CAPSTONE EDA PROJECT

On

PLAY STORE APPS REVIEW ANALYSIS

Submitted by :
Mujtaba Ali
Kiran Joshi
Prateek Sachdeva
Samrat Chowdhury
Subham Sinha

PLAY STORE APPS REVIEW ANALYSIS

Why review analysis is important?

Play Store place 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

App Market is set to grow 20% by 2023

Datasets

Two datasets are provided
1.   Basic Information
2.   User Review

It consist of 13 columns: *App, Category, Rating, Review, Size, Installs, Type, Price, Content, Rating, Genres, Last Update, Current Ver,* and *Android Ver*
and *with 10841 Rows.*

- ## **Overview Of Analysis**

Data Cleaning
Understand the structure of the dataset and clean data before analysis

Data Exploration
Uncover initial patterns, characteristics, and points of interest using visual exploration

Predictive Modeling
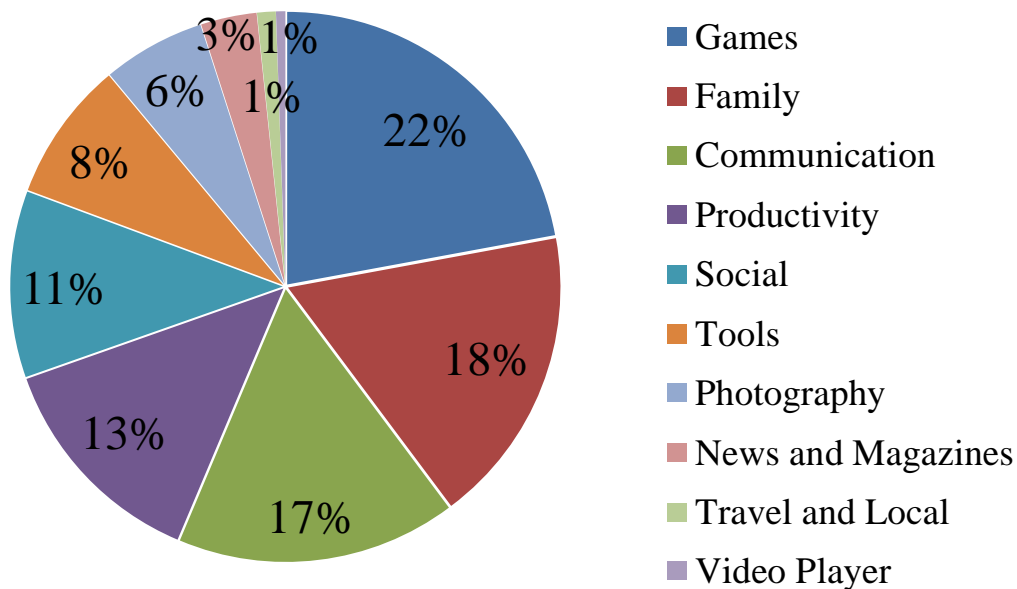Formulate a statistical model to forecast an outcome using relevant predictors

Problem Statement

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

Data Visualization

Top Installed Apps 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

Pie chart: Top Installed Apps By Category

- Games — 22%
- Family — 18%
- Communication — 17%
- Productivity — 13%
- Social — 11%
- Tools — 8%
- Photography — 6%
- News and Magazines — 3%
- Travel and Local — 1%
- Video Player — 1%

Legend:
- Games
- Family
- Communication
- Productivity
- Social
- Tools
- Photography
- News and Magazines
- Travel and Local
- Video Player

