A Comparison of Major Android Mobile App Stores using Meta-data

# 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 needed 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 two months, 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 significant characteristics of Tencent and Google Play app stores, 2) What are the significant differences between Tencent and Google Play app store metadata, especially on the apps released in both regions, 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? And if so, can it be used on either of the app stores of its ratings, and 4) Is there any correlation between the size of an app and its other characteristics?

# Approach

The first step in our project was to obtain mobile app meta-data. We searched Kaggle.com and UCI machine learning dataset repository but we only able to locate one 2015 Google play dataset and 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 meta-data we can possibly obtain and led to a few brainstorm sections Hence we resorted to crawling our own dataset and we spent 2 months 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 meta-data is not consistent with Google Play. Therefore, we decided to crawl the largest Android app store, Tencent, and Google Play in the US because their meta-data formats are nearly identical. After obtaining the datasets in late April 2019, we brainstormed potential analytics we can perform on the dataset. We then spend the entire month of May analyzing the datasets to answer our research questions.

# Datasets

We scraped two up-to-date datasets of app metadata from the internet, of Google Play Store (US region) and Tencent App Store (one of the largest in China).

For both app stores, we used a similar breadth first search (BFS) 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 BFS terminated upon reaching a computation budget.

Google app store has 642,774 apps 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 (1 - 5)and other similar meta data such as average rating, download counts, rating count, and picture number. 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 in general are much bigger than that of Google play.

# 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 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 to preprocess the data by importing, then properly casting and filtering raw data. Then, we spend the majority of time exploring the datasets by creating graph visualizations of different data columns into different types of graph (such as bar chart, scatter plots, etc.), across both datasets. We then arrive in critical analyses and speculations from these graphs, and then finally reaching conclusions based on aforementioned preliminary findings.

# Results

In this section, we present the major results of our study.

**1. Are the two app stores different with respect to its rating distributions and categories? (Figure 1- a, 1-b, 1-c, 1-d in visualization appendix)**

* Rating distribution (Fig 1-a, 1-b)

As shown above, we can notice that the most infrequent rating is 2 stars, 4 stars ratings are similar to 1 star, and that 5-star ratings are disproportionally higher than all the other options combined.

One reason for this 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 subsection 3 for further investigation.

* Category distribution (Fig 1-c, 1-d)

Google Play Store has much fewer unknown categories when compared with Tencent. The first three most frequent categories in Google Play are “education”, “tool”, and “entertainment”. While in Tencent they are “tools”, “lifestyle” and “shopping”. There are 30 distinct categories at Tencent while Google Play Store has 49 distinct categories.

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

* Sizes and average ratings are different across the two app stores for the same app (Fig 2-a).

This figure shows a shared set of apps having different rating across the two stores. The average file size for Google Store is 28.4M and the average for Tencent apps is 25.5M. But according to the figure categories are not consistent across the two stores. The fit line is **GfileSize = 0.23747\*TfileSize + 26M**. Though a majority of files have similar sizes, there are many outliers in the Tencent store (LEGO apps) that are much bigger than its counterpart in the Google play store.

Out of all the apps that appeared in both app stores, there are 1266 apps (27%) that have a higher rating at Tencent than Google Play. The rest 73% of the apps boasts higher ratings at Google Play Store. As such, rating-wise, 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. This interpretation of result can indirectly demonstrate that the publishers do the same amount of work for advertising and the possibility of fraud decreases.

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

Categories such as “tools”, “education” and “photography” act similar in both app stores and are both in the top 10 categories. However, for other categories in the bar chart, 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 varied.

Also, there may occurs fraud in some apps in Tencent as they have the second most number of records in Tencent but is exclude in the top 10 categories in google.

**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? (Fig 3-a and 3-b)**

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

Apps from Tencent are “Fast hand” (a livestreaming app), “Weibo” (similar to Twitter). Their ratings are very high, nearly 5 points. As the populations of the people who rates them are equally high, so the data is believable. The rating count of the Google Play is not very high, but the quantity of download is really high, which is about 14 billion. According to the picture, usually, the rating is higher, the population of the people who download and rate it is more.

