

# Mobile App Analytics



## Goal:

Analyzing and interpreting user behavior trends.

## Data Description:

### Source:

The data is taken from the **Firestore Analytics** public project:

*firebase-public-project.analytics\_153293282*.

This dataset contains events that represent user activity within a mobile app.

### Main Characteristics:

#### 1. Time Range:

The time range of June 12, 2018, to October 3, 2018, was chosen for analysis because it is the only available period in the dataset. This range provides almost four months of data, allowing for the identification of event dynamics and trends. Date filtering was applied to make the queries clear and universal, helping to avoid potential errors when updating or expanding the data. The sample includes all events within the available period, ensuring maximum representativeness and accuracy of the results. It also allows for a detailed analysis of user behavior in the short term.

#### 2. Event Types:

**Navigation:** screen\_view, which shows interactions with screens.

**User Engagement:** user\_engagement, reflecting overall activity.

**Gaming:** level\_start\_quickplay, level\_end\_quickplay (start and end of levels), post\_score (posting scores).

### 3. Users:

Users are identified by a unique `user_pseudo_id`, which allows tracking their behavior.

#### Format:

The data is structured as a table with the following main fields:

- **event\_name** — the name of the event.
- **event\_timestamp** — the time when the event occurred.
- **user\_pseudo\_id** — the unique user identifier.
- **event\_params** — parameters providing extra information about the event.

### Analysis Methods:

#### 1. BigQuery (SQL):

Purpose:

- Collecting, cleaning, and preprocessing large datasets.
- Executing complex queries to:
  - Filter data by dates (`event_date`).
  - Group events (`event_name`).
  - Calculate session durations (`TIMESTAMP_DIFF`).
  - Identify returning users.
- Optimizing queries for quick data retrieval.

#### 2. Python (Jupyter Notebook):

##### Pandas:

- For data manipulation after downloading from BigQuery.
- Used for:
  - Grouping and aggregating data (e.g., counting events, creating cumulative metrics).
  - Cleaning data (removing null or extreme values).
  - Example: Analyzing session durations, counting event occurrences.

## Matplotlib and Seaborn:

→ For creating visualizations such as:

- Line charts (user retention).
- Histograms (session duration distribution).
- Heatmaps (activity by hour and day of the week).

### Results of Using These Methods:

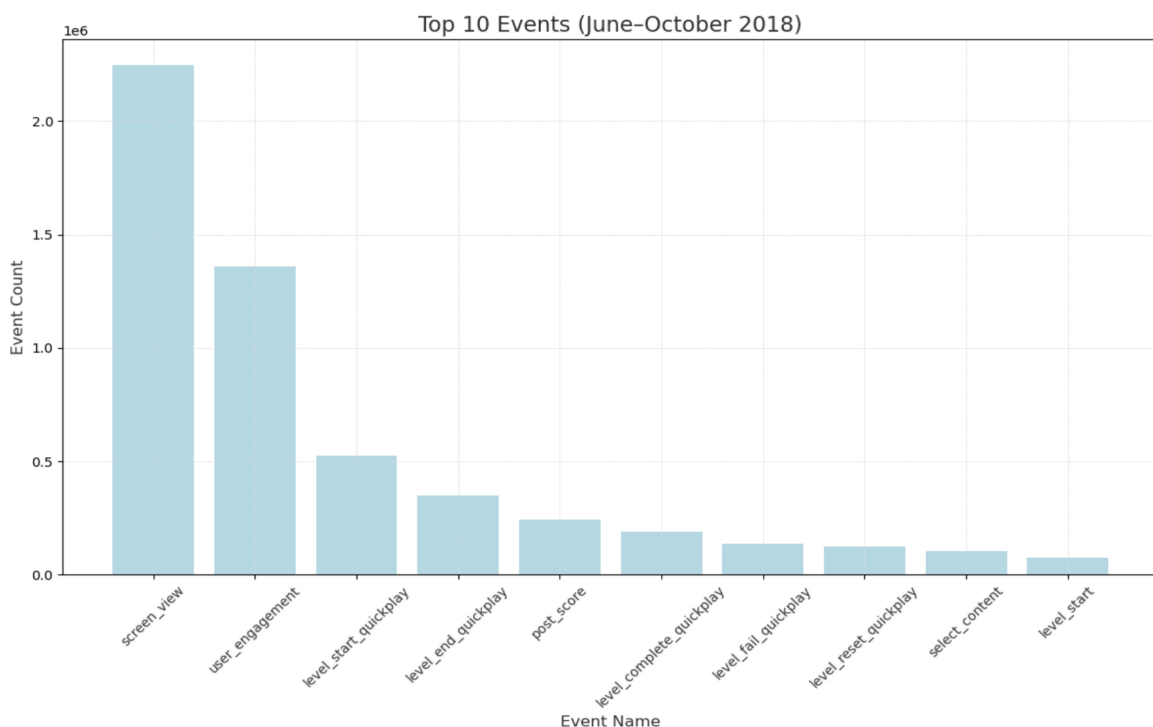
These tools supported the full analytics workflow — from data collection to visualization and reporting. BigQuery enabled efficient handling of large datasets, Python allowed in-depth analysis and chart creation.

