ESOGÜ Müh. Mim. Fak. Dergisi 2026, 34(2), 2310-2325 J ESOGU Eng. Arch. Fac. 2026, 34(2), 2310-2325

ANALYSIS OF THE MOBILE APPLICATION ECOSYSTEM ACROSS DIFFERENT REGIONS

Cagri SAHIN
Department of Computer Engineering, Gazi University, Ankara, Turkiye
ORCID No: https://orcid.org/0000-0001-5800-6243

Keywords
Mobile applications
User interaction
Regional differences
App ratings

Abstract

Mobile applications have become essential in daily life, influencing how people communicate, work, and entertain themselves, with millions of apps available on platforms. User feedback through ratings and reviews serves as a key measure of app success, shaping download decisions and overall user preferences. App ratings often vary across countries due to cultural, linguistic, and geographic factors, making localized insights crucial for understanding relevance and usability. Moreover, multiple determinants contribute to app popularity, underscoring the importance of conducting comprehensive cross-country analyses to better understand user engagement and satisfaction. However, despite this importance, prior studies are mostly based on limited or single-country data and lack large-scale comparative analyses across countries and app ecosystems.

This study analyzes the popularity, user engagement, and rating consistency of 1521 selected mobile applications across 86 Play Store regions to provide detailed insights into global app usage and interaction patterns. The research examines total ratings, total reviews, and review-to-rating ratios to identify patterns of user engagement and satisfaction, while also assessing rating consistency and preference similarities across countries. It further examines the monetization strategies of popular apps and their influence on app popularity across various categories, while also comparing user ratings between the Play Store and App Store to reveal platform-specific dynamics and feedback variations.

MOBİL UYGULAMA EKOSİSTEMİNİN FARKLI BÖLGELER AÇISINDAN ANALİZİ

Anahtar Kelimeler
Mobil uygulamalar
Kullanıcı etkileşimi
Bölgesel farklılıklar
Uygulama puanları

Öz

Mobil uygulamalar, insanların iletişim kurma, çalışma ve eğlenme biçimlerini etkileyerek günlük yaşamın vazgeçilmez bir parçası haline gelmiş ve farklı platformlarda milyonlarca uygulama kullanıcıların erişimine sunulmuştur. Uygulama başarısının temel göstergelerinden biri olan kullanıcı geri bildirimleri, puanlar ve yorumlar uygulama indirme kararlarını ve genel kullanıcı tercihlerini şekillendirmektedir. Uygulama puanları, kültürel, dilsel ve coğrafi faktörler nedeniyle ülkeler arasında farklılık gösterebilmekte, bu durum uygulamaların ilgi düzeyini ve kullanabilirliğini anlamada yerel içgörüleri önemli hale getirmektedir. Buna ek olarak, uygulama popülerliğinin farklı faktörlerden etkilenmesi, kullanıcı etkileşimi ve memnuniyetini daha iyi anlayabilmek için ülkeler arası karşılaştırmalı analizlerin gerekliliğini ortaya koymaktadır. Buna rağmen, mevcut çalışmaların büyük kısmı sınırlı veri setlerine veya tek bir ülkeye odaklanmakta, ülkeler ve uygulama ekosistemleri arasında geniş ölçekli karşılaştırmalı analiz eksikliği bulunmaktadır.

Bu çalışma, 86 Play Store uygulama mağazası bölgesindeki 1521 seçilmiş mobil uygulamanın popülerliği, kullanıcı etkileşimi ve puan tutarlılığını analiz ederek global uygulama kullanımı ve etkileşimleri hakkında ayrıntılı bilgiler sunmaktadır. Araştırma, kullanıcı etkileşimi ve memnuniyet örüntülerini belirlemek amacıyla uygulamaların toplam puanlarını, toplam yorum sayısını ve yorum-puan oranlarını incelemekte, ayrıca ülkeler arasındaki puan tutarlılığını ve tercih benzerliklerini değerlendirmektedir. Ek olarak, popüler uygulamaların gelir elde etme stratejilerini ve farklı kategorilerin uygulama popülerliği üzerindeki etkilerini incelemekte, Play Store ve App Store arasındaki kullanıcı puanlarını karşılaştırarak platforma özgü dinamikler ve geri bildirim farklılıklarını ortaya koymaktadır.

Araştırma Makalesi		Research Article	
Başvuru Tarihi	: 27.11.2025	Submission Date	: 27.11.2025
Kabul Tarihi	: 02.06.2026	Accepted Date	: 02.06.2026

* Sorumlu yazar: cagrisahin@gazi.edu.tr
https://doi.org/10.31796/ogummf.1831167

1. Introduction

Mobile applications have become an integral part of daily life, shaping how people communicate, work, and entertain themselves across the globe. With millions of apps available on digital platforms such as the Play Store (Google Play Store) and App Store (Apple App Store), understanding user engagement, app performance, and regional preferences has become increasingly important for developers, researchers, and businesses alike. Popular apps not only reflect user needs but also influence trends within the digital ecosystem, with major companies dominating global usage patterns.

A key measure of app success is user feedback, which is captured through ratings and reviews. These metrics serve as a proxy for user satisfaction and interaction, as user ratings significantly influence app selection across categories, with quantified satisfaction shaping overall user preferences (Mahmood, 2020; Sallberg, Wang and Numminen, 2023). For example, apps with a 2-star rating are avoided by 85% of users, while only 50% of users consider downloading apps with a 3-star rating (Alchemer, 2022). Recognizing the importance of these ratings, app stores incorporate localization strategies to enhance the user experience. Country or region specific ratings, submitted by users within the same geographical context, offer insights into an app's relevance, usability, and overall satisfaction within that particular region.

Despite the global reach of these apps, user behavior and app ratings often vary across countries due to cultural, linguistic, and geographical factors. While some apps maintain consistent performance worldwide, localized deviations highlight the importance of considering regional perspectives when evaluating app success. Revenue strategies also play a role in shaping app popularity.

In recent years, the mobile application ecosystem has undergone substantial structural changes driven by the rise of algorithmic recommendation systems and the growing importance of region-specific ranking mechanisms. App stores now rely more heavily on localized engagement indicators—such as country-level ratings, reviews, and usage patterns—resulting in substantial differences in app visibility and perceived quality across regions. At the same time, post-pandemic shifts in digital behavior have diversified user expectations, further amplifying cross-country variation in app evaluation. Consequently, relying solely on global averages is increasingly insufficient, highlighting the need for large-scale, cross-country analyses of user engagement and satisfaction in today's evolving digital landscape.

This study investigates how global mobile application popularity, user engagement, monetization, and rating behavior vary across countries and digital platforms, with a primary focus on the Play Store. Using a dataset of 1521 popular apps across 86 localized Play Store regions, the study first examines the structural characteristics of the app ecosystem, including category distribution and monetization strategies. It then evaluates the global reach of apps and identifies the apps with the highest levels of user interaction based on ratings, reviews, and review-to-rating ratios. Building on this foundation, the study analyzes whether user ratings remain consistent across countries and whether similarities exist between regional rating patterns, thereby revealing the potential influence of cultural, linguistic, and geographical factors on app evaluation. Finally, the study extends the analysis beyond Google Play by comparing ratings with the App Store to identify platform-specific differences in user feedback.

The overall question of how user ratings differ across apps, categories, regions, and platforms has been refined into the following specific research questions:

RQ1 - Category Distribution: What is the distribution of apps that meet the inclusion thresholds across the different Play Store categories?

RQ2 - Revenue Strategy: What revenue strategies underpin the popular mobile applications?

RQ3 - Global Usage: Do the selected apps qualify as global based on the presence of ratings across localized Play Store regions?

RQ4 - User Interaction: What are the top 30 global apps, ranked by total user ratings and reviews, that demonstrate the highest user interaction?

RQ5 - Rating-Review Correlation: Is there a statistical relationship between average user ratings and the review-to-rating ratio across global apps?

RQ6 - Consistency: Do global apps consistently maintain user ratings across all countries?

RQ7 - Similarity: Is there a significant similarity in app ratings between users from different countries?

RQ8 - Store Comparison: Do user app ratings differ significantly between the Play Store and the App Store?