* Category vs average rating (Fig 3-b)

In Tencent, game categories are among the most popular, whose rating is about 3.5 points. In Google, the most popular is the “personalization”, whose rating is about 4.4 points. 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 points, while no category of Tencent is higher than 4. The most three high-rating categories of Tencent are all games. This is possibly due to a systematic difference between the stores and its users. From this we may form further hypotheses that prompt further study: does one app store’s design motivate users to give higher rating? Are there any cultural differences at play?

* What is the relationship between rating count and download count (Fig 3-c)

In Tencent, most of apps' download count is positively correlated with rating count, but some of the tool 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 “Well Quickly” have a high rating count bat 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.

For Google Play, most of which are similar to the app behaviors in Tencent. However, some outliers do exist such as Google that has a high download count but a low rating count, while Facebook has a download count not that high but an extremely high rating count. Finally, YouTube has both a high download count and a high rating count.

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.

As we can see, there are many strange and abnormal points both in the graph of Google and Tecent. And most of which are in two condictions, high downloads with few ratings or few downloads with high ratings.

**4. Which kind of app category appears in different groups of file sizes? (Fig 4)**

# 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 <limit to three pages>

# 

Figure 1 - a

Figure 1 - b

<description>

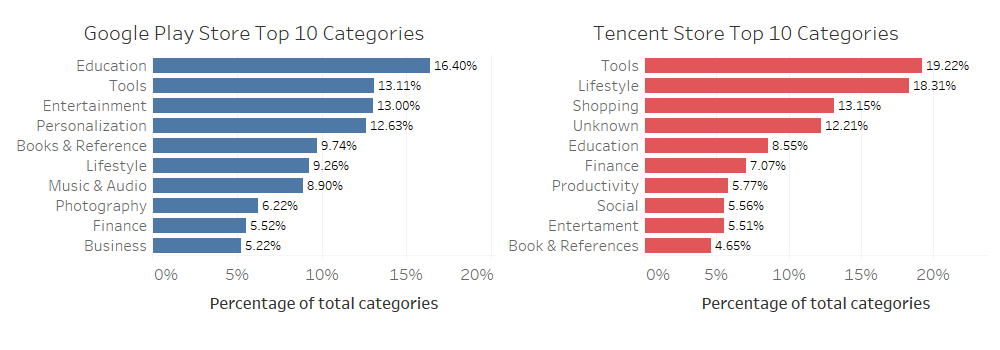
Fig 1 - c

Fig 1 - d

<description>

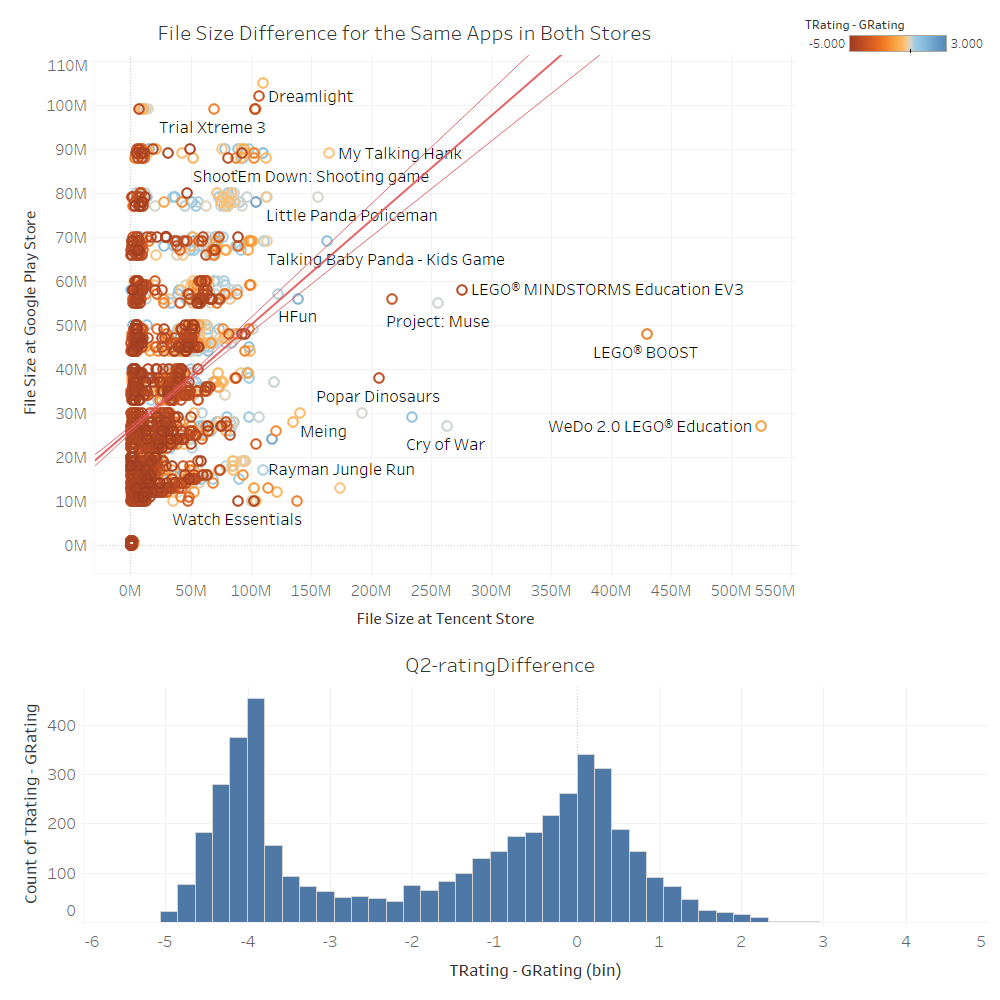


Fig 2-a

The graphs show the flie difference for same apps in Tecnet and Google store.There are three lines show the tendency of the size difference.Color is the symbol of rating.

