A Comparison of Major Android Mobile App Stores Using Metadata

# Authors (listed by alphabetical order)

Siyu Chen (Shanghai Experimental High School, Shanghai, China, [charleshx@163.com](mailto:charleshx@163.com))

Yuxuan Chen (ULink Cambridge School, Shanghai, China, [274380981@qq.com](mailto:274380981@qq.com))

Ziyi Huang (Shanghai Experimental High School, Shanghai, China, [963352472@qq.com](mailto:963352472@qq.com))

Ruijun Sun (Chuansha High School, Shanghai, China, [2731045743@qq.com](mailto:2731045743@qq.com))

Yuxiang Zhou (Shanghai Experimental High School, Shanghai, China, [952486298@qq.com](mailto:952486298@qq.com))

Xinyu Zhu (Shanghai Experimental High School, Shanghai, China, [zyjen163@163.com](mailto:zyjen163@163.com))

# Faculty Advisor

Junshen Kevin Chen, Computer Science, Stanford University, jkc1@stanford.edu

# Summary

<Provide a succinct paragraph describing your project (a maximum of 100 words) that can be used for promotion at the conference.>

# Problem and Motivation

With the ever-increasing availability of mobile electronic devices, the use of mobile apps has become an essential part of people's life. Users download applications from app stores and often rate them based on their usability and functionality. As daily users of mobile apps, we have always been interested in the overall trend of mobile apps. We believe that these mobile applications have fundamentally changed how our generation interacts with the world. In particular, we are interested in any similarity or difference of this trend across app stores of different regions.

To achieve our research goal, we need an update-to-date comprehensive mobile app data. Therefore, we decided to crawl the available meta-data from major app stores. Apple app store has a uniformed interface across the globe and has a sophisticated mechanism to prevent large-scale crawling. So we decided not to include it in this study. Instead, we focused on two major Android app stores: Tencent app store from mainland China and Google Play store in the US region. Over a two-month period, we crawled the meta-data from both app stores: 76,838 apps from Tencent, and 643,220 apps from Google Play. The datasets include the name, category, rating, size, release time, download times, and other related information.

We focus on the following questions in this study: 1) What are the major characteristics of Tencent and Google Play app stores with respect to rating and category distributions? 2) For the apps that appear in both stores, what are the major similarity and differences? and 3) What is the relationship between the number of ratings (or the average rating) and the quality of an app? Is there any strong link between them?

# Approach

The first step in our project was to obtain mobile app metadata. We searched Kaggle.com and UCI machine learning repository but were only able to locate one 2015 Google play dataset and one 2017 Apple store dataset. After examining these two datasets, we cannot verify the quality of the two datasets and due to their age difference, we cannot directly compare the two datasets. Our questions on regional differences cannot be answered using these publicly available datasets either. But these two datasets allowed us to see what type of metadata we can possibly obtain and led to a few brainstorm sections. Hence we resorted to crawling our own datasets and we spent one month learning web-crawling techniques using Google Chrome package in Jupyter notebook. We first crawled a small Android app store in China, Wandoujia, but later realized its metadata is not consistent with Google Play. Therefore, we decided to crawl one of the largest Android app store, Tencent, in China and Google Play in the US because their metadata formats are nearly identical. After we completed the data crawling process in late April 2019, we brainstormed potential analytics we can perform on the datasets. We then spend the entire month of May analyzing the datasets to answer our research questions.

# Datasets

For Google Play Store (US region) and Tencent App Store (one of the largest in China), we used a similar breadth first search strategy by seeding a queue with names of popular apps, then sending search queries to each app store and record the app metadata of apps returned by the website, then appending newly seen app names to the end of this queue. The Tencent dataset is a "complete search-able set" meaning that the queue is exhausted at the end of scraping, whereas the Google Play set is a "reasonably search-able set" with the searches terminated upon reaching a computation budget.

Google app store has much more apps (642,774 in total) than that of Tencent which has 76,838 apps. Given the fact that our dataset is a subset of all possible apps from Google, the two stores' app numbers are about a ratio of 10:1. However, for each app, the two stores have similar rating structures (including the number of ratings for 1-star, 2-star until 5-star) and other similar metadata such as average rating, download counts, rating count, file size, etc. There are 4,671 apps that appeared in both stores. Another big difference between the two app stores is the number of downloads for apps. Though Tencent has much fewer distinct apps compared with google play, the number of downloads for apps that appear in both app stores have much bigger download counts than that of Google play. It indicates that common apps in both stores tend to be popular apps in mainland China. We also translated the Chinese category names at Tencent dataset into English and standardized the units for file size for both app store metadata.

