

# Analysis of Mobile Game Data

Jere Arokivi

November 4, 2023

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Data Description</b>	<b>2</b>
2.1	Dataset Overview . . . . .	2
2.1.1	Account Table . . . . .	2
2.1.2	Account Date Session Table . . . . .	3
2.1.3	In App Purchase Table . . . . .	3
<b>3</b>	<b>Daily Active Users Analysis</b>	<b>4</b>
3.1	Temporal Trend in DAU . . . . .	4
3.2	Trends Identification . . . . .	5
3.3	Anomalies Detection . . . . .	6
<b>4</b>	<b>Sales Analysis</b>	<b>7</b>
4.1	Geographic Revenue Split . . . . .	7
4.2	Average Revenue Per User . . . . .	8
4.3	Observations . . . . .	10
4.3.1	User Activity . . . . .	10
4.3.2	Sales Analysis . . . . .	10
<b>5</b>	<b>Conclusions</b>	<b>10</b>
5.1	Strategic Implications . . . . .	10
5.2	Revenue Optimization . . . . .	10
5.3	Recommendations . . . . .	11

# 1 Introduction

This report presents a detailed analysis of user activity and sales data from a hypothetical free-to-play mobile game. Utilizing analytics, we extract trends, identify seasonal patterns, and detect anomalies within the data. The report offers actionable insights for game developers and business strategists aiming to enhance player engagement and revenue generation. The insights and visualizations were generated using Python.

## 2 Data Description

### 2.1 Dataset Overview

The dataset was very well cleaned. The structured and comprehensive nature of the data made it an efficient and thorough analysis. The data included user accounts, session details, and in-app purchase information, providing a broad view of user behavior and spending patterns.

Accounts were registered from a total of 187 countries, out of which 107 records were identified as having missing country values. The majority of these accounts were associated with a `create_app_store_id` of 2. The top three countries in terms of account creation were China, the United States, and Turkey, accounting for 33.4%, 10.2%, and 4.4% of the total, respectively. Collectively these contribute 48.5% of the overall account creation volume. Distribution-wise, 82% of the accounts were created via Android platforms, with the remaining 18% originating from iOS.

#### 2.1.1 Account Table

Account table contains following columns

- `account_id`: Contains account identifier
- `created_time`: Creation time of the account
- `created_device`: Device which the account was created from
- `created_platform`: Platform which the accounts was created from
- `country_code`: Country of the account
- `created_app_store_id`: App store version when the account was created

### **2.1.2 Account Date Session Table**

Account\_date\_session table contains following columns

- account\_id: Contains account identifier
- date: Date when session took place
- session\_count: Amount of sessions for the days user was active
- session\_duration\_seconds: How much time user spent during the sessions

### **2.1.3 In App Purchase Table**

Iap\_purchase table contains following columns

- account\_id: Contains account identifier
- created\_time: Time of the transaction
- package\_id\_hash: Package identifiers hash
- iap\_price\_usd\_cents: Transaction price in cents of USD
- app\_store\_id: App store identifier

### 3 Daily Active Users Analysis

#### 3.1 Temporal Trend in DAU

In-depth time series analysis was conducted to find the patterns in Daily Active Users (DAU). The analytical process involved decomposing the time series to outline underlying trends and seasonal fluctuations.