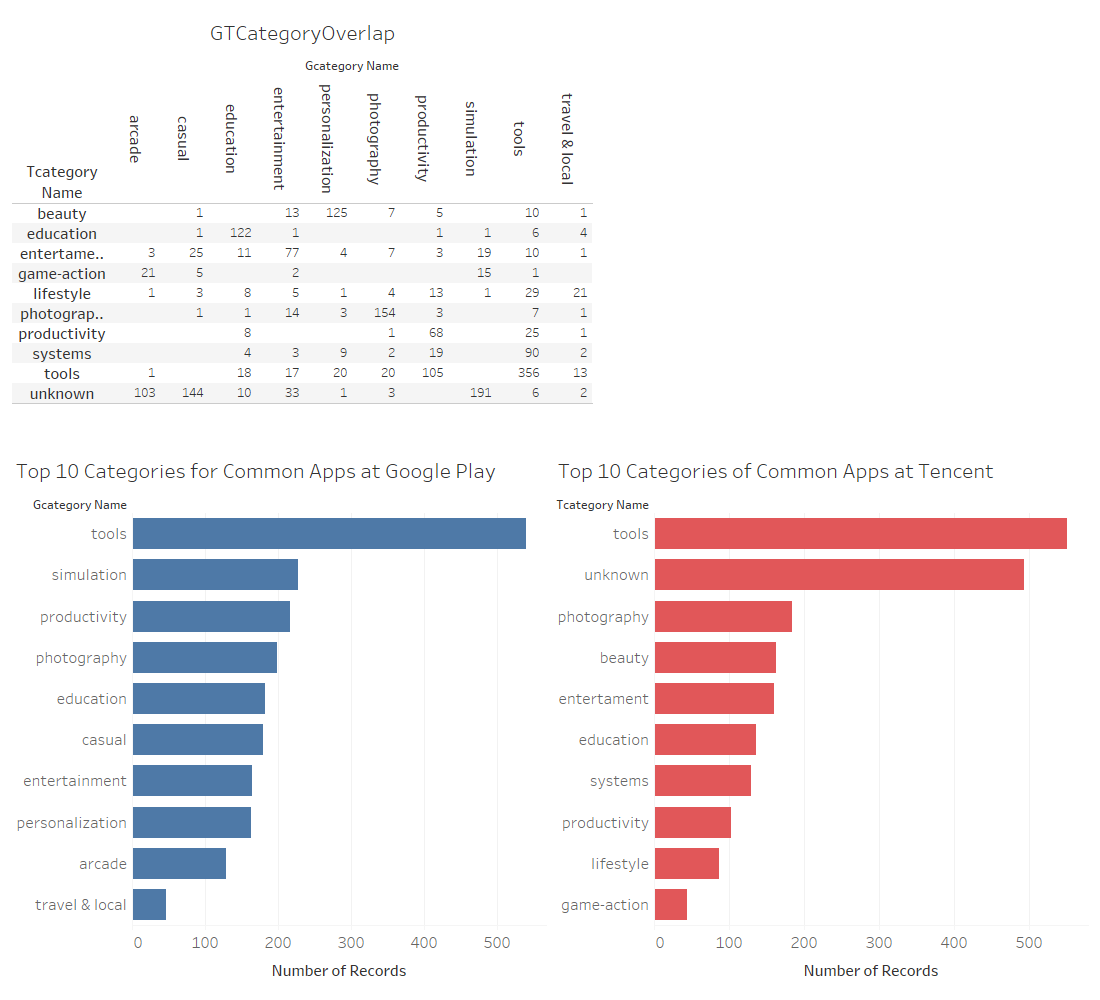


Fig 2-b

The two graphs illustrate the top 10 categories for common apps in both stores. Tools have the most number of records in common.

