GOOGLE PLAY STORE ANALYTICS

Tanmay Atkekar | IST 719 | APRIL 19, 2019

INTRODUCTION

This dataset contains information for about 10000 apps across all genres listed on the Google Play Store. This information includes the following –

App Category	Install Count
Rating	App Genre
Review Count	Last Updated
Size	Price

STORY

The Play Store data has enormous potential to drive app-making businesses to success, for indie developers to big companies alike.

Actionable insights can be drawn from this dataset to gauge the popularity of different types of applications. This chart shows the different correlations and interesting tidbits in this dataset.

WHO IS THIS PROJECT FOR?

My main audience is indie app developers. They are the ones who will be directly impacted by the findings from this data.

My secondary audience would be industry analysts and market watchers who can bet on companies that are in the business of making apps that are shown to be popular from this dataset.

CORRELATIONS

The most important question that this dataset can answer is that which variables matter the most to get the maximum number of installs. The following plot will attempt to answer just this question.

The main number that stands out from this plot is that the app ratings don't affect the number of installs as much, so focus on making the app great, and word of mouth will propel your app to success.
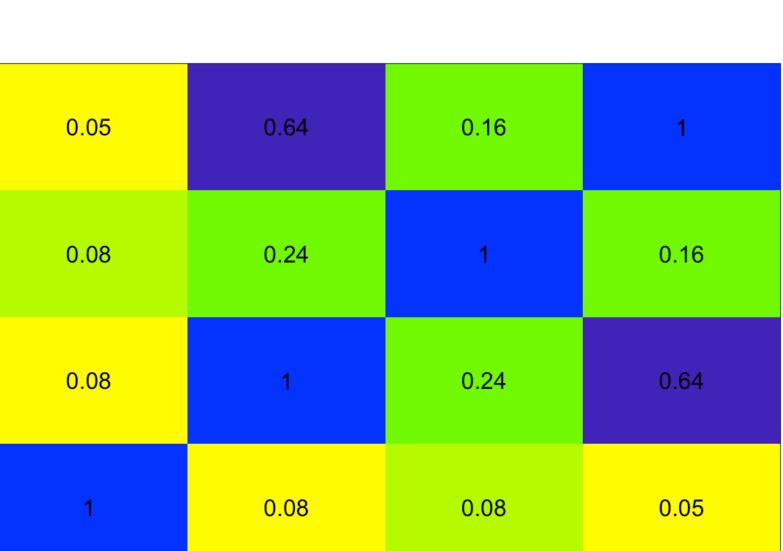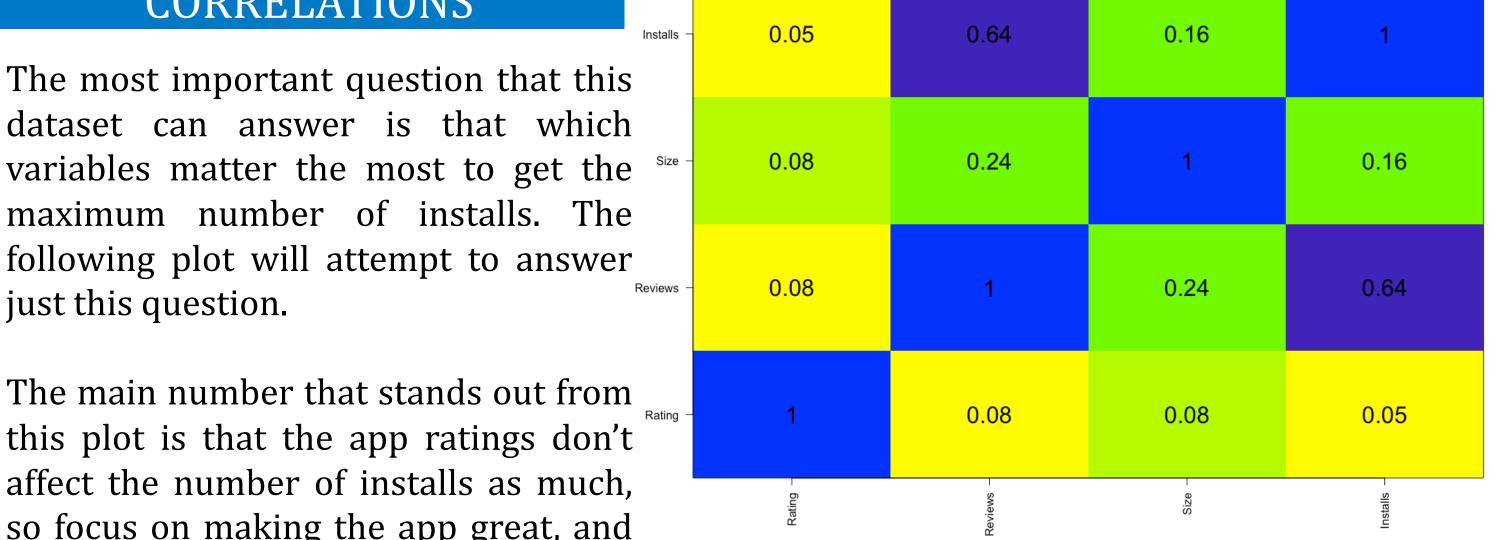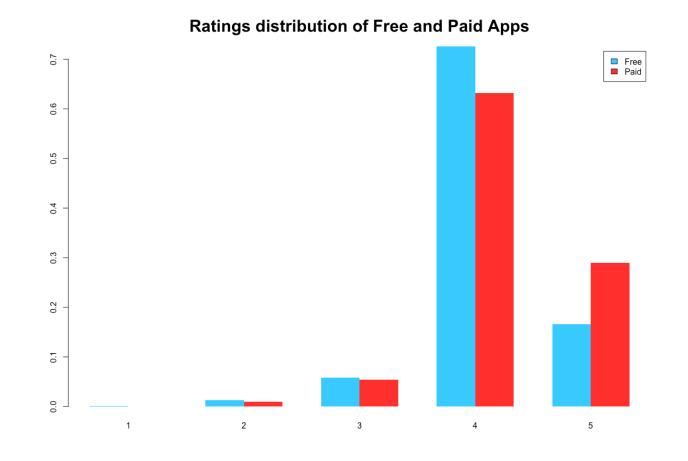

