Descriptive Analytics Approach to Determine Popularity of Games Played on Android Platform

Saifuddin Mohtaram[1*], Kalsum U. H.[2], Hakimi M.[3], Mariani Mohd Dahlan[4], Shima Sabri[5], Airuddin Ahmad[6]

Abstract*:* The method utilized in this study involves descriptive analytics, specifically employing data collected from Kaggle, a platform known for hosting benchmark data on various topics, including games. The study initially gathers data from Kaggle, followed by a rigorous data cleaning process to ensure the accuracy and reliability of the results obtained. Subsequently, the cleaned data is categorized and visualized to identify the most installed games and those with the highest rating of FIVE (5) stars. Furthermore, the study delves into previous approaches used to determine popular games, drawing from methodologies akin to those utilized in assessing game popularity. These approaches typically involve analyzing data on game usage, user engagement metrics, and feedback from users. Platforms like Kaggle provide valuable datasets for researchers to analyze and draw insights from regarding game popularity trends. By evaluating factors such as installation rates and user ratings, the study identifies and evaluates the most popular games among others listed in the dataset.

Keywords: Game, Most Installed Games, Highest Rating Games, Action Game, Android Game

1. Introduction

Games represent a multifaceted domain at the confluence of entertainment, engagement, and cognitive challenge, encompassing diverse forms ranging from digital interfaces to traditional tabletop setups. Within this academic discourse, researcher embark on an exploration of the nuanced landscape of games, delving into their manifold types, platforms, and inherent constraints (Mohtaram et. al, 2017)

The Definition of Games: Games are meticulously structured activities characterized by predefined rules, objectives, and anticipated outcomes, wherein players actively participate, shaping the course of events and determining success or failure. This spectrum of games is expansive, spanning digital genres such as action, adventure, and simulation, alongside traditional formats like board games, card games, sports simulations, and puzzles, each demanding distinct cognitive and motor skills from participants (Krath, 2021, Dahlan et. al, 2023).

Evolving Perspectives: While games historically served primarily as sources of leisure and amusement, contemporary perspectives recognize their broader socio-economic implications. Notably, games have emerged as significant sources of income for professional gamers, streamers, and e-sports athletes, while also

1Faculty of Computing & Multimedia, Universiti Poly-Tech Malaysia, Kuala Lumpur, Malaysia. saifuddin@uptm.edu.my
2Department of Computer Science and Mathematics, Kolej Poly-Tech MARA Batu Pahat, Johor, Malaysia
3Department of Computer Science and Mathematics, Kolej Poly-Tech MARA Batu Pahat, Johor, Malaysia
4Faculty of Computing & Multimedia, Universiti Poly-Tech Malaysia, Kuala Lumpur, Malaysia. mariani@uptm.edu.my
5Faculty of Computing & Multimedia, Universiti Poly-Tech Malaysia, Kuala Lumpur, Malaysia.
6Faculty of Computing & Multimedia, Universiti Poly-Tech Malaysia, Kuala Lumpur, Malaysia.

offering lucrative career avenues in game development, design, and marketing.

Platforms and Devices: The proliferation of gaming platforms across various technological mediums underscores the ubiquity and accessibility of games in modern society. From mobile platforms like Android and iOS to dedicated gaming consoles such as PlayStation and Xbox, personal computers, and web browsers, games cater to diverse user preferences and technological capabilities, fostering a democratized gaming culture (Pee, 2017; Cole, 2020). Since the launch of the Android operating system in 2008, it has gained immense popularity, and the platform has had a positive impact on the development area (Gotz, 2017; Wolniak, 2023). One of the significant advantages of the Android operating system is that all platform features are available for mobile app developers, making it the most frequently used platform for developing mobile games (Girdhar, 2024). The use of the Android platform in various devices and phones is a result of its great popularity, leading to a large number of Android gamers (Garcia & Patel, 2017).