The results indicate that global mobile applications exhibit heterogeneous patterns of popularity, monetization, and user engagement across categories and regions. While most popular apps are free and largely ad-supported, a substantial proportion demonstrate global reach and high interaction levels, although engagement metrics such as ratings and reviews do not always align consistently. User ratings show weak associations with engagement behavior and vary across countries, with most apps maintaining moderate consistency, while a notable share exhibits regional inconsistency likely influenced by cultural and linguistic differences. Finally, a clear platform effect is observed, with App Store ratings generally higher than

Play Store ratings in most regions, and these differences are statistically significant. Overall, the findings highlight that app performance and user evaluation are shaped by the combined influence of global usage patterns, regional behavioral differences, and platform-specific dynamics.

Overall, the findings provide actionable guidance for developers and researchers to optimize app design and marketing strategies while accounting for both global trends and localized user expectations, and are discussed in detail in the results section, where each research question is addressed systematically.

The remainder of this article is organized as follows: The next section presents related work. The methodology section details the study design and the process of dataset selection and generation. This is followed by the presentation of results and a discussion of the findings. Finally, the conclusion section summarizes the key insights and proposes directions for future research.

2. Related Work

Cultural and regional differences in user behavior are crucial because they shape how individuals perceive, adopt, and interact with digital technologies across contexts. Major technology acceptance frameworks reflect this by modeling culture as a key moderating factor. For example, the Technology Acceptance Model (TAM), Theory of Planned Behavior (TPB), and Innovation Diffusion Theory (IDT) are often extended with cultural variables to explain differences in attitudes, perceived usefulness, ease of use, and behavioral intentions across regions (Pratesi, Hu, Rialti, Zollo and Faraoni, 2021; Zhao, Wang, Li, Zhou and Li, 2020). Similarly, the Unified Theory of Acceptance and Use of Technology (UTAUT) incorporates cultural values, trust, and user experience to account for cross-country variations in technology adoption (Ahmad, Kim, Choi and Haq, 2021). Hofstede's cultural dimensions theory further demonstrates how societal values influence trust, risk perception, social influence, and engagement behavior, ultimately shaping satisfaction, continued usage, and overall interaction with digital platforms (Gupta, Pansari and Kumar, 2018; Thompson and Brouthers, 2021). Building on these insights, layered frameworks such as the Onion Model of Human Factors link cultural values to observable behaviors and interface preferences, providing a structured approach for designing culturally sensitive technologies and enhancing cross-regional adoption (Guo, Rau and Heimgärtner, 2022).

These theoretical insights provide a foundation for understanding user behaviors in digital ecosystems such as application stores, where cultural and regional differences directly influence interactions, reviews, and ratings. This is particularly relevant as application stores enable users to navigate and discover apps that align with their needs and interests. These stores also provide users with a review and rating mechanism, enabling them to provide feedback and share their opinions about the apps they have experienced. These reviews and ratings have an important impact on shaping the user's decision-making process (Liptrot, Pearson, Montazami and Dubé, 2024; Mahmood, 2020; Sallberg et al., 2023). While app reviews reflect user feedback based on direct experience, app ratings provide quantitative information about user satisfaction. The exploration of user app ratings and reviews provides valuable insights into the intricate dynamics of user experiences, as articulated through their feedback, contributing significantly to our comprehension of the digital ecosystem. Therefore, researchers have conducted studies on app information obtained from application stores (Chen, Lin, Hoi, Xiao and Zhang, 2014; Iacob, Veerappa and Harrison, 2013; Panichella et al., 2015; Phong, Nguyen, Pham and Nguyen, 2015).

A closely related area of research involves studies that have attempted to identify the cultural factors influencing the characteristics of app reviews, compare app-store attributes, explore the consistency of app rating scores in app stores, and investigate the perception of the same apps across different countries. Guzman, Oliveira, Steiner, Wagner and Glinz (2018) analyzed 2560 app reviews authored by users from eight countries, each characterized by diverse national cultures. Their study contributes evidence regarding the influence of cultural factors on the characteristics of app reviews. Srisopha, Phonsom, Lin and Boehm (2019) retrieved 300,643 user reviews for the 15 most downloaded apps in the App Store from nine English-speaking countries and the US. Their results indicate that all countries exhibit some factors that are proportionally inconsistent with the US.

In addition to these studies, surveys have been conducted to investigate country differences in mobile app user behavior. For example, Lim, Bentley, Kanakam, Ishikawa and Honiden (2015) collected survey data from 15 countries to analyze user behavior differences across countries. Their findings reveal that users from the USA are more inclined to download medical apps, while those from the United Kingdom and Canada are more likely to be influenced by pricing factors. Moreover, users from Japan and Australia exhibit a lower likelihood of rating apps. Another study revealed that user feedback, ratings, and screenshots from other apps are the three most crucial pieces of information developers gather from the open market (Al-Subaihin, Sarro, Black, Capra and Harman, 2021). They also concluded that app stores facilitate communication between users and developers, providing platforms for bug reporting.

Hu, Wang, Bezemer and Hassan (2019) examined 68 hybrid app-pairs, representing apps available in both the Play Store and App Store. The findings indicated that 33 out of the 68 hybrid apps did not demonstrate consistent star ratings across platforms. Ali, Joorabchi and Mesbah (2017) collected textual user reviews from both the Play Store and App Store to compare app-store attributes and investigate user complaints related to cross-platform development challenges that influence user-perceived ratings. Their findings suggest that apps are predominantly developed for a specific platform, with Android users being more inclined to rate apps compared to iOS users. Additionally, the majority of cross-platform apps lack consistent releases, yet most top-rated apps are accessible on both platforms. While Android users commonly express complaints about app features and non-functional properties, iOS users tend to provide post-update complaints.

Some studies provide a systematic literature review on application stores and review analysis, highlighting its role in supporting software engineering by offering insights into user needs and issues (Dabrowski, Letier, Perini and Susi, 2022; Martin, Sarro, Harman, Jia and Zhang, 2017).

There are also studies that rely on user app ratings and reviews. For example, a recent study collected 100,010 user reviews from nine prominent generative AI apps on the U.S. App Store, analyzing them through integrated VADER sentiment scores and LDA topic modeling (Meng, Park and Um, 2026). The results demonstrated that both the sentiment of reviews and titles have a significant positive influence on user ratings. This highlights the key role that user sentiment, as reflected in review content and titles, plays in shaping overall app evaluations.

Ngereza et al. (2026) examined the effectiveness of 511 systems in enhancing road safety and mobility, with a focus on user feedback through app ratings and reviews. Ethical concerns in user app store reviews were analyzed by examining five million reviews, developing a set of ethical issues based on user preferences, and manually labeling a sample to better understand and classify these concerns (Tjikhoeri, Olson and Guzmán, 2024). The results show that user reviews addressing ethical concerns are generally longer, more widely viewed, and often receive lower ratings.

Yu et al. (2023) propose a novel approach to localizing function errors in mobile apps using user reviews and semantic meanings, demonstrating better performance compared to existing tools. An automated issue analysis approach and framework have also been developed to efficiently detect and classify emerging issues in mobile apps, helping developers address user-reported concerns more effectively (Gao, Zeng, Lyu and King, 2018; De Lima and Marcacini, 2024).

Although existing studies highlight the importance of app store data analysis, they are often limited by small datasets, a narrow focus on a limited number of countries, or an isolated focus on specific aspects such as reviews or platform differences. In particular, there is a lack of unified, large-scale studies examining how user ratings and interaction patterns vary across countries and app ecosystems, as well as limited in-depth analysis of country-level ratings, category distribution, app globality, revenue strategies, and cross-store comparisons.

3. Methodology

This section outlines the study design, including the selection of application stores, identification of the specific apps considered, determination of the target countries/regions for the local application stores, and the approach used for data collection. All procedures were conducted in accordance with research and publication ethics, and the study relied on publicly accessible data without involving human subjects or personal information. This study accessed only publicly available, non-proprietary metadata from application stores. Although the Terms of Service of application stores generally restrict automated access to the platform, the data collection was limited to publicly available information.

