



# **Capstone Project - 1**

## **PLAY STORE APP REVIEW ANALYSIS**

### ***Team Members***

**Ankit Patil**  
**Mihir Kulkarni**  
**Ram Bakale**

# Let's Review Play Store:

1. Overview of Google Play Store
2. Available Data sets
3. Data Summary
4. Exploratory Data Analysis
  - a) Apps per Category
  - b) Average Rating across categories
  - c) Average Price according to Categories
  - d) Average Price according to Categories
  - e) Number of Installs per Category Type wise
  - f) Free Vs Paid Apps Analysis
  - g) Apps per Age Group
  - h) 5 Star Rated Apps per Category
  - i) Top 10 most Expensive Apps
  - j) Top 10 highest rated Apps
  - k) Top 10 Revenue making Apps
  - l) Distribution of Sentiments as per Installs
5. Summary
6. Conclusion



## Overview of Google Play Store

- The Google Play Store is an App that provides access to the collection of Apps that are available on Servers at google for Android devices.
- These apps further may be classified as per the category it falls under.
- Google keeps record for each User about apps installed, ratings given, reviews posted etc.
- This data is further used and analyzed by Google to enhance user experience by improving App engagement



## Available Data sets

1. **Play Store Data.csv:** In this file the 10840+ rows containing apps from free and paid category are described against 13 columns as characteristics comparable amongst them.
2. **User Review.csv:** It contains 3 columns as characteristics :
  - ❖ Sentiment (Positive, Negative or Neutral),
  - ❖ Sentiment Polarity
  - ❖ Sentiment Subjectivity.

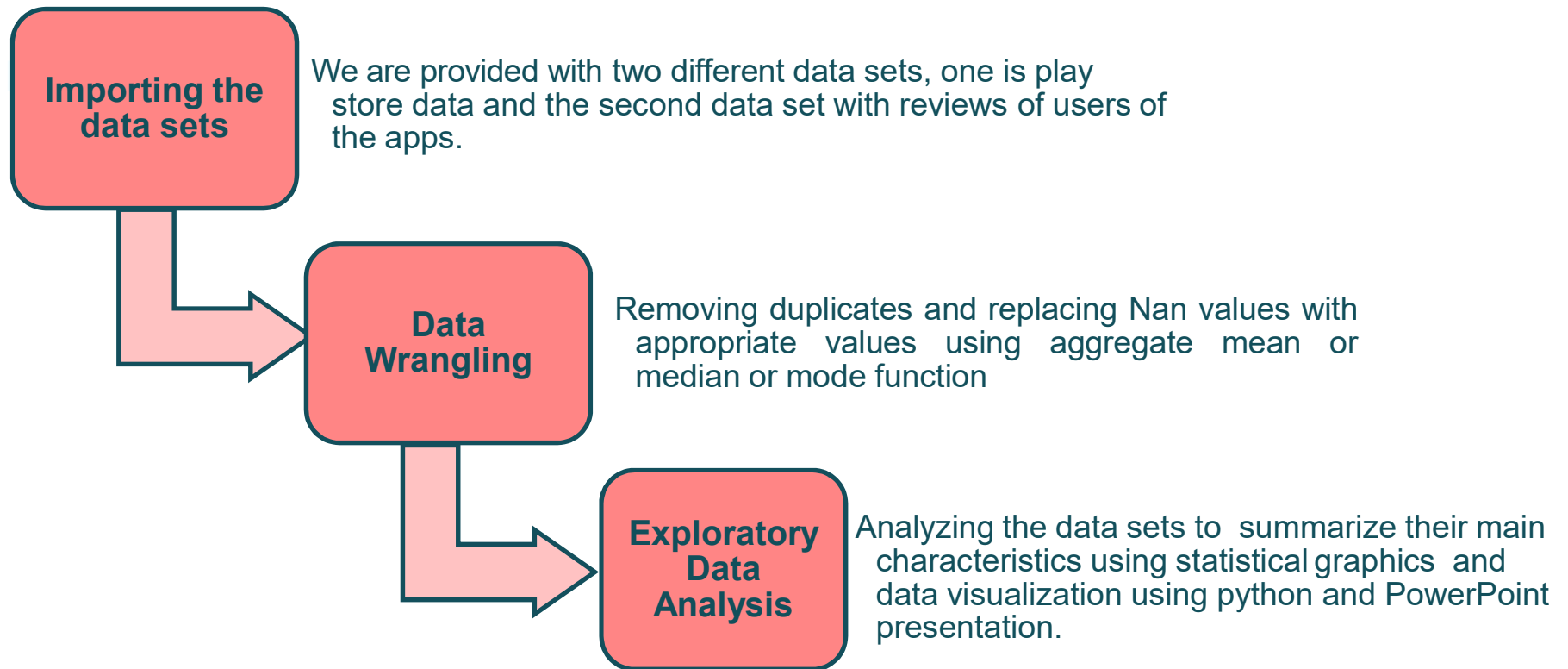
## Data Summary: Play Store

| Variable        | Description                          |
|-----------------|--------------------------------------|
| App             | Name of Application                  |
| Category        | Category to which the app belongs to |
| Rating          | Overall user rating                  |
| Reviews         | Number of user reviews for the app   |
| Size            | Size of the app                      |
| Installs        | Number of user downloads for the app |
| Type            | Paid or Free                         |
| Price           | Price of the app                     |
| Genres          | An App can belong to multiple genres |
| Last Updated    | Last updated date                    |
| Current Version | Current version of the App           |
| Android Version | Minimum Android Version required     |

