Capstone Project

Play Store App Review Analysis

**Team Members-
Aanchal Kankrecha
Sonika Baheti**

CONTENT

- ❑ Points of interest
- ❑ Objective
- ❑ Data Overview
- ❑ Data Cleaning
- ❑ Exploratory Data Analysis
- ❑ Conclusion

Point of Interest

- ➢ Correlation matrix

- ➢ Reviews vs installation

- ➢ Category wise installation

- ➢ Category wise apps

- ➢ Uninstalled apps

- ➢ Top 3 gaming, communication and tooling Apps?

- ➢ Content rating wise apps

- ➢ Content rating wise installation

- ➢ Revenue generated by paid apps

- ➢ Percentage of free vs paid apps

- ➢ Free and paid apps per category

- ➢ Apps with their type having rating 5

- ➢ Distribution of Rating

Point of Interest

➤ Distribution of size of apps

➤ Average rating of each category

➤ No. of apps in each genres

➤ Rating vs Price

➤ Rating vs Installs

➤ Last updated Apps

➤ No. of translated reviews of different apps

➤ Maximum translated reviews Apps

➤ Distribution plot of sentiment polarity

➤ Distribution plot of sentiment subjectivity

➤ Percentage of sentiment count

➤ Conclusion

Objective

Our goal is to explore the insights of Play store dataset and extract the useful information to meet the choice and demand of the users and their preferences of app installation on the basis of Reviews and Ratings.

Data Overview

Dataset 1:

Play store review analysis table is loaded in colab as df, this table contain following columns:

App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last updated	Current version	Android version

Dataset 2:

User Review dataset is loaded in colab as df2, this table contains the following columns:

App	Translated Review	Sentiment	Sentiment Polarity	Sentiment Subjectivity

Play store app review analysis Data Summary

AI

In given play store dataset, we have following features:

1. App – Name of the app
2. Category – Category of the app
3. Rating - Over all user rating of the app out of 5 on the Play Store
4. Size – Size of app
5. Reviews – Number of user reviews for the app
6. Installs – Number of user downloads/installs for the app
7. Type – Paid or free
8. Price – Cost of the app
9. Content Rating – Age group the app is targeted at
10. Genres - An app can belong to multiple genres
11. Last Updated - Date when the app was last updated on Play Store
12. Current Version - Current version of the app available on Play Store
13. Android Version – Minimum required android version

User Review Data Summary

● App – An app name
● Translated Review – Review given by users in words.
● Sentiment – Sentiment given to an app by users (i.e., Positive, Neutral, Negative).
● Sentiment Polarity – The polarity of sentiment measures how negative or positive the context is. In the data we have, the polarity ranges from +1(Positive) to -1(Negative).
● Sentiment Subjectivity - The subjectivity of a sentiment is how likely that sentiment is to be based on personal opinions or public notions.

Overview and understanding of dataset

-> info(): It informs about data columns and data types.

-> head(): It returns the first five data.

-> tail(): It returns the last five data.

-> columns : It returns data columns

-> shape : It gives number of rows and columns in a tuple.

Head of the Play store dataset

```
[ ] df.head(10)
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10,000+	Free	0	Everyone	Art & Design	January 7, 2018	1.0.0	4.0.3 and up
1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free	0	Everyone	Art & Design;Pretend Play	January 15, 2018	2.0.0	4.0.3 and up
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	August 1, 2018	1.2.4	4.0.3 and up
3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	June 8, 2018	Varies with device	4.2 and up
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Design;Creativity	June 20, 2018	1.1	4.4 and up
5	Paper flowers instructions	ART_AND_DESIGN	4.4	167	5.6M	50,000+	Free	0	Everyone	Art & Design	March 26, 2017	1.0	2.3 and up
6	Smoke Effect Photo Maker - Smoke Editor	ART_AND_DESIGN	3.8	178	19M	50,000+	Free	0	Everyone	Art & Design	April 26, 2018	1.1	4.0.3 and up
7	Infinite Painter	ART_AND_DESIGN	4.1	36815	29M	1,000,000+	Free	0	Everyone	Art & Design	June 14, 2018	6.1.61.1	4.2 and up
8	Garden Coloring Book	ART_AND_DESIGN	4.4	13791	33M	1,000,000+	Free	0	Everyone	Art & Design	September 20, 2017	2.9.2	3.0 and up
9	Kids Paint Free - Drawing Fun	ART_AND_DESIGN	4.7	121	3.1M	10,000+	Free	0	Everyone	Art & Design;Creativity	July 3, 2018	2.8	4.0.3 and up

Head of User Review Dataset

