Capstone Project – 1

Play Store App Review Analysis

By
Mohammed Saif Khan
Data Science Trainee, AlmaBetter

Problem Statement

Two datasets are provided, one with the **basic information** about the apps and the other with **user reviews** for the respective apps. We must examine and evaluate the data in both datasets in order to identify the important characteristics that influence app engagement and success.

So, what factors influence an app's success?
An app is said to be successful if it has:
- ❖ A high number of installs
- ❖ A high average user rating
- ❖ A good number of positive reviews
- ❖ High revenue per customer

Data Summary

Columns in Apps Dataframe:

- ❖ **App**
- ❖ **Category**
- ❖ **Rating**
- ❖ **Reviews**
- ❖ **Size**
- ❖ **Installs**
- ❖ **Type**
- ❖ **Price**
- ❖ **Content Rating**
- ❖ **Genres**
- ❖ Last Updated
- ❖ Current Ver
- ❖ Android Ver

Columns in Reviews Dataframe:

- ❖ **App**
- ❖ Translated_Review
- ❖ **Sentiment**
- ❖ Sentiment_Polarity
- ❖ Sentiment_Subjectivity

Agenda

- ❖ Correlation Heatmap
- ❖ Categorical Analysis
- ❖ App Rating Analysis
- ❖ App Reviews Analysis
- ❖ App Size Analysis
- ❖ Free and Paid Apps Analysis
- ❖ Revenue Analysis
- ❖ Content Rating and Genres Analysis
- ❖ Reviews and Sentiment Analysis
- ❖ Analysis Summary

Correlation Heatmap

- ❖ There is a strong positive correlation between Reviews and Installs.
- ❖ Price is slightly negatively correlated with Rating, Reviews, and Installs.
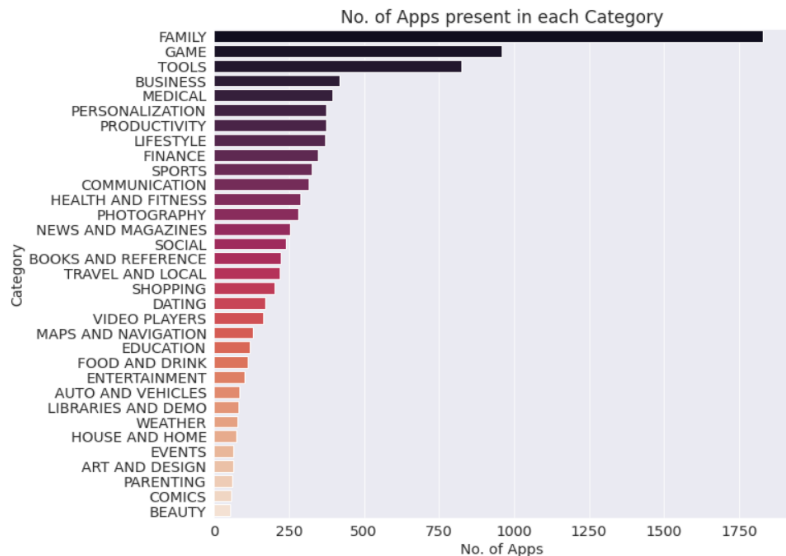- ❖ Rating is slightly positively correlated with Installs and Reviews.
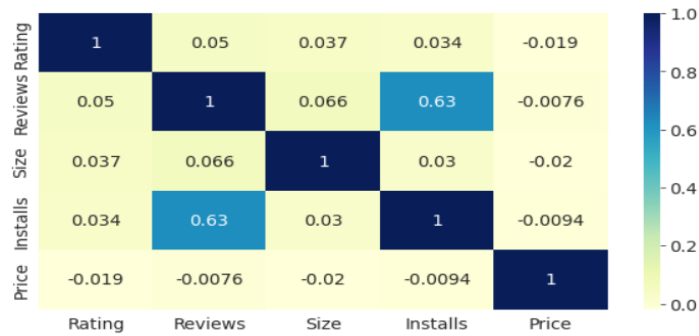
- ❖ Sentiment Polarity is slightly positive correlated with Sentiment Subjectivity.
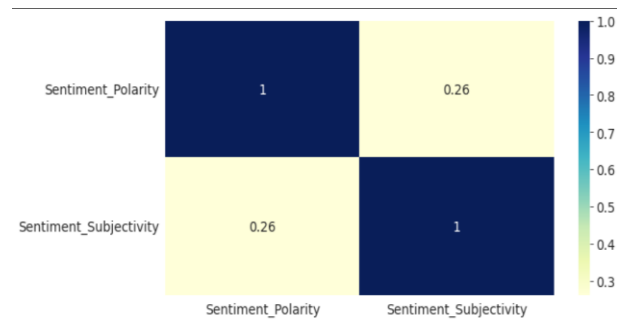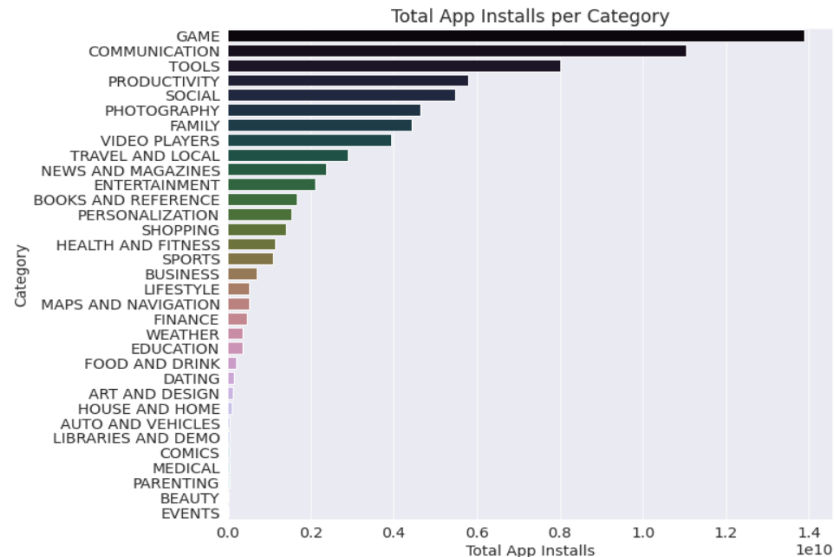
Categorical Analysis

