

Clash Royale Data Analysis Report

Clash Royale, the game that I was playing when it was launched in 2016, the case is really interesting for me since most probably my data is also in it. There are three datasets provided to us belonging to the data in 2016. These datasets are named as account, iap_purchase and account_date_session.

Account Dataset

This dataset encompasses the account creation times, devices, platforms, countries, and app store IDs of 112,792 users. There are only 107 missing values for the country codes of some users; aside from that, the dataset is complete with no other missing values. I have created a plot that showcases the top 10 devices used by players to create their accounts, as illustrated in the Figure 1. Interestingly, during that period, I was also an active player, using my iPhone 7. All of the top 10 devices used for account creation are iPhones and iPads, with the iPhone 7, 2 leading the list, associated with 2,764 accounts.

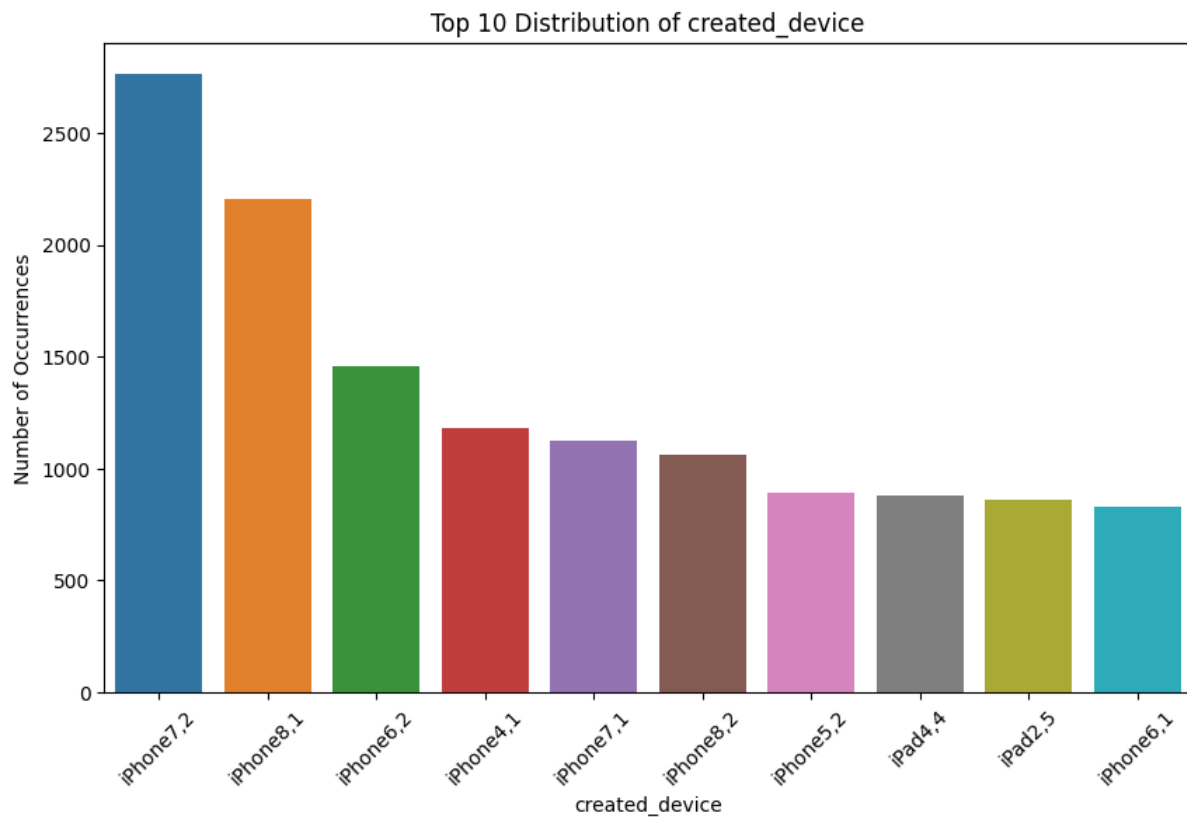


Figure 1: Top 10 Distribution of devices that users used to create their accounts

An additional visualization was created to analyze the distribution of account creations by day of the week. Saturday saw the highest number of new accounts, totaling 20,303, closely followed by Sunday with 18,549. Tuesday had the fewest, with 14,109 accounts created shown in Figure 2.

