	6%	8%	5%	3%	12%	10%	14%	12%	13%	22%
12/03/2017	3%	3%	4%	4%	7%	7%	6%	9%	7%	15%	6%	9%	12%	17%	22%
13/03/2017	3%	3%	4%	4%	7%	4%	5%	12%	0%	0%	10%	5%	15%	4%	6%
14/03/2017	3%	3%	3%	3%	6%	4%	7%	7%	0%	4%	16%	5%	15%	8%	11%
15/03/2017	3%	4%	2%	3%	5%	4%	7%	9%	3%	0%	13%	5%	19%	8%	11%
16/03/2017	1%	4%	4%	3%	6%	4%	5%	5%	3%	0%	10%	14%	8%	8%	17%
17/03/2017	2%	3%	2%	3%	5%	5%	4%	12%	3%	4%	6%	9%	8%	8%	6%
18/03/2017	3%	4%	3%	4%	7%	5%	4%	9%	0%	4%	10%	9%	8%	4%	17%
19/03/2017	3%	3%	2%	4%	6%	4%	4%	5%	3%	4%	13%	9%	8%	4%	11%
20/03/2017	3%	3%	2%	3%	4%	2%	4%	5%	0%	4%	10%	9%	8%	8%	11%
21/03/2017	3%	2%	2%	3%	5%	3%	4%	2%	0%	0%	6%	5%	8%	4%	17%

INTERIM ANALYSIS!!!

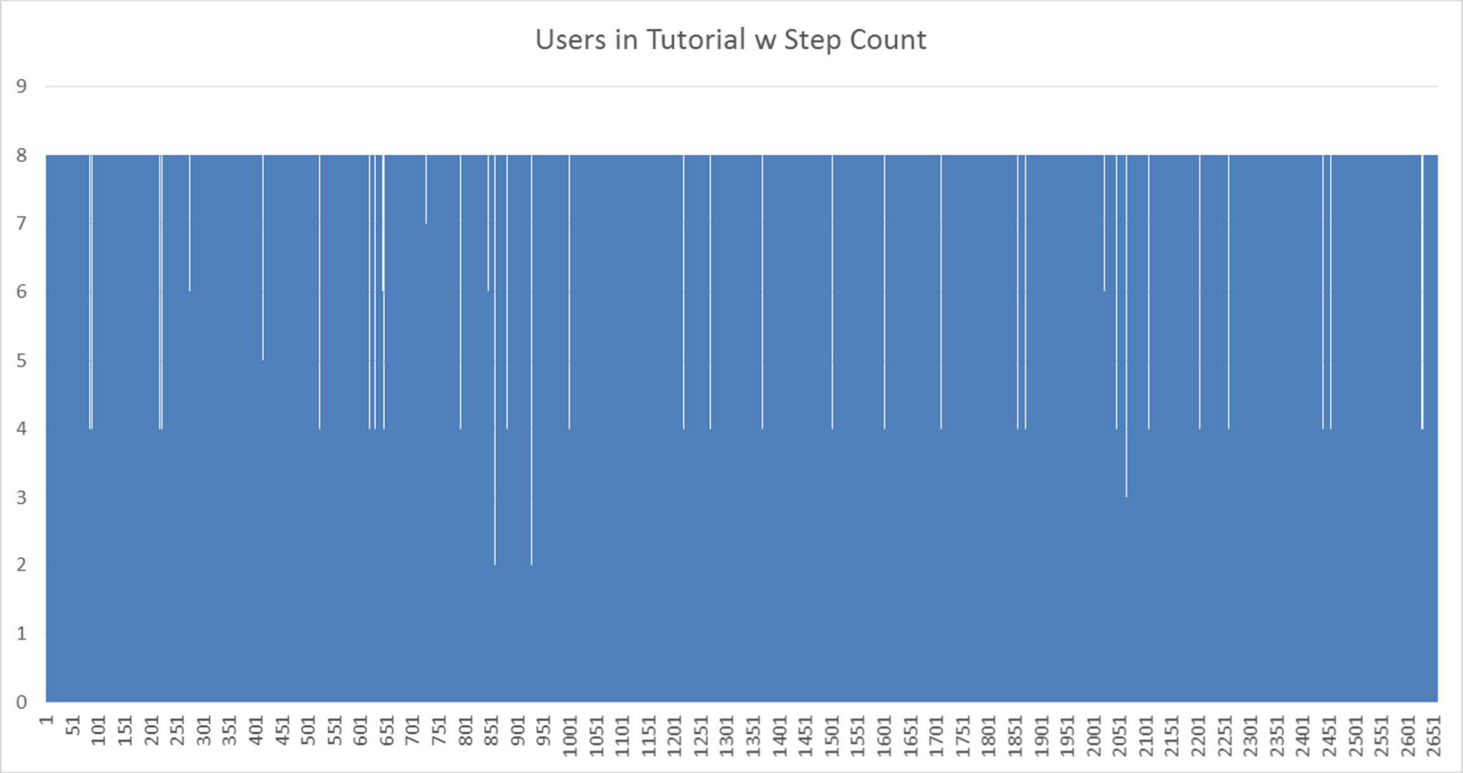
At this point of the analysis, it is noted that retention rates are significantly low. Although the initial idea was to go deeper into user performances in terms of on which aspects of the game they are good or bad, changing the approach towards understanding why the retention is very low seemed to be the sensible decision.