Exploring Game Categories: The taxonomy of games encompasses a diverse array of categories, each catering to specific player preferences and skill sets. These categories span action-packed adventures, strategic simulations, immersive role-playing narratives, and casual pick-up-and-play experiences, among others, reflecting the breadth and depth of gaming experiences available to enthusiasts (van den Berg, 2020; Easttom, 2023).

Acknowledging Limitations: Despite their manifold benefits, games are not immune to limitations and challenges. Variability in player preferences introduces subjectivity into gaming experiences, while technical constraints such as hardware capabilities and network infrastructure may impede seamless gameplay. Cultural factors also influence gaming preferences, and concerns regarding excessive gaming and its potential health ramifications warrant careful consideration. One of the key factors that determine the success of a mobile game is its popularity among

players. Highly rated video games are more likely to draw attention from players, who will download the game and recommend it to their friends. However, if a video game is not popular, it may not receive the same level of attention and may struggle to gain traction among players (Chen & Kim, 2018).

In synthesis, games emerge as dynamic cultural artifacts that transcend mere entertainment, offering profound insights into human cognition, social interaction, and technological advancement. As players continue to navigate the evolving landscape of gaming, critical reflection on the diverse facets and implications of games remains imperative within academic discourse (Pee, 2017; Van den Berg, 2020; Easttom, 2023). This study collects benchmark data from Kaggle. The data categorizes into most played games and most rating games. The expected result from this study is to determine the popular games played with the highest rating.

2. Materials and Methods

Figure 1 shown the process for cleaning dataset. The research process commences with essential stages, including data mining, data scraping, exploratory data analysis, and culminates in data visualization. Each of these stages serves distinct functions in addressing research questions. Specifically, the method employed is known as Descriptive Analytics, which entails the collection, cleaning, and summarization of data from diverse sources. Descriptive Analytics allows researchers to analyze data, uncover patterns, identify trends, and gain historical insights. Techniques

such as exploratory data analysis, summary statistics, and data visualization contribute to this comprehensive approach (Yu & Huang, 2022).
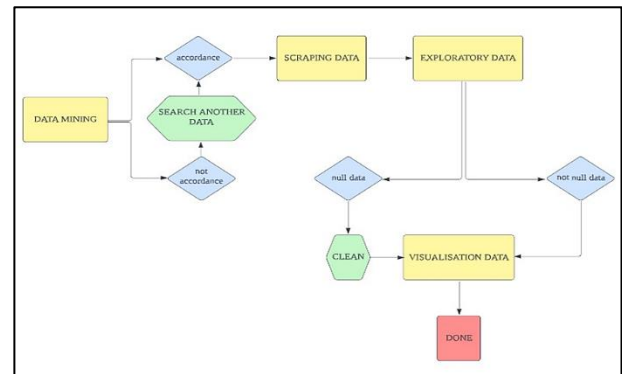

Fig.1 Dataset cleaning.

During the Data Mining phase, it is imperative to utilize benchmark data to ensure the success of the study. Researchers must identify relevant datasets related to their research topic. Subsequently, the process of data scraping involves extracting large datasets from websites and transforming them into structured formats, often using tools like Excel spreadsheets (Courtney, 2021). This step streamlines data processing and facilitates subsequent analyses such as looking for null and non-null data like shown in Figure 2.
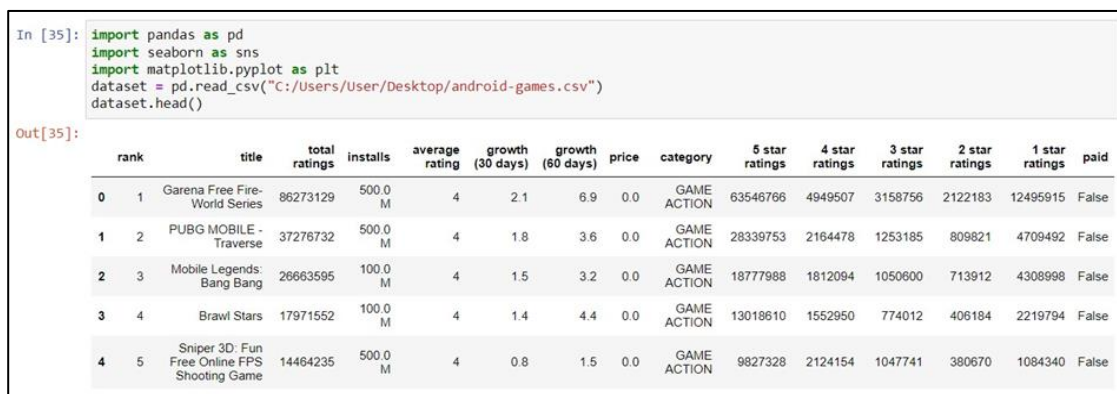