# Tools and Analytics

The two tools we used in this study are primarily Jupyter Notebook and Tableau. The programming language used is Python.

When crawling the datasets, we used Selenium WebDriver to simulate a user behavior by opening an automated browser, locating corresponding buttons, and extracting necessary elements from the webpages and then storing to a data file. We studied each app store's render structure and deployed a similar method for both app stores. We distributed this crawl to multiple virtual machines on Google Clout Platform via VPN to make the process fault-tolerant and work around IP bans. We also used Pandas in Python to format the data files into CSVs afterwards.

When analyzing the data, we used Tableau Prep to preprocess the data by importing, then properly casting and filtering raw data. Then, we used Tableau to visualize the datasets. We then arrive in critical analyses and speculations from these graphs, and then finally reaching important conclusions as presented in the following section.

# Results

In this section, we present the major results of our study for the three major research questions.

### 5.1. Are the two app stores different with respect to its rating distributions and categories?

* **Rating distribution is not uniform (Fig. 1-a)**

Fig. 1-a shows that the most infrequent rating is 2-star while percentage of 4-star ratings are similar to that of 1-star. And that 5-star ratings are disproportionally higher than all the other options combined. This is true for both Google Play and Tencent. One possible reason for this phenomenon may be that users tend to rate apps only when they really like it or dislike it, resulting in the distribution polarizing on both ends. Another reason for 5-stars being unusually high may be the amount of fake rating possibly existing in the app store (see 5.3 for further investigation).

* **Category distribution is different between the two stores (Fig. 1-b)**

Google Play Store has much fewer unknown categories when compared with Tencent. The first three most frequent categories in Google Play are *Education*, *Tools*, and *Entertainment*. In Tencent, they are *Tools*, *Lifestyle*, and *Shopping*. There are 30 distinct categories at Tencent while Google Play Store has 49 distinct categories.

### 5.2. For the apps that appear in both stores, what is the major similarity and differences?

* **Sizes and average ratings are not similar across the two app stores for the same app (Fig. 2-a).**

For the collection of 4,671 apps that appear in both stores, the average file size for Google Store is 28.4M and for Tencent apps, it is 25.5M. The fit line is **GoogleFileSize = 0.23747\*TencentFileSize + 26M**. There are many outliers in the Tencent store, especially the LEGO apps, that are much bigger than its counterparts in the Google play store.

Out of all the apps that appeared in both app stores, there are 1,266 apps (27%) that have a higher average rating at Tencent than Google Play. The rest 73% of the apps boasts higher ratings at Google Play Store. Despite this difference, the comparison between the average rating in two stores displays a similar behavior of app users as they mainly give the same rating for the same apps.

* **Category assignment between the two app stores are not consistent (Fig. 2-b)**

Category assignments are not consistent in the two app stores. For example, apps that are categorized as *Entertainment* is categories into many categories at Tencent. Categories such as *Tools*, *Education*, and *Photography* are similar in both app stores and are both in the top 10 categories. However, for other categories, they have different number of records in Google and Tencent, which reveals that app users have different behavior at each store or possibly the publisher spent variant amount of money for advertising such that the demand for the products may have varied.

### 5.3.What's the relationship between the amount of ratings (or the average rating) and the quality of an app? Is there any strong link between them?

* **Download count vs average rating (Fig. 3-a)**

Apps with high download counts from Tencent are Yingyongbao (Official app store app from Tencent), QQ, Wechat (both are social media apps), and Wifi key (an app to search for free wifi-hotspots). Their ratings are not high though. For Google, Google and Google Maps have much higher download counts than other apps and their ratings are not the highest either.

* **Category vs average rating (Fig. 3-b)**

In Tencent, game categories are among the most popular whose average rating is about 3.5 stars. In Google, the most popular is the *Personalization* whose rating is about 4.4 star. It is possible to infer that users of Tencent Store people love playing games, while users of Google Play don't enjoy games as so much. Almost all the ratings of any category in Google Play is more than 4 stars, while no category of Tencent is higher than 4. The distribution of average ratings for categories is more uniform at Google than Tencent.