## Analysis

**What are the most popular events among all users during the available period (June 12, 2018 – October 3, 2018)?**

### Goal:

**To identify the most popular user events during the available period (June 12, 2018 – October 3, 2018) to understand user engagement patterns and key interactions within the application.**



## Insights:

- The most frequent event is screen\_view, indicating that users primarily navigate through different screens.
- User\_engagement ranks second, showing a high level of active interaction with the app.
- Events such as level\_start\_quickplay and level\_end\_quickplay suggest frequent gaming activities, while post\_score indicates competitive participation.

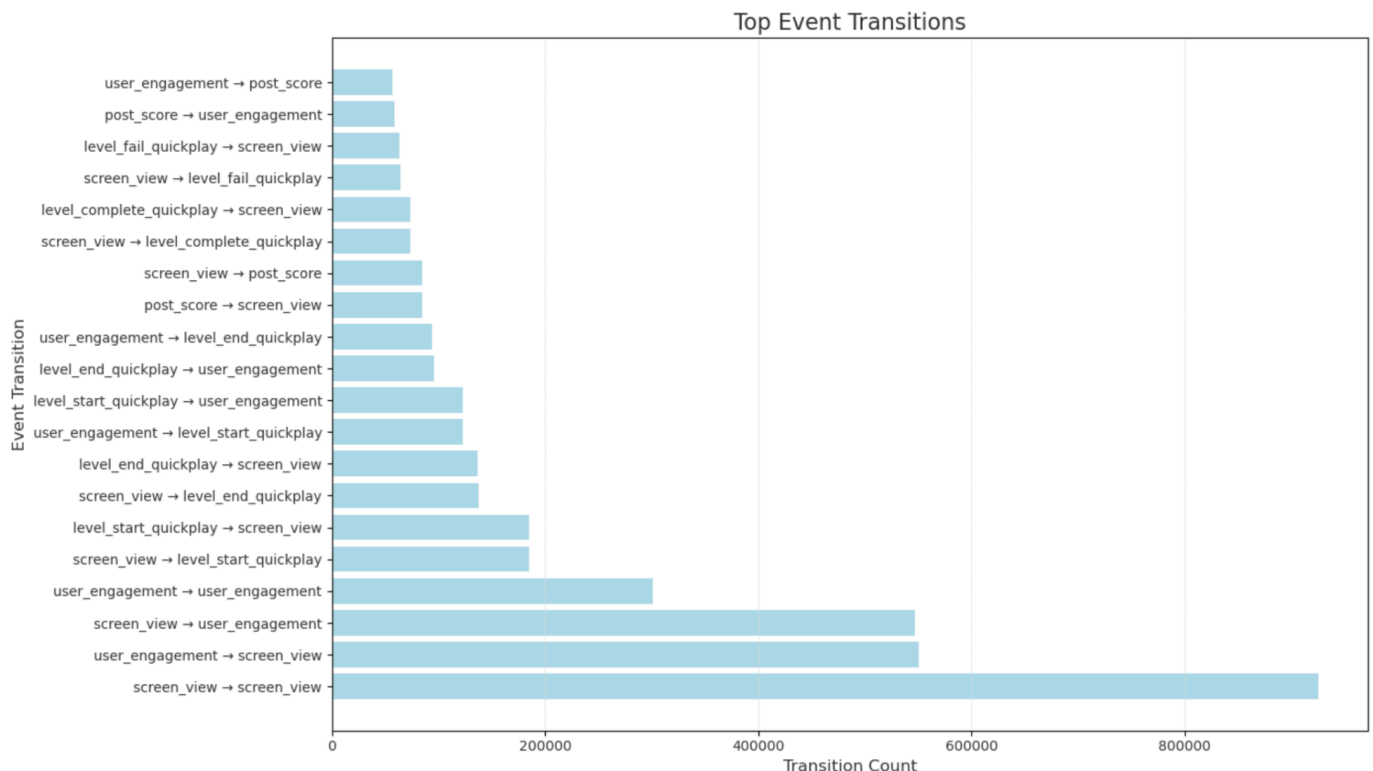
## Recommendations:

- ❖ Focus on optimizing the user interface for better navigation since screen\_view is the most common event.
- ❖ Enhance features related to user engagement to retain the active user base.
- ❖ Analyze gaming patterns further to improve retention and satisfaction among players, as gaming-related events are prominent.

## Which event is the most frequent continuation of other events?

### Goal:

To identify which event most frequently follows other events to understand user behavior and transitions between actions.



## Insights:

- The event `screen_view` is the most frequent continuation of other events, highlighting its central role in user navigation.
- Transitions such as `user_engagement` → `screen_view` and `screen_view` → `screen_view` indicate a strong focus on exploring different screens.
- Game-related transitions like `level_start_quickplay` → `screen_view` and `level_end_quickplay` → `screen_view` emphasize the importance of gaming activities in user interactions.

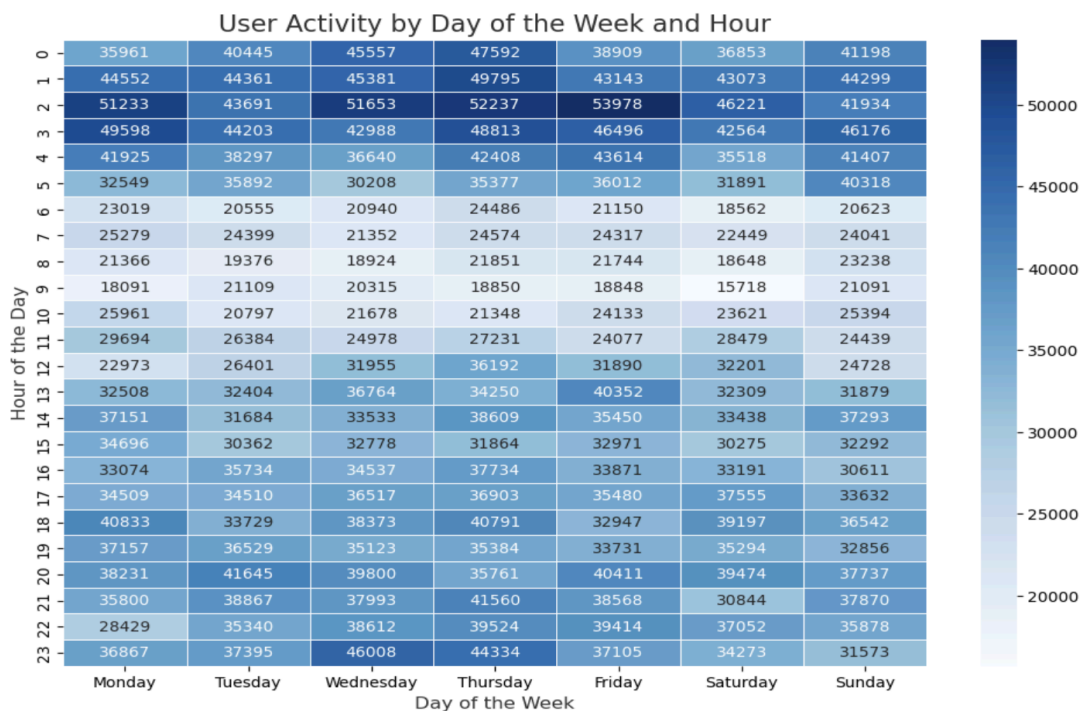
## Recommendations:

- ❖ Optimize the `screen_view` flow to ensure smooth navigation and enhance user experience.
- ❖ Investigate transitions from engagement events to `screen_view` to identify opportunities for improving engagement.
- ❖ Leverage game-related transitions to refine game mechanics and boost user satisfaction.

## Is user activity dependent on the day of the week and the time of day?

### Goal:

To determine whether user activity patterns vary based on the day of the week and the time of day to identify peak engagement periods and optimize content delivery.



## Insights:

- The heatmap shows that user activity is highest in the evening and nighttime, especially between 8:00 PM and 2:00 AM, indicating that most interactions happen during non-working hours.
- On weekdays (Monday to Friday), activity is lower in the morning (6:00 AM–9:00 AM) but increases gradually throughout the day, peaking in the evening.
- Weekend mornings (Saturday and Sunday) have the lowest activity, but it rises significantly in the evening.
- This distribution suggests that most users engage with the app for entertainment or during their free time.
- The high nighttime activity may indicate a global audience in different time zones or specific behavioral patterns among users.

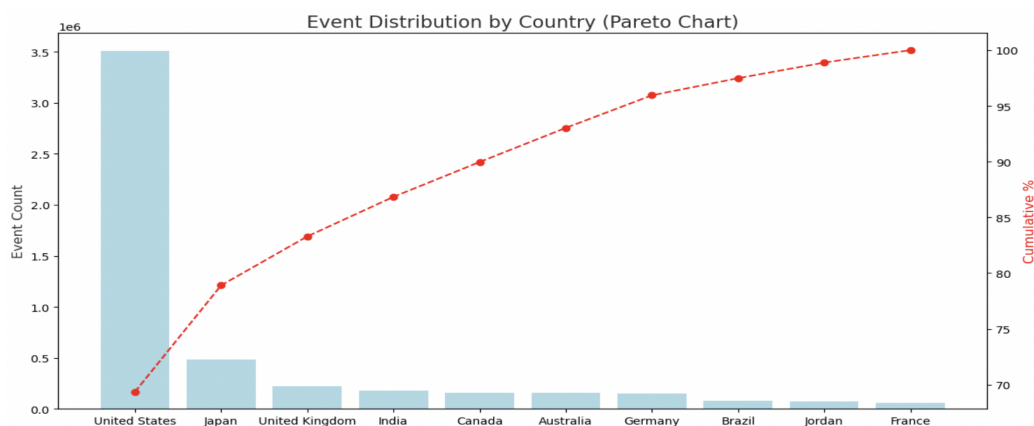
## Recommendations:

- ❖ Schedule marketing campaigns and push notifications for the evening and nighttime hours to reach users during peak activity.
- ❖ Introduce additional features or events that are appealing for evening or nighttime engagement.
- ❖ Develop content or offers that encourage higher activity in the morning to balance user interactions throughout the day.
- ❖ Analyze user behavior from different time zones to tailor content to their habits.
- ❖ Use this data to optimize server performance and allocate resources during high-traffic hours.

