Capstone Project

EDA ON Play Store App Review

The Play Store app data has enormous potential to drive app-making businesses to success.
Taking into account billion of Android users worldwide, mining this data has the potential to reveal user behavior and trends in the whole global scope.

We have a similar data set today obtained by scraping Google Play Store. Let's deep dive into the data set and try to gather some useful intel from it

❑ Problem Statement:

As we have discussed earlier, **Play Store has the potential to drive app making businesses to the sky but Play Store can only provide us with the Platform apart from that there certainly are a lot of other factors that affect the success of an Application, what are they?**

An app must have the following factors to hit the success mark:

➢ **Huge Number of Installs**
➢ **A lot of Positive Reviews**
➢ **High Revenue Generation**
➢ **High Ratings by Users**
➢ **And many more**

☐ Data Summary:

We were given two data sets

Play_Store_Data

- ➤ **App**: indicates the name of the app.
- ➤ **Category**: indicates category of the app.
- ➤ **Size**: indicates the size of the apps, in megabytes(M) and kilobytes(K).
- ➤ **Rating**: ratings received from the user, ranging from 0 to 5
- ➤ **Reviews**: the number of reviews given by the users .
- ➤ **Installs**: indicates the number of Installs.
- ➤ **Type**: displays whether the app is free or paid
- ➤ **Price**: indicates the cost of an app
- ➤ **Content Rating**: indicates the age group for particular apps.
- ➤ **Genres**: indicates the genres of apps
- ➤ **Last Updated**: indicates the last update date and time
- ➤ **Current Ver**: indicates current version of the app
- ➤ **Android Ver**: indicates the compatibility version for apps

User_reviews

- ➤ **App**: indicates the name of the app.
- ➤ **Translated Review**: indicates the reviews written by users.
- ➤ **Sentiment**: either the review is Positive, Negative, or Neutral.
- ➤ **Sentiment_Polarity**: sentiment in numerical form ranging from -1.00 to 1.00
- ➤ **Sentiment_Subjectivity**: explains how likely that sentiment is based upon data, personal opinions , feelings or public opinions.

➢ So we will divide our workflow into the following 3 Analysis:

Univariate Analysis	Bivariate Analysis	Multivariate Analysis

1) **Univariate analysis:** Univariate analysis is the simplest of the three analyses where the data you are analyzing is only one variable.

2) **Bivariate analysis:** Bivariate analysis is where you are comparing two variables to study their relationships.

3) **Multivariate analysis:** Multivariate analysis is similar to Bivariate analysis but you are comparing more than two variables.

Correlation Heatmap:

- A **Positive Correlation** was seen between **Reviews** and **Installs**

- Also a **Negative Correlation** was seen between **Ratings**, **Reviews**, & **Installs** with the Price.

CONCLUSION:

- The Positive Correlation between Reviews & Installs is no exception as it is Normal human behavior to only go after the things that are highly praised of.
- But the Negative correlation here clearly shows that People are not satisfied with the services of paid apps, therefore working on the Reviews and then offering somewhat discounts over those apps can be a great fix
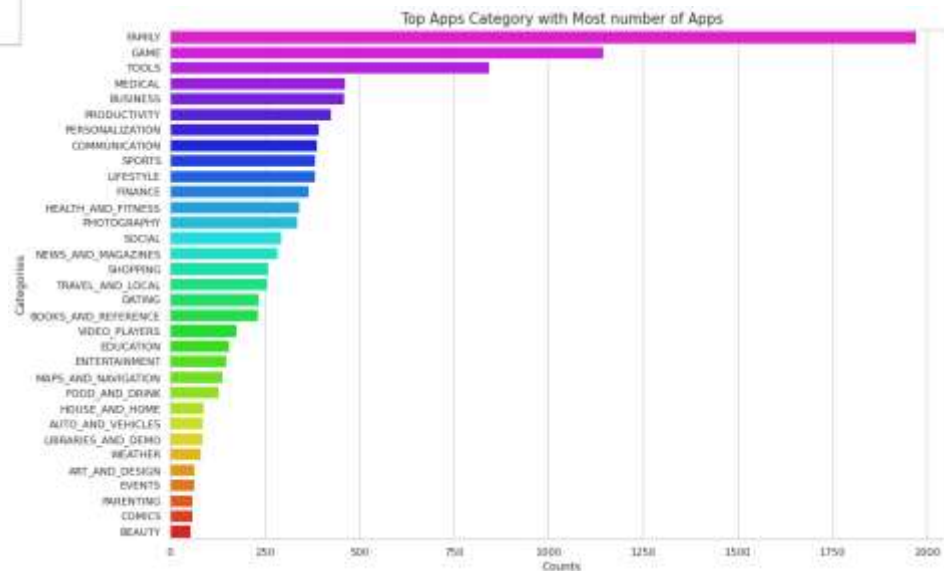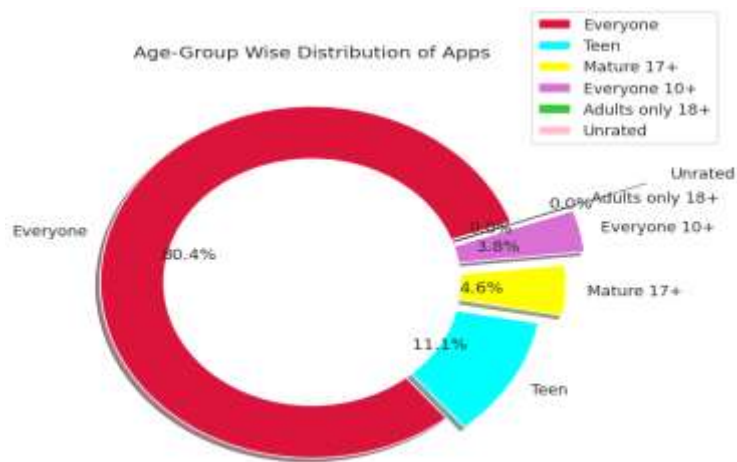
Top 10 Genres apps in Playstore
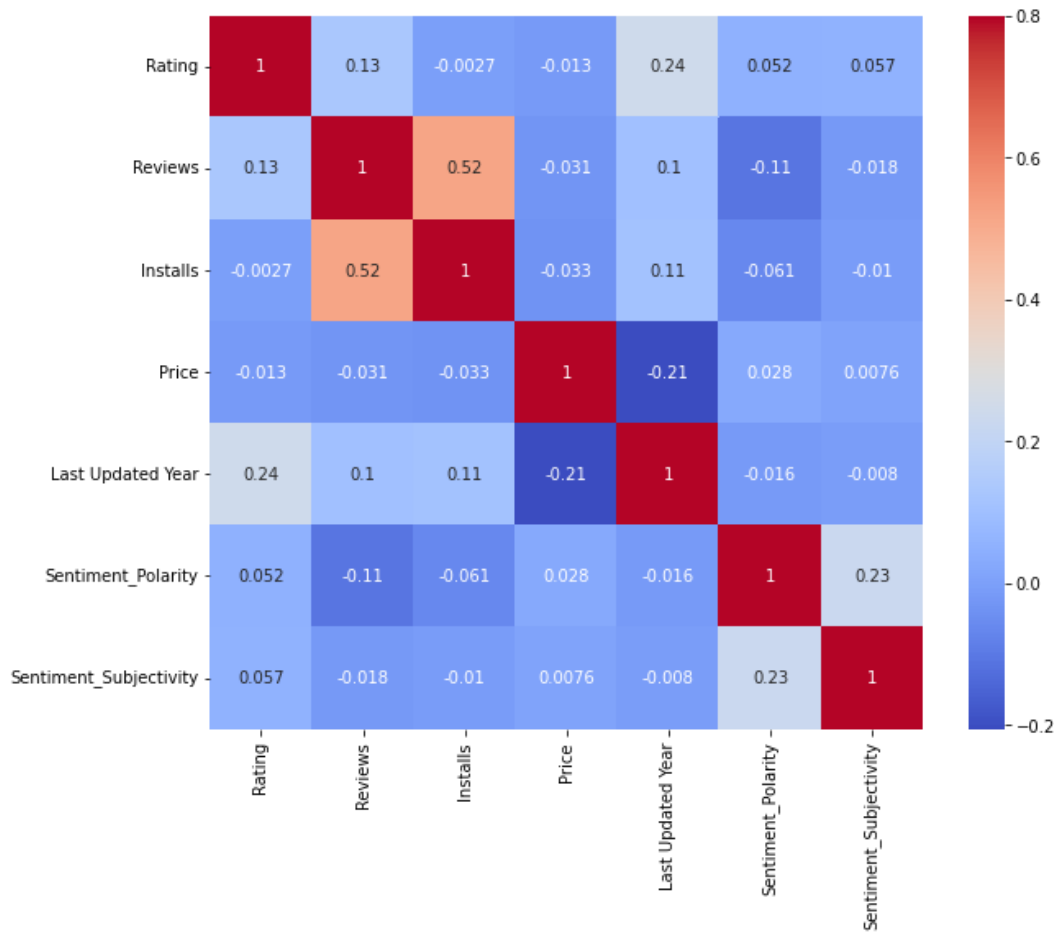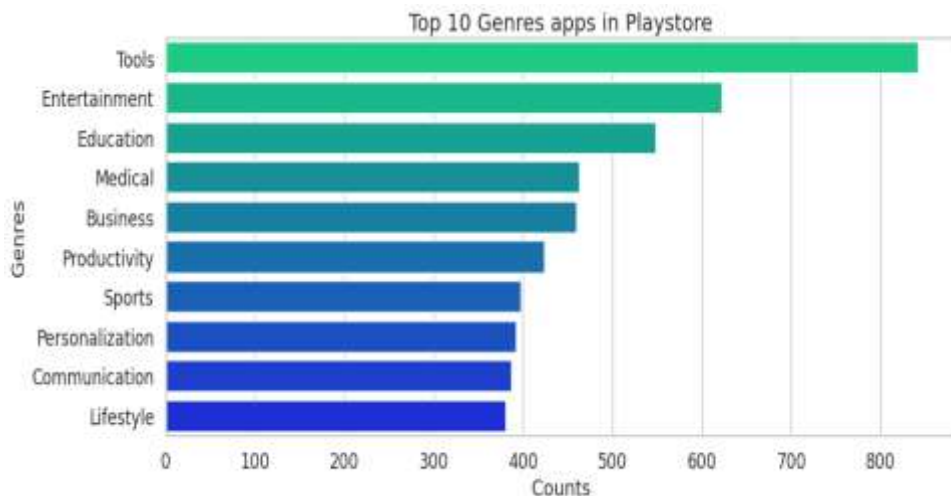
TOP GENRES IN PLAY STORE