GAME category apps has been installed maximum number of time

Category Vs No. of App Installation Bar Graph
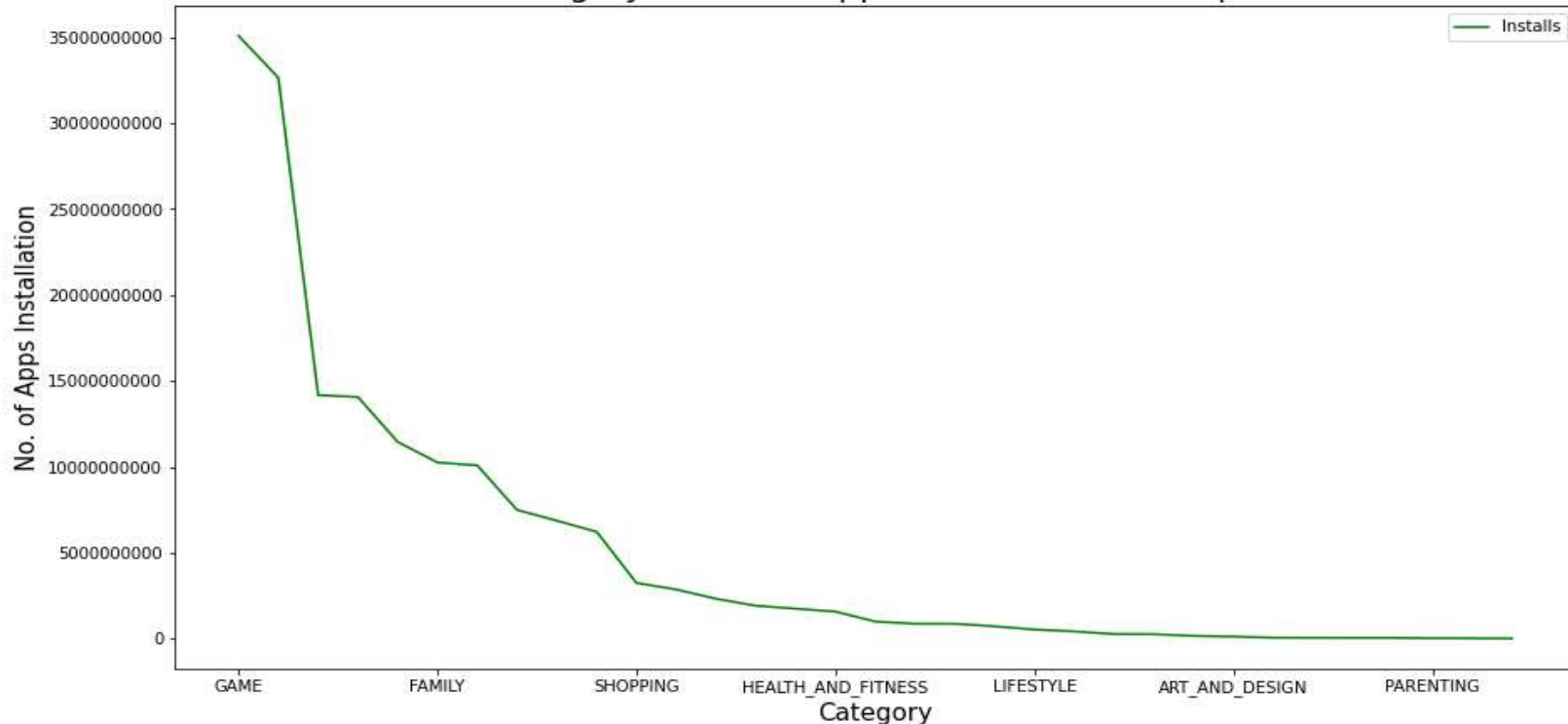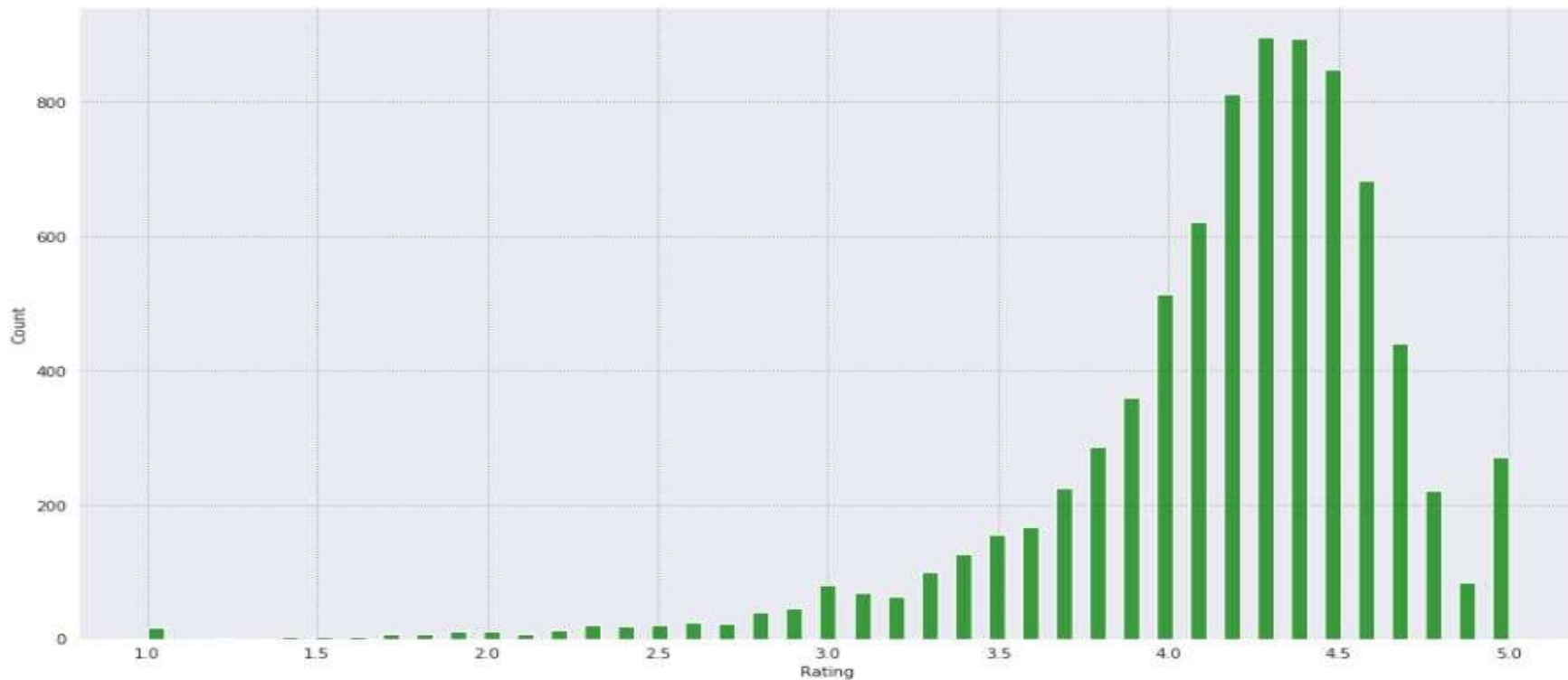
Factors Affecting Maximum Number Of Downloads

•The size of app indirectly proportional to No. of Installation.
• If the size of app is less than or equal to 10MB then there is high chance to install that application.


