



Google Play Store App Analysis

Bikon Kumar Das

Contents

- Why Google Play store?
- Exploring the dataset
- Candy, Vitamins or Painkillers? Maybe *addictive painkillers..*
- Exploring app categories
- Gaming Industry
- Untapped Category
- Paid apps vs Free Apps
- Rating and Reviews
- Predicting Installation Number
- Do I have enough space on my phone?
- Mystery of Editor's Choice
- Conclusion

Dataset

- Contains more than 1 million app data from google playstore
- 946k unique mobile app
- 19 unique variables
- For machine learning part, I have used 100k data due to lack of enough computational power



Why Google Playstore?

The mobile app industry has a value at 100 billion dollars and projected to grow 4 times by 2026. Everyone is thriving to create apps and take a bite out of this statistic. Google is bringing Android games to Windows in 2022.

The major factors driving the growth of the market are the increase in devices and the rise in customer base for the e-commerce industry. The fuel of the market is affected by the continuous growth in apps specifically focused on health and fitness, communication, social, and gaming apps. To make a profit, some apps are free to download and only require consumers, if they want, to purchase in-app products.

- If you are looking for **Drive, Growth and Retention**



Business Problem Scenario

Let's assume that you are an app developer. As a new developer, it is easy to download an app onto the google play store. With an one time fee of \$25 you can upload as many free apps as you want. Since about 3700 apps are releases daily, it's important to figure out what will make your apps stand out. Let's focus on what can contribute to the increase of installs on a specific app. Build a model that can determine the relationships between different features and the number of installs. ***How can developers make their apps more popular on Google App Store?***

Exploring The Dataset

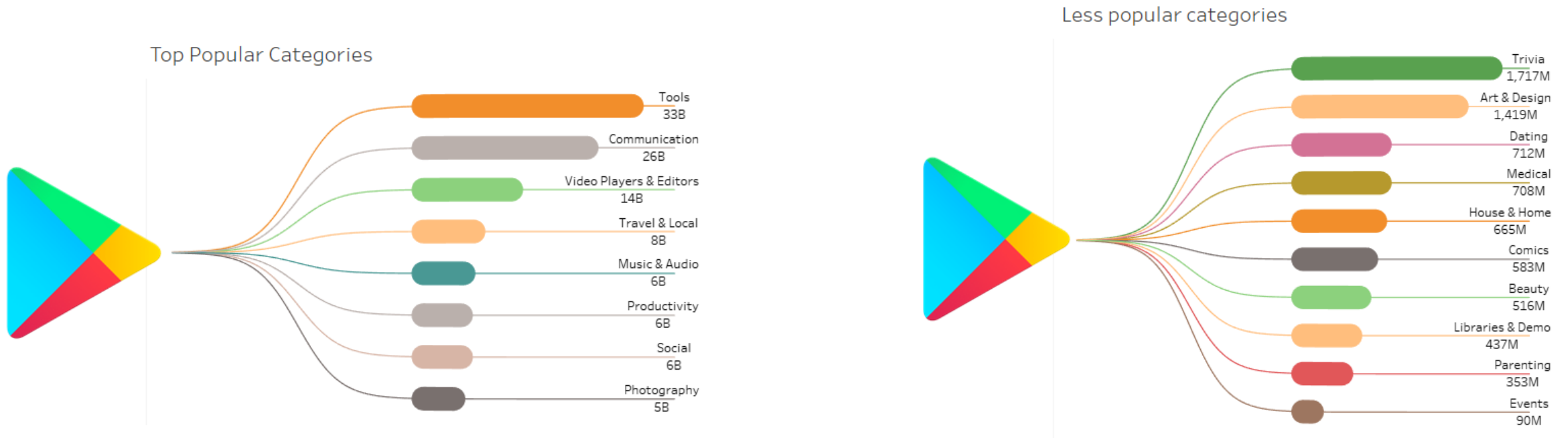
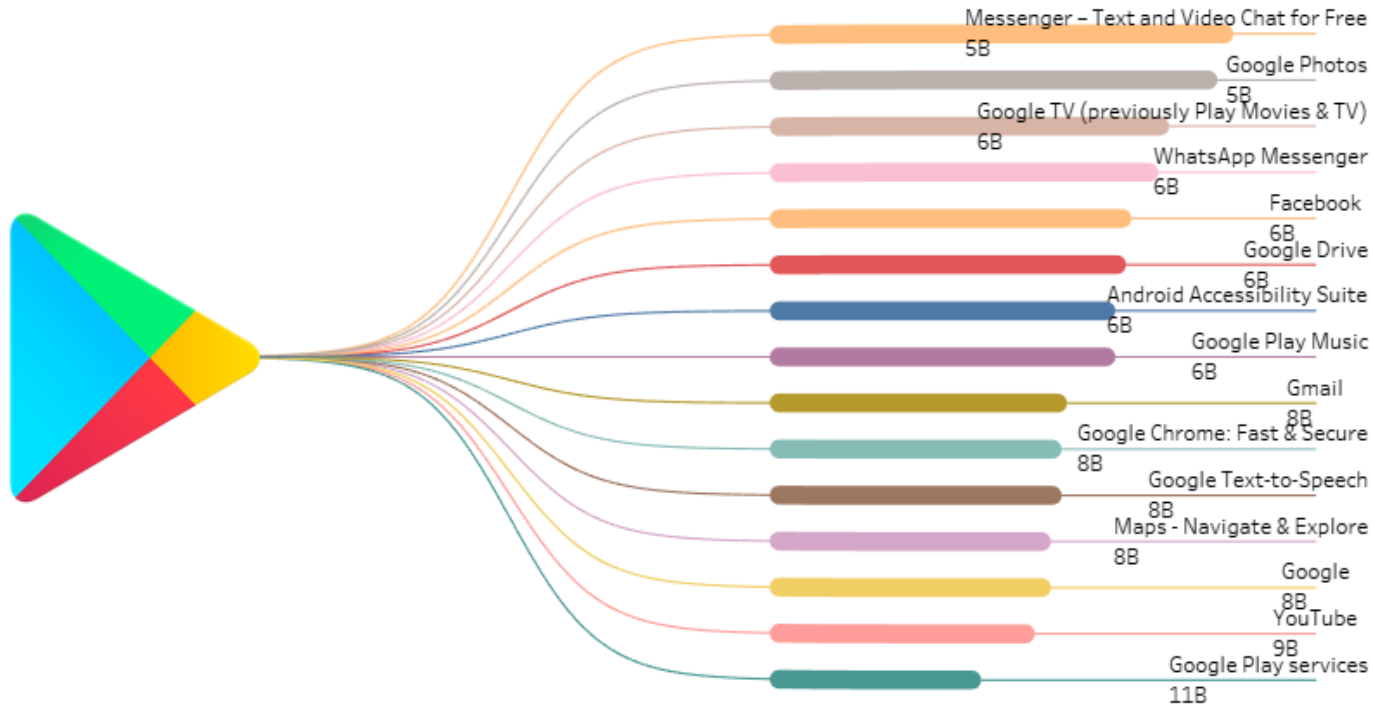


Tableau Dashboards of this analysis: 1)[Dashboard 1](#) 2)[Dashboard 2](#)

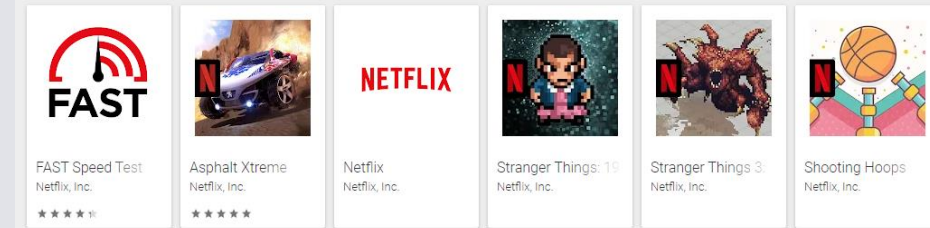
- Tools category is dominating playstore with more than 33B installs in total and Events category has the lowest number of installations
- Communication comes second with 26B installs and doing **justice to the purpose of mobile phone** which is mainly communication
- There are few **untapped Industry** we can find here. Like **Beauty, Medical, Comics** category which worth more than few billion dollars.

