

INTRODUCTION

In recent years, mobile apps became indispensable tools in tackling our daily routine issues. With the explosion in online interactive learning, hands-on coding resources, video tutorials, app development has become easier than ever. With thousands of applications produced monthly across different platforms, Google App store is swamped with countless new applications, with an increasingly enormous number of developers working tirelessly to achieve success in their craft. Since most Google App Store applications are free and not paid, the income expectation is incredibly uncertain and out-of-the-way no matter how the in-application purchases, adverts, and memberships boost the achievement of the app. Most of the time, an app's success is typically dictated by the number of installations and also the client reviews that it's gotten throughout its existence rather than the income. App ratings are feedbacks provided voluntarily by users and are a very important evaluation criterion for many apps. Nonetheless, these ratings can often be biased because of insufficient or missing votes. In the time past, the degree of downloads an app received determined its ranking within the Google App Store. Users who clicked the "install" button are the key influential factor to an app's success. Because of the dynamic nature of app store algorithms, 'installs' alone are not any longer a benchmark to live an app's ranking because they were within the past. The measuring parameters to observe how an app is performing is very important, but ultimately, the standard of users (in-app actions like ordering food on the app, booking a ride, opening a feature of your app, clicking on Ads on the app, and trying to find products, services, and content) that installs an app determines its success. In addition, by understanding the various varieties of users, developers can start to fully grasp why quality is more beneficial than quantity. Additionally, analyzing micro-transactions, subscriptions, reviews, ratings, and in-app actions will also help gather insight into who the foremost engaged users are.

THE OBJECTIVE OF THIS ANALYSIS IS TO;

1. show with actionable insights, the factors that drives the number of downloads and the average rating of an app on the google play store.
2. suggest what an app developer should focus on – “increasing number of downloads or average ratings.

DEFINITION OF ATTRIBUTES

App — Application

Category — The categories of the Application

Rating — The overall user rating of the Application

Reviews — The number of user reviews for each Application

Size — The size of the Application

Installs — The number of user downloads or installations

Type— Either free or paid

Price — The cost of the Application

Content Rating — The Age group of the Application, Everyone, Teen and Adults

Genres — The style or type of Application

Last updated — The date the App was last upgraded

Current Version — The latest edition of the Application

Android Version — The edition for Android users.

HYPOTHETICAL QUESTIONS TO CONSIDER

1. What insights can you draw from this data about the drivers of the number of downloads and average rating of apps on the Google app store?
2. Based on your Analysis, what do you think an app developer should focus on – “Increasing number of downloads or Average ratings”.

And to help find solution to the above questions, I will go through the following steps;

- ✓ Loading the Dataset
- ✓ Data Cleansing
- ✓ Data exploratory analysis and insights
- ✓ Summary
- ✓ Conclusion

LOADING THE DATASET

The dataset used for this project is the ‘Google Play store Dataset’ provided by Hult International Business School. The dataset was loaded onto Microsoft Excel, Rapid Miner and Microsoft Power Bi, to prepare it for the analysis

DATA CLEANSING

Data cleaning and preparation is the first most important step in Data Analysis project. As evidence shows, most data analyst spend most of their time up to 80% on data cleaning. The data cleansing of this project was carried out using Microsoft Power Bi, Python and Microsoft Excel, to carry-out the following tasks;

1. Remove duplicate values.
2. Remove the “+” symbol from ‘installs’ field and change it to numeric.
3. Remove the “\$” object attached to the prices in the ‘Price’ field and change it to numeric.
4. Change the Last updated field to a Date Format.
5. Change the ‘Current Version’ column to Decimal and remove all the null values and errors.
6. Remove the non-English characters from the App names.

7. Dealing with all the columns with 'NaN' or Null values (Size, Rating and Android Version) and change it to numeric.
8. Dropped all the rows with Missing values and "Values with device" data.
9. Change the 'Review' datatype to numeric

By the end of the Data cleansing processes we had a total data record of 7,029 from the original size of 10,842.