Distribution of Size Feature
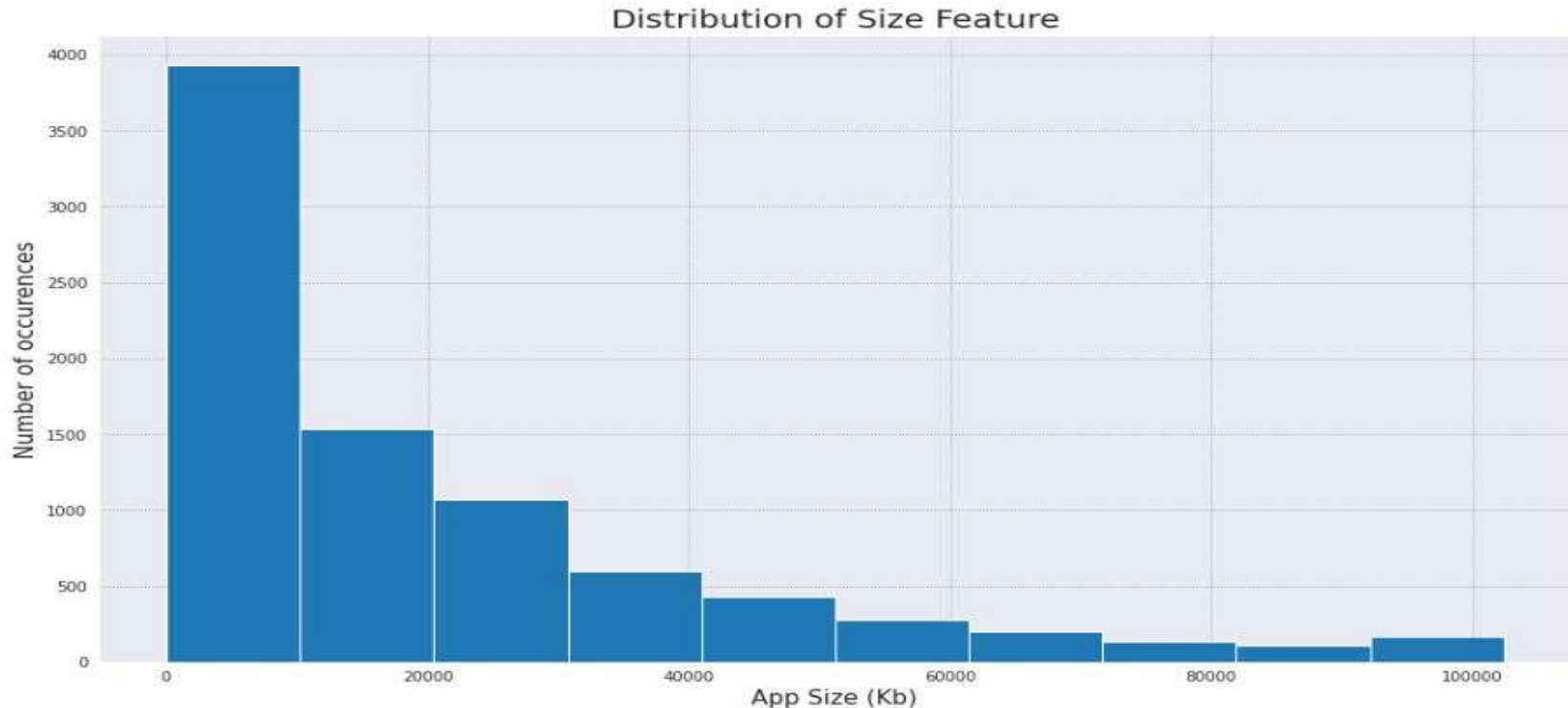
Maximum Downloads