## Which countries generate the largest share of events?

### Goal:

To identify the countries contributing the most to the total number of events and understand the dominance of these countries in overall activity.



## Insights:

- The majority of events are generated by a small number of countries, with the United States contributing the largest share, followed by Japan and the United Kingdom.
- The cumulative percentage curve shows that a small group of countries accounts for a significant portion of total events, highlighting a clear concentration of activity.
- Other countries, such as Australia, Germany, and France, contribute a smaller share, suggesting regional disparities in user engagement.
- The long tail of countries with minimal contributions indicates limited activity in those regions, which might be due to lower market penetration or user interest.

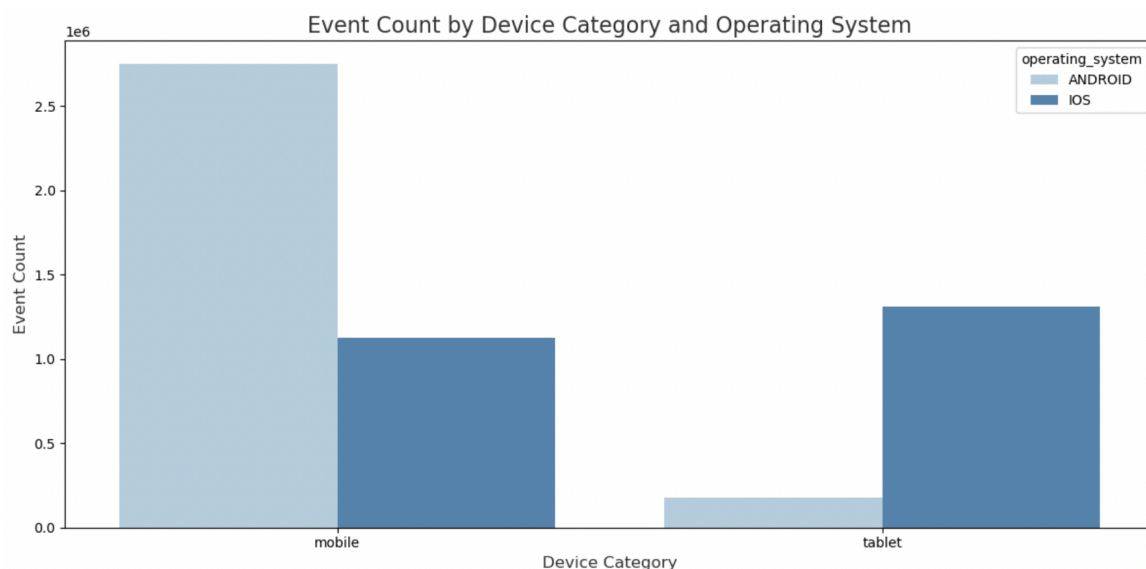
## Recommendations:

- ❖ Focus marketing and user acquisition efforts on countries with the highest contributions to reinforce existing engagement and capitalize on their potential.
- ❖ Analyze factors driving high activity in dominant countries and consider replicating successful strategies in lower-performing regions.
- ❖ Explore opportunities to increase activity in underrepresented countries through localized campaigns, new features, or partnerships tailored to regional preferences.
- ❖ Monitor trends over time to identify emerging markets that may require attention or investment.

## What are the most popular devices and operating systems?

### Goal:

Identify which device categories and operating systems generate the most events to optimize the application experience and marketing strategies accordingly.



## Insights:

- **Mobile devices** remain the most popular category, generating the majority of events in the dataset.
- Among operating systems, **Android** dominates on mobile devices, significantly outperforming **iOS**.
- On tablets, **iOS** generates more events than **Android**, suggesting a preference for this platform in the tablet category.