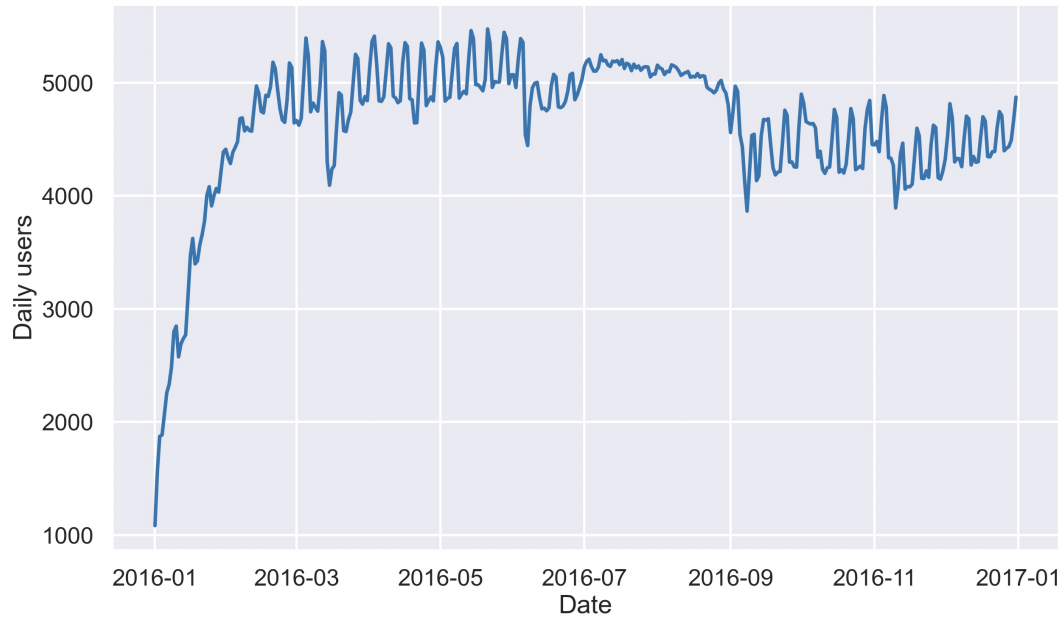


Figure 1: Trend of Daily Active Users Throughout 2016

A careful examination of the DAU time series revealed noticeable patterns of trend and seasonality, prompting the utilization of an additive time series decomposition model. This model was selected to isolate the trend, seasonal, and residual components, thereby facilitating a analysis of the natural patterns in the data.

The additive decomposition results, as illustrated in Figure (2), verify the robustness of the model fit. The RMSE of 107.82, coupled with a mean absolute error of 1.4, and a relative error margin approximating 2% at peak values, collectively confirming the model's efficiency in isolating the DAU trends with high accuracy.

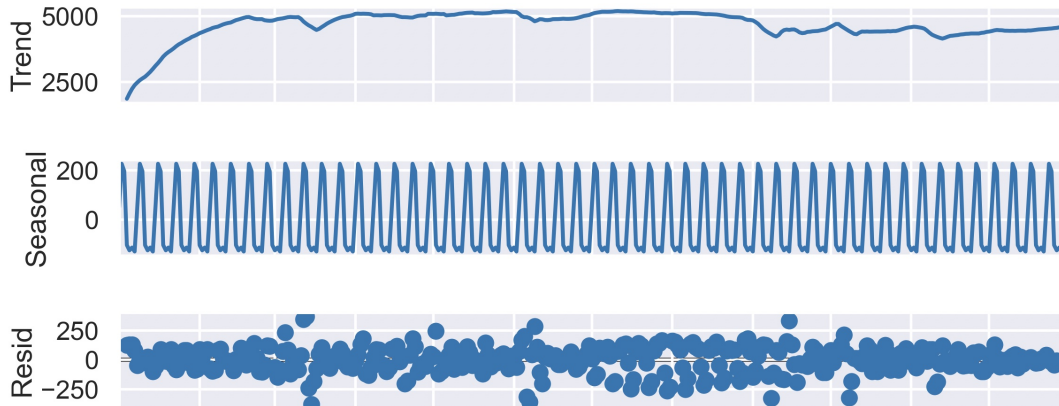


Figure 2: Decomposition of Daily active users

### 3.2 Trends Identification

Analysing the **trend** component of the decomposition from figure (2) we can see clear rise in users during first 2 months of the year, which after the daily users settle around 5000 users a day without much fluctuation in daily users apart from small drop in users mid March and around September the overall trend decreases to around 4500 daily users.

From **seasonal** component of the decomposition we notice clear seasonality and with further analyzing the seasonality happens to split into 52 cycles within the year indicating a weekly trend in data further. By observing 1 cycle from the seasonal component of the decomposition we get seasonal weekday trend shown in figure (3).