➢ **Tools** genre is at the top with around **842** counts
➢ Followed by **Entertainment** , **Education** and others

Top Apps Category with Most number of Apps

Most Listed Apps Over Play Store

➢ **Family Category** has the most number of apps listed on the Play Store i.e **1972**

➢ Followed by **GAME** with a total of **1144** apps

➢ And others...

Age-Group Wise Distribution of Apps

Legend: Everyone, Teen, Mature 17+, Everyone 10+, Adults only 18+, Unrated

Everyone 80.4%, Teen 11.1%, Mature 17+ 4.6%, Everyone 10+ 3.8%, Adults only 18+ 0.0%, Unrated 0.0%
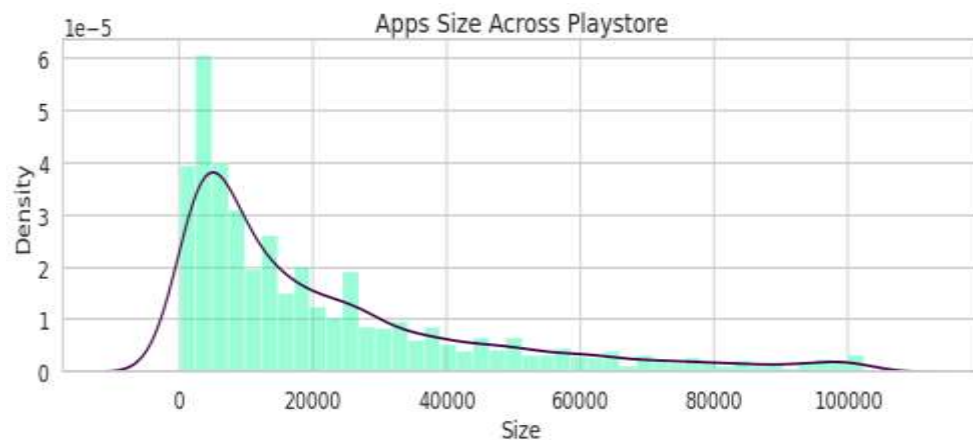

