

## Lab 3: Data Visualization Report

### 1. Data Set Analysis

### 2. Dashboard Design

### 3. Analysis Results

# Lab 3: Data Visualization Report

1952211 黄金坤

---

## 1. Data Set Analysis

**Selected data set:** Google play store

**Data set characterisation:**

- Relevant information includes: "name", "category", "number of reviews", "volume", "number of downloads", "free or paid", "specific price", "content rating (for groups)", "recent updates", "current version", "suitable models".
- The information dimension is relatively broad, but the core object is unique and the objective of the analysis is very clear.

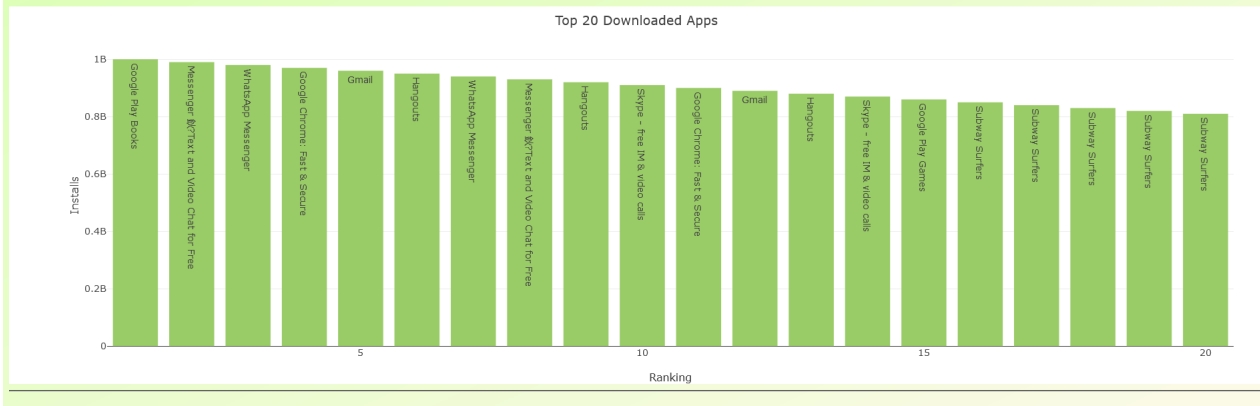
**Data analysis tasks:** Using the given Google Software Store data, the characteristics of popular (highly downloaded) software are explored to help companies determine the direction of their software design and optimisation.

## 2. Dashboard Design

- User Role Segmentation: Only one type of key role - corporate
- Display scenarios: mobile phones, computers, projectors, etc.
- Layout design: The above-mentioned "role division" and "display scenarios" were combined to create a "general division" layout, in which the core indicators are displayed first and then refined layer by layer.

