

# **The Economics of App Success: How Revenue Streams Influence Downloads and User Ratings**

Group Assignment - Strategy and Business Models

Group 5

28/11/2024



# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Academic Background . . . . .	3
1.2	Societal Background . . . . .	4
1.3	Research Gap . . . . .	4
<b>2</b>	<b>Theory and Hypotheses</b>	<b>5</b>
2.1	Literature Review . . . . .	5
2.1.1	Popularity . . . . .	5
2.1.2	Rating . . . . .	5
2.1.3	Revenue Estimation . . . . .	6
<b>3</b>	<b>Methods and Data</b>	<b>6</b>
3.1	Dataset Description . . . . .	7
3.2	Variable Selection . . . . .	7
3.3	Statistical Methods . . . . .	8
3.3.1	Handeling of Assumptions . . . . .	9
<b>4</b>	<b>Results</b>	<b>9</b>
4.1	Number of Downloads. . . . .	10
4.1.1	Distribution of app downloads. . . . .	11
4.1.2	Games in the Google Play Store . . . . .	13
4.1.3	Sample and premium Apps . . . . .	16
4.2	Ratings . . . . .	17
4.2.1	Quality of Free and Paid Apps . . . . .	18
4.2.2	Variance in Ratings . . . . .	20
4.2.3	Correlation between Sample and Premium Apps . . . . .	21
4.3	Conclusion . . . . .	22
<b>5</b>	<b>Discussion</b>	<b>22</b>
5.1	Reflection on the Findings . . . . .	22
5.2	Practical Implications for Businesses . . . . .	22
5.3	Future Research Directions . . . . .	22
<b>6</b>	<b>References</b>	<b>22</b>

# 1 Introduction

With the every-growing popularity of cellphones (Richter 2023), the popularity of mobile applications is also steadily increasing. In 2024, mobile applications are estimated to generate over \$900 billion in revenue (“Mobile App Revenue Worldwide by Segment (2019-2027)” 2023). Generally, mobile applications (*‘apps’* from here on) tend to be categorized in three different categories (Roma and Ragaglia 2016). Paid apps are the most transparent; they revenue is based on an up-front purchase by the user. Free apps, on the other hand, require no purchase by the user at any stage. According to Roma and Ragaglia (2016), these apps make their revenue from deals with third-parties, either through advertisement or other purposes such as market information.

Finally, freemium apps are, as the name suggests, a middle-ground between free and premium. Users get access to a basic version of the application first and can unlock more features through an in-app payment (Kumar 2014). Of these three revenue models, freemium is the most commonly used and the most (Salehudin and Alpert 2021) and leads to more downloads as well as revenue (Liu, Au, and Choi 2014).

## 1.1 Academic Background

Most research uses these three established categories—paid, freemium, and free—when discussing revenue models for apps. However, by limiting the discussion to these three terms, nuances within these categories might be missed.

In a review paper from 2023 (Djaruma et al. 2023), different levels of monetization are suggested based on previous literature. These levels provide a clear framework for the revenue models of mobile apps.

Strategy	Description
Level 5: Premium	Pay to use the application. This either happens up-front, or after a trial period.
Level 4: Semi-premium	Use a limited number of features for free. Unlock the app with all features through an in-app purchase.
Level 3: In-app advertisement and in-app purchases	Free application with ads, encouraging users to remove ads or to make in-app purchases.
Level 2: Sample and premium	Two different versions of the same app. One is a version with limit features and/or ads. The other version is a premium version.
Level 1: In-app advertisement	Only one version of the app, with only ads and no in-app purchases.
Level 0: Free	The app has no monetization. However, money can still be made through selling user information.

Strategy	Description
Table 1: Six levels of monetization for apps	

## 1.2 Societal Background

Currently, most apps utilize the freemium revenue model (Salehudin and Alpert 2021). However, as discussed in Djaruma et al. (2023), there are many revenue models between completely premium and completely free. A more fine-grained classification of app revenue models beyond the traditional “paid-freemium-free” framework holds significant societal and business implications.

For society, such distinctions enhance transparency. Some monetization models, such as free or ad-filled apps, may rely on selling user information as a source of revenue (Bamberger et al. 2020). Therefore, clearer distinctions regarding the revenue model will empower consumers to make informed choices. It may also enable policymakers to identify and regulate exploitative practices, such as manipulative microtransactions or intrusive ad models, ensuring all applications align with ethical and legal standards (Mileros and Forchheimer 2024).

For businesses, this paper should unlock more insight into the effectiveness of different revenue streams. This will allow developers to tailor monetization strategies to specific audiences. Furthermore, both consumers and regulatory bodies are growing more concerned with the privacy concerns of apps, especially ones that rely on market information (Mileros and Forchheimer 2024). A granular understanding helps businesses adapt, aligning profitability with sustainability and ethical considerations.

## 1.3 Research Gap

In short, apps play an increasingly important role in our techno-centric society. To improve the user experience and increase profits, consideration of revenue models is key. Despite the great depth of research on this topic, literature tends to be focussed on the three big categories of paid, freemium, and free. This lack of nuance prevents us from understanding the fine-grained details that may help improve future apps.

The levels of monetization as proposed by Djaruma et al. (2023) would allow for this nuance. However, their framework has never been used in an empirical setting, as the paper by Djaruma et al. (2023) was published only last year. Applying this framework to see how different revenue streams impact the popularity of an app may yield valuable insights into the preferences of consumers. Therefore, the question to answer within this paper will be: *How are the 5 different revenue models as proposed by Djaruma et al. (2023) correlated to the success of an app?*

## 2 Theory and Hypotheses

In this section, prior research into the topic of revenue streams and its correlation to success in apps will be discussed. As mentioned in the Introduction section, this paper will apply the 5 levels of revenue as proposed by Djaruma et al. (2023) to app data. The following section will contain a holistic overview of the existing research, as well as hypotheses that arise from this theoretical framework.

### 2.1 Literature Review

To answer the question “*How are the 5 different revenue models as proposed by Djaruma et al. (2023) correlate to the success of an app?*”, we must first define what constitutes to success. In this paper, success will be defined by a couple of factors: popularity, rating, and estimated revenue.

#### 2.1.1 Popularity

The popularity of an app can be measured by the number of downloads. It is important to note the popularity of an app is complex, and is not solely dependent on the chosen revenue model. Other features, such as whether an app is featured on charts, whether it has frequent updates, and word-of-mouth awareness, will also impact the popularity of an app (Aydin Gokgoz, Ataman, and Bruggen 2021). However, despite these other variables, to versions of the same app will still have drastically different performances with different revenue streams (Liu, Au, and Choi 2014).

H1a: Apps that allow the user to have free access to all features (level 0 and 1) will have the highest amount of downloads overall. However, the ratings may fluctuate, as quality can vary for free-to-access apps.

H1b: The apps with the most downloads will be level 1. Most social media platforms, which dominate our culture, tend to have this revenue stream (Djaruma et al. 2023).

H1c: For apps that utilize a sample and a premium version of the same app (level 2), the free versions of an app will have more downloads than their paid-for counterpart. Most, if not all, users will download the free version first, and then might upgrade. This means there should be a disparity between the number of downloads between the apps, as is also demonstrated by Liu, Au, and Choi (2012).

