# Play Store App Review Analysis

**Team members:** Dhruve Dubey , Sanket Kamble , Deepankar Singh,

Pratiksha Chimankar,Komal Sursawant

**INTRODUCTION:**

Google play store was launched on March 6, 2012,bringing together Android Market marketing a shift in the Google’s digital distribution strategy. As we know that google play store apps is the biggest market for Android app because of more than 5 millions application with more than 3 billion active users in more than 190 countries. Developers and users play an important role in this. Determine the impact of market interactions on advance technology However, both developers and users suffer from a lack of awareness of the inner workings and dynamics of well-known app markets. This project tells us effectual of google play store app and get insides from that for developers to upgrade it.

**METHODOLOGY:**

We analyze the data by following steps: About data, Description of data, Data cleaning, Inspect the null and missing values and data analysis and visualization. We need to start from get the data from Almabetter’s dashboard. Then doing data cleaning and data visualization. After that analyze the data by different plots .

**DESCRIPTION OF DATA:**

While looking to the Downloaded data sets we can see they provided two .CSV files. In first csv file playstore\_review details of all play store applications are given . There are thirteen features of apps are provided. In second csv file five columns are provide. App, Translated\_Review, Sentiment ,Sentiment\_Polarity, Sentiment\_Subjectivity. In first data set features of 13 columns is

* **App**: name of the application
* **Category**: category of application
* **Rating** : rating given to the application
* **Reviews**: Number of reviews given to the Application
* **Size** : Size of the application
* **Installs**: Number of downloads of the Applications
* **Type** : Free or Paid
* **Price** : Price of the applications
* **Content Rating** : it is related to the age of consumer
* **Genres** : Types of Genre of the Application on the playstore

In second csv file features of 5 columns is

* **App**: An app name
* **Translated\_Review:** Reviews of apps
* **Sentiments**: Sentiment given to an apps by consumer
* **Sentiment polarity**: The polarity of sentiments measures how negative and positive the context is.
* **Sentiment subjectivity:** The subjectivity of sentiment is how likely that sentiment is to be based on data or factual information.

**DATASETS BREAKDOWN**

Next step for data visualization is process with data

* Importing important libraries for analysis
* Mounting Google drive and read csv files
* viewing all data information
* Check the Null values
* Check the duplicate values
* Make the size column in one format
* Removing the special characters
* Changing the data type to float

**DATA CLEANING**

This is the most important part of project..Validate and cleanse the data is the essential for eliminating inaccurate data and completing gaps. Here are several key duties

* deleting unnecessary data
* completing blank values.

For a given application, a dataset might contain duplicate values we need to remove it.

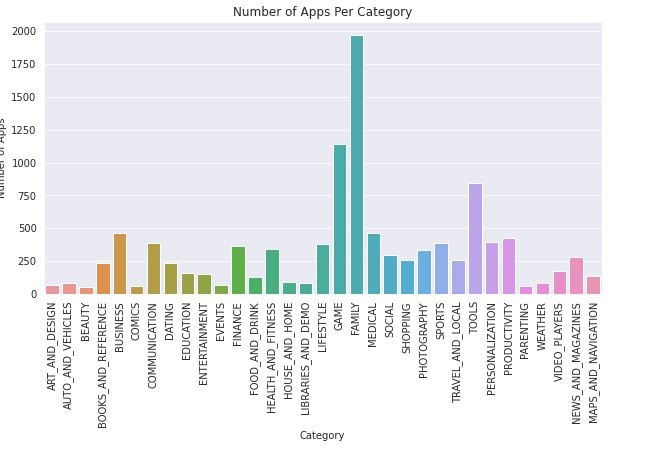
* In play store dataset we have entries of size , installs and price with special characters like 'M','+','$' respectively.
* This prevents columns from being purely numerical, making their use in furture mathematics difficult.
* we need to remove this special characters(M,+,$)
* Also we need Size column in on format

Now that we have removed all the null values, we would remove some labels and columns that are not necessary for actionable insights. As a result, we may move forward with the exploratory data analysis and make observations about the datasets.

**DATA ANALYSIS AND VISUALIZATION**

**Case 1:**

* What are the top categories on playstore?

