

Capstone Project-1



Google Play

PlayStore Review Analysis by

Amar Singh Chouhan

Ashwin Babu

Hitesh Malvi

Ninad Rathore

Vikas Chandra



Agenda



Google Play

- 1 *PlayStore*
- 2 *Introduction to dataset*
- 3 *Visualizing Dataset*
- 4 *Content rating distribution*
- 5 *Content Rating v/s Rating*
- 6 *Revenue Generation*
- 7 *Analyzing Sentiments*
- 8 *Correlation of play store and user review*
- 9 *Relation between paid and free apps*
- 10 *Conclusion*

Introduction

- Google play was launched on 6 March 2012.
- Play Store and formerly Android Market, is a digital distribution service operated and developed by Google.
- It has enormous potential to drive app-making businesses to success.
- Google playstore is the largest and most used app on the Store. (85%)
- There are more than 3.04 millions apps on playstore.
- provides developers thousands of continually updated APIs that enable them to deliver high-quality experiences in their apps

PlayStore Attributes

- Apps
- Category
- Rating
- Review
- Size
- Install
- Type

- Price
- Content Rating
- Geners
- Last update
- Current Version
- Andriod Version

User Review Attributes

- App
- Translated Review
- Sentiment
- Sentiment Polarity
- Sentiment Subjectivity

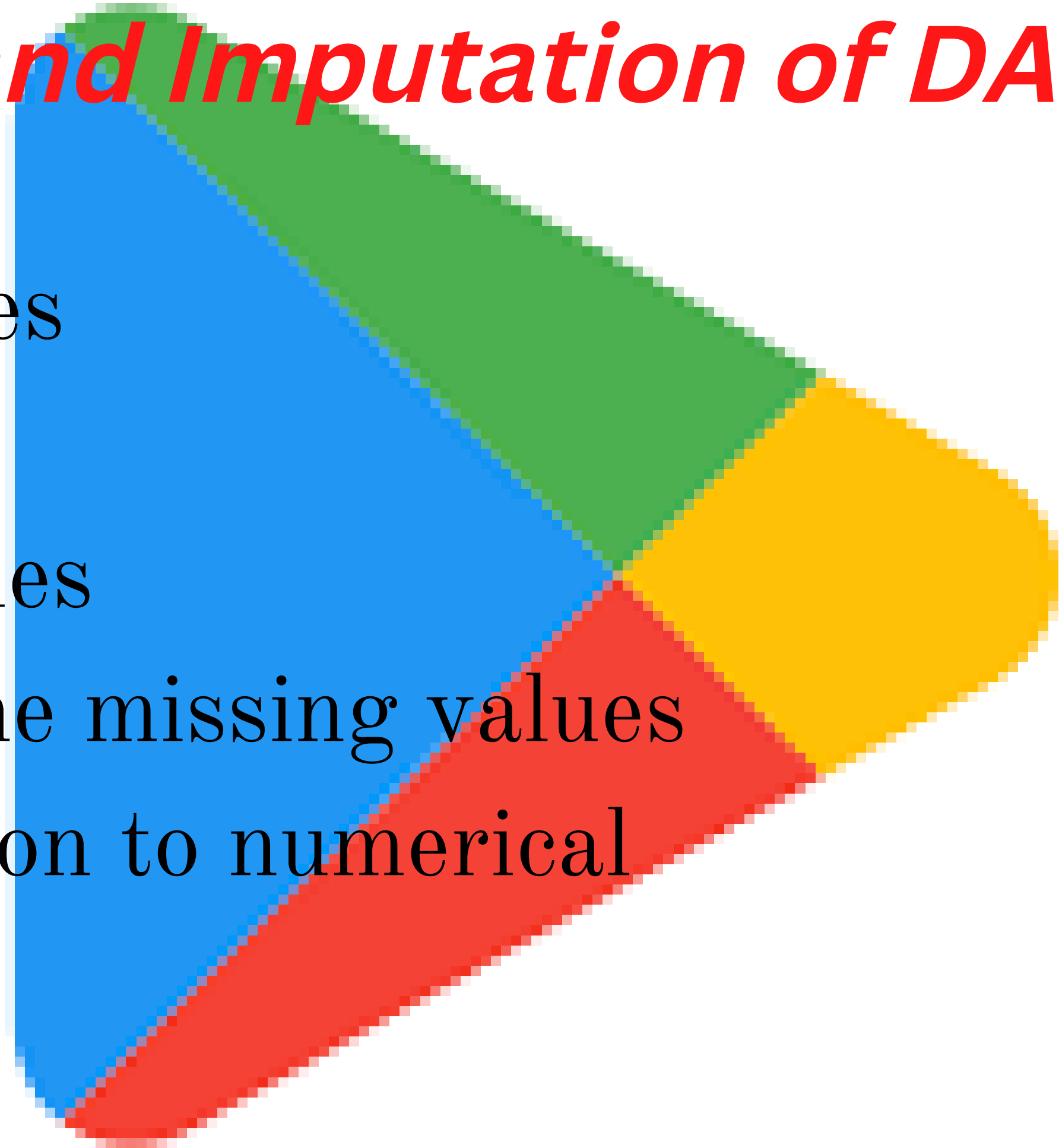


Visualizing Dataset

- Importing Dataset
- Importing libraries
- Cleaning data
- Data Imputation
- EDA

Cleaning and Imputation of DATA

- Duplicates
- Outliers
- Null values
- Filling the missing values
- Conversion to numerical