H1d: The most downloaded apps in the gaming category will likely fall under level 4. Many popular games use this type of “pay-to-win” mechanism (Nieborg 2016). Therefore, it would be expected this same pattern would arise from our data.

#### 2.1.2 Rating

The downloads of an app are not everything. An app can be downloaded often, but may not be highly rated.

H2a: Apps that require the user to pay to unlock features (level 2, 3, and 4) will tend to have lower ratings than the version that requires payment upfront (level 5). The main draw of a freemium model is to attract users, and have them update to a paid version (Kumar 2014). However, as Kumar (2014) points out, this can be a double-edged sword. Too few features, and it may not be attractive to users. Too many features, and the users will not update.

H2b: Fully premium apps (level 5) will have less variance in their ratings, while all other levels will have more. In the same vein as H2a, users have more realistic expectations of paid apps compared to apps that require you to unlock features (Kumar 2014). Therefore, more users downloading premium apps will be satisfied with their purchase, leading to less variance.

H3: For apps that utilize a sample and a premium version of the same app (level 2), the rating of the paid-for version is positively associated with the rating of the free version of the same app. This was true for the study on the most popular apps in the Google Play Store by Liu, Au, and Choi (2012), so it is expected a similar pattern should arise for this dataset.

### **2.1.3 Revenue Estimation**

It is important to point out downloads and ratings likely do not directly correlate to the actual revenue of an app. The revenue of apps “premium” apps that require an upfront payment, the revenue is relatively simple to track and compare. However, for apps that rely on advertisement, in-app purchases and/or selling market information, this is harder to track.

For apps that solely on advertisement, time retention can be a good measure of revenue (Ross 2018). However, this only works if the app solely relies on ads. An example of this given by Djaruma et al. (2023) is TikTok: this app relies not only on advertisement, but also on users purchasing products through its shop. Therefore, using solely the time retention would not accurately capture the revenue of an app with both revenue streams. Furthermore, the selling of user data is usually not publicized, meaning it is not possible to know the revenue from this.

Unfortunately, our data only contains the price of “premium” app versions. The data does not include any details regarding in-app purchases nor time-retention. Because of this lack of sufficient data, solely downloads and ratings will be taken into account as indicators of success.

## **3 Methods and Data**

In this section, we will discuss the dataset and methods used to test the hypotheses outlined in the previous section. The focus lies on providing a comprehensive description of the dataset, including its structure and the variables it contains, followed by an explanation of the variable selection process. Additionally, we outline the statistical methods applied and discuss how assumptions, such as missing values and potential biases, were addressed to ensure the robustness of our analysis.

### 3.1 Dataset Description

The dataset used for this research consists of 1,016,666 instances and 27 variables, representing a detailed overview of mobile applications across various revenue models. Each instance corresponds to an app, and the variables capture key attributes such as app downloads, user ratings, and monetization strategies. Below is an overview of some of variables:

Variable	Description
my_app_id (object)	Unique identifier for each app.
date_published (object)	The publication date of the app. Only three missing values (0.000295% null).
privacy_policy (object)	Information about the app's privacy policy, missing in 28.57% of cases.
rating_app (float64)	The average rating of the app, with 8.76% missing values.
nb_rating (object)	Number of ratings received by the app, missing in 8.76% of cases.
num_downloads (object)	The number of downloads for the app, nearly complete with only 15 missing values (0.001475% null).
price_gplay (object)	The price of the app as listed on Google Play, missing in 0.43% of cases.
in_app (bool)	Indicates whether the app has in-app purchases (no missing values).
has_ads (bool)	Indicates whether the app contains advertisements (no missing values).
content_rating_app (object)	The app's content rating, with three missing values (0.000295% null).
developer_name (object)	The name of the app developer, missing in only 16 instances (0.001574% null).

The dataset includes several additional variables related to app features, developer information, and user engagement metrics such as `visit_website`, `more_from_developer`, and `family_library`. However, some variables, such as `whats_new` (100% null) and `in_app_product` (89.57% null), were deemed unsuitable for analysis due to their high proportion of missing data.

The primary purpose of this dataset in this study, is to analyze app monetization strategies by categorizing apps into distinct revenue levels and evaluating their performance based on key metrics like downloads and user ratings.

### 3.2 Variable Selection

The dataset utilized in this study consists of 1,016,666 entries, encompassing a broad range of attributes related to mobile applications. For the purpose of our analysis, 13 variables

were selected, capturing critical information about app characteristics, user engagement, monetization strategies, and developer details. These variables include the app’s unique identifier (`my_app_id`), the total number of downloads (`num_downloads`), average user ratings (`rating_app`), and the number of ratings (`nb_rating`). Additionally, the dataset provides information on app pricing (`price_gplay`), the presence of in-app purchases (`in_app`), and whether the app includes advertisements (`has_ads`). Other variables, such as content ratings (`content_rating_app`), app categories (`categ_app`), and developer information (`developer_name` and `developer_info`), further enhance the richness of the dataset. This subset of variables allows us to comprehensively examine the interplay between monetization strategies and app success.

To systematically explore monetization strategies, we classified the apps into six distinct levels based on their monetization models. These levels reflect varying approaches to generating revenue, ranging from completely free apps to fully premium paid apps.

Level 0 represents apps with no monetization, offering free services without ads or in-app purchases. At the opposite end, Level 5 includes premium apps requiring upfront payment, free from ads or in-app purchases, delivering a premium experience.

In between, Level 1 consists of free apps monetized solely through ads, while Level 3 combines ads and in-app purchases, offering additional features for users willing to pay. Level 4 refines the freemium model by removing ads and relying entirely on in-app purchases to monetize.

Level 2 employs a dual-version strategy, featuring both free sample apps with limited functionality (and potentially ads) and paid premium apps with comprehensive features and no ads or in-app purchases.

This classification is grounded in theoretical frameworks, such as the monetization levels proposed by Djaruma et al. (2023) and the App business models of (CITE), and allows for a nuanced analysis of how different revenue models impact app success metrics like user ratings and downloads. Our systematic categorization facilitates a deeper understanding of the relationship between monetization strategies and app performance.

### 3.3 Statistical Methods

To test our hypotheses, we employed a combination of descriptive statistics, text processing, and machine learning techniques. Descriptive statistics were utilized to analyze distributions and trends in metrics such as `num_downloads`, `rating_app`, and `price_gplay`. We categorized applications into six monetization levels based on binary indicators: `is_free`, `in_app`, and `has_ads`. Price values were processed to distinguish between free and paid applications.

To identify paired sample and premium applications within level 2, we applied Term Frequency-Inverse Document Frequency (TF-IDF) vectorization combined with cosine similarity on application names. This approach is effective for measuring textual similarity between documents (CITE Source 2). Subsequent filtering involved prefix matching and the identification of indicative terms (e.g., “Free,” “Pro”) to ensure logical pairing based on naming conventions and shared developers.



