

What App is That?

Alex Yom

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

Overview

- Introduction
- Dataset
- Analysis
- Discussion

Introduction

- With new cell phone applications coming out everyday, is there a way to predict a successful app? This project will seek to find the driving factors of an app's success and popularity.
- App developers and tech companies would be able to use this information to target a specific criteria for new potential apps.
- My initial hypothesis is that factors such as the specific category of an app will have an impact on the amount of times the app is downloaded.

Dataset

- This dataset consists of all the apps in the Google Play Store. Along with the name of the app, the category, number of installs, payment method, content rating, etc are included.

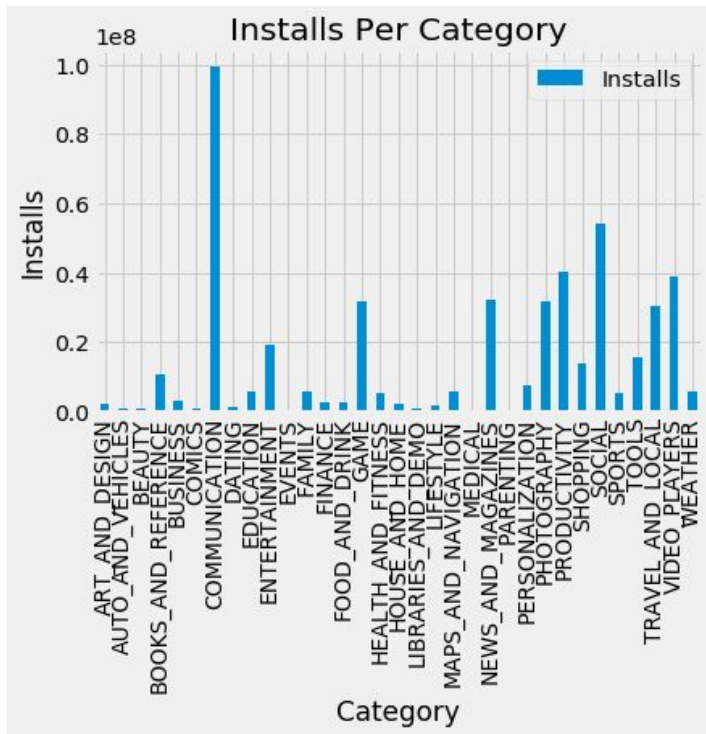
1	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content R	Genres	Last Upda	Current V	Android Ver
2	Photo Editor & Candy Camera	ART_AND_DESIGN	4.1	159	19M	10,000+	Free		0 Everyone	Art & Desi	7-Jan-18	1.0.0	4.0.3 and up
3	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free		0 Everyone	Art & Desi	15-Jan-18	2.0.0	4.0.3 and up
4	U Launcher Lite â€” FREE Live	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free		0 Everyone	Art & Desi	1-Aug-18	1.2.4	4.0.3 and up
5	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000	Free		0 Teen	Art & Desi	8-Jun-18	Varies wit	4.2 and up
6	Pixel Draw - Number Art Color	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free		0 Everyone	Art & Desi	20-Jun-18	1.1	4.4 and up
7	Paper flowers instructions	ART_AND_DESIGN	4.4	167	5.6M	50,000+	Free		0 Everyone	Art & Desi	#####	1	2.3 and up
8	Smoke Effect Photo Maker - Si	ART_AND_DESIGN	3.8	178	19M	50,000+	Free		0 Everyone	Art & Desi	#####	1.1	4.0.3 and up
9	Infinite Painter	ART_AND_DESIGN	4.1	36815	29M	1,000,000+	Free		0 Everyone	Art & Desi	14-Jun-18	6.1.61.1	4.2 and up
10	Garden Coloring Book	ART_AND_DESIGN	4.4	13791	33M	1,000,000+	Free		0 Everyone	Art & Desi	#####	2.9.2	3.0 and up
11	Kids Paint Free - Drawing Fun	ART_AND_DESIGN	4.7	121	3.1M	10,000+	Free		0 Everyone	Art & Desi	3-Jul-18	2.8	4.0.3 and up
12	Text on Photo - Fontee	ART_AND_DESIGN	4.4	13880	28M	1,000,000+	Free		0 Everyone	Art & Desi	27-Oct-17	1.0.4	4.1 and up

- When the dataset is ranked by the most popular apps being first, the communication apps dominate the top 25.
- 9 of the top 25 apps are communication apps while the news and magazine apps have 4 apps within the top 25.