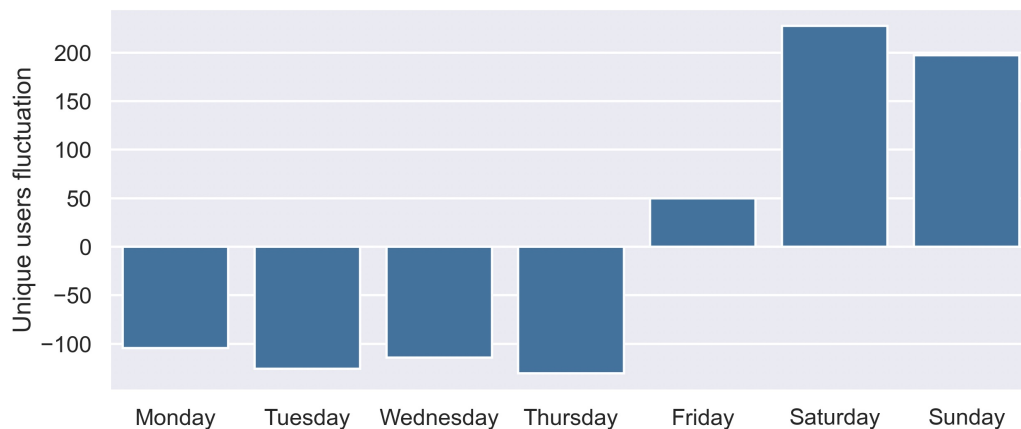


Figure 3: Fluctuation over weekdays from moving average

This indicates that there is less active daily users during workdays and more during weekends and this observation is as expected, users have more time and energy to play games during weekends.

### 3.3 Anomalies Detection

In-depth **residual** analysis of figure (2) has revealed a significant shift in user behavior in the early to mid-March period. This is corroborated by the DAU metrics (1) and the decomposition trend line (2). The fluctuation may be attributed to the launch of the Clash Royale, in March 2016, which could have diverted attention from our subject game. Additional factors such as competing game events or less favorable game updates during this period may have also contributed to the observed decline in user engagement.

A following anomaly is observed from June to August, characterized by a marked reduction in user activity commencing in June. This may correspond with Tencent’s acquisition of a significant share in Supercell Securities, which occurred in June 2016. Simultaneously, the onset of the summer holiday season could have influenced this trend, as user engagement patterns traditionally fluctuate during this period.

Further observation within this anomaly period reveals diminished volatility in the weekday imbalances, particularly during July and August. This pattern aligns with the typical summer holiday season, suggesting that users may have more discretionary time to engage with games during traditional working days.

The final noteworthy anomaly detected occurs at the onset of November. Preliminary analysis suggests a potential correlation with the U.S. Presidential elections, an event that could potentially impact digital consumer behavior due to its significant media coverage and public involvement.

## 4 Sales Analysis

### 4.1 Geographic Revenue Split

The scope of the revenue analysis was confined to the top 20 contributing countries, which account for 98.8% of the total revenue. This strategic focus allows for a detailed examination of the markets that are most impactful to revenue streams. The analysis divided the revenue sources into two primary classifications: by individual country and by continent.

