

Capstone Project on

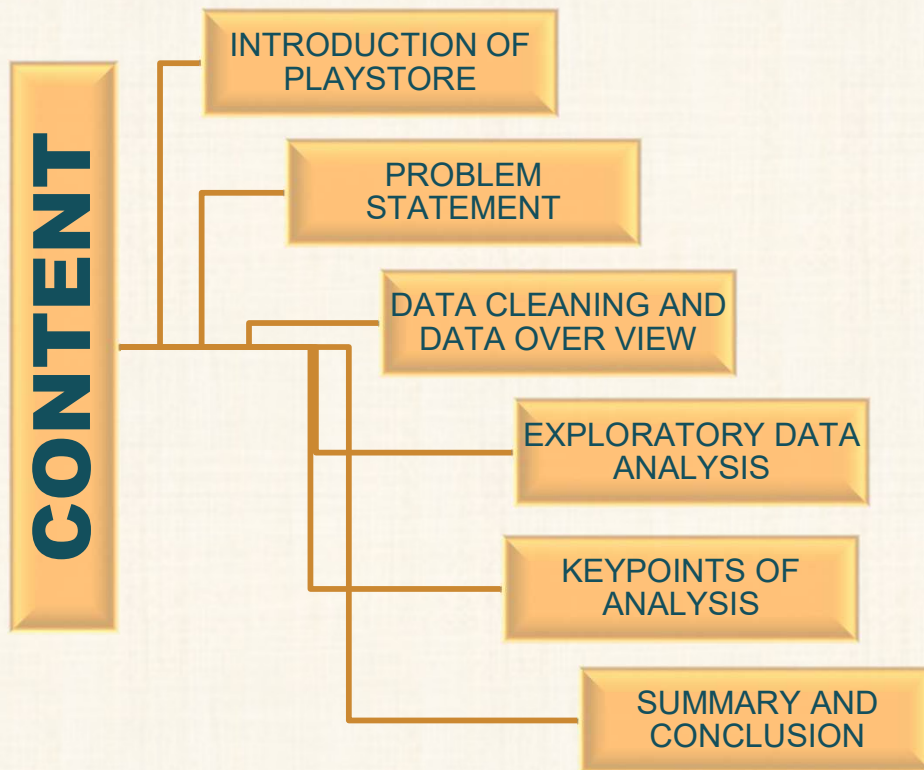
Play Store App Review Analysis



Individual done by Monica Patel

(Under the Guidance of Team Alma better)

Content



INTRODUCTION OF PLAYSTORE

- **Google Play**, also known as the **Google Play Store** and formerly the **Android Market**, is a digital distribution service operated and developed by Google. It serves as the official app store for certified devices running on the Android operating system and its derivatives, as well as ChromeOS, allowing users to browse and download applications developed with the Android software development kit (SDK) and published through Google. Google Play has also served as a digital media store, offering games, music, books, movies, and television programs.^[4] Content that has been purchased on Google Play Movies & TV and Google Play Books can be accessed on a web browser and through the Android and iOS apps.

PROBLEM STATEMENT

- **The Play Store apps data has enormous potential to drive app-making businesses to success. Android is expanding as an operating system and Mobile app industry is increasing significantly and thus giving rise to more competitions to the one's that are creating applications. Due to the competition in the market and also expansion in order to help our developer understand what kinds of apps are**
- **We aim on providing doing sentimental analysis on the apps that generated most positive and negative sentiments and sustainability of app in market on basis of previous data and current market.**

DATA OVERVIEW AND DATA CLEANING

- Dataset shape of Playstore –(10841 Rows and 13 columns)
- Dataset shape of user reviews –(64295 Rows and 5 columns)

PLAYSTORE DATA

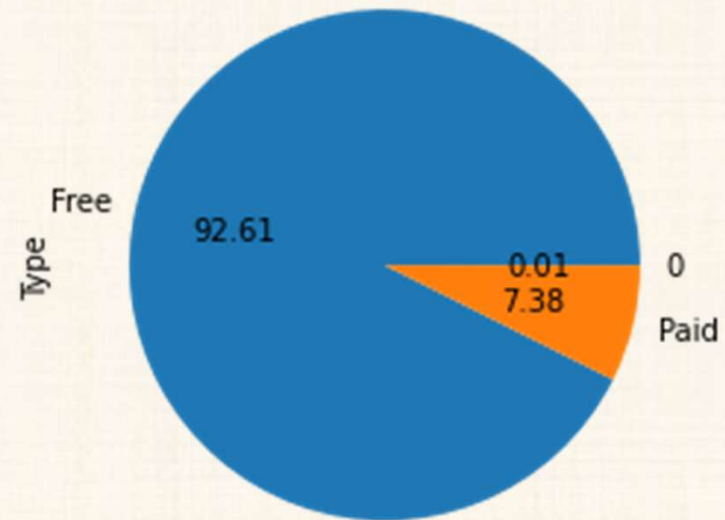
1 APP
2 CATEGORY
3 RATING
4 REVIEWS
5 SIZE
6 INSTALLS
7 TYPE
8 PRICE
9 CONTENT RATING
10 GENRES
11 LASTUPDATED
12 CURRENT VER
13 ANDROID VER