Apps Size Across Playstore

Age-Group-wise Distribution of Apps

➤ Most of the apps i.e **80.4%** in Play Store are for **Everyone**

CONCLUSION

➤ Most of the apps do not have any age limitations, which indeed is a great move to attract people of every age group.
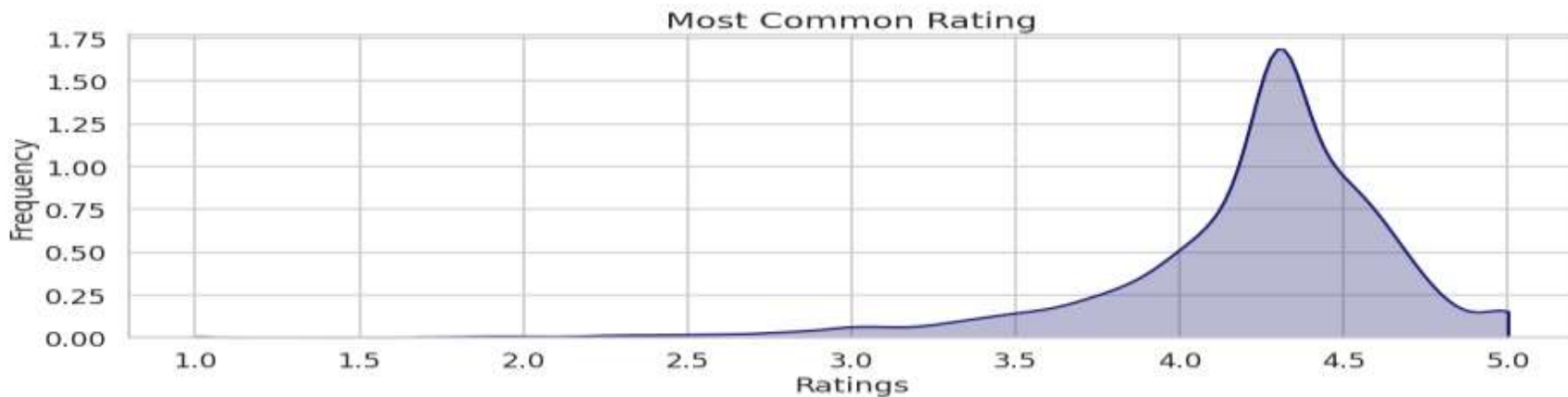
Average Size of Apps on Play Store

➤ The **Average Size** of Apps on the Play store is around **20MB**

CONCLUSION

➤ Lower size results in people downloading those apps without worrying about any storage issues, which may sometime be useful to bring in more Installs.

Most Common Rating

Most Common Rating

So, the average Rating for the Majority of Apps in the Play store is from **4.3** to **4.6/4.7** out of 5.

CONCLUSION
➤ **These Ratings clearly show the success rate of apps over the play store as clearly more than 70% of the apps are highly rated by the users**

Ratings and reviews

4.4
★★★★½
9,693,354
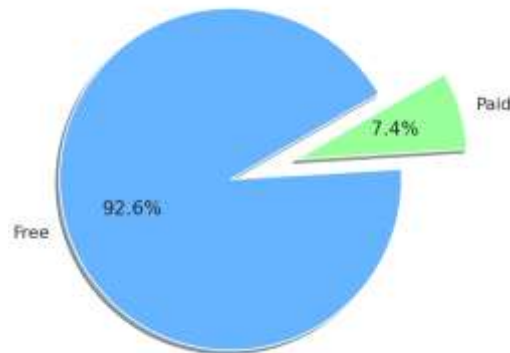
Gameplay 4.2 ★   Graphics 4.2 ★   Controls 4.1 ★