WHICH CATEGORIES ARE THE MOST POPULAR?

Family Apps dominate the Google Play Store. It is important to note that a lot of the family apps include multiplayer games, and as you can see, games are the second most popular category. Thus if you are planning to make an Android game, you are on solid territory.
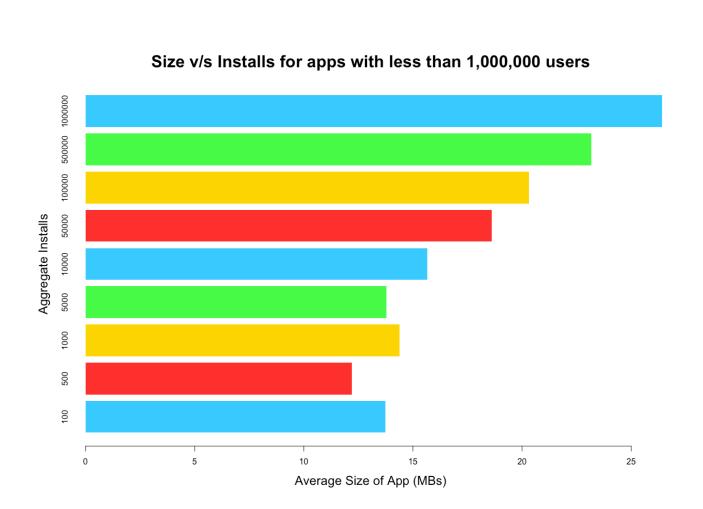

DO PEOPLE LIKE PAID APPS?

Yes! People seem to give higher ratings to paid apps, as compared to free ones. There are almost no 1 star ratings to paid apps, so if you make a good paid app, people will reward it with their positive reviews.

DOES APP SIZE MATTER?

App size might matter in the very beginning, as till the apps reach 10,000 users, there does seem to be a slight trend towards reducing app sizes. After that point, the size doesn't matter, so developers can and should focus on features rather than optimizing the app size.
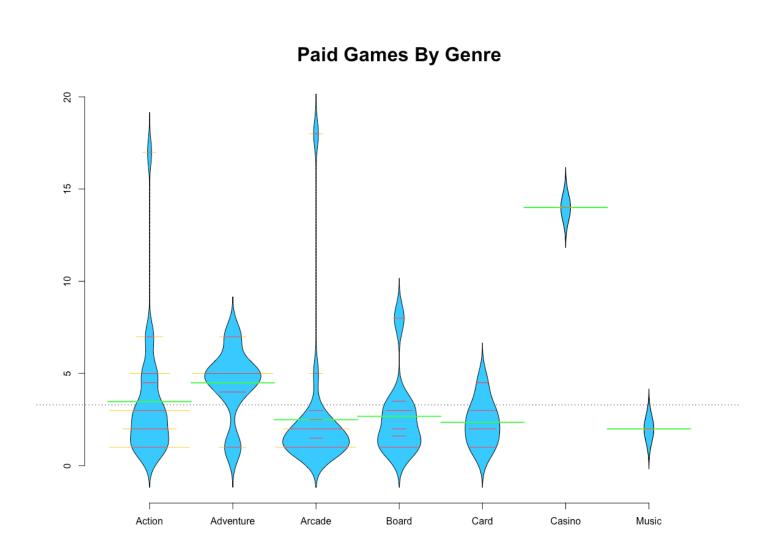

...AND WHICH GAME TYPES MAKES THE MOST MONEY?

The following plot is a combination of a density plot and a box plot, called as a bean plot.

MOST COMMON OCCURING WORDS IN POSITIVE AND NEGATIVE REVIEWS

Positive

Negative

DATA CLEANING & SOURCES

I cleaned the original dataset by removing NA values, converting columns such as Install Count, Rating, Size to numeric values by removing numbers and letters.

Packages used – tidyverse, lubridate, tm, wordcloud, graphics, beanplot, SnowballC

Sources –
https://www.kaggle.com/lava18/google-play-store-apps for the data
https://www.schemecolor.com/google-play-new.php for the color palette