```
[7] df2.head(10)
```

|	App	Translated_Review	Sentiment	Sentiment_Polarity	Sentiment_Subjectivity
0	10 Best Foods for You	I like eat delicious food. That's I'm cooking ...	Positive	1.00	0.533333
1	10 Best Foods for You	This help eating healthy exercise regular basis	Positive	0.25	0.288462
2	10 Best Foods for You	NaN	NaN	NaN	NaN
3	10 Best Foods for You	Works great especially going grocery store	Positive	0.40	0.875000
4	10 Best Foods for You	Best idea us	Positive	1.00	0.300000
5	10 Best Foods for You	Best way	Positive	1.00	0.300000
6	10 Best Foods for You	Amazing	Positive	0.60	0.900000
7	10 Best Foods for You	NaN	NaN	NaN	NaN
8	10 Best Foods for You	Looking forward app,	Neutral	0.00	0.000000
9	10 Best Foods for You	It helpful site ! It help foods get !	Neutral	0.00	0.000000

Data Cleaning

Price, Installs → Removed special characters('+','$',',')
Filled null values by 0.

Size → Removed special characters
Converted in Kbs.

Rating → Replaced null values with 0.

Reviews → Changed data type to float

Type → Filled nan values with Free

Last updated
Android version → Replaced varies with device values with
their respective modes.
Current version

Exploratory Data Analysis

Let the DATA tell the story ☺

Heatmap
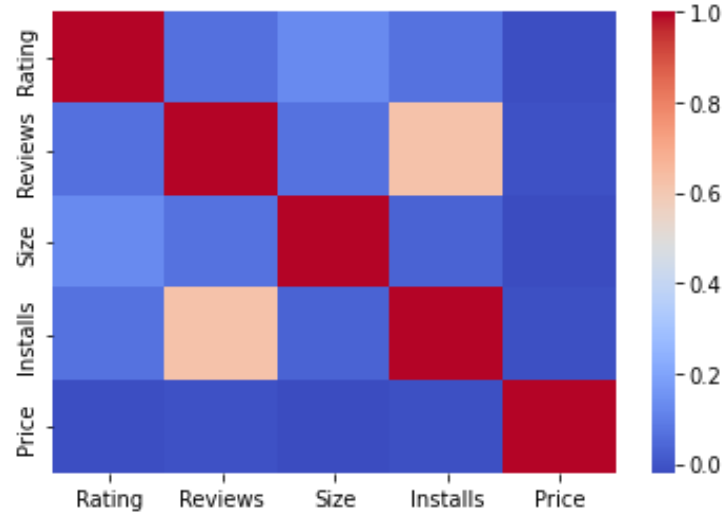
This shows that there exist positive correlation between Installs and Reviews and very less correlation among other Variables. Maximum no. of reviews are **8871** from the range 0 to 20,00,000 having maximum no. of installs as **7666**. Thus validating the conclusion of heatmap that there exist **positive correlation** between Reviews and Installs

To check no. of reviews of apps having 1 Billion installs

This graph suggests that **Facebook** has the maximum no. of reviews when installs is 1B.

Review of apps with installation of 1Billion
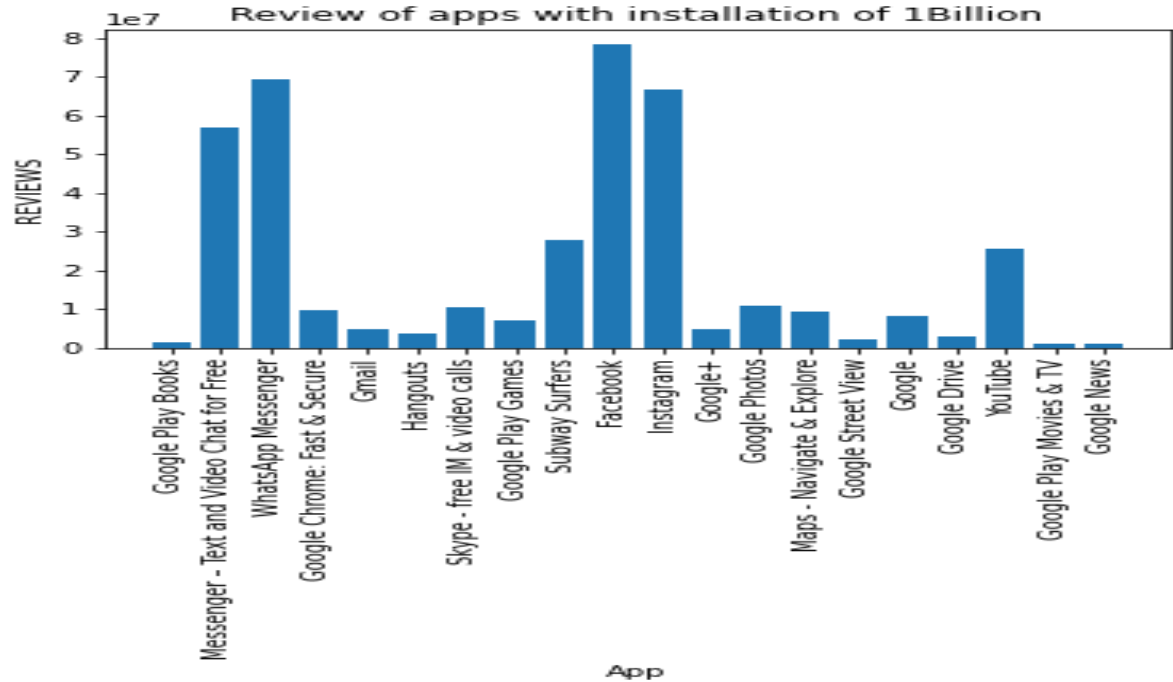
	App	Reviews
0	Facebook	78158306.0
1	WhatsApp Messenger	69119316.0
2	Instagram	66577313.0
3	Messenger – Text and Video Chat for Free	56642847.0
