Jiang, Keyi
Li, Tina
Nguyen, Noah

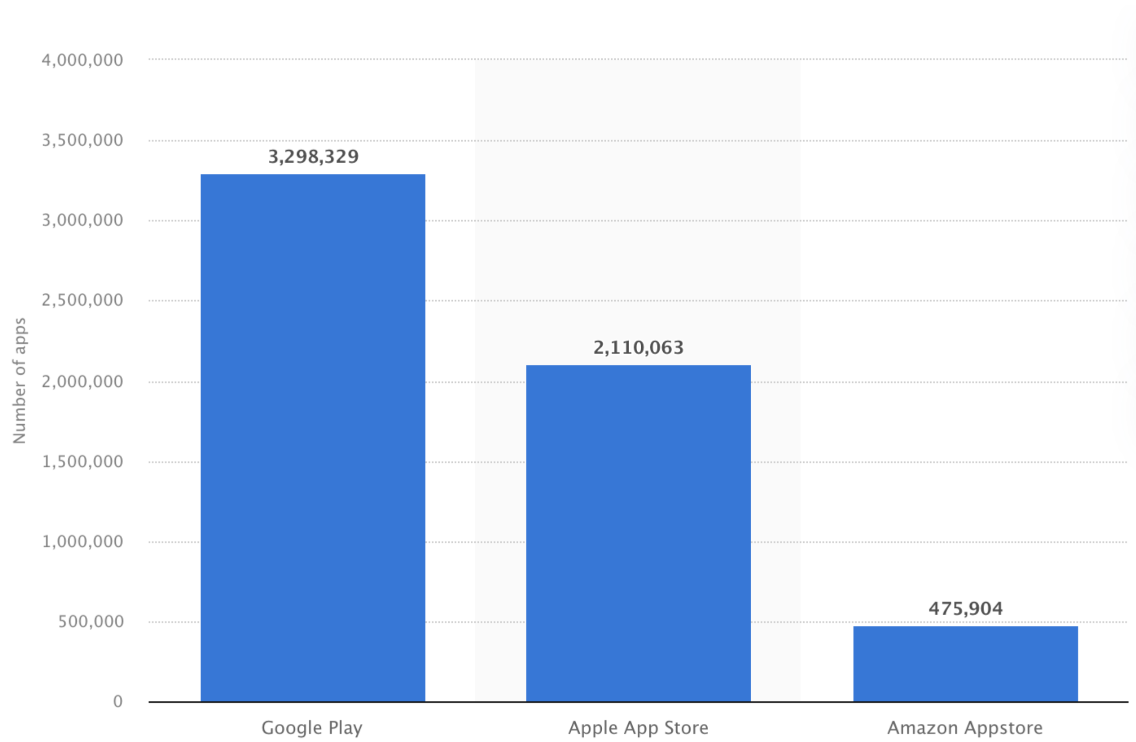
BUS 256A Marketing Analytics (Spring 2022)
Final Project

Google Play Store: An Analysis of Factors behind A Successful Android App

With the rapid development of technology and the increasing number of developers, there are hundreds of apps created daily. However, not all apps are successful. Our report will analyze the factors behind a successful android app. This dataset was acquired from Kaggle. This report is broken down into an introduction, project overview, business model metrics, explanation of data cleansing, descriptive data analysis, regression analysis, and recommendations. Through the study, we found the importance of high ratings, a large number of reviews, a small app size, low price, non-restricting content rating, and frequent updates to the number of installs an app can get. Likewise, apps packed with more content (bigger size), at a lower price, accessible to more people, and frequently updated would have higher ratings. Consumers are more likely to review apps when they have bad experiences with them, so app developers must pay special attention to what their users are saying about their apps, keep the app updated and relevant, and troubleshoot any issues that arise.

1. Introduction to Google Play Store

Created and operated by Google, Google Play Store (also known as Google Play) is the official app store of the Android operating system. It is pre-installed along with the Google Mobile Services (GMS) in most Android mobile devices and cannot be uninstalled from them. When Google first launched the service in late 2008 as Android Market, it offered only a variety of application categories such as social, game, education, etc. However, in 2012, Android Market became Google Play Store, and the platform began allowing users to download music, books, and movie content, making it a digital media store, too.