Most Popular Apps (Top 15)



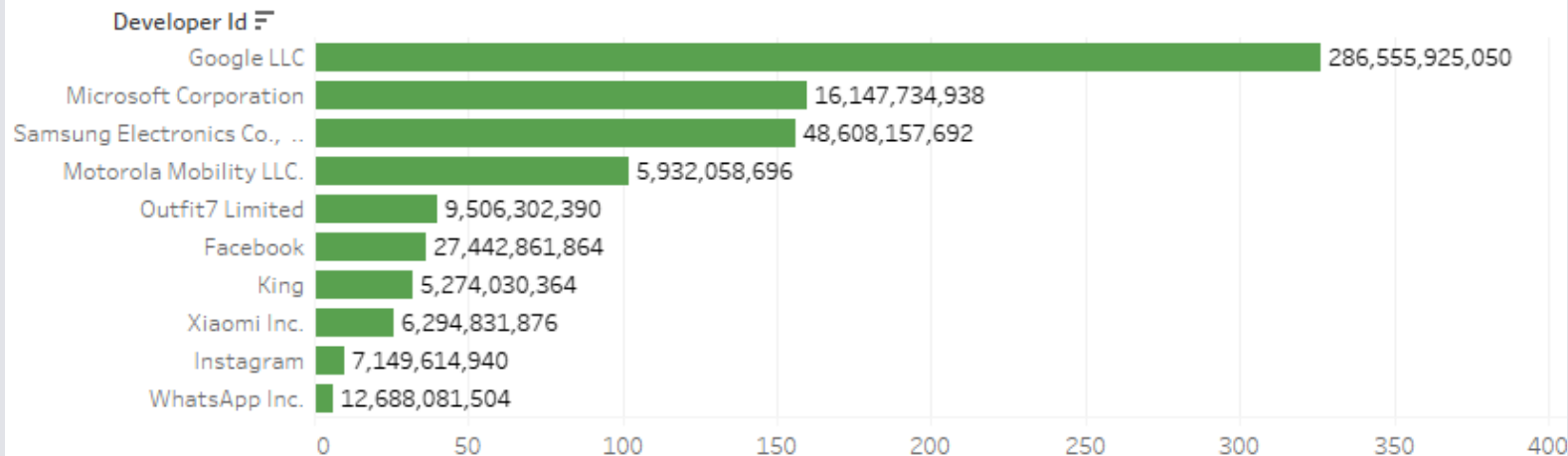
- 12 of the top 15 apps are from Google LLC and they have the most number of top downloaded apps
- Only 3 apps are from Facebook Inc.

More by Netflix, Inc.



#Insight 1: The feature of “More by [Company Name]” helps to get more downloads on playstore

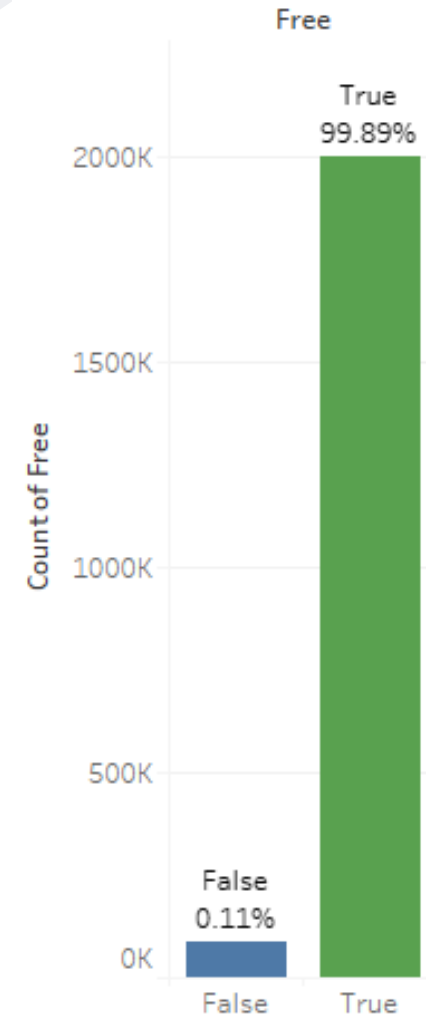
Top App Development Companies



Which dominates playstore?

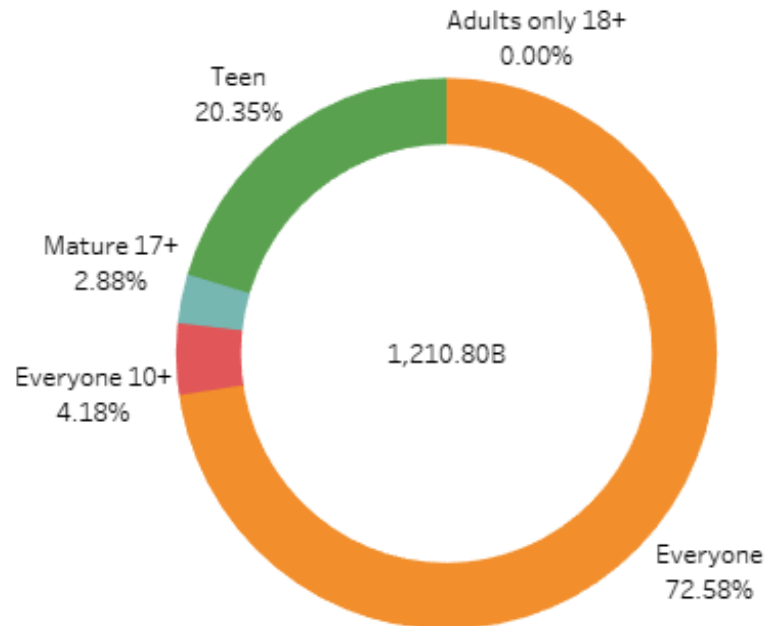
- Only 0.11% apps on playstore comes under paid app category.
- 99.89% of the apps are free to use and that is why the number of the apps on playstore in this category is more than 2000k.

It means people tend to install free apps most.



72% of the apps are made for every users on android. This is understandable since “Everyone” tend to mean kids, young adults and older people. Parents tend to let kids borrow their phones so not only do they have apps for themselves but they will download apps for their kids as well. Nowadays kids under 13 are getting more involved with technology from Leap Frog pads to owning their own phones/tablets.