Bivariate Analysis:

Percentage of Free apps with the Paid ones

- ➤ **≈93%** of the apps in the google play store are **Free**
- ➤ And around **≈7%** are the **Paid** ones

CONCLUSION
- ➤ Reason for this behavior is the Negative Correlation between **Ratings**, **Reviews**, & **Installs** with the **Price** as we have analyzed earlier. We can conclude that the Success Rate of Paid Apps is less but never negligible.

Percentage of Free apps and Paid apps

Categories with the Most Number of Installs

GAME category is at the top with the Most Number of **Installs** closely followed by **COMMUNICATION** then **PRODUCTIVITY** and others

CONCLUSION
- ➤ As we have seen before, the FAMILY Category even After holding the most number of apps listings over the Play store, couldn't beat GAME Category as Most of the apps are downloaded from this category...clearly indicating the success ratio of GAME Category
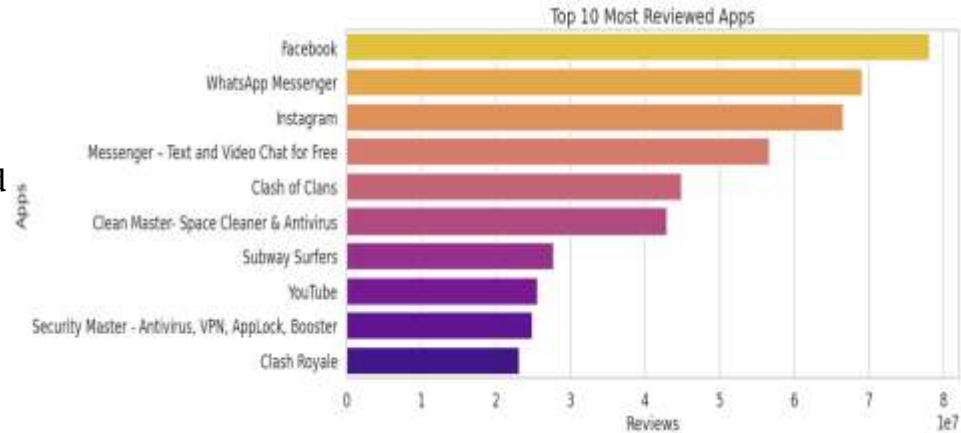

Top Apps Category with Most number of Installs
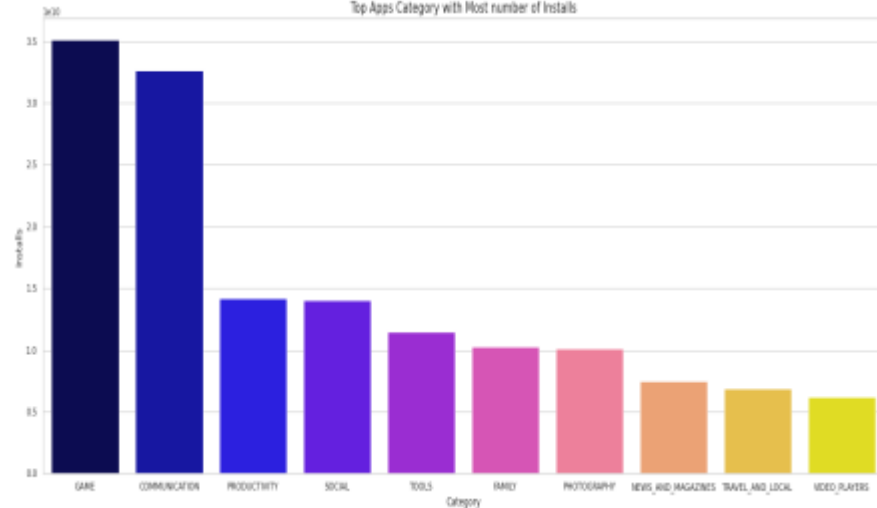
Apps with the Most Reviews