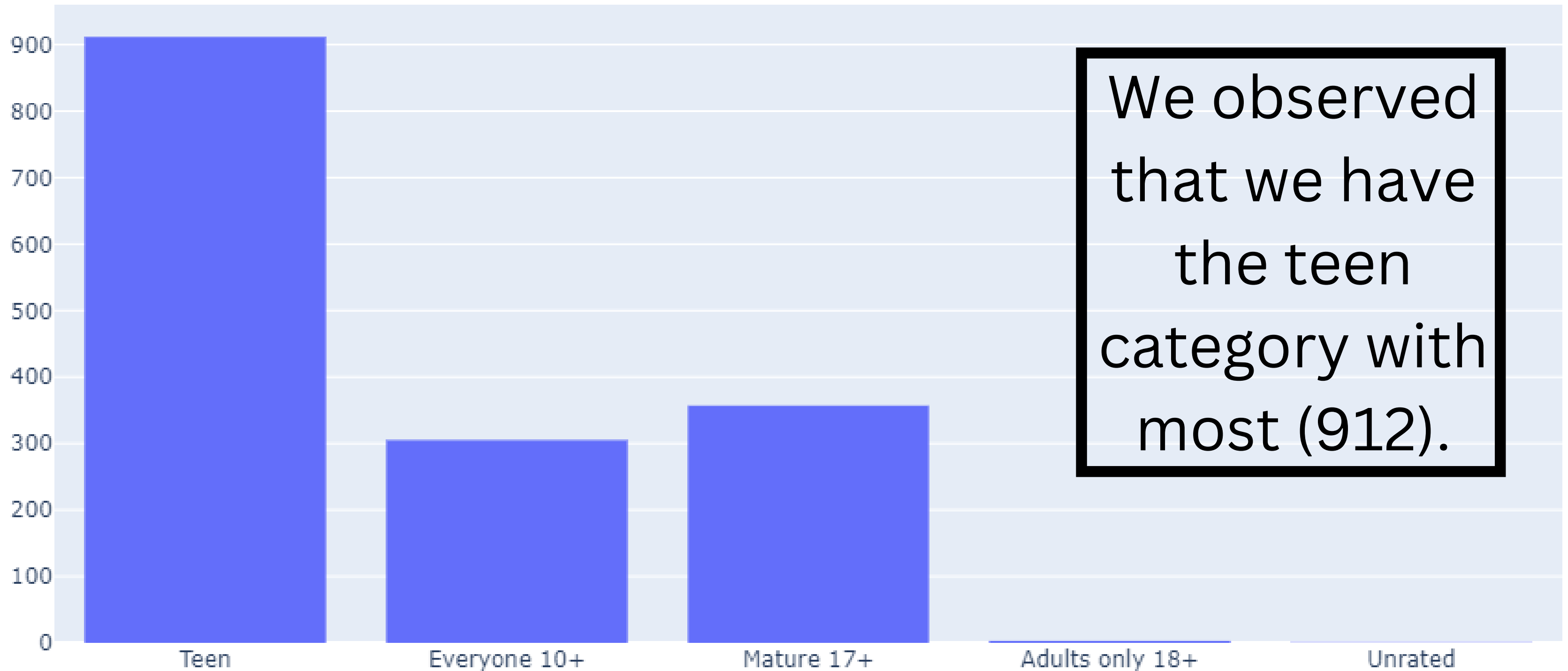


Exploratory Data Analysis(EDA)