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.
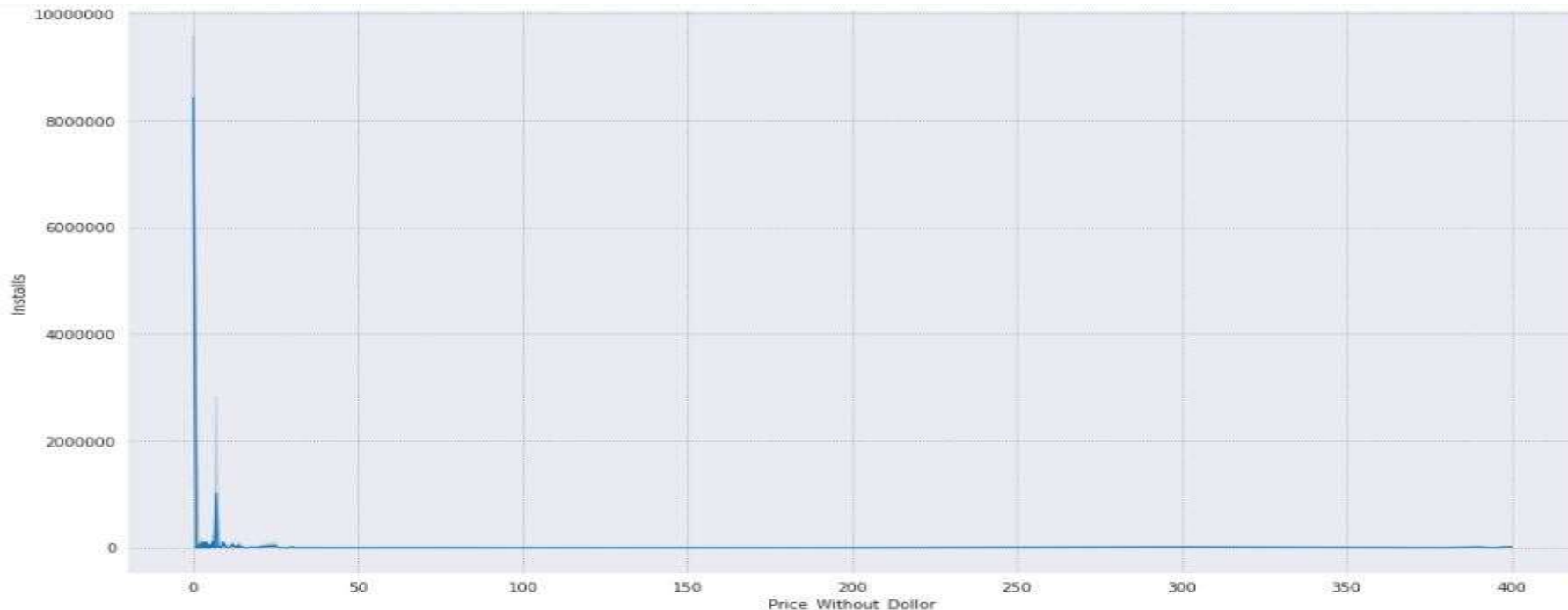
- If Developers are updating theirs app regularly then there is high chance to install.
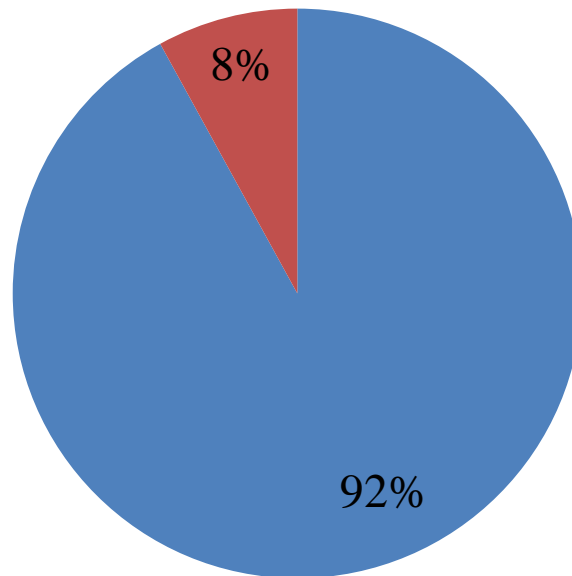- As you can see in Histplot, developer should give updates within 1 to 2 years.
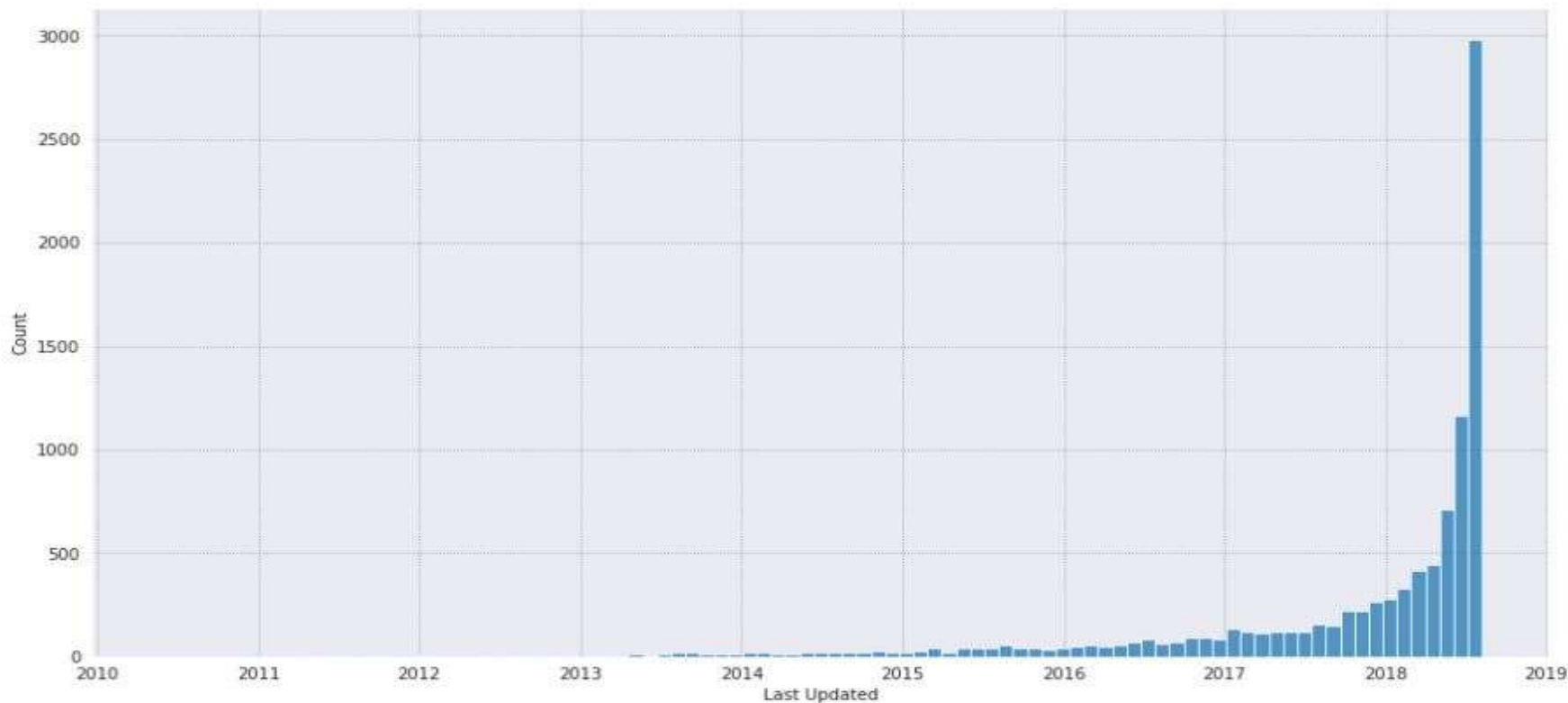
Free vs Paid Apps