➤ **Facebook** is the **Most Reviewed** app on the play store with a total of over **78 Million** reviews

➤ Followed by **WhatsApp Messenger**, **Instagram** and others..
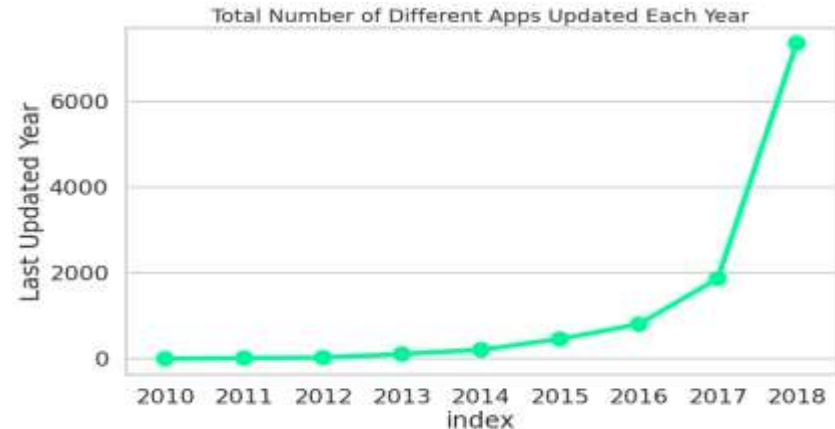
CONCLUSION

➤ Connecting and Entertaining People is always Fruitful as we can see that TOP apps serve their purpose of connecting people.

Top 10 Most Reviewed Apps

App Updates Each Year

➤ The Year **2017 & 2018** has the most number of different app updates...

➤ And can also be considered the **Busiest year** the for play store!

CONCLUSION

➤ Year Updation plot shows the number of apps being updated is increasing with the years, and from 2017 to 2018 it has increased drastically, explaining the interaction of these apps with the people in the upcoming years.

Total Number of Different Apps Updated Each Year

☐ **Multivariate Analysis:**

Top 10 Most Profitable Apps
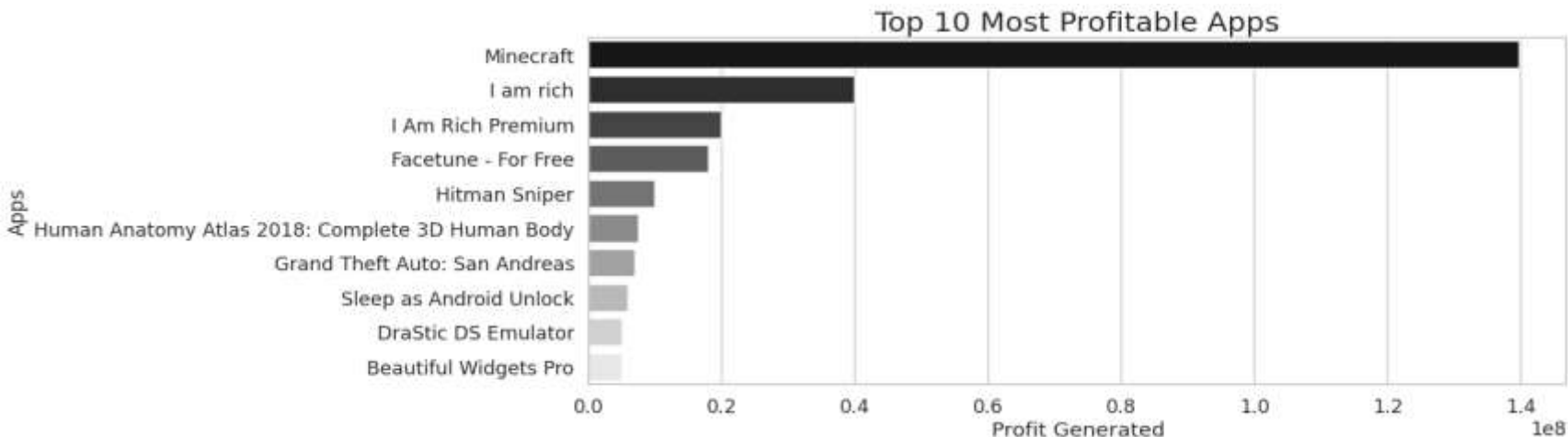
Most Profitable Apps on the Play Store