4	Clash of Clans	44891723.0
5	Clean Master- Space Cleaner & Antivirus	42916526.0
6	Subway Surfers	27722264.0
7	YouTube	25655305.0
8	Security Master - Antivirus, VPN, AppLock, Boo...	24900999.0
9	Clash Royale	23133508.0
10	Candy Crush Saga	22426677.0
11	UC Browser - Fast Download Private & Secure	17712922.0
12	Snapchat	17014787.0
13	360 Security - Free Antivirus, Booster, Cleaner	16771865.0
14	My Talking Tom	14891223.0
15	8 Ball Pool	14198297.0
16	DU Battery Saver - Battery Charger & Battery Life	13479633.0
17	BBM - Free Calls & Messages	12842860.0

Facebook has the highest number of reviews among all the apps ,i.e., **78,158,306.0**

List of apps having Zero installs

App
Popsicle Launcher for Android P 9.0 launcher
Ak Parti Yardım Toplama
AP Series Solution Pro
Ain Arabic Kids Alif Ba ta
cronometra-br
Pekalongan CJ
CX Network
Sweden Newspapers
Test Application DT 02
Command & Conquer: Rivals
EG	Explore Folegandros
EP Cook Book
Eu sou Rico
Eu Sou Rico
I'm Rich/Eu sou Rico/أنا غني/我很有錢

There are total 15 uninstalled apps.

Which category has the highest Installs

Category v/s Installs
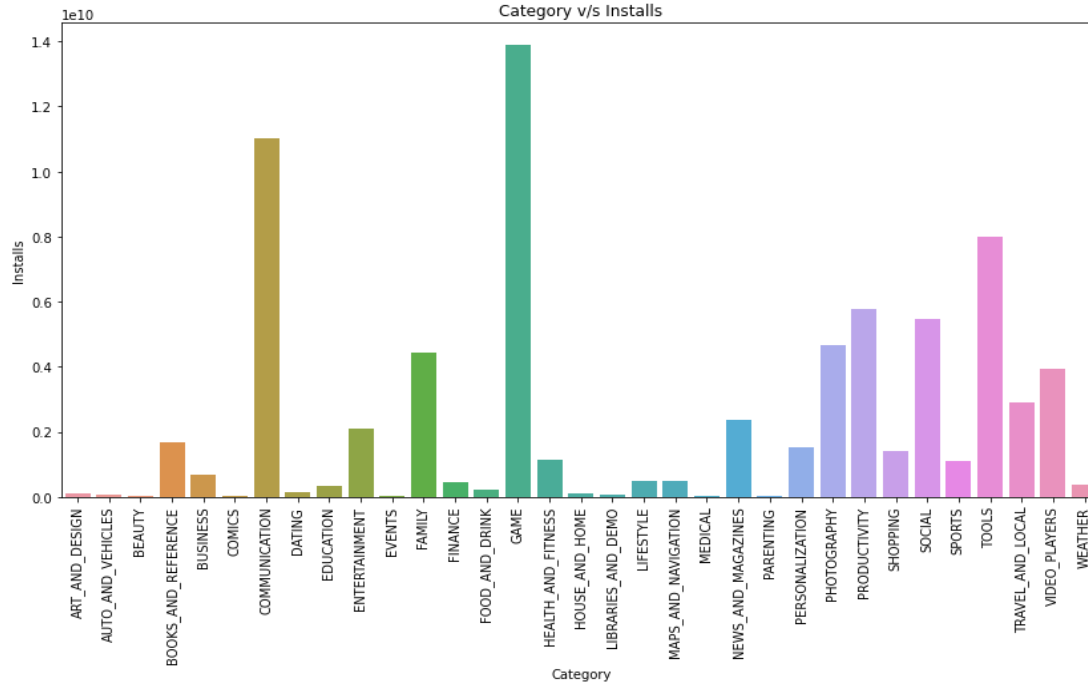
Games are the most downloaded category followed by communication and tools.

Top 3 Apps of categories having maximum installs.

game	communication	tool
Subway Surfers	Messenger – Text and Video Chat for Free	Google
Candy Crush Saga	Skype - free IM & video calls	Google Translate
Temple Run 2	WhatsApp Messenger	Clean Master- Space Cleaner & Antivirus

Number of apps belonging to each Content Rating

	Content Rating	No of Apps
0	Everyone	7903
1	Teen	1036
2	Mature 17+	393
3	Everyone 10+	322
4	Adults only 18+	3
5	Unrated	2
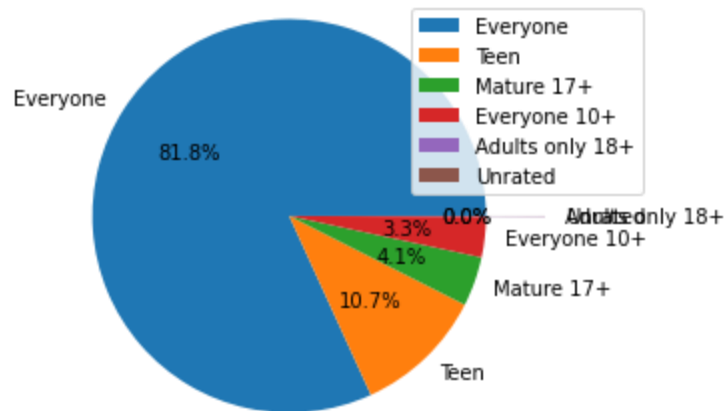
How much Installs are made for each type of content Rating ?

Content Rating	Installs
Everyone	52179352961
Teen	16487275393
Everyone 10+	4016271795
Mature 17+	2437986878
Adults only 18+	2000000
Unrated	50500
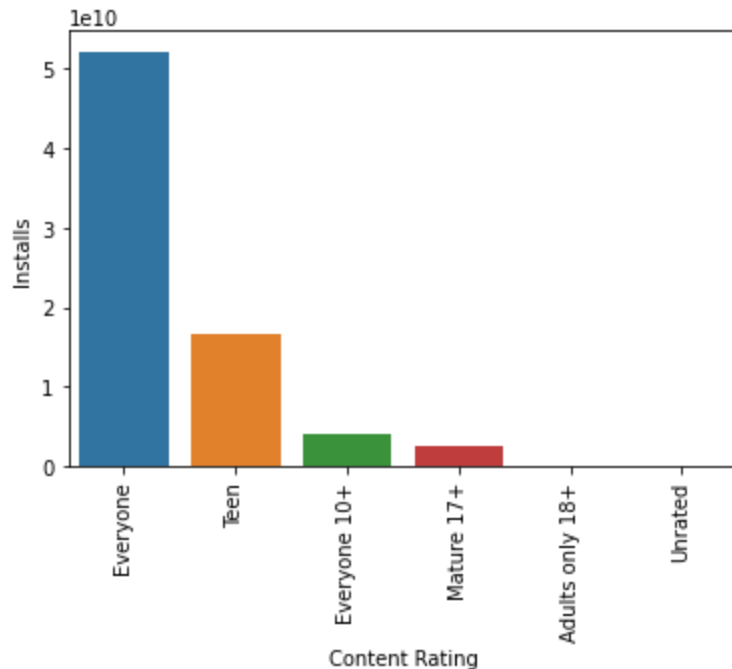
