

# **CAPSTONE EDA PROJECT**

## **ON**

# **PLAY STORE APPs REVIEW ANALYSIS**

**MUJTABA ALI, KIRAN JOSHI, PRATEEK SACHDEVA,**  
**SAMRAT CHOWDHURY & SUBHAM SINHA**

**DATA SCIENCE TRAINEES,**  
**ALMABETTER, BANGALORE**

### **Abstract:**

The Play Store apps data has enormous potential to drive app-making businesses to success. Actionable insights can be drawn for developers to work on and capture the Android market.

Each app (row) has values for category, rating, size, and more. Another dataset contains customer reviews of the android apps.

So, we need to explore and analyze the data to discover key factors responsible for app engagement and success.

- Find out the app category which application has been installed maximum number of times.
- What are the factors affecting the maximum downloads?
- Exploring the top 5 highest rated free and paid apps.
- Exploring the average rating for each category.
- Understanding the sentiment polarity of the users.

### **1.Problem Statement**

Explore and analyze the data to discover key factors responsible for app engagement and success.

In this project, we investigate the different variables of Apps on Google Play Store that why and how certain applications succeed and others. Also, what is required for an application to be considered as successfully topping the charts. We attempt to use our analysis to answer the following problem statements:

- **App** : Name of application.
- **Category** : Category of application.
- **Rating** : Rating of an Application (between 0.0 to 5.00).
- **Reviews** : Reviews given by users.
- **Size** : Size of app (Kb or Mb).
- **Installs** : Total No. Of installations.
- **Type** : Type of app (Free/ paid).
- **Price** : Price of app.
- **Content Rating** : the rating given by targeted audience.
- **Genres** : Genre of App.

- **Last Updated** : Date when app is updated last time.
- **Current Ver** : Current version of App.
- **Android Ver** : Android version of device.

## 2. Introduction

PlayStore, is where you can download or buy millions of apps, games, and other media onto your Android device. You can find programs for a wide array of interests. You can download apps or games by navigating to the app or game page within the Play Store and tapping Install. It's easy to search for things you need online, from apps to media; Google is the authenticated space for everyone. Developers are free to host their apps and can improve its performance as per users' feedback.

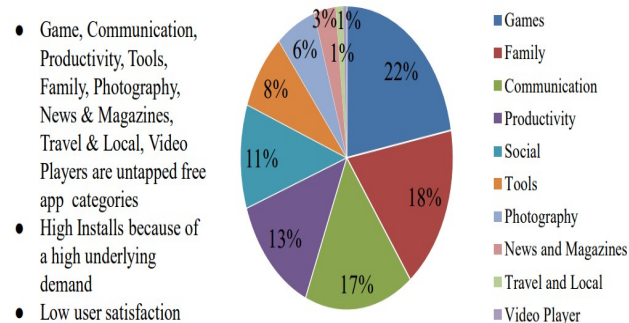
## 3. Reasons for doing Analysis

The reasons for doing analysis are:

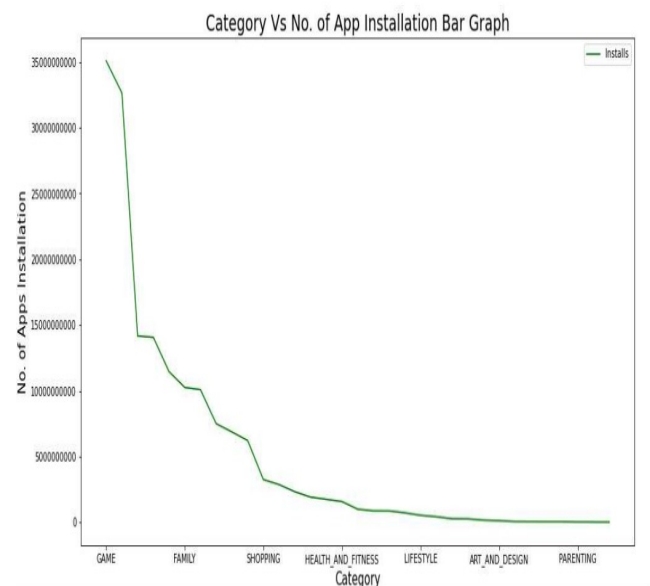
- Play Store has a greater level of importance on ratings and reviews than ever before. Apps with higher ratings and reviews rank high in search. If an app ranks high then there's a better chance of it being found and downloaded. The more downloads the more it will be kept high in the rankings
- **Exploratory Data Analysis**  
After loading the dataset

## Data Visualization

### Top Installed Apps By Category



We Explore and analyze the data to discover key factors responsible for app engagement and success. Unique values from the given datasets. We applied methods like unique (), shape() and groupby() methods to understand to figure out the following inferences that game cat app has been installed max no of times.



- **Null values Treatment**

Our dataset contains a large number of null values which might tend to disturb our accuracy hence we dropped them at the beginning of our project inorder to get a better result.