To adjust ratings for applications with few reviews, we calculated a Bayesian average. This method provides a more robust measure of user satisfaction by accounting for the number of ratings and the overall average rating across all applications (CITE Source 3). Visualizations, including scatter plots and box plots, were employed to explore relationships between monetization levels and user engagement metrics which will be displayed in the results section.

### 3.3.1 Handling of Assumptions

We addressed missing values by removing rows with critical nulls, such as those in `num_downloads`, to maintain data integrity. Text-based variables like `content_rating_app` were standardized to ensure consistency. For `price_gplay`, currency symbols were removed to facilitate the classification of applications into free or paid categories.

Outliers in metrics like `num_downloads` were retained if they represented industry-leading applications, as their exclusion could skew the analysis. The use of Bayesian averages mitigated bias in `rating_app` due to low review counts, providing a more accurate reflection of user satisfaction. Covariance checks were conducted to ensure the absence of multicollinearity among numerical variables, thereby enhancing the reliability of correlation and regression analyses.

Some applications exhibited rare combinations of `is_free`, `in_app`, and `has_ads` that did not fit within the predefined monetization levels. These applications were excluded from the analysis but documented as a limitation. Edge cases in level 2 application pairing were flagged for potential mismatches due to naming ambiguities, ensuring transparency in the classification process.

These methodologies facilitated a systematic and accurate exploration of monetization models and their impact on application performance.

## 4 Results

In this section, we will visualize the data through tables and visualizations. These plots largely explore the data around the hypotheses and research question, we discussed in the previous sections. The aim of this section is to present possible evidence in supporting a hypothesis. This will be discussed and concluded upon in the next section.

Before we look at the results, let's revisit the research question: *"How are the 6 different revenue models as proposed by Djaruma et al. (2023) correlate to the success of an app?"*. Where we identify 6 different revenue models, described in this Section 2.1 as (revenue) levels. For reference see Table 1.

The dataset is divided into these different levels. More than 75% of all the apps belong to level 0 and 1 (see Figure 1). With the smallest population being level 2 with two different version of the same app.

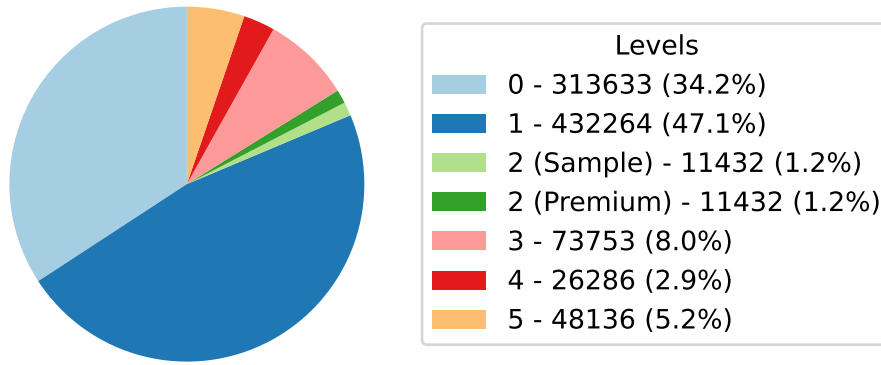
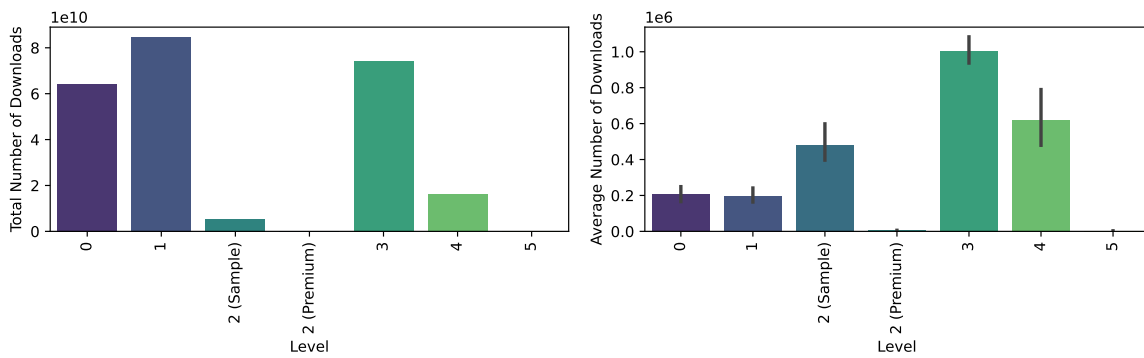


Figure 1: Distribution of the amount of apps across the revenue levels

To find the success of an app, as discussed in Section 2.1. Metrics like popularity, the rating and the revenue estimation of an application can be used. The popularity measured by the number of downloads and rating by the average rating given by users. (Note: Revenue estimation is excluded, because this is not measurable with the data.) These metrics are used to evaluate the hypotheses (see Section 2.1).

#### 4.1 Number of Downloads.

Figure 2 provides insights into the number of app downloads categorized by revenue levels. Figure 2a displays the total amount of downloads, while the Figure 2b displays the average amount of downloads. In the next following three subsections, we will look into the key takeaways from these two graphs: distribution of app downloads, games in the Google Play Store and sample and premium apps.



(a) Total Number of Downloads by Level

(b) Average Number of Downloads by Level

Figure 2: Comparison of Total and Average Downloads by Level

#### 4.1.1 Distribution of app downloads.

There is a significant difference between the total and average number of downloads for level 0 and 1, in Figure 2a and Figure 2b respectively. This is largely a result of the fact that 75% of the total apps fall under these levels (see Figure 1.) By testing the hypothesis 1b, in Section 2.1.1, it claims that apps in level 1 are downloaded the most, like social media platforms that dominate our culture. At a first glance the claim that this level “dominates our culture” doesn’t appear to be the case.

Level 0 and 1 seem to have the highest overall downloads, as shown in Figure 2a. Figure 3 illustrates just how big the difference between the total downloads is. With a difference of over 40 billion downloads ( $0.4e11$ ).

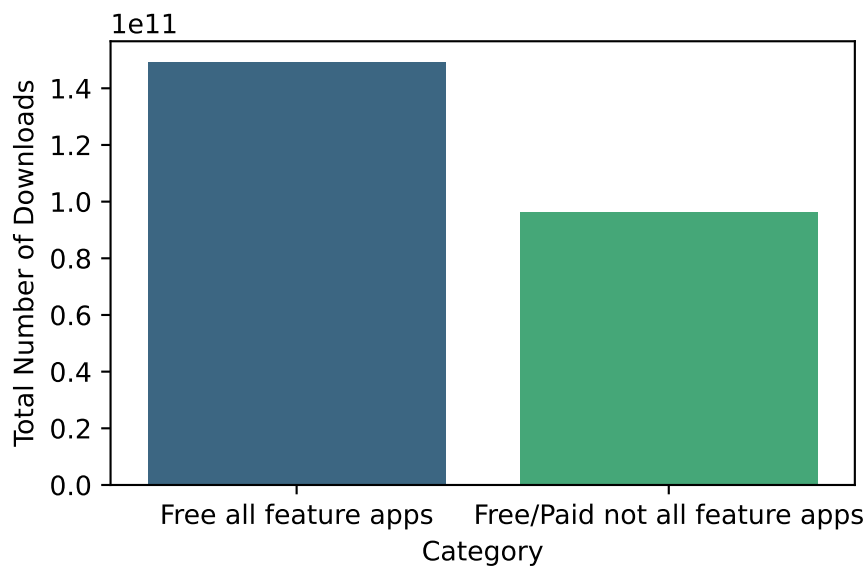


Figure 3: Difference Total Number of Downloads

Figure 4 provides insights into distribution inequality. The distribution of the app downloads are highly concentrated, meaning a small proportion of the apps make up for the vast majority of the downloads. Figure 4a shows relatively the same inequality as Figure 4b. Which indicates that a small amount of apps dominate the numbers, however this might also be present in other levels.