92%

of apps on play store **are free to download** and rest are paid

Top 5 Highest Rated Free Apps

DF Glue Board

CN Resident

CL Strength

Color CL

CP Trivia

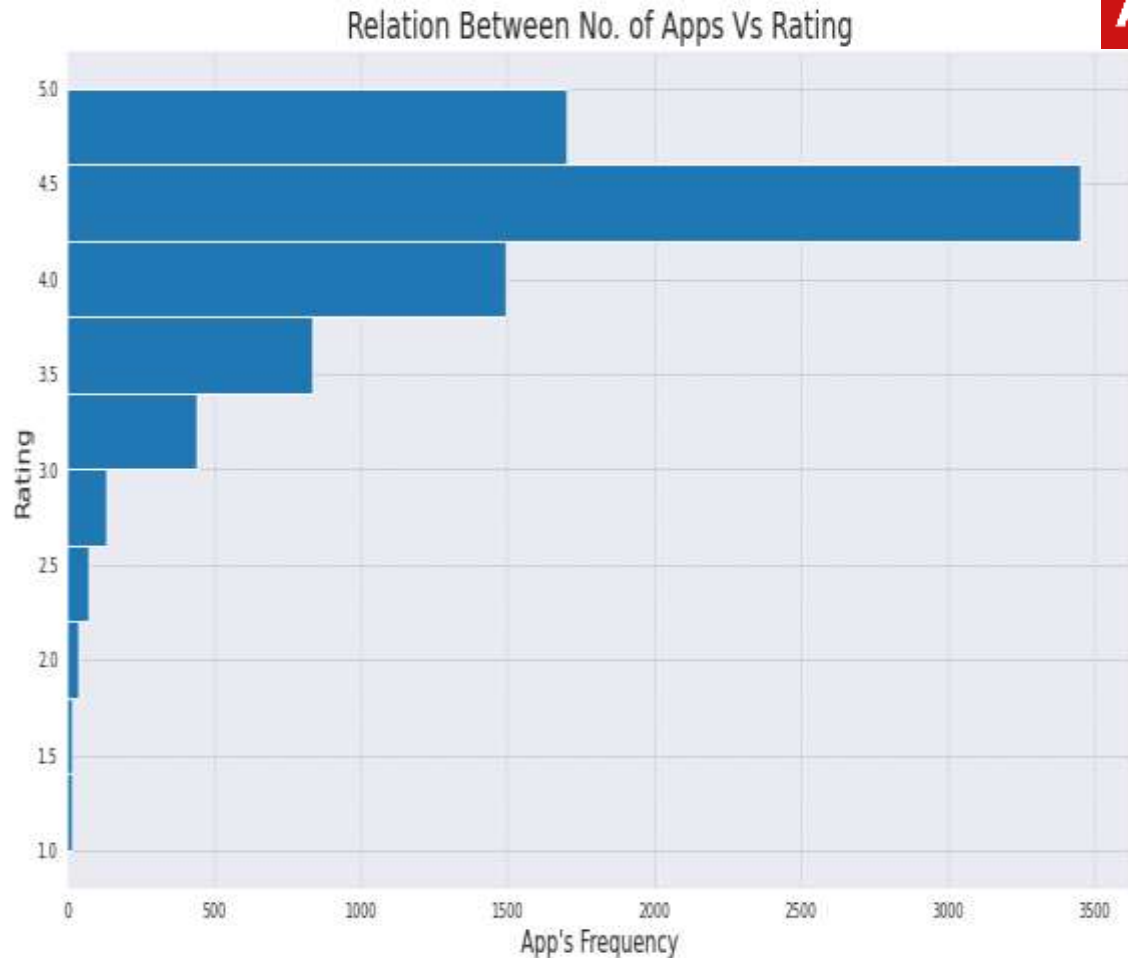## Top 