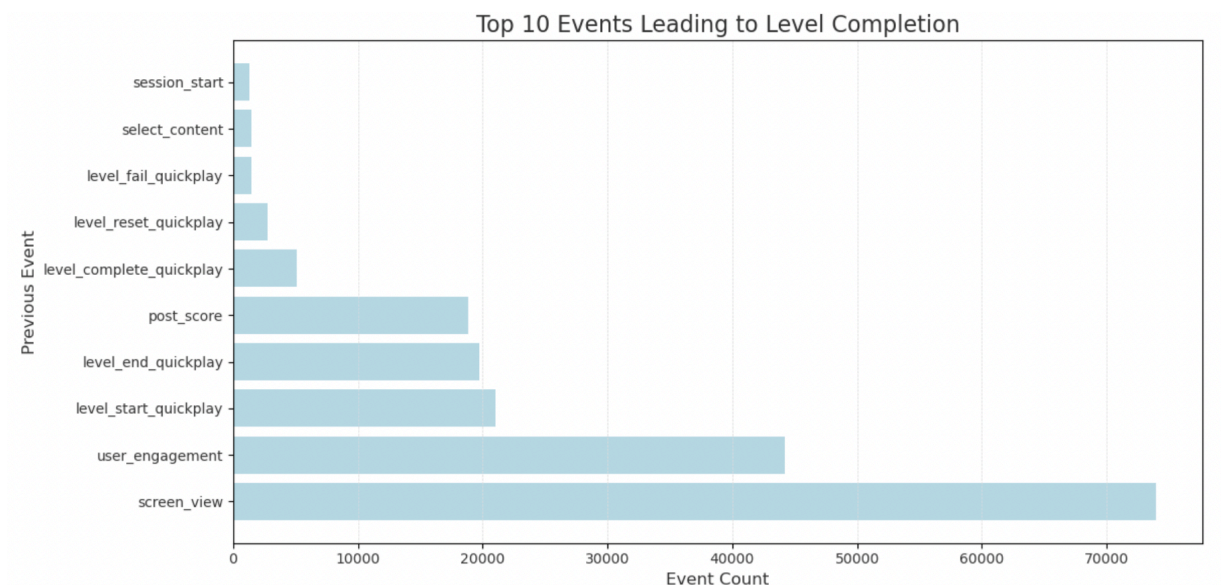
## Recommendations:

- ❖ Optimize the application for mobile devices, particularly for the **Android** platform, as it dominates in this category.
- ❖ Focus on improving the user experience for **iOS** users on tablets, given their higher engagement compared to **Android** users in this category.
- ❖ Adjust marketing strategies to target **Android** users on mobile devices and **iOS** users on tablets, leveraging platform-specific preferences to enhance engagement and retention.

## What are the most common events preceding a level completion (level\_complete\_quickplay)?

### Goal:

To identify the events that frequently occur before a level completion event. This analysis aims to understand user behavior leading to successful level completion and uncover patterns or key events that drive this outcome.





## Insights:

- **Screen view** and **user engagement** are the most common events preceding a level completion, indicating these actions are vital in guiding users to complete a level.
- **Level start quickplay** and **level end quickplay** suggest that most users progress systematically through the game from starting a level to completing it.
- Events like **post\_score** and **level reset quickplay** show additional behaviors that might either support or challenge users in their gameplay experience.