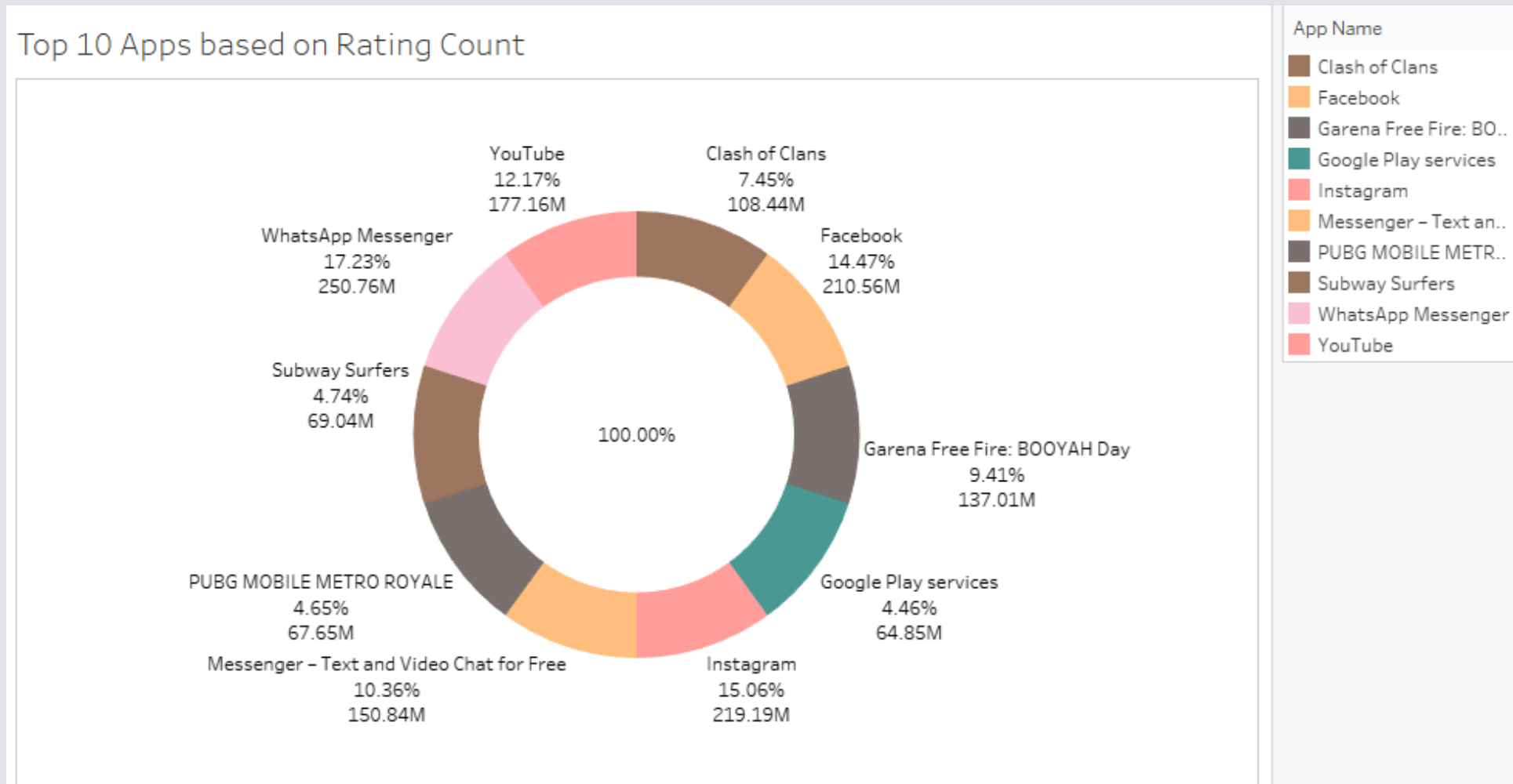
Content Rating



#Insight 2:

- Keep your app **free of cost**
- Targeted audience should be **“Everyone”** for higher reach

- We can see **4 games** among these top 10 apps based on rating count which are Clash of Clans, Garena Freefire, PUBG Mobile and Subway Surfers.



Candy, Vitamins or Painkillers? Maybe addictive painkillers..

#Why “Tool” ,“Communication” and “Gaming” category is dominating the play store?

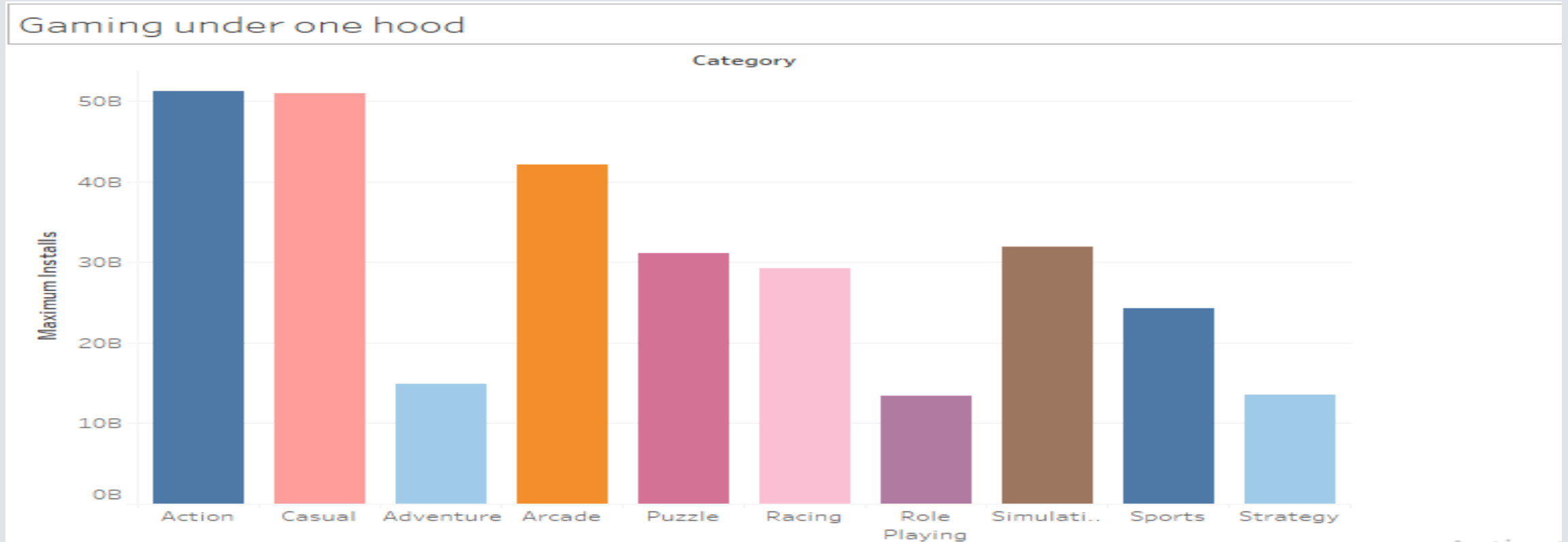
-Because they are the solutions of our problems. Gaming is like Candy, it is fun to play. Tools like google sheet/gmail is very much needed in our life. Messenger/Whatsapp is solving our communication problem. They work like vitamins. But, the main factor is if it is a addictive or not..

Advice:

#The more it solves our daily life problem, the more it gets installations.

So, my advice to a new developer is that the developer should solve real world problems by an app.

Gaming Industry



- For example, **casual** and **hypercasual games** are typically shorter in gameplay, and players don't really see themselves as gamers. They engage with rewarded video, and they typically don't mind interstitial ads. In some cases, banner ads work there as well.
- Gaming category has a few billion downloads and a huge market which is a green signal for game deveopers.

Insights of Gaming category

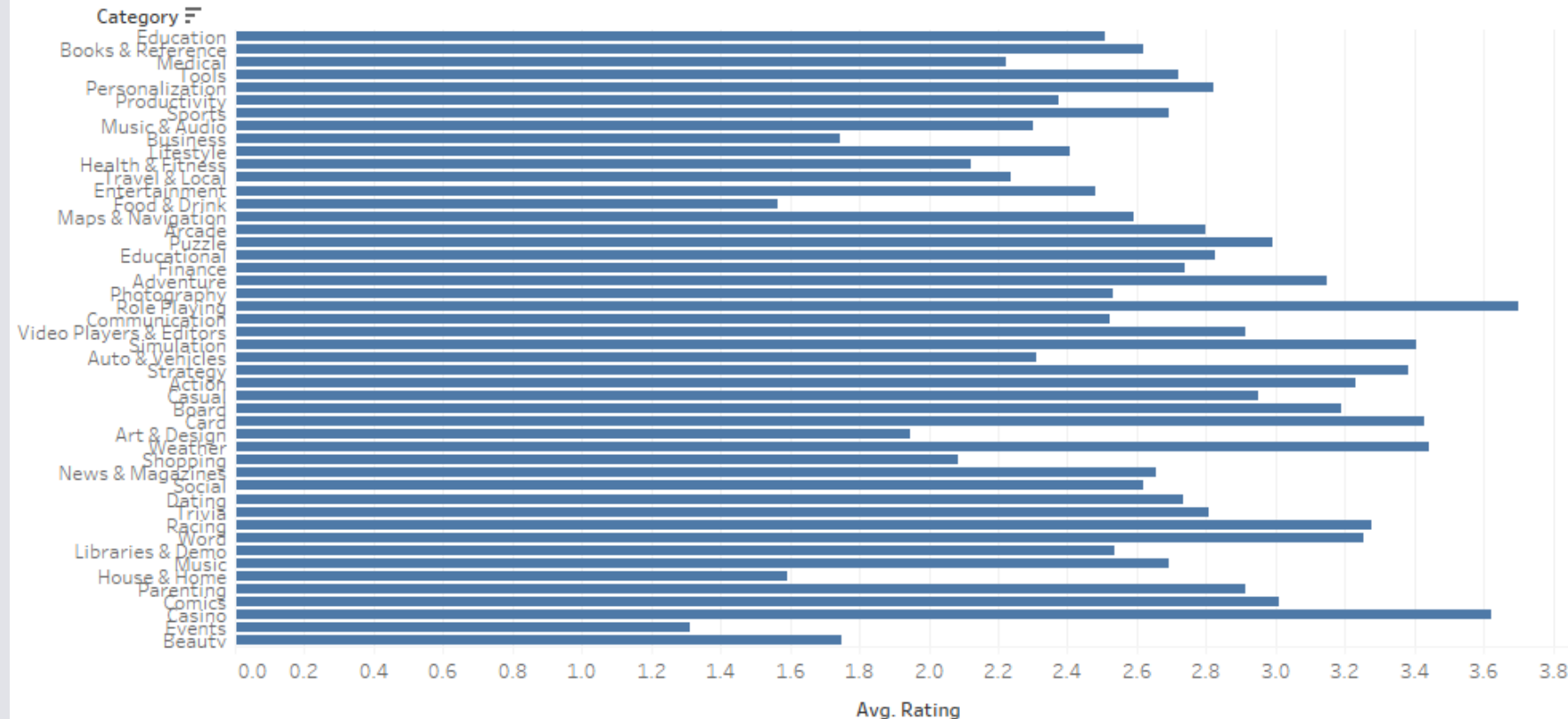
- The best way of achieving success with paid apps is to already have a following elsewhere and then transitioning the following to the mobile landscape a The category of app that was most successful at doing this is video games A **prebuilt following** also **enables developers to charge more** than the median app cost //
- Added benefit is you get a **fanbase to secure more revenue**