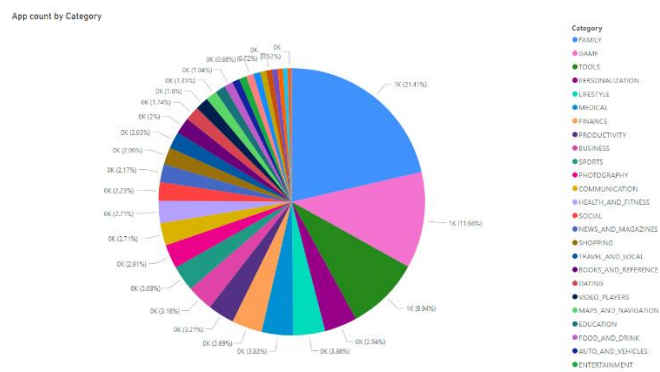
App	Category	Rating	Reviews	Size	Installs	Type	Price (\$)	Content Rating	Genres	Last Updated
aa	FAMILY	4.2	886418	6.5M	50000000	Free	0	Everyone	Strategy	30-Jun-18
AabeHayatFullNovel	BOOKS_AND_REFERENCE	4.3	1476	41M	100000	Free	0	Everyone	Books & Reference	21-Feb-17
AAFP	MEDICAL	3.8	63	24M	10000	Free	0	Everyone	Medical	22-Jun-18
AajBanglaeisamayerkhobor	NEWS_AND_MAGAZINES	4.9	34	3.8M	1000	Free	0	Everyone	News & Magazines	15-Jan-17
ab	FAMILY	3.8	10	22M	1000	Free	0	Everyone	Casual	27-Oct-16
ABblastMatch	FAMILY	4.3	111	48M	10000	Free	0	Everyone	Puzzle	19-Jun-18
ABClickShop	SHOPPING	3.8	454	18M	100000	Free	0	Everyone	Shopping	26-Jun-18
ABcmousecom	FAMILY	4.3	50887	91M	5000000	Free	0	Everyone	Education;Education	03-Jul-18
ABCNewsUSWorldNews	NEWS_AND_MAGAZINES	4	18976	35M	1000000	Free	0	Everyone 10+	News & Magazines	06-Jun-18
ABCPreschoolFree	EDUCATION	3.8	27572	25M	5000000	Free	0	Everyone	Education;Education	25-Oct-17
ABCrushMatch	FAMILY	4.3	96	54M	10000	Free	0	Everyone	Casual	23-Feb-17
abeoCoder	MEDICAL	3.1	83	3.2M	10000	Free	0	Everyone	Medical	30-Jul-18
ABMatch	FAMILY	4.2	9	30M	1000	Free	0	Everyone	Casual	19-May-18
ABMobileApp	HEALTH_AND_FITNESS	3.4	2105	8.2M	100000	Free	0	Everyone	Health & Fitness	21-Aug-17
ABrepeater	VIDEO_PLAYERS	4.4	32	239k	5000	Free	0	Everyone	Video Players & Editor	18-May-18
ABRepeatPlayer	VIDEO_PLAYERS	4.3	1628	18M	100000	Free	0	Everyone	Video Players & Editor	06-Mar-18
AbsCoreBackWorkoutChallenge	HEALTH_AND_FITNESS	4.8	3908	33M	100000	Free	0	Everyone	Health & Fitness	09-Jul-18
ABScreenRecorder	TOOLS	4.2	533	5.0M	10000	Free	0	Everyone	Tools	02-Oct-17
AbsoluteRCHeliSimulator	SPORTS	3.9	654	50M	10000	Paid	4.99	Everyone	Sports	20-Dec-17
AbsTrainingBurnbellyfat	HEALTH_AND_FITNESS	4.7	2071	6.9M	100000	Free	0	Everyone	Health & Fitness	24-Jul-18
AbsWorkoutBurnBellyFatwithNoEquipment	HEALTH_AND_FITNESS	4.8	69279	10M	10000000	Free	0	Everyone	Health & Fitness	23-Jul-18
AbsworkoutDayFitnessChallenge	HEALTH_AND_FITNESS	4.8	18604	15M	1000000	Free	0	Everyone	Health & Fitness	02-May-18
AbsWorkoutDaysFitnessAppforSixPackAbs	HEALTH_AND_FITNESS	4.6	299	11M	100000	Free	0	Everyone	Health & Fitness	03-Aug-18
ABTaxi	MAPS_AND_NAVIGATION	4.1	41	25M	1000	Free	0	Everyone	Maps & Navigation	04-Jul-18
AbWorkouts	HEALTH_AND_FITNESS	4.1	27501	7.0M	1000000	Free	0	Everyone	Health & Fitness	05-Jan-18
AbWorkoutsAbGenerator	HEALTH_AND_FITNESS	4.7	21	3.8M	1000	Free	0	Everyone	Health & Fitness	31-Dec-16
ACAirConditionerRemote	TOOLS	3.4	7816	5.1M	1000000	Free	0	Everyone	Tools	30-Jul-16
ACAirconditionTroubleshootRepairMaintenance	BOOKS_AND_REFERENCE	4.2	27	3.1M	5000	Free	0	Everyone	Books & Reference	07-Feb-18
ACallFromSantaClaus	FAMILY	3.9	3720	11M	1000000	Free	0	Everyone	Entertainment	04-Dec-16

