**PLAYSTORE APP REVIEW ANALYSIS**

**Team members  
AKSHAY KUMAR SAINI**

**KUMKUM SINGH**

**Abstract:**

Google Play Store, also referred to as the Play Store and as Google Play, is a platform created and managed by Google that enables the downloading of applications for mobile devices with the Android operating system. The system provides access to free and paid apps.

We were given two data set- one contains the data related to app available in play store i.e., reviews, rating, size, install etc. and another data set contained the customer reviews of a particular app. We were supposed to figure out the factors responsible for the success of particular app.

**1.Introduction**

Play store is digital distribution service developed by google. Play store was launched in 2012. It acts as an official platform for certified devices running on the android operating system and its derivatives. Play store provides many applications for the android users. Most of these applications are charges free and some of them are have some cost. These applications can be directly downloaded from the play store.

The Google Play Store provided over 82 billion app in 2016 and reached over 3.5 million apps in 2017, while after a purge of apps is back to over 3 million. It has been the subject of multiple issues concerning security, in which malicious software has been approved and uploaded to the store and downloaded by users, with varying degrees of severity.

Google play provides applications for movies, gaming, T.V. shows, e-books, magazines, etc. It has become the biggest and most important part of android.

The purpose of our project is to analyse the data set provided and to find out the factors responsible for the success of the applications provided by the playstore.

**2.Problem Definition:**

The play store apps data has the enormous potential to app-making business es to success. Actionable insights can be drawn to developers to work on and capture the android market.

Analysis of data set provided will help the developers of the application to see what are the factors which helps the application to become successful in the market. What are the elements of an application which attracts the customers and make them download that application.

The Problem statements are

1. Does the size of an app affects the rating?
2. Does the price of an app affects the number of people downloading the application?
3. To find out the highest rated application.
4. Does the rating of the app affects the number of people downloading the app?
5. To find out which category has highest app?

**3.EDA ON THE GIVEN DATASET**

We were given two type of data set: -

1. **Play store data set**: -

which provides the data like name of apps, category, rating, content rating, review, size, installed, type, price, genres, last update, current version, android version.

By diagnosing the dataset we came to know: -

* Category, type, latest update, genres data was provided in the form of string which need not to be converted into any other data type.
* Rating was provided in the integer data type and need not to be converted into any other data type.
* Installed, Review, price and size was provided in the form of string and was converted into float data type.

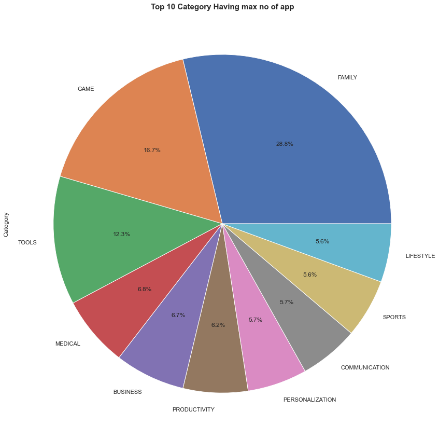
1. **User review data set**: -

which provides data like name of apps and translated reviews and the type of the review i.e Positive, Negative, Neutral.

By diagnosing the dataset we came to know:

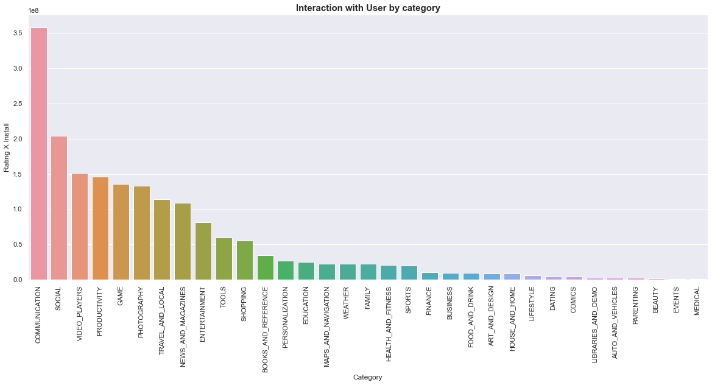
* Every data was given in a specific datatype, so didn’t need to change the datatype

**4.** **TOP 10 CATEGORIES HAVING MAXIMUM NUMBERS OF APPS.**



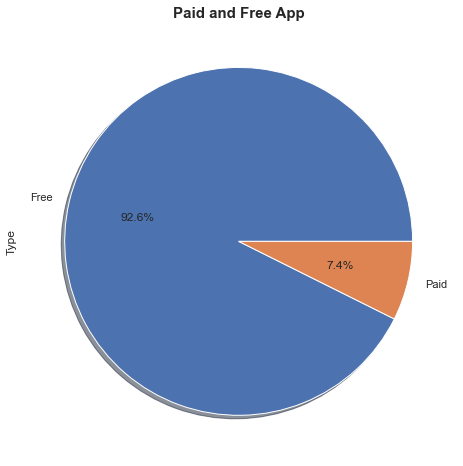
This graph shows the category having maximum number of apps. Here Family category has maximum numbers of apps, accounts for 28.8%. Second comes the game category accounting for 16.7%. Among these sports and lifestyle category has minimum number of apps which accounts for 5.6%.