3.1. Selection of Application Stores

In the contemporary mobile ecosystem, the distribution and accessibility of mobile applications are predominantly determined by a small number of leading digital platforms. Among these, the Play Store and App Store have established a global duopoly in terms of both app availability and user base. As of June of 2025, the Play Store, serving Android users, hosted an extensive collection of approximately 4.5 million apps, positioning it as the largest app repository by volume (Statista, 2025b). In comparison, the App Store, catering to iOS users, offered approximately 2.4 million apps during the same period, thereby securing its status as the second-largest platform. The Amazon Appstore, with a significantly smaller catalog of approximately 800,000 apps, occupies a niche position in the market.

As of the 2025, Android-based devices commanded a dominant share of the global mobile operating system market, accounting for 71.88%, while iOS held a significant 27.65% share (Statista, 2025a). Together, these two platforms represent over 99% of the global smartphone market, with Android maintaining a clear lead. This data underscores the widespread adoption of Android and iOS, demonstrating that the vast majority of smartphones in circulation operate on one of these two platforms. Given the overwhelming market share of these ecosystems, it is crucial to focus on them when analyzing user ratings for apps, as they represent the primary environments in which mobile apps are downloaded, used, and rated. In this study, the Play Store was selected to identify specific apps under consideration, determine the local regions of the application stores, and analyze user ratings, given its dominant position in the global market and its wide array of apps. To compare user app ratings across different application stores, the App Store was also included in the analysis, providing a broader understanding of user preferences in both ecosystems. The Amazon Appstore was excluded from the main analysis due to its smaller market share.

3.2. Identification of Applications

To identify a representative and high-quality sample of apps for this study, a comprehensive analysis of the Play Store was conducted, examining apps across all 32 distinct categories available on the Play Store. It is important to note that Games is not considered a category in the Play Store, as it is placed in a separate section. The selection process was guided by two specific and quantifiable criteria designed to ensure the inclusion of data from a large and active user base. To accurately represent a broad and engaged user base, only apps with at least 5 million downloads were included. This approach is common practice in studies analyzing apps on the Play Store, which focus on popular applications and broad user engagement, although no universally accepted threshold exists (Chang, Zaeem and Barber, 2020; Gao and Dube, 2024; Liu et al., 2024; Rahaman, Farooq, Neamtiu and Zhao, 2023; Ravichander, Black, Wilson, Norton and Sadeh, 2019; Teles et al., 2019). This threshold serves as a reliable proxy for global market penetration, effectively reducing the influence of niche or unrepresentative apps. In addition, the criterion of 500,000 user ratings was set to minimize the risk of manipulation or sampling bias, ensuring that the rating data represented a wide and diverse user base. This requirement aimed to improve the reliability and validity of user feedback by ensuring a substantial volume of ratings. The distinction between these two metrics is critical: while downloads indicate initial user interest, ratings indicate sustained interaction. Consequently, these thresholds also suggest that the selected apps are more likely to be widely used and evaluated by users, thereby strengthening the representativeness of the dataset.

To identify apps with at least 5 million downloads, the public website providing reliable data for Android apps was relied upon (AndroidRank, 2025). The app data was then verified using a generated script to ensure the actual download count met the criteria from the Play Store. Afterward, apps with at least 5 million downloads were further filtered based on the next criterion of having at least 500,000 ratings. In total, 1521 apps that met these specific criteria were identified across 32 distinct categories on the Play Store.

Table 1. Determined Country Regions

Country/Region	Country/Region	Country/Region	Country/Region	Country/Region
Algeria	Argentina	Australia	Austria	Azerbaijan
Belarus	Belgium	Bolivia	Brazil	Bulgaria
Cambodia	Canada	Chile	Colombia	Costa Rica
Czech Republic	Dominican Republic	Ecuador	Egypt	El Salvador
France	Georgia	Germany	Ghana	Greece
Guatemala	Honduras	Hong Kong	Hungary	India
Indonesia	Iraq	Ireland	Israel	Italy
Japan	Jordan	Kazakhstan	Kenya	Kyrgyzstan
Laos	Lebanon	Libya	Malaysia	Mexico
Morocco	Myanmar	Nepal	Netherlands	New Zealand
Nicaragua	Nigeria	Norway	Oman	Pakistan
Panama	Paraguay	Peru	Philippines	Poland
Portugal	Republic of Korea	Romania	Russia	Saudi Arabia
Senegal	Serbia	Singapore	South Africa	Spain
Sri Lanka	Sweden	Switzerland	Taiwan	Thailand
Tunisia	Turkiye	Ukraine	United Arab Emirates	United Kingdom
United States	Uruguay	Uzbekistan	Venezuela	Vietnam
Yemen				

3.3. Determination of Application Store Regions

Application stores are localized into different regions or countries to offer region-specific content, including region-based reviews and app ratings. This means users can access the version of the store associated with their location, allowing them to leave reviews and ratings for apps based on their regional experience. As a result, an app can have different rating scores depending on the country or region. For example, on the Play Store in August 2025, the Instagram app has a rating score of 4.198 in Germany and a rating score of 3.556 in the United Kingdom.

To determine the regions for application stores, all countries worldwide were evaluated. For each country, it was confirmed whether a localized store exists. As a result, 86 distinct application store regions were identified on the Play Store. Table 1 shows the list of these countries, representing the 86 regions considered for application store localization. These countries span across multiple continents, highlighting where localized application stores, such as those on the Play Store, are available. Note that some countries, such as China, are not included in the table because the Play Store is not available in these regions due to governmental regulations and restrictions. In the case of China, the country operates its own app store ecosystem, with local platforms dominating the market. These regulatory barriers prevent the availability of Google services, including the Play Store, in certain countries. As a result, the list of countries provided only includes those where the Play Store is officially accessible, excluding regions like China where alternative app distribution methods are used.

3.4. Data Collection

To collect the necessary data for analysis, information was gathered for the 1521 identified apps in the Play Store, including the number of user ratings, number of reviews, and app pricing. Additionally, various data points, such as country-specific rating scores, were collected for each app in the selected regions. The *google-play-scraper* tool was used to systematically retrieve this information from the Play Store.

Instead of relying solely on overall app rating scores and total review counts, this data includes localized rating scores and reviews, reflecting user feedback in different countries. The data represents a static snapshot from the second half of 2025. It is important to note that app ratings and review counts are dynamic and can change over time.

4. Data Analysis and Discussion

This section presents the study's results in relation to the research questions, providing an analysis of how the findings address each key area of interest.

Table 2. Distribution of Apps Across Categories

Category	Apps	With Ads	Global
Art and Design	9	2	8
Auto and Vehicles	12	2	3
Beauty	2	1	1
Books and Reference	33	17	24
Business	38	12	19
Comics	2	2	2
Communications	69	42	57
Dating	13	12	12
Education	40	18	30
Entertainment	98	79	69
Events	0	0	0
Finance	215	29	28
Food and Drink	38	7	7
Health and Fitness	46	24	39
House and Home	4	3	2
Libraries and Demo	0	0	0
Lifestyle	46	24	26
Maps and Navigation	30	12	17
Medical	8	2	1
Music and Audio	92	82	76
News and Magazines	17	17	10
Parenting	3	2	3
Personalization	38	33	37
Photography	74	57	74
Productivity	85	39	75
Shopping	120	37	24
Social	53	39	44
Sports	15	15	11
Tools	195	123	166
Travel and Local	42	18	26
Video Players and Editors	63	51	61
Weather	21	18	19

4.1. Category Distribution

Application stores provide categories to users for organizing, discovering, and comparing apps based on their functionality, purpose, or content type. The Play Store has 32 different categories, and investigating apps within each category can reveal patterns of user interests and growth potential across different sectors. Table 2 presents the distribution of the 1521 apps that met the inclusion thresholds of at least 5 million downloads and 500,000 ratings, hereafter referred to as popular apps. The ''Category'' columns list the names of the Play Store categories, while the ''Apps'' columns indicate the number of popular apps in each category that satisfy the inclusion criteria.

According to Table 2, the Finance category leads with 215 apps that meet inclusion thresholds, significantly ahead of Tools (195) and Shopping (120). This concentration reflects strong demand for digital financial services, widely used in everyday activities such as banking, investments, and payments, and indicates that Finance apps are major drivers of user engagement on the Play Store. Tools follows with 195 apps, highlighting demand for productivity and utility apps that support daily tasks, while Shopping ranks next with 120 apps, emphasizing the widespread use of e-commerce platforms and their key role in mobile user activity. However, some categories have very few or no apps. For instance, Libraries and Demo and Events contain no qualifying apps, while Comics, Medical, House and Home, and Parenting include only a small number of apps. This suggests that these categories have smaller user bases or attract limited user engagement, making it challenging for apps in these sectors to achieve widespread adoption and receive high-volume user feedback. Overall, the distribution of apps across categories is uneven, with a small number of dominant categories accounting for the majority of user adoption and engagement on the Play Store.