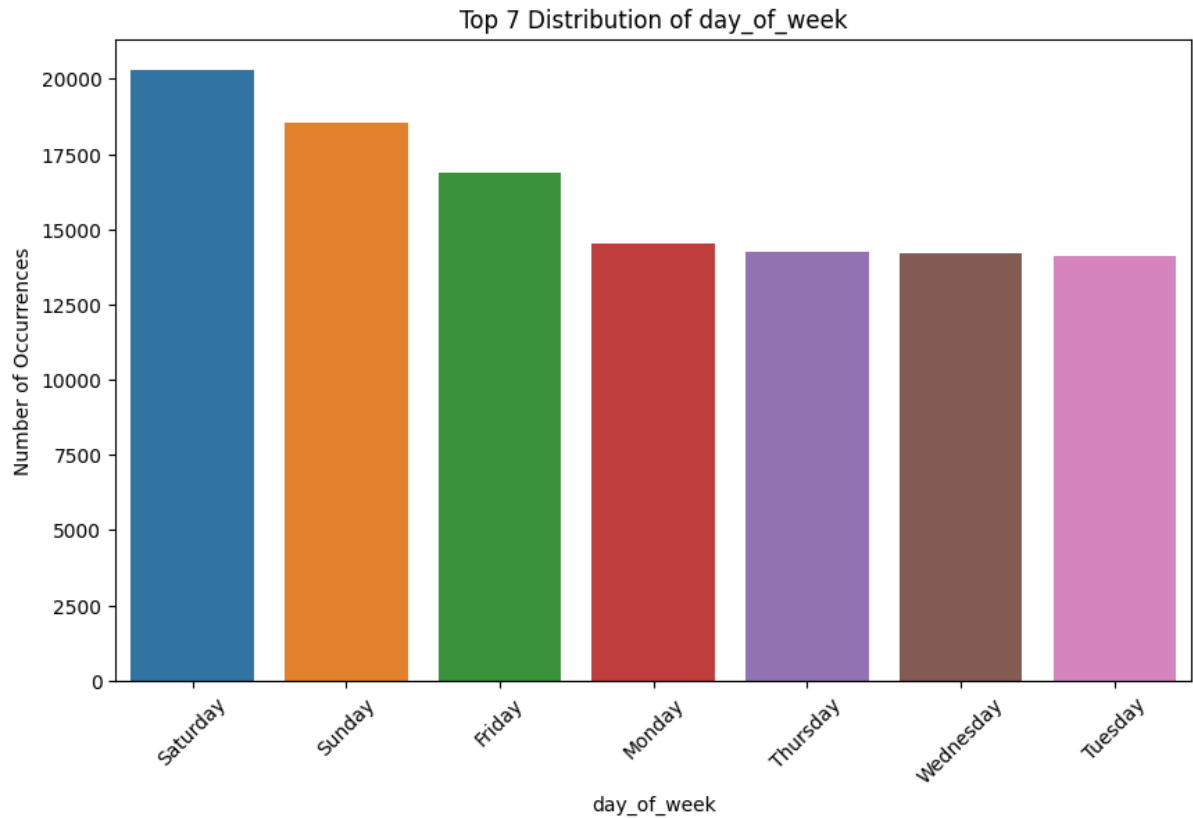


Figure 2: Day of week distribution on account creation

Among all the registered accounts, China boasts the highest number of users at 38,044, followed by the USA with 11,553 users, and Turkey with 5,060 users, one being me. To facilitate a better geographical analysis, I have created a world map visualization shown in Figure 3. This map is logarithmically scaled to accurately represent the wide range of user counts, and countries depicted in gray indicate regions where no player accounts were created in 2016. Interestingly, the map reveals that, with the exception of some African nations, people from virtually every part of the globe downloaded and enjoyed Clash Royale in its first year.

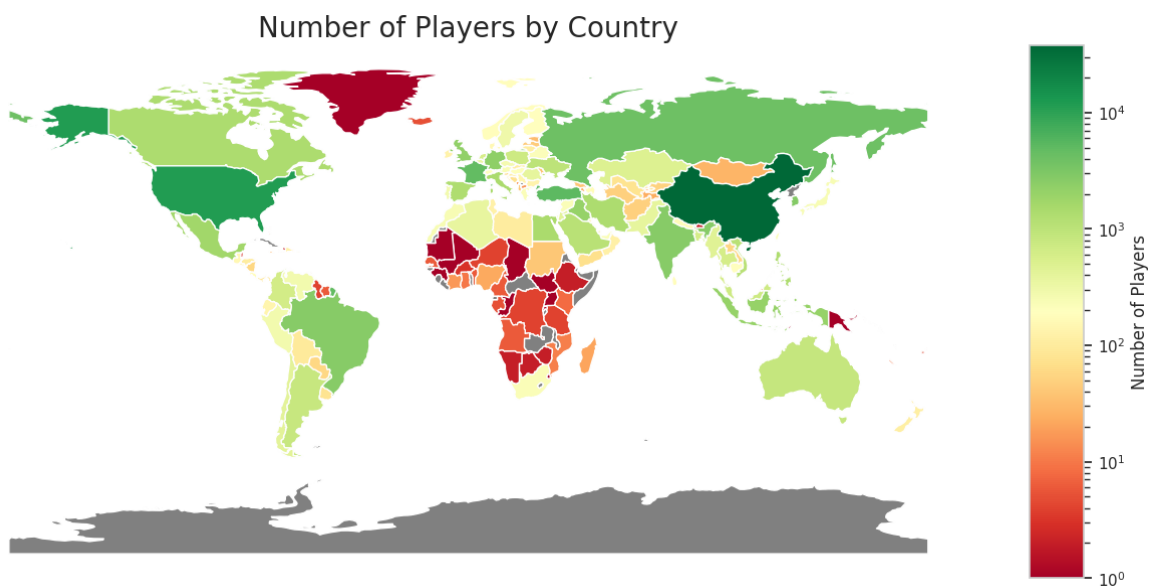


Figure 3: Number of players distribution all over the world

Given that the dataset includes date and time information, it allows for an observation of the times of day when players are most likely to create their accounts. The pattern of account creation times resembles a normal distribution, peaking around 12-13 hours (midday to early afternoon) and tapering off as the time moves away from this peak period shown in Figure 4.