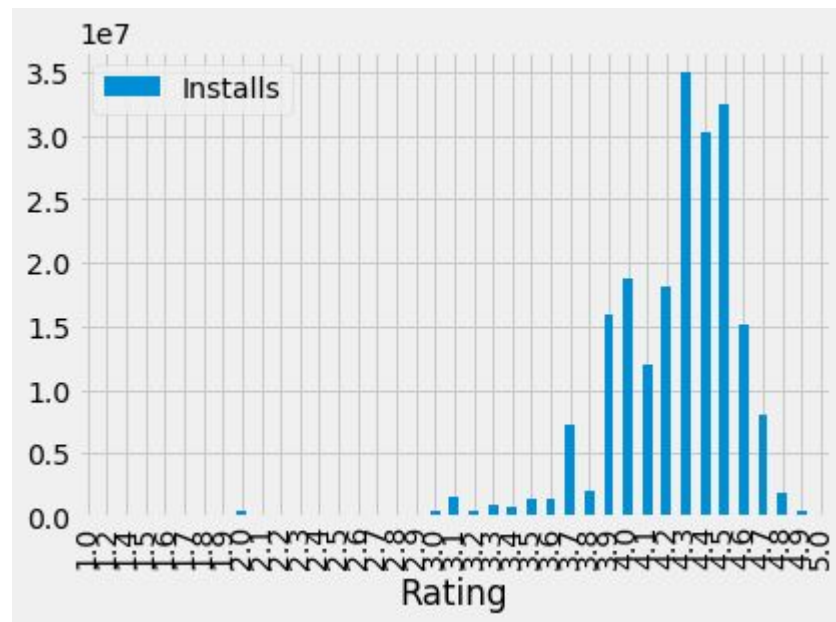
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Update	Current Version	Android Version	
2	451	Gmail	COMMUNICATION	4.3	4604483	Varies with device	1E+09	Free	0	Everyone	Communication	2-Aug-18	Varies with device	Varies with device	
3	3665	YouTube	VIDEO_PLAYERS	4.3	25655305	Varies with device	1E+09	Free	0	Teen	Video Players & Editors	2-Aug-18	Varies with device	Varies with device	
4	382	Messenger	COMMUNICATION	4	56646578	Varies with device	1E+09	Free	0	Everyone	Communication	1-Aug-18	Varies with device	Varies with device	
5	386	Hangouts	COMMUNICATION	4	3419433	Varies with device	1E+09	Free	0	Everyone	Communication	21-Jul-18	Varies with device	Varies with device	
6	391	Skype - free IM & video calls	COMMUNICATION	4.1	10484169	Varies with device	1E+09	Free	0	Everyone	Communication	3-Aug-18	Varies with device	Varies with device	
7	3687	Google Play Movies & TV	VIDEO_PLAYERS	3.7	906384	Varies with device	1E+09	Free	0	Teen	Video Players & Editors	6-Aug-18	Varies with device	Varies with device	
8	411	Google Chrome: Fast & Secure	COMMUNICATION	4.3	9643041	Varies with device	1E+09	Free	0	Everyone	Communication	1-Aug-18	Varies with device	Varies with device	
9	3736	Google News	NEWS_AND_MAGAZINES	3.9	877635	13M	1E+09	Free	0	Teen	News & Magazines	1-Aug-18	5.2.0	4.4 and up	
10	3765	Google News	NEWS_AND_MAGAZINES	3.9	877635	13M	1E+09	Free	0	Teen	News & Magazines	1-Aug-18	5.2.0	4.4 and up	
11	865	Google Play Games	ENTERTAINMENT	4.3	7165362	Varies with device	1E+09	Free	0	Teen	Entertainment	16-Jul-18	Varies with device	Varies with device	
12	3816	Google News	NEWS_AND_MAGAZINES	3.9	877643	13M	1E+09	Free	0	Teen	News & Magazines	1-Aug-18	5.2.0	4.4 and up	
13	464	Hangouts	COMMUNICATION	4	3419513	Varies with device	1E+09	Free	0	Everyone	Communication	21-Jul-18	Varies with device	Varies with device	
14	468	Skype - free IM & video calls	COMMUNICATION	4.1	10484169	Varies with device	1E+09	Free	0	Everyone	Communication	3-Aug-18	Varies with device	Varies with device	
15	3896	Subway Surfers	GAME	4.5	27711703	76M	1E+09	Free	0	Everyone 10+	Arcade	12-Jul-18	1.90.0	4.1 and up	
16	3904	WhatsApp Messenger	COMMUNICATION	4.4	69109672	Varies with device	1E+09	Free	0	Everyone	Communication	3-Aug-18	Varies with device	Varies with device	
17	3909	Instagram	SOCIAL	4.5	66509917	Varies with device	1E+09	Free	0	Teen	Social	31-Jul-18	Varies with device	Varies with device	
18	3928	YouTube	VIDEO_PLAYERS	4.3	25623548	Varies with device	1E+09	Free	0	Teen	Video Players & Editors	2-Aug-18	Varies with device	Varies with device	
19	381	WhatsApp Messenger	COMMUNICATION	4.4	69119316	Varies with device	1E+09	Free	0	Everyone	Communication	3-Aug-18	Varies with device	Varies with device	
20	2808	Google Photos	PHOTOGRAPHY	4.5	10858556	Varies with device	1E+09	Free	0	Everyone	Photography	6-Aug-18	Varies with device	Varies with device	
21	9844	Google News	NEWS_AND_MAGAZINES	3.9	878065	13M	1E+09	Free	0	Teen	News & Magazines	1-Aug-18	5.2.0	4.4 and up	
22	3565	Google Drive	PRODUCTIVITY	4.4	2731211	Varies with device	1E+09	Free	0	Everyone	Productivity	6-Aug-18	Varies with device	Varies with device	
23	152	Google Play Books	BOOKS_AND_REFERENCE	3.9	1433233	Varies with device	1E+09	Free	0	Teen	Books & Reference	3-Aug-18	Varies with device	Varies with device	
24	5856	Google Play Games	FAMILY	4.3	7168735	Varies with device	1E+09	Free	0	Teen	Entertainment	16-Jul-18	Varies with device	Varies with device	
25	1700	Subway Surfers	GAME	4.5	27723193	76M	1E+09	Free	0	Everyone 10+	Arcade	12-Jul-18	1.90.0	4.1 and up	
26	2884	Google Photos	PHOTOGRAPHY	4.5	10859051	Varies with device	1E+09	Free	0	Everyone	Photography	6-Aug-18	Varies with device	Varies with device	