## Data Summary: User Reviews

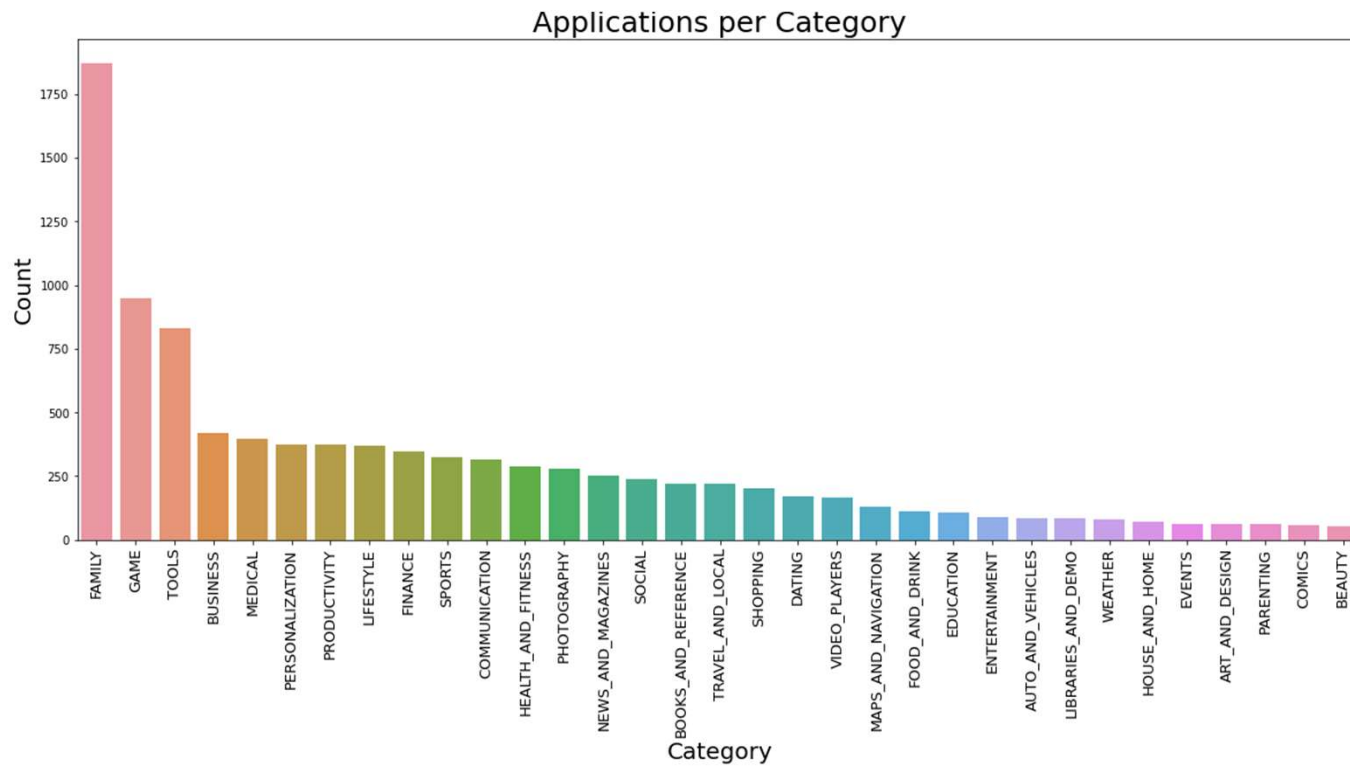
| Variable               | Description                  |
|------------------------|------------------------------|
| App                    | Application name             |
| Translated Review      | User review                  |
| Sentiment              | Positive/Negative/Neutral    |
| Sentiment Polarity     | Sentiment polarity score     |
| Sentiment Subjectivity | Sentiment subjectivity score |

## Process flow



# EXPLORATORY DATA ANALYSIS(EDA)

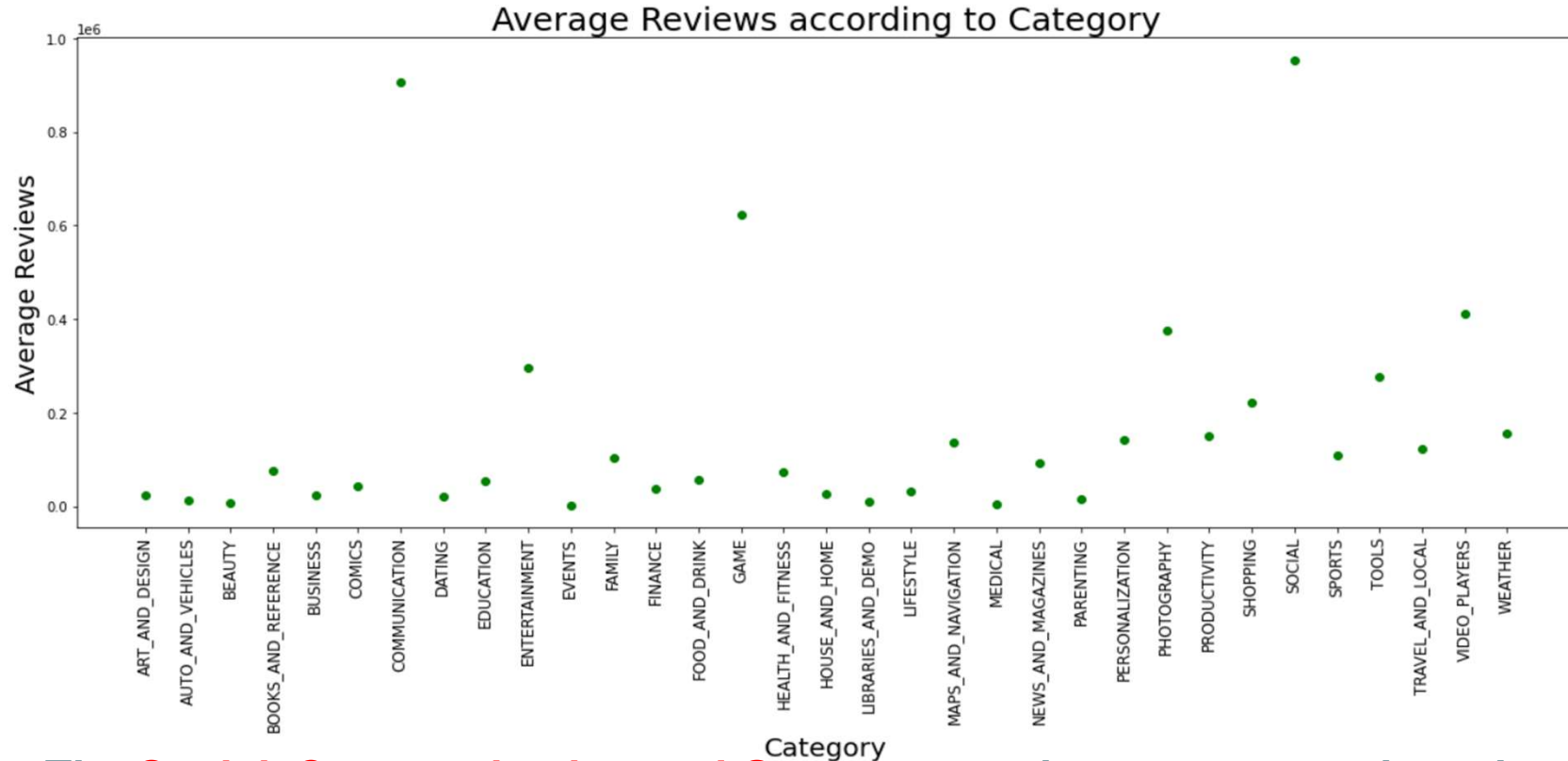
## Count of applications for each category



Family has most number of apps present on Play Store, followed by Game and then Tools  
Beauty has least number of apps present on Play Store



## EDA(continued)



The **Social, Communication and Game** categories are most reviewed on the play store.