	App Name	Category	Maximum Installs	Price
1	Mind Games Pro	Puzzle	1128674	2.99
2	Where's My Wa...	Puzzle	3097437	1.99
3	The Room Two	Puzzle	1153973	1.99
4	Hitman GO	Puzzle	1343139	5.99
5	Lara Croft GO	Puzzle	1335605	5.99
6	Monument Valley	Puzzle	4394254	3.99

#You can charge a high price and can get a huge revenue also.

Untapped Categories

Category By AVG Rating



Beauty, Education, Finance are few of the promising categories on playstore. By looking at the avg rating, we can see there are few categories which are **still untapped and can grow in future.**

	App Name	Category	Maximum Installs
1	Google Pay (ol...	Finance	258358826
2	PayPal Mobile ...	Finance	164212774
3	Paytm -UPI, M...	Finance	270220054
4	???????? ?????	Finance	127580373
5	CAIXA	Finance	101414306
6	PhonePe UPI, ...	Finance	224971086

	App Name	Category	Maximum Installs
1	Photomath	Education	125948618
2	Duolingo: Lear...	Education	231646278
3	Samsung Globa...	Education	101652670
4	Google Classro...	Education	164456738
5	U-Dictionary: O...	Education	109993812

	App Name	Category	Maximum Installs
1	YouCam Nails -...	Beauty	15143771
2	Fashion Nails 3...	Beauty	15041411
3	Beauty Camera ...	Beauty	23427188
4	Beauty Makeup...	Beauty	22607870
5	Perfect365: One...	Beauty	52321809

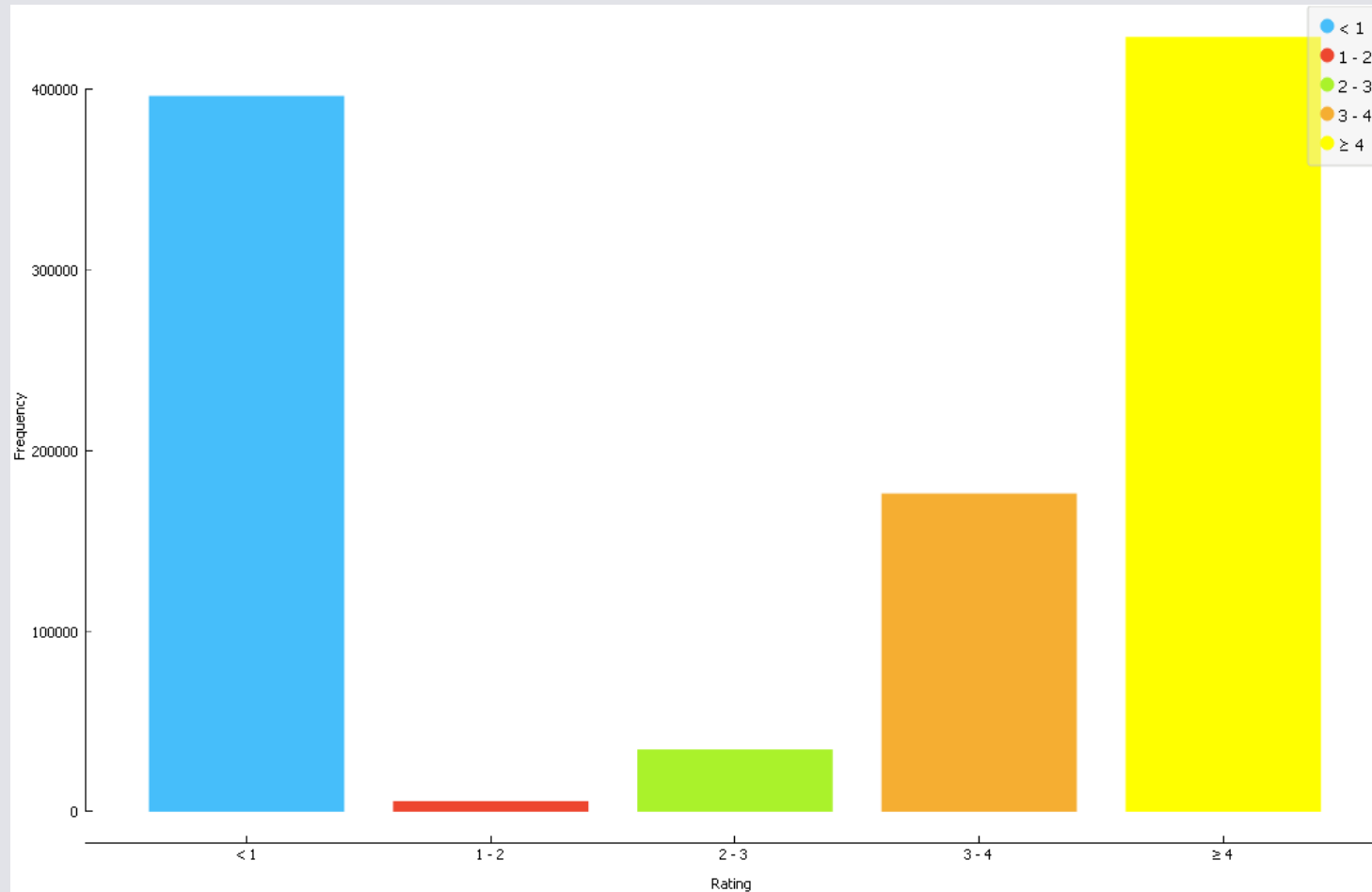
FinTech **Market** was valued at USD7301.78 **billion** in 2020.

Personal Finance category is growing vastly. There is a huge demand of **Mobile Financial Services** right now.