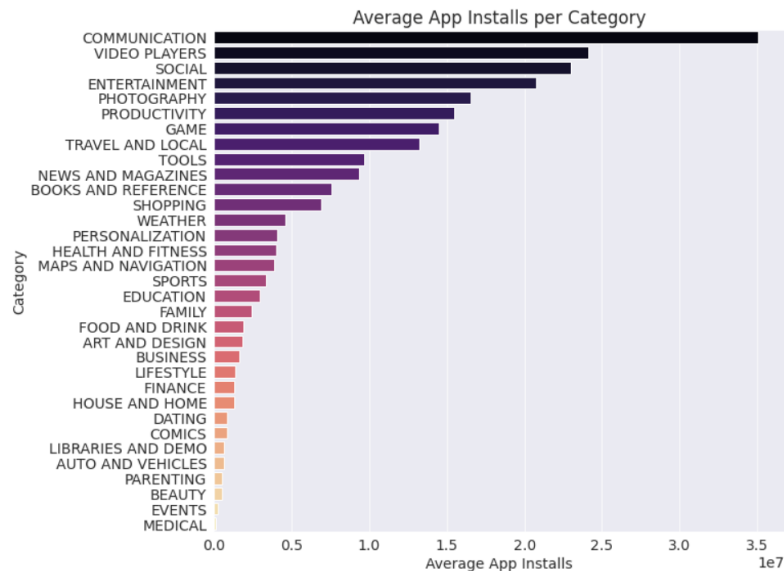❖ Family, Game and Tools category have the highest number of apps.

❖ Game, Communication and Tools category have the highest number of app installs.

No. of Apps present in each Category

Total App Installs per Category

Categorical Analysis (Contd.)