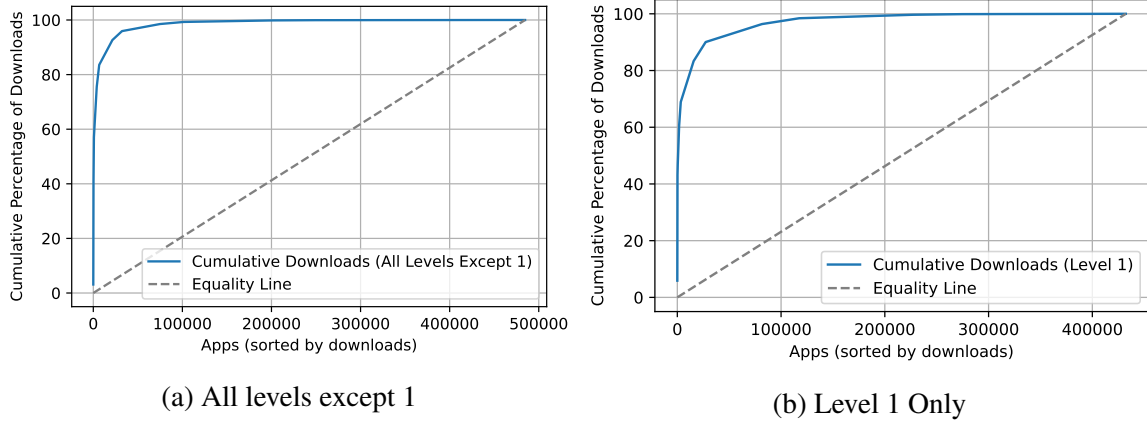


Figure 4: Distribution of App Downloads (Top-Heavy Analysis)

The download range of apps in Figure 5 illustrates that level 1 dominates in terms of apps with over 1 billion+ downloads. Furthermore, level 0 also has a few apps with over 1 billion downloads.

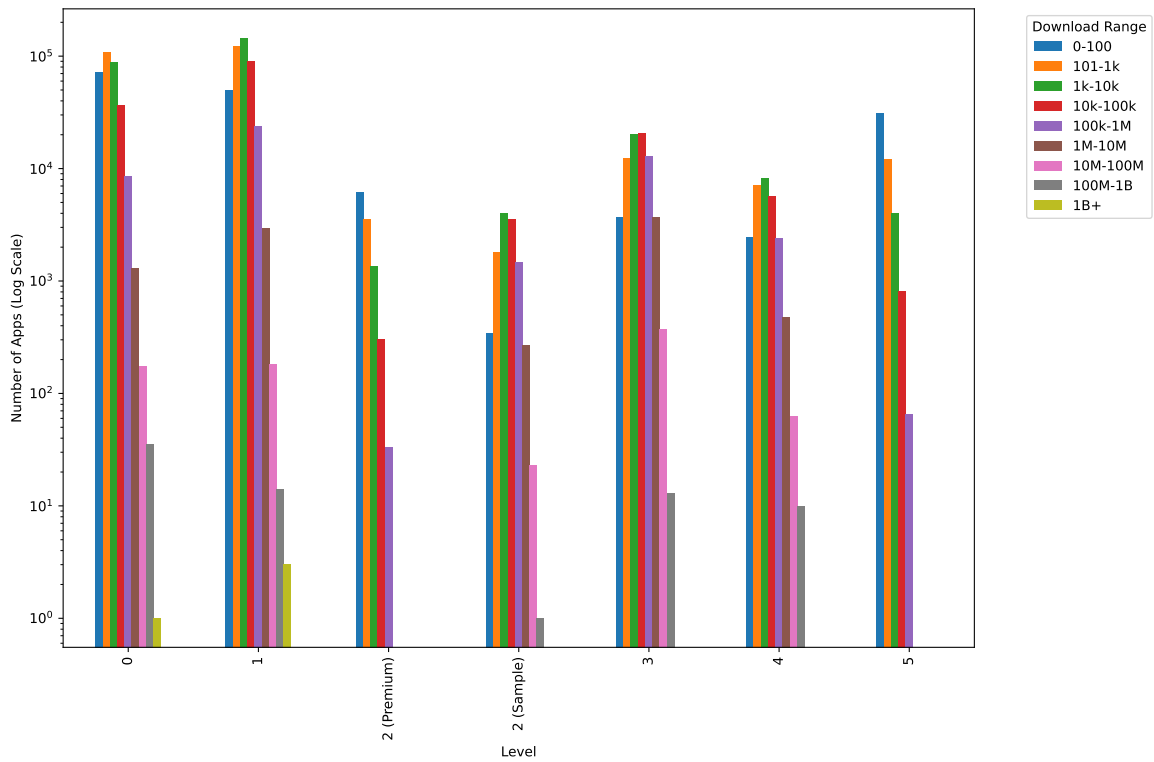


Figure 5: Number of Downloads by Level (Log Scale)

The apps categories Communication and Social classify as social platforms as shown in Figure 6. Attributing to 40% of all the apps with more than 1 billion downloads.

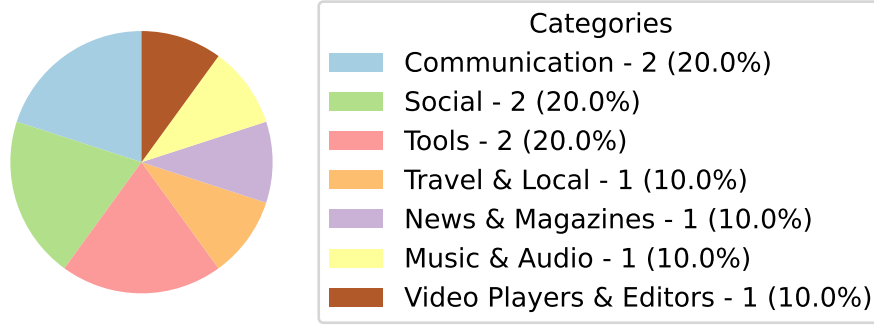


Figure 6: 1 Billion+ Downloads in Level 1

Taking a closer look at the apps provided by (Djaruma et al. 2023) in Table 3. We do see that Facebook, Instagram, Spotify, Snapchat and Amazon Shopping fall under level 1. (Disclaimer: TikTok didn't exist up until 2019).