USER REVIEWS DATA

1 TRANSLATED
REVIEW
2 SENTIMENT
3 SENTIMENT
POLARITY
4 SENTIMENT
SUBJECTIVITY

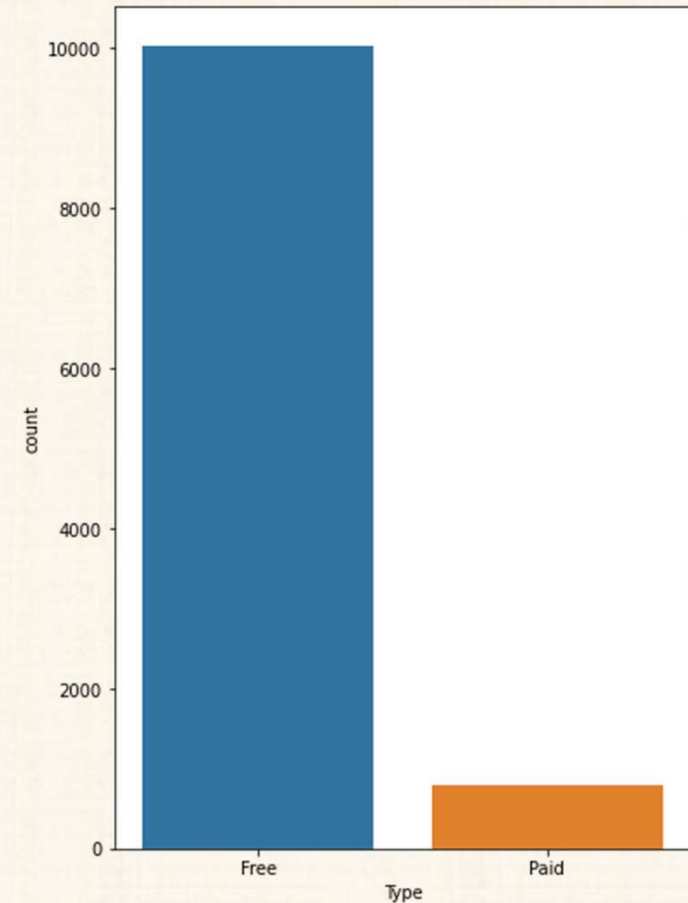
EDA (EXPLORATORY DATA ANALYSIS)

- **Percentage of free and paid app**



- Free apps are ruling the market by huge percentage that is 92.6%
- Percentage of paid apps present in Play store is 7.38% which is very low.

- **Size of Free and Paid Apps**
The size of paid apps are comparatively lower than Size of free apps, it may be because of paid apps developers charging for efficiency of apps and free app developers want to cut the cost of development by neglecting size efficiency.



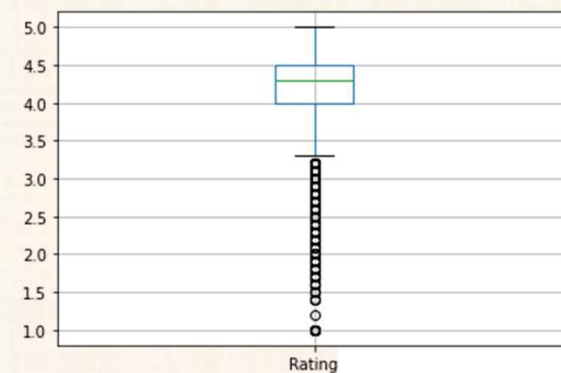
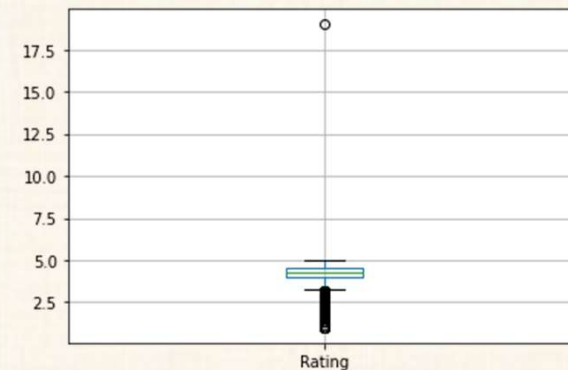
Num of apps In playstore for each category