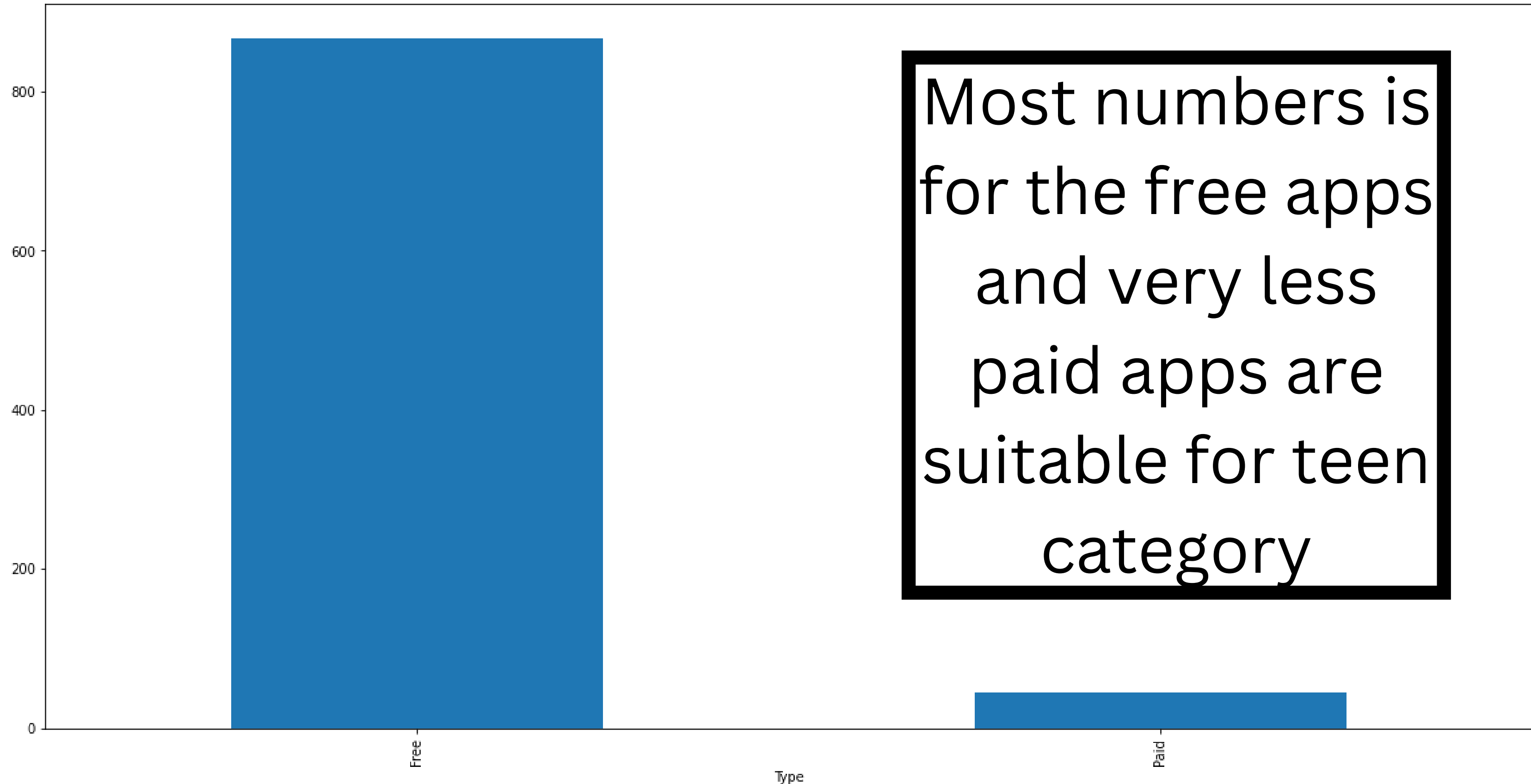


Google Play

Content rating

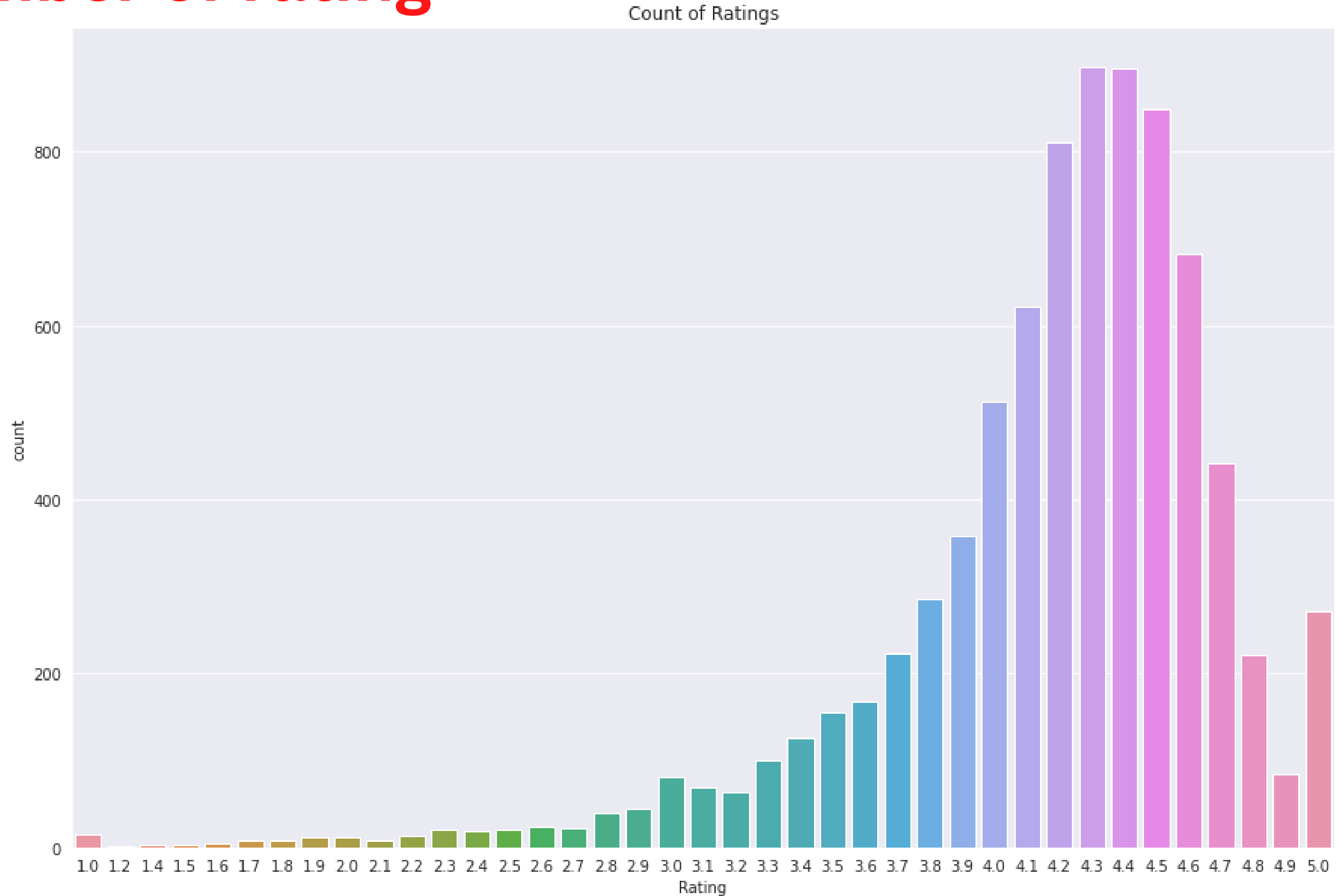


Top Content Rating in teen of paid and free type app