	App Name	Downloads (in B)	Category	Level
0	WhatsApp	1	Communication	0
1	Facebook Lite	1	Social	1
2	Instagram	1	Social	1
3	Facebook Messenger	1	Communication	4
4	Spotify	0.5	Music & Audio	1
5	Snapchat	0.5	Social	1
6	Netflix	0.5	Entertainment	4
7	Amazon Shopping	0.1	Shopping	1

Table 3: Top 8 Apps by Downloads (in Billions)

It can be concluded that the most downloaded apps fall under level 1 (see Figure 2a). But this requires a nuanced view, since 50% of the apps fall under level 1 (see Figure 1) and it has a relatively low average among other levels (see Figure 2b). So, it requires further analysis to understand the popularity of level 1. As seen from Figure 4 it shows that there is high inequality, resulting in a small proportion of the apps accounting to a majority of the downloads. Furthermore, Figure 5 shows that apps in level 1 are considered among the top downloaded apps. In particular, the categories under which social media platform fall (see Figure 6). Looking at the apps provided by Djaruma et al. (2023), most social network platforms indeed fall under level 1 (see Table 3). Given these observations, it can be concluded that there is enough evidence to support h1b (see Section 2.1.1).

#### 4.1.2 Games in the Google Play Store

The difference between total and average number of downloads in level 3 and 4 is not relatively significant as shown in Figure 2. Furthermore, it is relatively high in both instances. Hypothesis 1d, which takes a closer look into these 2 revenue levels, claims that many popular games fall under these 2.

According to Figure 7b, 15.5% of the total apps are games. While these games attribute to 26.7% of the total app downloads, as shown in Figure 7a.

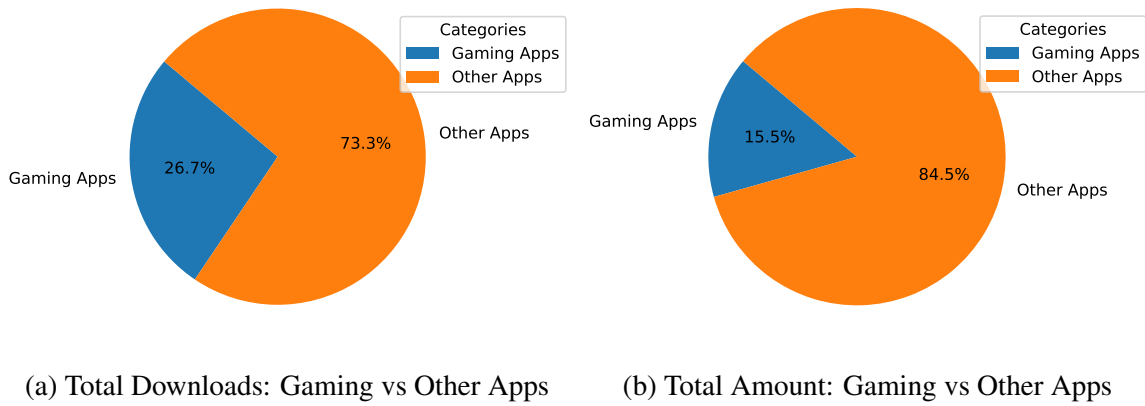


Figure 7: Proportion of Games compared to other apps

The average downloads of games is 98.6% higher than the average downloads of the other apps (see Table 4).

	Category	Avg Downloads	% Diff Gaming vs Other
0	Gaming	460828	98.6
1	Other	232041	-98.6
2	All	267450	-

Table 4: Average Downloads Games vs Other Apps

Up until now, these findings conclude that games are hugely popular across all levels. Figure 8 shows that games, especially in level 3 and 4, are even more popular. Averaging around 1 million downloads, compared to close to 0.5 million in Table 4.

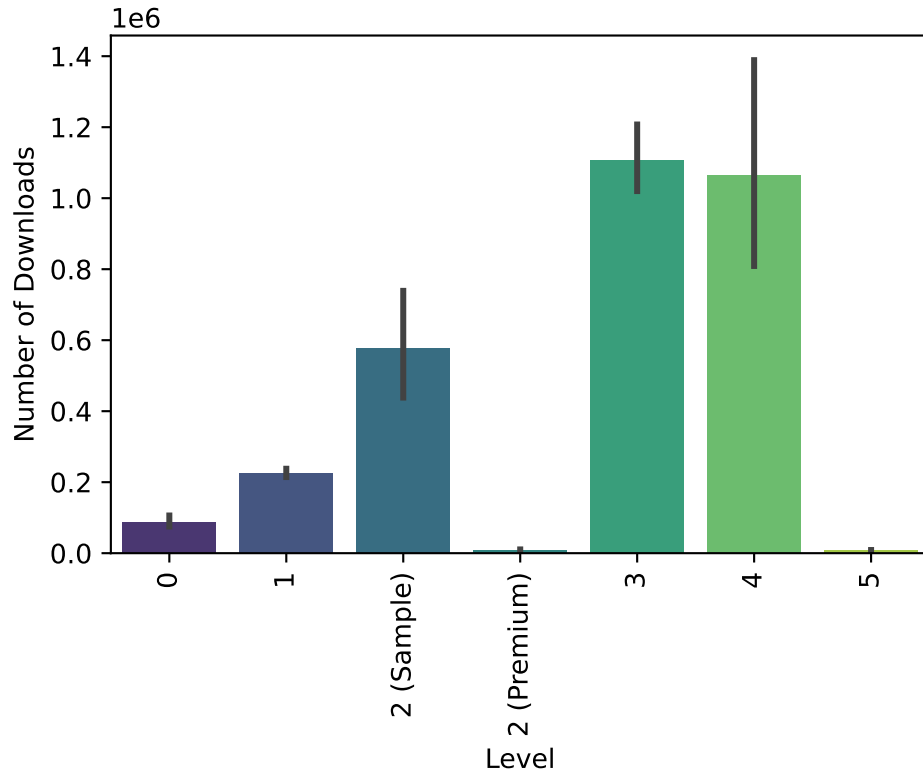


Figure 8: Average amount of Downloads for Gaming Apps by Level

The variance of level 4 is higher than level 3 as shown in Table 4. In Table 5, it reveals a notable variance in downloads between the 2 revenue levels. The variance in level 4 is much larger than in level 3, even if it only displays the top 5. For instance, the downloads in level 4 varies from 0.5 to 0.1 billion. While the downloads in level 3 range from 1.0 to 0.5 billion.

	App Name	Downloads (in Billions)	level
0	Subway Surfers	1	3
1	Hill Climb Racing	0.5	3
2	Temple Run 2	0.5	3
3	My Talking Tom	0.5	3
4	Pou	0.5	3
5	Candy Crush Saga	0.5	4
6	Clash of Clans	0.5	4
7	Pet Rescue Saga	0.1	4
8	Farm Heroes Saga	0.1	4
9	Candy Crush Soda Saga	0.1	4

Table 5: Combined top 5 games of level 3 and 4

The games in the Google Play Store play a significant role. As shown in Figure 7a, it attributes to 26.7% of the total downloads. Furthermore, the average downloads of games

is 98.6% more than any other apps with an average close to 0.5 million downloads (see Table 4). In level 3 and 4, the average downloads jumps to over 1 million average downloads (see Figure 8). Given these observations, it can be concluded that there is enough evidence to support h1d (see Section 2.1.1).