- Number of Apps per Category
 - Most Apps belongs to 'Family' Category
 - Top 3 categories are- 'Family', 'Game' & 'Tools' for most of apps are developed
 - least number of apps are made for category 'Beauty' and 'comic'



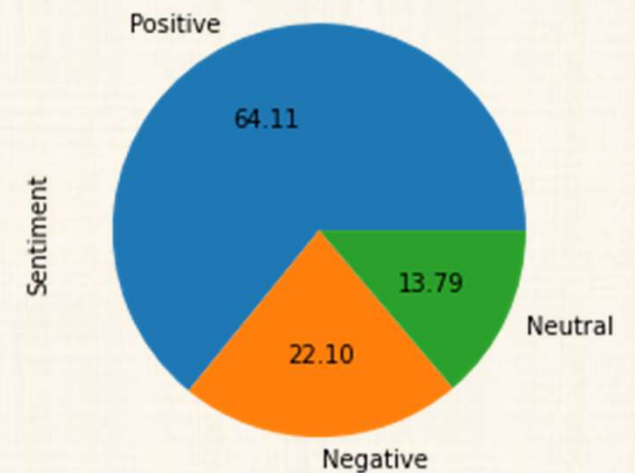
Play store outliers and after the outliers

- This is the outlier of playstore which has shown
 - Above the 17.5 so we clean the data and the
 - Result is below .
-
- After cleaning the outlier been removed .
 - `playstore_df.drop([10472], inplace=True)`



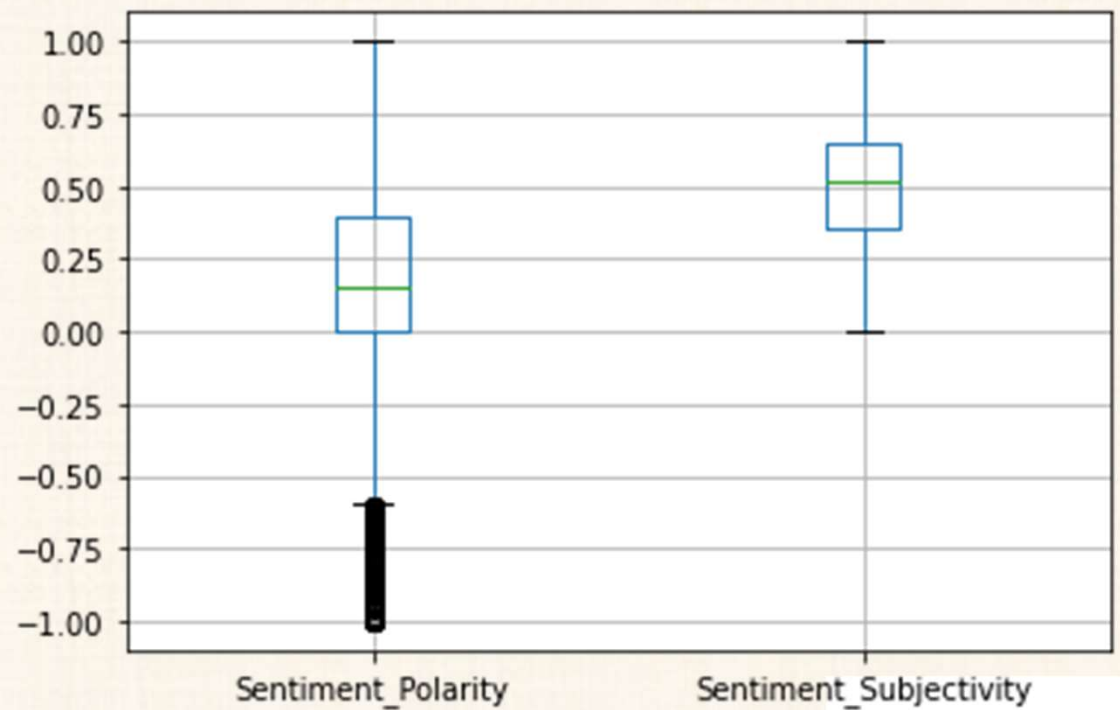
USER REVIEWS

- POSITIVE SENTIMENT – 64.11
- NEGATIVE -22.10
- NETUTRAL -13.79
- Sentiment analysis is a technique with which we can identify and determine if data indicates a positive, negative or neutral. It aids in understanding customer feedback and is mostly used by several companies to analyse brand and product reviews. It also helps in finding out the underlying sentiment in a text.
-