Analysis

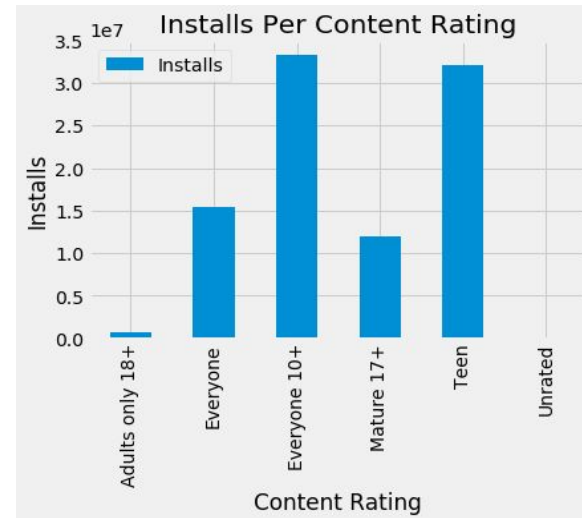
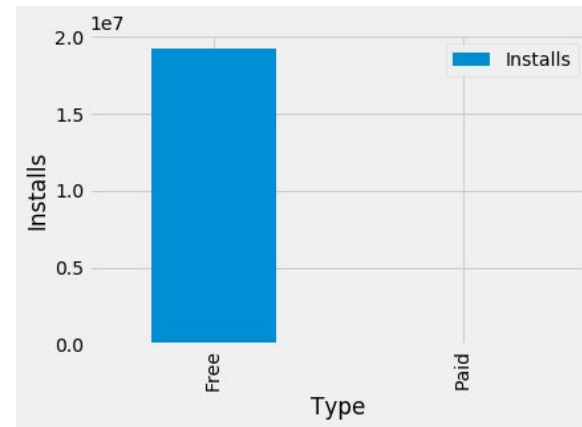
- After cleaning the data, it was time to explore the data.
- The communication apps appear to be more downloaded than any other type of application by a wide margin. Social, productivity, and video playing apps following behind respectively.