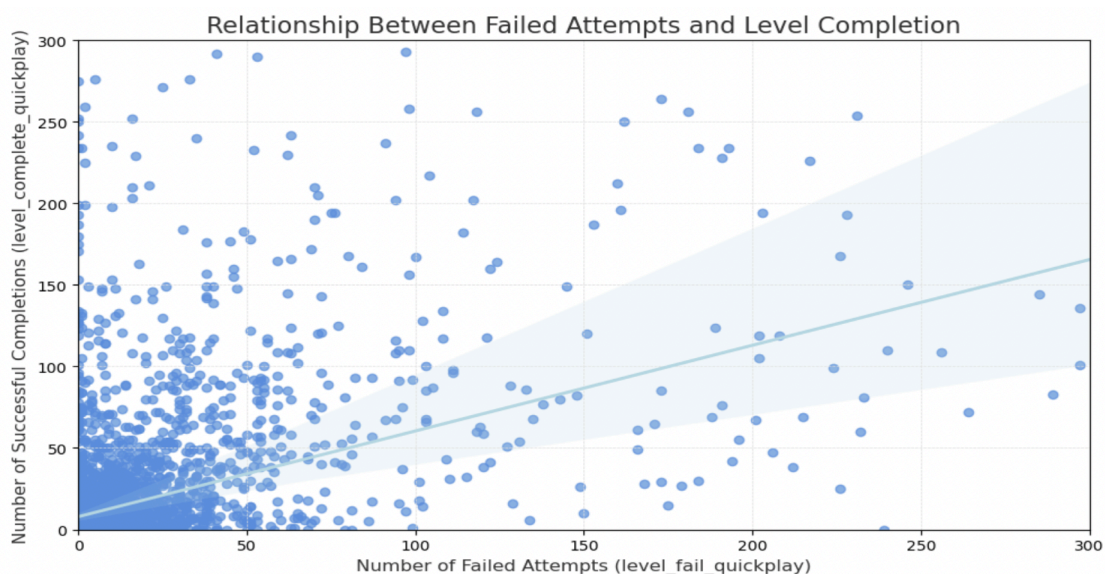
## Recommendations:

- ❖ Enhance the user experience in the **screen view** and **user engagement** stages to keep users engaged and motivated to proceed with the game.
- ❖ Focus on improving the **level start quickplay** to **level complete quickplay** flow by ensuring these stages are seamless and engaging.
- ❖ Use insights from **post\_score** and **level reset quickplay** events to identify and address obstacles that may hinder users from completing levels.
- ❖ Consider adding rewards, hints, or tutorials at key stages to support users, especially during challenging moments.

**Is there a relationship between the number of failed attempts (level\_fail\_quickplay) and successful level completions?**

## Goal:

To investigate whether repeated attempts (failures) are correlated with successfully completing a level. This analysis helps to understand if persistence plays a role in success or if users tend to disengage after failures.



### Insights:

- There is a **positive correlation** between failed attempts and successful completions. This indicates that users who fail more often are likely to eventually succeed.
- However, the trend suggests that most users complete levels after a relatively low number of failed attempts, while very high failure counts are rare.
- The data also shows significant variance, meaning there are exceptions to this trend where some users fail multiple times but never complete the level.

### Recommendations:

- ❖ Provide motivational feedback or hints to users after a certain number of failures to encourage them to keep trying.
- ❖ Consider analyzing **failure patterns by level type or difficulty** to better understand where users face challenges.
- ❖ Design interventions for users who fail multiple times in quick succession, such as offering in-game assistance, tutorials, or simplified versions of the level.

## What is the probability of users returning to the game after a few days?

### Goal:

To analyze user retention over several days and understand how many users continue playing after their first day.



### Insights:

- **Sharp Drop in Retention:** The retention rate drops to about 50% by the second day and decreases further to around 30% by the third day. By the seventh day, the retention is very low.
- **Standard Retention Pattern:** This follows the common “Retention Curve,” where most users leave after the first few days, and only a small percentage continue to engage.

