Data Mining Analysis of the Global Mobile Game Market

Ji Weng

weng.j1@northeastern.edu

# Abstract

# Keywords

Mobile game, market trend, data mining, clustering

# Introduction

The rise of smartphones and their continuous performance improvements have significantly transformed the gaming industry. Mobile gaming has attracted a wide spectrum of users, including many who were not traditionally part of the gaming audience or who did not own dedicated gaming consoles. These users have now become active participants in mobile gaming, fueling content consumption and market growth.

According to a report by Tianyi Gu, the global mobile gaming market grew by 13% in 2020 due to the impact of COVID-19, with revenues exceeding $77 billion (Tianyi Gu, *COVID-19’s Impact on the Mobile Games Market: Consumer Engagement Spikes as Revenues Exceed $77 Billion*). However, as the pandemic subsided, the market experienced a 5.6% decline in 2022 (Louise Wooldridge, *Games Market Enters a New Era as Pandemic Years Reset Growth Dynamic*).

Over the past five years, the mobile gaming industry has undergone rapid development, characterized by fluctuations in downloads and revenue driven by evolving market trends, technological advancements, and shifting player preferences. Understanding these dynamics is essential for developers, publishers, and investors aiming to navigate this highly competitive landscape.

This project aims to explore global mobile gaming trends from 2020 to 2024. We begin by identifying the major regions contributing to downloads and revenue globally, examining their trends in user acquisition and monetization, and forecasting potential download and revenue figures for 2025.

Next, we conduct a country-level clustering analysis of the global mobile gaming market. We interpret the significance of each cluster and provide in-depth analysis of representative countries within each group, focusing on their Android and iOS market shares, genre-specific performance, and growth patterns.

Additionally, we perform clustering analysis at the game level, grouping mobile games based on key metrics. We then explore the meaning of these clusters and conduct a cross-analysis between game clusters and game genres.

Through these analyses, we aim to uncover macro-level trends in the global mobile gaming market, highlight the characteristics and market shares of key regions, and identify opportunities and emerging patterns within various game genres. These insights will provide industry stakeholders with a comprehensive understanding of the current landscape, enabling them to refine their product development and publishing strategies accordingly.

# Dataset

## Data Sources

Dataset link：

https://github.com/ponyo2024/mobile\_game\_market\_analysis/

This project relies on third-party data platforms to provide comprehensive statistics on the mobile gaming market, particularly:

**Sensor Tower**: For global mobile game downloads and revenue data, including regional and genre-level breakdowns.  
Sensor Tower is the industry’s leading and most widely recognized third-party analytics platform for mobile applications. All raw data used in this project was directly exported from Sensor Tower in CSV format, followed by data cleaning and preprocessing.

## Data Description

The dataset includes:

**Top 10,000 apps globally by revenue for each of the past five years**

5 files in total, one per year

Each file contains the top 10,000 apps by annual revenue

Total: 50,000 rows

**Top 10,000 apps globally by downloads for each of the past five years**

5 files in total, one per year

Each file contains the top 10,000 apps by annual downloads

Total: 50,000 rows

**Top 10,000 apps by revenue in 21 high-GDP countries from 2020 to 2024**

105 files in total (21 countries × 5 years)

Each file contains the top 10,000 apps by revenue for a given country and year

Total: 1,050,000 rows

**Grand Total**:

**115 CSV files**

**1.15 million rows**

|  |  |
| --- | --- |
| **Column** | **Description** |
| App Name | Name of the app |
| App ID | Unified ID of the app |
| Publisher Name | Name of the app’s publisher |
| Publisher ID | Unified ID of the app’s publisher |
| Platform | Android Google Play or iOS App Store |
| Downloads (Absolute) | Downloads |
| Revenue (Absolute, $) | Revenue by US dollar |
| Average DAU (Absolute) | Average daily active user |
| RPD (All Time, WW) | Revenue per download, all time, world wide |
| Game Genre | puzzle/strategy/shooter/simulation/casino…etc |
| Most Popular Country by Downloads | The country which has most downloads of the app |
| Earliest Release Date | Earliest release date of the app |
| Contains Ads | Contains in-game ads or not |

# Methodology

This project applies a variety of data mining techniques introduced in the IE7275 course, including data cleaning and importation, exploratory data analysis (EDA), and visual analytics. For clustering analysis, we utilize both K-means and hierarchical clustering methods, supported by the elbow method to determine the optimal number of clusters. Principal Component Analysis (PCA) is employed for dimensionality reduction and to enhance cluster interpretability. Additionally, we apply linear regression models to identify trends and make predictions for future mobile game downloads and revenues.

# Results

Global Macro Data

**Discussion**

# References

<https://newzoo.com/resources/blog/mobile-games-market-engagement-revenues-covid-19-gaming>

<https://www.ampereanalysis.com/insight/games-market-enters-a-new-era-as-pandemic-years-reset-growth-dynamic>

# Appendix

https://github.com/ponyo2024/mobile\_game\_market\_analysis/