**5.INTERACTION WITH THE USER**



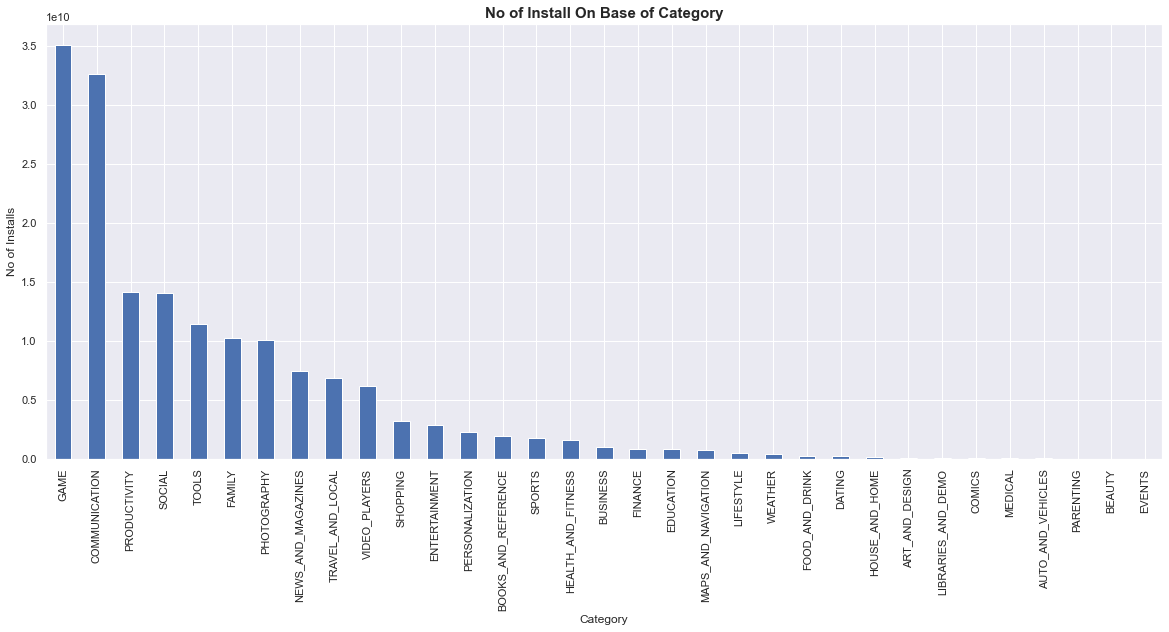
This analysis shows a comparison among all the categories based on the interaction of the user. Communication category has highest interaction with the users, which accounts for approx. 3.5 million whereas categories like events and medical have lowest interaction with the user.

**6.PAID AND FREE APPS**



This graph shows percentage of free and paid apps. Number of free applications accounts for 92.6% whereas number of paid applications accounts for only 7.4%. This data shows that apps provided by playstore are mostly free of cost and few are paid apps.

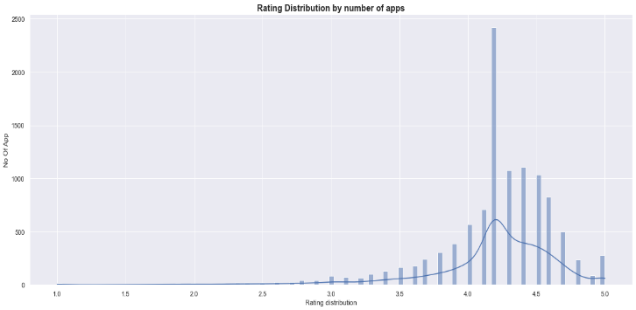
**7.No. of installed based on category: -**

****

This analysis shows the comparison among all the categories based on number of apps installed. Data shows that people prefer to install apps belonging to gaming most which accounts for 3.5 million.

On the other hand, categories like arts and design, libraries and demo, comics, medical have least number of install i.e., 0 number of installs.

**8.RATING DISTRIBUTION: -**



This analysis shows the range of rating a app lies between 3.0 – 5.0, where 3 is considered a poor rating and 5 is considered as excellent. Mostly apps are rated by users in range of 4 – 4.5.

9. **SENTIMENTAL ANALYSIS: -**

Chart

Description automatically generated

This analysis shows the sentimental reviews of various apps which belongs To different-different category. This provides the apps which have positive , negative and neutral sentimental reviews.

According to this analysis most of the apps including education, Dating and health and fitness app have good number of positive reviews i.e. more than 50% whereas game app have received 50% positive reviews and 50% negative reviews. Neutral reviews accounts for very less percentage in all the apps.

**10.Conclusion**

The analysis on the data set of google play store throws the light on the which category has the maximum applications, which category of apps people prefer to install more. Which apps has highest rating.

Many libraries like klib, pandas, ast, matplotlib, numpy, seaborn, datetime and wordcloud were used to analyse the data.

Analysis showed that**: -**

* family category has maximum numbers of apps, accounts for 28.8%.
* Communication category has highest interaction with the users, which accounts for approx. 3.5 million.
* Number of free applications accounts for 92.6% whereas number of paid applications accounts for only 7.4%.
* Mostly apps are rated by users in range of 4 – 4.5.
* people prefer to install apps belonging to gaming most which accounts for 3.5 million.
* most of the apps including education, Dating and health and fitness app have good number of positive reviews i.e. more than 50% whereas game app have received 50% positive reviews and 50% negative reviews.

**11.References**

**1.** [**www.lifwire.com**](http://www.lifwire.com)

**2.** [**www.androidauthority.com**](http://www.androidauthority.com)

**3.** [**www.geeksforgeeks.org**](http://www.geeksforgeeks.org)

**4.** [**www.stackoverflow.com**](http://www.stackoverflow.com)