5 Highest Rated Paid Apps

USMLE Step 2 CK Flashcards

211:CK

30WPM Amateur ham radio Koch

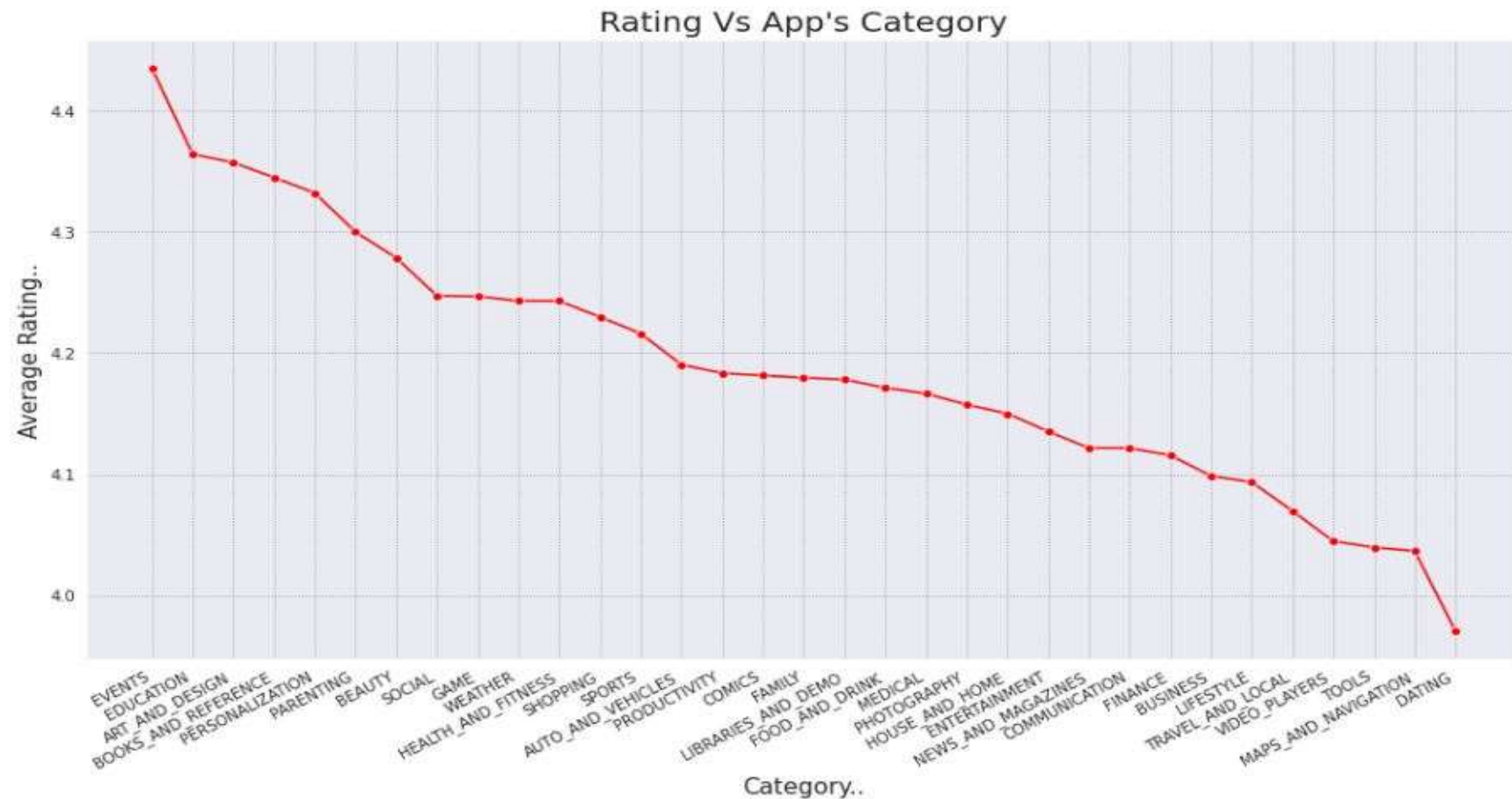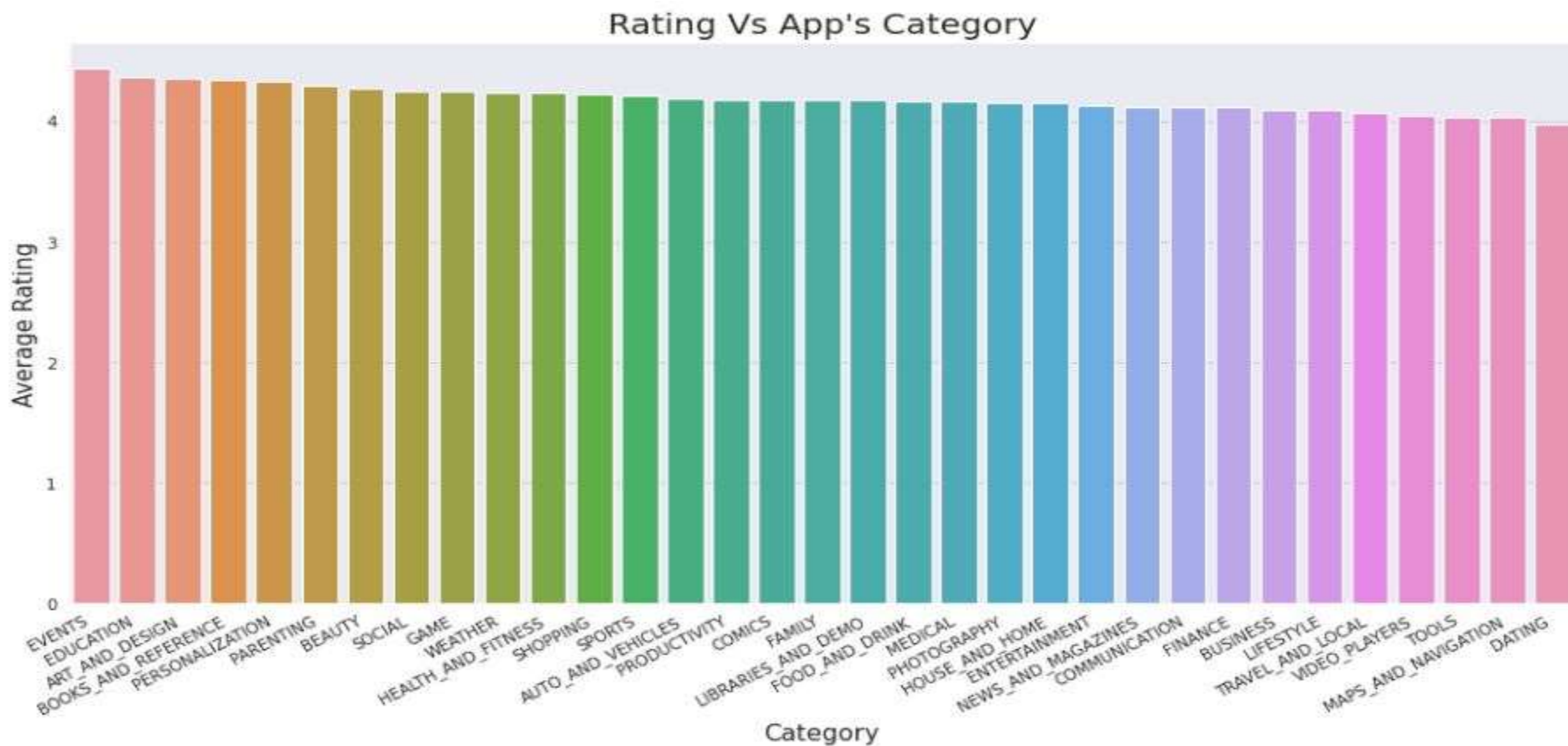CW Morse code tra...