❖ Communication, Video Players and Social category have the highest number of average app installs

❖ The paid apps in the Finance, Lifestyle, and Events category are on average significantly more expensive than the paid apps in other categories.
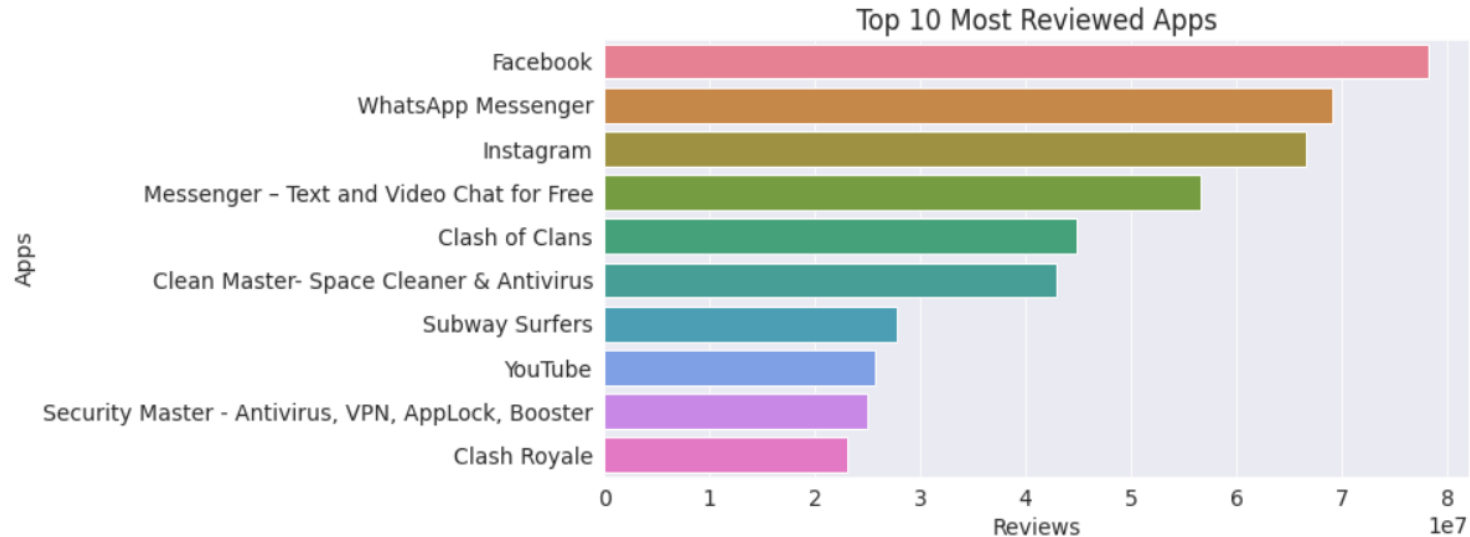

Average App Installs per Category

Average price of paid apps in each category
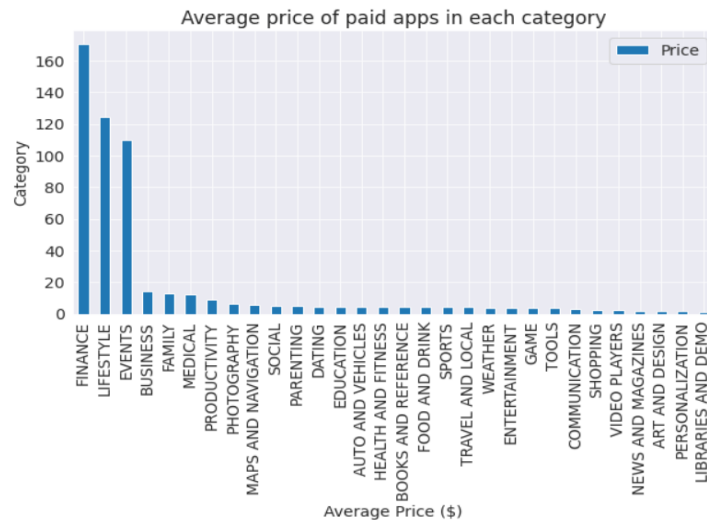
App Rating Analysis

❖ The Distribution of the Rating is the most between 3.5 to 4.8. The peak of the distribution is at 4.3.