Fig.2 View of exploratory data.

Furthermore, exploratory data analysis delves deeper into the data. Researchers scrutinize variables, identify missing values (null data), and assess data quality. If null values exist, thorough cleaning is necessary to eliminate useless or erroneous data. Conversely, if the dataset is devoid of null values (as depicted in Figure 3), researchers proceed to the subsequent stages.
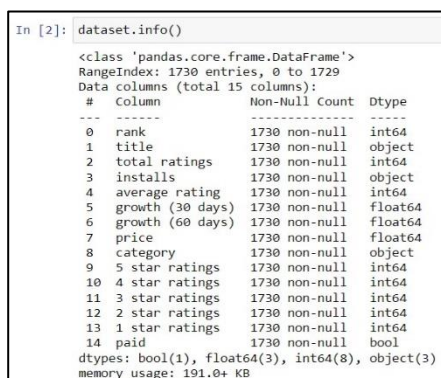

Fig.3 Information of data whether the value is null or not.

Finally, data visualization represents the culminating step in this study. By translating complex big data into visual representations (such as graphs or charts), researchers enhance understanding and accessibility for diverse audiences. For instance, game developers can utilize visualizations to discern demand patterns across different game categories, thereby informing advanced decision-making in their field.

In summary, Kaggle, as a prominent dataset platform, provides valuable resources for researchers across various domains. Its extensive collection of open datasets facilitates rigorous analysis and contributes to advancements in research and practice.

3. Results and Discussion

The evaluation process has yielded affirmative outcomes. These results are predicated on the installation metrics of various games by players. Simultaneously, game ratings are contingent upon player feedback, spanning a spectrum from one-star to five-star ratings. Notably, Figure 4 illustrates that the game genre with the highest player installations is Game Action, boasting 7.4 billion installations, closely followed by Game Arcade with 7.2 billion

installations. Additionally, Game Casual and Game Racing exhibit substantial player adoption, with 6.4 billion and 4.7 billion installations, respectively. Evidently, the most installed games align with player preferences, signifying their popularity. Conversely, Game Trivia lags behind, garnering a mere 0.7 billion installations, indicative of its comparatively lower popularity.

Figure 5 sheds light on the highest-rated games, specifically those receiving a coveted five-star rating. Game Action emerges as the frontrunner, with an impressive 287 million players awarding it a five-star rating. Game Casual follows suit, securing 179 million five-star ratings. In contrast, Game Strategy and Game Arcade receive 133 million and 130 million five-star ratings, respectively. Notably, Game Arcade experiences a relative dearth of five-star ratings, implying diminished popularity, which, in turn, impacts its installation numbers.