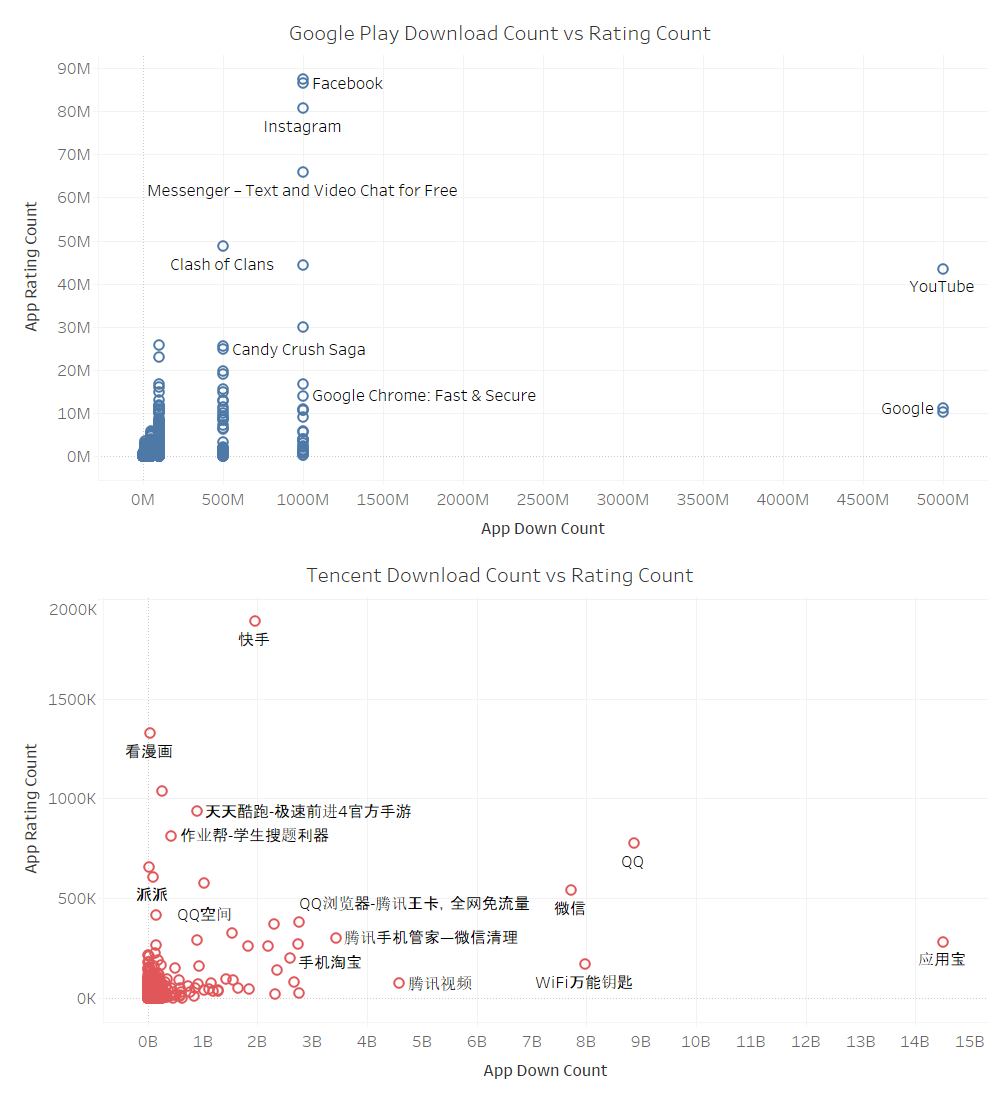


Fig 3-a

These two graphs shows the different apps' download count vs app average rating.There are four apps need to be concerned-------Google maps, YouTube, WIFI Wannengyaoshi and Yingyongbao.