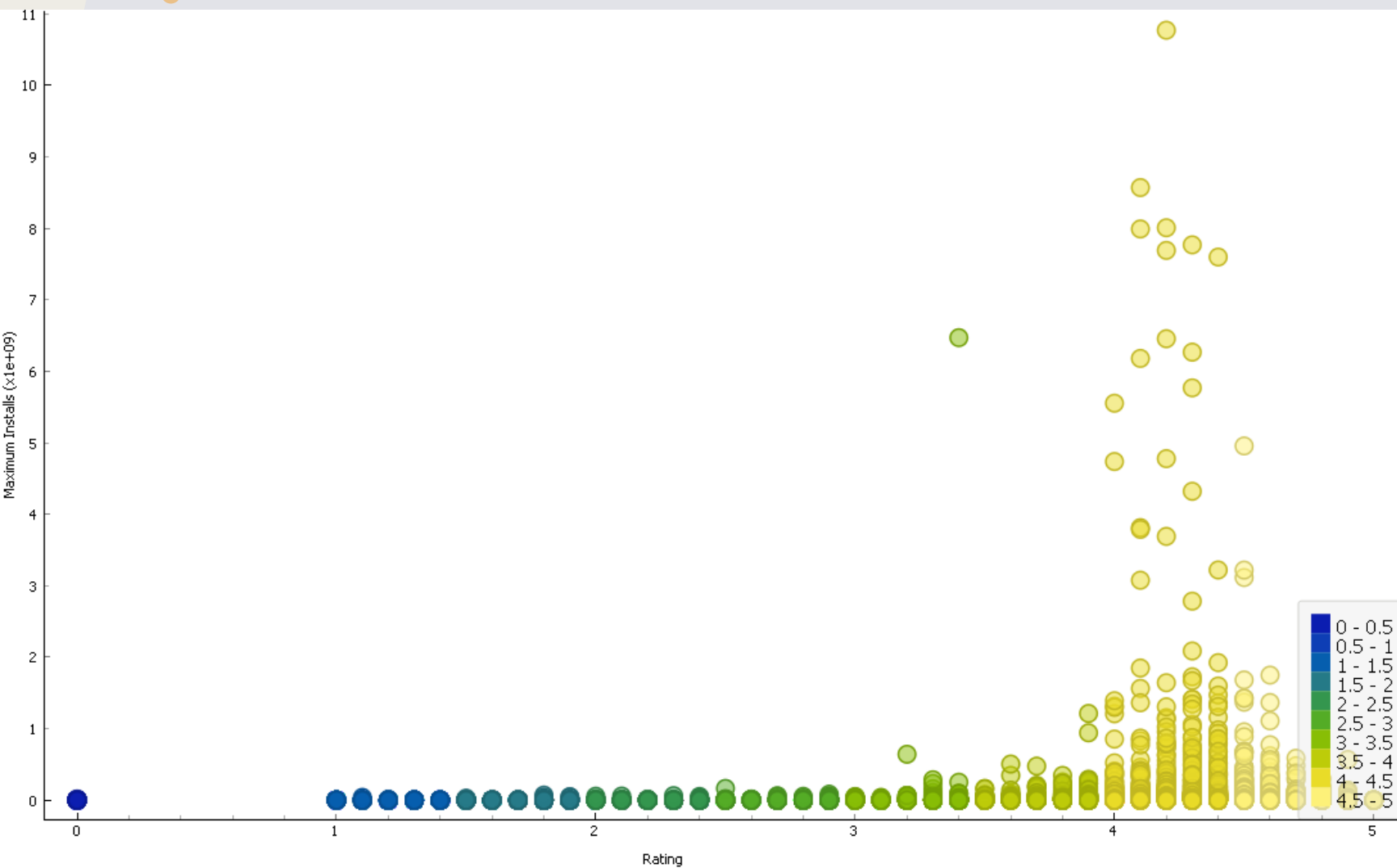
The Global EdTech & Smart Classroom Market size was estimated at **USD 83.61 Billion in 2020** and expected to reach USD 99.50 Billion in 2021. There is a huge demand of linguistics apps on playstore.

- The Beauty industry is worth over \$500 billion
- It is the category with the one of the fewest apps
- Developing an app that caters to the beauty industry has the potential for great success since the category has yet to receive large investments .

Ratings and Beyond



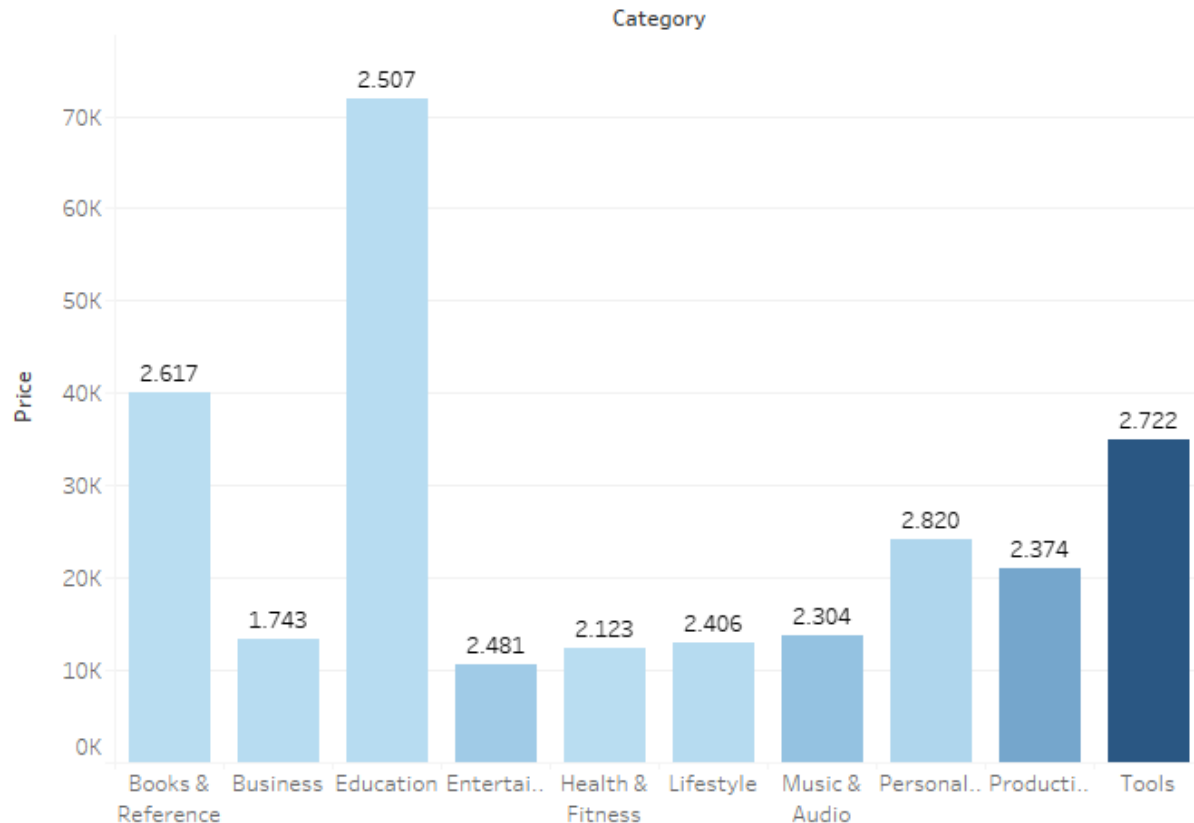
- Most of the apps on this dataset is rated equal or over 4.00
- The number of 1 or less rated app is very close to 4 rated apps which is quite alarming.
- There is a significant number we can see as between 3 to 4 rated apps.



#If the Rating is between 4-5 which can be counted as positive ratings, are tend to get more number of installs. So, it directly correlated with Number of installations.