Minecraft is the **Most Profitable App** on the play store with a total of around One-Hundred Forty Million (**≈140**) Revenue generated
followed by **I am rich** and others

CONCLUSION
➢ GAMING Category again dominates here, as, including Minecraft the, majority of the Most profitable Apps are from GAME Category
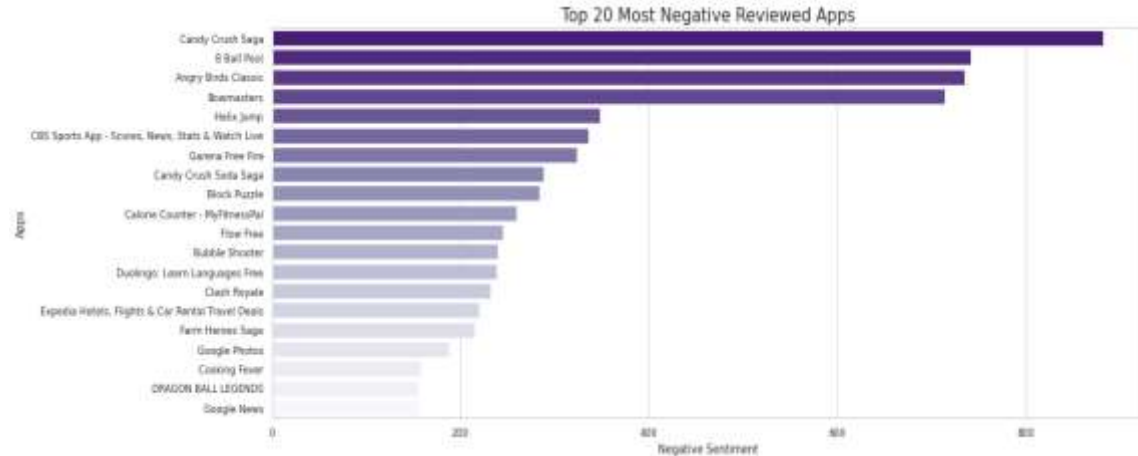
Apps with the Most Positive Reviews

Duolingo: Learn Languages Free is the **Most Positive Reviewed** app in the play store followed by **Helix Jump**, **Bowmasters** and others

Top 20 Most Positive Reviewed Apps
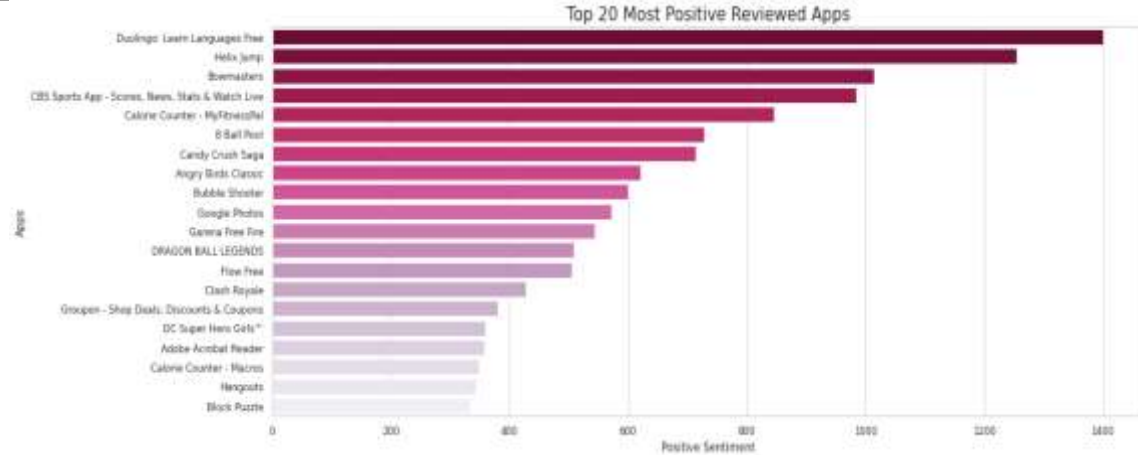
Apps with the Most Negative Reviews

Candy Crush Saga is the most negatively reviewed app on play store followed by **8 Ball Pool**, **Angry Birds Classic** and others...