Distribution of Rating
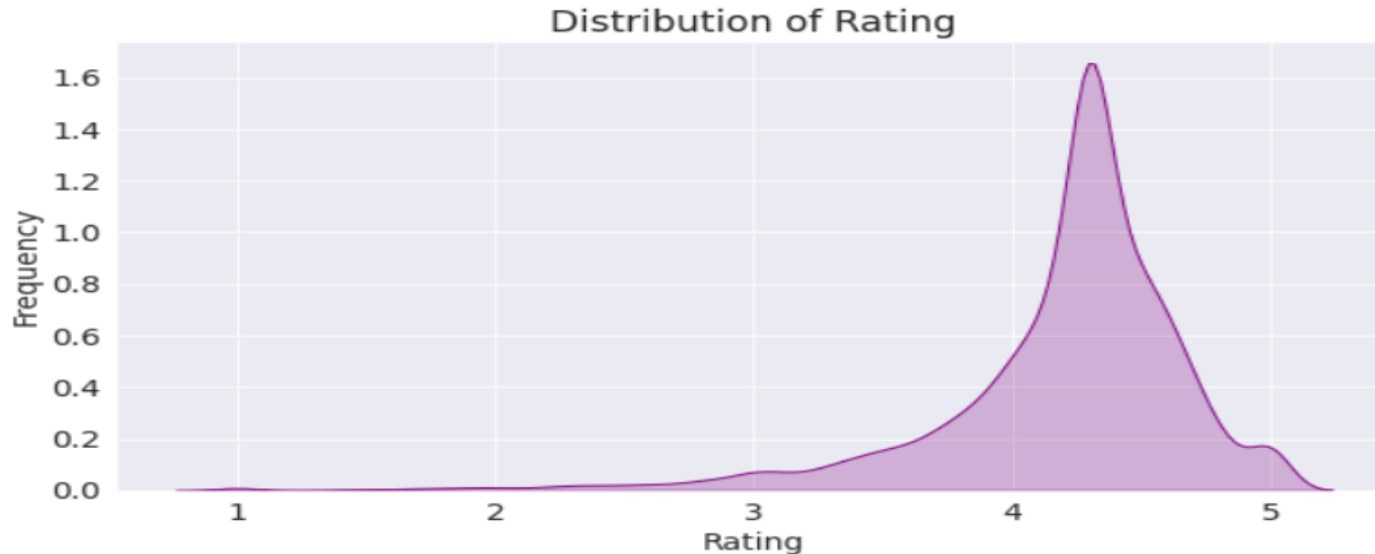
App Reviews Analysis

❖ Facebook, WhatsApp Messenger and Instagram are the most reviewed apps.

Top 10 Most Reviewed Apps

App Size Analysis

- ❖ The apps are categorized based on its size between ~0 to 100 MB in the intervals of 10 MB each.
- ❖ The total number of apps in each size category indicates the competition.
- ❖ Average number of user reviews and average app installs in each size category indicates the popularity of the respective app.
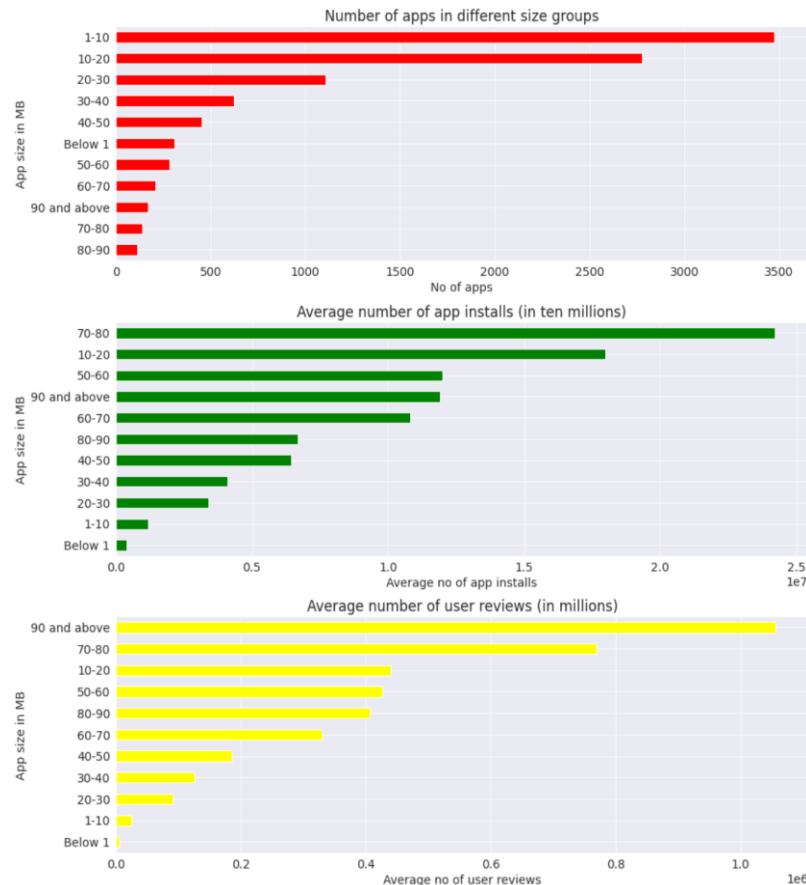
Free and Paid Apps Analysis