Number of Apps Available in Leading App Stores as of Q1, 2022 (Source: Statista)

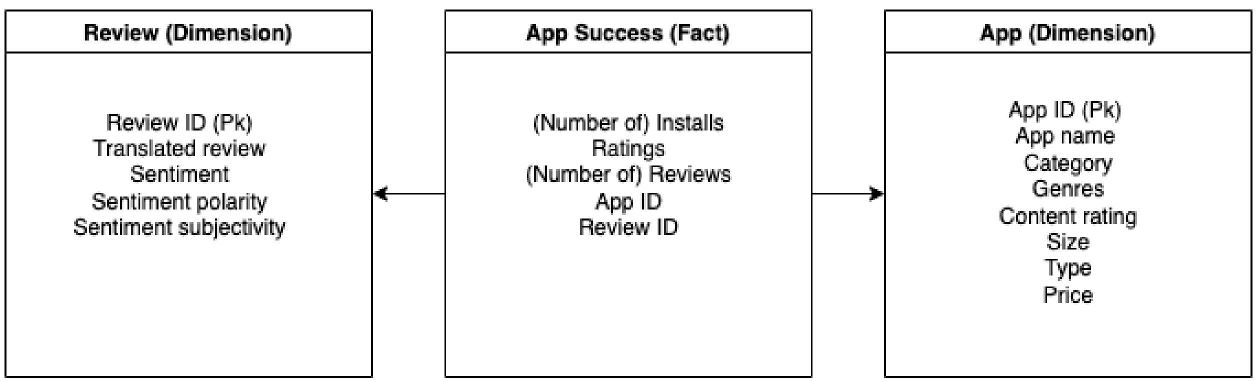
Google Play Store is the current biggest app store by the number of apps available. As of the first quarter of 2022, the platform offers 3.3 million apps, 1.2 million more than the second leading app store, Apple App Store (Statista, 2022). Yet there is still competition for Google as other Android phone manufacturers launch their app stores (e.g. Samsung Galaxy Store and Huawei AppGallery), and alternative app stores such as XDA Labs and F-Droid offer users appealing features that are banned from the Google Play Store (e.g. allowing YouTube video to run in the background) (Android Headlines, 2020).

2. General Project Overview

This project seeks to understand the factors behind a Google Play Store app's success or failure by analyzing and predicting the impact of several app features on the number of installs and number and sentiment (polarity) of ratings and reviews of each app on the platform.

The initial hypotheses of the project are: (1) all positional features of an app (e.g. category, age rating, price, and/or subscription model) have an influence on its success and popularity, and (2) users are more likely to leave reviews for apps that they have a negative experience with. We then test those hypotheses using two data frames, Google Play Store App and Google Play Store User Reviews data frames, which are publicly accessible on Kaggle. The former data frame shows information about almost 10,000 Google Play apps, including category, rating, number of reviews, number of installs, type, price, content rating, and genres. The latter pulls the 100 most relevant reviews from over 1,000 apps on Google Play Stores with the sentiment, sentiment polarity, and sentiment subjectivity. The finding of this project could serve as a guideline for future app developers to design successful and well-received apps that will thrive with Android users.

3. Business Model & Metrics



Dimensional Table of Google Play Store Apps

The main metrics used to assess an app's success are the number of installs, ratings, and reviews.

The popularity of an app can be shown directly through the number of installs as the more people know about it, the more downloads would be made. We are also mindful of the fact that there are niche apps that may target a small population but can be very successful within that community (e.g. apps for patients of specific medical conditions, apps for people with learning difficulties, etc.). Yet the scope of the project and the data frames do not allow us to account for that, so the consensus view of this project is that more installs indicate bigger success.

Ratings show consumer satisfaction with the app. There are apps with very interesting content but low usability, making users frustrated. That means the app developers have not delivered the app values well. Since an app's rating is very visible on Google Play Store, it also determines whether future consumers would install it or not.

Lastly, reviews could be controversial, but outstanding apps should have more positive reviews than negative ones. That demonstrates more people have had good experiences with the apps than those with bad experiences. This also affects potential users' decisions to download the app or not. Controlling this is, however, tricky for app developers because most consumers tend to leave reviews only when they have an extreme belief or feeling about the product. Besides, reviewers who leave negative comments could have deleted the app from their devices, so this could be an indicator of the uninstall rate for the said app, too.

4. Data Cleansing & Standardization

The dataset is downloaded from Kaggle and originally updated in 2019 by LAVANYA. It consists of two tables - ‘googleplaystore.csv’ and ‘googleplaystore_user_reviews’.

a. 'Googleplaystore.csv' contains 10,841 entries and 13 columns in total and, according to Kaggle, it has 9,660 unique apps. It mainly includes important features of the app that we shall use later to predict what kind of app is favored by the public and more likely receives a higher rating in general.
b. ‘Googleplaystore_user_reviews’ contains 64,295 entries and 5 columns. This table records actual reviews of apps and measures the reviews’ sentiment, including polarity and subjectivity.