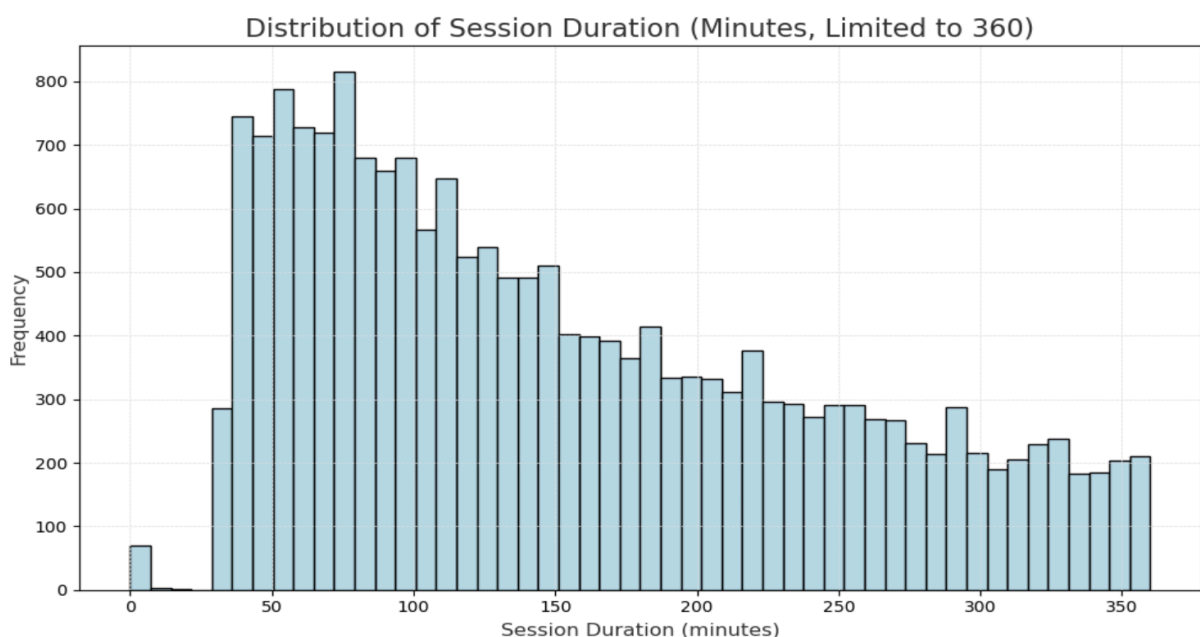
### Recommendations:

- ❖ **Improve First Day Experience:** Make the first experience for users engaging and easy to understand. Add tutorials or features that motivate users to play more.
- ❖ **Introduce Incentives for Returning:** Send reminders through push notifications or emails to bring users back on the second and third days. Offer bonuses or rewards for returning users.
- ❖ **Analyze Engaged Users:** Study the behavior of users who return after the first day and replicate successful strategies for other players.
- ❖ **Enhance Long-Term Engagement:** Add new features or game mechanics to maintain user interest over a longer period.

## What is the distribution of session durations for sessions lasting up to 360 minutes (6 hours)?

### Goal:

To analyze the session duration distribution for sessions under 6 hours to identify user engagement trends and recommend strategies for improving session quality.



## Insights:

- The majority of session durations are concentrated between **50 and 150 minutes**, indicating that most users engage in medium-length sessions.
- The frequency of sessions decreases steadily for durations longer than 150 minutes, but there is still a noticeable number of sessions extending to the limit of 360 minutes.
- Users appear to engage actively in sessions of moderate length, with longer sessions being less common but still significant.

## Recommendations:

- ❖ **Focus on improving medium-length session experiences:** Since most users have sessions between 50 and 150 minutes, enhance features that maintain engagement during this time frame, such as content updates or in-game rewards.
- ❖ **Investigate long sessions (300–360 minutes):** Evaluate user behavior during long sessions to ensure sustained engagement and rule out potential usability or technical issues.
- ❖ **Incentivize longer session durations:** Introduce mechanics like rewards for completing specific milestones that encourage users to extend their sessions.
- ❖ **Monitor drop-off points:** Identify where and why users tend to end sessions within the 50-150 minute range to address any usability issues or optimize retention strategies.