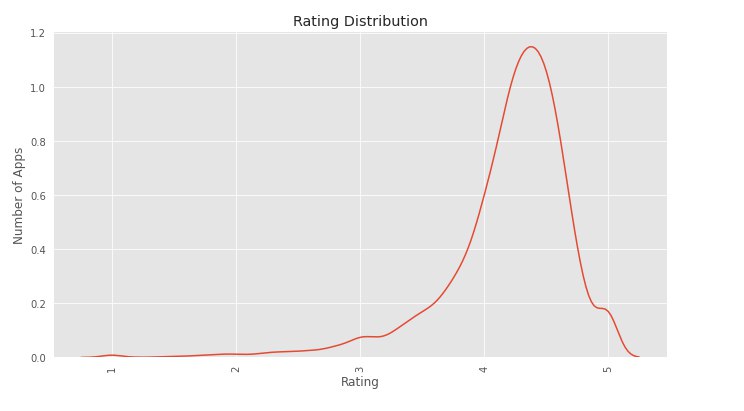


* From the above graph,we can conclude that Family category has highest number of apps in playstore.
* Also top 5 category, which has the large number of apps in playstore are Family,Game,Tools,Business, Medical
* Family which category has maximum apps in play store more than 1950

**Case 2:**

* How importance is the rating of application?

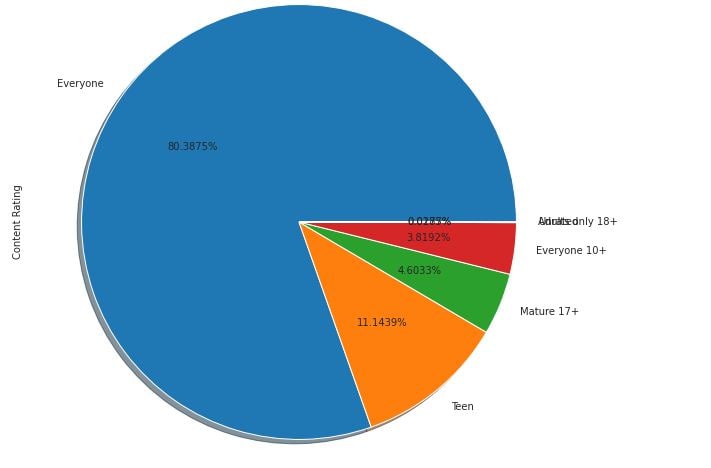




* As we can see clearly from above bar graph that approx more than 1100 apps have a 4.5 rating. And
* According to the below diagram most of the apps is highly rated and less number of apps is low rated . 4.4 is highest number of apps rating

**Case 3:**

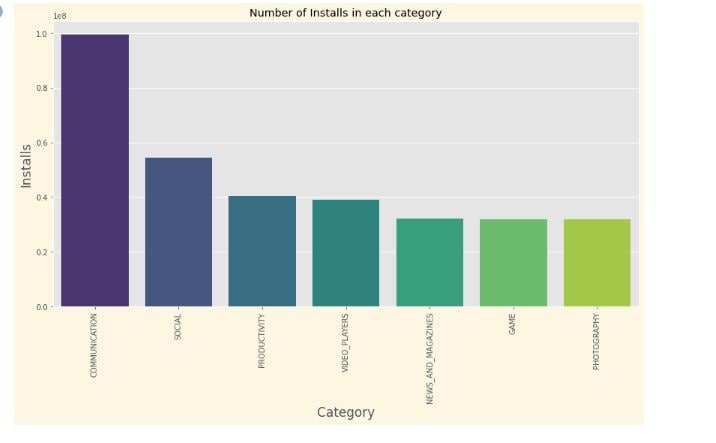
* mostly percentage of content on playstore



* From the above pie chart we can conclude that most of the apps installed on the playstore is for everone group with 80%(approx) ,followed by Teen, mature17+,etc

**Case 4:**

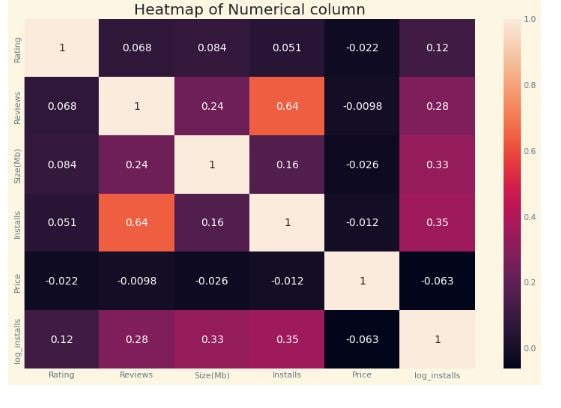
Which category has the most number of installation ?



* From the above graph we can conclude that a large number of installation is done by consumer in communication category ,followed by social ,productivity,etc.
* This means that most of the consumer likes the communication apps so,Developer needs to make more apps in this domain.

**Case 5:**

Check correlation



* There is a strong positive correlation between the Reviews and Installation
* The price is negatively correlated with Rating,Reviews, and Installs.