### 4.1.3 Sample and premium Apps

Levels 2 (premium) and 5 have the least amount of total and average downloads, as shown in Figure 2. Furthermore, we also see a high disparity in average downloads between level 2 sample and premium, as shown in Figure 2b.

As shown in Figure 9 , just 1.1% of the downloads are attributed to premium apps within all apps in level 2. Illustrating in depth just how high this offset is.

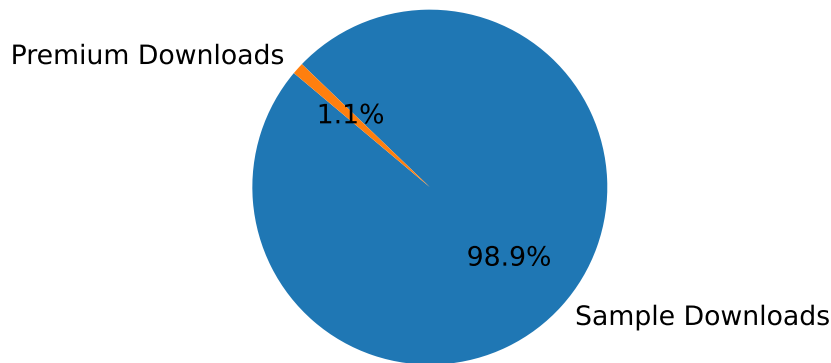


Figure 9: Proportion of Total Downloads: Sample vs Premium

High disparity is also evident in the average downloads, as shown in Table 6. Where the average download difference is close to 500.000. With an average download ratio of nearly 27%. Meaning that about one in fourth users that download the sample app, also download the premium app.

	Metric	Value
0	Average Sample Downloads	480358
1	Average Premium Downloads	5213.22
2	Average Download Difference	-475145
3	Average Download Ratio	0.265126

Table 6: Sample and Premium download metrics

Hypothesis 1c claims that the sample version indeed has more downloads than their paid-for counterpart. This claim is supported by examining Figure 2 and Figure 2b. However, for a more detailed perspective like the proportion of the sample vs premium downloads (see



Figure 9), this further strengthens this argument. Additionally, the high disparity across metrics in Table 6 is even more evidence to support h1c (see Section 2.1.1).

## 4.2 Ratings

The average and Bayesian average (adjusting for total reviews) are relatively the same across all revenue levels, as shown in Figure 10.

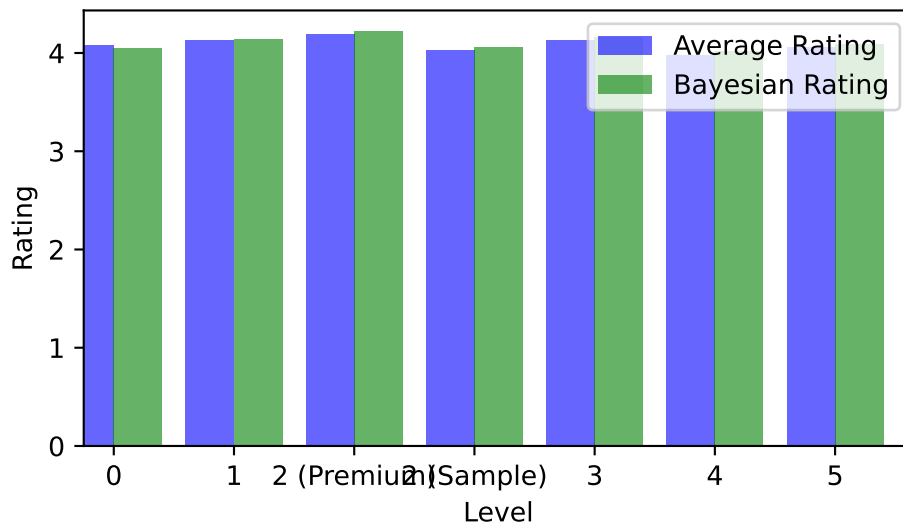


Figure 10: Proportion of Ratings Across Levels

The variance of the ratings across the different levels is where we do see lot of variance, as illustrated in Figure 11. However, when adjusting for total reviews: the Bayesian average rating, the variance diminishes. This is logical, since it smooths out the fluctuations and provides a more consistent view of the ratings. This shows that levels especially the paid apps (level 2 premium and 5) have strong outliers.

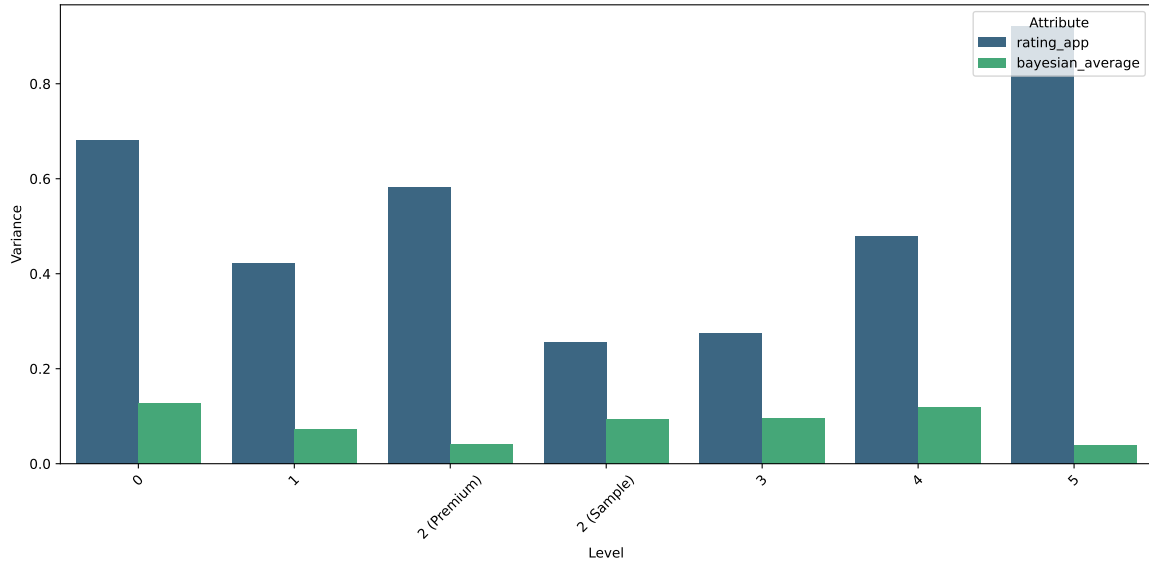


Figure 11: Variance of Different Attributes Grouped by Level

In the next two subsections we investigate further on the variance and the difference in rating between paid and free apps.

In the next following three subsections, we will investigate further into:

- Quality of Free and Paid Apps: Fluctuations in ratings caused by the quality.
- Variance in Ratings: Less fluctuation in premium apps than their free counterpart.
- Correlation between Sample and Premium Apps: Correlation in terms of rating between two app versions.