See the revenue of Top paid apps

App	Revenue
Minecraft	69900000.000000
I am rich	39999000.000000
I Am Rich Premium	19999500.000000
Hitman Sniper	9900000.000000
Grand Theft Auto: San Andreas	6990000.000000

Minecraft generates the maximum revenue.

Category wise app distribution
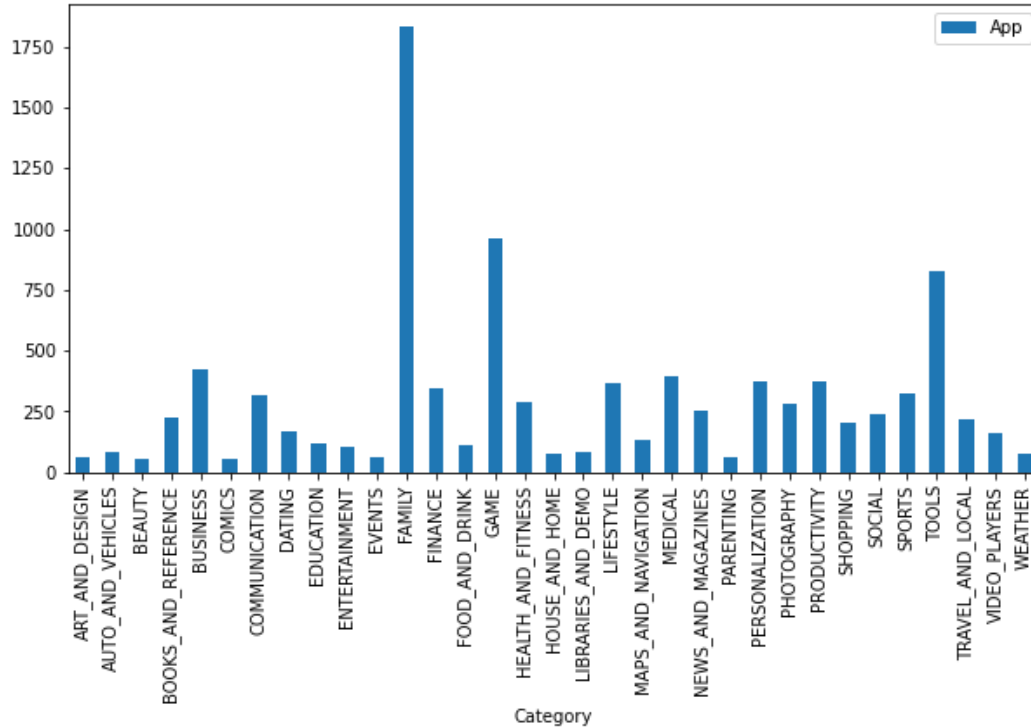
From graph it can be clearly seen that play store has most apps for family followed by games and tools

Free v/s Paid Apps

Type	App
Free	8903
Paid	756

Free 92.2%

7.8% Paid
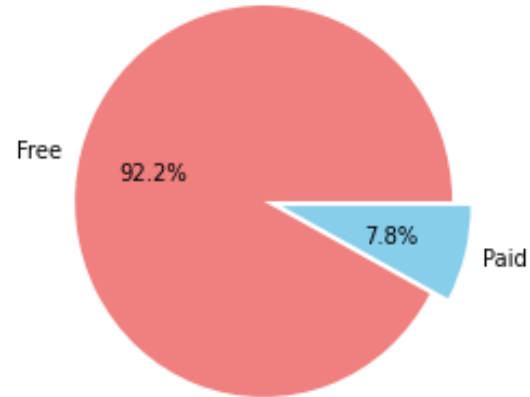
More free apps are available on play store.
Approximately 92% of the apps on play store are free while 8% are paid apps.

No. of free and paid apps for each category.

Count of applications in each category differentiated by their type
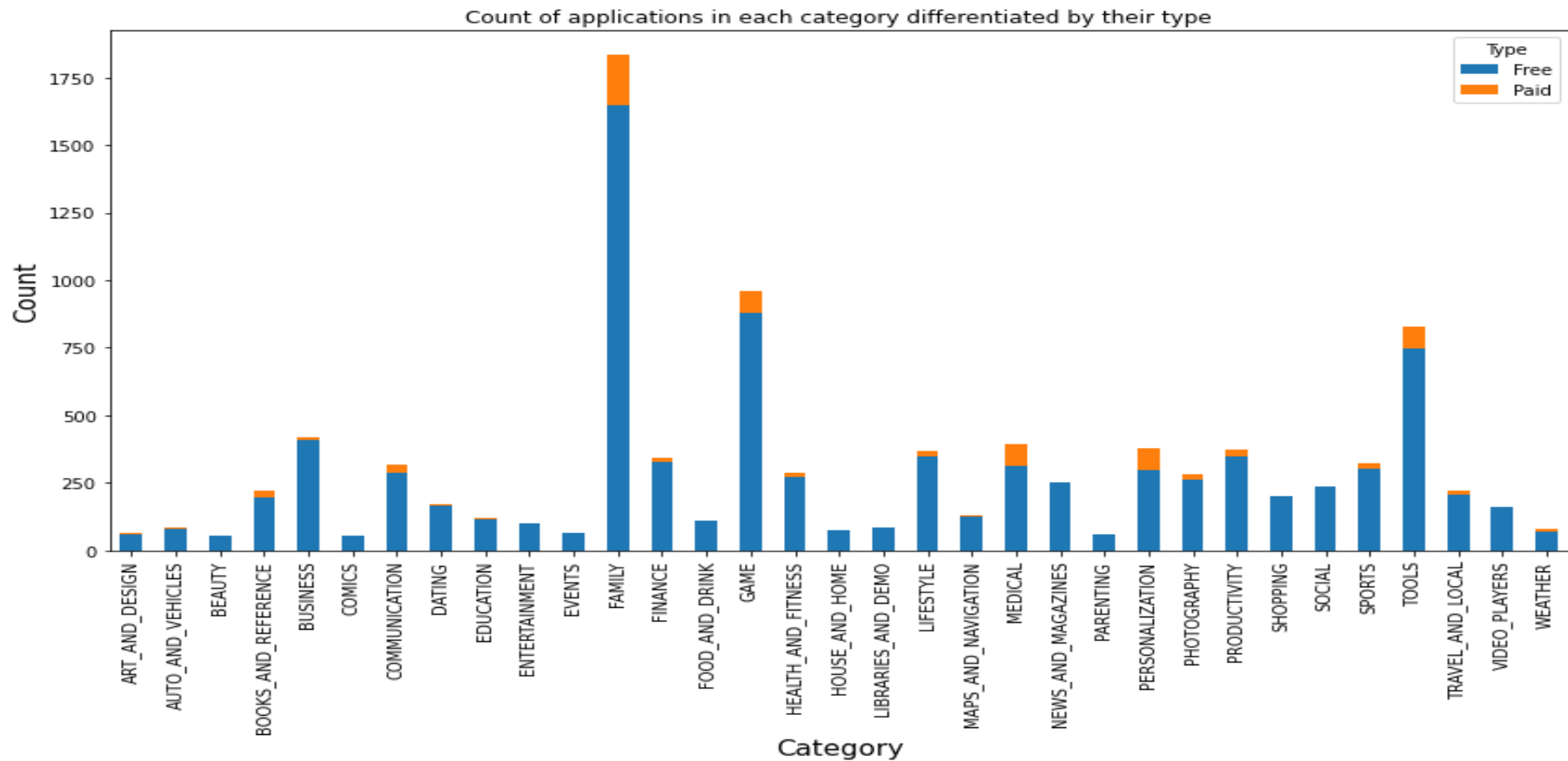
Histogram of rating
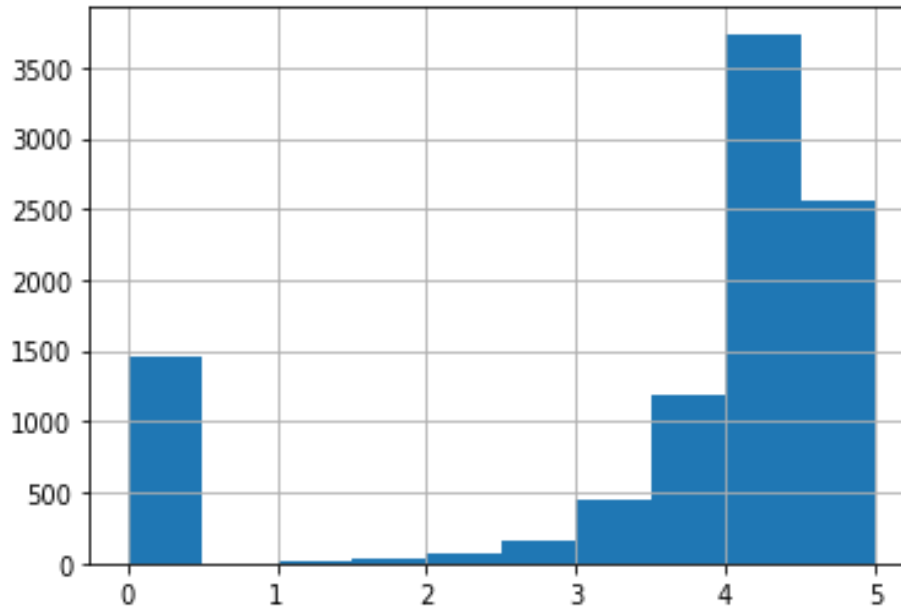
Most of the apps has rating between 3 and 5.
Average rating of app is 4.2 on play store.

Distribution of Size