There were five steps taken to do the data cleansing and standardization work:

a. MetaData Explanation: it is very essential to understand each feature before doing any specific change to the data. I will explain each later.
b. Standardize Field Names: renaming the columns and checking the index for easier coding and clearer table formats. For our dataset, I simply removed the space that appeared in the middle of column names.
c. Keep Relevant Data Fields: First, we removed the unnecessary columns of app versions. Then since the percentage of missing values is small, we simply deleted NaN values in the data frame and dropped duplicates.
d. Correct Feature Type: For better understanding and further analysis, we have to carefully correct each feature.
    i. Reviews: Convert to int.

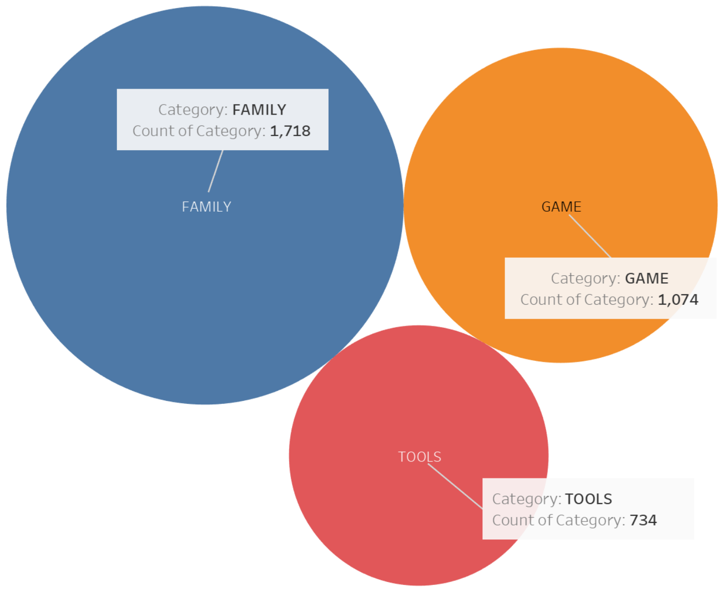
ii. Size: Remove the character M or K. K is the smallest measurement unit in this case and assuming 1M= 1000K. Size ending with M should multiply 1000. Then fill the 'varies with device' field with the average size computed.
iii. Installs: Remove characters ‘,’ or ‘+’ and convert to int.
iv. Type: Index it - 0 for free and 1 for paid.
v. Price: Remove ‘$’ and convert to float.
vi. Genres: Instead of specific genres, convert them into genre numbers.
vii. Last_updated: Convert it to date-time.

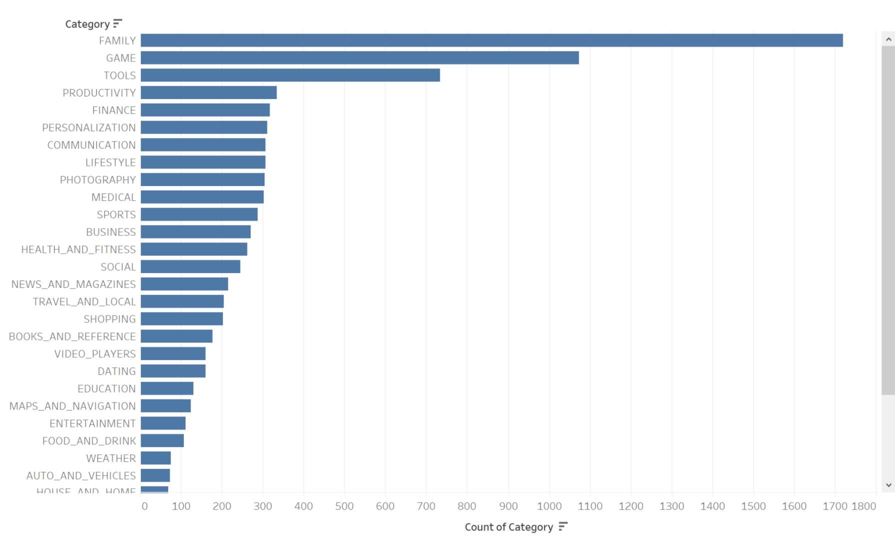
e. Adjust/Add necessary columns:
   i. Days_after_last_updated: Since the data only contains those before 2019, we set '2019-01-01' as the current date and calculated the number of days since the app last updated. Therefore, the small it is, the more recent the app got updated.
   ii. Merge data frame: to make user_reviews a feature and analyze its relationship to high ratings or high reviews, we decide to merge two data frames on the app name. After merging, the new data frame contains 40,376 reviews of 816 unique apps.