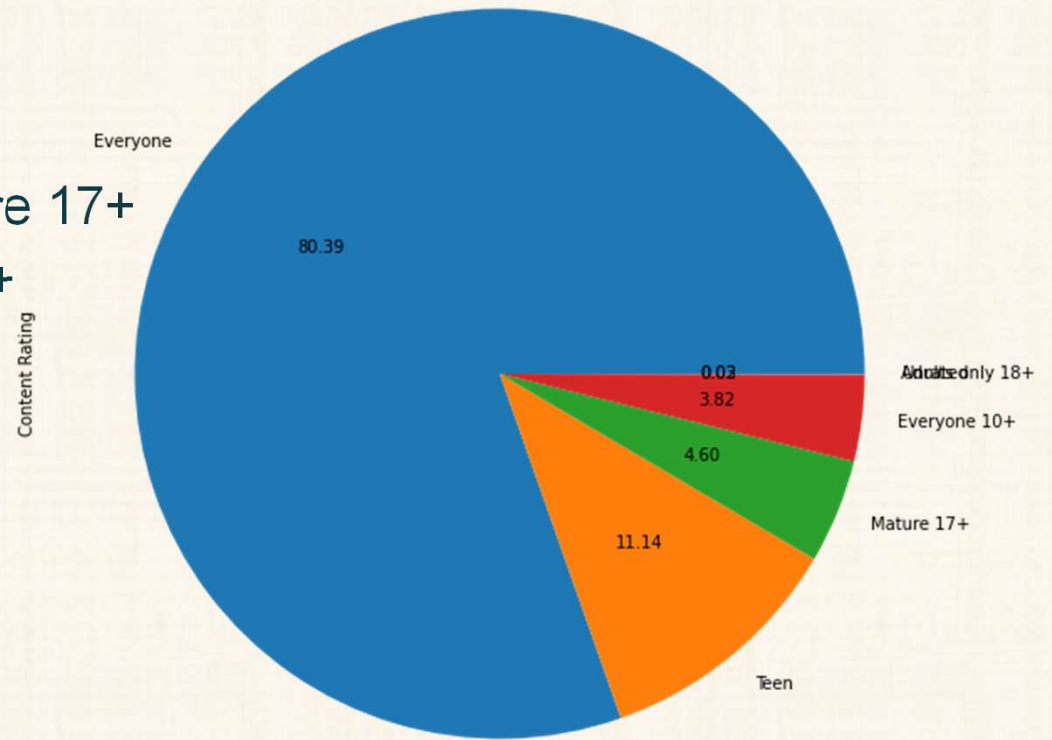
USER REVIEWS OF SENTIMENT POLARITY AND SENTIMENT SUBJECTIVITY

- Positive and Negative sentiments are more and Neutral are least

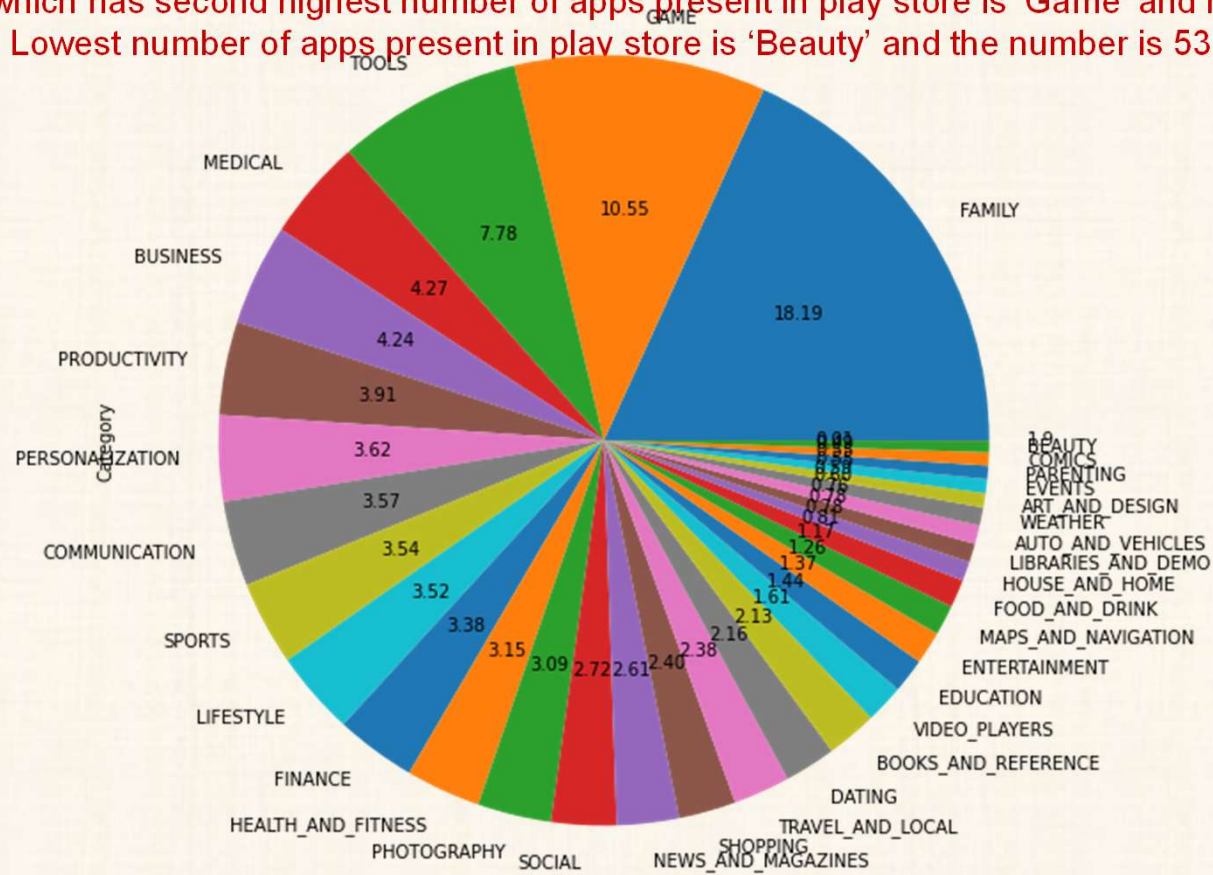


Distribution App of content rating