❖ In the dataframe there are 92.17% Free apps and 7.83% Paid apps.

❖ The top categories in which top free apps fall are Communication(6), Social(3), Video Players(2), Travel and Local(2).

❖ The top categories in which top paid apps fall are Finance(7), Lifestyle(6), Family(4).

Categories in which the top 20 free apps belong

Categories in which the top 20 paid apps belong

Percentage of Free and Paid Apps
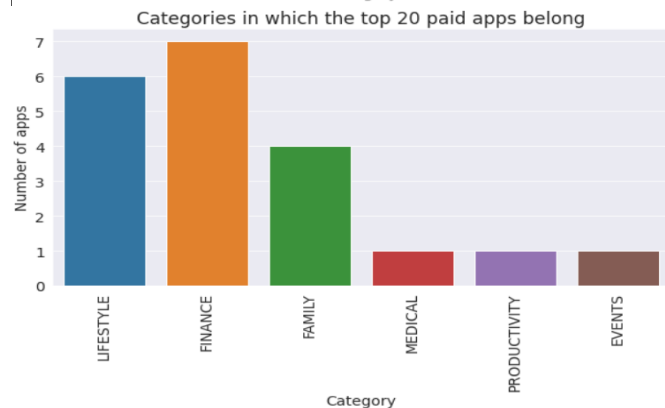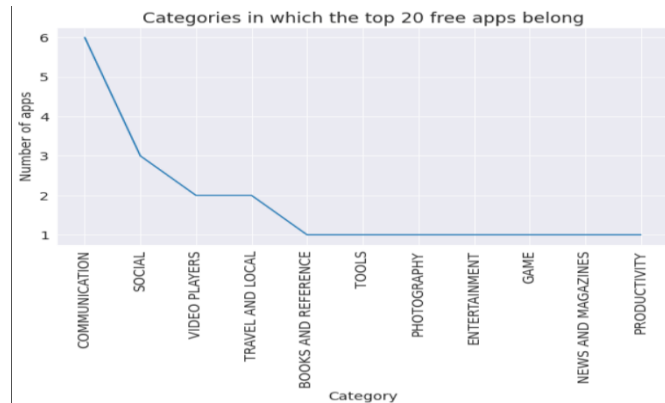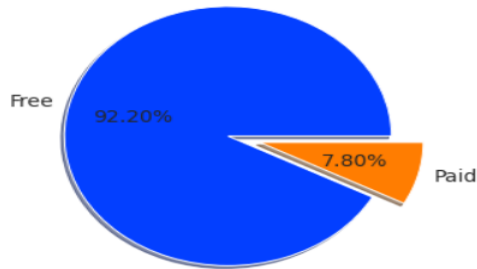
Revenue Analysis