Free vs Paid

Tools has the highest Avg Rating (paid category)

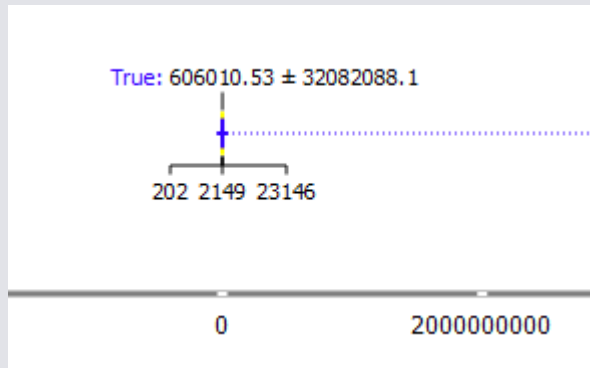


- Tools is the highest rating for paid apps category
- Adult Category has no Paid app
- 8253 is the median of Minimum Installs for Paid apps and has 2.19 average rating
- 359229 is the median of Minimum Installs for Free apps and 2.54 average rating
- **In top 100 apps, there are only 30 apps which has in app purchases**

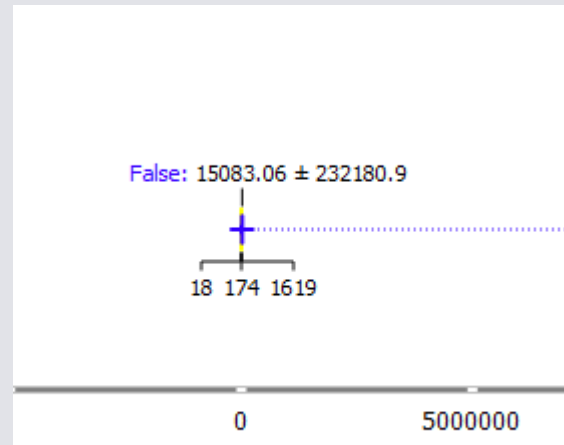
- When designing an app it's important to consider the amount of people that could be interested in downloading your app
- When designing for a niche market it's better to make the app cost money up front a Buyers are more willing to pay since there are less resources available

Predicting Number of installs

Median of Installation for Free app



Median of Installation for Paid app



Silhouette Scores	
2	0.987
3	0.586
4	0.608
5	0.613
6	0.617
7	0.591
8	0.603

Pearson correlation		
Installs		
Filter ...		
1	+1.000	Installs
2	+0.947	Installs
3	+0.514	Installs
4	+0.013	Installs
5	+0.009	Installs
6	-0.001	Installs
		Minimum Installs
		Maximum Installs
		Rating Count
		Rating
		Size
		Price

The number of installation has a positive correlation with Minimum installs followed by Maximum install, Rating count, Size and Price.

Evaluation Results

Model	MSE	RMSE	MAE	R2
kNN	48096212500005.680	6935143.293	96225.098	0.448
Stack	0.000	0.000	0.000	1.000
Random Forest	15963956325319.479	3995492.000	44045.033	0.817
Linear Regression	0.000	0.000	0.000	1.000

Model Comparison by RMSE

	kNN	Stack	Random Forest	Linear Regressi...
kNN		0.864	0.873	0.864
Stack	0.136		0.149	0.908
Random Forest	0.127	0.851		0.851
Linear Regression	0.136	0.092	0.149	

Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

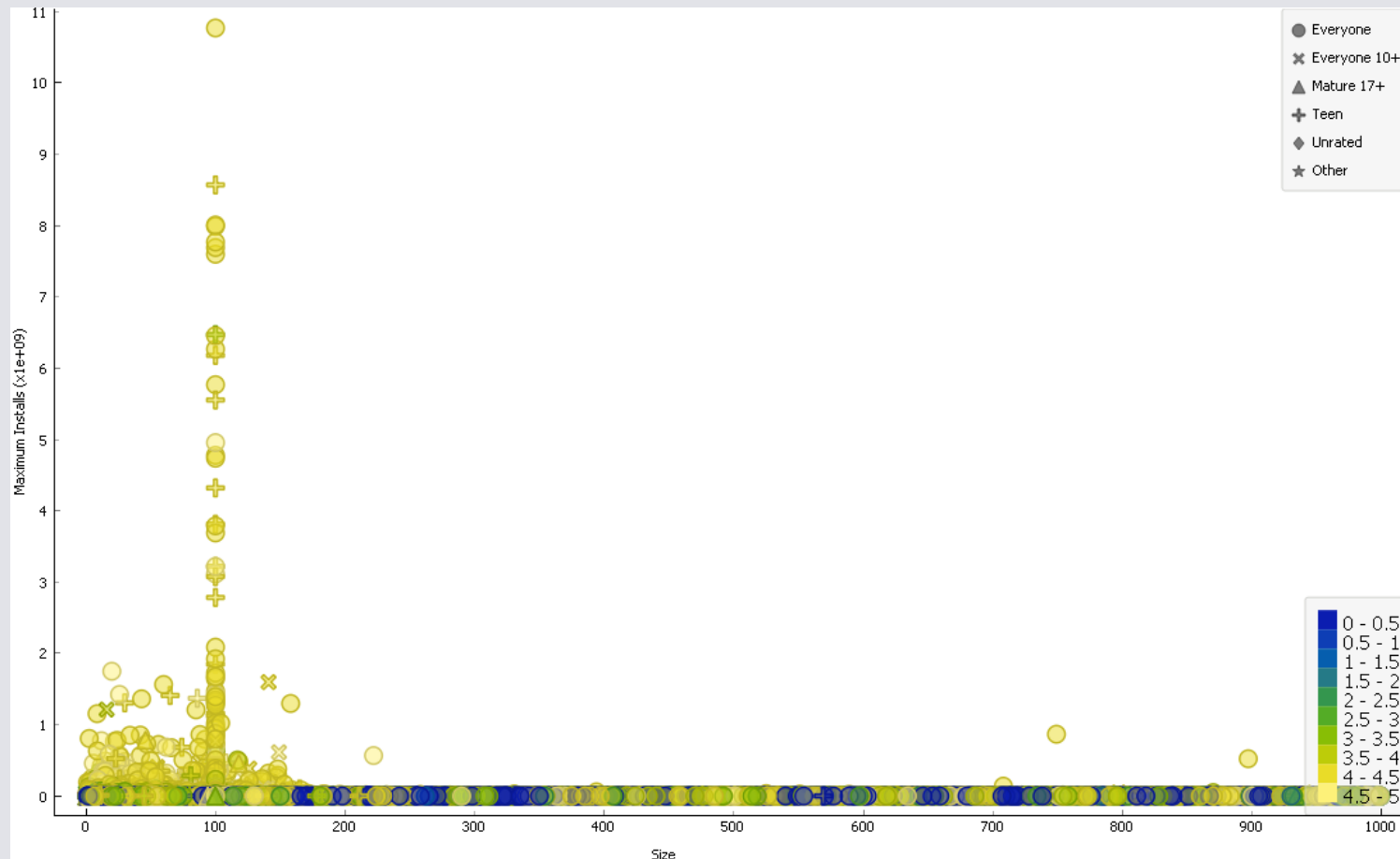
Here, I have tried to predict the the number of installs with KNN, Stack, Random Forest and Linear Regression algorithm.

Stack and Linear Regression overfits the model but **Random Forest works good for this model.**

Based on the model recommendations..

- **Reviews** are crucial to getting more installs. Since word of mouth is very powerful when marketing a product, it's best to focus on getting as many people talking about your products.
- **Updating** can lead to getting more people to stay with your app. Updates can help with fixes and bugs that an app might have, and people seem to love when an app has been updated recently. 'Out with the old, in with the new'.
- **Size** can affect if people have the necessary room to house your app. Though having a large size can represent having more content, it can possibly affect the performance on some platforms.
- **Price** can affect if a person will download an app. Consumers will most likely see if they can get a free app over a paid app.
- **Rating** can be a factor if an app is rated higher than the other. Higher ratings tend to mean the app is good. It also affects the ranking of an app in the app store.