**The final result is shown:**

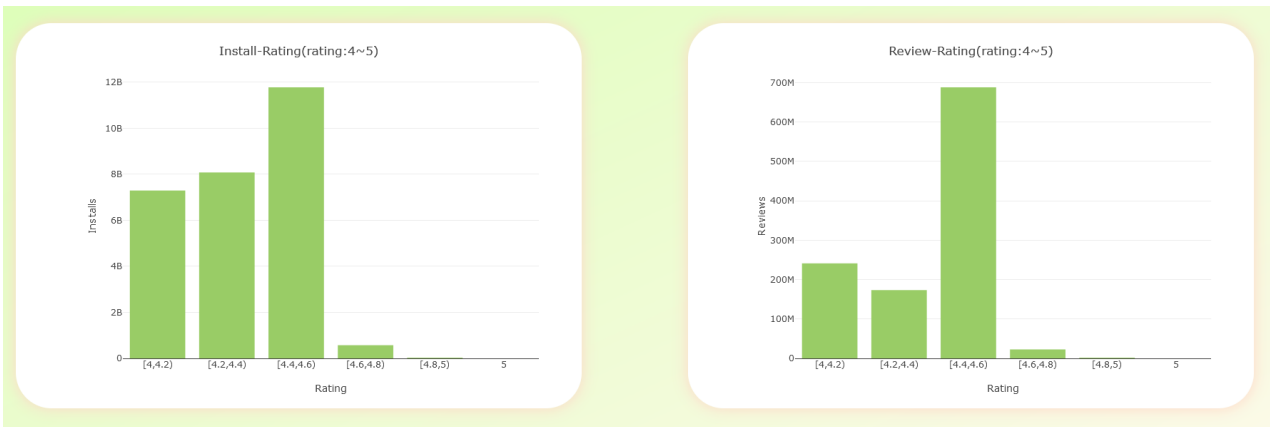
## Google Play Store DashBoard



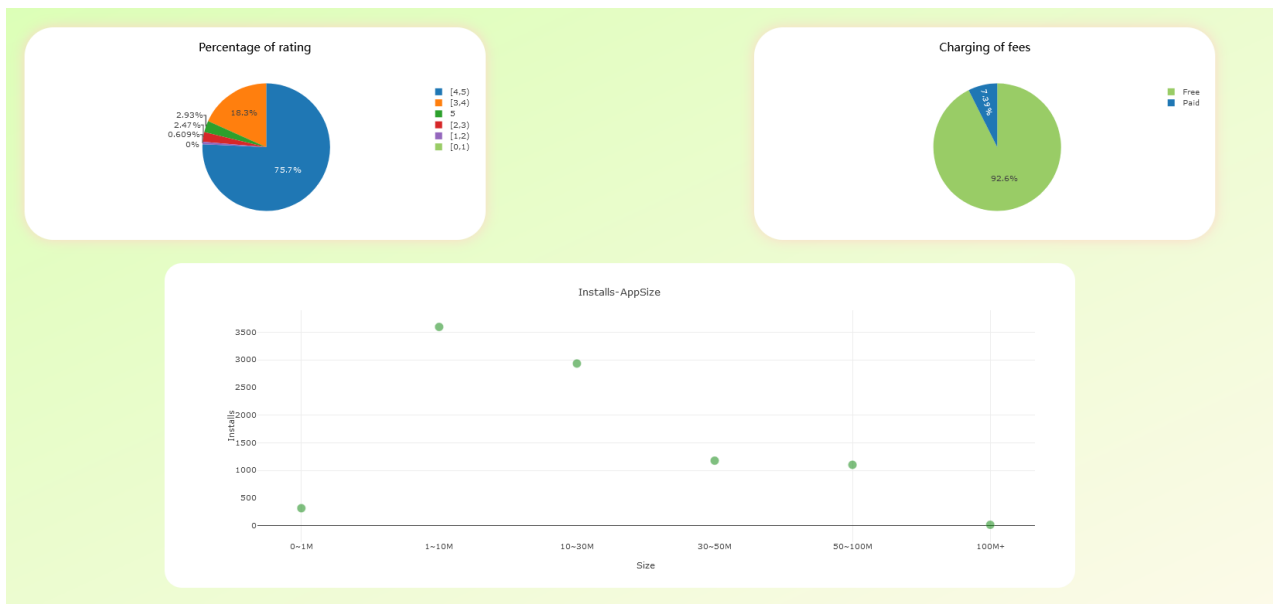
The first thing users will see is the top 20 most downloaded apps on the Google Play Store and their specific downloads. Knowing the top performers in the current competitive market is the first thing to know, as they are largely a direct reflection of the market and the first place to look!



Of course, a cursory look at a good software product needs to be followed by a broader look at the characteristics of other good products. As shown in the graph above, the relationship between downloads and product ratings, reviews and product ratings is shown immediately afterwards. As you can see, almost all of the popular products have a rating of 4 or more. The options at the top of the chart also allow users to check the downloads and reviews of software for different groups, so that companies can understand which groups are the main audience.



We then present a more granular breakdown of the 4 to 5 score range to help companies further understand the specific score range of the most popular products. Here we can see that the best performing software scores lie between 4.4 and 4.6, rather than between 4.6 and 5.



In the final section, further details of the market are described in terms of other dimensions, for example here I have chosen to look at the distribution of software scores, the share of free vs. paid, and the relationship between downloads and the size of the software. It is easy for companies to understand the most accepted scores, pricing, and size across the market.

### 3. Analysis Results

Following the above icon viewing process, companies will receive the following information: the top 20 best performing products in the current market, the most popular products have a score range of 4.4 to 4.6 (rather than those with exceptionally high ratings), most products are free, downloads drop significantly when the software is larger than 30M, and the market is dominated by young users.