- ❖ Revenue generated is given by the formula: Revenue = Installs * Price
- ❖ Note that in this case, revenue refers to the money earned only from paid app installs.
- ❖ The top categories in which these apps fall are Lifestyle(5), Family(5), and Game(4).
- ❖ Minecraft, I am rich, and I am rich premium are the top paid apps based on revenue generated.
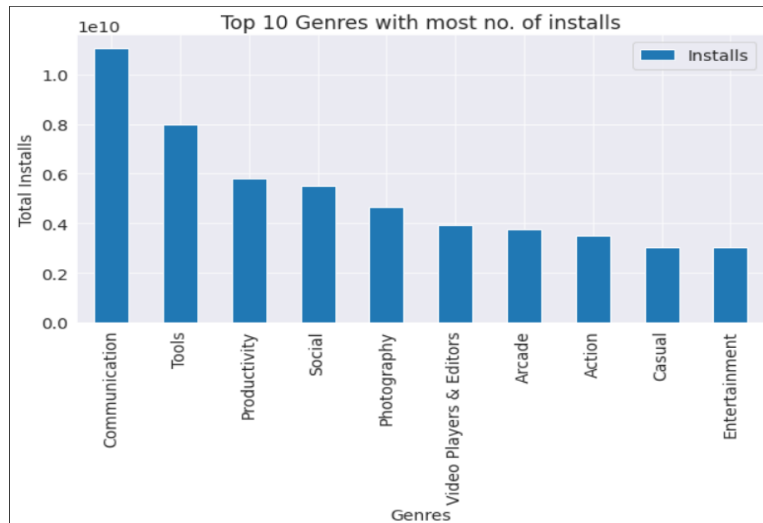- ❖ Minecraft is the only app that has over 10M installs.

Most Revenue Generating Apps
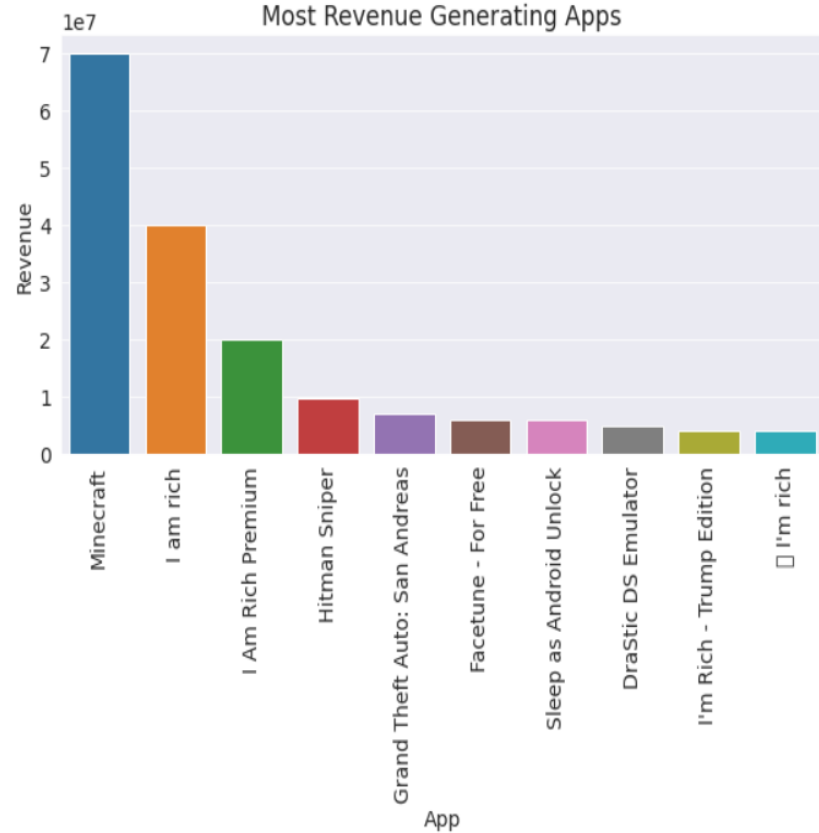
Content Rating and Genres Analysis