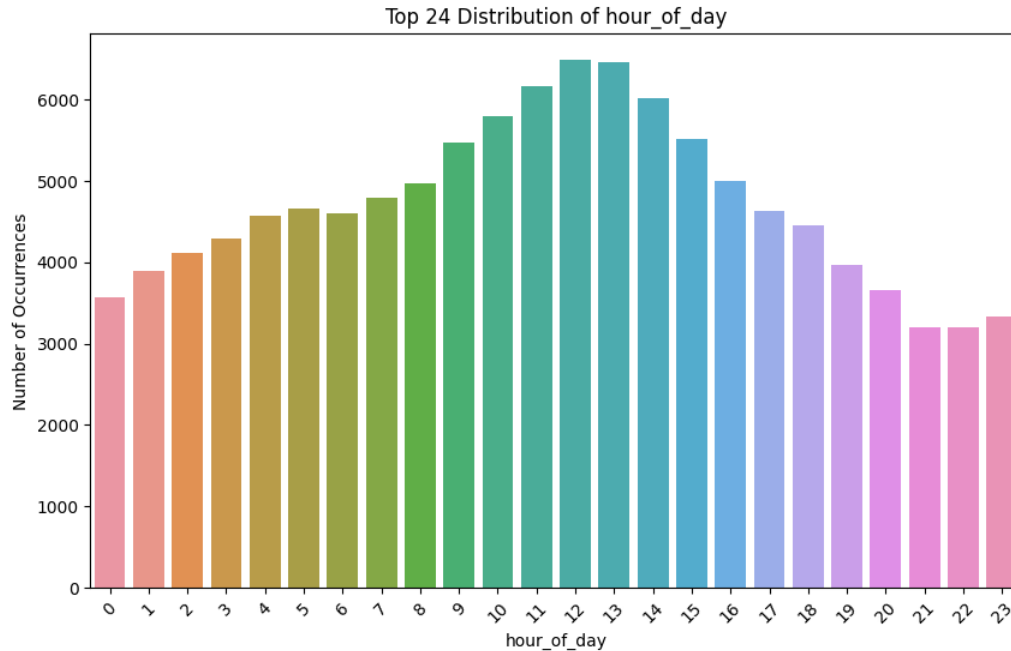


Figure 4: Time of the day for account creations

Iap_purchase Dataset

The dataset records 9,909 transactions, detailing account IDs, timestamps, amounts in US cents, and app store IDs. Notably, there are only seven distinct transaction amounts, which suggests that these can be categorized and visualized effectively using a bar chart shown in Figure 5. Among these transactions, 292 were valued at \$36.99, 369 at \$18.49, 1,426 at \$7.39, and 2,073 transactions were 3,69 dollars.

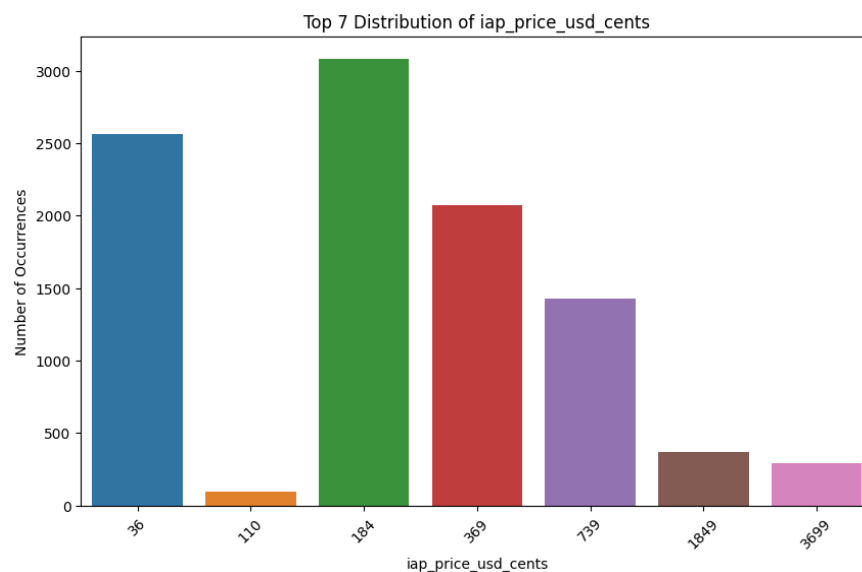


Figure 5: Distribution of transaction amounts

Although China has 38,044 users and the USA has 11,553 users, an examination of the total amount of in-app purchases reveals that the USA ranks first with \$13,148.18, while China follows with \$11,365.07. South Korea, Canada, and Turkey also feature prominently in the rankings with \$4,635, \$2,545 and \$1,933 respectively shown in Figure 6. This time, I don't have my data in it since I was a broke student during those times.

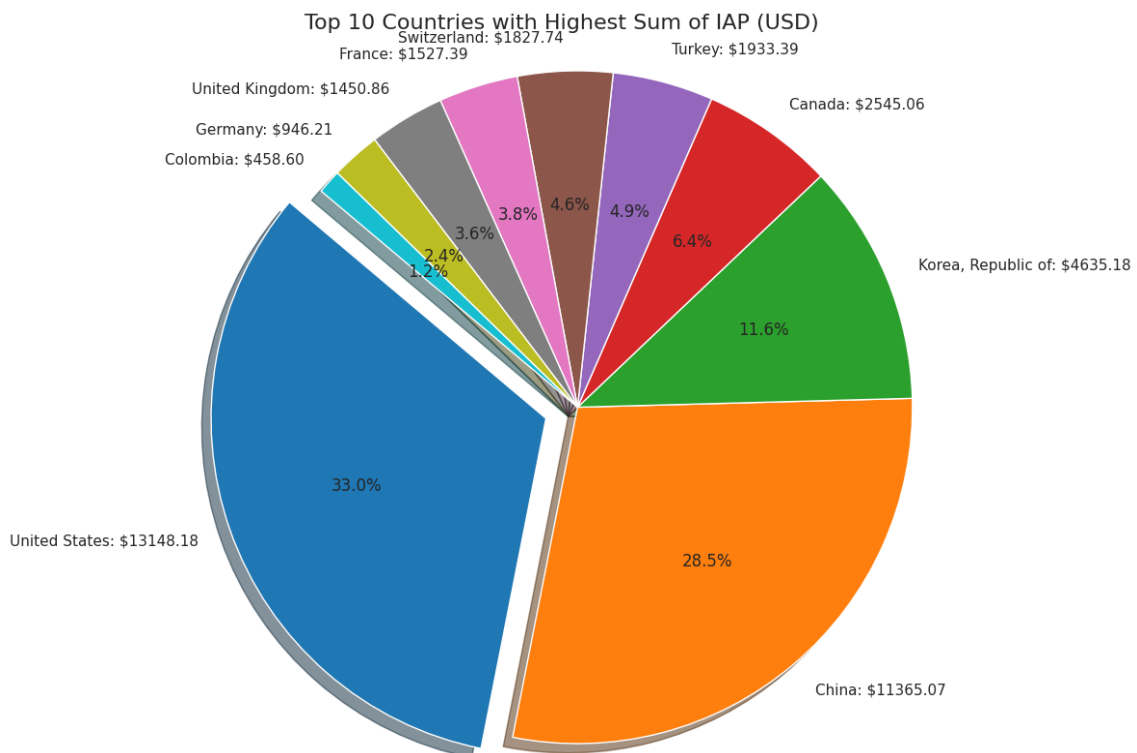


Figure 6: Pie chart for top 10 countries with highest sum of IAP

Upon visualizing the dataset on a world map, it becomes apparent that many countries with lower economic status, including nearly all of Africa and several South American nations, have no recorded in-app purchases.

