**Project 1 Write-Up**

Group: Sarah Long, Vito Iosue, James Meuschke

Data: Google Play downloads

Source: Kaggle <https://www.kaggle.com/lava18/google-play-store-apps>

**Context**

**Sample Size**

This data set is a 1% sample of the overall Google Play app offering. It includes over 267,000 apps and 13 attributes of these apps. Based on some research, it appears that this is because users will only see a portion of all apps based on geographical location and preferences. Since Google uses its own recommendation and user behavior learning strategies, all users will see different apps when they visit the Google Play Store.

**Background on app stores**

Interesting to note: while the Google Play store currently contains over 2.6 million apps, which is slightly larger than the Apple App store, containing just over 2 million apps. The Windows, Amazon, and Blackberry World app stores follow far behind. This could be for several reasons: Google has fewer restrictions and qualifications for apps, it may be easier to develop in Java than Swift, and Android users are typically more diverse (global, income level).

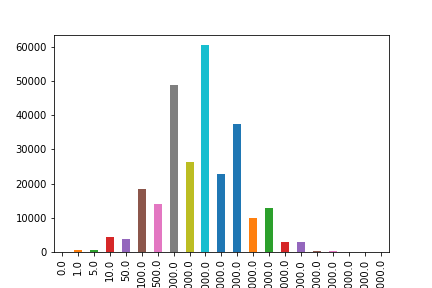
<https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/>

**Variables used**

**Dependent variable: Installs**

* Definite dropoff point on Installs level: it seems that after a certain point, it is difficult to get more users. To better understand this number, consider that Android OS has 2 billion monthly active users, or 85% of the global mobile OS market. Again, this shows how vastly user preferences differ, which means that there is still room for innovation in the market.

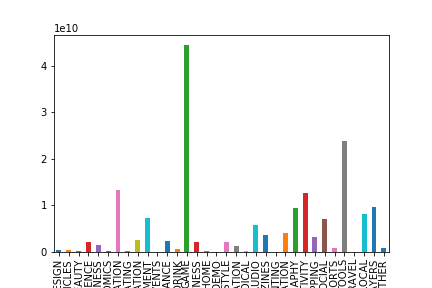
<https://expandedramblings.com/index.php/android-statistics/>



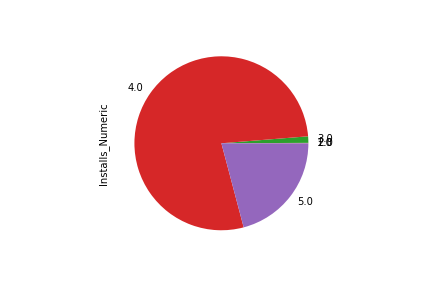
**Independent variables: Category, Rating, Reviews**