findings underscore a significant correlation between high ratings and installation figures. Popular games consistently exhibit both robust installations and favorable ratings. Figures 4 and 5 collectively affirm that Game Action stands as the most installed and highest-rated game. In conclusion, the action genre enjoys widespread player engagement, solidifying its status as a popular choice among players.
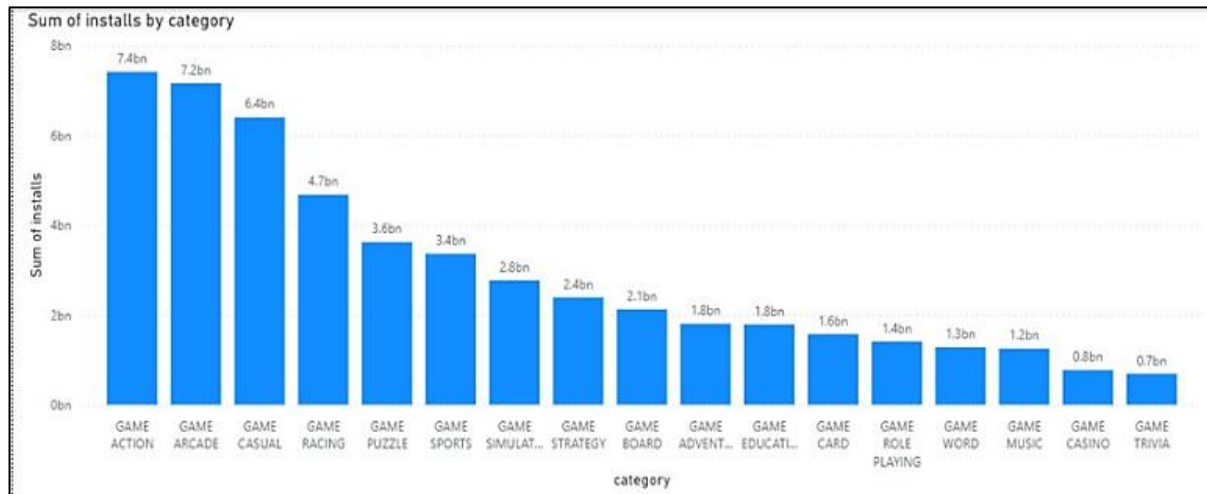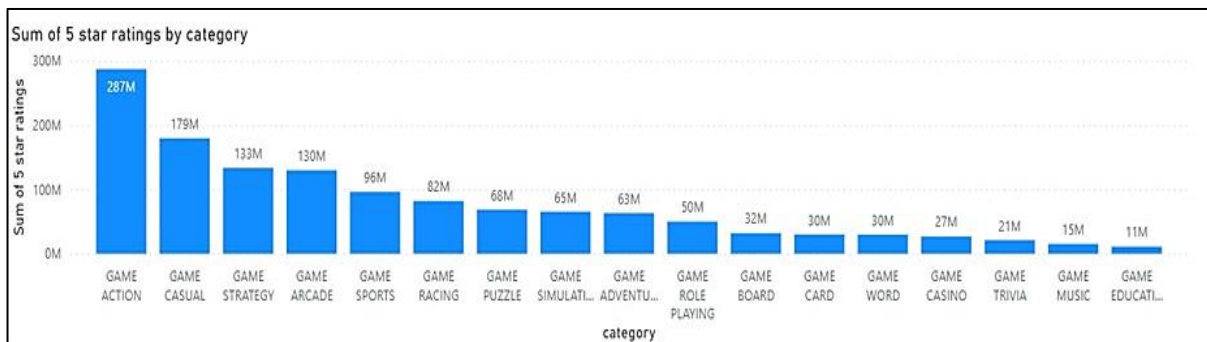

Fig.4 Most installed game by players.

Fig.5 Rating of games with FIVE (5) stars.

For future research, exploring emerging game genres and analyzing player engagement patterns across different demographics could provide valuable insights. Additionally, investigating the impact of in-game monetization strategies on player behavior and retention warrants attention. Furthermore, understanding the dynamics of cross-platform gaming and the influence of social networks on game adoption remains an intriguing avenue for future studies.

4. Conclusion

In the current landscape, an extensive array of games has been developed. For game developers, possessing a comprehensive understanding of specific game categories is pivotal for achieving success in their business endeavors. Consequently, this research study aims to assist developers in identifying and creating appropriate game categories that align with the demands of this new era.