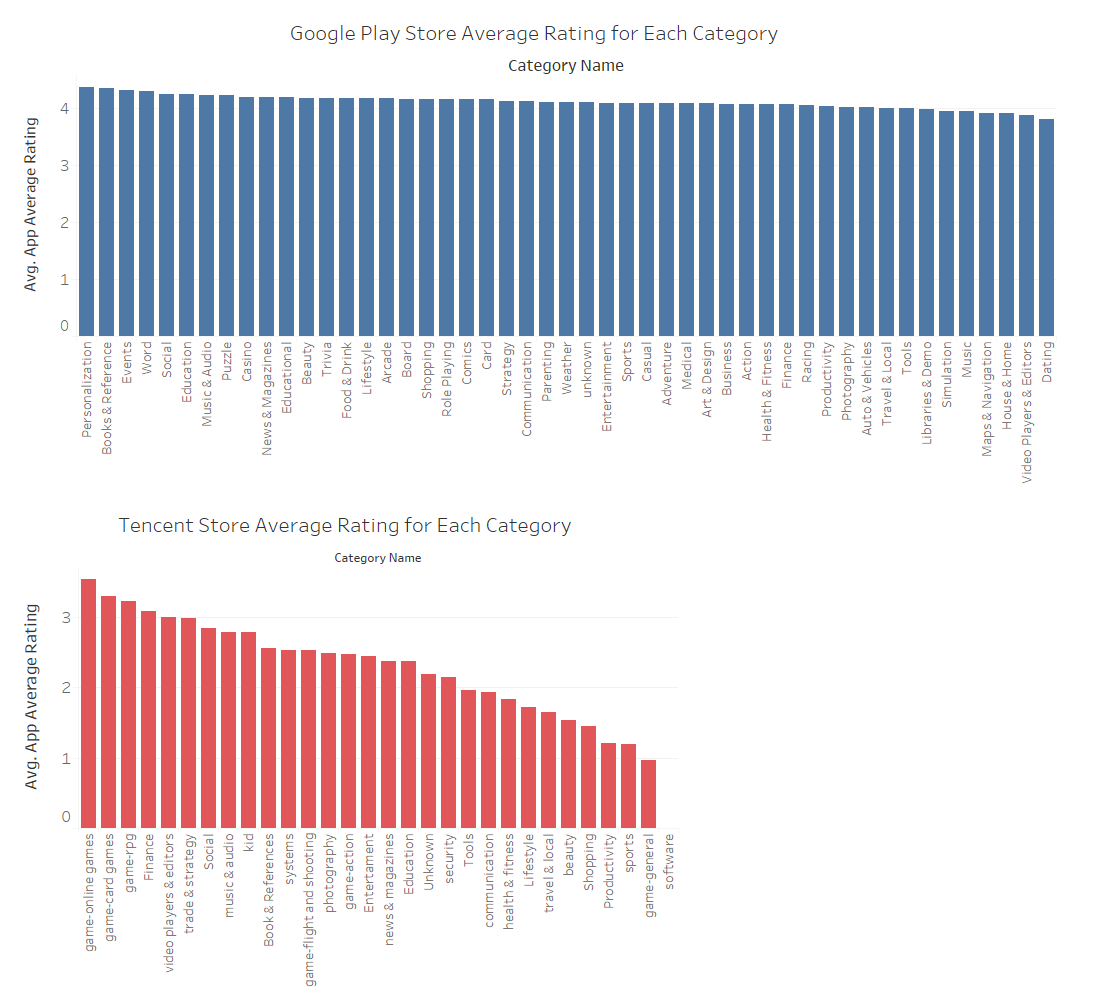


Fig 3-b

Both two charts are about the average rating of each kind of applications. The first chart is about the Tencent, while the second one is the Google.