#### 4.2.1 Quality of Free and Paid Apps

Level 0 and 1 have the highest overall downloads, as shown in Figure 3 by quite a margin (40 billion downloads (0.4e11)) compared to the other apps. These apps under level 0 and 1 have free access to all features. Hypothesis 1a (see Section 2.1.1) suggests that the variance in rating between “free access to all feature apps” and “free access to not all features or paid apps” is a sign of quality disparity.

Table 7 shows the variance and standard deviation between “free access to all feature apps” and “free access to not all features or paid apps”. From which we can conclude that there is not a big statistical difference. Roughly, 0.02 and 0.01 difference in variance and standard deviation rating respectively.

category		variance in rating	standard deviation in rating
0	Free all feature apps	0.531282	0.728891
1	Free/Paid not all feature apps	0.511318	0.715065

category	variance in rating	standard deviation in rating
----------	--------------------	------------------------------

Table 7: Variance (and std) difference. Between free access to all feature apps and free access to not all features or paid apps

Figure 12 visualizes variance with outliers. It shows that both categories have a similar median rating close to 4.5, with “Free/Paid not all feature apps” being slightly lower. Both categories seem to have a lot of outliers under 3, with ‘free all feature apps’ having more outliers in 2 and 1.

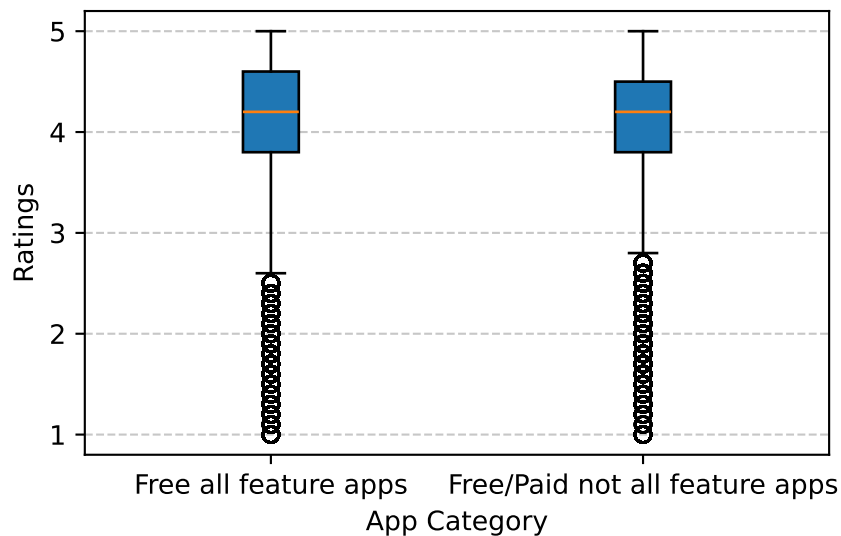


Figure 12: Boxplot of Ratings by App Category

The variance and standard deviation is minimal between the two categories, as shown in Table 7. Figure 12 illustrates that both categories receive high ratings, with significant number of outliers. These outliers can be interpreted as variability in user satisfaction. The slightly higher median in ‘Free all feature apps’ may imply that users prefer this over ‘Free/Paid not all feature apps’. However, this is marginal and would probably require statistical validation. Thus inconclusive as evidence to support h1a (see Section 2.1.1).

Conversely, hypothesis 2a (see Section 2.1.2) suggests that the apps where users need to pay to unlock features tend to have a lower rating than apps that require payment upfront.

As shown in Figure 13, premium apps have a higher rating on average compared to free apps with users needing to pay to unlock features. The free apps show a greater spread in ratings.

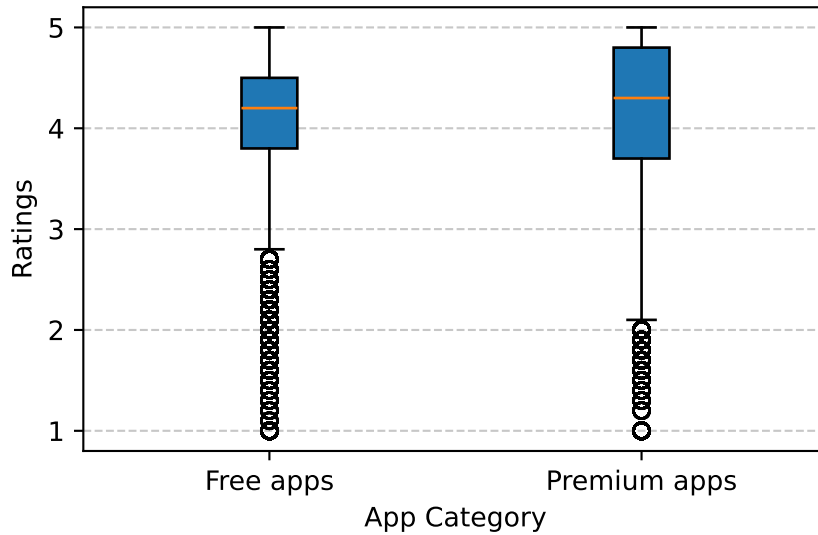


Figure 13: Boxplot of Ratings for Free and Premium Apps

The apps that require upfront payment receive a higher rating on average than free apps, shown in Figure 13. This can indicate that users perceive as being of higher quality. The greater spread of ratings for free apps, may indicate that the app quality varies more. This suffices as enough evidence to support h2a (see Section 2.1.2).

#### 4.2.2 Variance in Ratings

Figure 14 illustrates a significant difference in the variance of ratings. Premium apps tend to have a higher variance, with more outliers.

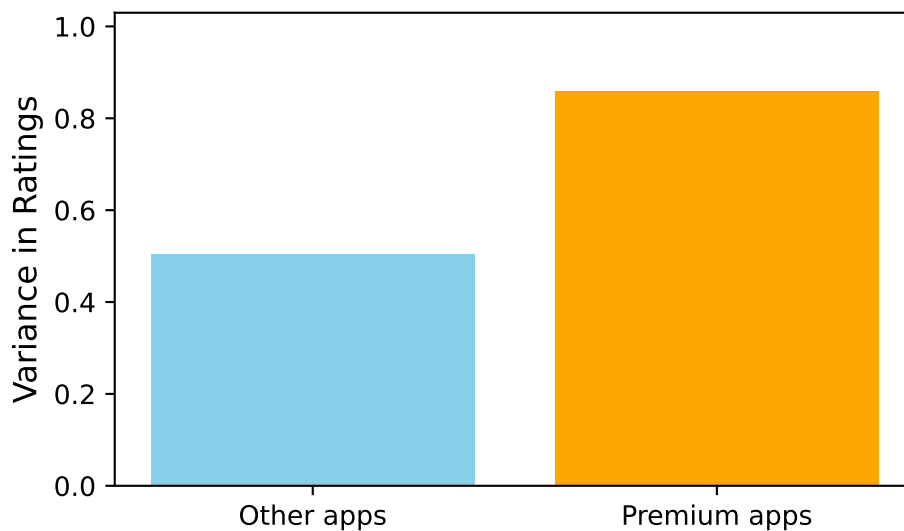


Figure 14: Variance in Ratings for Other vs Premium Apps

Figure 15 illustrates the variance of the ratings and bayesian variance. For level 5 we see

that the variance is one of the lowest when adjusted by Bayesian average. This difference may arise because users who are satisfied with the app tend to rate the app highly, while users who are not satisfied, tend to rate it much lower, having paid for the app.