5. Descriptive Data Analysis (Picture in presentation):

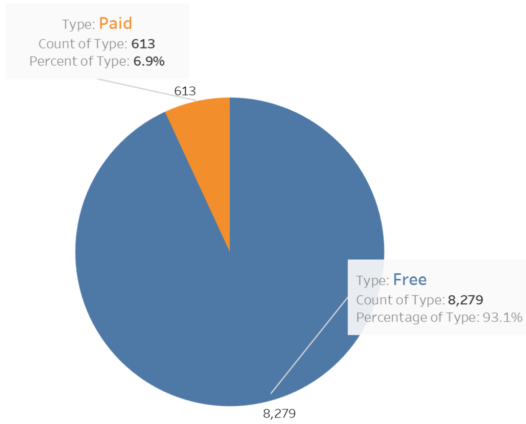
There are 8,196 unique apps left in the first table.

a. Category: 33 distinct categories were addressed amongst all the apps. The top 3 categories that most apps fall into are Family, Game, and Tools, which together account for almost 40% of the total apps.
   When segmenting the categories, categories that are only available for ‘Mature 17+’ are similar to those which apps on average reviewed negatively, such as ‘Social’ and ‘Game’.

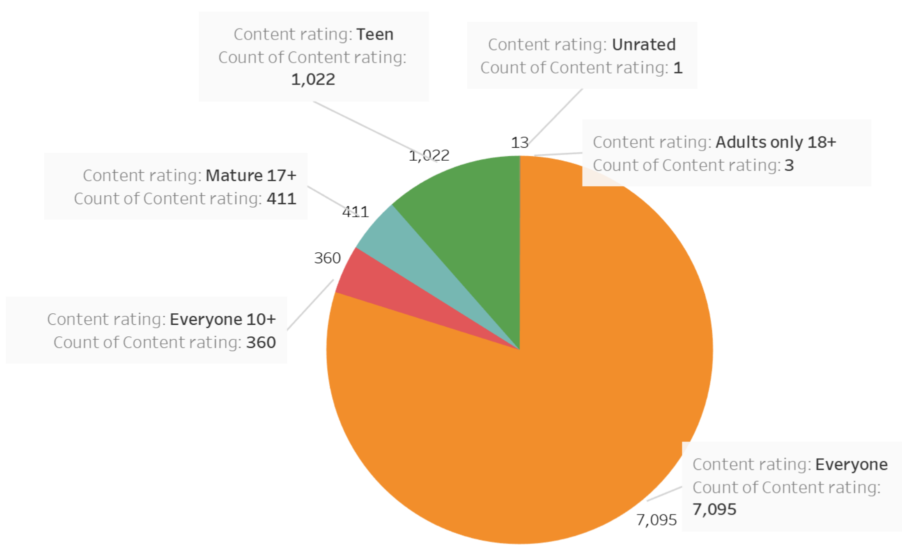




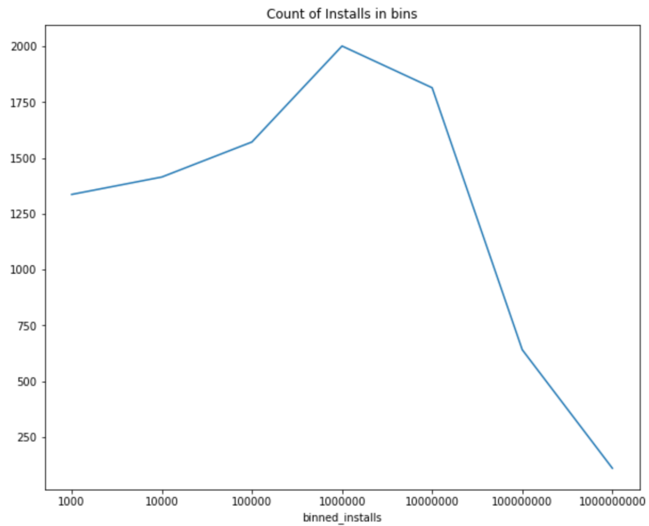
b. Type: Only 6.9% of Google Play Store apps require payment, ranging from $0.99 - $400.