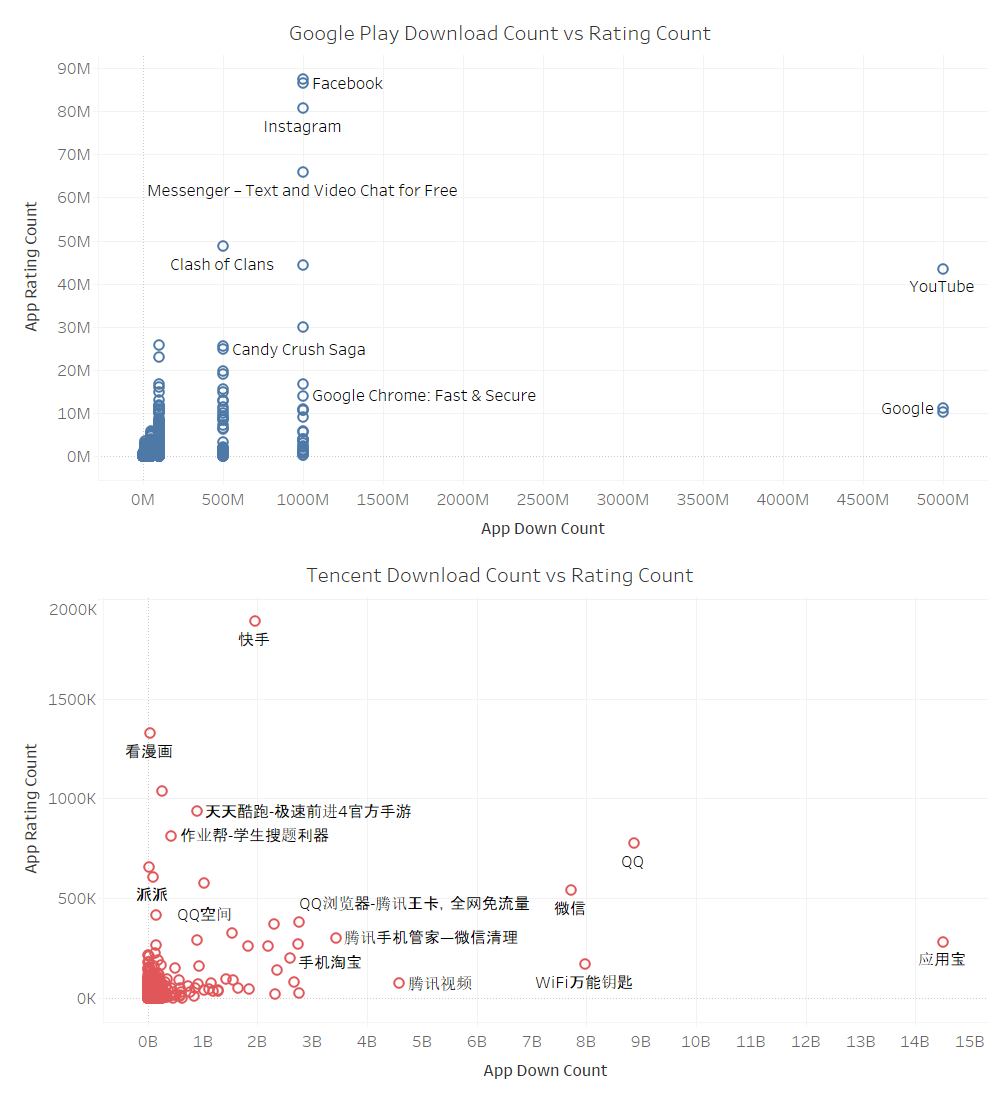


Fig 3-c

These charts show the ralationship between Download Count and Rating Count for each appstore. The abscissa axis is for the Doenload Count and the vertical axis is for the Rating Count. Each app gets its point on the charts.