## EDA(continued)

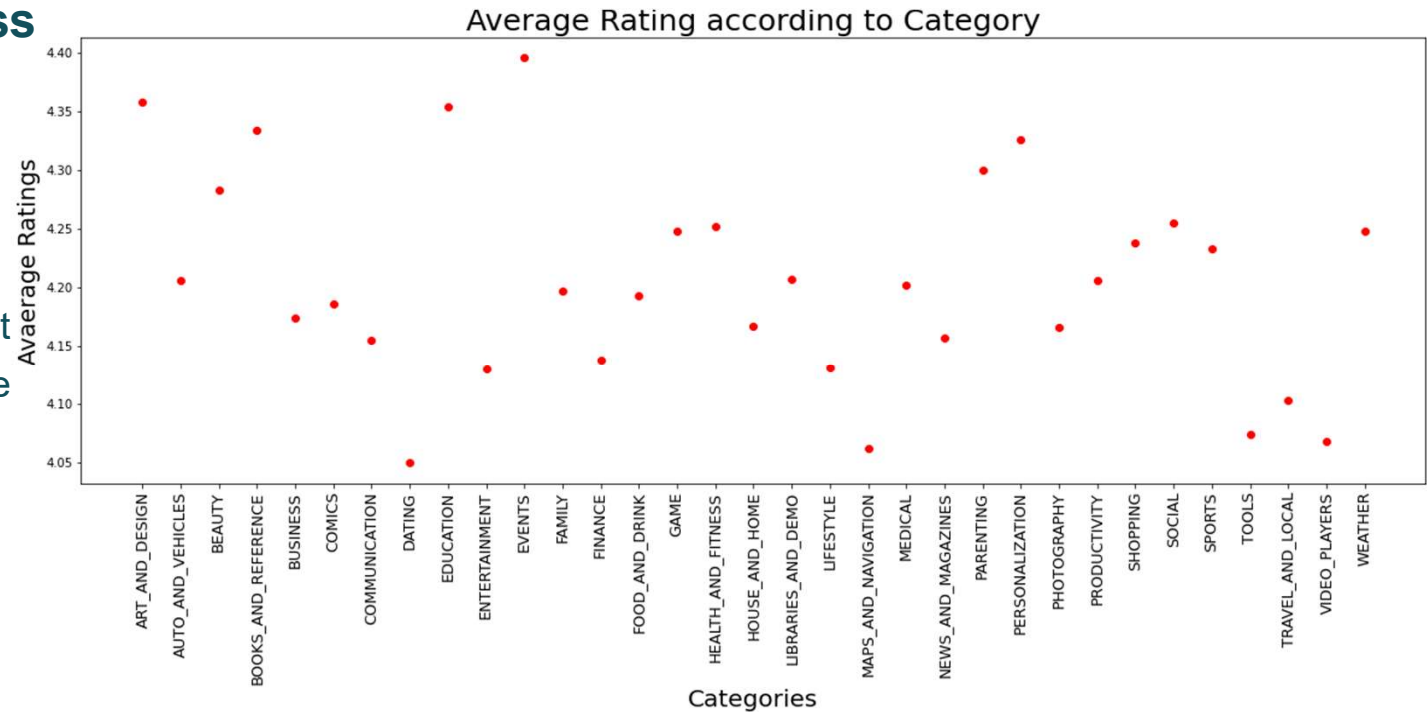
### Average Rating Across Categories:

#### Highest Rated apps:

Events, Art and Design, Education Categories has got maximum i.e. 4.35 and above ratings.

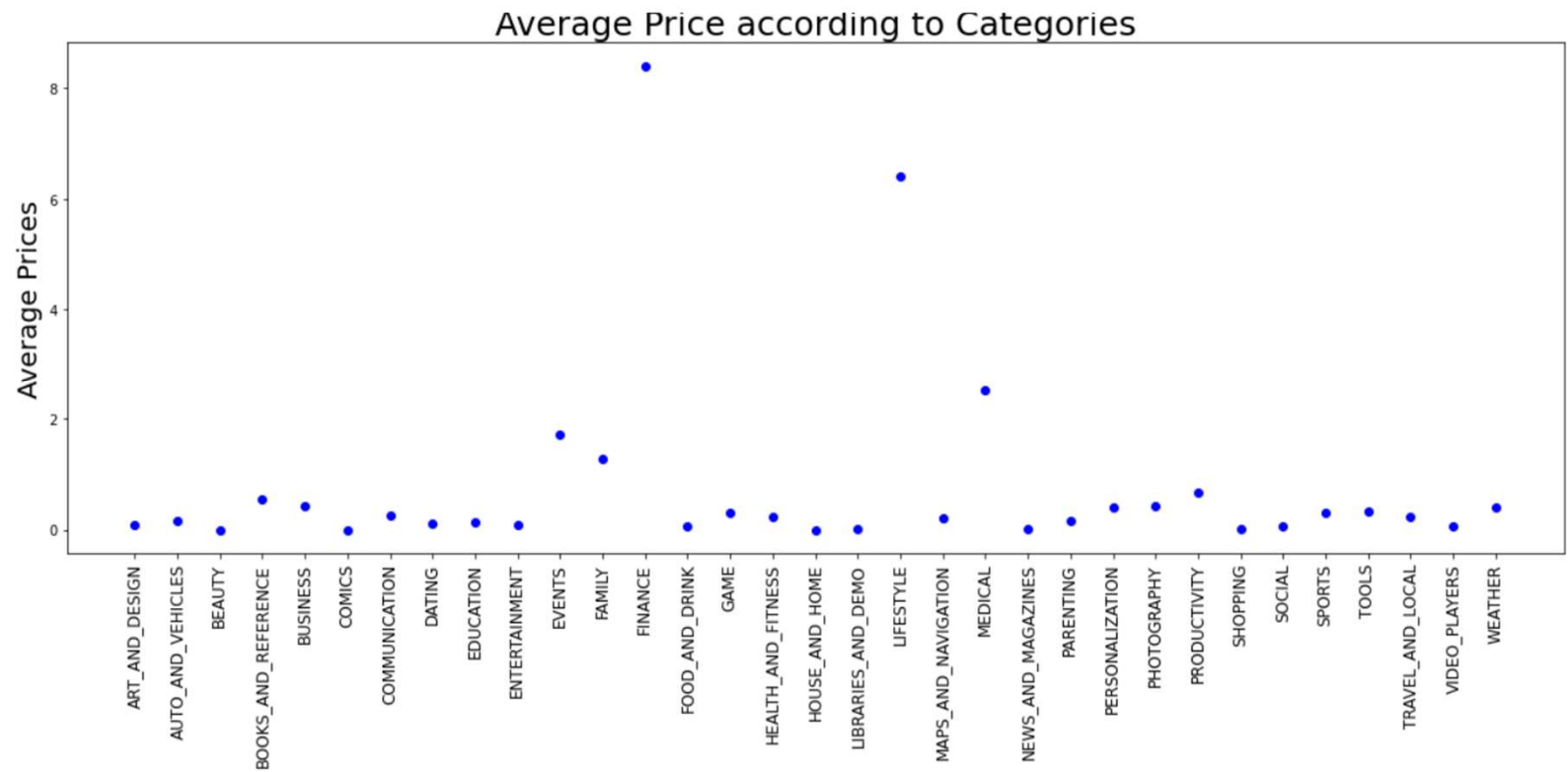
#### Lowest Rated apps:

Dating, Maps and navigation, Video players category has got least ratings as per the available data.

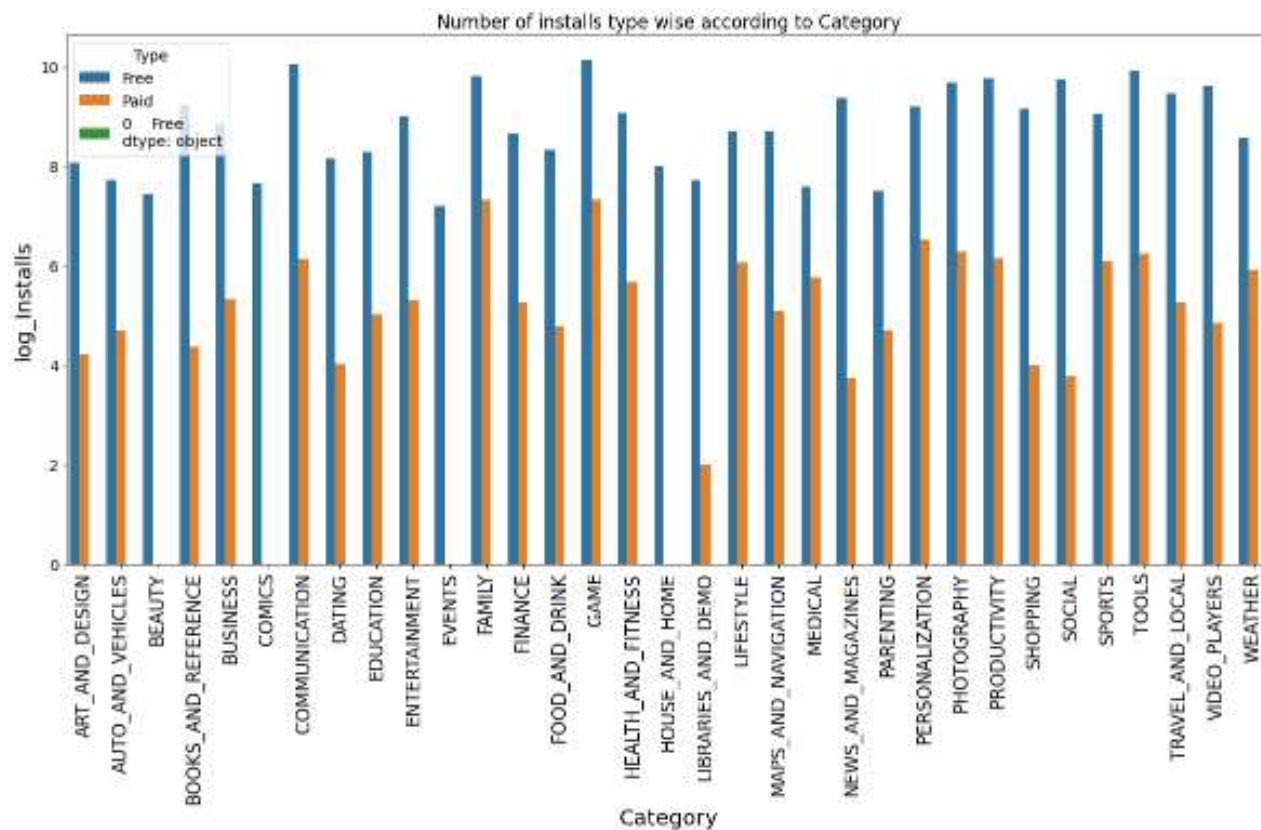


## EDA(continued)

Finance has highest average price followed by Lifestyle and then Medical Category for apps on Play Store.



We have converted numeric values to logarithmic values to compensate for skewness in number of installs for different categories



In our dataset, the majority of apps in **Family, Game** and **Communication** categories are free to install.

At the same time **Family, Game, and Personalization** categories had the biggest number of paid apps available on play store.

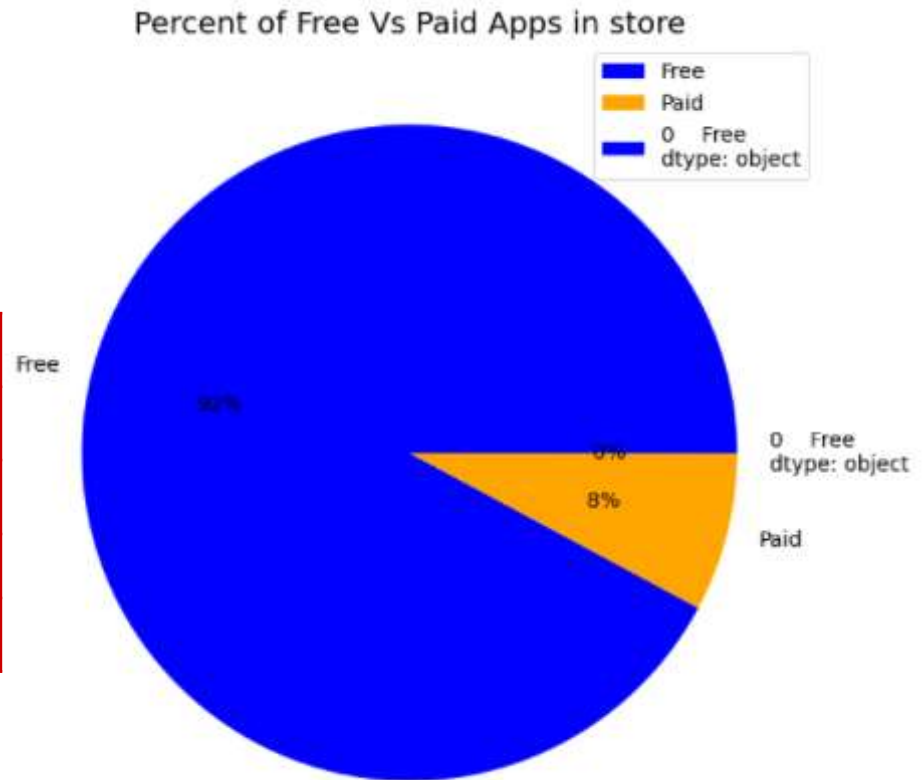
**Comics, Events** and **House And Home** Category has no paid apps