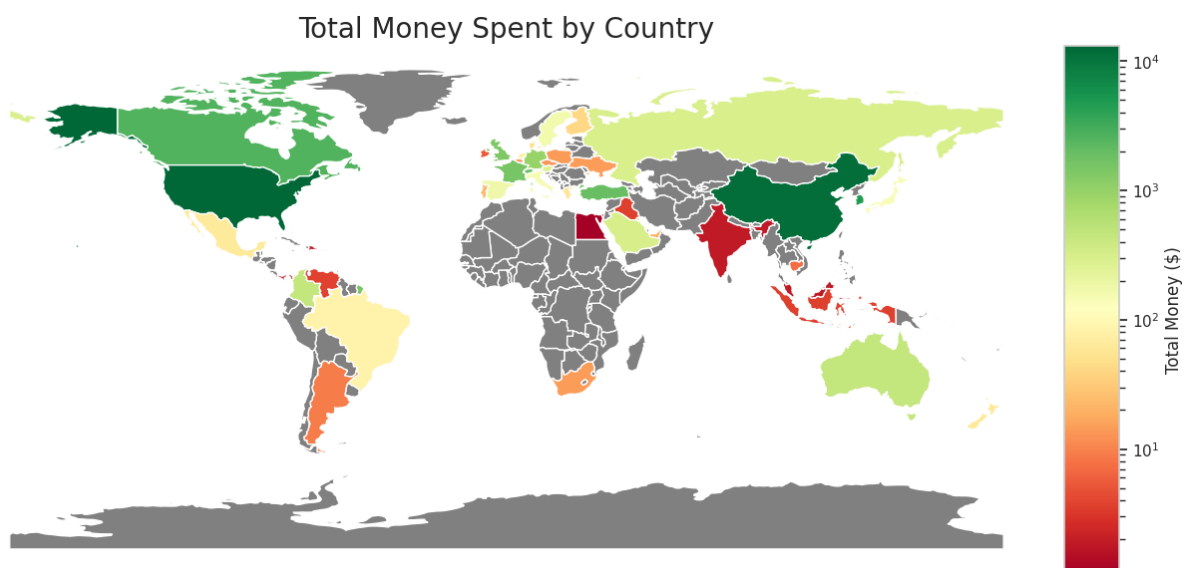


Figure 7: Total in app purchases distribution over the world map

Account_date_session Dataset

This dataset, comprising 1,698,974 entries, is instrumental for analyzing daily active users and their session durations. Each record corresponds to an individual user, identified by their account ID, and includes data on the date, number of sessions, and the length of the session duration.

DAU Changes Over Time

The dataset detailing the account sessions spans from January 1st, 2016, to December 31st, 2016. There is a noticeable surge in the number of daily active users from January to March, after which the count stabilizes at approximately 5,000, exhibiting weekly fluctuations. This weekly pattern fades during the summer, particularly in July and August, and re-emerges in September. The graph is shown in Figure 8.

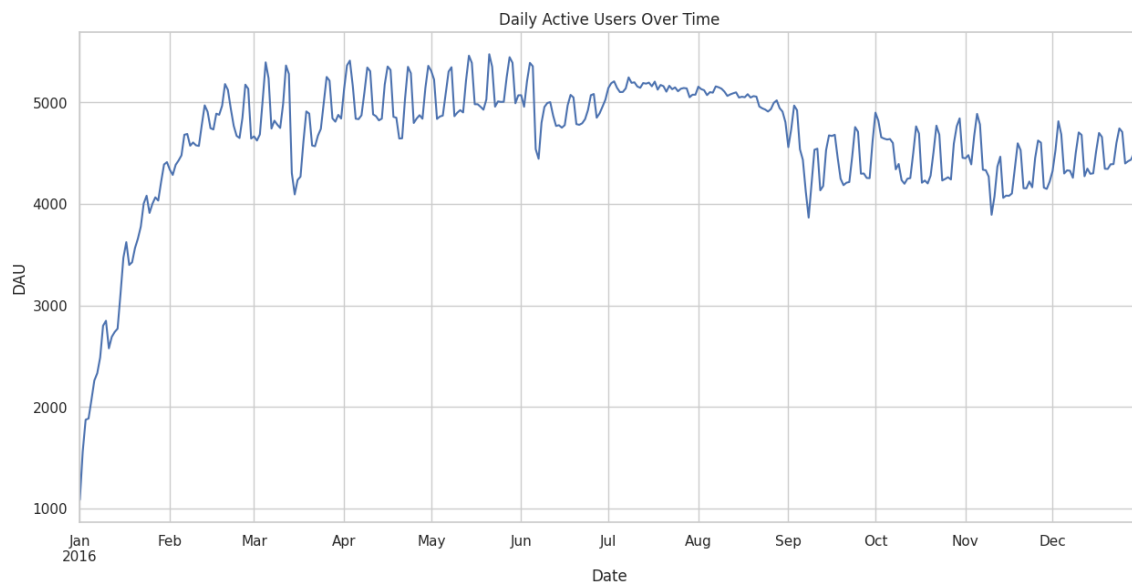


Figure 8: Daily Active Users over the year

To understand the weekly fluctuations and the summertime change, I analyzed the data by day of the week. It became apparent that Saturdays and Sundays attract significantly more daily active users compared to weekdays, with Fridays also showing higher activity than other weekdays shown in Figure 9. However, during the summer, the daily user count tends to level out across the week. My hypothesis for this trend is that the game is predominantly played by students who, during the school year, play mostly on weekends and Fridays. In contrast, during the summer break, the day of the week has little impact on their gaming habits.