4.2. Revenue Strategy

An important aspect of user preference is the revenue strategy of apps, as it can directly influence both adoption and satisfaction. To balance user access and developer profitability, mobile applications typically adopt one of three models: free, free with ads, or paid. Free apps allow users to access the app at no cost to attract a large user base, while free-with-ads apps also provide core functionality for free but generate revenue through in-app advertisements, balancing accessibility with developer income. Paid apps, on the other hand, require an upfront purchase, usually promising premium features or an ad-free experience to justify the cost.

To determine the revenue strategies underlying popular apps, the price information of the app was analyzed. The main finding is that all 1521 popular apps considered are free; none are paid. This suggests that to achieve popularity, apps are more likely to be free, as users tend to prefer free apps over paid ones. The popular apps were then further examined to identify whether they are completely free or free with advertisements. In Table 2, the "With Ads" column shows the number of apps in each category that are free with advertisements. The results show that 53.85% of the apps are free with advertisements, while 46.15% are completely free without ads. In addition to overall trends, revenue strategies vary across app categories. Some categories have a higher proportion of apps with advertisements, and in certain cases, all apps are ad-supported, while other categories consist mostly of apps that are completely free without ads. For example, in the Comics, News and Magazines, and Sports categories, all apps are supported by ads. Additionally, categories such as Dating, Entertainment, Music and Audio, Personalization, Video Players and Editors, and Weather have more than 80% of apps with advertisements, indicating a strong reliance on ad-based revenue in these domains. In contrast, most apps in the Finance category are completely free without advertisements, with only 29 out of 215 apps (approximately 13.5%) containing ads. This is likely because financial apps require a clean, professional interface and uninterrupted user experience, as users expect privacy, security, and reliability, making ad-based revenue less suitable for this category.

After examining whether apps are free, free with ads, or paid, it is also important to consider the role of in-app purchases (IAP) within free apps. Free apps may still generate revenue by offering optional paid features, subscriptions, or virtual goods, reflecting a freemium model. Among the free apps analyzed, 153 apps (approximately 22%) offer in-app purchases (IAP), demonstrating that popular free apps can generate revenue through user spending within the app.

4.3. Global Usage

The distribution of apps across categories provides useful insights; however, not all apps are necessarily global, as population size and the number of active smart phone users vary across country regions. Therefore, it is important to assess how many of these apps are truly global, they are available in the localized versions of the Play Store and receive user ratings. To analyze this, the user ratings of all apps were examined in the respective local stores. An app is considered global if it is available in the store and has ratings from at least 81 of the 86 localized Play Store regions. The threshold of 81 countries (approximately 95% coverage) was chosen to account for minor regional restrictions or temporary unavailability, while still ensuring that the app is widely accessible and adopted across the majority of countries. This threshold is not intended to imply absolute universality, nor does it represent an arbitrary selection; rather, it approximates near-global availability via a principled operational rule. Specifically, 81 corresponds to the integer floor of a 95% coverage level, a commonly used benchmark for high completeness in empirical research (Moore, McCabe and Craig, 2017). A lower threshold (<=90%) would risk including apps with substantial regional gaps, whereas a stricter threshold (>=97%) would be overly sensitive to minor or temporary absences in a small number of regions.

Table 2 shows the number of global apps from popular apps in the Global column. Based on the table, 971 apps, which is around 64% of popular apps (out of a total of 1521 apps), are determined to be global. This means that about two-thirds of the popular apps are available and used by people globally. However, the result also highlights the need for a more nuanced understanding of success in the mobile app ecosystem. It demonstrates that being ‘‘popular’’ in some markets does not necessarily equate to an app being truly ‘‘global’’. One reason for this is that regions with a high number of users can drive app popularity, but this does not necessarily ensure widespread international availability or usage. This is evidenced by the data in the Finance, Food and Drink, Shopping, and Auto and Vehicles categories. For instance, in the Finance category, there are 215 apps, yet only 28 of them are classified as global.

This disparity reflects structural constraints inherent to financial services, including fragmented regulatory frameworks, as well as technical dependencies on local banking APIs and payment infrastructures. Cultural factors and strong domestic competition further limit scalability, as users often prefer localized solutions tailored to regional financial practices. Consequently, finance apps tend to remain region-specific rather than globally scalable. In the Food and Drink category, only 7 out of 38 apps are global, indicating that these apps are often designed to meet local consumer preferences and needs. Similarly, in the Shopping category, only 24 out of 120 apps are global, implying that these apps are typically tailored to specific regional markets, with limited international accessibility. The Medical category also demonstrates this trend, with 8 apps, only 1 of which are globally available. Apps that meet popularity criteria in categories like Art and Design, Communication, Dating, Health and Fitness, Music and Audio, Parenting, Personalization, Photography, Productivity, Social, Tools, Video Players and Editors, and Weather are more likely to be global due to their universal appeal and broad user demand, making them accessible to a diverse range of users worldwide.

4.4. User Interaction

User interaction with mobile applications can be effectively assessed through two key indicators: the total number of ratings and the total number of reviews. The total number of ratings reflects the number of users providing feedback on a scale of 1 to 5, indicating the app’s interaction with its overall user base. Similarly, the total number of reviews provides deeper insights into user experiences, since writing a review requires more effort and often includes detailed opinions. While these metrics highlight user interaction with the app, the average rating offers a complementary measure of overall user satisfaction. Another indicator for evaluating the top apps could be the total number of downloads, which reflects how many times an app has been installed from the store. However, for pre-installed apps, this metric can be misleading when identifying the top apps, since many installations occur automatically during device setup and may not accurately represent active user adoption.

Table 3 presents a combined list of the top 30 global apps based on total ratings and the top 30 global apps based on total reviews. When these two lists are merged, 35 unique apps appear due to overlaps in the rankings. The table is ordered by total ratings, with each app’s name (the number in parentheses indicates its rank based on total reviews), followed by the average rating, total number of ratings, total number of reviews, and the review-to-rating ratio. The table highlights a clear dominance of the top four apps — WhatsApp Messenger, Facebook, YouTube, and Instagram — over the rest serves as a strong indicator of deep and widespread user interaction. Another key point is that these four dominant apps are developed by different companies. Furthermore, Google LLC developed 9 of the apps listed in the table, while Meta Platforms Inc developed 3. Notably, no other companies have more than one app featured on the list, which emphasizes the substantial influence of Google and Meta in shaping mobile app trends and the direction of the global digital ecosystem. The average user ratings for top global apps range from 3.75 to 4.808, with most apps scoring above 4 and an overall average of 4.398, reflecting high user satisfaction. This suggests that users generally find these apps valuable and that they meet their needs.

4.5. Rating-Review Correlation

The review-to-rating ratio in relation to average user ratings can provide valuable insights into how users interact with apps, particularly in terms of whether they are more likely to leave feedback when they have complaints or issues, or when they also have positive experiences. For example, a high review-to-rating ratio with low average user ratings could indicate that users are more likely to leave feedback when they are dissatisfied or facing issues with the app. Conversely, a low review-to-rating ratio with high average ratings may suggest that users are generally content and do not feel the need to provide feedback. Therefore, the review-to-rating ratio and average ratings are analyzed to explore patterns in user behavior. In Table 3, an example of the review-to-rating ratio is provided for the top global apps.

