CS5346 2023-24 OTOT Tasks B1

Zhao Shanhe*

March 28th, 2024

1   Introduction

Data Name	Description	Data Type
App	App name	Norminal
Category	Category the app belongs to	Discrete
Rating	Overall user rating of the app	Ordinal
Reviews	Number of user reviews for the app	Continuous
Size	Size of the app	Textual
Installs	Number of user downloads /installs for the app	Discrete
Type	Paid or Free	Discrete
Price	Price of the app	Textual
Content Rating	The age group the app is targeted at - Children / Mature 21+ / Adult	Discrete
Genres	An app can belong to multiple genres (apart from its main category). For eg, a musical family game will belong to	Discrete
Last Updated	Date when the app was last updated on Play Store	Temporal
Current Ver	The current version of the app available on Play Store	Textual
Android Ver	Min required Android version	Textual

Data source: https://www.kaggle.com/datasets/lava18/google-play-store-apps
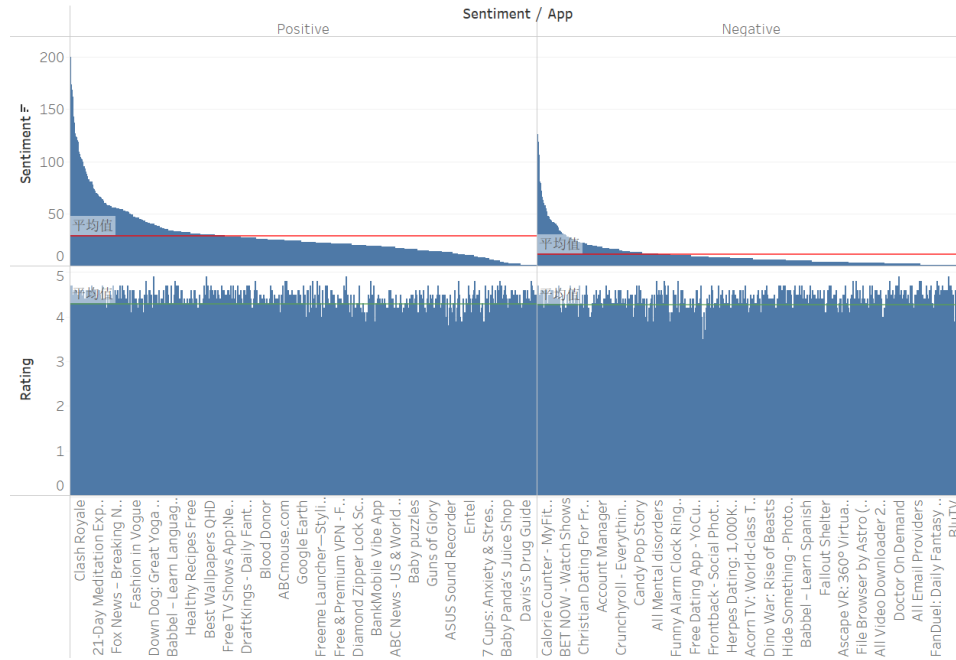
---

*Student ID: A0287084U

This task is done through **_Tableau_**.

The purpose of this visualization task is to find out which types of software are more popular in the market and which ones are more recognized by looking at App downloads, ratings, etc.

Task on Github: https://github.com/Edward-EH-Holmes/CS5346-2023-24-Own-Time-Own-Target-OTOT-Tasks/tree/main/Task%20B/B1

2 Visualization

Figure 1: Sentiment-Rating Relationship in different App (Bar chart)

If you want to see the bar chart (Figure 1) more clearly, take a look at this chart.