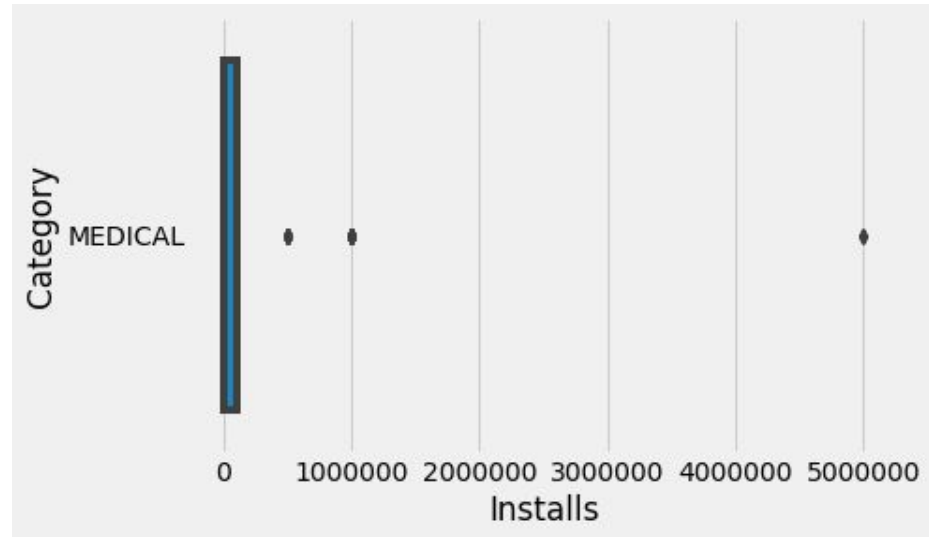
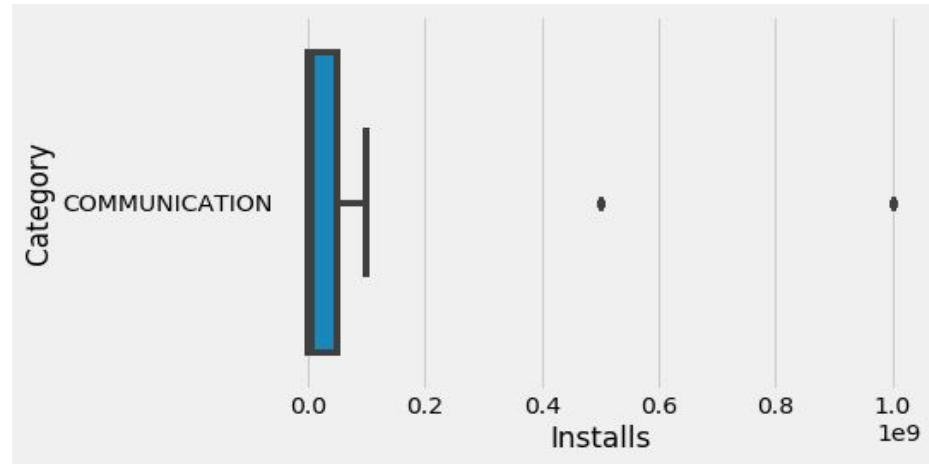
- Interestingly, the most downloaded apps seem to be rated between 4.3-4.5 out of 5.