Morse Player

AC DC Power Monitor

Relation Between No. of Apps Vs Rating

Average Rating For Each Category

Rating Vs App's Category

In the below graph we found the average rating.
The maximum average rating of category is Events

Rating Vs App's Category

Sentiment Polarity Of Users
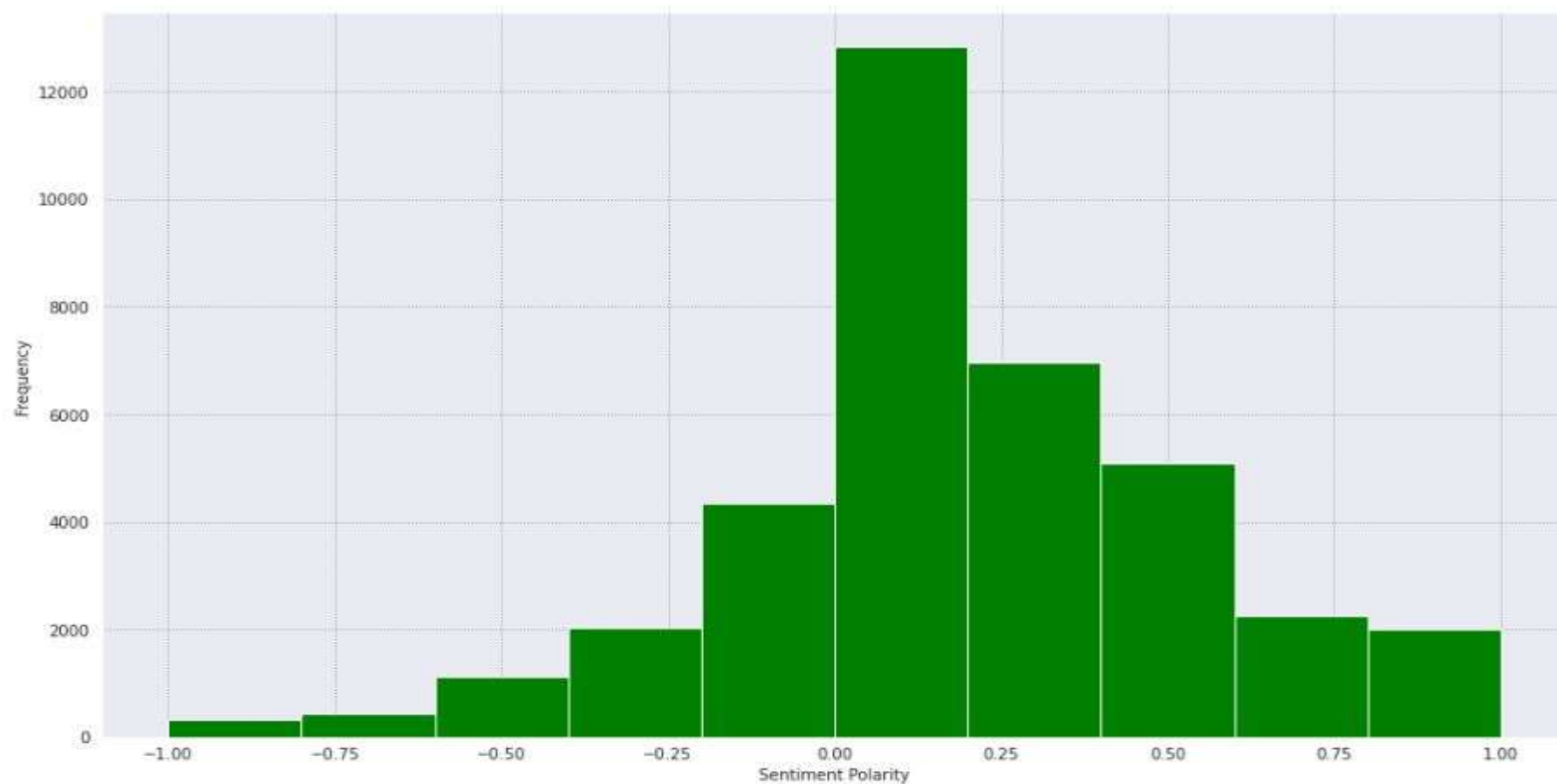
• Maximum Apps around 40000 apps reviews Sentiment polarity lies between 0.0 to .18 (**Neutral to towards Positive**)