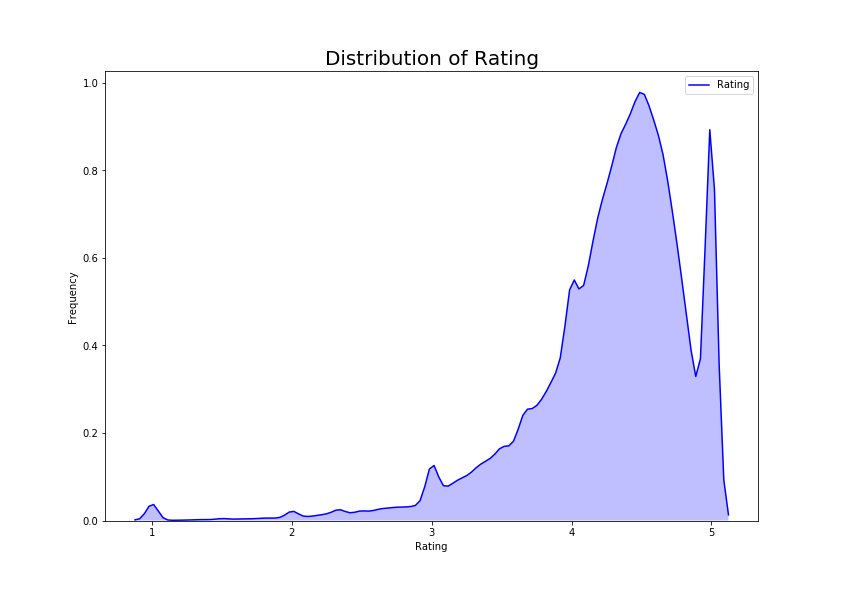
**Independent Variable #1: Category**

**The data included a variety of categories. The largest categories were Games, Tools, and Communication.**

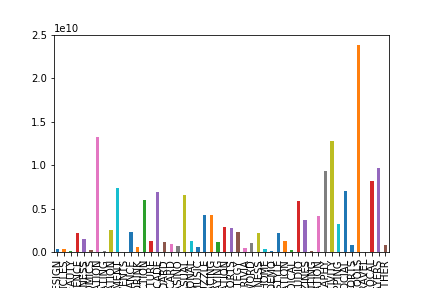
****

**Independent Variable #2: Rating**

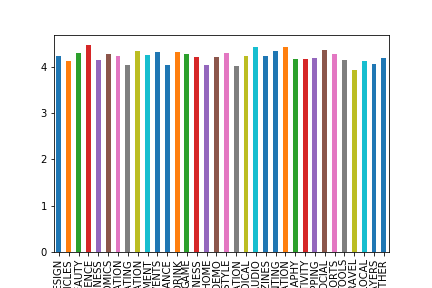


****

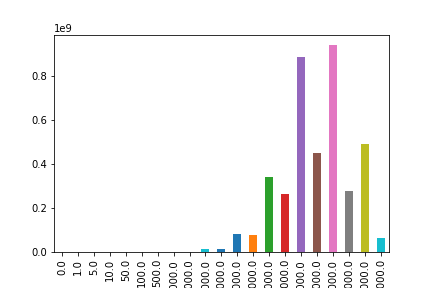
**Independent Variable #3: Reviews**



**Ratings per level of Installation**

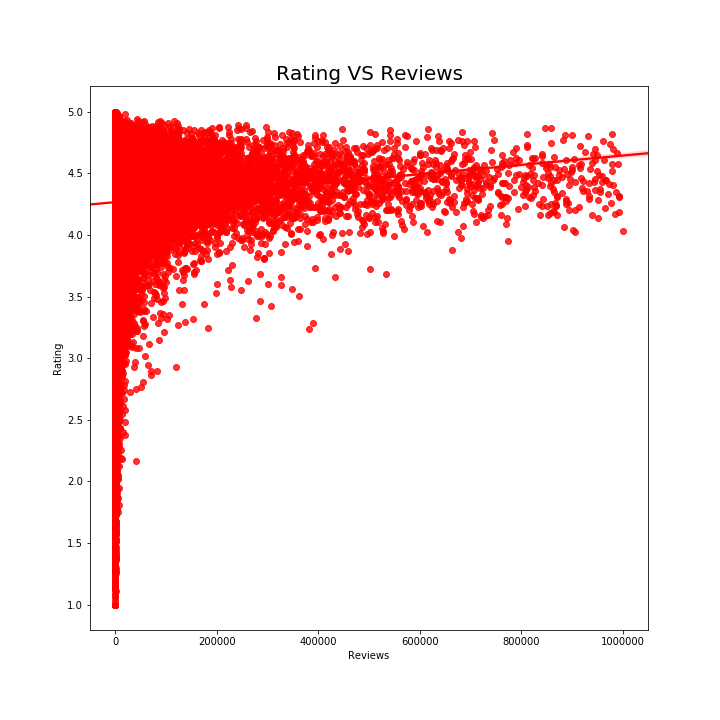


**Correlation between Installations and Reviews**



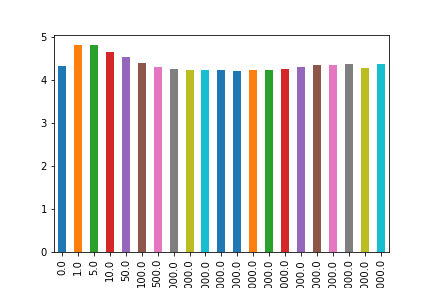
**Correlation between Ratings and Reviews**

Unsurprisingly, it appears that most people give an average of about a 4.5 rating.



**Correlation between Category and Rating**

Overall, we did not see a strong indicator that a particular category had higher ratings.



**Conclusion**

1. There seems to be an inflection point where it becomes increasingly more difficult to get more downloads.
2. It seems to be the case that the more installs you get, the more reviews your app tends to receive, and consequently, a higher ratings.
3. If you want to make an app, aim for making either a game, an apps that solves some sort of problem (i.e. tools), an app that improve productivity, or one that facilitates communication.