An Extract of the Cleansed Dataset

DATA EXPLORATORY ANALYSIS

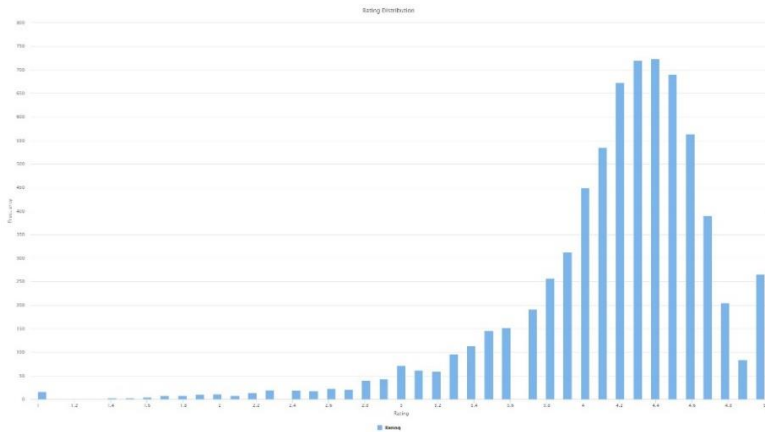
The main features used for this Analysis are; *App, Categories, Rating, Review, Size, Installs, Type, Price, Content and Genres.*

App count by Category



In all of the 33 categories, *Game*, *Family*, and *Tools* seem to be dominating Google Play store, accounting for more than 42.01% of all the apps [Game (21.41%), Game (11.66%), 8.94%).]

App Rating Distribution

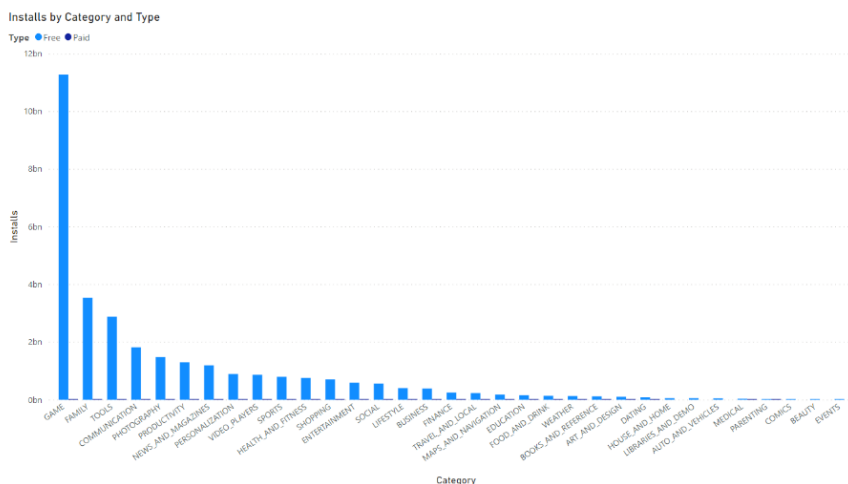


The app rating distribution resembles a skewed bell-shaped curve with the mean value (average) of 4.161.

The major factors that drives number of downloads are;

- i. Rating
- ii. Reviews
- iii. Installs (No of downloads)
- iv. Types (Free or paid)
- v. Category

Installs by Category and Type

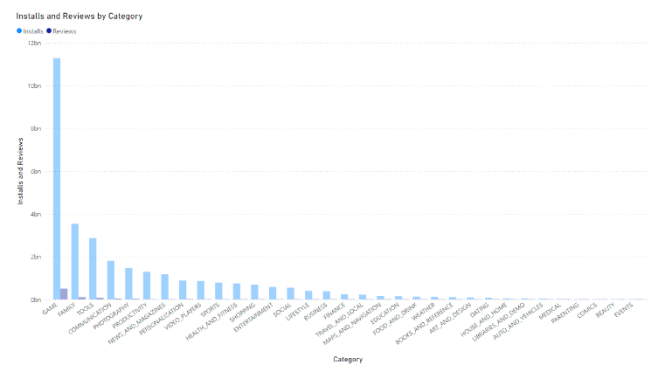


From the graph we see that the ‘Game’ apps have the highest number of downloads with over 11billion ‘installs’, it is not surprising as they are also the category with the highest number of free

