Investigating Rolling Player Retention in Mobile Game Company (SQL Project)

Spring Chhom, Amy Wong and Saba Hakimpour

Juno College of Technology, October 2021

Project Objective

Investigate retention rate of a mobile game company.

About Dataset

Dataset was in CSV format and included 4 tables of item, matches, player and purchase information.

Business Questions

---

1. **investigate the percentage of rolling retention of players over 1 year and calculate fractional retention.**

First we needed to define a query to determine retained and not-retained players. My partners and I used max day to see the latest day players played a game. Then used if statement to define retained players as 1 and non-retained one as 0.
To get the fractional retention we divided total retained players by the total number of distinct players. So this query gave us a table for the day players joined, total players joined on the specific day, total players retained and fractional retention.

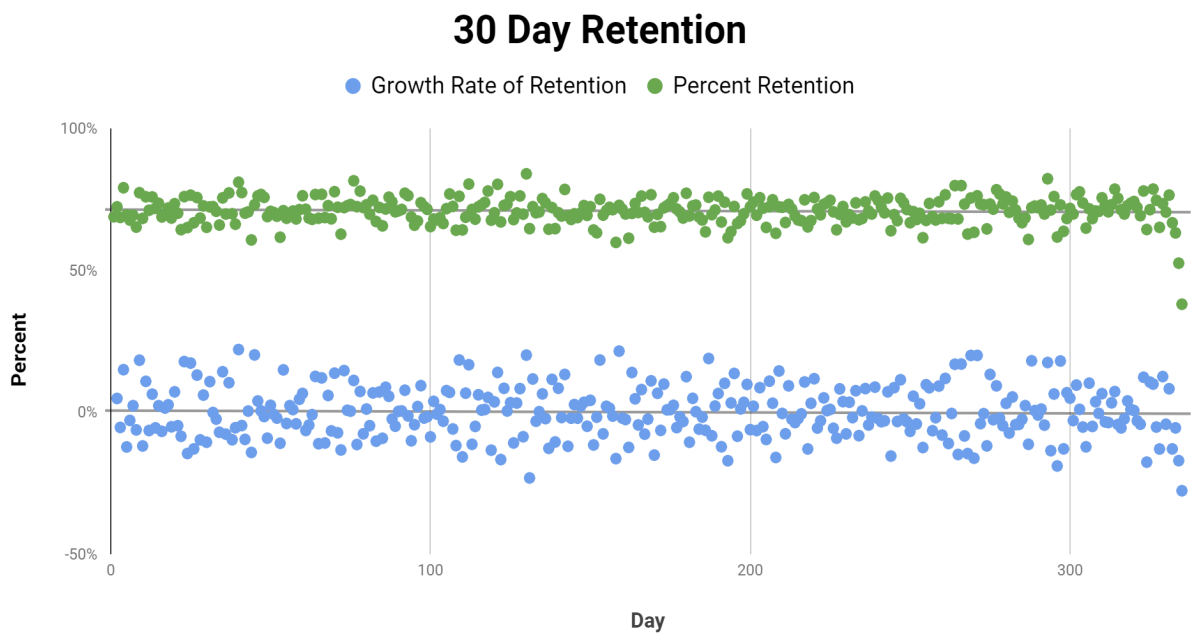
Answer: Of the 40,452 players who joined the game over one year, 65.62% of them were considered retained. The range of fractional retention was between 56% to 80% and they averaged 68% overall.

2. The percentage of growth rate.

For this analysis we used the fractions retention query to calculate growth rate per day: first used the window function to be able to get past fraction rate then subtract past fractional rate from current fractional retention , then divide the result by past fraction rate to get the growth rate for each day. We used safe divide to prevent any zero division error.

Answer: Average retention growth rate per day was 0.24% (basically 0% - no real change)

Visualization (All the visuals were made in Google Sheets)



3. Calculate average purchase by retained and non-retained group.

For this analysis we had to join all tables. First, we used the sum function along with the partition to get how much each player spent. Then, define a query regarding retention status of each player and use the partition to be able to get the status per player. Finally, based on this query we used the AVG function to get the result of the average amount spent by retained and non-retained groups.

Answer: Average spend among non-retained players is higher than retained players.