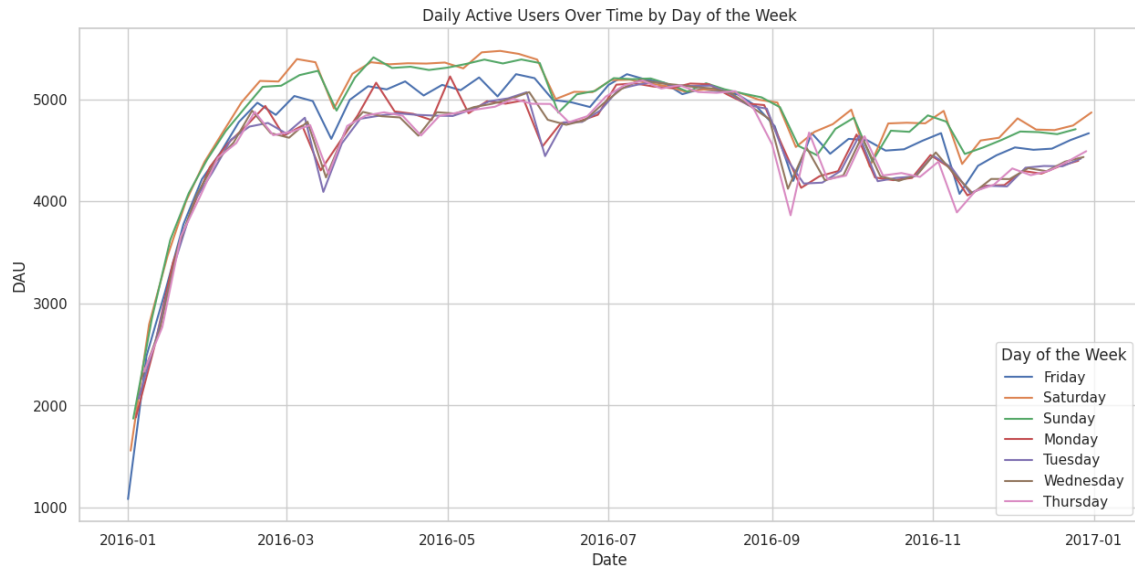


Figure 9: Daily active users in days of the week over the year

Upon conducting the same analysis with a focus on total session duration rather than daily active users, a similar seasonal trend is observed as shown in Figure 10. Notably, certain Saturdays stand out with the longest session durations, a pattern that is particularly pronounced in certain weeks, especially the first week of July and the last week of October. This could potentially be due to the introduction of tournaments or significant updates in the game during those specific weeks.

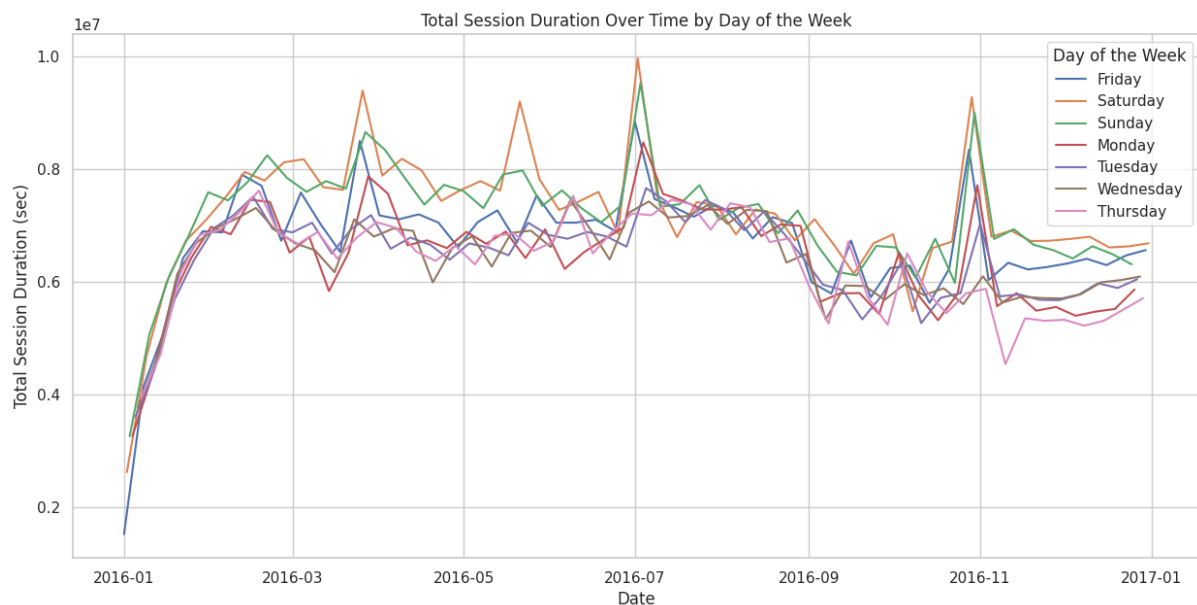


Figure 10: Session Duration in days of the week over the year

There are also some unusual decreases, and I suspect the reason for that may be maintenance periods in the game. Therefore, I looked into Clash Royale's social media pages and discovered three instances of maintenance in 2016, the screenshots from those dates shown in Figure 11.



Figure 11: Maintenance times shared in social media in 2016

Nevertheless, the maintenance appears to have little impact on the overall playtime, as indicated by the graph in Figure 12. As a Clash Royale player, I'm aware that we receive a notification once maintenance concludes, prompting us to jump right back into the game. This likely means the maintenance breaks were brief, and players continued playing afterwards, resulting in play sessions that are comparable in length to those on non-maintenance days.