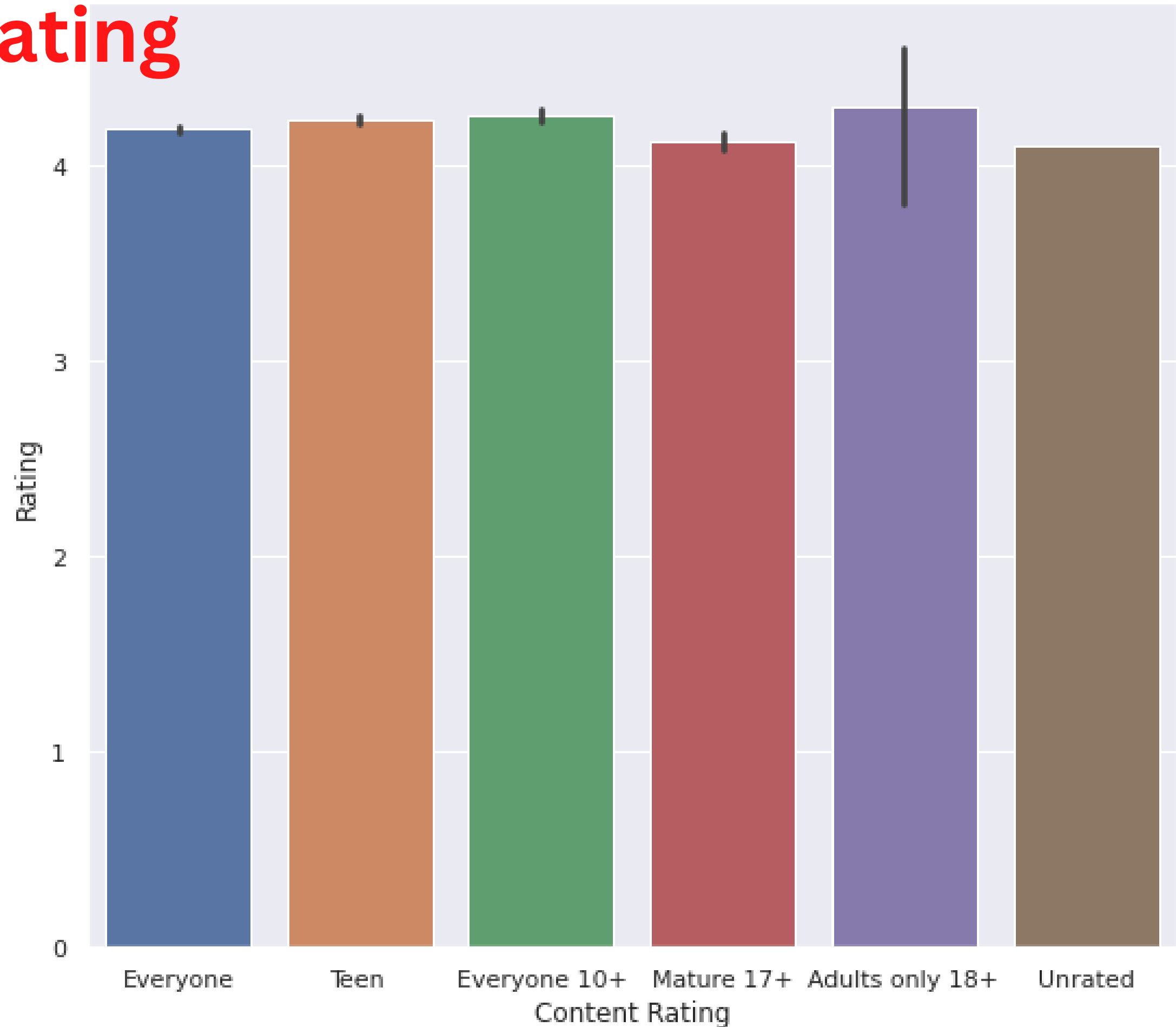
Highest number of rating

- Highest rating 4.3
- Lowest rating 1.2
- Max Rating 4.2 to 4.6

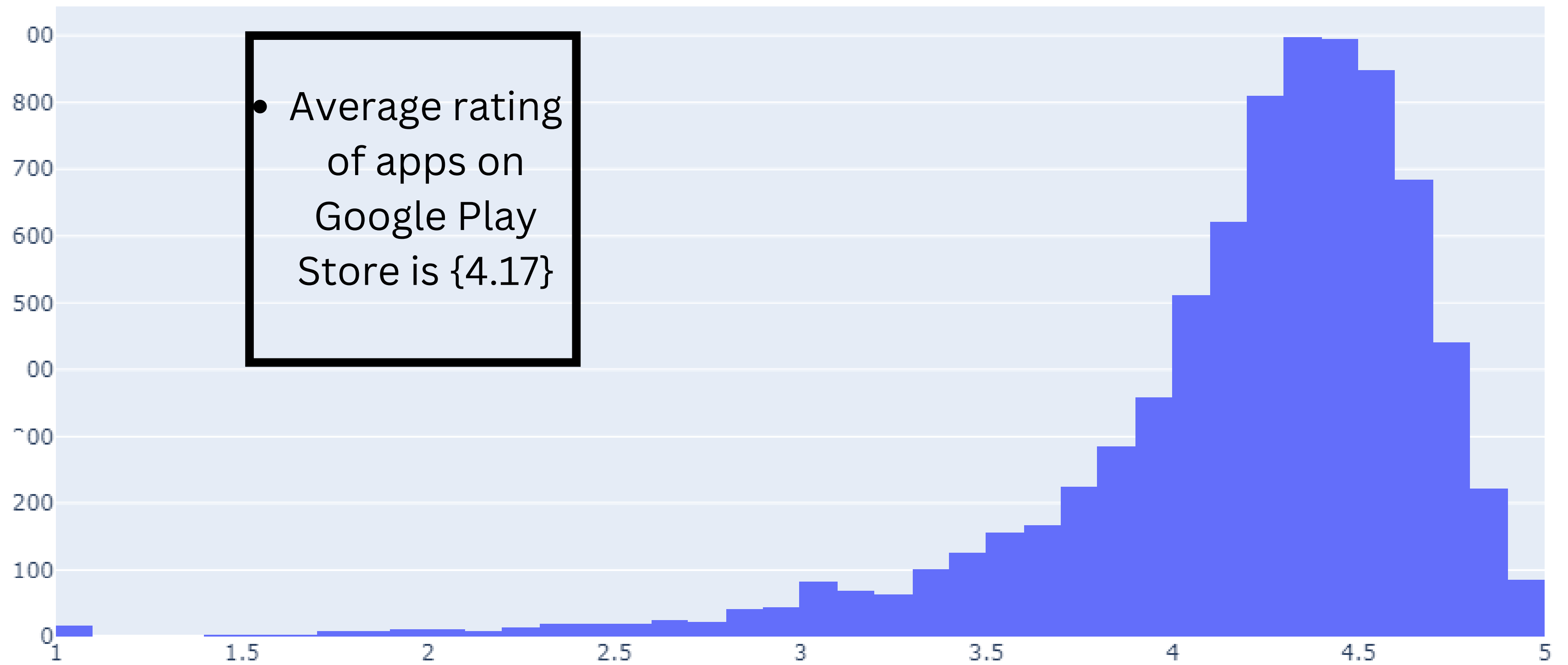


Content Rating v/s Rating