* **Rating count vs download count (Fig. 3-c)**

At Tencent, most apps' download count is positively correlated with rating count, but some of the Tools app have a high download count but a low rating count, possibly due to their tool-like nature of being quickly used and not given much thought, and therefore less likely to be rated. However, some of the apps such as *Fast Hand* (a video blog software) have a high rating count but a low download count, and therefore it is highly possible to have fraudulent ratings there; i.e. ratings made by / paid for by the app maker. The most obvious outlier of high downloads with few ratings in Tencent is Yingyongbao (i.e. the app store itself). The app store is the prerequisite for users to download any other app. Therefore, we speculate that the “high download count and low rating count” phenomenon is due to people only using it as a channel to downloads apps, as a mandatory measure but without given much thought.

For Google Play, the positive relationship for most apps is similar to that of Tencent. However, some outliers do exist such as Google that has a high download count but a low rating count, while Facebook and Instagram have a download count not that high but an extremely high rating count.

# Contributions and Uniqueness

<Point out the novelty of your approach and your main contributions for practitioners and/or researchers. Which key and/or surprising insight you have gained from your analytics?>

# Appendix: Data Visualizations

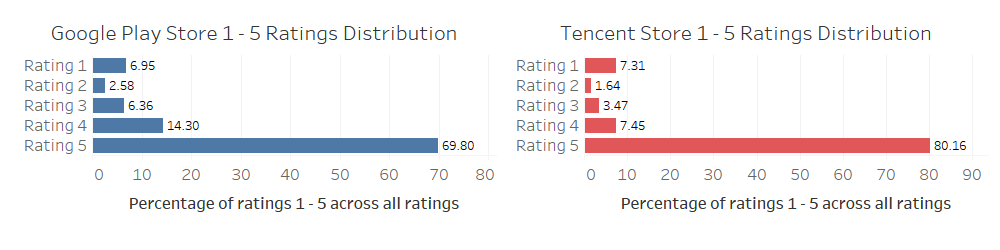
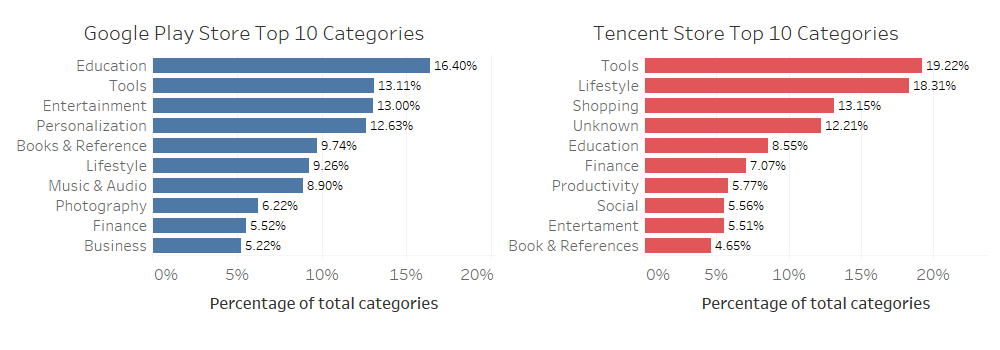


Figure 1-a: Distribution of ratings for all apps at both app stores.