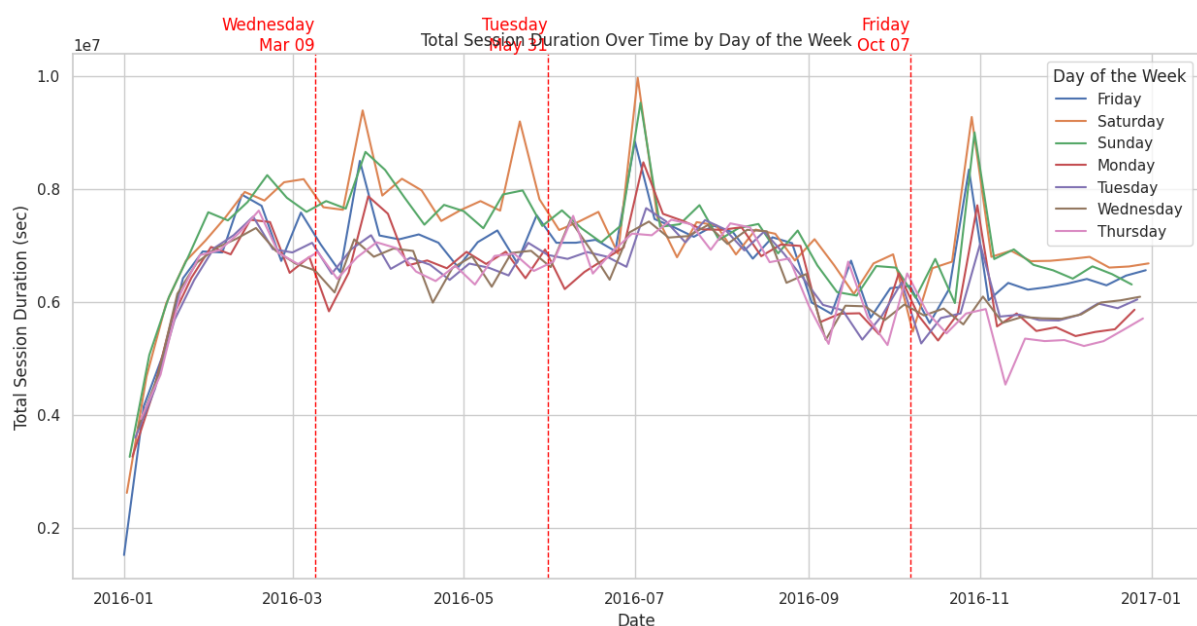


Figure 12: Session Duration in days of the week over the year maintenance times highlighted

IAP Changes Over Time

The in-app purchase (IAP) data for the year 2016 presents a fascinating insight into user spending behaviour and the game's revenue patterns shown in Figure 13. The data reveals significant variability in daily revenue, with some days showing modest earnings, while others exhibit substantial spikes. Notably, the highest revenue day falls on January 31st, with a remarkable \$1,202.61, a figure that dwarfs the surrounding days' earnings. This peak could be attributed to special events, in-game promotions, or updates that typically drive higher user engagement and spending. The days leading up to this spike show a gradual increase in revenue, suggesting a buildup of user anticipation or the successful culmination of a marketing campaign.

Analyzing the data for actionable insights, it's evident that certain days and periods are more lucrative than others. For instance, the beginning of February shows a sharp decline post the January peak, but there is a quick recovery with consistent high points, such as on February 1st and 4th, where revenue again surpasses the \$700 and \$300 marks, respectively. This pattern indicates that maintaining user interest through regular updates or events could be a key strategy for sustained revenue generation. Moreover, the end of each month seems to show an uptick in spending, which could guide the timing of new content releases or special offers. Overall, the data underscores the importance of strategic planning in in-game event scheduling and targeted promotions to capitalize on high-spending periods and maximize revenue potential for Clash Royale.

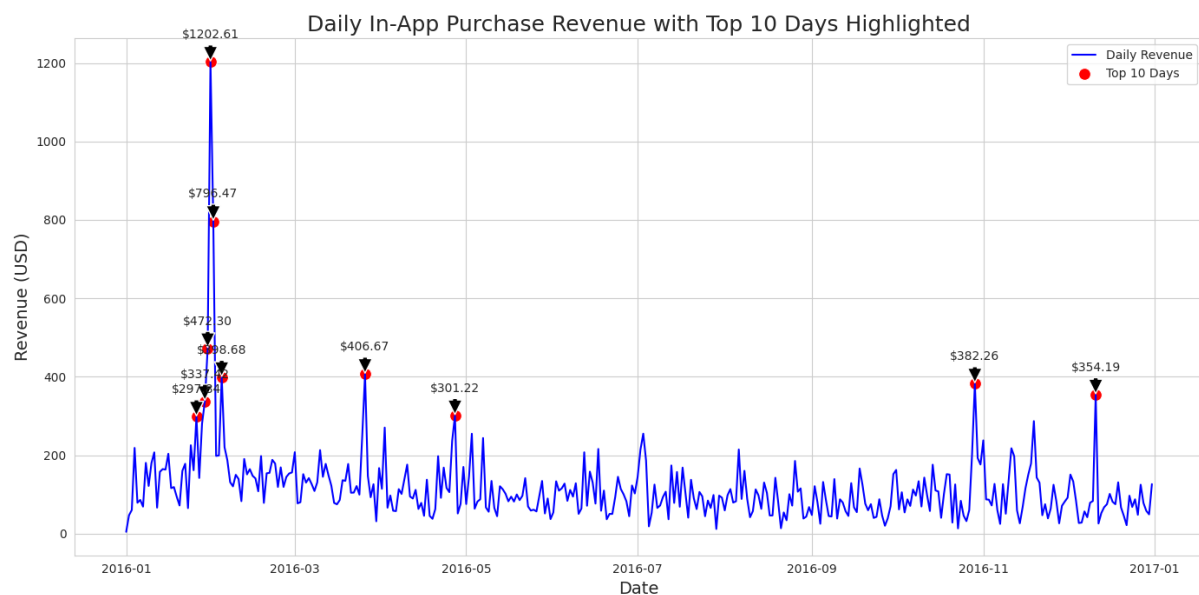


Figure 13: Daily IAP revenue over the year with top 10 highest revenue days highlighted

Average Revenue Per User Per Market

The data on Average Revenue Per User (ARPU) for in-app purchases in Clash Royale reveals a striking contrast in spending habits across different countries. High ARPU values in nations like Switzerland (\$9.23), Canada (\$1.85), and the USA (\$1.14) indicate a greater capacity or willingness of users in these areas to spend on in-app items shown in Figure 14. On the other hand, many countries, including large markets such as India and China, show very low to no ARPU, which could point to a preference for non-spending gaming experiences or different financial priorities. The

reasons for these differences could include economic conditions, cultural attitudes towards gaming expenses, competition from other forms of entertainment, and disposable income levels.