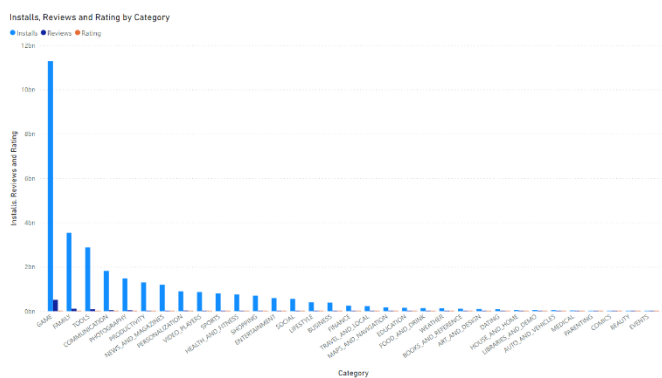
apps. A further insight into the ‘Review’ and ‘Rating’ of this category will tell us if this is the category most developers will invest more effort and time on.

The number of downloads (96.86%) far much outweighs the number of ‘Reviews’ (3.14%) as show in the chart and it further shows us the huge gap that exist between the number of users who downloaded the apps and those that did a review on the app which shows they are satisfied and will most likely to continue to use this app for a long time to come.

Installs and Reviews by Category

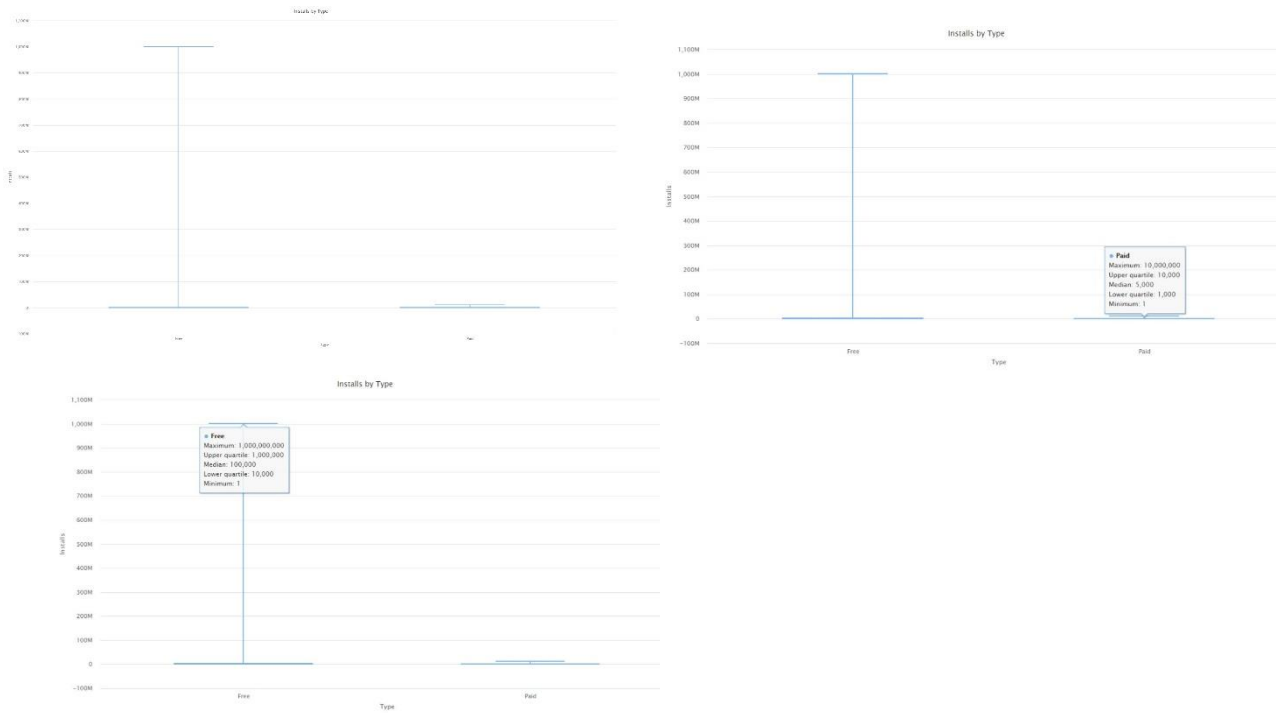


Installs, Reviews and Rating by Category



Okay things are getting a little bit clearer here as the hottest category of apps ‘Game’ comes off with not so much on Reviews, with a Review of 514million as against the downloads (installs) of 11billion. One would expect that with it very high downloads that a lot of users would have done lots of reviews on these apps. This shows us that majority of the apps in this category are ‘bad market’ for the developer because the user retainership or loyalty would be very low. How can the app get the attention of investors when many users are not even on the app? So what it means is that people install the apps, get dissatisfied and probably discard them for another. I mean if I am a user and I am satisfied using an app I will be more than delighted to write a nice review to encourage my friends to go check it out. Bottom-line, majority of the apps in this category are low quality apps.

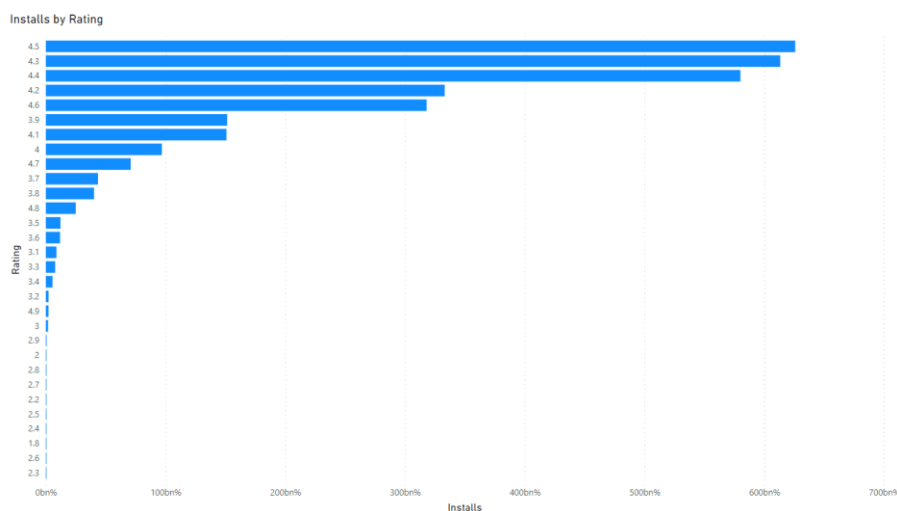
Installs by Type



The gap in the number of downloads is significant. The median value of free apps is a hundred times larger than that of paid ones with the interquartile range being significantly broader for the former one. Also the maximum outlier of a free app category can reach values above one billion, whereas only a few paid apps can hit the 10 million download benchmark.

An app developer possibly wants to know how customers feel about paid and free apps, respectively. A prevalence of negative reviews can decrease your profits in the long run, so before embarking on any specific route, you need to ensure that the category and the type go in line with your business strategy.