## EDA(continued)

### Free Vs Paid Apps

Here we can see that there are 92% of apps are free and only 8% of Apps are paid on Play store.

|                | Free Apps   | Paid Apps |
|----------------|-------------|-----------|
| Number         | 8903        | 755       |
| Average Rating | 4.186027    | 4.269934  |
| Installs       | 75216062646 | 57363881  |
| Reviews        | 2087014255  | 6595973   |

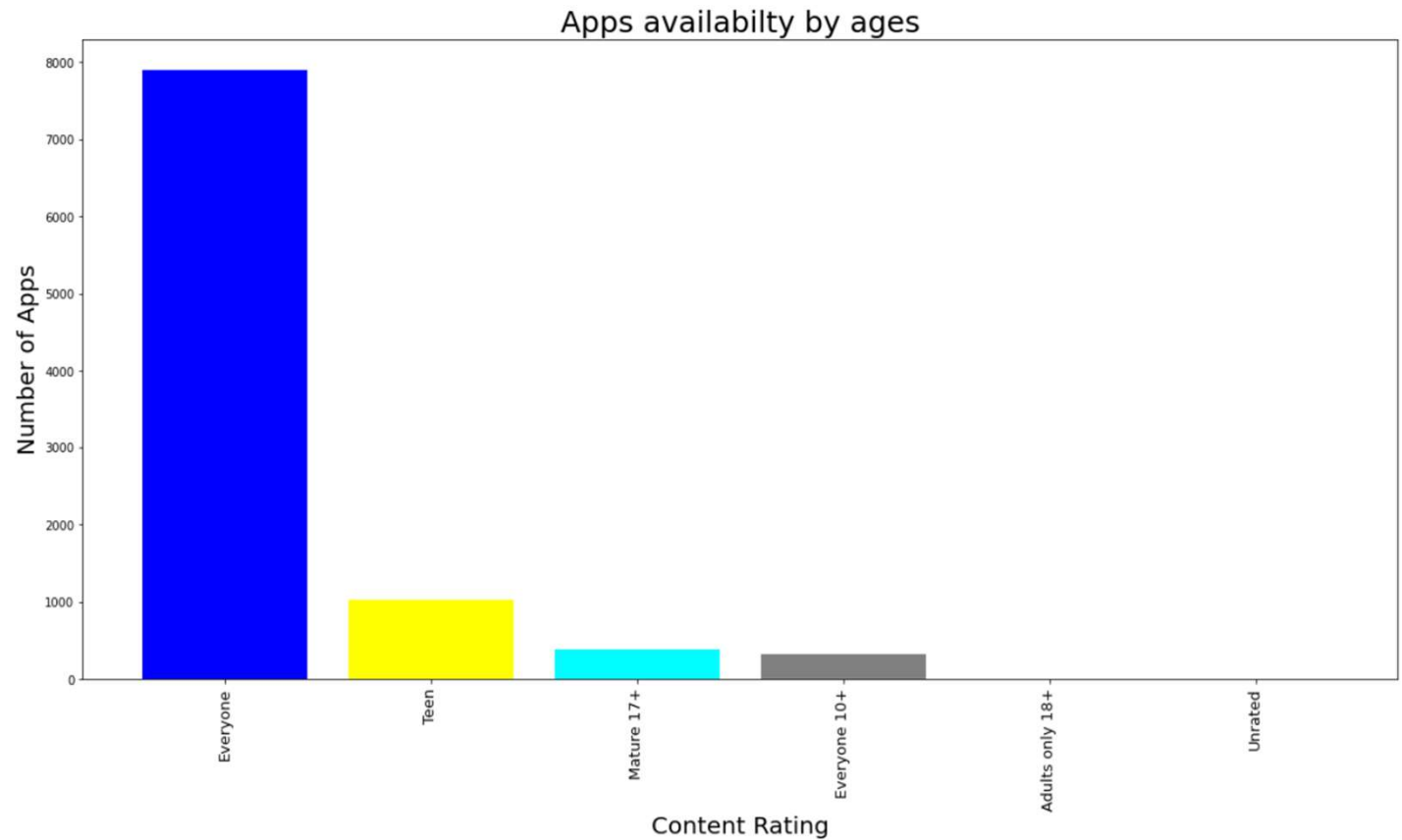