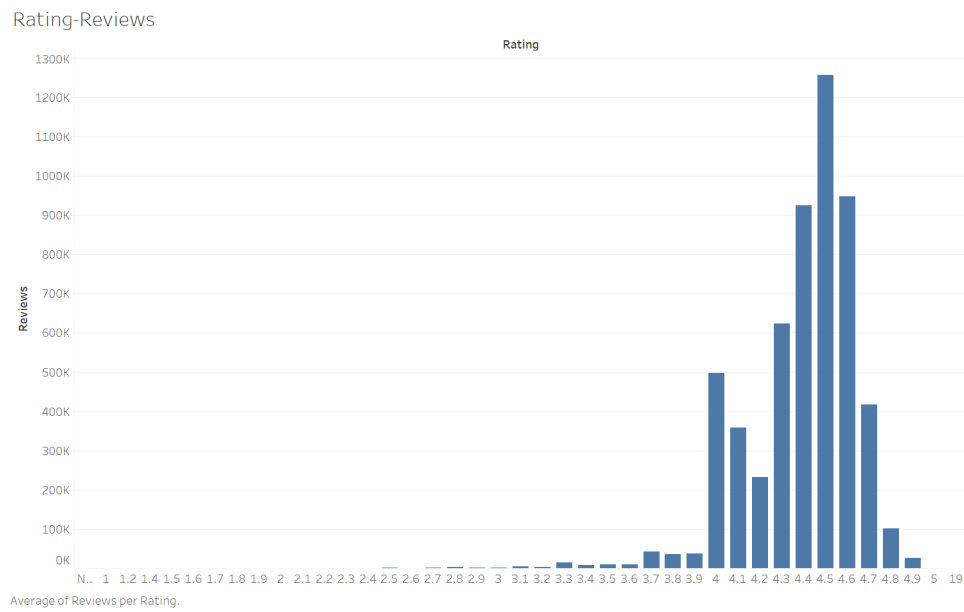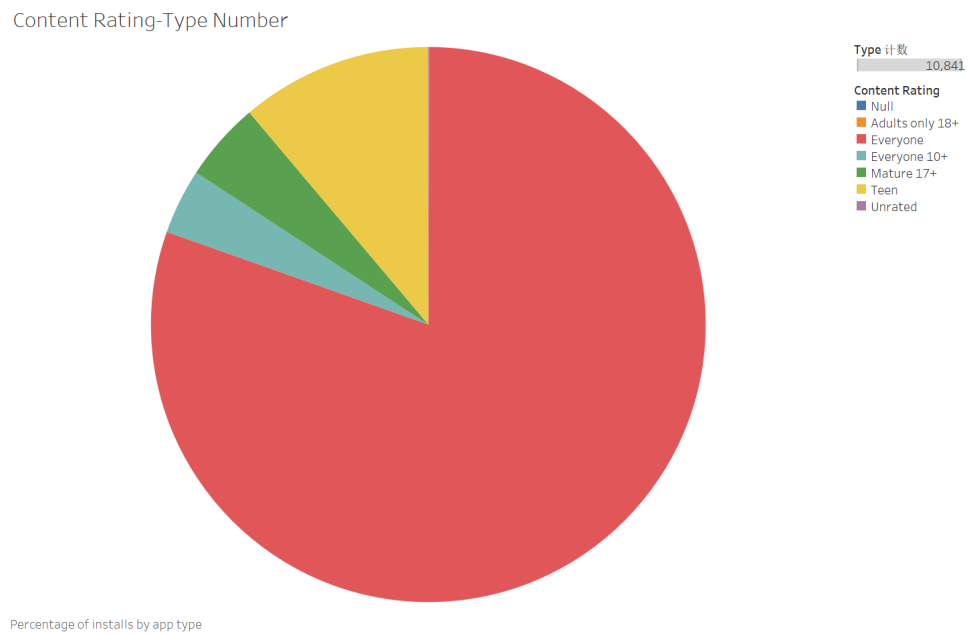Figure 2: Average of Reviews for per Rating (Bar chart)

Content Rating-Type Number

Type 计数
10,841

Content Rating
■ Null
■ Adults only 18+
■ Everyone
■ Everyone 10+
■ Mature 17+
■ Teen
■ Unrated

Percentage of installs by app type

Figure 3: Percentage of installs by app content rating (Pie chart)

Rating of software with installed volume of 1,000,000,000+
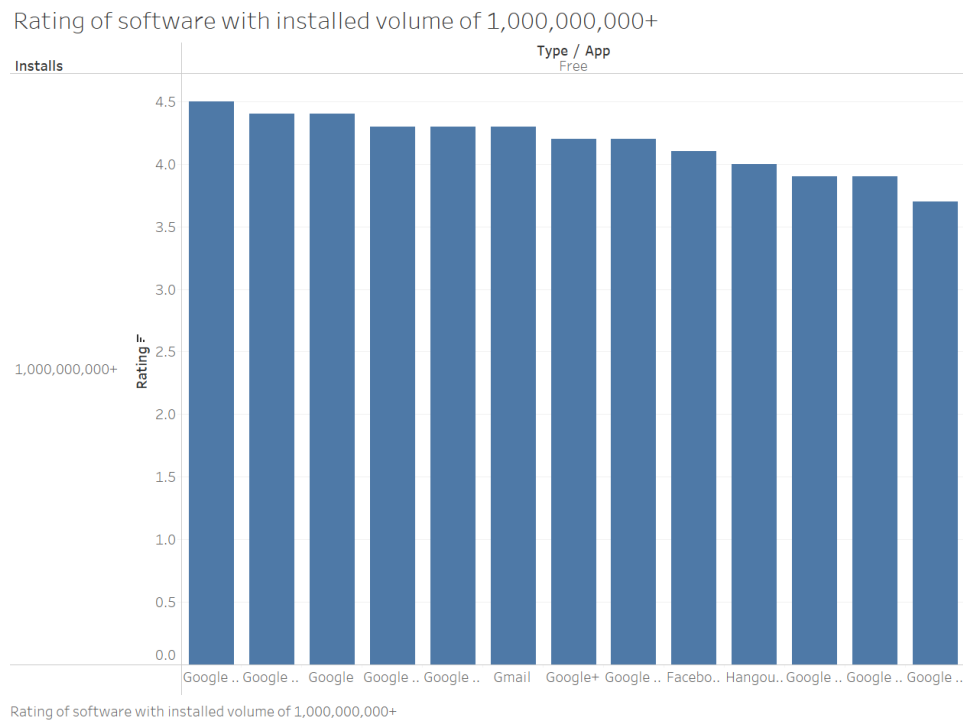