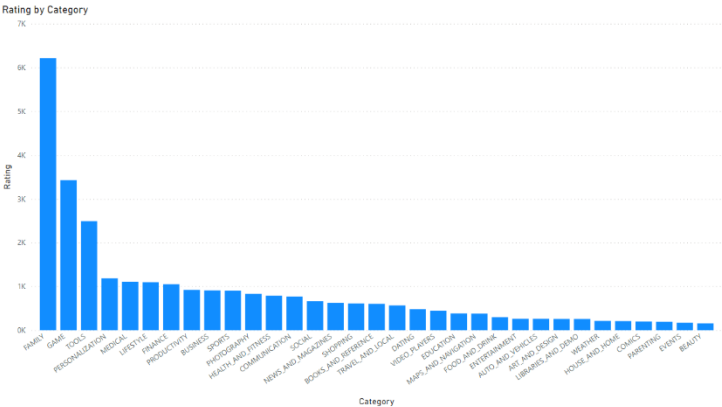
Installs by Rating



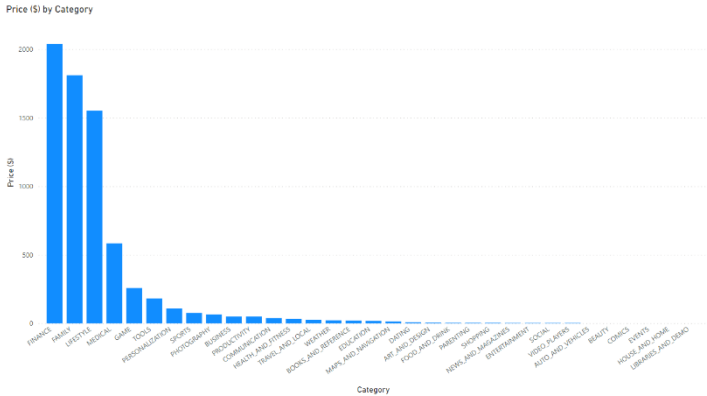
It is clear now from the chart above that the maximum installs (number of downloads) stems from the top rated apps in the Google App store data set. Rating indeed influences the number of users who will eventually install

an app. This better buttress the point that developers should work more on buiding quality brands that will in turn give the app better rating that focuses on only the number of downloads gathered from the app.