Number of Free Apps present on Play Store is in multiple folds of Paid Apps

## EDA(continued)

### Apps classified by different age group

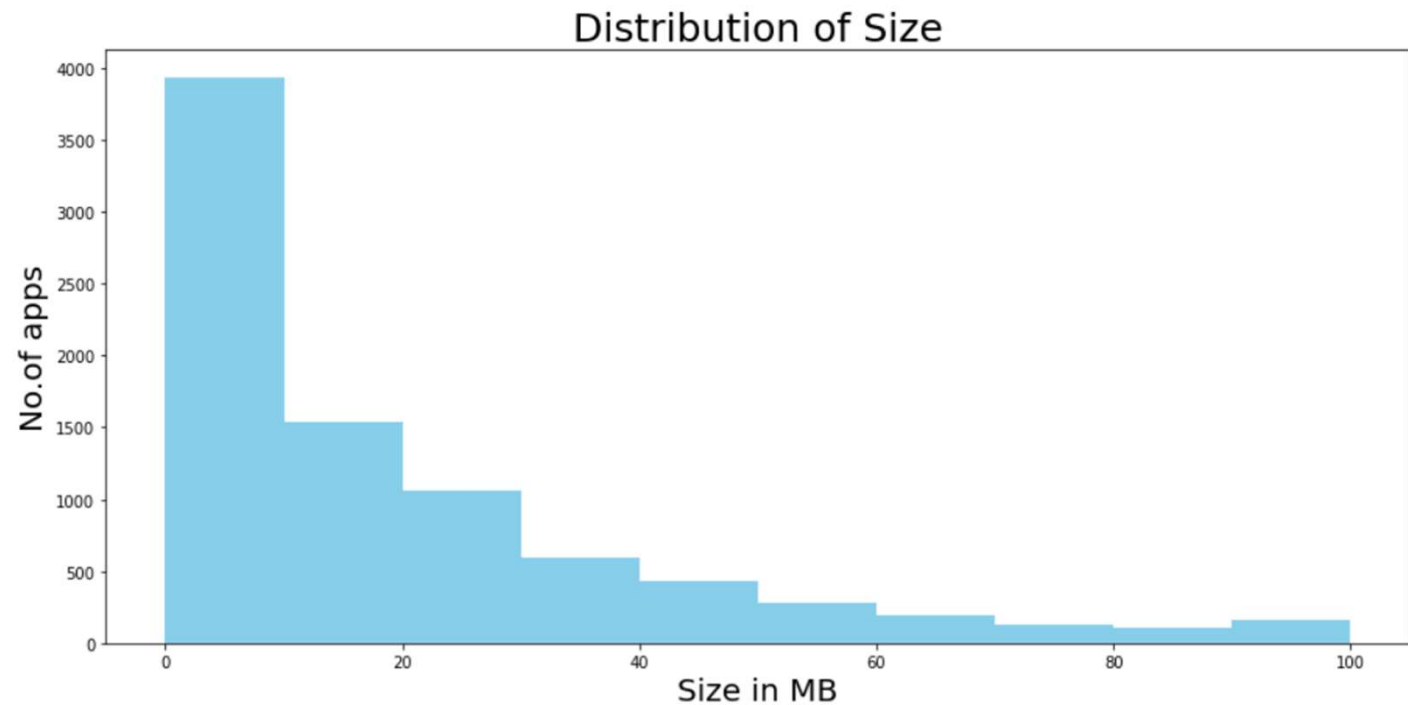
Apps rated for Everyone are the highest number of Apps followed by Apps rated for Teens present on Play Store



## EDA(continued)

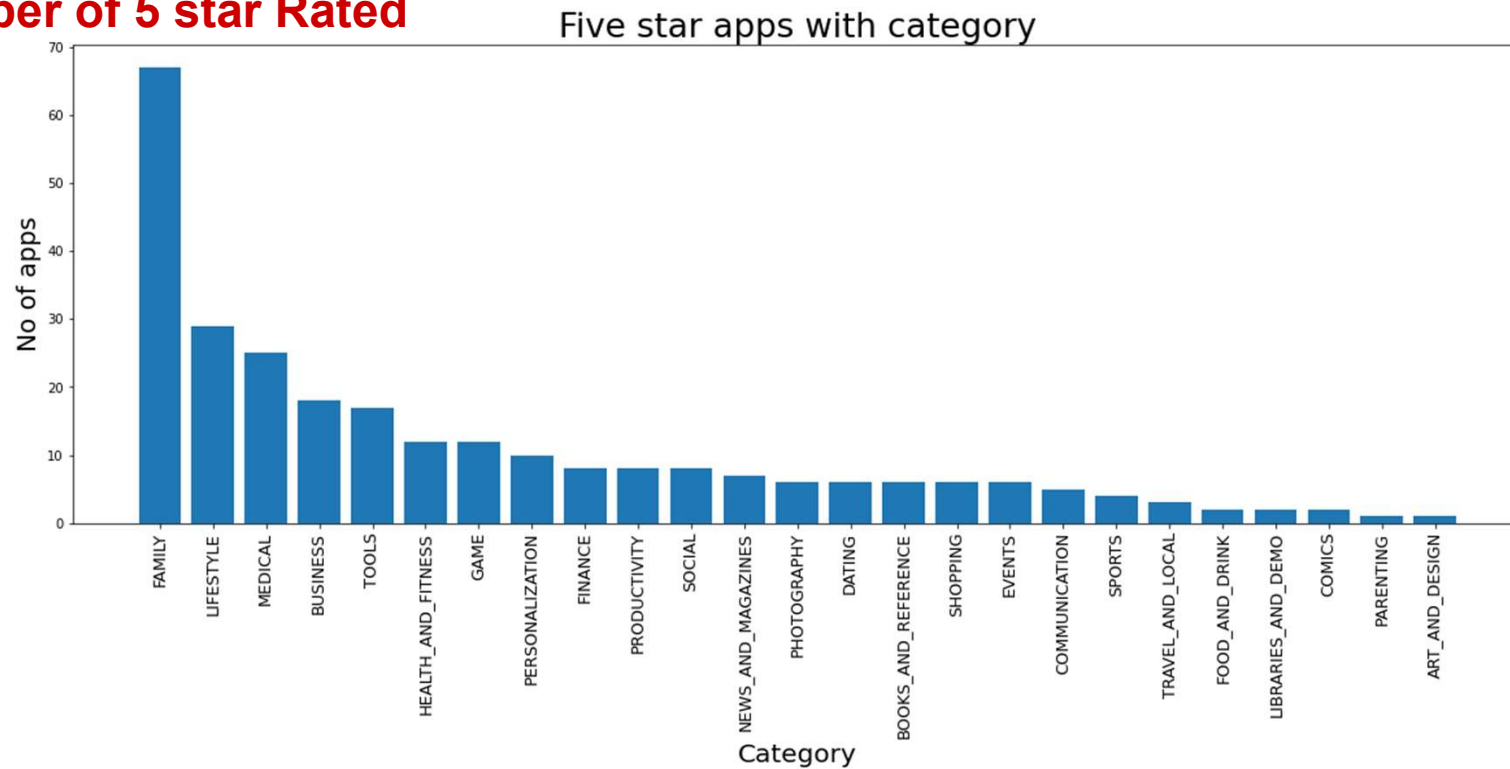
### Distribution of App sizes

It can be seen from the distribution size plot that maximum applications' size is between 0–10 MB.