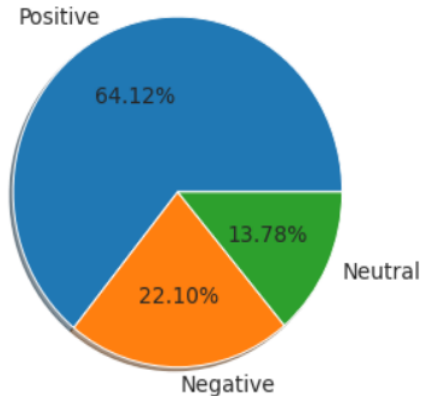❖ 81.8% of the apps in the dataframe are from Content Rating 'Everyone'.
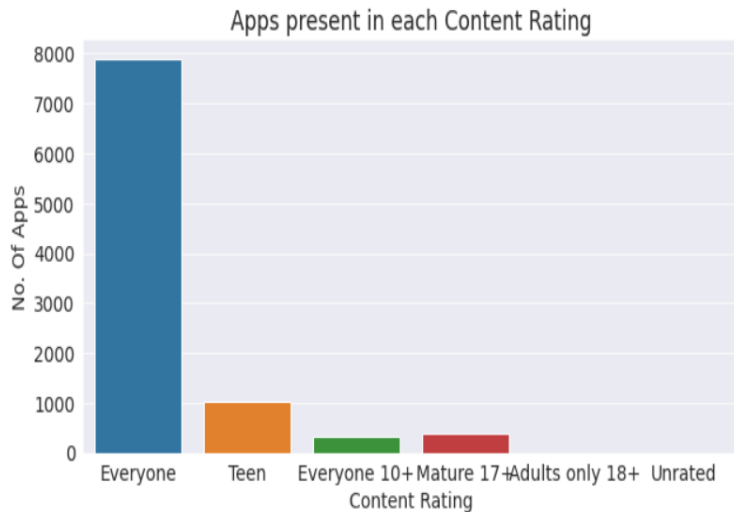
❖ Communication, Tools and Productivity genres are the genres with most no. of installs.

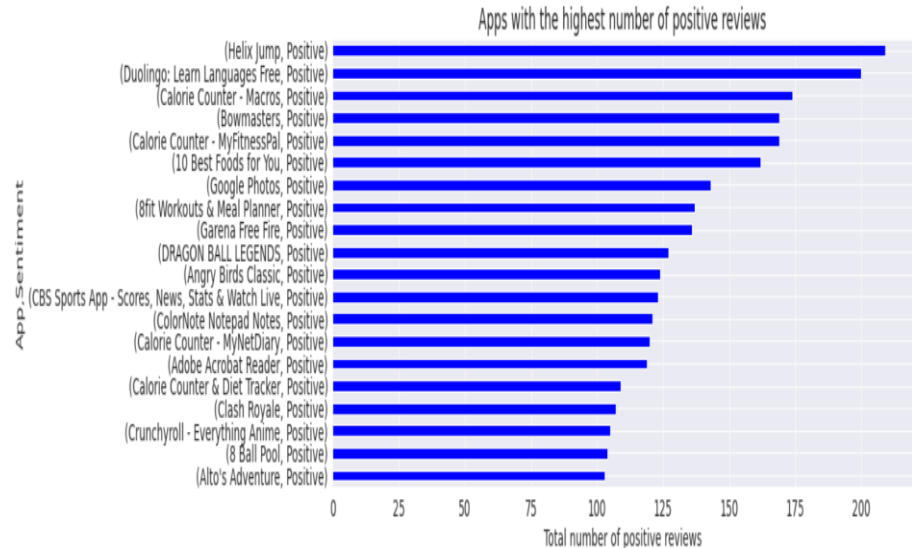# Reviews and Sentiment Analysis

❖ In the dataframe there are 64.12% Positive Sentiments, 22.10% Negative Sentiments and 13.78% Neutral Sentiments.

❖ Helix Jump has the most no. of positive reviews.