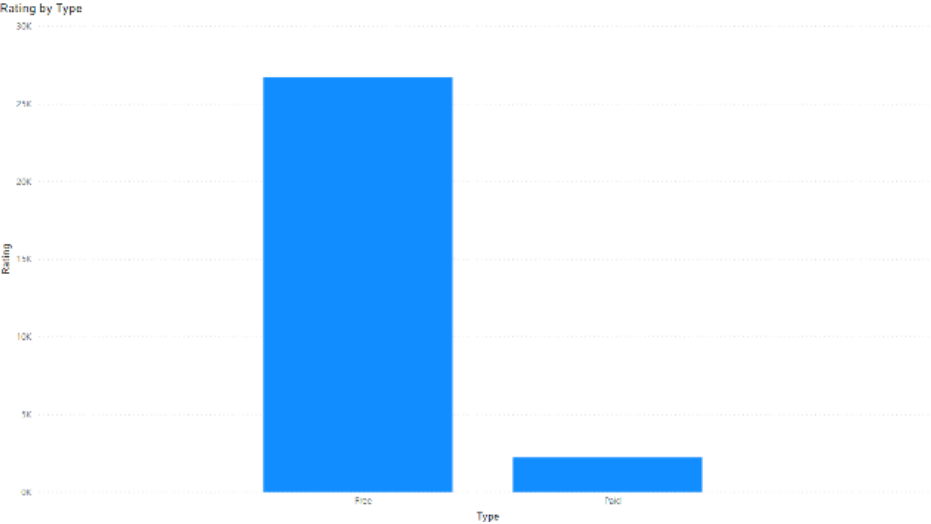
Rating by Category



Price by Category

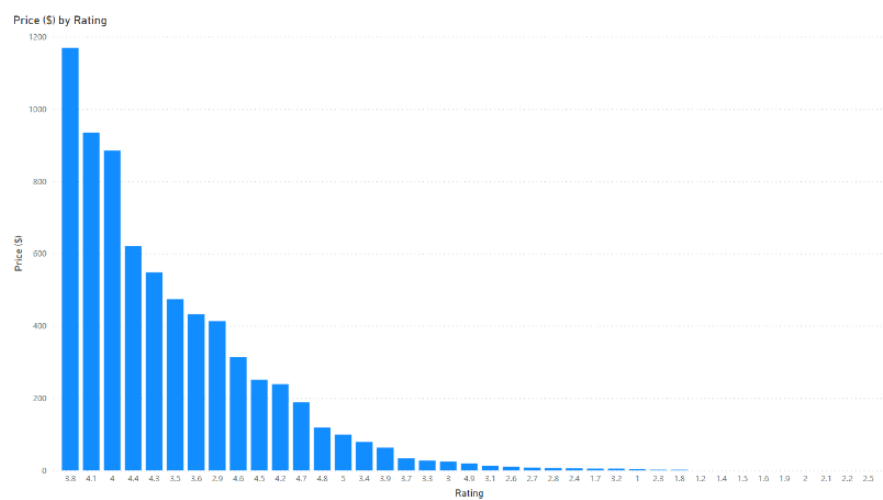


Rating by Type



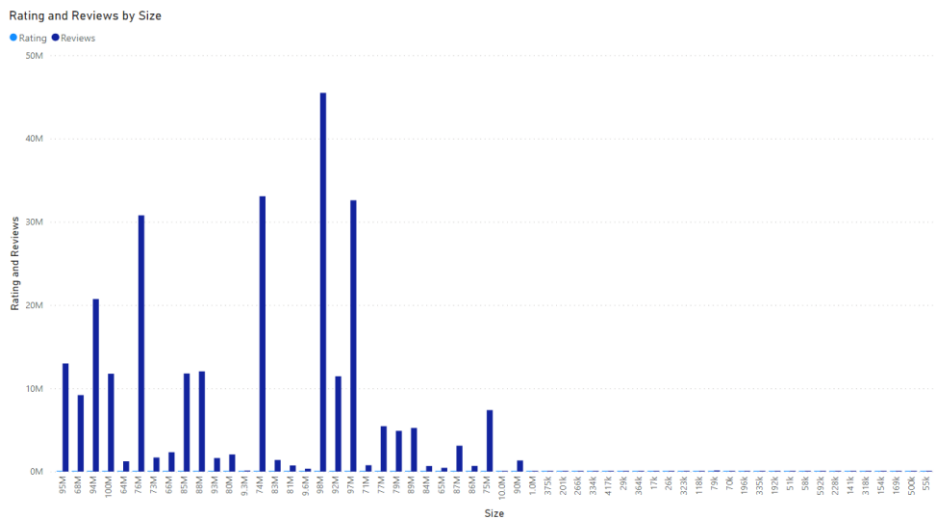
It is interesting to see how app category is greatly influenced by the app ‘Rating’ and ‘Price’. It also shows that paid users are very quick to rate apps, it is understandable that one will not relent to show its satisfaction or dissatisfaction after paying such a premium on an app, in the insights to come we will find out if the apps really were worth the dollar spent on them as we take look at specific ratings they all got.

Price (\$) by Rating