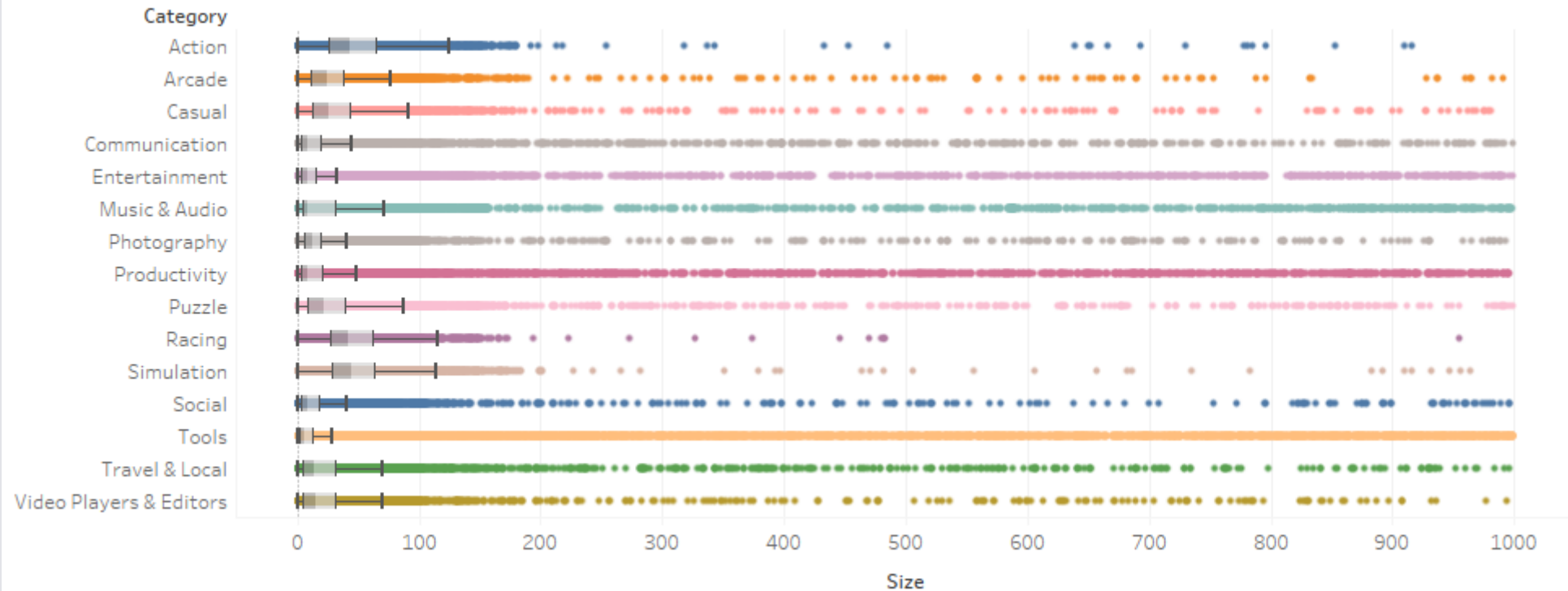
Do I have enough space on my phone?



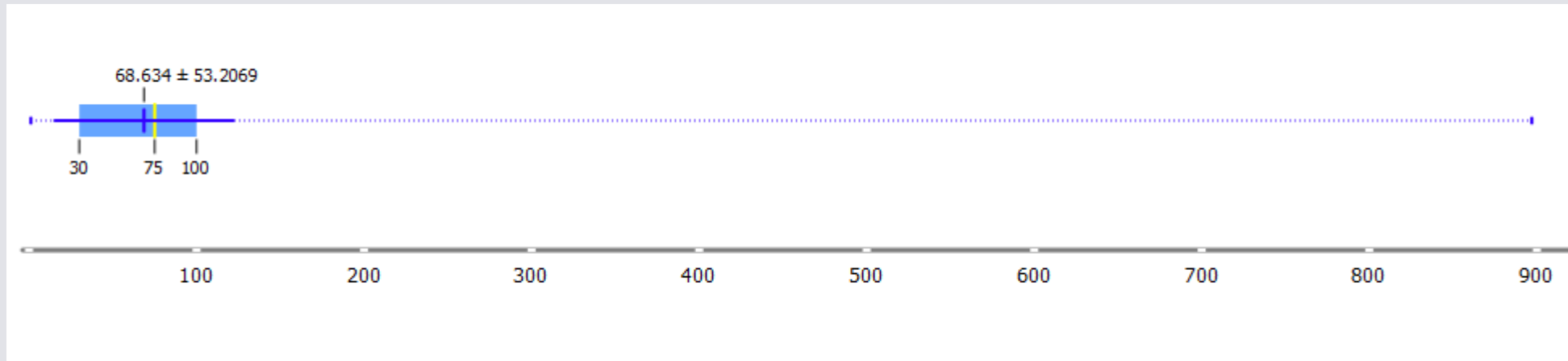
- Lower the Size, better the installation number
- In Everyone and Teen category, we can see that lower size has a big impact over there

Category-wise size via Boxplot

Category-wise Size



Size is a great predictor for Installs. Majority of the apps installed are no more than 200 megabytes in size. This shows that size does matter. Most consumers don't want to download apps that takes up too much space on their device storage. Moreover, Google only permits apps that are less than 500mb. You can still upload a bigger app but it will be broken into two parts. Diving deeper you can see that Gaming apps are between 25-100 megabytes, and they have the most apps that are big in size. This could possibly be the result of updates happening in most games