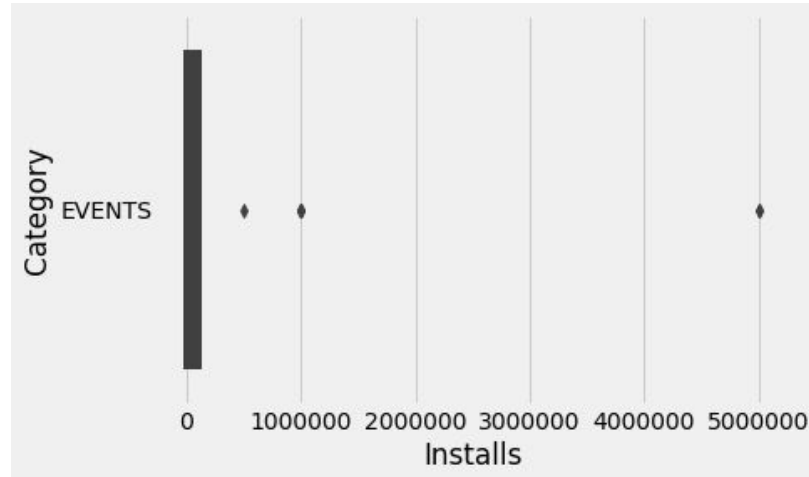
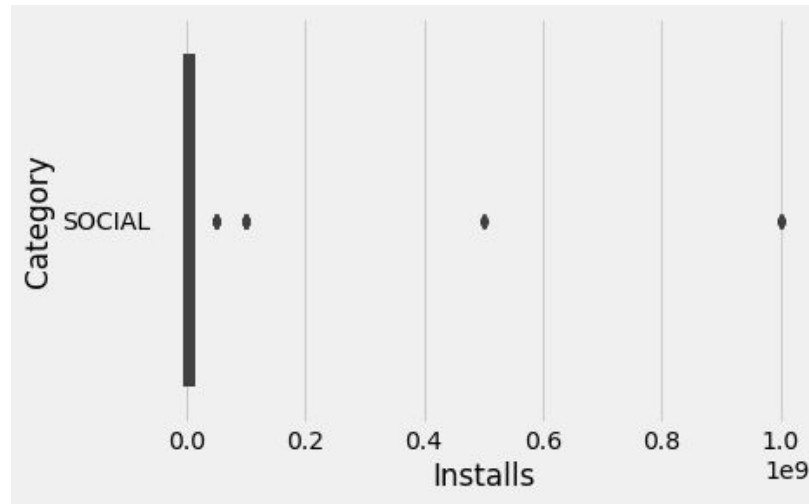
- Apps that are free are downloaded substantially more than apps that are paid on average.
- Apps that are rated everyone 10+ and teen seem to be the downloaded type of app.



- After initial data exploration, it seemed like the specific category of an app could dictate whether or not an app would receive downloads.
- The communication apps saw the most popularity on the google play store with an average of 99,534,273 downloads.
- On the other hand, medical apps saw the lowest average amount of downloads coming in at 152,016 downloads.

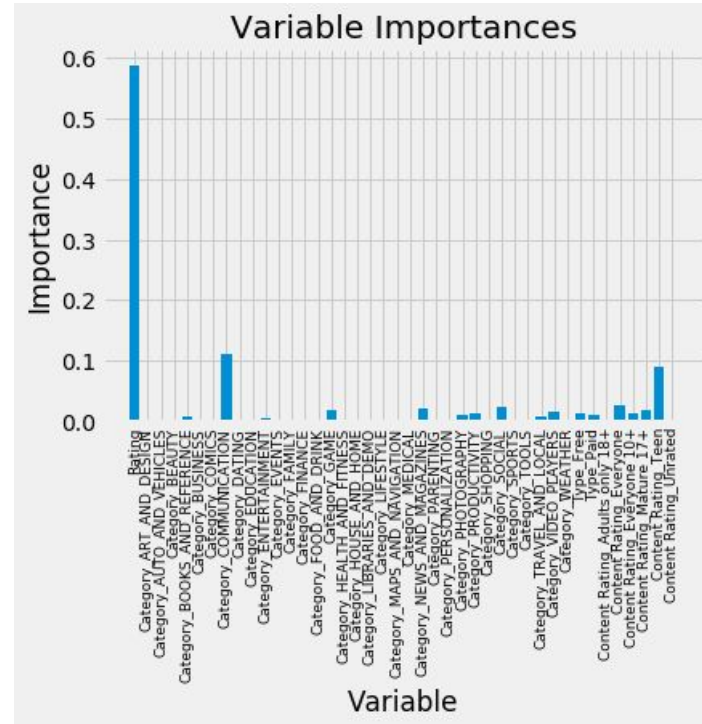


- The social apps were on average the second most downloaded category of apps with an average of 54,323,712 downloads.
- The event apps were the second lowest downloaded apps with an average of 354,431 downloads.