For this purpose, first analysis was on the average game duration per user and the results are as following:

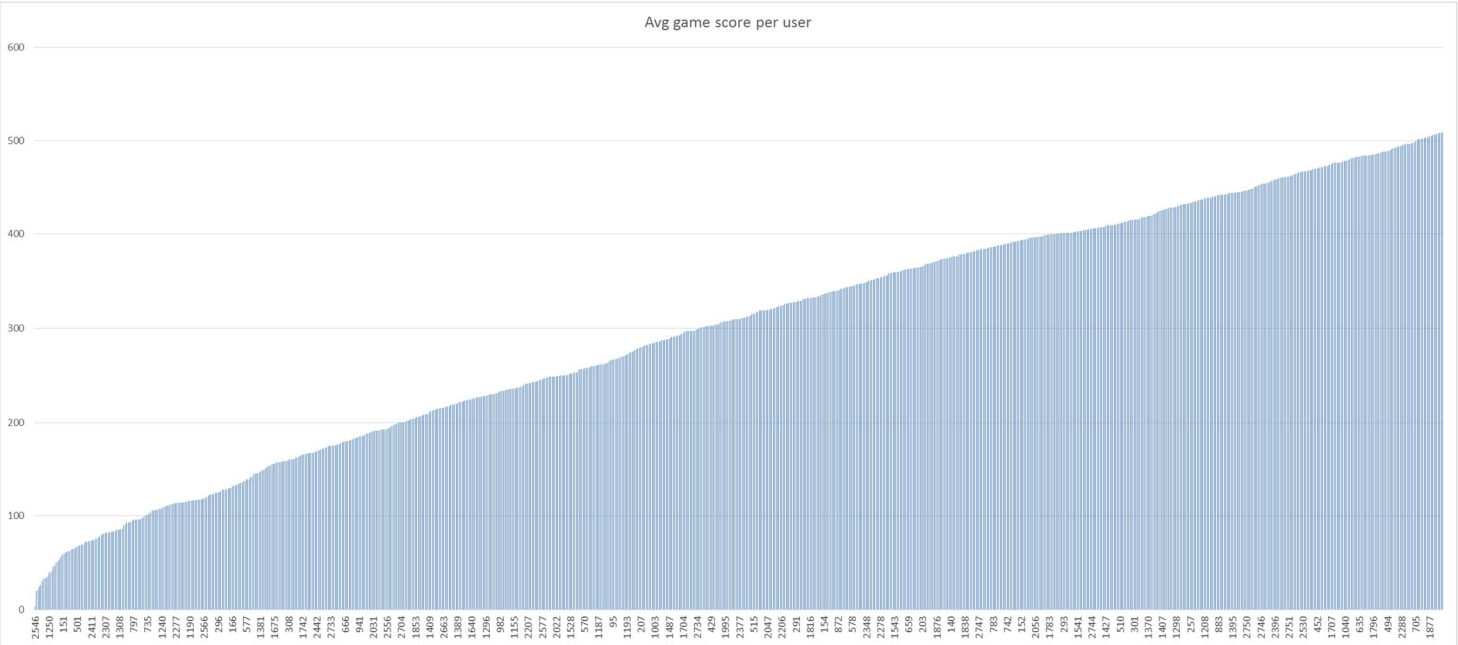


It can be seen that except for a couple of users, the average game duration per user seems to be similar for all users.