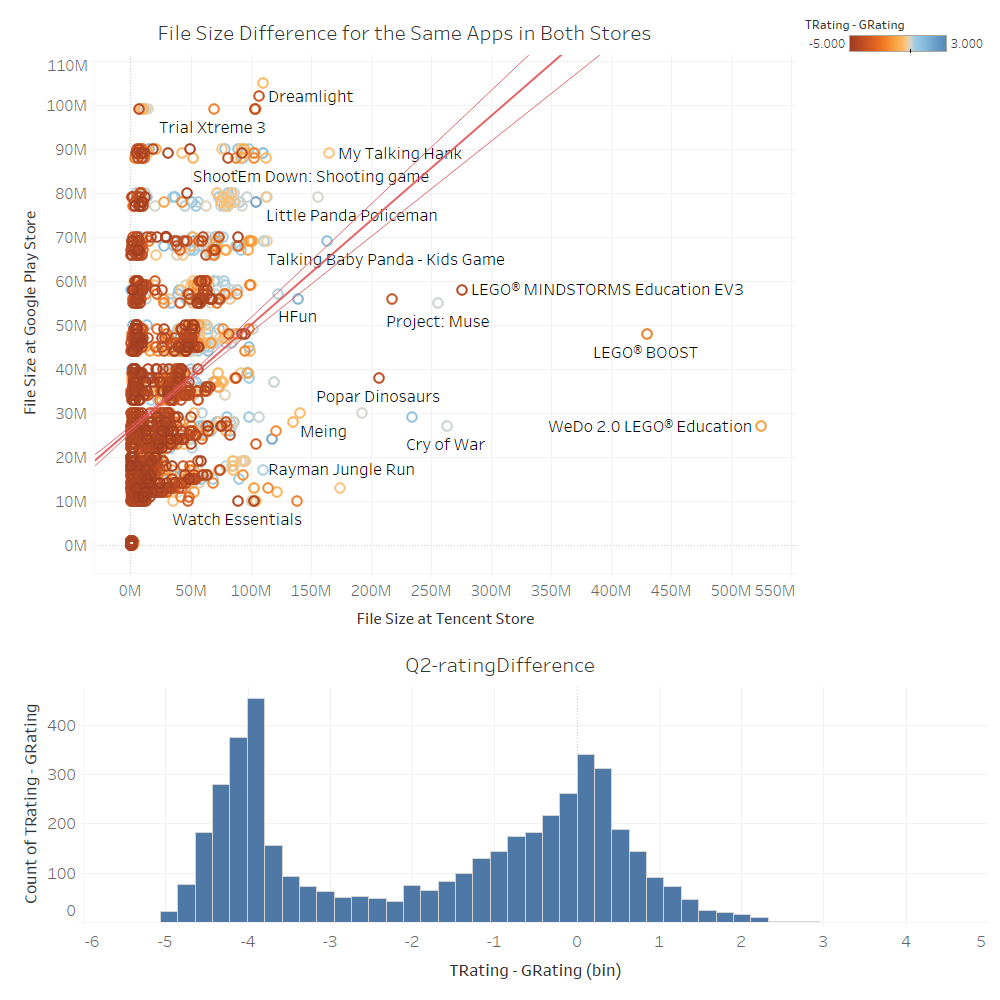
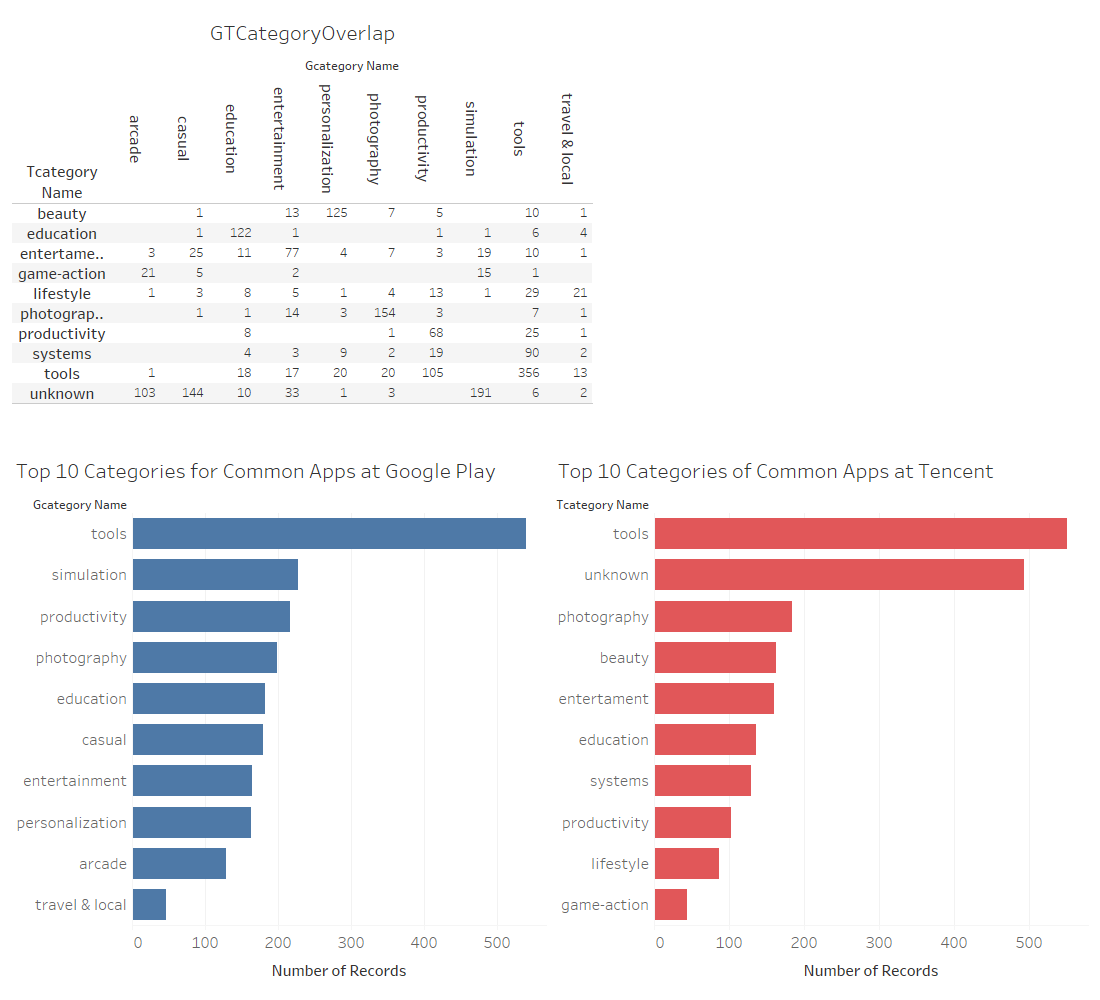
Figure 1-b: Distribution of app categories at both app stores