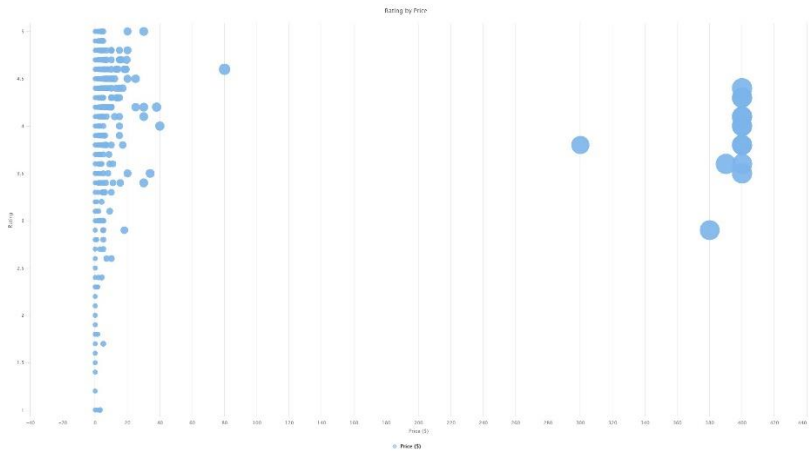
It is clearly seen that the majority of the paid apps have very high ratings which shows that the apps are really great apps and this means good business for the developer as more businesses are willing to put their Ad's at a cost on such apps. There will also be a very high in-app actions or activities on the app, and generating more revenues for the developer.

Rating and Reviews by Size



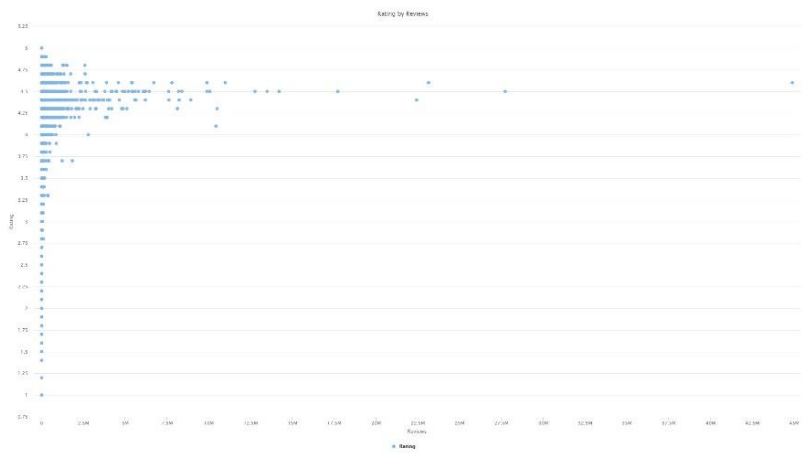
written codes otherwise known as legacy code etc. The cluster of well-rated apps occurs in the left corner signifying the importance of the small to medium size applications.