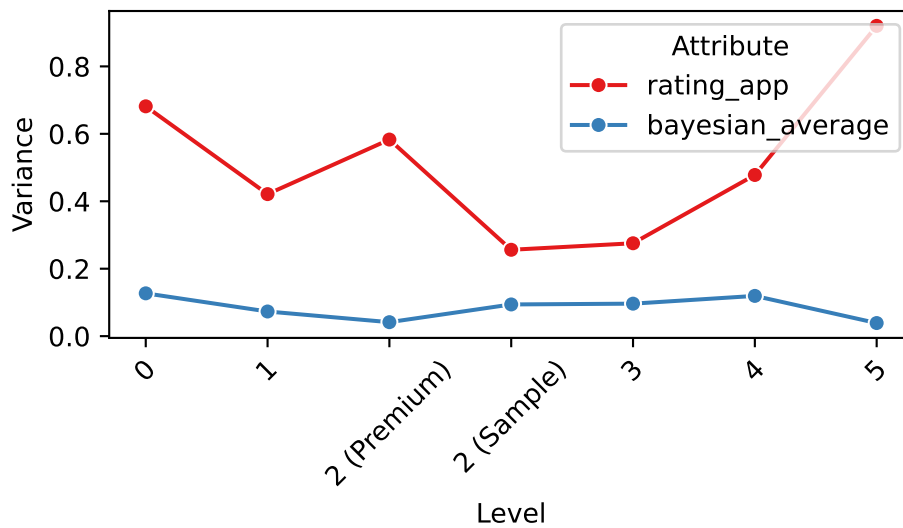


Figure 15: Variance of Attributes Across Different Levels

In conclusion, the variance across all levels show that level 5 has the highest variance for rating and the lowest variance for bayesian rating, as shown in Figure 15. The bayesian average seems more reliable regarding the overall app performance. The premium apps may have polarized feedback due to high user expectations. In that case it's enough evidence to support h2b (see Section 2.1.2).

#### 4.2.3 Correlation between Sample and Premium Apps

Revenue level 2 has two versions of the same app. A sample and a premium version.

Table 8 shows that the correlation between ratings of sample and premium apps is considered moderate positive with 0.35. The correlation between the Bayesian average rating of sample and premium apps is considered moderate to strong positive correlation with 0.5.

Metric		Correlation Coefficient
0	Sample Rating vs Premium Rating	0.352868
1	Sample Bayesian Average vs Premium Bayesian Average	0.505431

Table 8: Correlation ranking between sample and premium

Table 8 shows moderate to strong positive correlation when using the bayesian average rating. This provides evidence to support H2c (see Section 2.1.2).

### 4.3 Conclusion

Hypothesis	Support	Sufficient evidence?
H1a	Section 4.2.1	Inconclusive
H1b	Section 4.1.1	Yes
H1c	Section 4.1.3	Yes
H1d	Section 4.1.2	Yes
H2a	Section 4.2.1	Yes
H2b	Section 2.1.2	No
H2c	Section 4.2.3	Yes

The research question: “How are the 6 different revenue models as proposed by Djaruma et al. (2023) correlate to the success of an app?”. This success was measured by popularity and rating.

HOE MOETEN WE DE RESEARCH QUESTION BEANTWOORDEN???

## 5 Discussion

### 5.1 Reflection on the Findings

Downloads do not necessarily indicate revenue for freemium models (Djaruma et al. 2023). The time the user spends on an app and the purchases made within this app (Ross 2018) are better measures of the revenue for freemium applications.

### 5.2 Practical Implications for Businesses

### 5.3 Future Research Directions

## 6 References

- Aydin Gokgoz, Zeynep, M. Berk Ataman, and Gerrit H. van Bruggen. 2021. “There’s an App for That! Understanding the Drivers of Mobile Application Downloads.” *Journal of Business Research* 123 (February): 423–37. <https://doi.org/10.1016/j.jbusres.2020.10.006>.
- Bamberger, Kenneth A., Serge Egelman, Catherine Han, Amit Elazari Bar On, and Irwin Reyes. 2020. “Can You Pay for Privacy? Consumer Expectations and the Behavior of Free and Paid Apps.” *Berkeley Technology Law Journal* 35: 327. <https://heinonline.org/HOL/Page?handle=hein.journals/berktech35&id=339&div=&collection=>.
- Djaruma, Heryan, Zaki Widyadhana Wirawan, Alexander Agung Santoso Gunawan, and Karen Etania Saputra. 2023. “2023 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT).” In, 577–83. <https://doi.org/10.1109/COMNETSAT59769.2023.10420748>.

- Kumar, Vineet. 2014. "Making "Freemium" Work." *Harvard Business Review*, May. <https://hbr.org/2014/05/making-freemium-work>.
- Liu, Charles Zhechao, Yoris A. Au, and Hoon Seok Choi. 2012. "An Empirical Study of the Freemium Strategy for Mobile Apps: Evidence from the Google Play Market." In *Proceedings of the 2012 International Conference on Information Systems (ICIS)*, 1–19. <http://aisel.aisnet.org:80/cgi/viewcontent.cgi?article=1050&context=icis2012>.
- . 2014. "Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play." *Journal of Management Information Systems* 31 (3): 326–54. <https://doi.org/10.1080/07421222.2014.995564>.
- Mileros, Martin D., and Robert Forchheimer. 2024. "Free for You and Me? Exploring the Value Users Gain from Their Seemingly Free Apps." *Digital Policy, Regulation and Governance* ahead-of-print (ahead-of-print). <https://doi.org/10.1108/DPRG-01-2024-0009>.
- "Mobile App Revenue Worldwide by Segment (2019-2027)." 2023. Statista. <https://www.statista.com/forecasts/1262892/mobile-app-revenue-worldwide-by-segment>.
- Nieborg, David B. 2016. "Free-to-Play Games and App Advertising: The Rise of the Player Commodity." In. Routledge.
- Richter, Felix. 2023. "Charted: There Are More Mobile Phones Than People in the World." World Economic Forum. <https://www.weforum.org/stories/2023/04/charted-there-are-more-phones-than-people-in-the-world/>.
- Roma, Paolo, and Daniele Ragaglia. 2016. "Revenue Models, in-App Purchase, and the App Performance: Evidence from Apple's App Store and Google Play." *Electronic Commerce Research and Applications* 17 (May): 173–90. <https://doi.org/10.1016/j.elerap.2016.04.007>.
- Ross, Nicholas. 2018. "Customer Retention in Freemium Applications." *Journal of Marketing Analytics* 6 (4): 127–37. <https://doi.org/10.1057/s41270-018-0042-x>.
- Salehudin, Imam, and Frank Alpert. 2021. "No Such Thing As A Free App: A Taxonomy of Freemium Business Models and User Archetypes in the Mobile Games Market," December. <https://papers.ssrn.com/abstract=4001100>.