Only Adult 18+
is been rated the
most among all
the categories

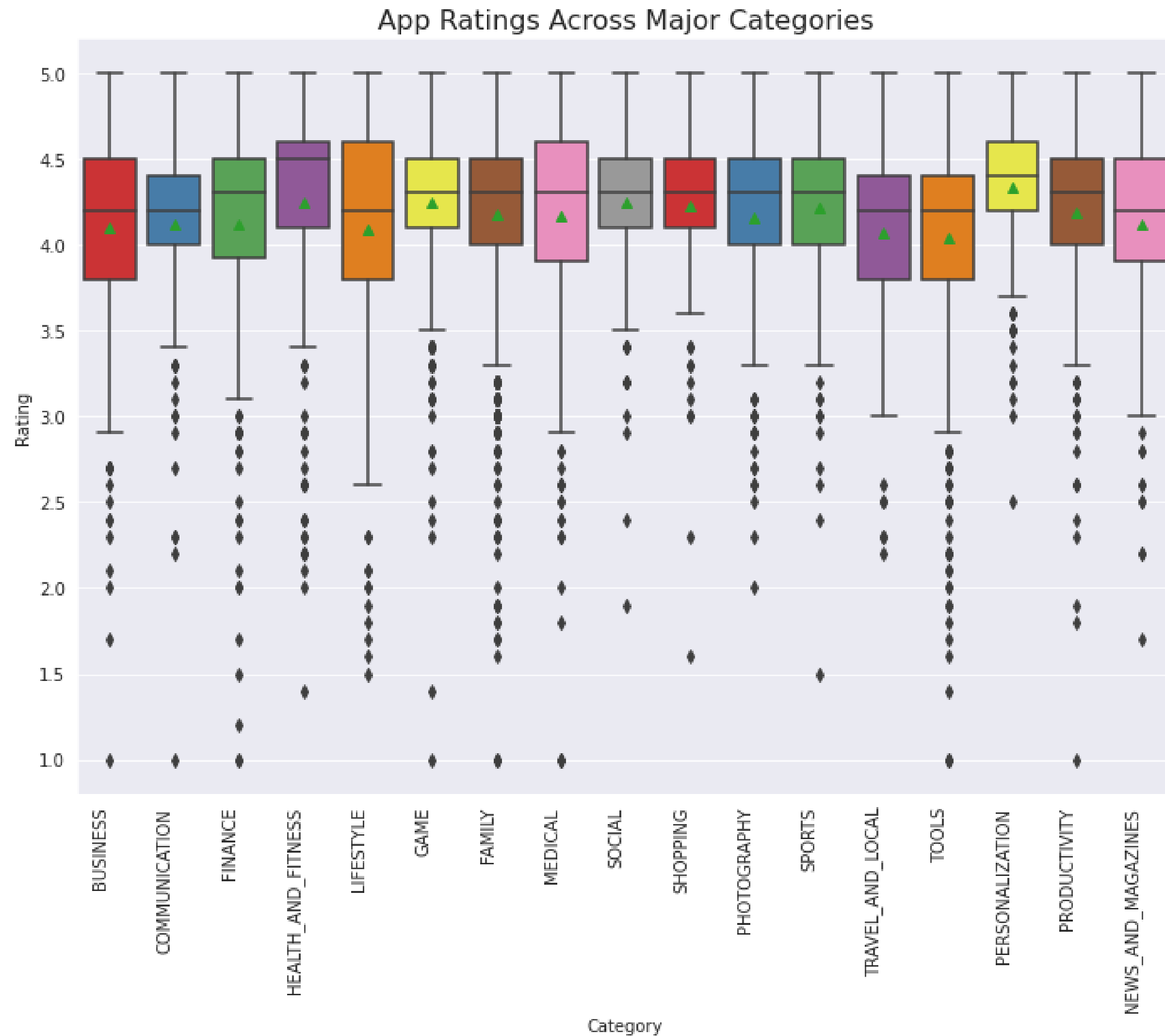


Average app rating



App rating with different category