To explore whether a relationship exists between the review-to-rating ratio and average user ratings, the Pearson correlation coefficient was calculated for 971 global apps. Pearson’s correlation coefficient assesses both the strength and direction of the linear relationship between two variables. A value of (+1) indicates a perfect positive correlation, (-1) indicates a perfect negative correlation, and (0) suggests no linear relationship between the variables. The calculated coefficient value of -0.1226 suggests a weak negative correlation, indicating a slight tendency for higher review-to-rating ratios to be associated with lower average ratings. However, the correlation is very weak, meaning that the relationship is not strong or consistent. The p-value of 0.00013, which is much lower than the standard threshold of 0.05, indicates that the observed correlation is statistically significant and unlikely to have occurred by chance. To further evaluate the practical significance of this relationship, an effect size measure was computed using Cohen's d, derived from the Pearson correlation coefficient. Using the observed correlation value ($r = -0.1226$), the resulting Cohen's d is approximately −0.25, which corresponds to a small effect size (values around 0.2 indicate a small effect, 0.5 a medium effect, and 0.8 a large effect). This indicates that the practical impact of the relationship is limited, suggesting that other factors likely influence app ratings.

Table 3. Top 30 Global Apps by Total Ratings and Reviews

App Name	Average Ratings	Total Ratings	Total Reviews	Review-to-Rating Ratio
WhatsApp Messenger (1)	4.284	210,849,113	58,162,237	27.585
Facebook (4)	4.53	172,171,474	42,420,972	24.639
YouTube (2)	4.044	165,209,259	53,056,066	32.114
Instagram (3)	4.33	164,330,576	52,099,639	31.704
Messenger (6)	4.591	106,165,602	25,684,784	24.193
TikTok (5) - zhiliaoapp.musically	4.213	67,108,209	25,936,930	38.649
Flipkart Online Shopping App (13)	4.327	66,260,174	9,947,009	15.012
Google Photos (7)	4.456	52,348,686	15,415,489	29.448
Google Chrome (9)	4.255	47,643,912	13,744,678	28.849
Google Play services (8)	4.368	44,850,068	13,746,664	30.650
Phone by Google	4.519	41,173,168	1,746,766	4.242
Snapchat (10)	4.29	38,274,075	12,274,446	32.070
Duolingo: Language Lessons (12)	4.656	35,917,936	11,034,025	30.720
Spotify: Music and Podcasts (11)	4.411	34,046,853	11,375,925	33.413
Google Messages	4.56	33,117,184	4,385,061	13.241
Facebook Lite (14)	4.465	32,927,180	9,083,558	27.587
Lazada 9.9 | Mega Brands (24)	4.718	28,819,626	5,513,185	19.130
Google (15)	4.325	28,368,049	8,779,474	30.948
MyJio: For Everything Jio (28)	4.469	26,829,928	5,203,000	19.393
Truecaller: Phone Call Blocker (23)	4.483	26,338,611	5,542,355	21.043
ChatGPT	4.646	25,973,607	2,795,645	10.763
Video Editor & Maker - InShot (17)	4.79	23,052,505	7,227,891	31.354
X (Formerly Twitter) (20)	3.967	22,683,294	6,293,310	27.744
TikTok (16) - ss.android.ugc.trill	4.111	22,407,666	8,375,503	37.378
UC Browser-Safe, Fast, Private (19)	3.818	22,250,429	6,842,741	30.753
Paytm: Secure UPI Payments	4.638	22,183,859	3,961,151	17.856
Canva: AI Video & Photo Editor	4.724	21,949,379	2,890,662	13.170
Mercado Libre: Compras Online (18)	4.808	20,713,360	7,076,522	34.164
Google Maps (26)	4.06	19,137,805	5,370,718	28.063
SHAREit: Transfer, Share Files (21)	4.394	18,266,290	5,770,269	31.590
Gboard - the Google Keyboard (25)	4.532	17,275,461	5,498,790	31.830
AliExpress - Shopping App (27)	4.584	16,193,120	5,347,297	33.022
Uber - Request a ride (30)	4.547	15,558,775	4,843,402	31.130
CapCut - Video Editor (29)	3.75	11,929,309	4,907,492	41.138
Likee - Short Video Community (22)	4.269	11,403,931	5,593,442	49.048

4.6. Consistency

Global apps often face challenges in maintaining consistency in user ratings across different countries due to various factors, such as cultural differences, regional preferences, and varying levels of user experience. While the core features of an app may remain the same, the user experience can vary significantly, leading to wide disparities in ratings across different countries. To analyze global app consistency, the

Coefficient of Variation Across Countries (CV) was calculated for each app.

$$CV = \left(\frac{Standard\ Deviation}{Mean\ Rating}\right) x\ 100 \quad (1)$$

In Equation (1), the Standard Deviation measures the spread or variability of ratings, while the Mean Rating represents the average rating across all countries. Dividing the standard deviation by the mean gives a measure of relative variability, and multiplying by 100 converts it into a percentage for easier interpretation. A high CV indicates significant rating variation across countries, suggesting inconsistent user experiences, while a low CV implies consistent performance across regions.

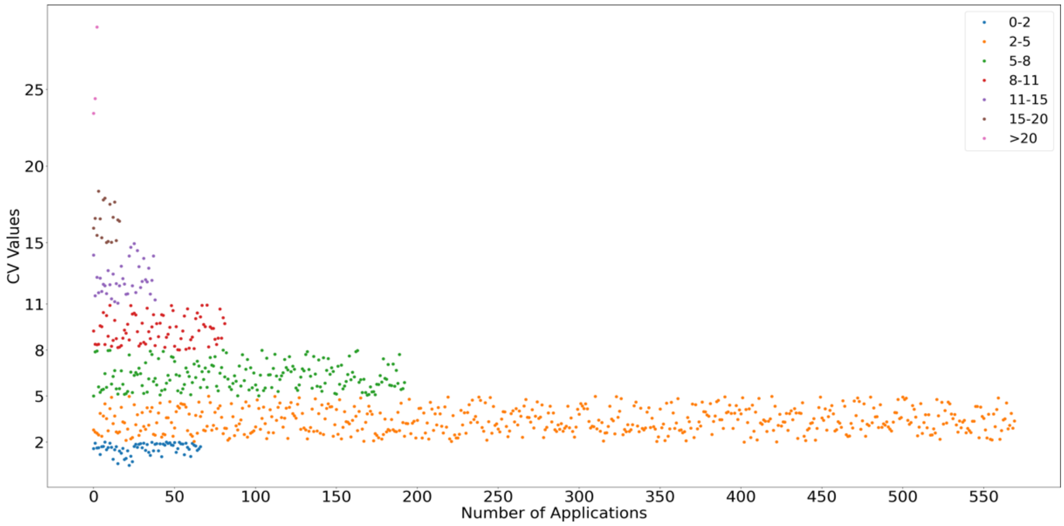

Figure 1. Consistency of Global Apps

Figure 1 represents the CV values of global apps, with the x-axis showing the number of apps and y-axis representing the CV values, which range from 0.482 to 29.072, with an overall mean of 5.002. The distribution is highly asymmetric, with a concentration of observations at lower values and relatively few observations at higher values, indicating that most apps exhibit low-to-moderate variability in user ratings across countries, while a smaller subset shows substantially higher inconsistency. Inspection of the empirical distribution shows a strong concentration of observations at lower values alongside a small number of higher values, indicating clear non-normality and supporting a distribution-based interpretation. Quartile analysis yields Q1 = 2.884, Q2 = 3.978, and Q3 = 6.088, confirming that most observations lie below approximately 6 and that variability increases substantially in the higher range of the distribution. Given the absence of established theoretical thresholds for interpreting CV in cross-country app rating consistency, a descriptive classification scheme is defined based on the empirical structure of the data. Category boundaries are aligned with key features of the distribution to ensure interpretability without imposing arbitrary cutoffs. Accordingly, values below 2 are classified as "consistent," representing exceptionally low variability below Q1. Values from 2 to 5 are classified as "likely consistent," capturing the central portion of the distribution around Q2. Values from 5 to 8 are classified as "unlikely consistent," covering the transition from typical to higher variability and extending beyond Q3 (6.088). Values from 8 to 11 are classified as "inconsistent," reflecting elevated variability above the main concentration of observations, while values above 11 are classified as "highly inconsistent," capturing extreme observations with substantially elevated cross-country divergence. This discretization follows established approaches for interpreting continuous variables when no externally validated thresholds exist (Altman, 1990; Hyndman and Fan, 1996). It is intended as a descriptive framework grounded in the empirical distribution rather than a universal classification standard. Based on the results, 6.9% of the apps fall into the "consistent" category, 58.7% are "likely consistent", 19.9% are "unlikely consistent", 8.4% are "inconsistent", and 6.1% are "highly inconsistent". This indicates that most global apps maintain a reasonably consistent user experience between regions. However, approximately 35% of the apps demonstrate variability in user experience, indicating that regional factors influence user ratings.