**Case 6:**

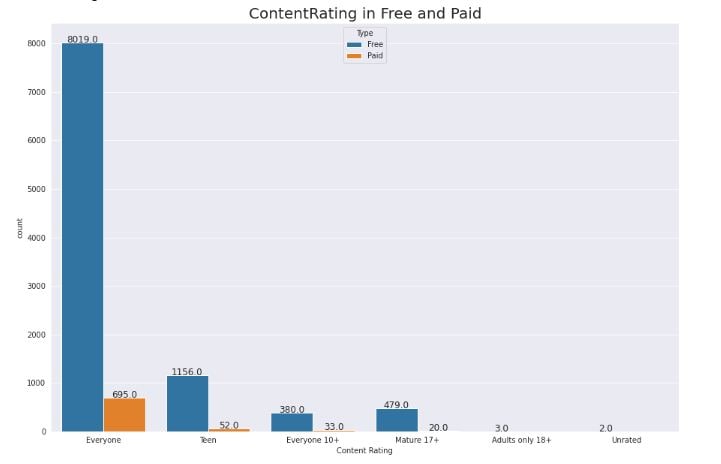
Mostly used words on playstore reviews?



* Here ,food,best,app,help,work words are used most of the times by the consumer in reviews.
* So,we can coclude that most of consumer like food domain apps therefore if develpoper works in this domain then there is a high chance of successful apps.

**Case 7:**

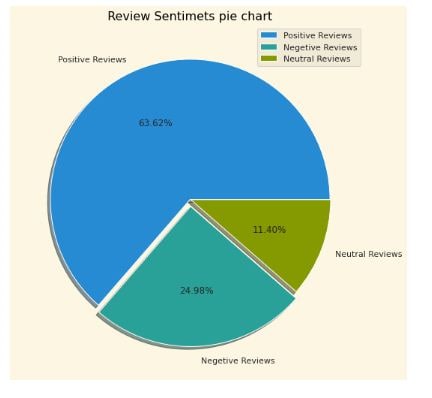
Relationship between content Rating in Free and paid apps?



* From we can conclude that ,most of the consumer given the maximum rating in free apps of everyone category than paid apps.
* Everyone category has maximum number of paid apps compared other category paid apps.

**Case 8:**

What are the sentiments across database?

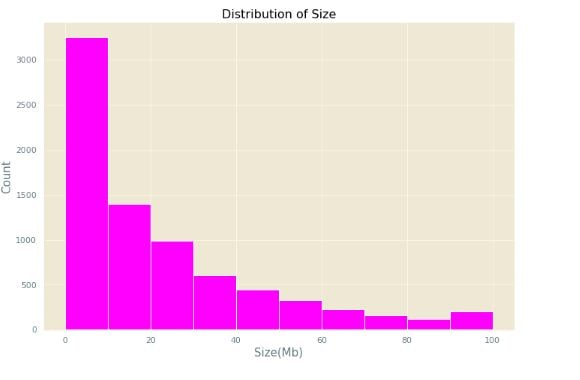


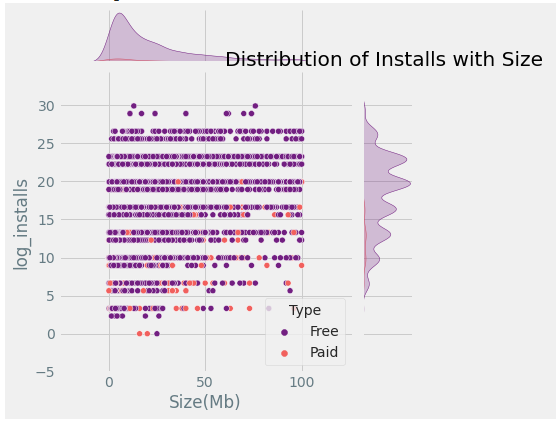
* Most of the sentiments reviews given by the user are positive with 63.62%
* But also there is a negative sentiments percentage of 24.98%
* This mean developer had to do their hardwork to change this negative sentiments into positive or neutral.

**Case 9:**

How size impacts on installation of apps?

* Here, we can say that high number of apps which are installed on playstore are between 0 to 20 Mb size
* This means that if the size of apps is less than 20 MB size then it is very high chance to downloading that apps.

****

****

**CONCLUSION**

As a result, the app development companies could choose the type of programme that should be created and could also evaluate how it would perform in the future. They may also view the categorised reviews of every programme in one interface, which will enable them to determine which apps are well-liked by customers and which ones require further development.

**ACKNOWLEDGEMENT**

This capstone project completed by Dhruve Dubey, Sanket kamble , pratiksha chimankar,Komal sursawant,Deepankar Singh. I want to express my sincere gratitude to each and every one of them. I owe Almabetter a debt of gratitude for their direction, continual oversight, and provision of the project's vital information as well as for their assistance in seeing it through to completion.