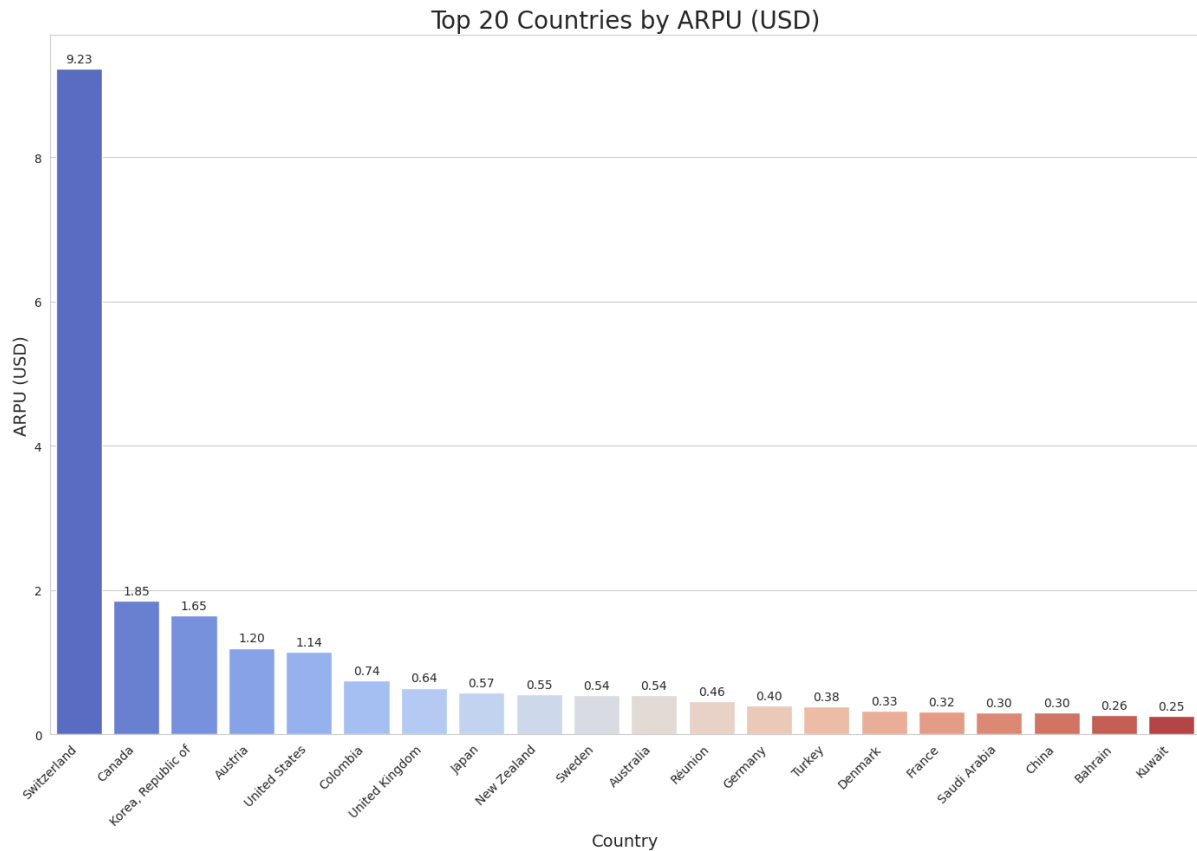


Figure 14: ARPU values for top 20 highest countries

Regarding strategic initiatives, these findings serve as a critical foundation for tailoring marketing initiatives and pricing frameworks to specific regional contexts. For territories demonstrating elevated ARPU, developers might consider introducing higher-tier in-app purchase options or exclusive subscription services that align with the observed consumer spending patterns. Conversely, in locales where ARPU is minimal, a strategy that emphasizes revenue generation through advertising or microtransactions, which may be more frequent yet financially accessible, could be more appropriate. This aligns with the purchasing capacity of the demographic in these regions. It has been noted that Clash Royale has adapted its financial strategies to accommodate regional differences, a practice that was not evident during the game's inaugural year. This initial oversight resulted in significantly lower ARPU in economically challenged countries, with India's ARPU exemplifying a mere \$0.0006 and more red countries with very low ARPUs can be seen in Figure 15. A thorough examination of the variables that affect these expenditure patterns is crucial. Factors such as payment infrastructure accessibility, local legislative frameworks, and the relative value of local currencies must be considered to tailor monetization strategies effectively for each distinct market. The primary objective is to maximize revenue potential while concurrently fostering a gratifying and immersive gaming experience for a global audience.

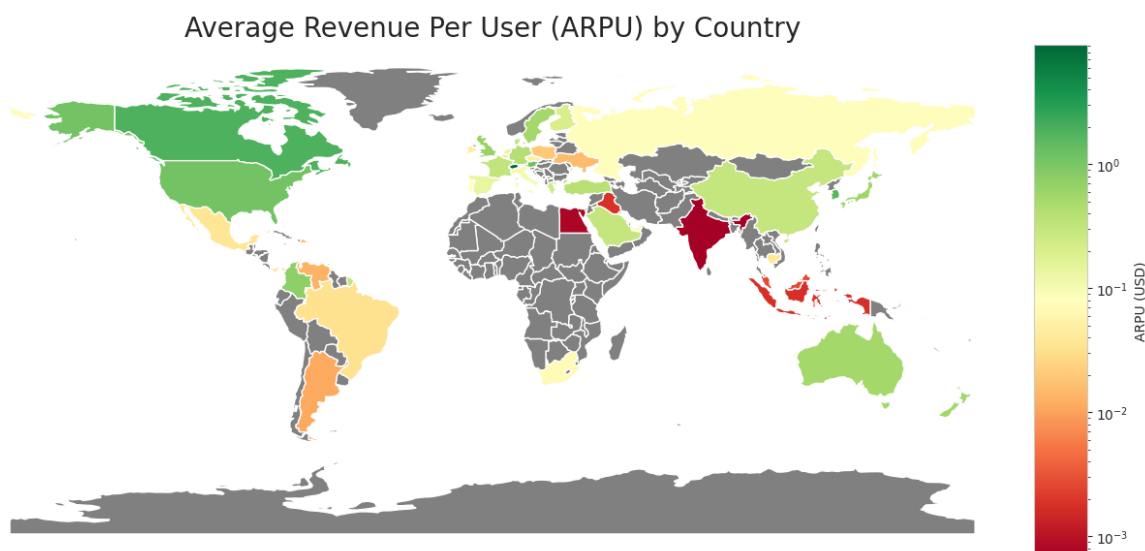


Figure 15: ARPU distribution on the world map visualization

Top Spenders

Number of users who spent over \$100 for Clash Royale in 2016 has been shown in the following bar plot in Figure 16. The user who spent the most amount of money is from Korea with the value 2,470.76 dollars, followed by a user from the USA with 1,468 dollars and a user from China with 1,269.11 dollars. The total number of users who spent over \$100 was 85.