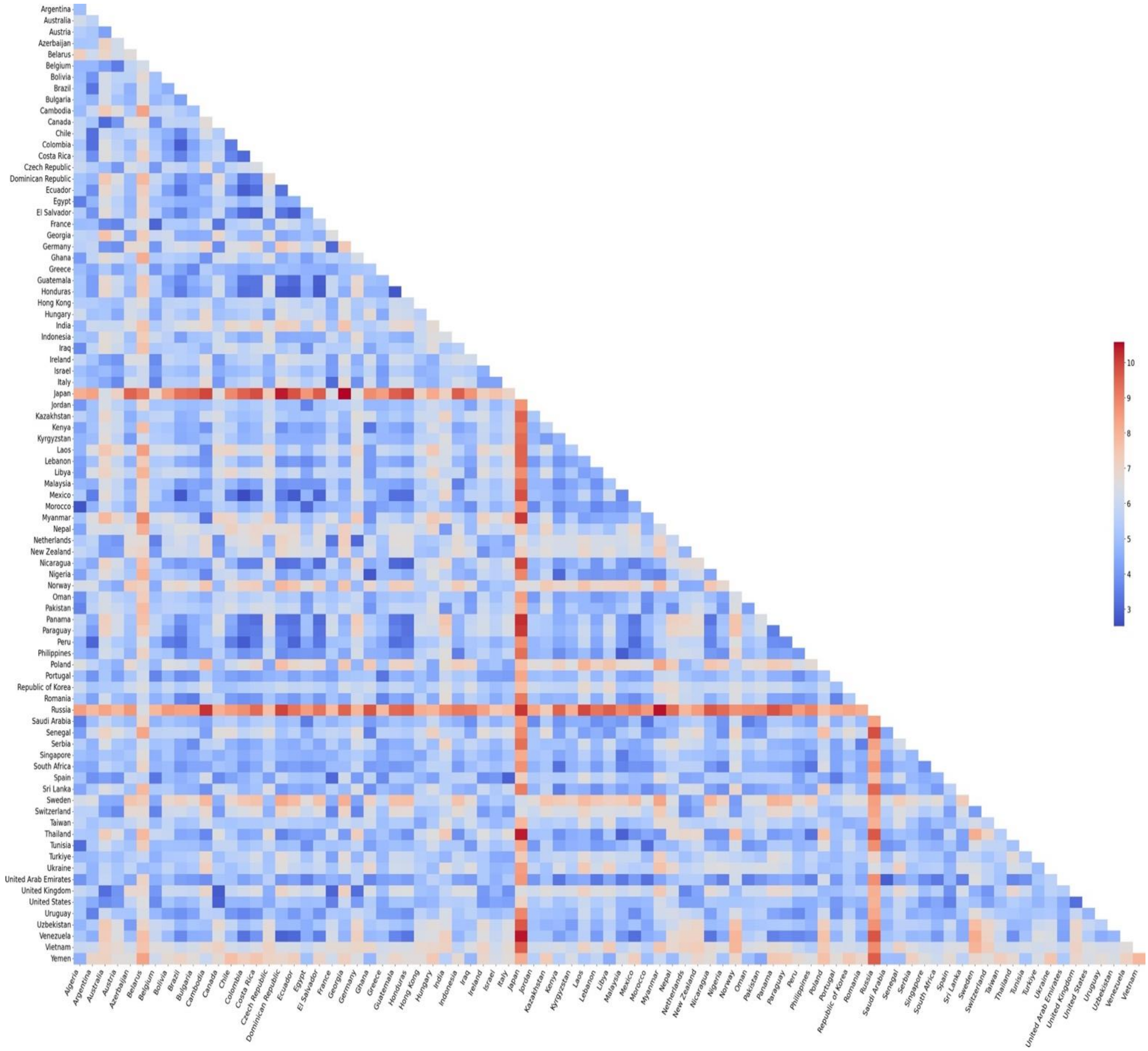

Figure 2. Country Similarities Based on App Rating

4.7. Similarity

Users from different regions can have varying perspectives and expectations for an app. Analyzing app rating similarities across different countries can provide insights into the commonalities or differences in users across these regions. Observing a high similarity in app ratings between countries may indicate preferences that users in these regions tend to share similar opinions. Conversely, a low similarity might suggest differences in user preferences or experiences. Using the knowledge gained from this analysis can help identify commonalities, enabling strategic decision-making or targeted improvements in user experience across diverse countries.

To investigate the significant similarity in app ratings among users from different countries, the mean absolute percentage difference was calculated for each pair of countries to compare their similarity in app ratings on the Play Store using the following formula:

$$MAPDpair = \frac{1}{n}\sum_{i=1}^{n}\frac{|CR1_i - CR2_i|}{(CR1_i + CR2_i)/2}\ x\ 100 \quad (2)$$

In the given equation (2), ‘‘n’’ represents the total number of apps, and ‘‘i’’ represents the app being considered in the calculation. CR1 and CR2 denote distinct countries/regions within the comparison pair, and they represent the app ratings for the i-th app in each respective country/region. In total, the computation produces 3655 mean absolute percentage difference ($MAPD_{pair}$) values, offering insights into the overall similarity or divergence in user preferences for apps across diverse geographical locations on both the Play Store.

Table 4. Apps Locally Available on Play Store and App Store

App Name	
WhatsApp Messenger	Facebook
YouTube	Instagram
Messenger	Google Chrome
Snapchat	Duolingo
Google	InShot
X	Google Maps
WhatsApp Business	Viber Messenger
Telegram	Uber
Microsoft Word	Gmail
Piscart	Google Meet
Shazam	Pinterest
Google Drive	Microsoft Outlook
Google Translate	Waze Navigation
Microsoft Teams	Adobe Acrobat Reader
Microsoft OneDrive	FaceApp

Figure 2 shows the matrix plot of these MAPD values, providing a visual representation of similarity in country-based user app ratings, where dark blue represents the highest similarity and dark red represents the lowest similarity. The findings suggest that cultural, linguistic, and geographical factors may play a pivotal role in shaping similarities in app ratings across countries. For instance, the top five pairs with the lowest MAPD values — Colombia and Mexico, Algeria and Morocco, Ghana and Nigeria, Brazil and Mexico, and Guatemala and Honduras — all exhibit significant cultural and geographical ties. These countries share similar languages, cultural practices, and regional influences, which likely contribute to their closely aligned app preferences. Although Canada, the United Kingdom, and Australia are geographically distant, they also exhibit high similarity in app ratings, likely due to their deep historical and cultural ties, which significantly influence their app rating patterns. An interesting observation from the data is that users from Japan and Russia exhibit distinct app rating patterns compared to those from other countries. Both countries have notably higher MAPD values, suggesting a significant divergence in user preferences compared to the rest.

4.8. Store Comparison

The same app can be available on different application stores for different mobile operating systems. To investigate whether user app ratings differ significantly between the Play Store and the App Store within the same country, user ratings were analyzed for each platform across various countries. The aim is to determine whether platform-specific dynamics results in significant differences in user feedback for the same apps within the same geographical region. For this analysis, the previously determined country regions in Table 1 were used as the basis for comparison, except for Myanmar. However, finding apps that are available in all countries on both the Play Store and the App Store is challenging. Some apps are unavailable on the App Store in certain countries, such as the Spotify app in Russia, the ChatGPT app in Venezuela, CapCut app in India, and the AliExpress Shopping app in Georgia. Therefore, the classified global apps on the Play Store were reviewed, starting with the app that has the highest number of ratings, until 30 apps were identified. A sample size of 30 is often considered the minimum for a reasonably reliable analysis, particularly for descriptive statistics and simple hypothesis testing, based on the Central Limit Theorem (Field, 2013).