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

• Mostly Apps Sentiment polarity lie between **-0.47 to +0.8**
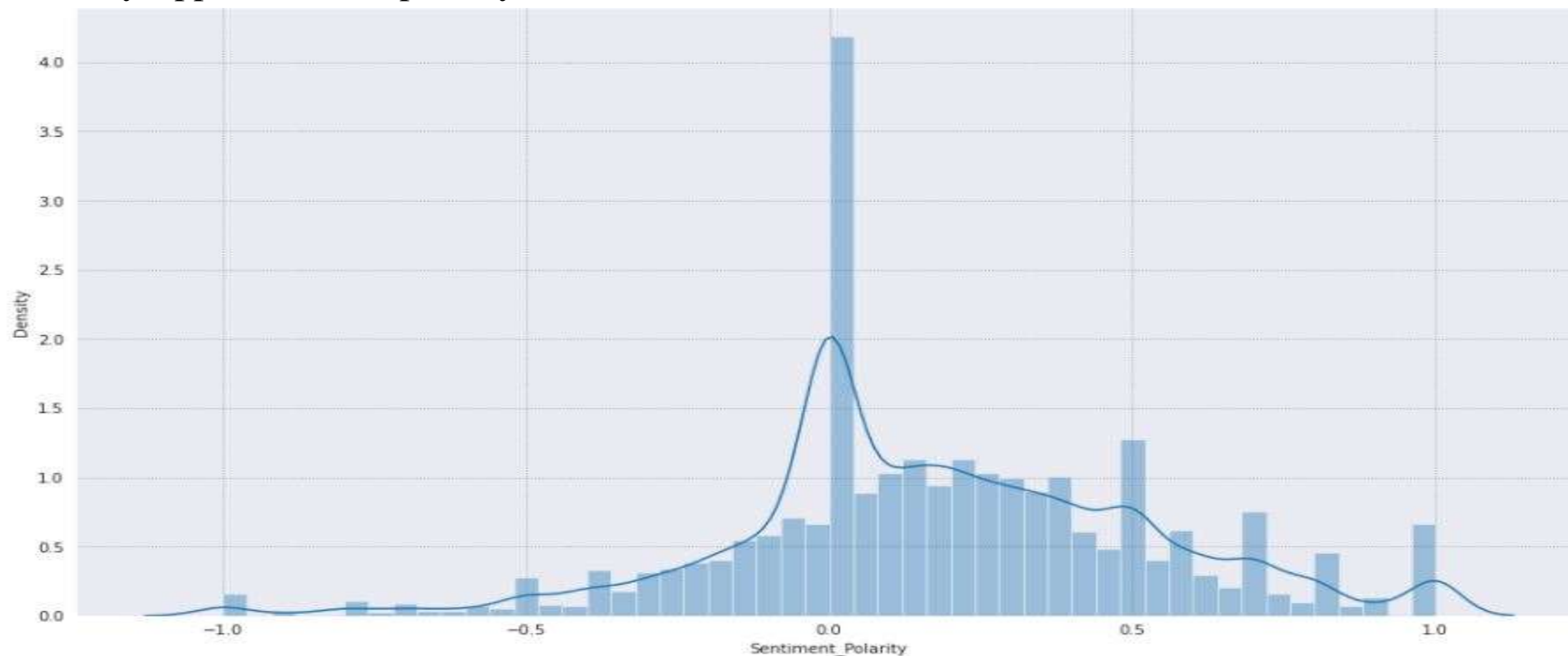
Correlations

We found that there is no correlation between app features like size, rating, number of installs and price. Surprisingly, it appears that there is no correlation between price and rating as well as between rating and whether an app contains ads or not.

<u>Summary</u>

Size
- Highly rated apps were optimally sized between 2MB to 40 MB
- Paid apps with specific functionality were lighter

Price
- Free apps outperform paid apps
- If an app is paid, the likelihood of it being popular is higher if it's priced under $10

Rating
- Positive relation between installs and rating.
- However, highly installed apps are not always highly rated.

Review
- Issues like loading time and positive features like usability were revealed by sentiment analysis.

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 decreases.

Thank You