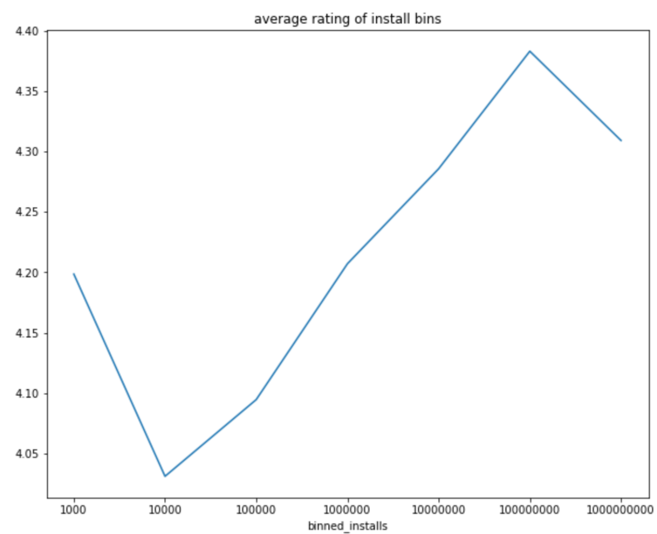


c. Content_rating: 6 different kinds of content rating, including 'Everyone' (0), 'Teen' (1), 'Everyone 10+' (2), 'Mature 17+' (3), 'Adults only 18+' (4), 'Unrated' (5). These content ratings are indexed from 0 to 5 based on how accessible they are. Among them, almost 80% of apps are accessible to everyone, following 11.5% for Teen. The number of apps falling in 'Adults only 18+' and 'Unrated' is small enough to neglect.

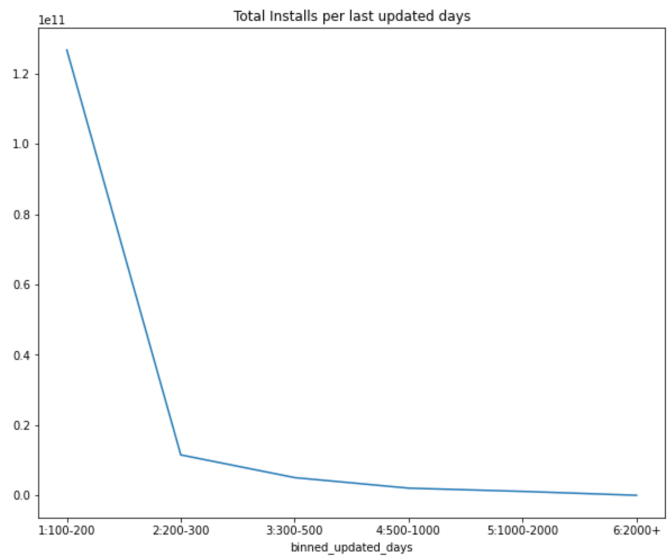


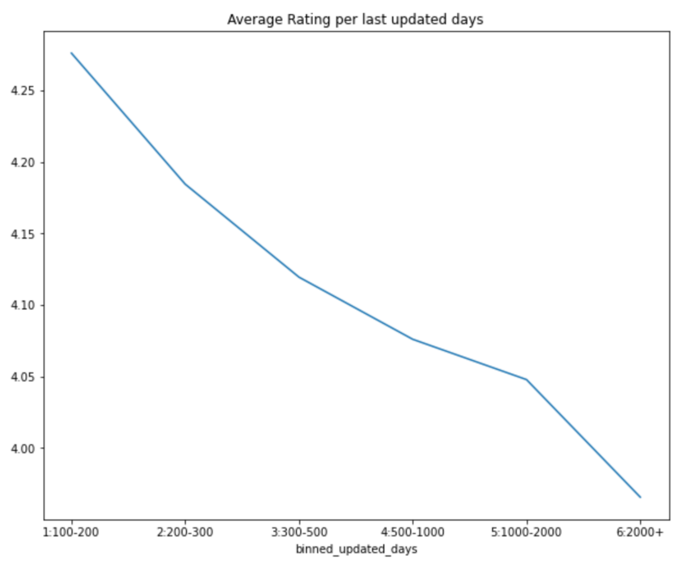
d. Installs: Installs numbers were separated into several bins from 1,000 to 1,000,000,000. The number of apps increases when the installs number reaches high but the turning point occurs when hitting 10,000,000 installs. However, the average rating of each install bin increases until reaching 100,000,000. Thus, in our data, most apps have been installed on a large number of basis and if an app has been installed more times, it tends to have a high rating. But the best app that accumulates over 1 billion installs would probably be rated lower.