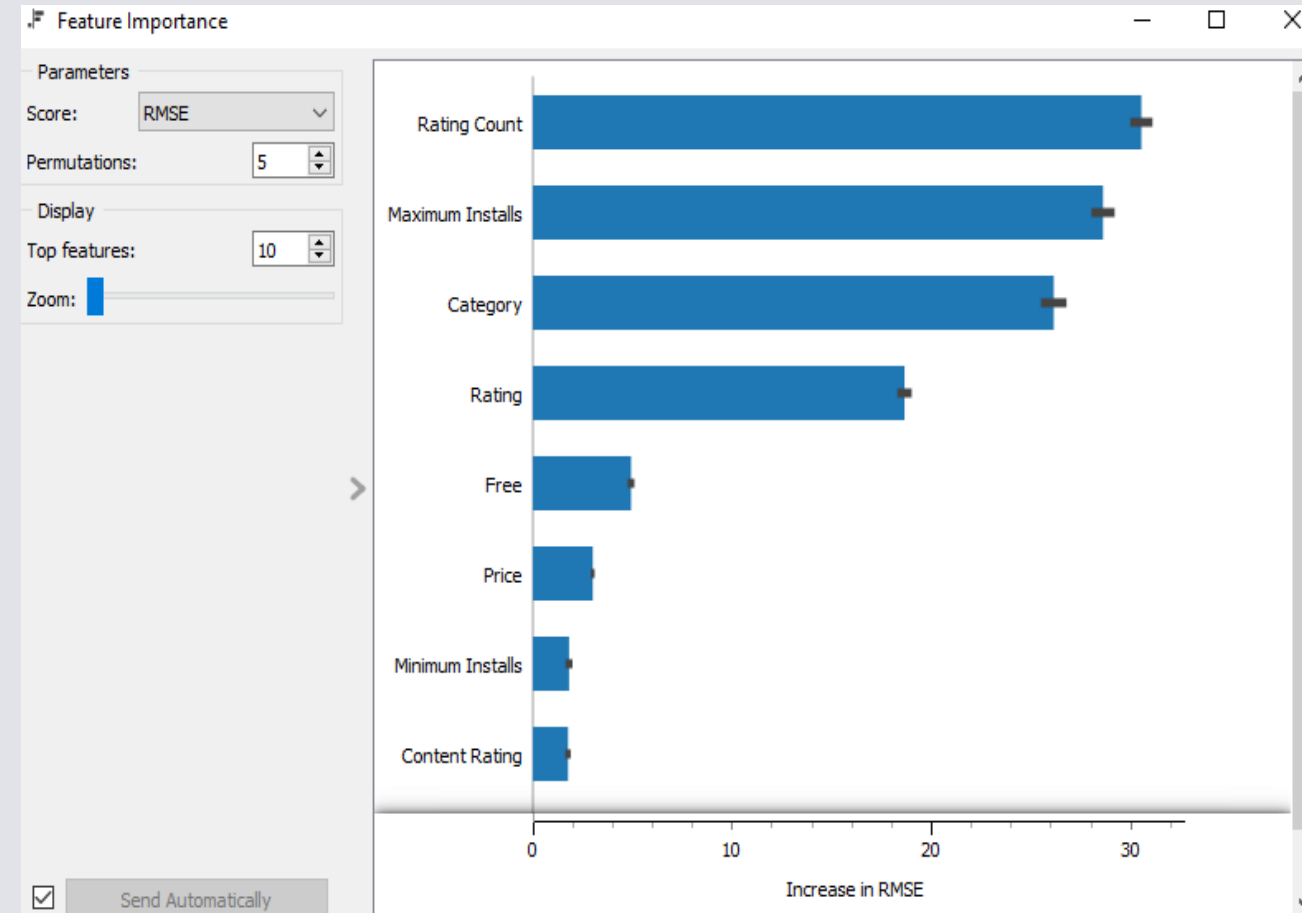


This is the median size for apps which got more than **10 million downloads**

Predicting Size

	kNN	Random Forest	Linear Regression	Stack	Size	Released
1	15.15	21.9912	22.6739	25.2...	4	12/15/2017
2	15....	9.36667	22.3417	24.5...	16	1/26/2018
3	21....	42.6524	29.8515	30.3...	76	8/20/2019
4	31....	48.9365	44.8624	33.7...	56	3/23/2020
5	122.8	58.4951	26.699	27.8...	4.4	12/4/2018
6	27....	80.255	43.6302	33.1...	81	6/10/2019
7	21....	9.04192	28.7796	28.0...	3.8	9/1/2014
8	18.12	40.956	24.5725	26.5...	6.3	7/24/2017
9	26....	20.5529	25.3103	25.7...	6.1	5/2/2020
10	34.9	14.174	25.6755	25.1...	8.1	9/13/2017
11	16.23	30.1842	25.3102	28.2...	25	11/6/2020
12	23.66	4.61855	27.5613	25.8...	2.7	5/1/2013
13	56.5	15.3632	36.0612	29.9...	3.9	11/14/2017
14	42	8.66121	22.2417	24.1...	2.5	11/12/2020

Model	MSE	RMSE	MAE	R2
kNN	4719.573	68.699	27.209	0.059
Random Forest	1950.087	44.160	16.417	0.611
Linear Regression	4915.655	70.112	26.660	0.020
Stack	4462.486	66.802	25.889	0.110

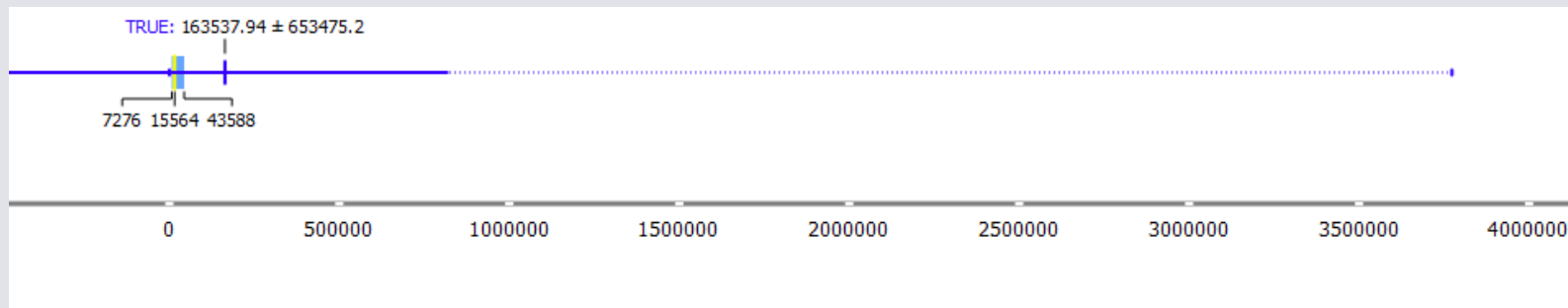


Here, Random Forest is having the lowest RMSE and the highest R2 value. Rating count is the most important feature for this model followed by Maximum install, Category, Rating, Free, Price, Minimum Installs.

Mystery of Editor's Choice

	App Name	Minimum Android	Developer Id	Released	Last Updated	Category	Rating	Rating Count	Installs	Minimum Installs
1	Slack	Varies with device	Slack Technolo...	7/22/2013	11/9/2020	Business	4.1	89946	10000000	10000000
2	Podcast Player	4.4 and up	Castbox.FM - R...	1/4/2016	10/27/2020	News & Magazi...	4.6	213624	10000000	10000000
3	Bubble Witch 3 ...	4.1 and up	King	1/11/2017	11/16/2020	Puzzle	4.6	2441529	50000000	50000000
4	Hungry Shark E...	4.1 and up	Ubisoft Entertai...	2/22/2013	10/27/2020	Arcade	4.4	7010986	100000000	100000000
5	MMX Hill Dash ...	5.0 and up	Hutch Games	4/23/2018	11/13/2020	Racing	4.1	104241	10000000	10000000
6	Top Drives Car ...	6.0 and up	Hutch Games	8/29/2017	11/10/2020	Racing	4.4	276478	10000000	10000000
7	BASEBALL 9	4.1 and up	playus soft	5/15/2018	11/23/2020	Sports	4.4	204134	10000000	10000000
8	Captain Tsubas...	4.4 and up	KLab	12/5/2017	11/20/2020	Sports	3.9	158081	5000000	5000000
9	DIGIMON ReArise	5.0 and up	BANDAI NAMC...	10/9/2019	10/2/2020	Role Playing	3.9	75112	1000000	1000000
10	MazM: The Pha...	4.4 and up	MazM	3/29/2018	11/11/2020	Adventure	4.6	12542	500000	500000
11	OPUS: Rocket o...	4.1 and up	Sigono Inc.	9/13/2017	6/16/2020	Adventure	4.5	52939	1000000	1000000
12	Florence	5.0 and up	Annapurna Inte...	3/5/2018	6/17/2020	Puzzle	4.7	11656	50000	50000

The number of installation is not related to getting into editor's choice list. 12th app of this list got only 5000 installs but still made it to it's way to editor's choice list. You can get it to with lower amount of installation!

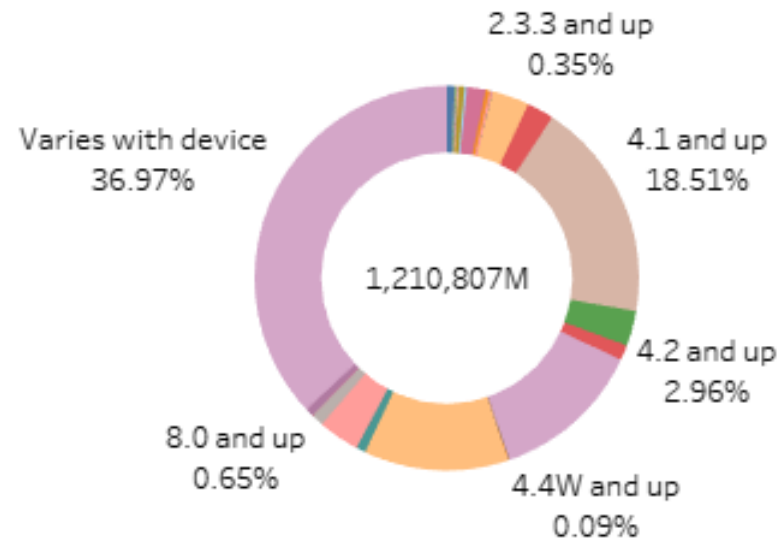


163537.94 Median of rating count for Free apps which made to the choice list



1643432.84 Median of rating count for Free apps which made to the choice list

Does Android Version Matter?



The apps with “Varies with device” has the most installations .Developer needs to keep it in mind when developing an app, so that the app can get a **large number of audience**.

Conclusion

- Find a way to market your app. It is a great way to get more installs which can lead to more reviews. If those reviews happen to be positive it can increase the chances of new consumers noticing/downloading your app.
- Updates - Keep your app updated and free of pesky bugs. Keep fresh and new content.
- When creating an app, it's best to research and see what category is popular in the market. So focusing on family and gaming apps could be boost the amount of installs and create a buzz for your company. However, if we're interested in getting more installs we could focus on communications apps since that's the main reason of having a phone in the first place.
- When creating apps, it best to manage the size of the app. Keeping the app under 200 megabytes, let's the consumer know that your app has substance but it won't take up too much of their device storage.

The slide features a light blue background with decorative elements in the corners. The top-left corner contains several small circles in yellow, orange, and green. The top-right corner features larger green circles and a small orange dot. The bottom-right corner is decorated with a cluster of circles in teal, yellow, and brown. The text "Thank You!" is centered in a dark blue, sans-serif font.

Thank You!