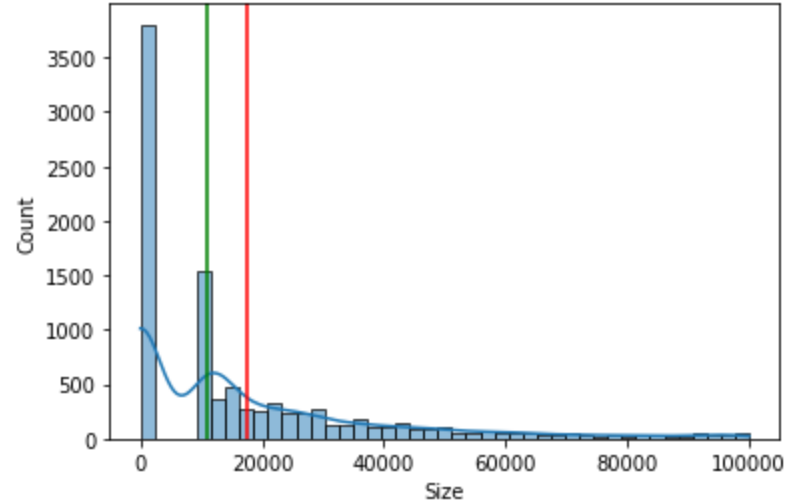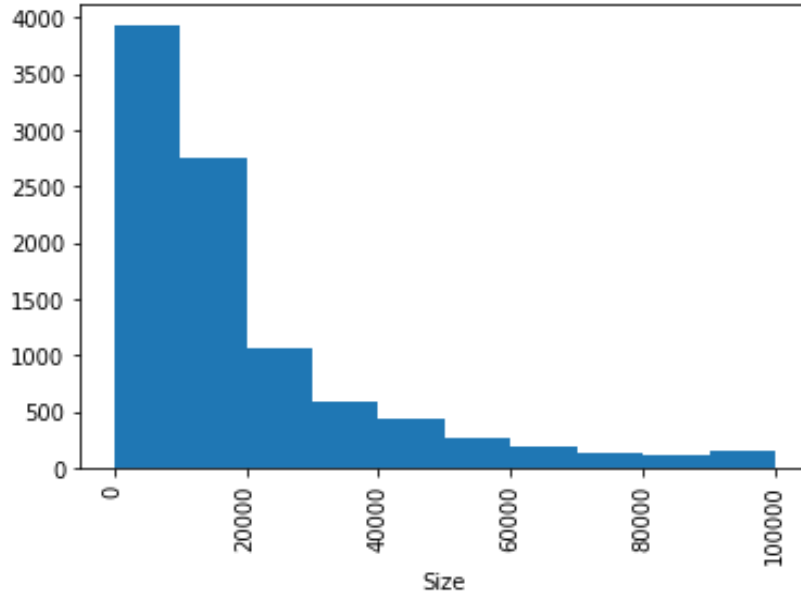
- From the above plot we can infer that data of size is **right tailed** because **mean>median**

  Most of the apps are of size between 1kb to 20 mb

- Mean Size of app is : 17435.69264934258 kb Median Size of app is : 11000.0 kb

Average rating of each category
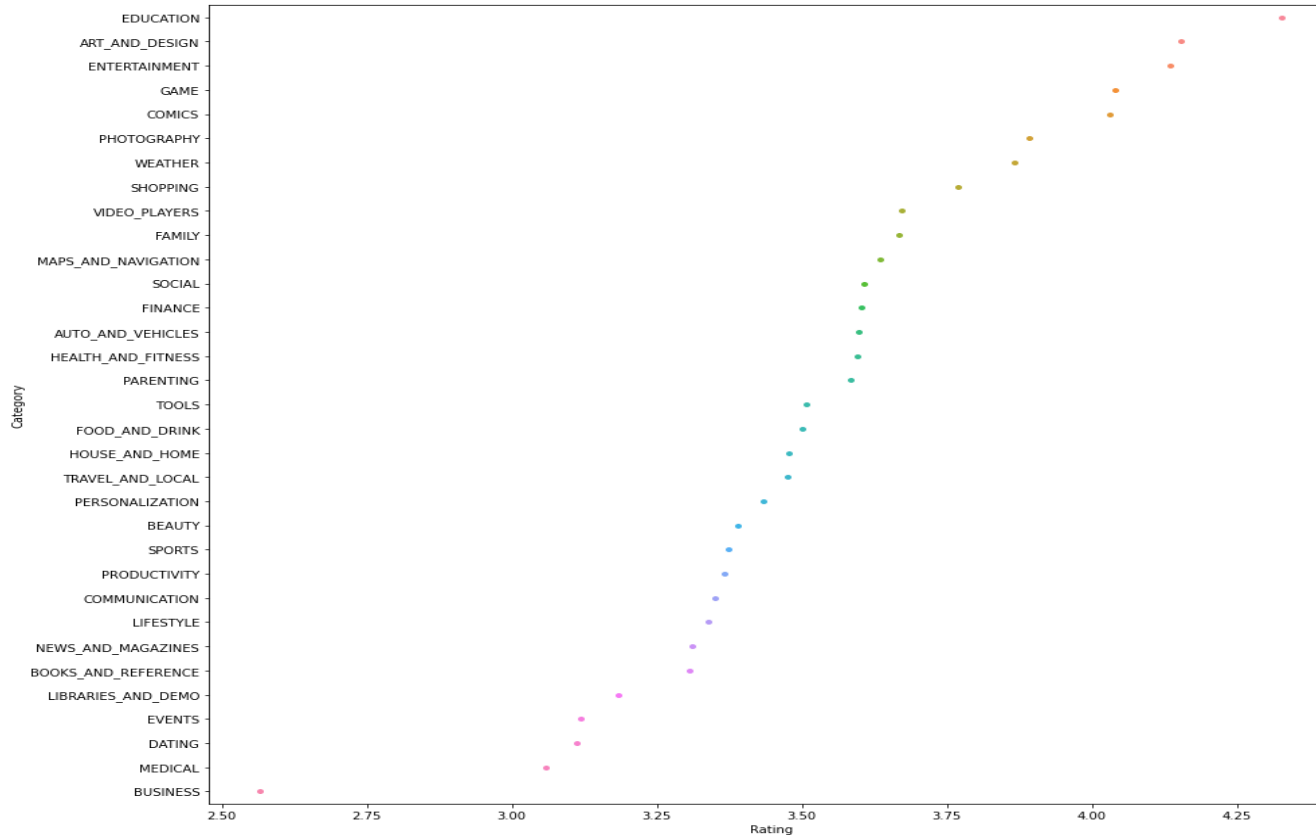
Education has the maximum average ratings of 4.3 followed by ART_AND_DESIGN 4.153125 & ENTERTAINMENT 4.135294
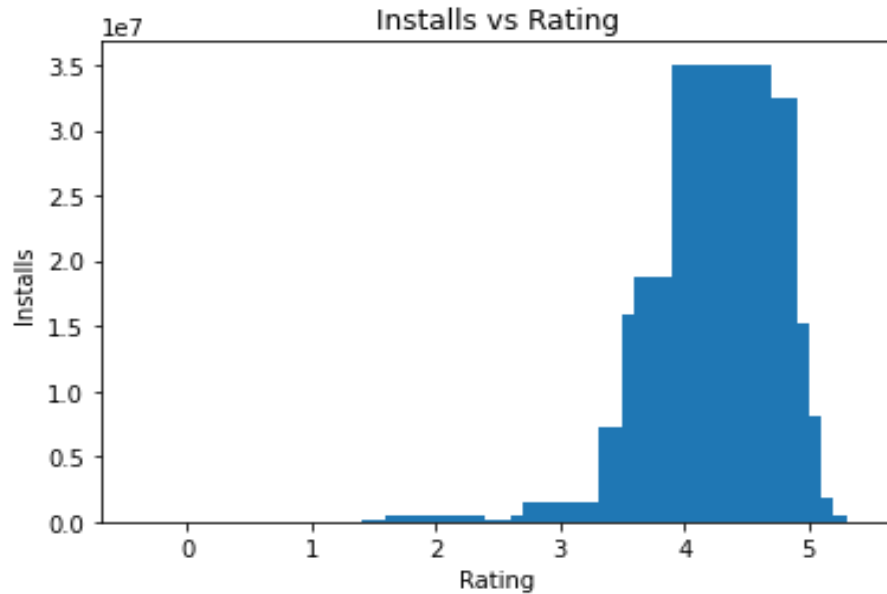
Genres with Number of Apps
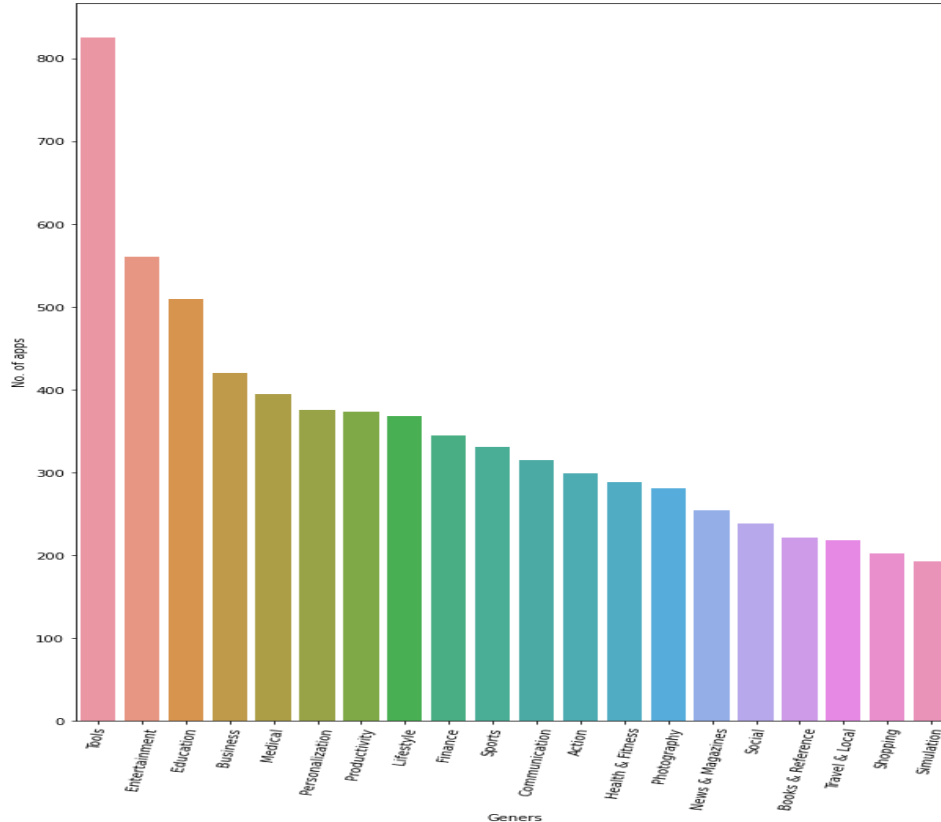
Tools are the Most trending Genres among all genres , followed by Entertainment and Education.
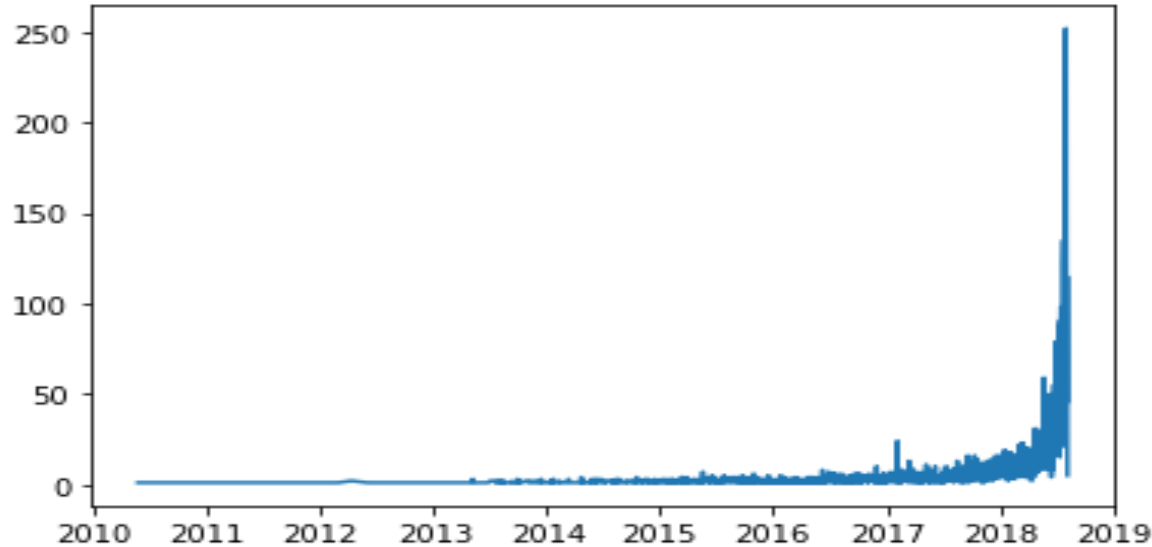
How Rating Changes with Price
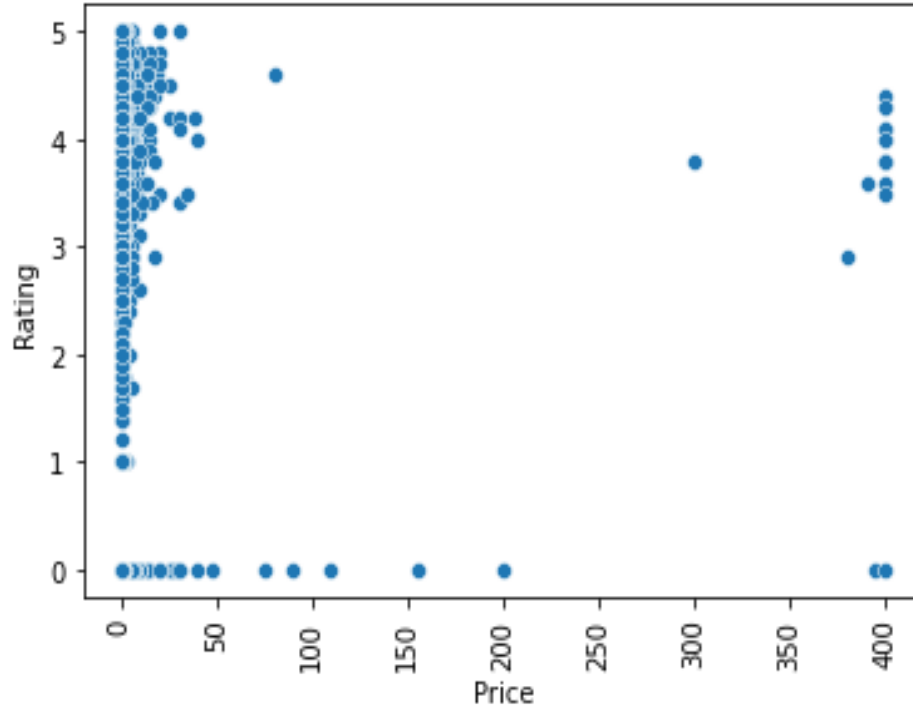
Different Ratings are given to free apps while high paid apps are given much good ratings.

How Rating and Installs are related?

Higher rated apps have more installs.
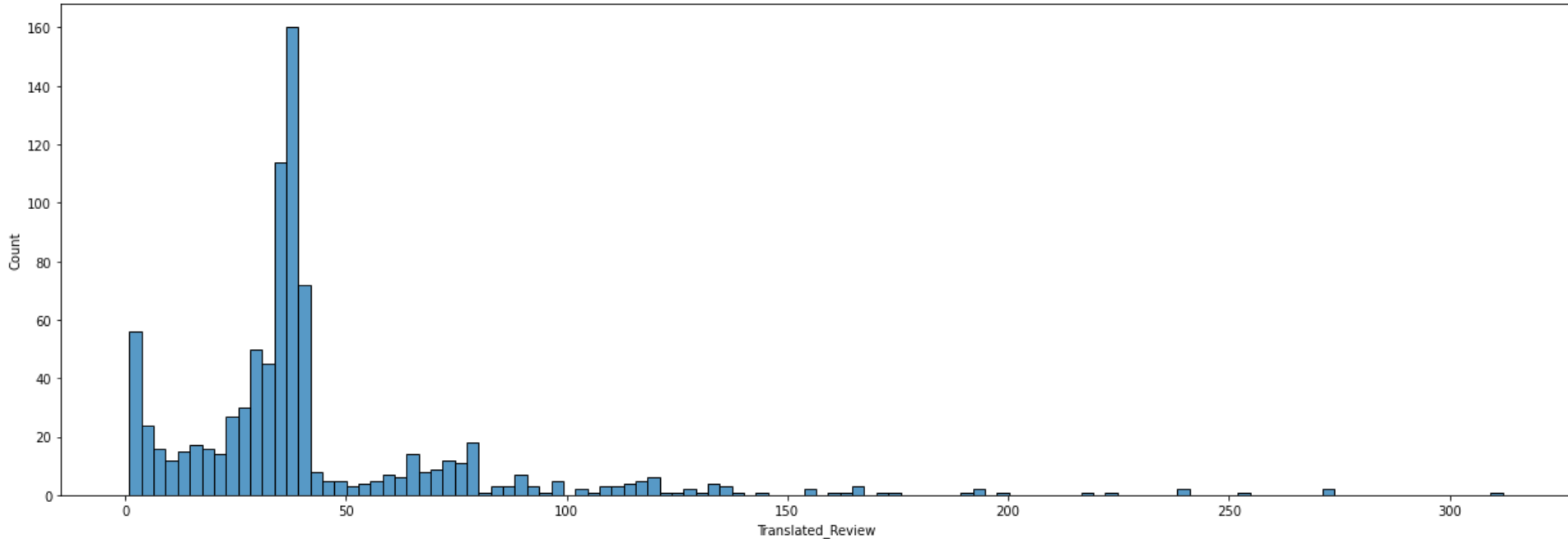