Percentage of Positive, Negative and Neutral Sentiments

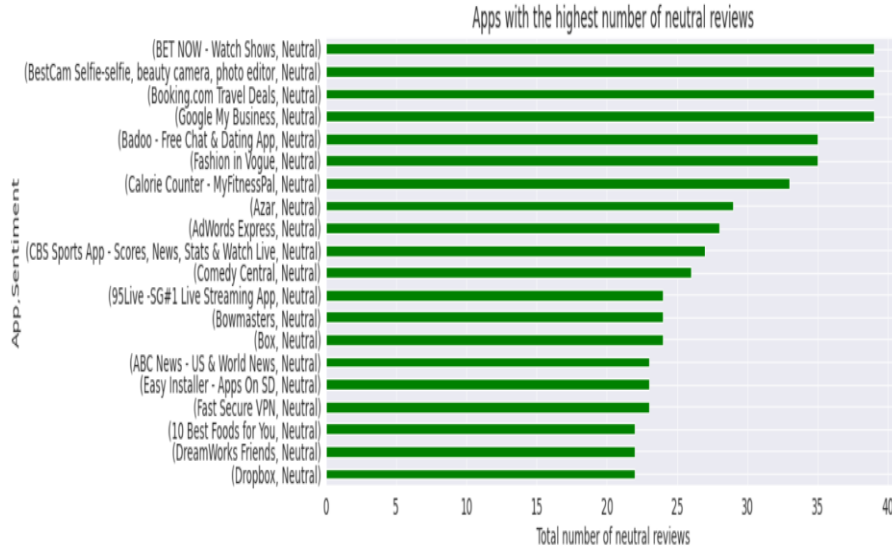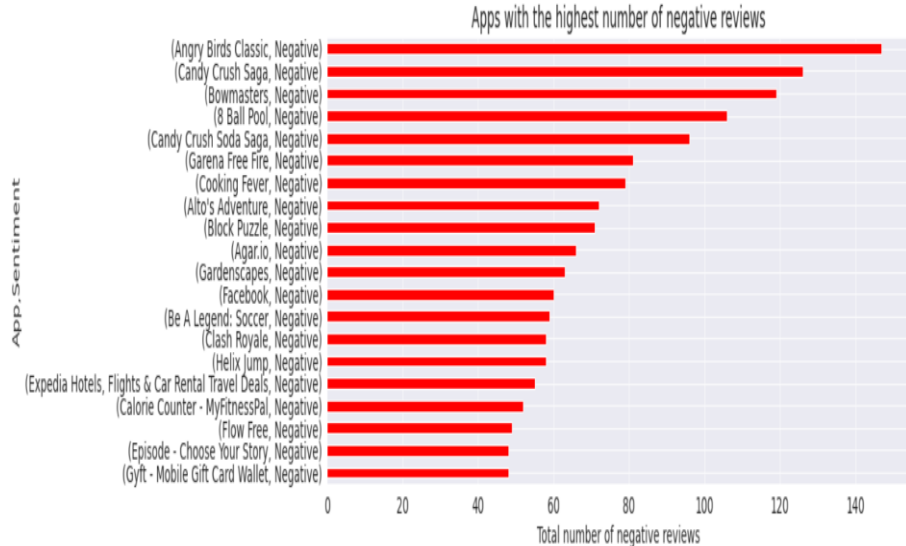Apps with the highest number of positive reviews

Reviews and Sentiment Analysis (Contd.)

❖ BET NOW – Watch Shows has the most no. of neutral reviews.

❖ Angry Birds Classic has the most no. of negative reviews.

Apps with the highest number of neutral reviews

Apps with the highest number of negative reviews

Final Takeaway

- ❖ Most competitive category: Family
- ❖ Category with the highest number of installs: Game
- ❖ Category with the highest average app installs: Communication
- ❖ Percentage of free apps = ~92%
- ❖ Percentage of apps with no age restrictions = ~82%
- ❖ Most distribution of rating is between: 3.5 to 4.8.
- ❖ Median rating of the apps in the play store is 4.3.
- ❖ There are 20 free apps that have been installed over a billion times.
- ❖ The median size of all apps in the play store is 12 MB.
- ❖ App(Paid) that generated the most revenue: Minecraft

Final Takeaway (Contd.)

❖ On average the most expensive category: Finance

❖ Most popular app in the Play Store based on the number of reviews: Facebook

❖ Genre with the most number of installs: Communication

❖ The apps whose size is greater than 90 MB has the highest number of average user reviews, i.e, they are more popular than the rest.

❖ Percentage of Positive app reviews = ~64%

❖ Helix Jump has the highest number of positive reviews and BET NOW - Watch Shows has the highest number of neutral reviews and Angry Birds Classic has the highest number of negative reviews.

AI

Thank You