Figure 4: Rating of software with installed volume of 1,000,000,000+ (Bar chart)

Rating of software with installed volume of 500,000,000+
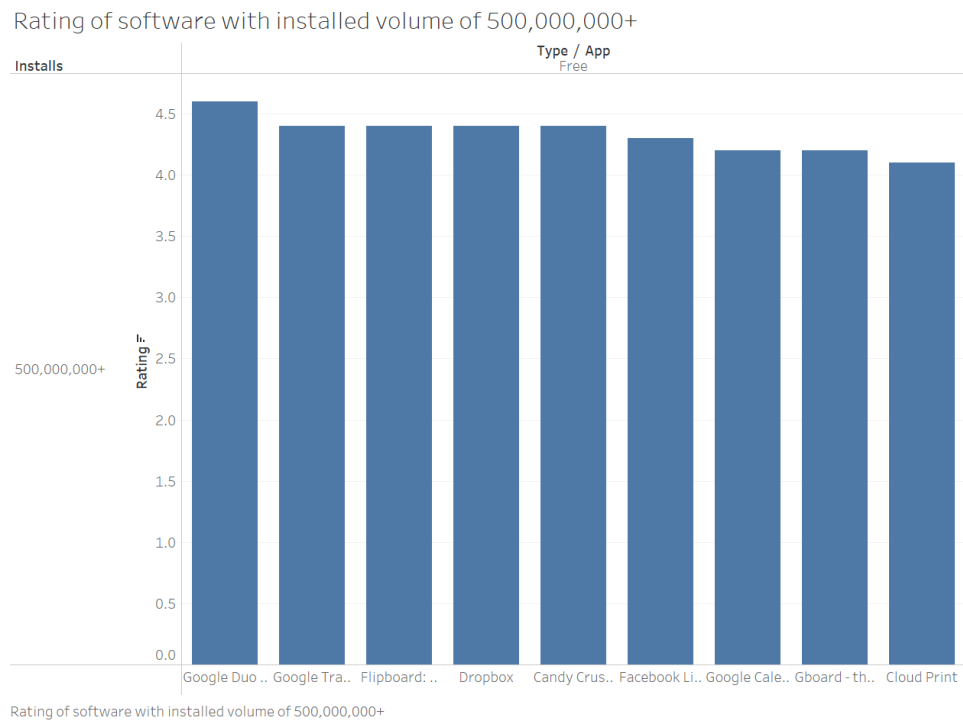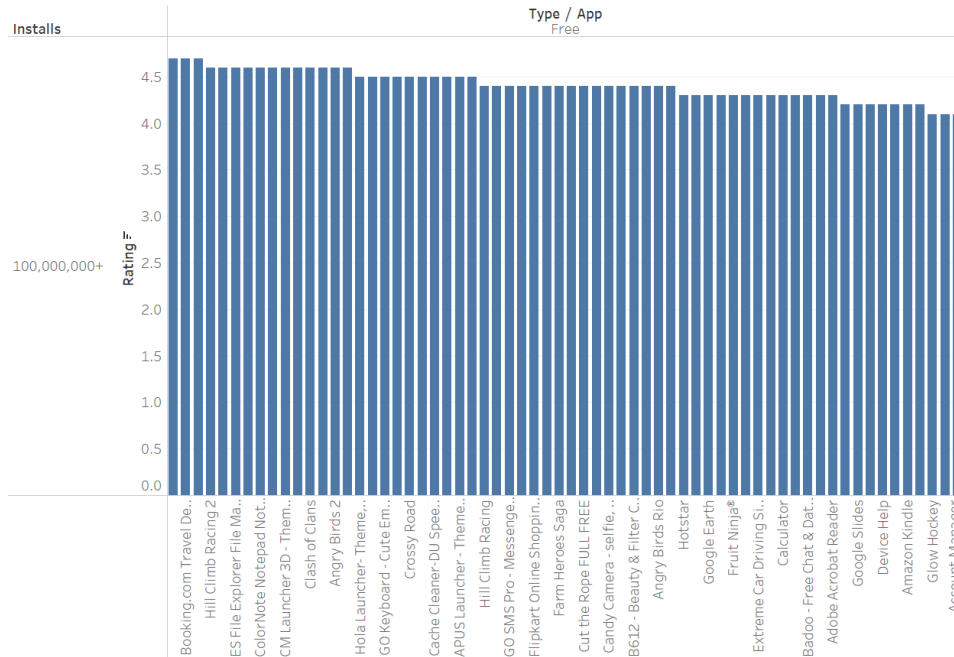