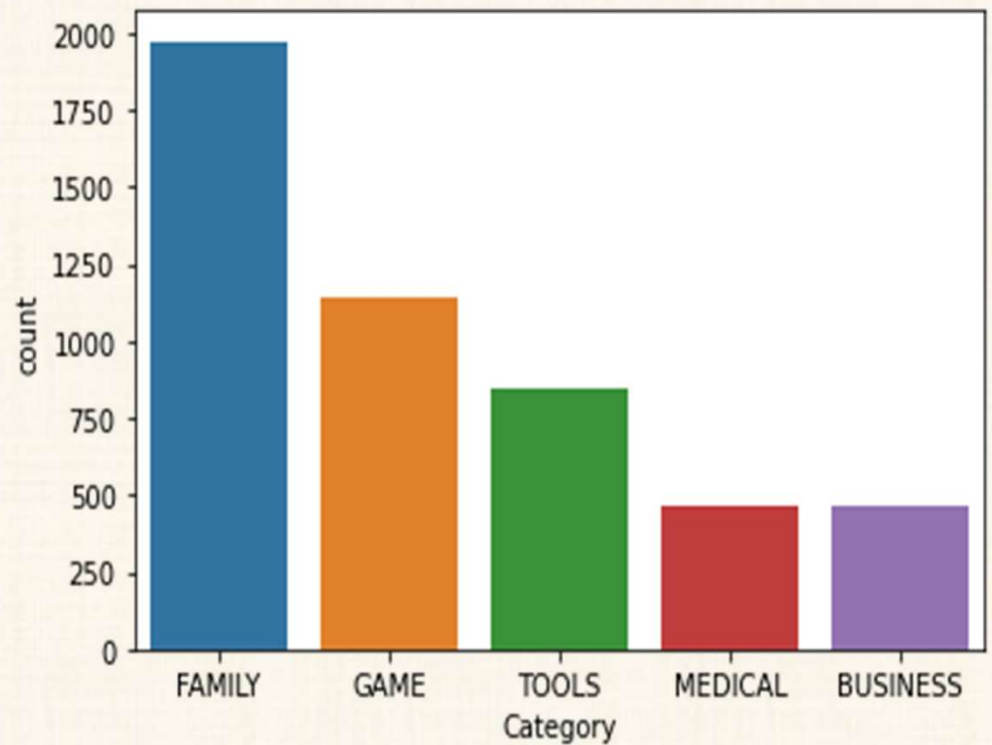
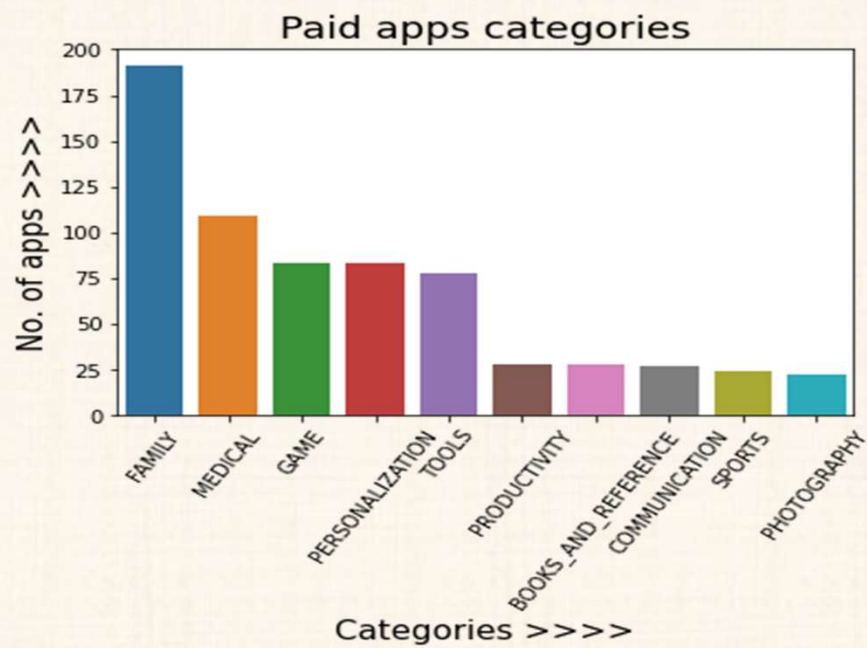
- 80.39 is highly used by everyone.
- 11.14 is for teen
- 4.60 mature person used which are 17+
- 3.82 use by everyone who are 10+
- 0.02 is for adult.