Table 4 presents a list of 30 mobile apps available on both the Play Store and the App Store across the considered countries. When comparing the user ratings for these apps within the same country on both stores, 2103 out of 2550 user ratings from the App Store (82.5%) were higher than those from the Play Store, while only 447 user ratings from the Play Store (17.5%) were higher than those from the App Store. In addition, the paired t-test was applied to analyze whether there are statistically significant differences in user ratings between the Play Store and the App Store. The paired t-test compares the ratings for the same apps across the two platforms within each country. Specifically, a p-value of less than 0.05 indicates that the difference in ratings between the Play Store and the App Store is statistically significant, meaning that the observed differences are unlikely to have occurred randomly. Of the 72 countries where a significant difference was observed, 71 showed that the App Store had better ratings than the Play Store, while only 1 country (Libya) had a significant difference in favor of the Play Store. In contrast, there were 13 countries (United Arab Emirates, Lebanon, Morocco, Egypt, Malaysia, Yemen, Nigeria, Oman, Jordan, Senegal, Tunisia, Iraq, Saudi Arabia) where no statistically significant difference in ratings was found between the two platforms.

Cohen's *d* is applied as a measure of effect size to assess the practical significance of the results beyond statistical significance. The magnitude of the difference is practically meaningful, with a Cohen's *d* of 0.7139, indicating a medium-to-large effect. This suggests that the difference in user ratings between the App Store and Play Store is not only statistically significant but also substantial in practical terms. In other words, the difference is large enough to be noticeable in real-world settings rather than trivial. This effect size indicates a consistent tendency for users to rate apps more favorably on the App Store across the sampled countries, reinforcing the robustness of the observed platform-level disparity in user satisfaction.

The results suggest that users tend to rate apps more favorably on the App Store compared to the Play Store, indicating potential platform-specific dynamics that influence user feedback. One possible explanation is that developers often prioritize iOS, since Apple devices are more standardized and easier to optimize. In contrast, Android apps must function across a wide range of devices, which can increase the likelihood of bugs and performance issues. As a result, iOS apps may offer more stable performance and fewer crashes, which can contribute to higher user ratings.

5. Limitations

Despite its contributions, this study has several limitations. One potential limitation is its focus on popular apps in both the Play Store and App Store. App availability can vary across countries and regions due to factors such as government regulations, content restrictions, company policies, or decisions by app store operators, which may affect the generalizability of the findings. Additionally, the dataset is limited to apps with at least 5 million downloads and 500,000 ratings across 32 Play Store categories, which confines the analysis to highly popular apps in the upper segment of the ecosystem. As a result, mid-tail and long-tail apps are excluded, further restricting the generalizability of the findings to the broader mobile app market. Another limitation is that apps are dynamic, and changes over time, such as shifts in revenue strategies (e.g., moving from paid to free or ad-supported models, or vice versa), might impact the findings.

6. Conclusion

This study conducted a comprehensive comparative analysis of mobile applications across multiple countries and platforms, focusing on variations in user ratings, review patterns, and monetization strategies. By collecting region-specific data from both the Play Store and App Store, the research provided insights into how app popularity and user satisfaction differ across diverse markets. The key findings identified through the analysis, based on the research questions listed below, are as follows:

*The uneven distribution of popular apps across Play Store categories suggests that competition is concentrated in specific domains, making category selection a strategic factor for visibility.

*All popular apps in the dataset are free, with 53.85% supported by advertisements, while 22% of the popular free apps use a freemium model, indicating that accessibility is essential for large-scale adoption.

*Approximately 64% of popular apps are considered global, being available and rated in at least 81 of 86 localized Play Store regions, highlighting the importance of scalability in successful apps.

*Top global apps exhibit extremely high user engagement, suggesting that sustained interaction is a key driver of app dominance, with WhatsApp, Facebook, YouTube, and Instagram leading overall. However, rankings based on total ratings and total reviews do not always align.

*A weak negative correlation ($r = -0.1226$) exists between the review rating ratio and the average user rating, indicating a slight tendency for more critical users to provide reviews; however, the effect is weak and of limited practical relevance. Cohen's d (≈ -0.25) confirms a small effect size.

*Most global apps maintain consistent user ratings across countries. Approximately 6.9% are fully consistent, 58.7% are likely consistent, and about 35% show varying degrees of inconsistency, highlighting the influence of regional factors on user experience.

*Cultural, linguistic, and geographic factors significantly influence rating similarities. Countries with shared language or culture, such as Colombia–Mexico or Algeria–Morocco, display higher similarity, while Japan and Russia exhibit distinct patterns.

*User ratings tend to be higher on the App Store than on the Play Store in most countries. In 82.5% of cases, App Store ratings exceeded Play Store ratings, and these differences are statistically significant with a medium-to-large effect size (Cohen's d = 0.7139), indicating systematic differences in user expectations across platforms.

The present research extends existing work by providing a unified, large-scale comparative analysis that integrates user ratings, engagement patterns, monetization strategies, and app distribution across countries and platforms. Unlike prior fragmented studies that focus on isolated dimensions or limited regions, this research demonstrates that app outcomes are shaped by a combination of app-level characteristics, platform ecosystems, and regional-cultural factors. In particular, the findings empirically show consistent cross-country variation, platform-based rating differences, and the role of cultural and linguistic proximity in shaping user evaluations. This contributes to a more integrated, ecosystem-oriented understanding of mobile app dynamics, moving beyond single-dimension or country-specific analyses.

These findings have several implications for developers. Category-level differences indicate uneven competitive pressure across domains, suggesting that strategic category selection is critical for achieving visibility. Monetization patterns confirm that free-to-download models are essential for scaling, while advertising and freemium strategies provide flexible and sustainable revenue pathways depending on app design and user engagement. Scalability potential varies systematically across app types, particularly between globally oriented applications (e.g., communication, entertainment, productivity) and regionally focused apps. Global applications require early integration of localization and scalable infrastructure, while regional apps benefit more from culturally specific design adaptation. In addition, engagement patterns indicate that retention and sustained interaction are stronger predictors of long-term success than initial download volume. Platform-specific differences further indicate that user expectations are not uniform across ecosystems. Higher ratings on iOS suggest that developers should adapt quality benchmarks and release strategies according to platform-specific behavior rather than applying identical standards. Finally, cultural and linguistic differences highlight that localization should be treated as a design-level adaptation process rather than simple translation, particularly for apps operating across diverse regions.

Possible research directions include the temporal evolution of these patterns, the role of user sentiment beyond quantitative ratings, and the interaction between localization strategies and long-term app success. Building on these concerns, future research needs to move beyond static comparisons and examine how these dynamics evolve over time and across emerging digital platforms. In particular, future work will explore temporal changes in app availability, integrate sentiment analysis of user reviews, and consider additional dynamic indicators to achieve a more comprehensive insight of user behavior and evolving trends within global digital ecosystems.

Conflict of Interest

No conflict of interest has been declared by the author.

References

Ahmad, W., Kim, W., Choi, H., & Haq, J. (2021). Modeling behavioral intention to use travel reservation apps: A cross-cultural examination between US and China. *Journal of Retailing and Consumer Services, 63*, 102689. https://doi.org/10.1016/j.jretconser.2021.102689

Al-Subaihin, A. A., Sarro, F., Black, S., Capra, L., & Harman, M. (2021). App store effects on software engineering practices. *IEEE Transactions on Software Engineering*, *47*(2), 300–319. https://doi.org/10.1109/TSE.2019.2891715

Alchemer (2022). Mobile app ratings and reviews: Where to start and how to win. Retrieved 10.12.2023, from https://www.alchemer.com/resources/blog/mobile-app-ratings-and-reviews/

Ali, M., Joorabchi, M. E., & Mesbah, A. (2017). Same app, different app stores: A comparative study. *2017 IEEE/ACM 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft)*, 79–90. https://doi.org/10.1109/MOBILESoft.2017.3

Altman, D. G. (1990). *Practical statistics for medical research*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429258589

AndroidRank (2025). Android market data, history, rankings. Retrieved 09.01.2025, from https://www.androidrank.org/

Chang, K. C., Zaeem, R. N., & Barber, K. S. (2020). A framework for estimating privacy risk scores of mobile apps. In W. Susilo, R. H. Deng, F. Guo, Y. Li, & R. Intan (Eds.), *Information security (ISC 2020)* (Lecture Notes in Computer Science, Vol. 12472). Springer. https://doi.org/10.1007/978-3-030-62974-8_13

Chen, N., Lin, J., Hoi, S. C. H., Xiao, X., & Zhang, B. (2014). AR-miner: Mining informative reviews for developers from mobile app marketplace. In

Proceedings of the 36th International Conference on Software Engineering (pp. 767–778). Association for Computing Machinery. https://doi.org/10.1145/2568225.2568263

Dabrowski, J., Letier, E., Perini, A., & Susi, A. (2022). Analysing app reviews for software engineering: A systematic literature review. *Empirical Software Engineering*, *27*(2), Article 43. https://doi.org/10.1007/s10664-021-10065-7

De Lima, V. M. A., & Marcacini, R. M. (2024). Opinion mining for app reviews: Identifying and prioritizing emerging issues for software maintenance and evolution. In *Proceedings of the XXIII Brazilian Symposium on Software Quality* (pp. 687–696). Association for Computing Machinery. https://doi.org/10.1145/3701625.3701700

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications.