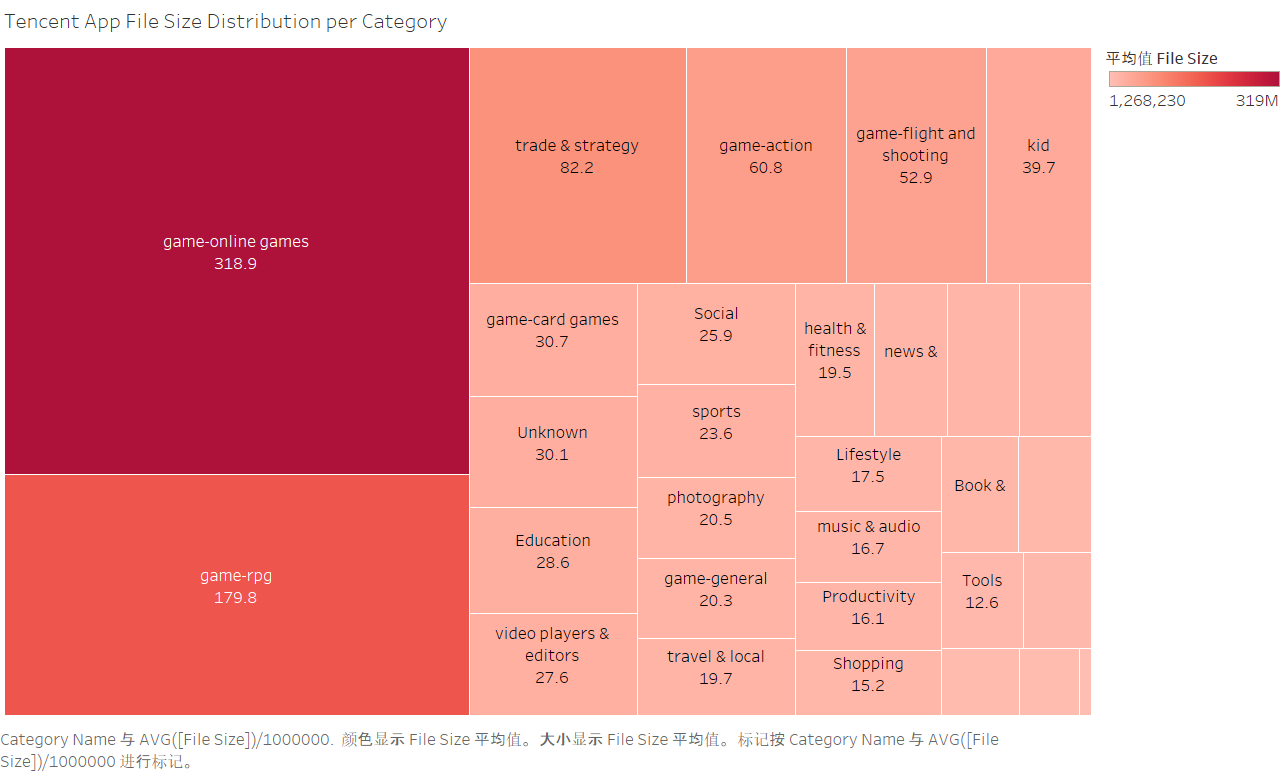
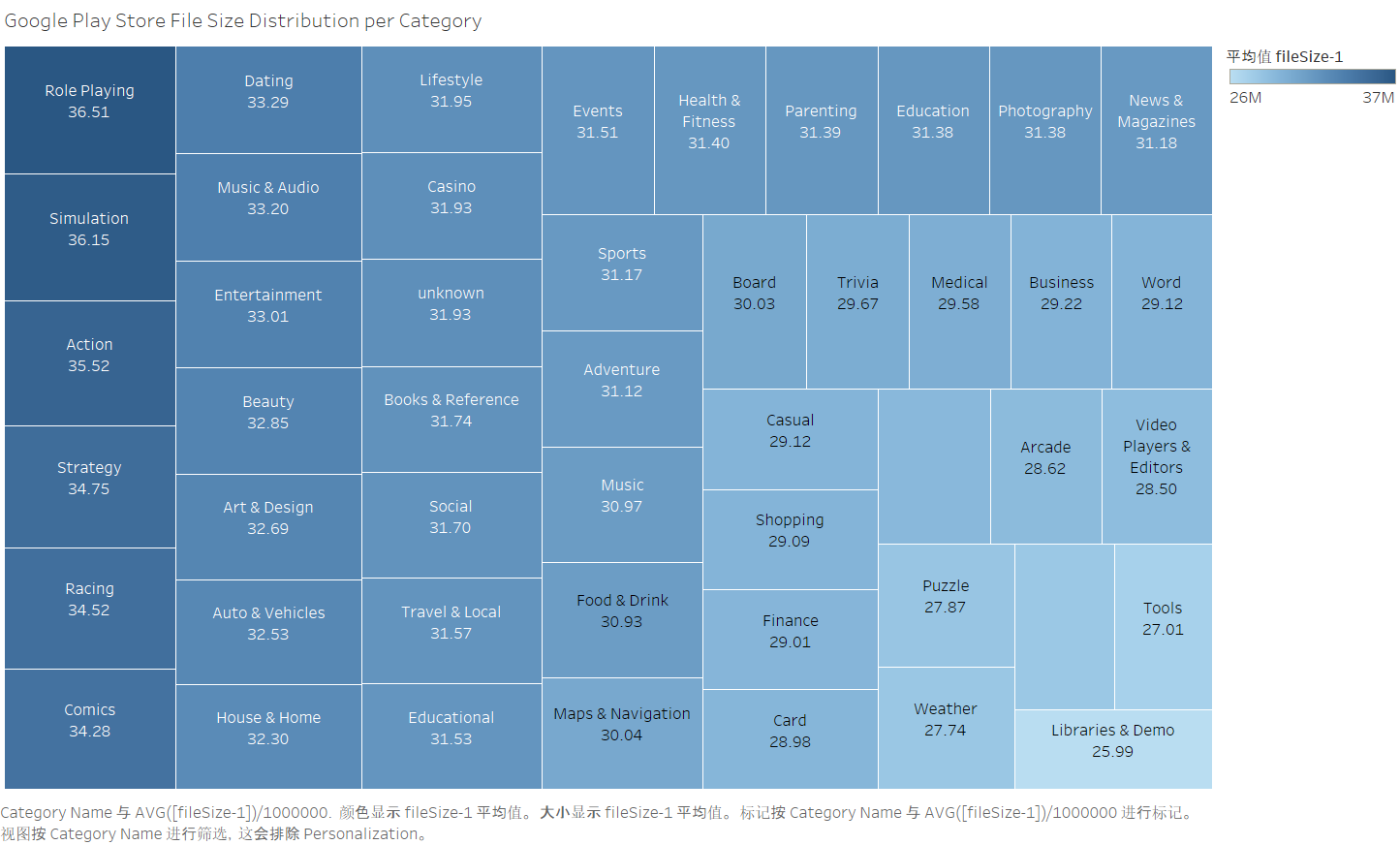


Figure 4

This chart shows the relationships between the average file size and the category in Google Play Store.

The different colour indicate different average file size of each category,colour blue means the average file size is large,colour orange means the average file size is small and colour yellow and light brown mean the average file size is of the middle size.In the middle of each block,we can see the name of category and we have transferred the average file size into a clear,small number.

# References

<Insert your references/citations here.>