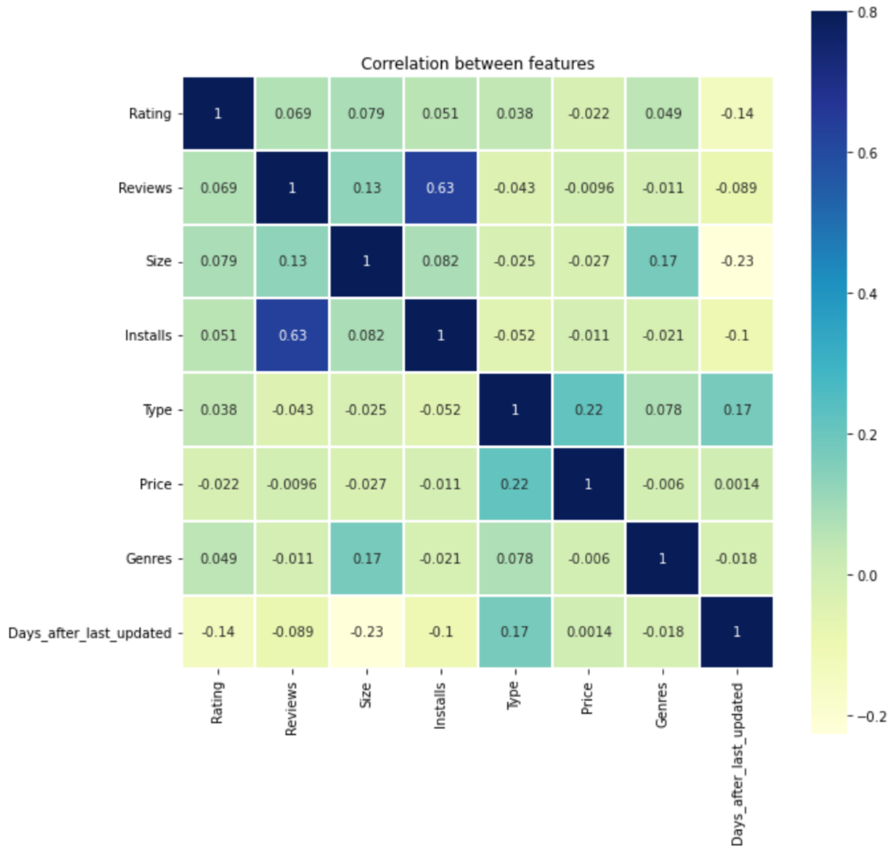


e. Days after last updated: As mentioned before, the fewer days there are since the app was last updated, the more recently updated it is. Most recently updated apps show a better score in rating and a much higher installs number in total.


Total Installs per last updated days
1e11
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1:100-200
2:200-300
3:300-500
4:500-1000
5:1000-2000
6:2000+
binned_updated_days


Average Rating per last updated days
4.25
4.20
4.15
4.10
4.05
4.00
1:100-200
2:200-300
3:300-500
4:500-1000
5:1000-2000
6:2000+
binned_updated_days

f. Heatmap:



The heatmap explains many relationships that are approved by our common senses. For example, Reviews are strongly positively correlated to Installs; Rating is negatively correlated to Days_after_last_updated; Price is positively correlated to Type; etc.

6. Regression Analysis

Simple Regression:

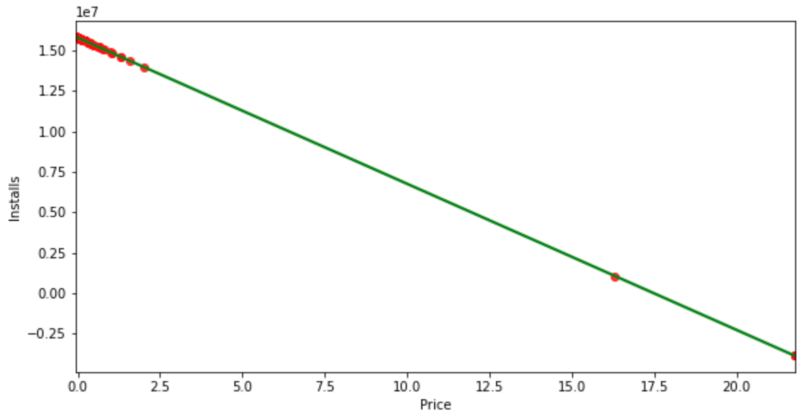
We conducted three simple regression tests to test our hypothesis.

1) Price & installs

The dependent variable y is installs and the independent variable is price.