Family' category has highest number of apps in play store i.e. 1971 apps (18.2%).
 The category which has second highest number of apps present in play store is 'Game' and number is 1144 apps.
 Lowest number of apps present in play store is 'Beauty' and the number is 53 apps.



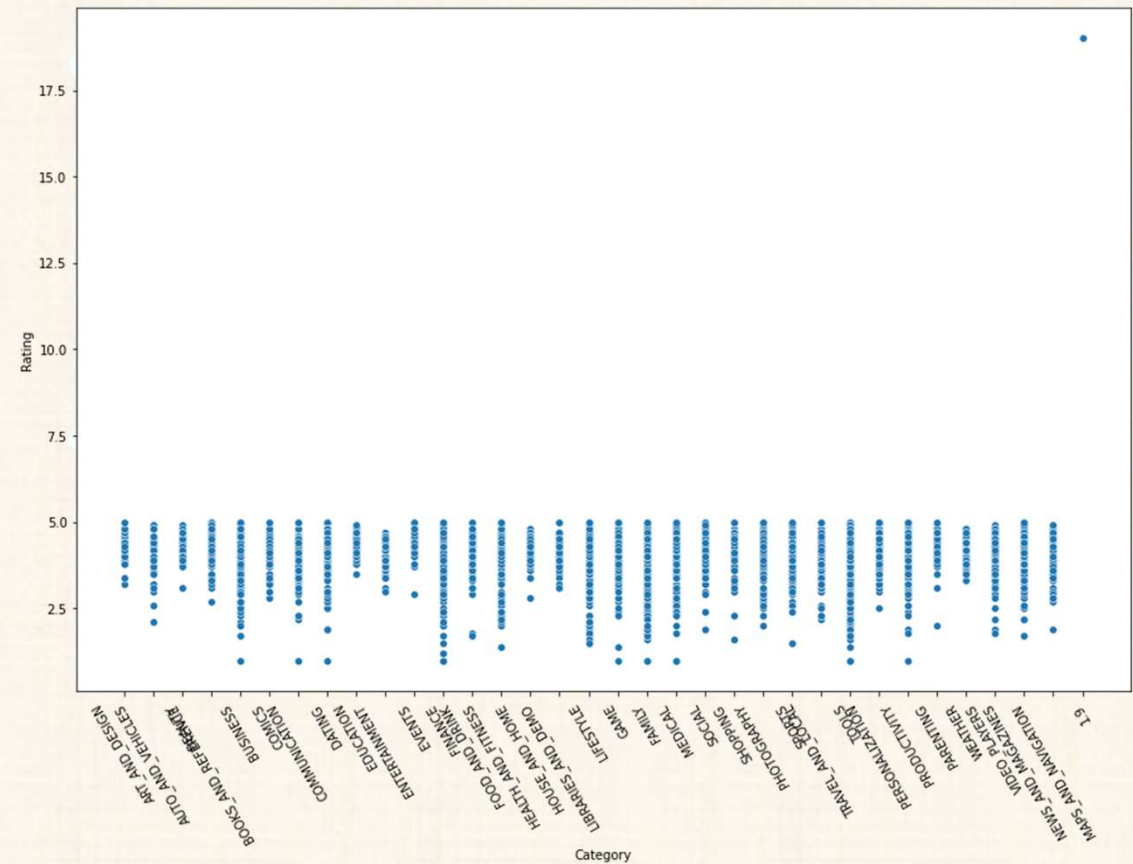
3 category from top 5 categories, Medical, Games and Tools are of specific micro-niche