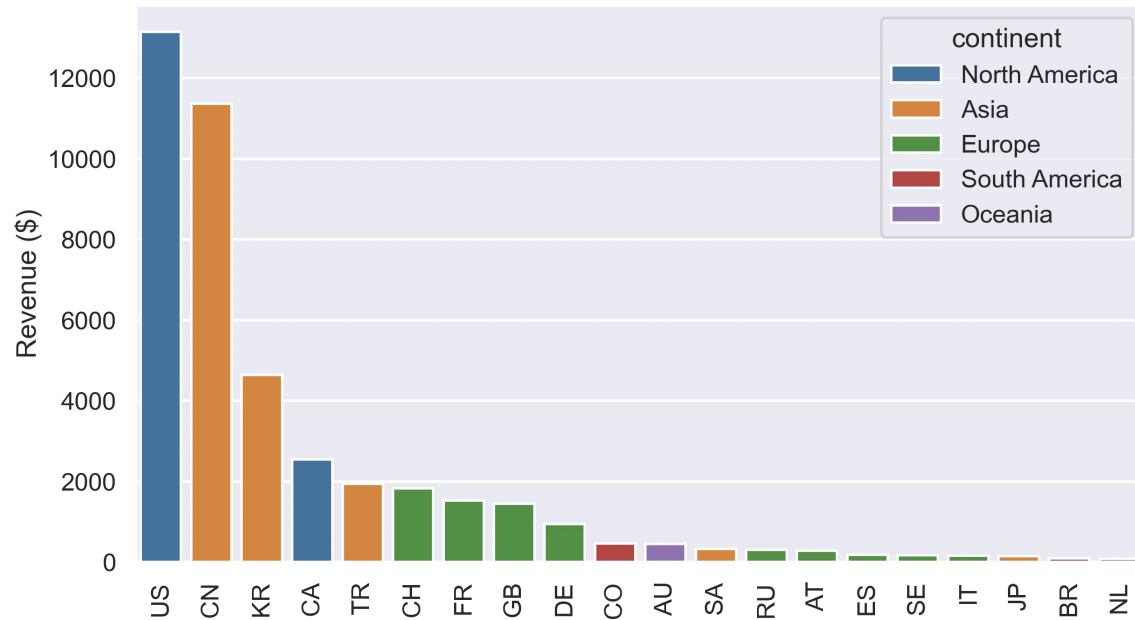


Figure 4: Distribution of Revenue Among Top 20 Countries

An examination of country-specific revenue reveals a highly concentrated distribution, with the top three countries—China, the United States, and South Korea—contributing 30.9%, 26.7%, and 10.9% of total revenue, respectively. Collectively, these countries are responsible for 68.5% of total revenue, highlighting a significant market concentration.

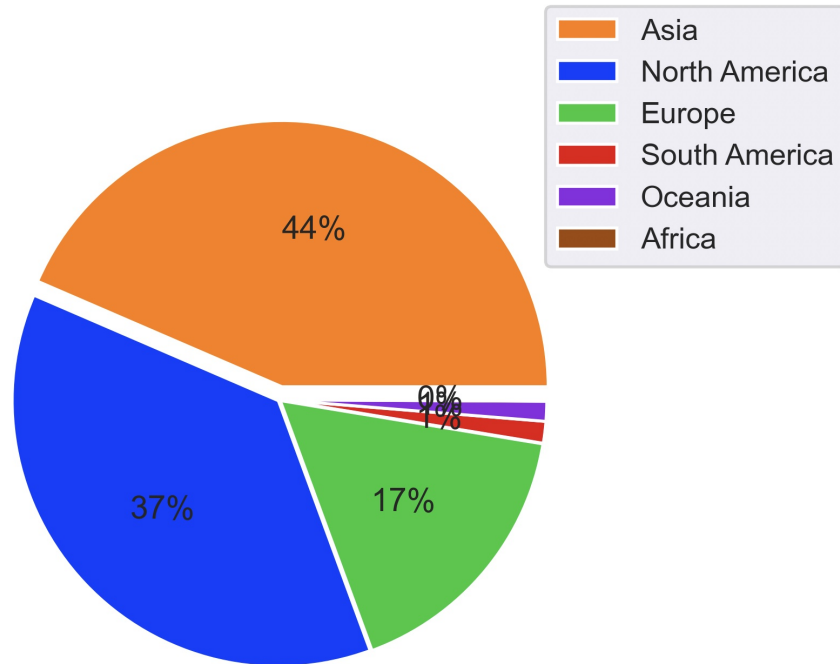


Figure 5: Proportional Revenue Contribution by Continent

Continental analysis confirms that the majority of revenue is sourced from Asia and North America, comprising 44% and 37% of the total revenue, respectively. Europe follows with a substantial 17% contribution. The remaining markets South America, Oceania, and Africa collectively make up a minor 2% of sales, suggesting a potential area for market development and growth strategies.

## 4.2 Average Revenue Per User

The first step in the analysis involved validating the suitability of the data for assessing ARPU by examining the mean and standard deviation of the revenue metrics for users. The presence of significant standard deviations suggested potential outliers within the dataset. Therefore, a distribution analysis was conducted, revealing an exponentially decreasing trend in the revenue density function per user. To address the outliers, quartile-based range filtering was implemented, utilizing the equation for the upper bound:  $upper\_bound = Q3 + 1.5 \times IQR$ , where  $Q3$  denotes the third quartile and  $IQR$  is the interquartile range.