Rating by Price



The relationship between ratings and prices can be very vital to consider by the app developer. App developers need to know the various strategies to monetise their app so as to enable it to maximise profits. Though, the costs of apps depend on innovation, features, Ease of navigation, great user experience design, complexity and the market platform. The simple understanding of price dynamics will go a long way in helping the app developer to make the right decision.

Rating by Reviews



The chart shows that apps with high number of reviews usually have higher ratings. These are usually the apps which have a very high number of installs and are owned by bigger companies who try to improve upon their apps.

SUMMARY

The importance of quality over quantity in this study can not be over emphasized as seen across all the insights. From the insights that shows ‘installs by Rating’ in this analysis it is observed that over 90% of app users go through ratings and app reviews before they install an app.

Therefore, the app ratings and reviews should be top priority for an app developer for these following reasons;

- i) It increases the number of users who will eventually download the app and install it. With the insights already shown this proves to be true as a maximum number of installs is observed from the apps with over 4.5 ratings which in turn reflects the number of reviews.
- ii) As the app ratings and reviews increases, it also increases in-app revenue through Ad’s placements and promotional prospects. Not forgetting, as an app developer, the primary aim must be to make money which is the goal of every business.
- iii) The success of any app depends on your users’ loyalty and retention. The quality of the app and not quantity is a key factor. Quality app entails good User Experience Design that offers seamless navigation, great app features among others. When users are satisfied with an app, they are most likely to stay and give better app reviews.
- iv) The result of an excellent review and rating will also boast the apps visibility on the Google App store as the app will be among the top ranked app and making it easier for users to find. Having a good visibility will also push your app among the top apps in the google search engine.
- v) It offers a very high conversion rate as most users who end up installing an app go through ratings and reviews of that app. So as for the assurance of people downloading and installing an app, app ratings and reviews are extremely important to the app developer.

CONCLUSION

This is an analysis of the Google App store trends and the information shown by the charts in the course of this analysis is often useful to app developers. The analysis presented here is not enough for one to decide the effectiveness and success of any given app in the future, but it provides meaningful insights with simple charts for illustration. If the developer is uncertain about the worth of a given app category, the statistics provided can provide him with the range distribution. App developer priority among the app features would be 'Ratings' and 'Reviews', a developer cannot afford the frequency of negative reviews on their products as this could ruin their reputation in the long run if they plan on growing. App developers can experiment in new categories by implementing new monetization strategies, capturing a large proportion of the market, and moving from paid apps to free apps that would give one a broader audience to undertake new ideas.

The results show that 'Reviews', 'Installs', 'Rating', 'size', 'Type (Free or Paid)' and 'Category' are the foremost influential variables and have high impact on the amount of app downloads. Likewise, some genres are rated higher and are highly loved by the users. On the opposite hand, some genres are highly criticized by the users, users expect way more from the app, or the users are highly diverse from different backgrounds which leads in mixed or lower kind of ratings. The more a user opens an app and invests time or money into it, the more likely they are to give a high rating and helpful review. Whether or not a high-quality user isn't completely satisfied with an update or new feature, they often provide feedback that also highlights what proportion they enjoy in the app. Reviews like this are a social proof which will cause the next organic download rate. High Quality users are key, curating a marketing strategy that's built on the muse of quality, not quantity, will help improve visibility and increase revenue.

References

1. <https://www.linkedin.com/pulse/data-analysis-android-apps-google-play-using-python-iaroslav-kazakov/> Retrieved on 26th July, 2021
2. <https://www.irjet.net/archives/V7/i12/IRJET-V7I1248.pdf> Retrieved on 26th July, 2021
3. <https://medium.com/analytics-vidhya/cleaning-the-google-playstore-dataset-235d9754b8d>
Retrieved on 26th July, 2021
4. <https://customerthink.com/quality-vs-quantity-of-mobile-app-users/> Retrieved on 26th July, 2021
5. <https://www.linkedin.com/pulse/data-analysis-android-apps-google-play-using-python-iaroslav-kazakov/> Retrieved on 26th July, 2021
6. https://www.researchgate.net/publication/337547293_Identifying_the_influence_of_various_factor_of_apps_on_google_play_apps_ratings Retrieved on 26th July, 2021
7. <https://www.mobileappdaily.com/importance-of-mobile-app-reviews/amp> Retrieved on 26th July, 2021