Gao, C., Zeng, J., Lyu, M. R., & King, I. (2018). Online app review analysis for identifying emerging issues. In *Proceedings of the 40th International Conference on Software Engineering* (pp. 48–58). Association for Computing Machinery. https://doi.org/10.1145/3180155.3180218

Gao, J., & Dube, A. K. (2024). Comparison of learning experience among informal learning apps. In *ICERI2024 Proceedings: 17th Annual International Conference of Education, Research and Innovation* (pp. 337–341). IATED. https://doi.org/10.21125/iceri.2024.0173

Guo, Z., Rau, P. L. P., & Heimgärtner, R. (2022). The "onion model of human factors": A theoretical framework for cross-cultural design. In P. L. P. Rau (Ed.), *Cross-Cultural Design. Interaction Design Across Cultures* (pp. 20–33). Springer Cham. https://doi.org/10.1007/978-3-031-05434-1_2

Gupta, S., Pansari, A., & Kumar, V. (2018). Global customer engagement. *Journal of International Marketing*, *26*(1), 4–29. https://doi.org/10.1509/jim.17.0091

Guzman, E., Oliveira, L., Steiner, Y., Wagner, L. C., & Glinz, M. (2018). User feedback in the app store: A cross-cultural study. In *Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Society* (pp. 13–22). Association for Computing Machinery. https://doi.org/10.1145/3183428.3183436

Hu, H., Wang, S., Bezemer, C.-P., & Hassan, A. E. (2019). Studying the consistency of star ratings and reviews of popular free hybrid Android and iOS apps. *Empirical Software Engineering*, *24*(1), 7–32. https://doi.org/10.1007/s10664-018-9617-6

Hyndman, R. J., & Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician, 50*(4), 361–365. https://doi.org/10.1080/00031305.1996.10473566

Iacob, C., Veerappa, V., & Harrison, R. (2013). What are you complaining about? A study of online reviews of mobile applications. *Proceedings of the 27th International BCS Human Computer Interaction Conference*, Article 29, 1–6.

Lim, S. L., Bentley, P. J., Kanakam, N., Ishikawa, F., & Honiden, S. (2015). Investigating country differences in mobile app user behavior and challenges for software engineering. *IEEE Transactions on Software Engineering*, *41*(1), 40–64. https://doi.org/10.1109/TSE.2014.2360674

Liptrot, E., Pearson, H. A., Montazami, A., & Dubé, A. K. (2024). Why this app? How user ratings and app store rankings impact educators' selection of educational apps. *Computers & Education*, *218*, Article 105080. https://doi.org/10.1016/j.compedu.2024.105080

Liu, X., Li, W., Hou, Q., Yang, S., Ying, L., Diao, W., … Duan, H. (2024). From promises to practice: Evaluating the private browsing modes of Android browser apps. *Proceedings of the ACM Conference*. https://doi.org/10.1145/3589334.3645320

Mahmood, A. (2020). Identifying the influence of various factor of apps on Google Play apps ratings. *Journal of Data, Information and Management*, *2*(1), 15–23. https://doi.org/10.1007/s42488-019-00015-w

Martin, W., Sarro, F., Harman, M., Jia, Y., & Zhang, Y. (2017). A survey of app store analysis for software engineering. *IEEE Transactions on Software Engineering*, *43*(9), 817–847. https://doi.org/10.1109/TSE.2016.2630689

Meng, Y., Park, S., & Um, G. (2026). How review sentiment influences ratings in generative AI applications: Insights from VADER and LDA analysis. *Journal of Retailing and Consumer Services*, *88*, Article 104560. https://doi.org/10.1016/j.jretconser.2025.104560

Moore, D. S., McCabe, G. P., & Craig, B. A. (2017). *Introduction to the practice of statistics* (9th ed.). W. H. Freeman.

Ngereza, A., Muller, V., Lippu, C. O., Kutela, B., Kidando, E., & Kitali, A. (2026). Exploring benefits and challenges of 511 applications in enhancing mobility and safety by evaluating user reviews. *Transportation Research Record*, *2680*(2), 598–613. https://doi.org/10.1177/03611981251366256

Panichella, S., Di Sorbo, A., Guzman, E., Visaggio, C. A., Canfora, G., & Gall, H. C. (2015). How can I improve my app? Classifying user reviews for software maintenance and evolution. *Proceedings of the 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 281–290. https://doi.org/10.1109/ICSM.2015.7332474

Phong, M. V., Nguyen, T. T., Pham, H. V., & Nguyen, T. T. (2015). Mining user opinions in mobile app reviews: A keyword-based approach *Proceedings of the 30th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 749–759. https://doi.org/10.1109/ASE.2015.85

Pratesi, F., Hu, L., Rialti, R., Zollo, L., & Faraoni, M. (2021). Cultural dimensions in online purchase behavior: Evidence from a cross-cultural study. *Italian Journal of Marketing, 2021*, 227-247. https://doi.org/10.1007/s43039-021-00022-z.

Rahaman, S., Farooq, U., Neamtiu, I., & Zhao, Z. (2023). Detecting potential user-data save & export losses due to Android app termination. *2023 IEEE/ACM International Conference on Automation of Software Test (AST)*, 152–162. https://doi.org/10.1109/AST58925.2023.00019

Ravichander, A., Black, A. W., Wilson, S., Norton, T., & Sadeh, N. (2019). Question answering for privacy policies: Combining computational and legal perspectives. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 4947–4958. https://doi.org/10.18653/v1/D19-1500

Sallberg, H., Wang, S., & Numminen, E. (2023). The combinatory role of online ratings and reviews in mobile app downloads: An empirical investigation of gaming and productivity apps from their initial app store launch. *Journal of Marketing Analytics*, *11*(3), 426–442. https://doi.org/10.1057/s41270-022-00171-w

Srisopha, K., Phonsom, C., Lin, K., & Boehm, B. (2019). Same app, different countries: A preliminary user reviews study on most downloaded iOS apps. *Proceedings of the 2019 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 76–80. https://doi.org/10.1109/ICSME.2019.00017

Statista (2025a). Market share of mobile operating systems worldwide from 2009 to 2025. Retrieved 07.30.2025, from https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/

Statista (2025b). Number of apps available in leading app stores as of august 2024 and june 2025. Retrieved 07.30.2025, from https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/

Teles, A., Rodrigues, I., Viana, D., Silva, F., Coutinho, L., Endler, M., ... Rabelo, R. (2019). Mobile mental health: A review of applications for depression assistance. *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, 708–713. https://doi.org/10.1109/CBMS.2019.00143

Thompson, F., & Brouthers, K. (2021). Digital consumer engagement: National cultural differences and cultural tightness. *Journal of International Marketing, 29*, 22-44. https://doi.org/10.1177/1069031x211005729

Tjikhoeri, N., Olson, L., & Guzmán, E. (2024). The best ends by the best means: Ethical concerns in app reviews. *Empirical Software Engineering*, *29*(6), Article 138. https://doi.org/10.1007/s10664-024-10463-7

Yu, L., Wang, H., Luo, X., Zhang, T., Liu, K., Chen, J., ... Xiao, X. (2023). Towards automatically localizing function errors in mobile apps with user reviews. *IEEE Transactions on Software Engineering*, *49*(4), 1464–1486. https://doi.org/10.1109/TSE.2022.3178096

Zhao, Y., Wang, N., Li, Y., Zhou, R., & Li, S. (2020). Do cultural differences affect users' e-learning adoption? A meta-analysis. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13002