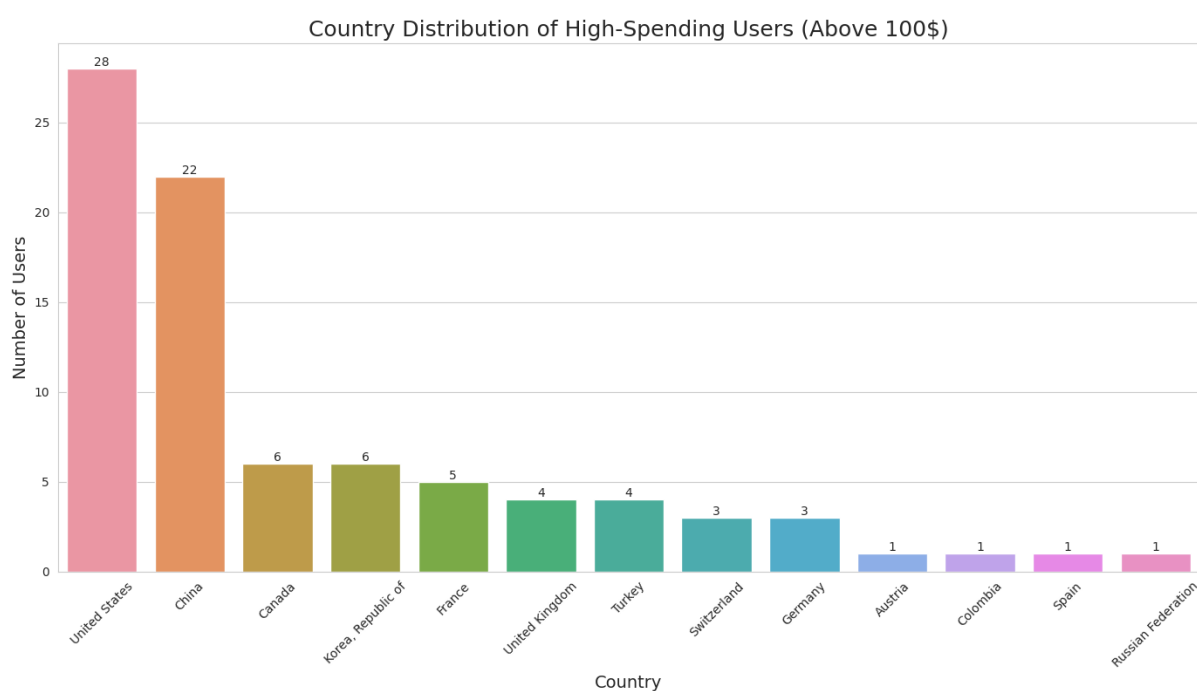


Figure 16: Number of players who spent over 100\$ IAP

Correlation Between User Activity and In-App Purchases

My initial assumption posited that a direct correlation existed between user engagement—as measured by session counts and durations—and in-app spending. However, upon constructing a correlation

matrix shown in Figure 17, it became evident that the relationship between in-app purchases and both session count and duration is surprisingly tenuous. The data indicates that the predominant factor influencing in-app purchase behavior is not user engagement, but rather the geographical location of the users. This insight suggests that regional factors may have a more substantial impact on spending habits within the app than previously understood.

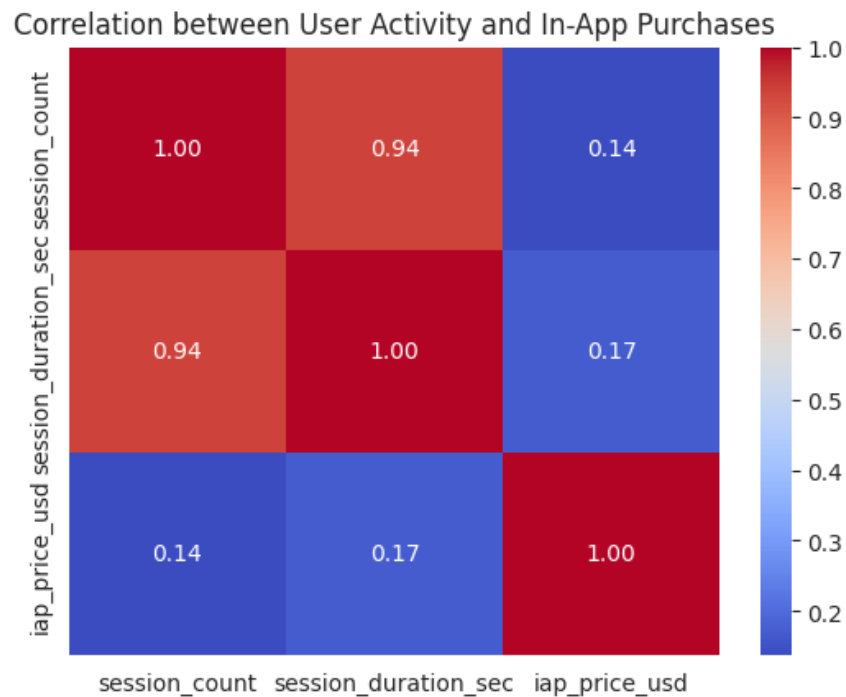


Figure 17: Correlation matrix over User Activity vs IAP

Implementation

I endeavored to maintain clarity and readability in the code through the use of comprehensive comments. Additionally, I have employed function definitions to minimize redundancy, ensuring that the code is not only efficient but also accessible for future reference and modifications.

 Clash_Royale.ipynb