## EDA(continued)

Number of 5 star Rated apps



Family Category has most 5 star rated Apps followed by Lifestyle, whereas Art and Design has least 5 star rated Apps



## EDA(continued)

### Top 10 Expensive Apps

| Application                         | Category  | Rating | Reviews | Size | Installs | Type | Price  | Content Rating | Genres        |
|-------------------------------------|-----------|--------|---------|------|----------|------|--------|----------------|---------------|
| I'm Rich - Trump Edition            | LIFESTYLE | 3.6    | 275.0   | 7.3  | 10000.0  | Paid | 400.00 | Everyone       | Lifestyle     |
| I Am Rich Pro                       | FAMILY    | 4.4    | 201.0   | 2.7  | 5000.0   | Paid | 399.99 | Everyone       | Entertainment |
| I am Rich                           | FINANCE   | 4.3    | 180.0   | 3.8  | 5000.0   | Paid | 399.99 | Everyone       | Finance       |
| I'm Rich/Eu sou Rico//انا غني/我很 有錢 | LIFESTYLE | 4.3    | 0.0     | 40.0 | 0.0      | Paid | 399.99 | Everyone       | Lifestyle     |
| most expensive app (H)              | FAMILY    | 4.3    | 6.0     | 1.5  | 100.0    | Paid | 399.99 | Everyone       | Entertainment |
| I Am Rich Premium                   | FINANCE   | 4.1    | 1867.0  | 4.7  | 50000.0  | Paid | 399.99 | Everyone       | Finance       |
| I am rich (Most expensive app)      | FINANCE   | 4.1    | 129.0   | 2.7  | 1000.0   | Paid | 399.99 | Teen           | Finance       |
| I AM RICH PRO PLUS                  | FINANCE   | 4.0    | 36.0    | 41.0 | 1000.0   | Paid | 399.99 | Everyone       | Finance       |
| I am Rich Plus                      | FAMILY    | 4.0    | 856.0   | 8.7  | 10000.0  | Paid | 399.99 | Everyone       | Entertainment |
| I am Rich!                          | FINANCE   | 3.8    | 93.0    | 22.0 | 1000.0   | Paid | 399.99 | Everyone       | Finance       |

I am rich – Trump Edition is the most expensive app followed by I am Rich Pro on the google play store.

## EDA(continued)

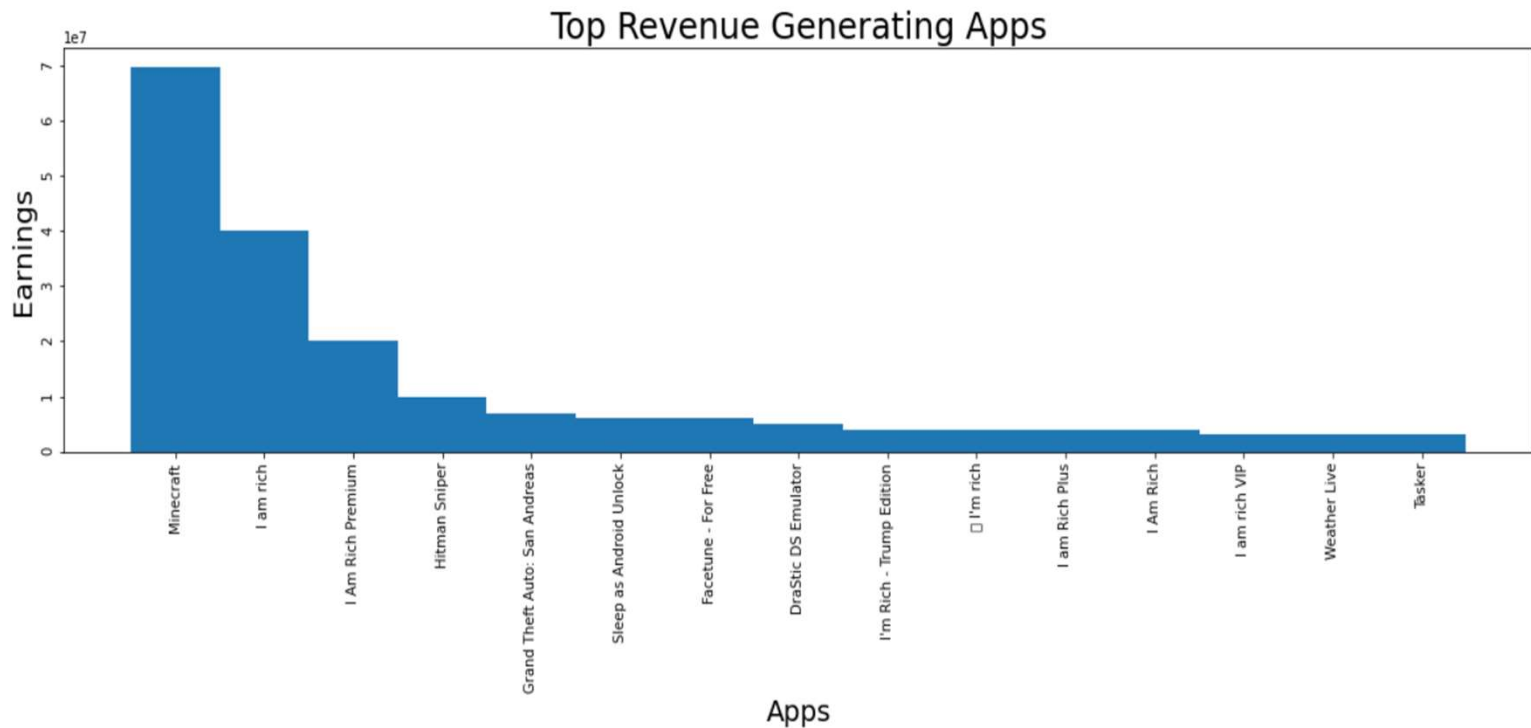
**Top 10  
Highest  
Rated Apps**

| Application                                      | Category           | Rating | Reviews | Size | Installs | Type | Price | Content Rating | Genres           |
|--|--------------------|--------|---------|------|----------|------|-------|----------------|------------------|
| Ríos de Fe                                       | LIFESTYLE          | 5.0    | 141.0   | 15.0 | 1000.0   | Free | 0.0   | Everyone       | Lifestyle        |
| FD Calculator (EMI, SIP, RD & Loan Eligibility)  | FINANCE            | 5.0    | 104.0   | 2.3  | 1000.0   | Free | 0.0   | Everyone       | Finance          |
| Oración CX                                       | LIFESTYLE          | 5.0    | 103.0   | 3.8  | 5000.0   | Free | 0.0   | Everyone       | Lifestyle        |
| Barisal University App-BU Face                   | FAMILY             | 5.0    | 100.0   | 10.0 | 1000.0   | Free | 0.0   | Everyone       | Education        |
| Master E.K                                       | FAMILY             | 5.0    | 90.0    | NaN  | 1000.0   | Free | 0.0   | Everyone       | Education        |
| CL REPL  | TOOLS              | 5.0    | 47.0    | 17.0 | 1000.0   | Free | 0.0   | Everyone       | Tools            |
| AJ Cam   | PHOTOGRAPHY        | 5.0    | 44.0    | 2.8  | 100.0    | Free | 0.0   | Everyone       | Photography      |
| AI Today : Artificial Intelligence News & AI 101 | NEWS_AND_MAGAZINES | 5.0    | 43.0    | 2.3  | 100.0    | Free | 0.0   | Everyone       | News & Magazines |
| CS & IT Interview Questions                      | FAMILY             | 5.0    | 43.0    | 3.3  | 1000.0   | Free | 0.0   | Everyone       | Education        |
| Ek Vote  | PRODUCTIVITY       | 5.0    | 43.0    | 6.2  | 500.0    | Free | 0.0   | Everyone       | Productivity     |

Rios de Fe is the Highest Rated app followed by FD Calculator (EMI, SIP, RD & Loan Eligibility) the google play store by Reviews.