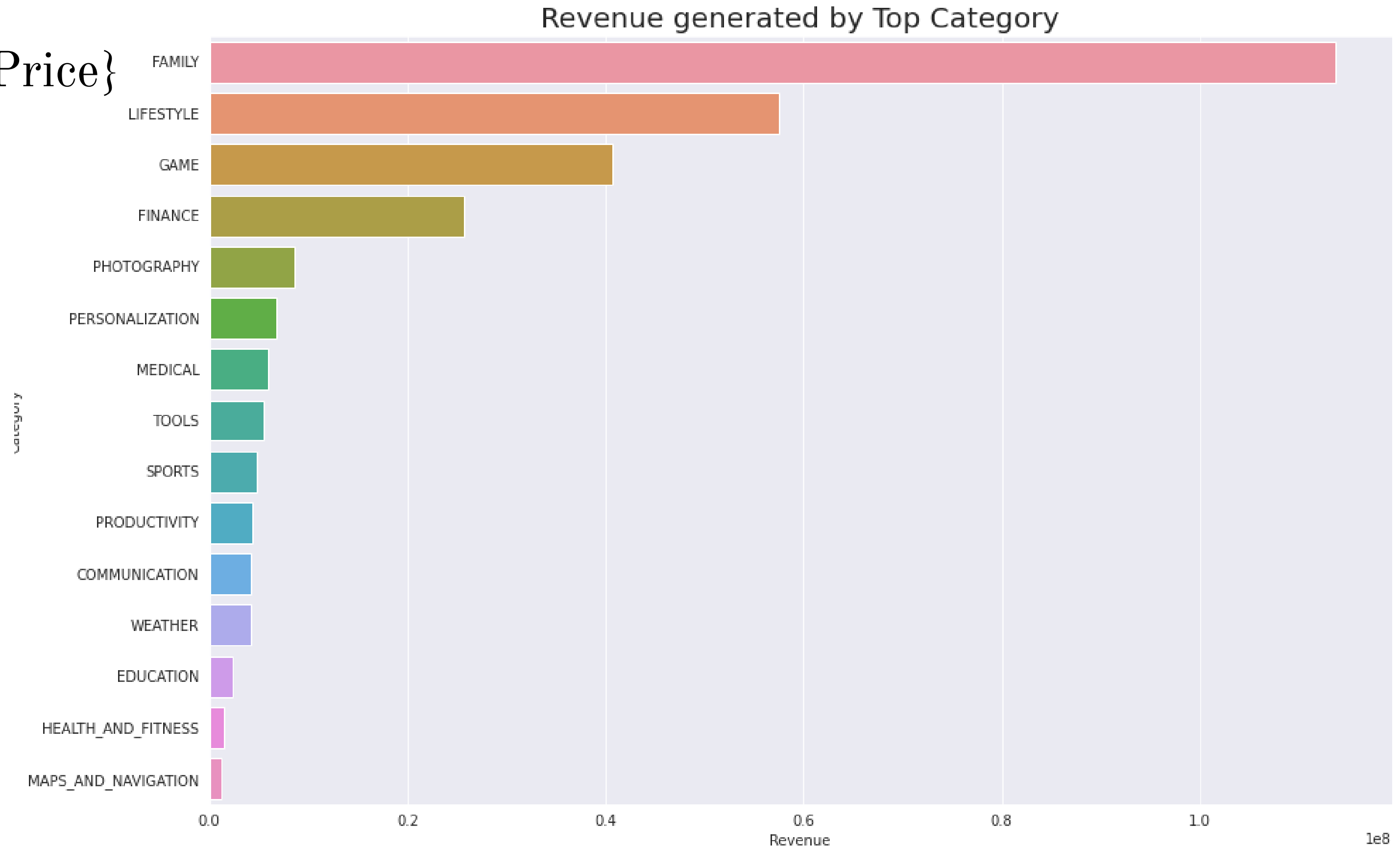
- The Health and Fitness and Books and Reference is having the most rating greater than 4.5
- Half of the Dating apps have a rating lower than the average.