- **Tops Apps Installed By Category:**

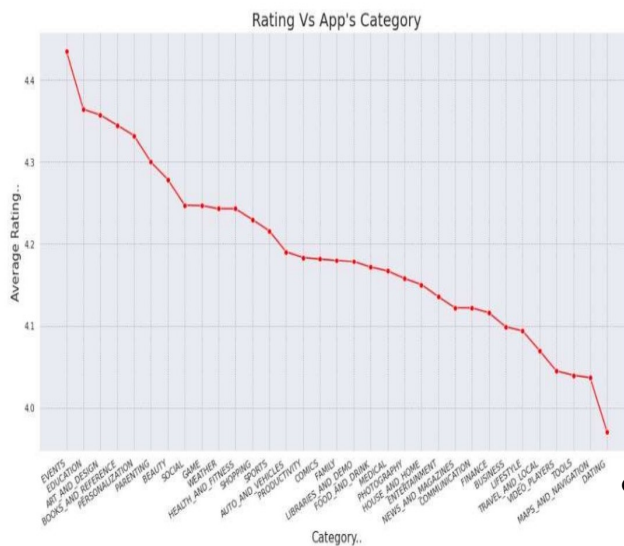
- Game, Communication, Productivity, Tools, Family, Photography, News & Magazines, Travel & Local, Video Players are untapped free app categories.
- High Installs because of a high underlying demand.
- Low user satisfaction.

- The maximum no. of downloads is between 4.0 to 4.6.
- If the app is free to use then there is high probability to install.
- Paid apps are less likely to be installed.
- For the paid apps category maximum Installations are upto 20 dollars.
- Users prefer to pay for apps that are light-weighted. As app size get larger conversion rate decreases.
- If Developers are updating theirs app regularly then there is high chance to install.
- As you can see in Histplot, developer should give updates within 1 to 2 years.

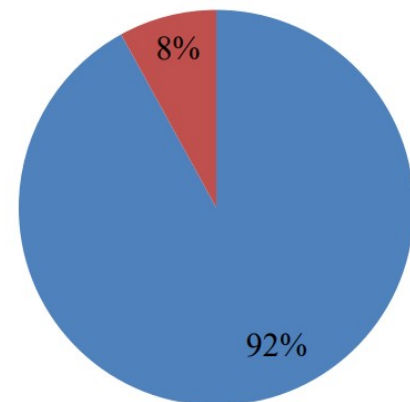
- **Free vs Paid Apps**

- 92% of apps on play store are free to download and rest are paid.

**Average Rating For Each Category**



■ Free apps ■ Paid apps



- **Top 5 Highest Rated Free Apps**

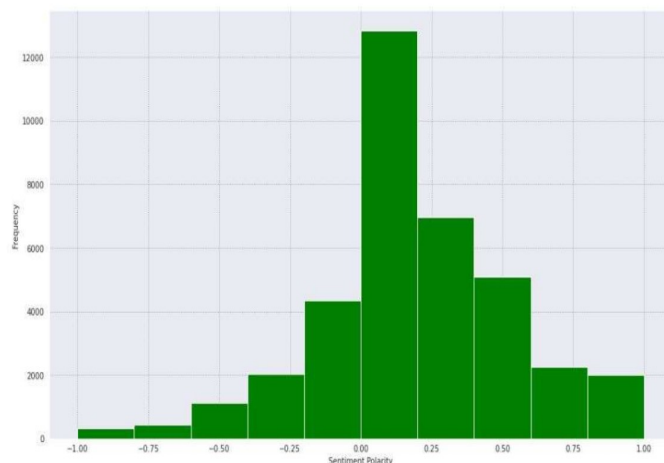
- DF Glue Board
- CN Resident
- CL Strength
- Color CL

- Maximum Downloads

- CP Trivia
- **Top 5 Highest Rated Paid Apps**
  - USMLE Step 2 CK Flashcards
  - 211:CK
  - 30WPM Amateur ham radio Koch CW Morse code tra...
  - Morse Player
  - AC DC Power Monitor
- Rating
- The average ratings across categories is statistically different.
- The Events, Education, Art and Design, Books and Reference produce the best apps with 50% apps having a rating greater than 4.5.

[Dating apps have a rating lower than the average](#)

#### Sentiment Polarity Of Users



- Sentiment Polarity
- Maximum Apps around 40000 apps reviews Sentiment polarity lies between
- to .18 (Neutral to towards Positive)

- Around 3000 apps reviews Sentiment Polarity lies between -1 to 0.0 means About 3000 apps have negative sentiment polarity.

#### • **4. Conclusion:**

Results we got as per our analysis :

- Game category has the most number of downloads and ratings, and Event category has the maximum number of average rating.
- The size and free/paid apps category affected the maximum number of downloads.
- We did not find any correlation between paid apps and number of downloads or ratings.
- Users prefer to pay for apps that are light-weighted. As app size get larger conversion rate Decrease.