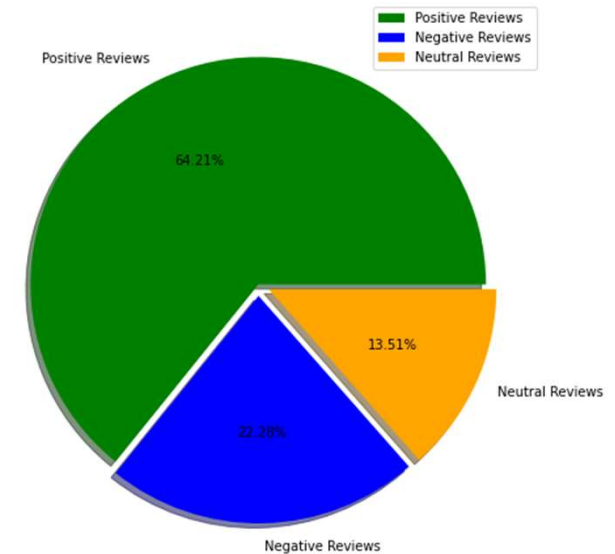
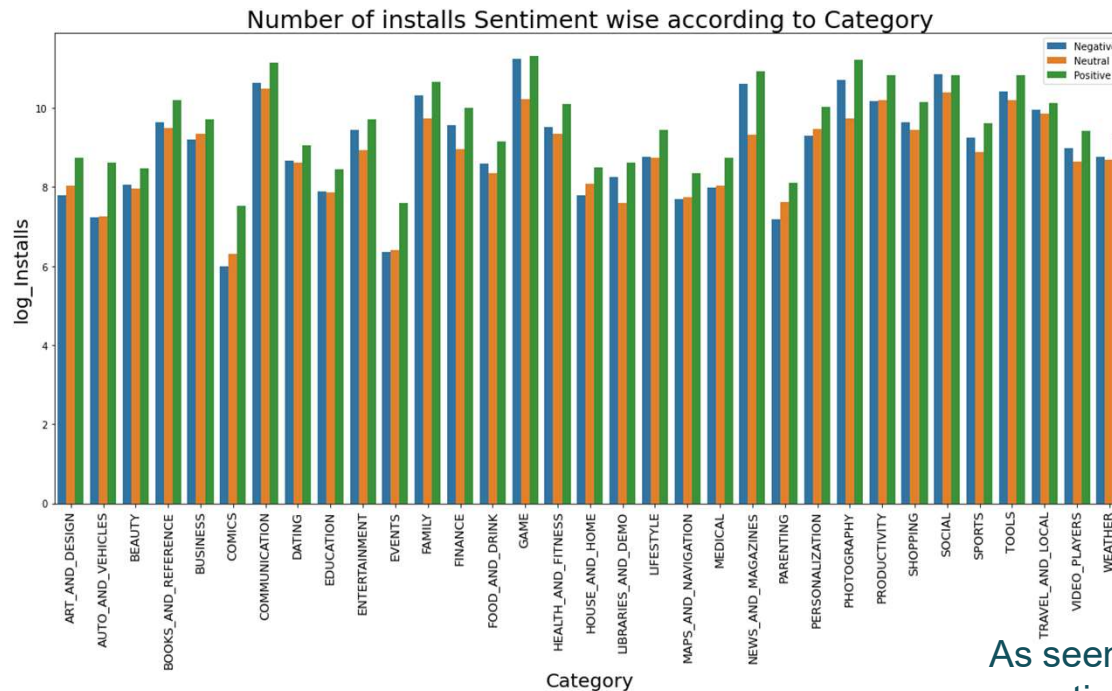
## EDA(continued)

**Top 10  
Revenue  
generating  
Apps  
Present on  
Play store  
based on  
their prices  
and  
installs.**



**Minecraft** has made the most revenue compared to other apps present on Play Store

## Distribution of Sentiments according to Installs



As seen from the pie chart, positive reviews are 64.08%, negative reviews are 22.85% and neutral reviews are 13.06% Majority of users are satisfied with the apps

\* We have converted Installs values to Log Installs values to compensate for wide range in installs across all categories

## SUMMARY

|   |   |
|---|---|
| Average Rating  | <b>Free: 4.19</b><br><b>Paid: 4.27</b>                                    |
| Number of Installs                                      | <b>Free: 75216062646</b><br><b>Paid: 57363881</b>                         |
| Total 5 Star Rated Apps                                 | <b>271</b>  |
| Category with most number of apps                       | <b>Family</b>   |
| Category with avg Highest rating                        | <b>Events</b>   |
| Category with highest avg Price                         | <b>Finance</b>  |
| Category with highest avg Reviews                       | <b>Social</b>   |
| Category with highest number of 5 star rated Apps       | <b>Family</b>   |
| Most Expensive App                                      | <b>I'm Rich - Trump Edition</b>   |
| Highest Rated app sorted with Highest number of Reviews | <b>Ríos de Fe</b>   |
| Highest Revenue making app                              | <b>Minecraft</b>  |
| Sentiments in Percentage                                | <b>Positive: 64.21</b><br><b>Neutral: 13.51</b><br><b>Negative: 22.58</b> |



## CONCLUSION:

Goal is to provide a better understanding of how App's ratings and reviews measure up to the industry as a whole.

1. User mostly prefers more of free apps. Most of the apps present in playstore are more or less of same size so size doesn't affect their decision much
2. The apps which are available for everyone are having the ratings 4 and above out of 5.
3. Most popular category of apps are on two basis - Number of Installs and number of reviews.
4. Family, Game, Communication, Social, Tools, Medical, Events Apps are the highest earning apps and at the same time Medical, Family, game, Tools apps are the expensive apps. This helps us to understand the application market and invest money in the right categories to make good profits.
5. Also, customer ratings plays a huge role while releasing the application.
6. The sentiment does not influence the final rating of the app because of the proportional behaviour.
7. Popular categories have a more positive sentiment.
8. Positive reviews are higher than negative and neutral but the sentiment subjectivity is not always proportional to sentiment polarity.



**THANK YOU**