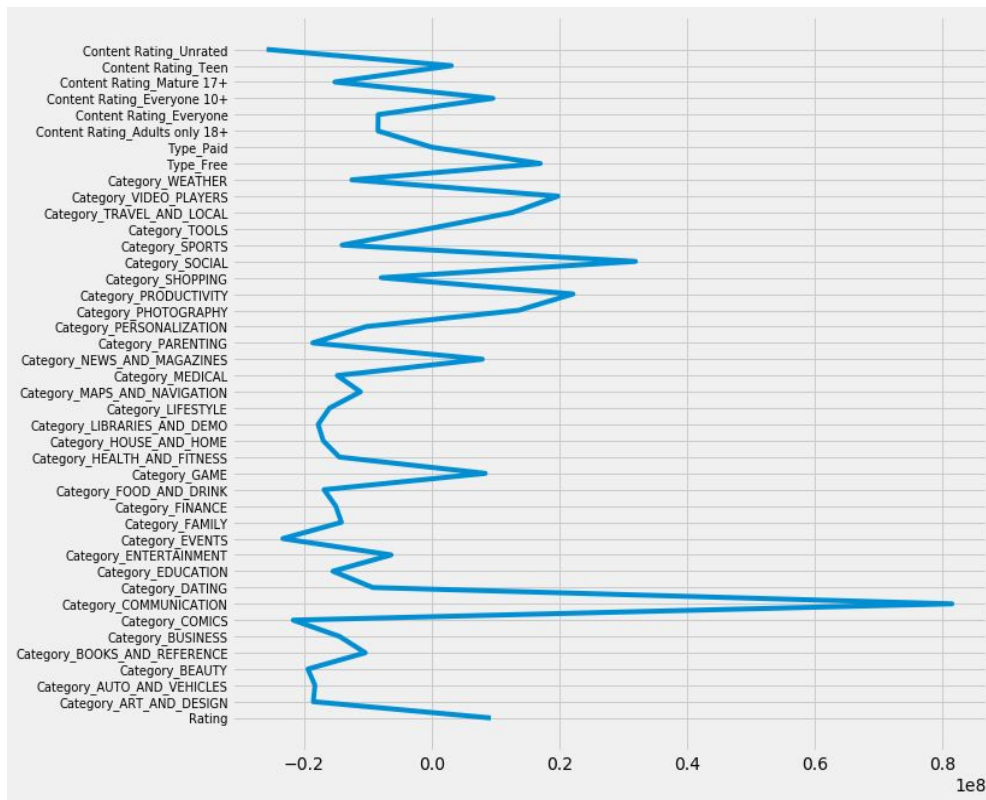


Machine Learning

- To take the given data, and see if it would be possible for us to predict downloads for future apps, machine learning was formulated and applied to the data.
- The results of random forests show that 'rating' and 'communication' seem to be the features that affect installations most.
- The prediction model calculations came out to a mean standard error of 26,079,931 downloads.



- A lasso regression model was also fit to the data to assess and supplement the results of the random forest predictor model results.
- The results show that this model chose the communication and social categories as the most influential indicators of installation numbers.



Discussion

- The results from the machine learning and the illustrations of the data visualizations highlight the unpredictable nature of the app market.
- Although there does seem to be some categories that garner more downloads on average, it is still possible that a popular medical app gets more downloads than an average communication app.
- To shine some different perspective on the results of the random forest predictor model which produced a mean absolute error of about 26,000,000 downloads. If a developer is looking to introduce a new communication app, which average above 99,000,000 downloads on the current market, the plus/minus of 26,000,000 downloads may be the negligible number for the developer.