Figure 5: Rating of software with installed volume of 500,000,000+ (Bar chart)

Rating of software with installed volume of 100,000,000+

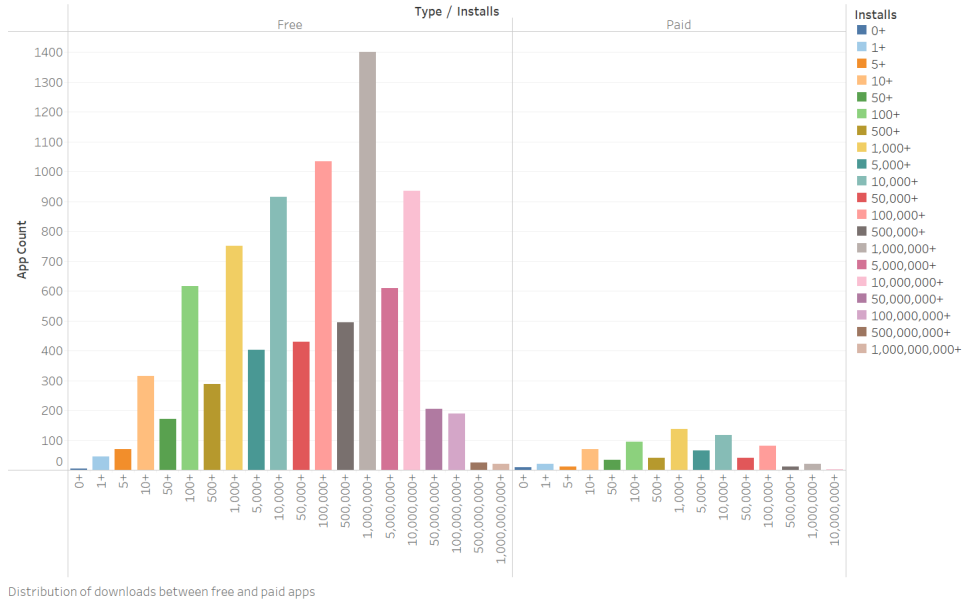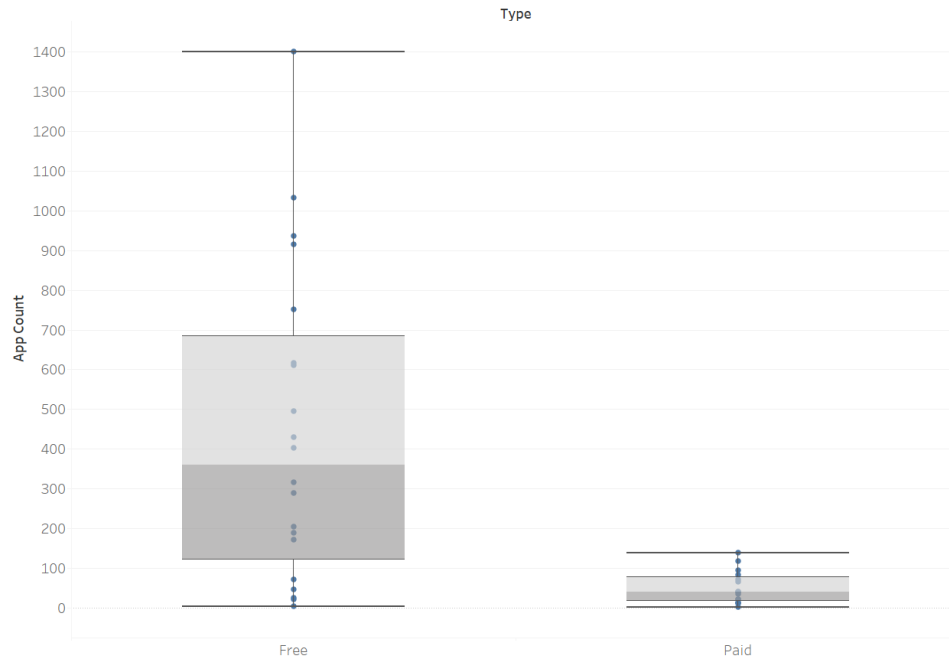Figure 6: Rating of software with installed volume of 100,000,000+ (Bar chart)