Figure 2-a: Scatter plot of app size comparisons between two app stores and the rating difference between Tencent and Google play is also color coded. The distribution of rating difference is shown at the bottom of the figure.



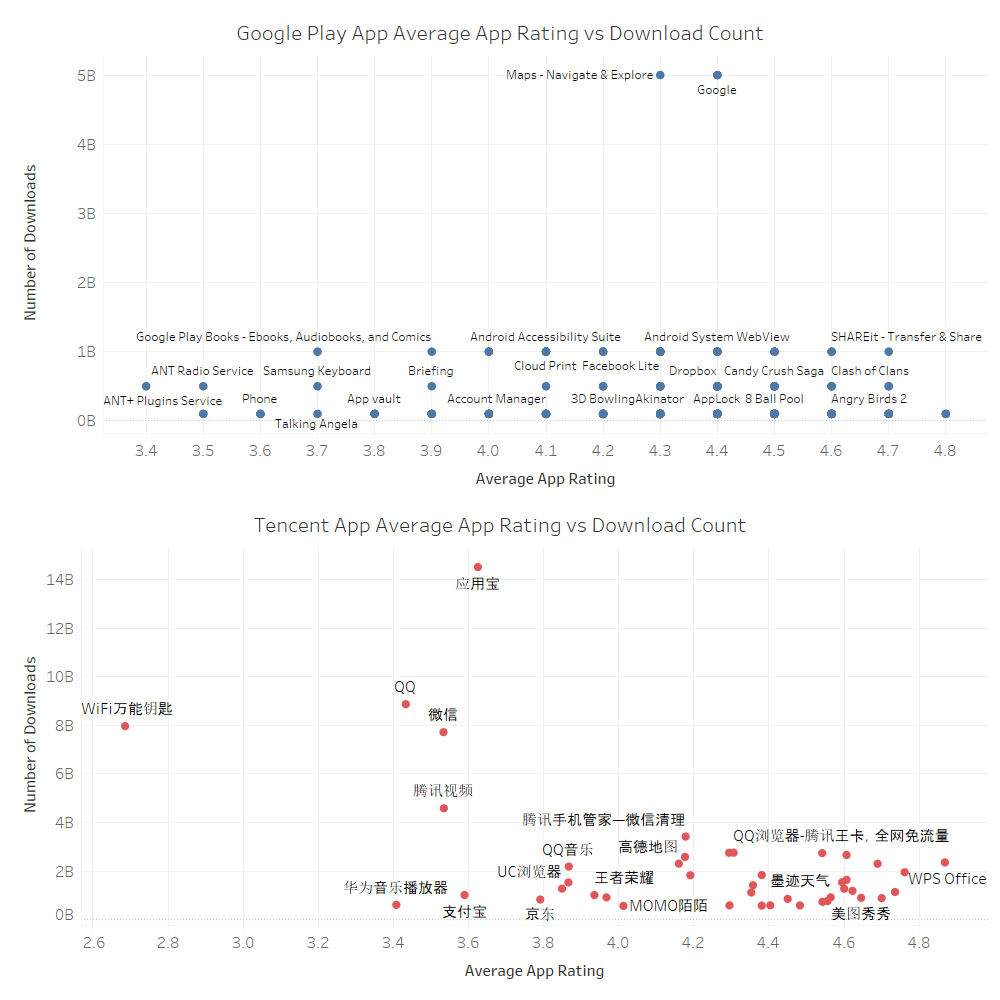
Figure 2-b: Top 10 categories of common apps in both app stores. The figure on the top shows the cross-tab category assignments in the two app stores.