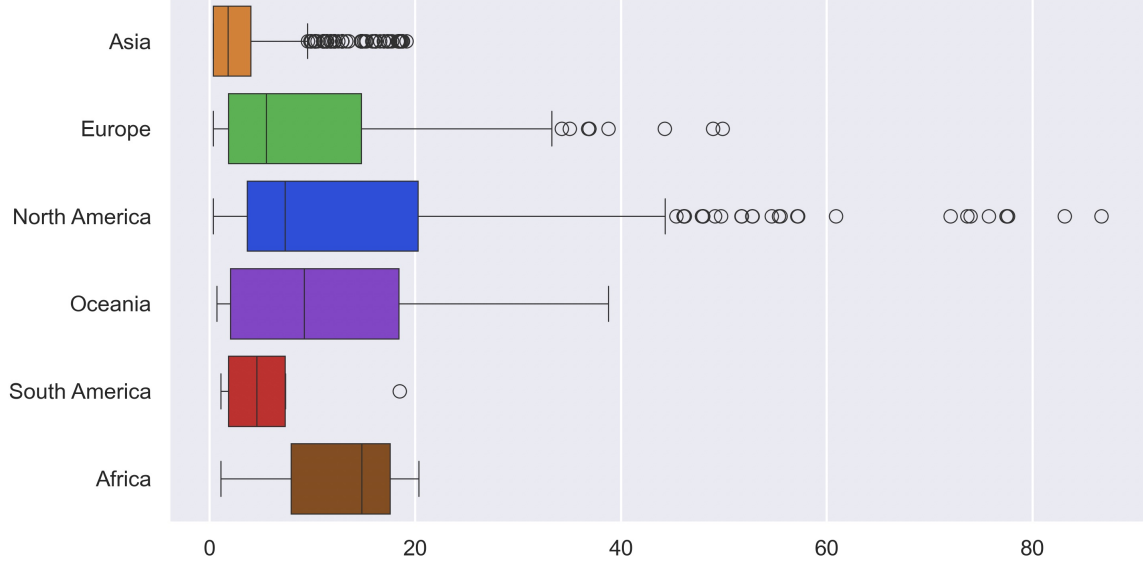


Figure 6: Distribution of Revenue(\$)<sup>2</sup> Per User by Continent

The data, as visualized in Figure (6), along with the statistical overview in Table (1), indicate that users from Asian markets typically engage in transactions of lesser value yet with greater consistency. In contrast, European and North American markets exhibit a rightward skew in the distribution, suggesting a subset of users with significantly higher spending, despite the median remaining on the lower end. Caution is advised when interpreting the data from South American and African markets due to the smaller sample sizes. Nonetheless, it is noteworthy that the median revenue per user in African markets exceeds that of other continents, while South America presents a comparatively lower median.

Continent	Count	Mean	Std	Min	25%	50%	75%	Max
Asia	858.0	3.376	4.283	0.36	0.36	1.82	4.03	19.14
Europe	186.0	10.031	10.641	0.36	1.84	5.52	14.77	49.87
North America	264.0	15.517	18.857	0.36	3.68	7.36	20.272	86.70
Oceania	23.0	10.883	9.787	0.72	2.02	9.22	18.425	38.77
South America	8.0	5.900	5.649	1.10	1.84	4.605	7.365	18.47
Africa	3.0	12.070	9.897	1.10	7.94	14.780	17.555	20.33

Table 1: Statistical Analysis by Continent

## **4.3 Observations**

### **4.3.1 User Activity**

The time series analysis of daily active users revealed a seasonality, with peaks typically on weekends, reflecting a standard engagement pattern in mobile gaming. The significant dips observed in March and June may correlate with external events such as game releases or macroeconomic factors, suggesting the game's performance is sensitive to market dynamics.

### **4.3.2 Sales Analysis**

The revenue analysis indicates a high concentration of income from a few countries, with China, the USA, and South Korea being the most prominent contributors. This suggests a focused user base where efforts on localization and targeted marketing could drive further revenue growth. The ARPU analysis highlighted variability across regions, with Asian markets spending less per transaction but more consistently, while North American and European users showed higher expenditure but with greater variability.

## **5 Conclusions**

The findings from this report offer actionable insights into user engagement patterns and monetization strategies that can improve future game development and marketing initiatives.

### **5.1 Strategic Implications**

The identified trends and seasonal behaviors provide a opportunity for optimizing content release and promotional schedules to maximize user engagement. The observed anomalies in user activity should prompt further investigation into external factors that may affect user retention and acquisition.

### **5.2 Revenue Optimization**

The revenue analysis underscores the importance of geographic market analysis in revenue optimization strategies. A targeted approach, focusing on high-value markets and user segments, could lead to increased ARPU and overall revenue growth.

### 5.3 Recommendations

Based on the analysis, some possible actions to ensure success in the future

- Enhance user engagement strategies to leverage the identified weekly and seasonal patterns.
- Investigate and respond to external market events that correlate with significant changes in DAU.
- Tailor monetization tactics to regional spending behaviors, optimizing pricing and in-app purchase strategies.
- Consider localized content and marketing efforts in high-revenue markets to capitalize on the existing user base.