Figure 7: Distribution of downloads between free and paid apps (Bar chart)

Distribution of downloads between free and paid apps

Figure 8: Distribution of downloads between free and paid apps (Box plot)

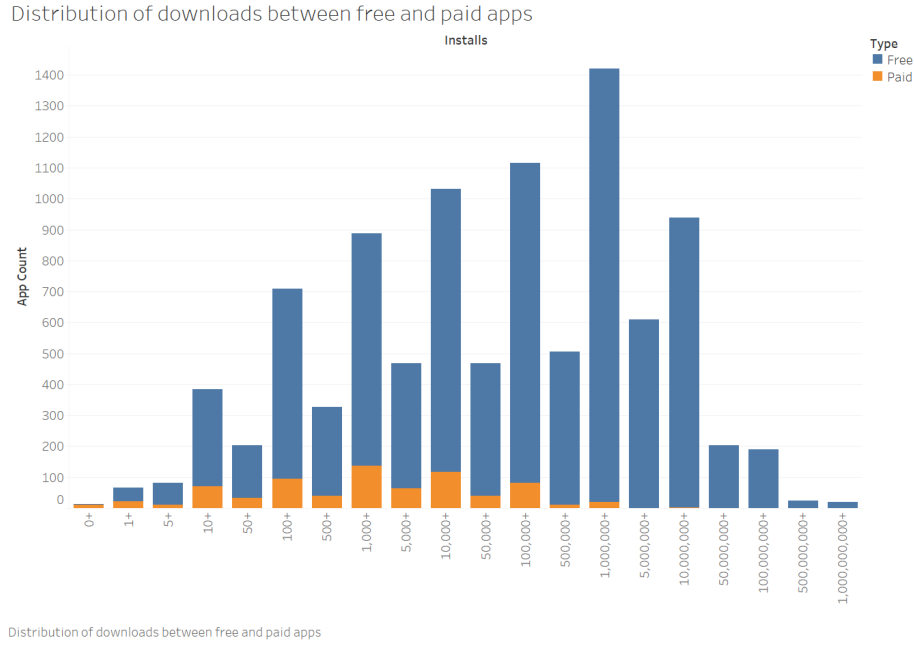Distribution of downloads between free and paid apps

Figure 9: Distribution of downloads between free and paid apps (Stacked Bar chart)

3 Conclusion

From this, we can draw the following conclusions:

- As can be seen from Figure 1, with the decline of sentiment (positive or negative), the rating of App did not show correlation (on the contrary, it showed chaotic and disorderly fluctuations). Therefore, sentiment is not related to rating.

- According to Figure 2, the number of reviews reaches its peak when the score is 4.5. Thus, the statement "the higher the rating of a software, the higher the heat of discussion" is a false proposition.

- According to Figure 3, apps for everyone are more popular with users, and App developers are equally willing to develop apps for everyone.

- According to Figure 4, 5 and 6

1. The number of App installs is not significantly related to the rating (for example, there is no significant difference between 1,000,000,000+ App installs, 500,000,000+ App installs, and 100,000,000+ App installs).

2. According to the number of apps installed, the series of software produced by Google has relatively high user engagement, almost topping the list of 1,000,000,000+ installs and 500,000,000+ installs.

3. Almost all apps with more than 100,000,000 installs have a rating above 4.0 (except Google Play Books, Google News, and Google Play Movies & TV).

- According to Figure 7, 8 and 9

  1. The number of apps with 1,000,000+ installs is the largest, while the number of apps with 500,000,000+ installs and 1,000,000,000+ installs is relatively small. It can be seen that under the choice of the market, there are not many apps that can become explosive. In addition, if an App wants to become a popular one, it must consider the user stickiness under the App ecosystem (according to the previous conclusion, we can know that Google's App has a high user stickiness).

  2. The number of paid apps is much lower than the number of free apps. Therefore, both in terms of the development willingness of App developers and the choice willingness of App users, free apps are more popular among Android users and developers.