Figure 3-a: Download count and average app ratings in both app stores.

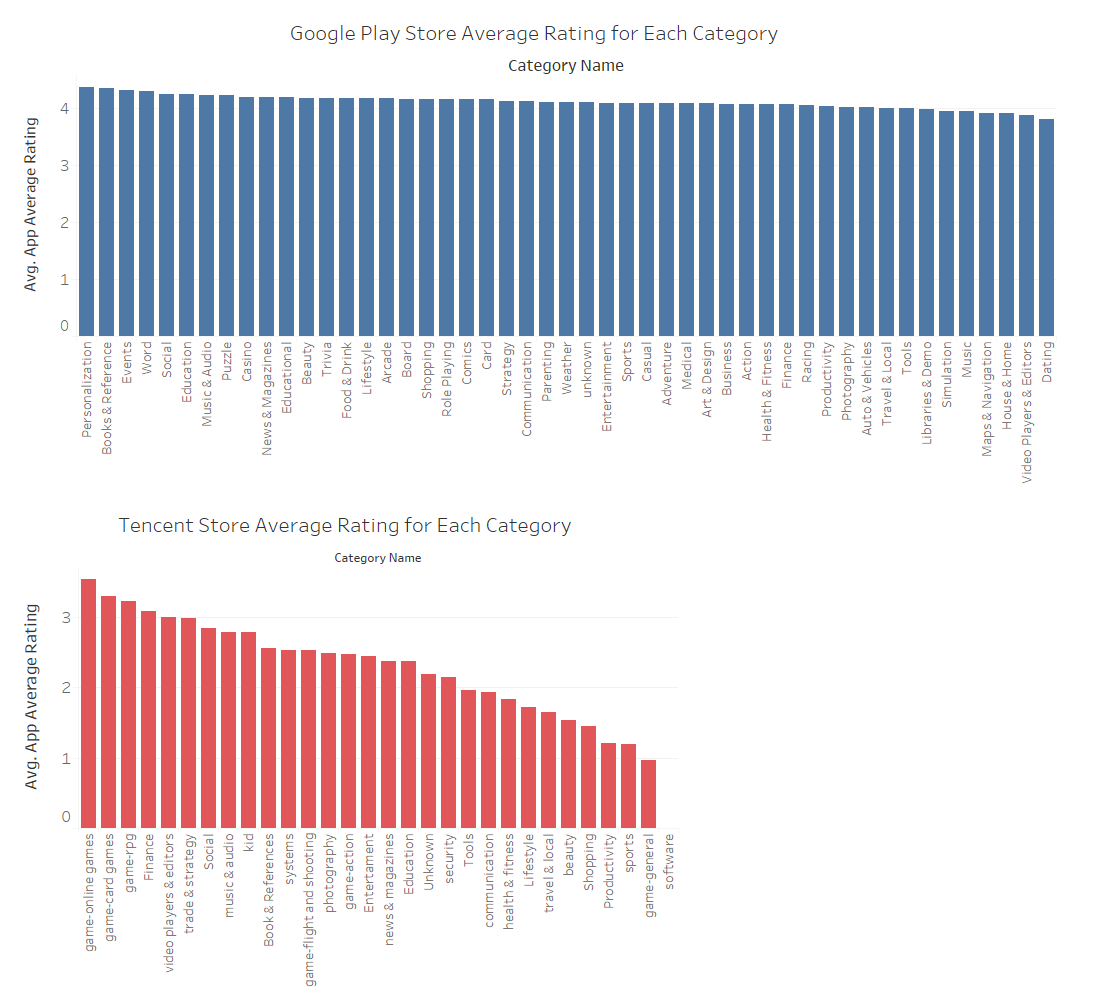


Figure 3-b. Average of ratings across all categories for both app stores.