Revenue Generation

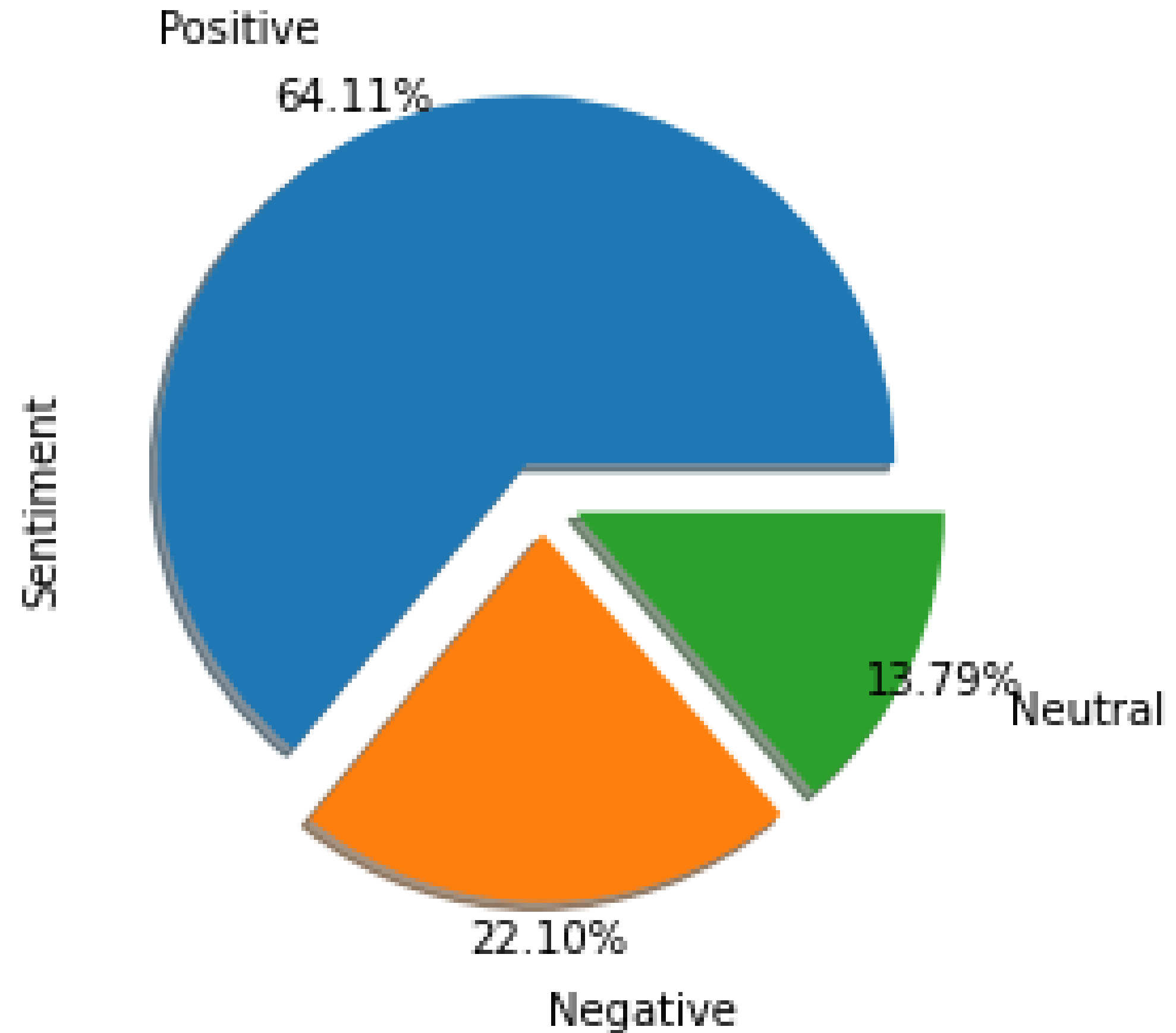
{Revenue = Install*Price}

- Family is the highest revenue generated category
- Maps are the lowest



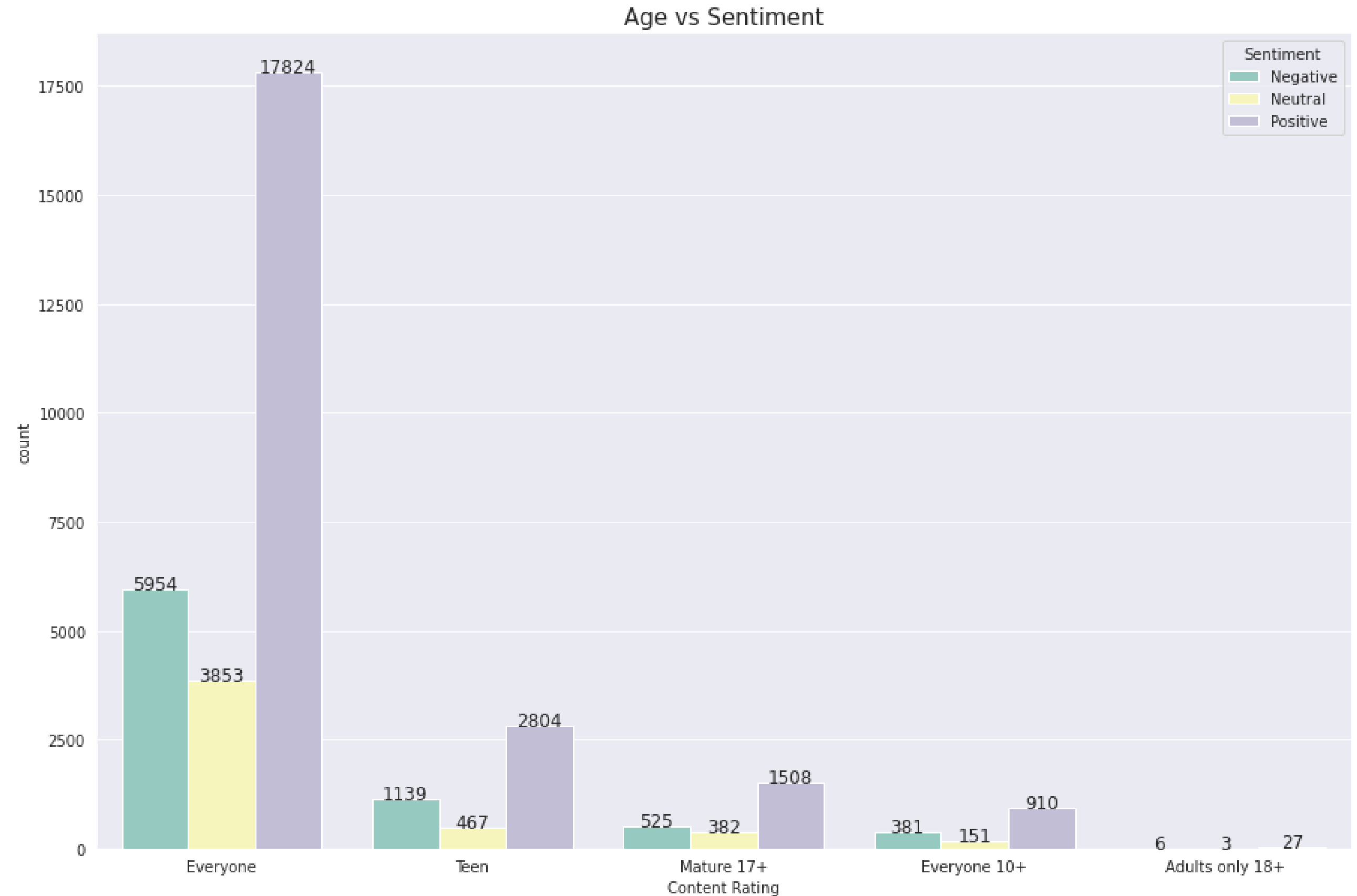
Average Sentiments

- More than 50% of the sentiments are positive.
- Only 22.10% are the negative sentiments.