CONCLUSION

➤ Negative Reviews here never mean that these Apps are not Profitable as Reviews are only given after Installs in most cases and these apps hold those numbers. People are still using these but these apps need to be fixed following those reviews ASAP

Top 20 Most Negative Reviewed Apps

Multivariate Analysis:

Also, there are a total of 20 Apps with over a Billion Installs..

CONCLUSION
- ➤ GOOGLE and FACEBOOK are ruling the world as out of these 20 almost 75% of the Apps are solely from these two Companies

	Category	App	Installs
7988	SOCIAL	Google+	1000000000
9426	TRAVEL_AND_LOCAL	Maps - Navigate & Explore	1000000000
9580	VIDEO_PLAYERS	Google Play Movies & TV	1000000000
4852	GAME	Subway Surfers	1000000000
7507	PRODUCTIVITY	Google Drive	1000000000
9395	TRAVEL_AND_LOCAL	Google Street View	1000000000
2698	FAMILY	Google Play Games	1000000000
1182	COMMUNICATION	WhatsApp Messenger	1000000000
1100	COMMUNICATION	Messenger – Text and Video Chat for Free	1000000000
1056	COMMUNICATION	Gmail	1000000000
9651	VIDEO_PLAYERS	YouTube	1000000000
8920	TOOLS	Google	1000000000
1146	COMMUNICATION	Skype - free IM & video calls	1000000000
7999	SOCIAL	Instagram	1000000000
341	BOOKS_AND_REFERENCE	Google Play Books	1000000000
1065	COMMUNICATION	Hangouts	1000000000
1059	COMMUNICATION	Google Chrome: Fast & Secure	1000000000
7963	SOCIAL	Facebook	1000000000
1535	ENTERTAINMENT	Google Play Games	1000000000
6468	NEWS_AND_MAGAZINES	Google News	1000000000

Challenges Faced:

➢ Reading and comprehending the problem statement.

➢ Converting columns to suitable formats

➢ Dealing with error, duplicate and NaN values in the dataset.

➢ To come up with the best suited visualizations to summarize the information in the dataset

☐ **Overall Conclusion:**

AI

Here is what we have concluded after analyzing the data set:

➤ A **Positive Correlation** was seen between **Reviews** and **Installs**…
   Also a **Negative Correlation** was seen between **Ratings**, **Reviews**, & **Installs** with the Price.

➤ **Most of the Apps** over Play store are in **FAMILY** Category followed by **GAMES**, **TOOLS** and others……

➤ **Most Installations** are from **GAME Category** then **COMMUNICATION** then **PRODUCTIVITY** and others.

➤ **TOOLS**, **ENTERTAINMENT**, and **EDUCATION** are some of the **TOP Genres** in Play store.

➤ **80.4%** of the apps are for **Everyone** i.e they do not have any age limitations

➤ The **Average Size** of Apps on Play store is around **≈20MB**

➤ Most of the Apps in play store hold a Rating of **4.3** to **4.6/7** out of 5.

➤ The **Most Compatible Android Version** is **4.1 and above**

Overall Conclusion:

- Around **93%** of the Apps on Play store are **Free** and only **7%** are **Paid**.

- From the **Year Updation plot,** we can see that most of the apps are updated in the year 2017-2018

- The **Most Successful App** on Play store is **Facebook** and then **WhatsApp Messenger** then **Instagram** and others.

- There are a total of **20 Apps** with over a **Billion** Installs registered from Play store.

- The Most Profitable App is **Minecraft** holding the throne with total revenue of **140 Million.**

- The Most Loved App is **Duolingo: Learn Languages Free** followed by **Helix Jump** then **Bowmasters** and others.

- **Candy Crush Saga** holds the section with Most Negative Reviews followed by **8 Ball Pool** and **Angry Birds Classic** and others.

AI

☐ **THANKS !**