The research commences by assembling a dataset comprising the most widely installed games, all of which have garnered the highest rating of five stars. Subsequently, rigorous data cleansing procedures are applied to ensure the accuracy of results. The dataset is then meticulously classified and visually represented, shedding light on the games that enjoy the highest installation frequency. Additionally, the data is stratified based on the coveted five-star rating.

The research findings reveal that the game category labeled 'Game Action' stands out as the most widely installed game, boasting an astonishing 7.4 billion installations by players. Furthermore, 'Game Action' also claims the highest rating, with 287 million players awarding it a perfect five-star score.

In summary, 'Game Action' emerges as an immensely popular game category among players, underscoring its significance in the contemporary gaming landscape.

Conflicts of Interest

In the world of group gaming, the rating of the game is not only dependent on the game, but the rating also depends on the satisfaction of playing the game as a group. This study not included

judgment from group gaming. The data collect based on various type of players.

Funding Statement

Acknowledgments

References

[1]. Chen, Y., & Kim, J. (2018). *The impact of video game popularity on player attention and traction. Journal of Interactive Entertainment,* 12*(3), 245–260*

[2]. Cole, A., & Barker, D. (2020). Games as Texts: A Practical Application of Textual Analysis to Games. Games as Texts.

[3]. Courtney, M. B. (2021). Exploratory Data Analysis in Schools: A Logic Model to Guide Implementation. International Journal of Education Policy and Leadership, 17(4), 1041.

[4]. Dahlan, M. M., Mohd Zanial, H. A., Amir Hussin, A. A, Yahya, M. F., Saifuddin Mohtaram, S. & Shima Sabri. (2023). Interactive Video Learning: A Comprehensive Review of Engaging Educational Approaches. Tuijin Jishu/Journal of Propulsion Technology, 44 (4), 6836-6847.

[5]. Easttom, C. (2023). In-depth guide to mobile device forensics. Accessed: Nov. 14, 2023. [Online]. Available: [An In-Depth Guide to Mobile Device Forensics (https://www.routledge.com/An-In-Depth-Guide-to-Mobile-Device-Forensics/Easttom/p/book/9780367633004)

[6]. Garcia, A., & Patel, R. (2017). The impact of mobile gaming on the development of the Android platform. *Journal of Mobile Technology, 5*(2), 78–92

[7]. Girdhar, A. (2024). *Mobile app development, Custom Application Development Company*. Appy Pie. https://www.appypie.com/mobile-application-development

[8]. Götz, F. M., Stieger, S., & Reips, U.-D. (2017). Users of the main smartphone operating systems (iOS, Android) differ only little in personality. *PLoS ONE, 12*(5), Article e0176921. https://doi.org/10.1371/journal.pone.0176921

[9]. Kaggle. (n.d.). *Find Open Datasets and Machine Learning Projects*. Retrieved from Kaggle Datasets

[10]. Krath, J., Schürmann, L., & von Korflesch, H. F. O. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games, and game-based learning. Computers in Human Behavior, 125, 106963.

[11]. Mohtaram, S., Shibghatullah, A. S., & Che Pee, N. (2017). DleksiaGame: A Mobile Dyslexia Screening Test Game to Screen Dyslexia Using Malay Language Instruction. Asian Journal of Information Technology. 16(1), 1-6.

[12]. Pee, N. C., et al. (2017). The framework of mobile dyslexia screening test using multiple-deficit theories. Advances in Science Letters, 23(5), 8344.

[13]. van den Berg, Y. H. M., Lansu, T. A. M., & Cillessen, A. H. N. (2020). Preference and popularity as distinct forms of status: A meta-analytic review of 20 years of research. Journal of Adolescence, 84, 7–10.

[14]. Wolniak, R. (2023). Article in Scientific Papers of Silesian University of Technology Organization and Management Series, 172, 42. doi: 10.29119/1641-3466.2023.172.42

[15]. Yu, N., & Huang, Y. T. (2022). Why do people play games on mobile commerce platforms? An empirical study on the influence of gamification on purchase intention. Computers in Human Behavior, 126, 106991.