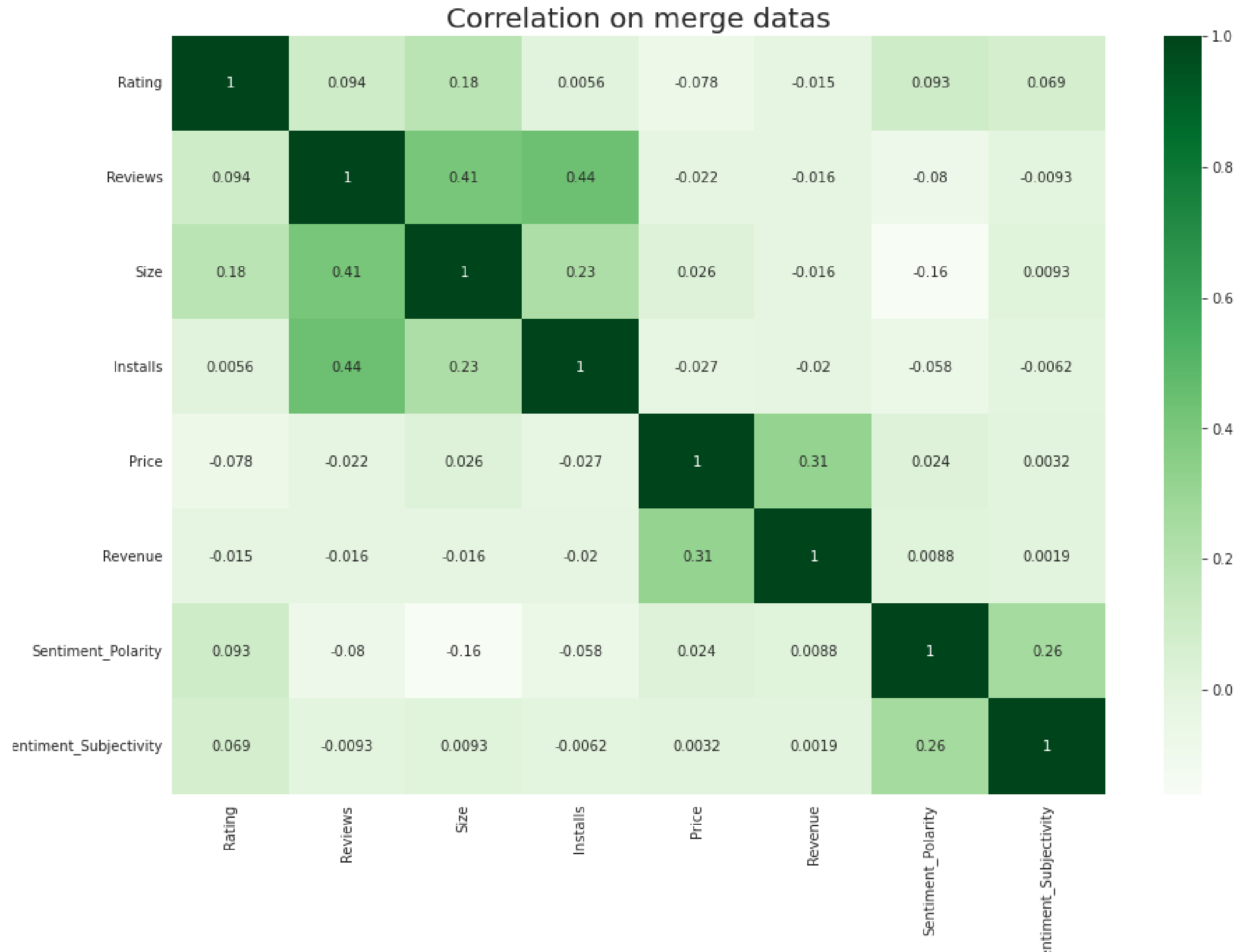


Comparision of age with sentiments

- The most positive sentiments (17824)
- The most negative are 5954



Correlation heatmap



Relation between paid and free apps

More than 90%
apps are free
on the
playstore.



Conclusion

- TEEN contains the most values in the content rating with the total number of 912 values.
- Most of the content rating given by teen is for the free apps that is something around 850.
- Most of the apps ratings are between 4.2 and 4.6. Maximum numbers of apps rated is 4.3 followed by 4.4
- Every Category under Content Rating is getting rating of 4+ which is good.
- Average rating of apps on Google Play Store is {4.17}
- The Health and Fitness and Books and Reference produce the best apps with 50% apps having a rating greater than 4.5. Interestingly, half of the Dating apps have a rating lower than the average.
- The Category who generated the highest revenue is 'Family' followed by 'Lifestyle' and 'Game'.
- More than 50% of the total sentiments are positive in the dataset user review.
- That most sentiments are from 'everyone'(17824) which is positive.
- 92.61% of the apps are free in the PlayStore app.

T

H
Y

A
O

N
U

K