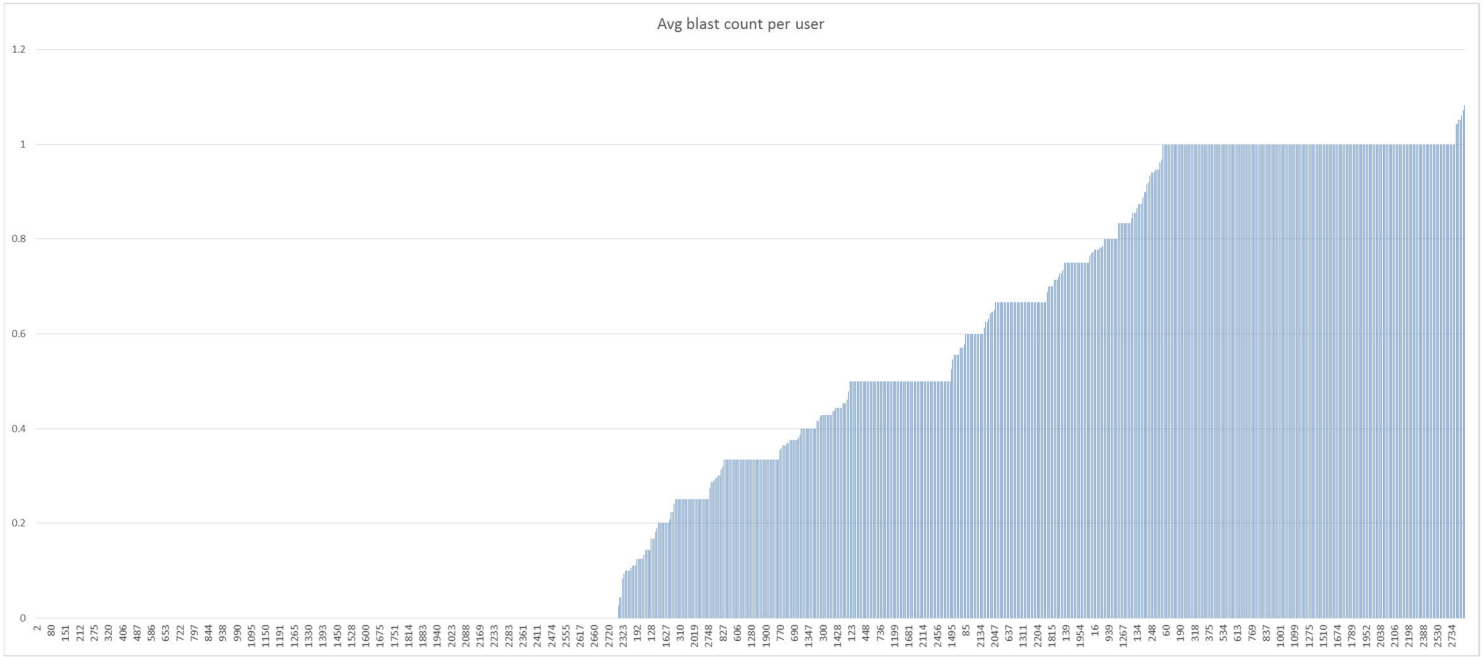
Another reason why the retention rates are low could also be the tutorials being not effective and/or the users didn't pay much attention to the tutorials. However, it is seen that out of 2746 users who installed the game, 2658 went through the tutorials, i.e. 96.7% of the total users. And after a simple analysis on how far they went in the tutorials, it is seen that the vast majority went all the way to the last level, as shown in the following graph:



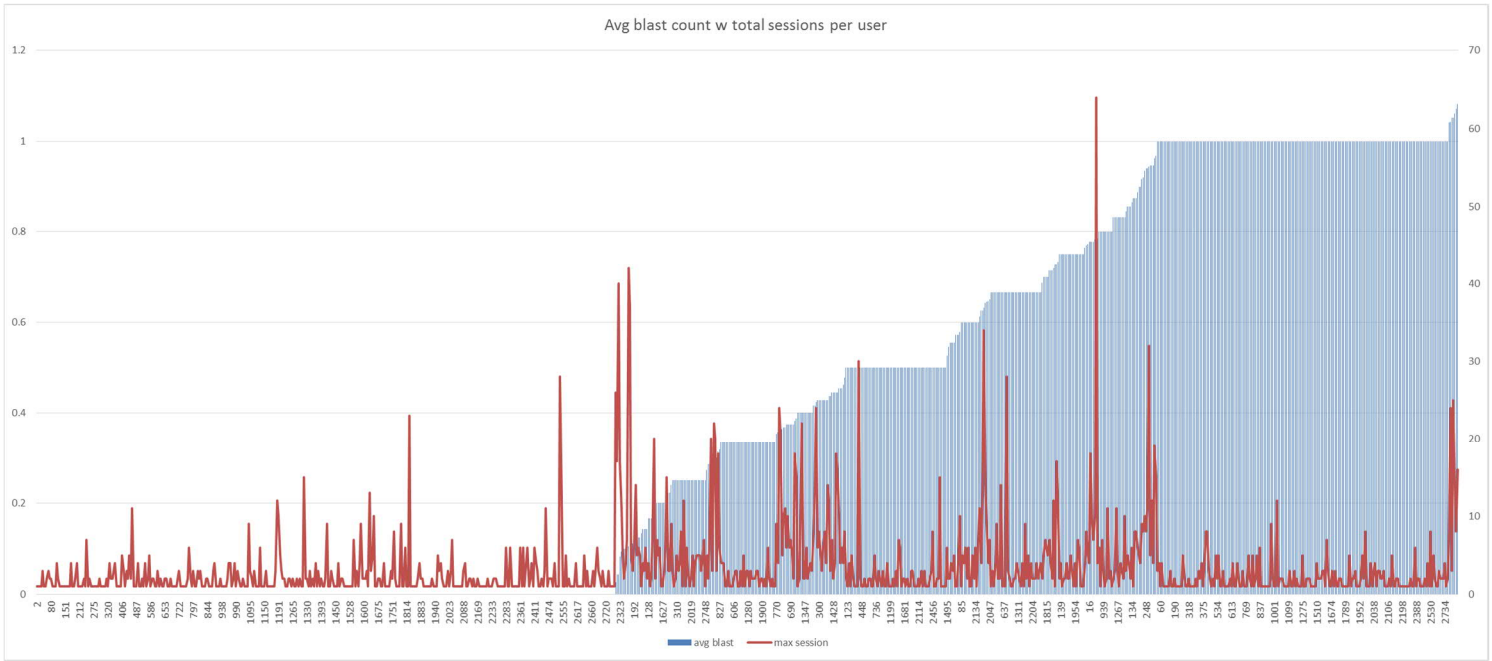
Final reason that could be found out from the data available could be the game being too hard. In order to understand how the players performed, all features of the game were analyzed and all of them (except for one!!!), seemed to have a similar distribution over the users as the following example, and these were exactly as expected:



The exception was the “blast count” feature, which is the explosion that happens when the user merges the paw piece, which is a big move in the game. It is seen that almost half of the users couldn’t even get to 1 (one) blast, and those who made it could only make it for once (except for a few very good users):



This could be the answer: the game is too hard to reach a certain level satisfaction, and the majority of the users lost their appetite towards it. To test this idea, and see whether those “unsuccessful” users are the ones who didn’t come back to play more, total number of sessions is distributed on the graph above, and if this hypothesis is correct, then we should see correlation between the two metrics. Results are as following:



It is clear that this hypothesis was wrong. We see that some users played the game a lot, but couldn’t manage to get a blast(s), and some made it with very few tries. The only possible insight from this graph could be the game is so imbalanced that it is too easy for some players, and too difficult for others. To test this hypothesis, “the average time spent per game per user” can be a good analysis, but it is out of the scope of this study.

3) Results

In conclusion, based on the analysis and insights above, it is clear that the game should NOT BE LAUNCHED as it is right now.

(If the final analysis suggestion of “the average time spent per game per user” distributed over the graph “average blast count w total sessions per user” show a correlation and the game seems to be imbalanced in terms of difficulty among different users, then the main suggestion would be to rearrange the difficulty accordingly. On the other hand, this result is very unlikely because the distribution of average score per user is pretty much as expected.)

The fact that tangible results show no clear indication of why the prototype is not successful implies that there are reasons cannot be seen through the data available, and the usual suspects might be that the users found the game not “fun” to play, the design was not attractive, or maybe the users don’t want the leave the similar games they are currently playing because this one does not offer something new, and so on. However, before embracing these ideas and attacking on the aspects mentioned, user feedback on these aspects of the game should be collected.