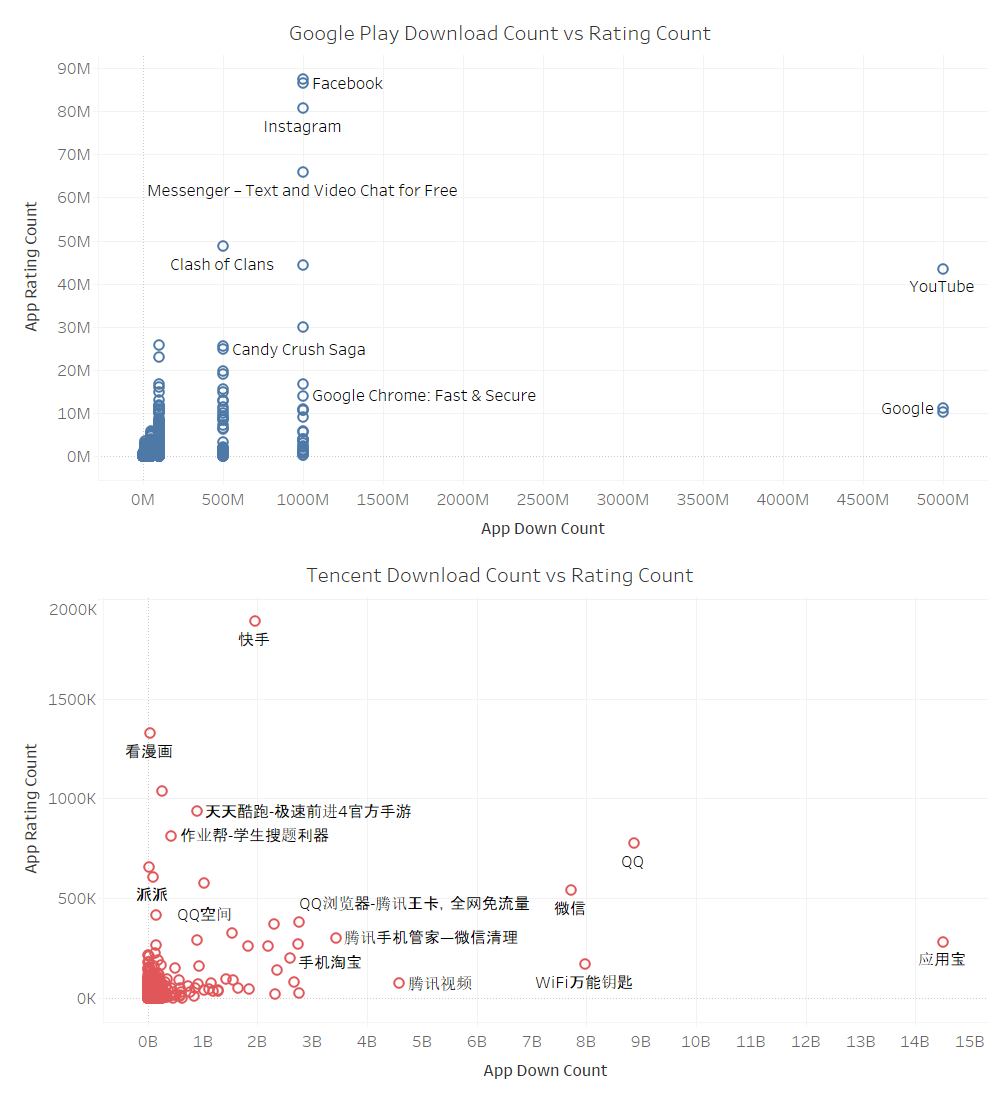


Figure 3-c: App download count vs Rating count in both app stores.

# References

<Insert your references/citations here.>