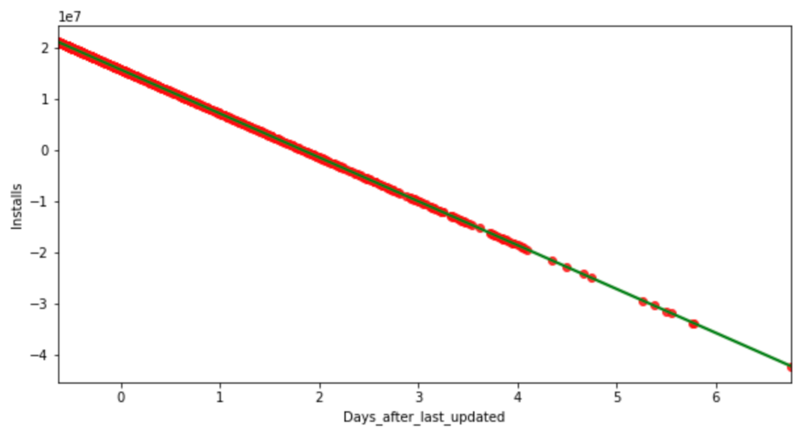
R2 score is -10816.48, this indicates that price and installs are not closely correlated.

Coefficients are -902,702.51 indicating that with every price increase, there is a proportional decrease in installs. Because the coefficient is negative, this is an indirect relationship meaning that as price increases, installs decrease. The intercept is 15,786,886.96.



2) Days after last updated & installs

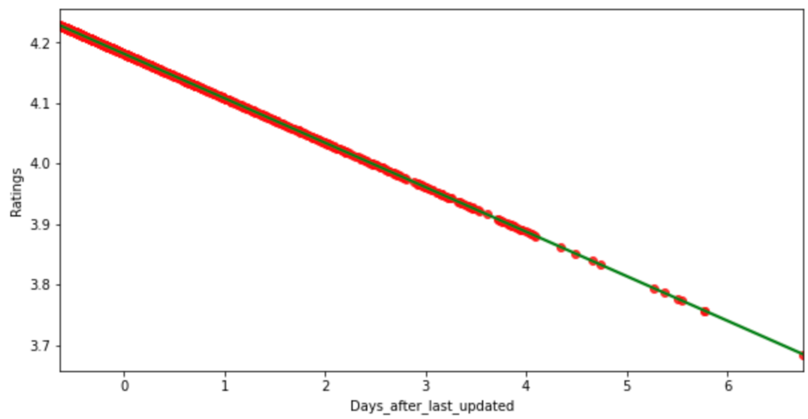
From our descriptive data analysis, most recently updated apps show a higher installs number in total. Thus, we decided to assess the relationship between the two variables. The dependent variable y is installs and the independent variable is days after last updated.



R2 score is -117.08. Coefficients -8,597,225.02 indicate that with every increase in days after the last update, there is a proportional decrease in installs. The further away from when the app last updated, the few installs there will be. The intercept is 15,786,886.9. This is a negative correlation that further supports our descriptive data analysis, there is a higher install rate when apps are more recently updated.

3) Days after last updated & rating

From our descriptive data analysis, most recently updated apps also show a higher rating score. Then assessed the relationship between the two variables with the dependent variable y being ratings and the independent variable being days after last updated.



R2 score is -47.08. Coefficients -0.07, indicating that with every increase in days after the last update, there is a proportional decrease in installs. The further away from when the app last updated, the few installs there will be. The intercept is 4.18, meaning at zero days after last updated the rating is around 4.18. There is a negative relationship between ratings and days after the last update, the more recent the update, the higher the rating.

Multiple Regression:

1) MLR on Rating with Size, Price, Content rating, and Days after last updated

R2 score is -40.012

Residual sum of squares (MSE): 0.26

Co-efficients are [[0.028 -0.011 -0.0095 -0.068]]

Intercept is [4.18]

Surprisingly, the regression result shows that apart from the size, all the above features have a negative relationship with rating. This means that the heavier the app is, the more likely it is rated higher, which could be interpreted as that users who can afford the space occupancy prefer apps with more content and capability. Meanwhile, they are more critical of apps that cost higher, which is relatively understandable as there are many free apps, so people would only pay for another app when they expect outstanding capabilities and performance. As for content rating, 0 means the app is accessible to all ages, and as the number increases, the target population shrinks, which means the user communities of the apps are also more mature, and hence more critical, when the content rating rises. Thus, the overall ratings of those apps may be lower due to the stricter perspectives of their users. Lastly, as discussed before, the longer the period since the last update is, the more outdated the app is, and there may be unresolved issues, which cause users to be frustrated.

2) MLR on Installs with Rating, Size, Reviews, Price, Content rating, and Days after last updated

R2 score is -0.0296

Residual sum of squares (MSE): 6140738581940644.00

Co-efficients are [[111249.81, -541221.86, 54704717.72, -446653.27, -587169.81, -3750014.12]

Intercept is [15993799.37]

Rating and number of reviews have a positive correlation with the number of installs. This is because potential users indicate high ratings as an indicator of previous users' good experience with the apps, so they are more likely to download the app to their devices, too. A large number of reviews means that there are a lot of people who have used those apps, validating the ratings as legitimate (as opposed to a 5-star rating by only a few people).