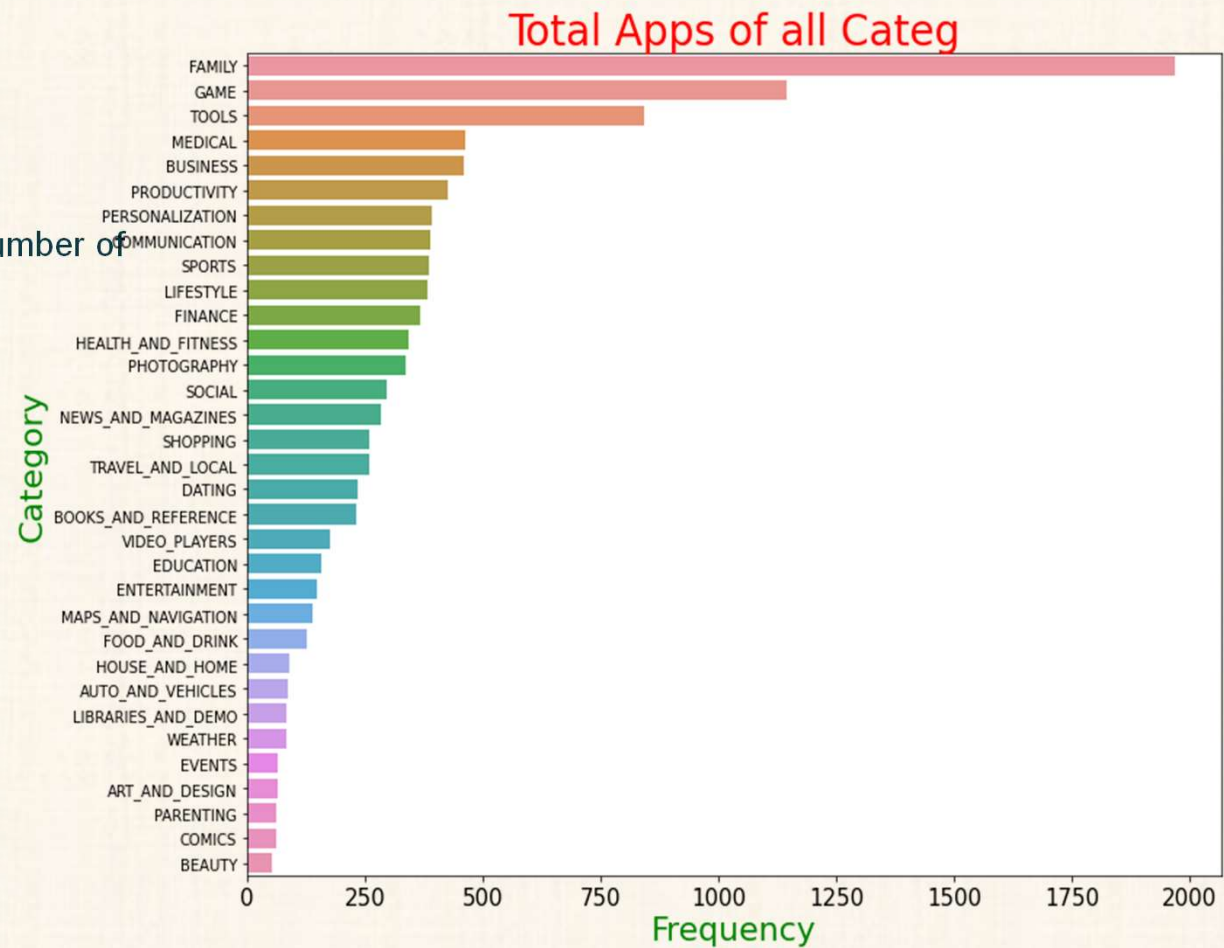
Category wise avg size of app

1. Game
2. Family
3. Parenting

1. Education
2. Travel And Local
3. Game

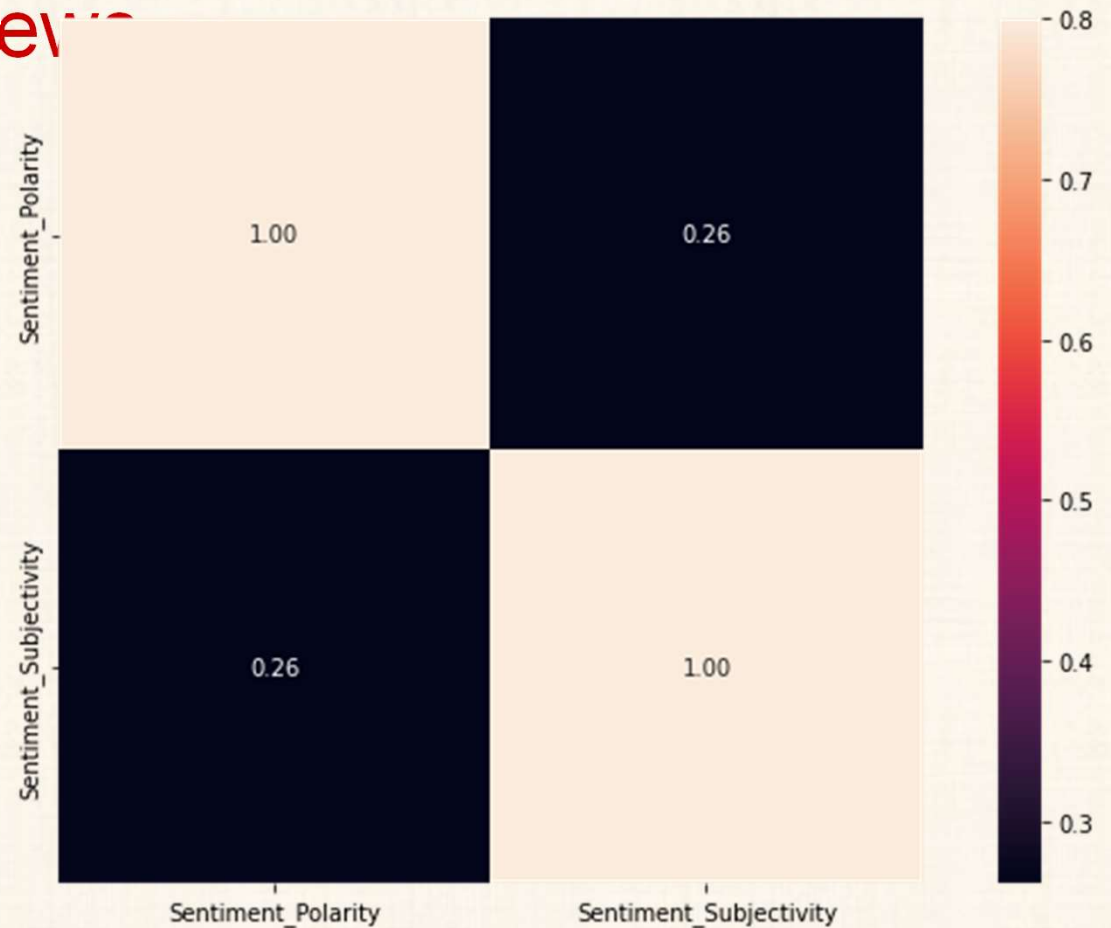


- Apps with 1,000,000 Installs
 - Game and Family category has most number of apps with 1000000+ installs
 - 'Photography' category also has higher downloads but it has low apps compare to other top categories



Correlation of user review

- Correlation between Sentiment polarity and sentiment subjectivity is
- 0.26 and 1.00 are positively as well as strong correlated with each other.



- **92.62% applications are present in play store is free and only 7.38% paid applications**
- **Android version required for the applications is increasing year by year.**
- **The average size of the application for all categories is lies between 5 – 10MB.**
- **18.2% of applications present in play store is from 'Family' category.**
- **The applications which belonging 'Social' category are most reviewed applications.**
- **Strong positive correlation in between reviews and installs.**
- **'64% users are positive sentiment and 21% are negative sentiment.**
- **'Comics', 'Beauty' and 'House and home' in this categories, only free applications are present.**

Summary and Conclusion

- **'Game' and 'Family' category are already dominating the market, but there are some low competitive categories which have lots of potential are- 'Beauty', 'Comic' and 'Photography' developers should focus on that.**

- **Paid apps are not suitable for broad users, they are meant to be made for a specific niche or segment and focus only on particular segment for better profits.**

- **For paid games, we seen a pattern that if it was already available for different platforms and is popular, then if it launch on play store, it will receive tremendous response from user. For example;- Minecraft, GTA vice city, Generated highest revenue.**

- **Size of app will impact user 's behaviour, so its best choice for developers to c with low size as possible.**

Thank You

shutterstock.com • 1263463690

AImaBetter