On the other hand, size, price, content rating, and days after the last update have negative correlations with the number of installs. Bigger apps require more available space on the devices, so this could be a hindrance to people who have limited data storage available. Likewise, most people usually opt for the free apps instead of paying extra for paid apps due to cost incentives. And as aforementioned, the more accessible an app is, in terms of age rating, the bigger its target user population is, and hence the more downloads it can have. And users would prefer apps that are frequently updated, which shows how much effort the developers put into their products after launch and how much care they have for their customers. Another explanation is that, as with how Google search engine optimization rating works, the Google Play Store app may push results (apps) that are more recently posted or updated to the top results to keep the content fresh and relevant.

3) Review Number & Sentiment Polarity, Subjectivity

R2 score is -15.86.

Residual sum of squares (MSE): 16,334,983,438,286.82, showing that there is a huge difference between the estimated point and the actual point.

Co-efficients are -992,562.94 for Sentiment Polarity and 223,890.3 for subjectivity.

Intercept is [1,405,517.66].

There is a negative correlation between review number and sentiment polarity, confirming our hypothesis that users who did not enjoy the apps are more likely to leave a review on them, while there is a positive correlation between review number and subjectivity. There are a total of 40,376 reviews and a large negative coefficient for sentiment polarity, meaning there is a large difference between positive and negative reviews. Generally, apps have more negative reviews than positive ones.

7. Conclusion & Recommendations

Overall, the data sample used in this study was not representative of the whole population of apps and user reviews on the Google Play Store, so the statistics and regression results drawn from them do not explain the whole picture of the apps' successes and failures as we desired. The data frames have thousands of unique values, but those IDs were not randomly selected from Google Play Store's 3.3 million apps and all the reviews. In fact, the provider of those data frames never mentioned how they chose the 10 thousand apps in the data frames, and the reviews included are just the top 100 reviews that are deemed most relevant. This affects the result of our analyses, as shown in the negative R2, and thus our first hypothesis could not be statistically significantly proved. Nevertheless, there are some key findings:

a. Categories such as 'Social' and 'Game' are more likely to contain violations or inappropriate topics, which would easily be reviewed negatively too. In general, 'Tools' is a very safe choice for a category as users review them objectively based on usability and efficiency and not content. App developers should closely monitor user reviews of their apps on the Google Play Store to adjust app content and fix issues.
b. More recently updated apps tend to have more installs, reviews, and better ratings. So developers need to continue maintaining the app, fixing bugs, and adding new features would contribute to the popularity and quality of the app. Similarly, ratings and reviews are also important predictors of an app's number of installs.
c. Size, price, and content rating could significantly restrict the number of people who will be downloading an app, so developers should only increase these

features if their apps belong to a niche category and can offer very competitive capabilities that users are willing to disregard limitations for. Examples of those apps are specialty apps (medical apps for specific patients, complex programs for specific professional occupations or tech addicts, and/or apps targeting wealthy consumers).

d. Users are more likely to leave reviews when their experiences with the apps are on an extreme level on either side of the spectrum–positive or negative. However, most extreme experiences are negative, and people who enjoy using the app usually do not want to write an extra review to repeat the advantages that other users have pointed out in their reviews. Hence, the overall trend is that once an app has run for a considerable time, most reviews they get are usually criticism. It could be detrimental to an app if all recent reviews, which users care the most about since it reflects the app's current performance, are flooded with negative comments. If that does happen, developers should redeem by paying special attention to every bad review (especially following an update) and being responsive by either answering their complaints or fixing the issues in the next updates.

Appendix

1. https://drive.google.com/drive/folders/1xmDZ4_GpjqK6f8gTghwpQnrIGGVEi9WA?usp=sharing
2. The google drive folder shared above includes the original data tables, Coding files, and a Tableau package file. If there are any problems, please contact: keyijiang@brandeis.edu.

Bibliography

1. "Biggest App Stores in the World 2022." Statista, April 27, 2022. https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/#:~:text=During%20the%20first%20quarter%20of,million%20available%20apps%20for%20iOS.
2. Lucic, Kristijan. "What Is Google Play Store? Everything You Need to Know." Android Headlines, April 3, 2020. https://www.androidheadlines.com/what-is-google-play-store.html.
3. Lavanya. "Google Play Store Apps." Kaggle, February 3, 2019. https://www.kaggle.com/datasets/lava18/google-play-store-apps?select=googleplaystore_user_reviews.csv.