Have a view on Last Updated of Apps

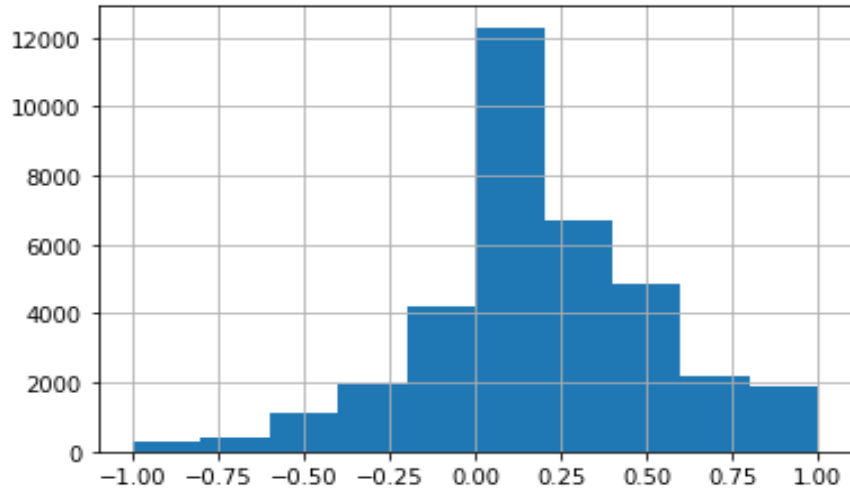| Last updated date	Max No. of apps updated
2018-08-03	252

- *Mostly Apps are updated in the year 2017,2018,2019.*

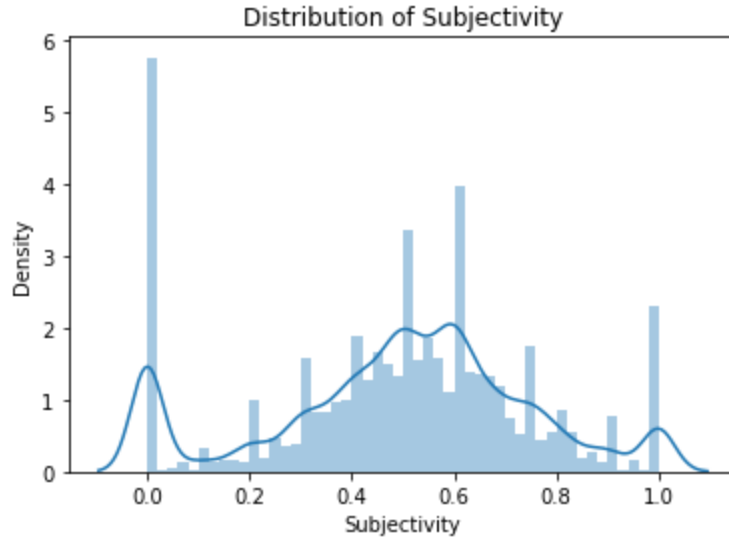# Distribution plot of translated reviews

App name	Translated_Review
Bowmasters	312

Distribution of sentiment polarity

Sentiment polarity is almost normally distributed.

Distribution of sentiment subjectivity
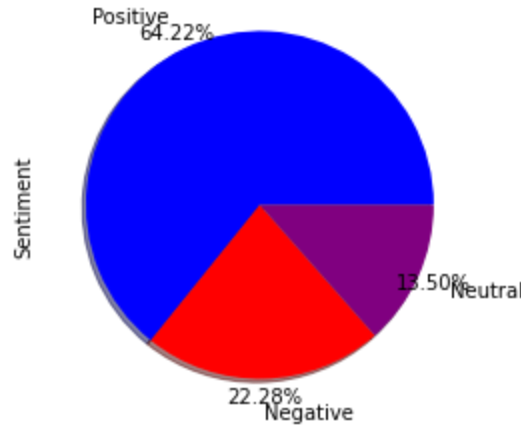

Distribution of Subjectivity

- *From the above plot sentiment subjectivity mostly lies between 0.4 to 0.8*

See the percent of Sentiment Counts

Positive	23073
Negative	8005
Neutral	4851

Here we can see that there are almost 64% positive Reviews given to apps ,i.e., most of the apps have positive effect on users.

Sentimental analysis for free and paid apps
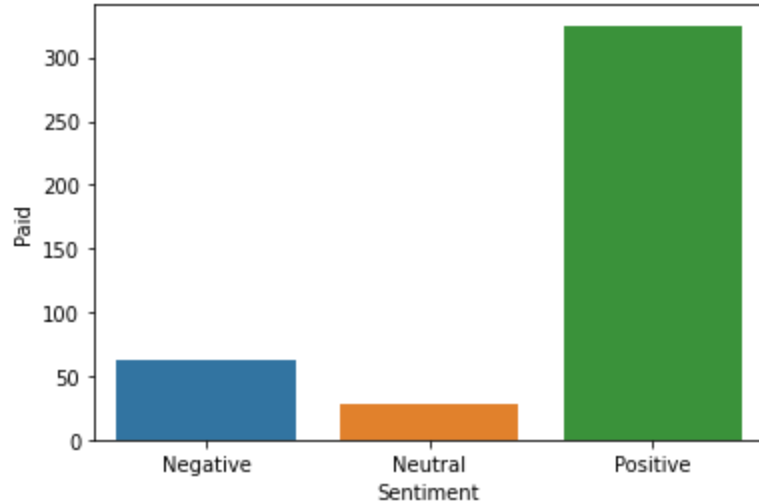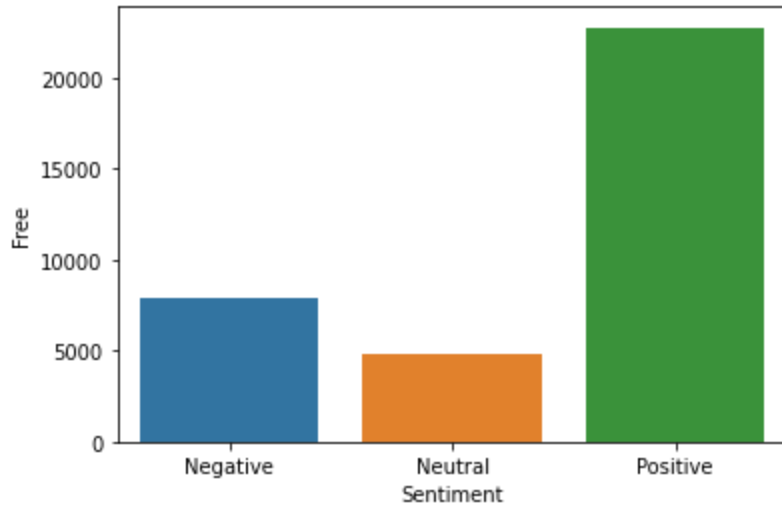
- Free and paid apps mostly have positive sentiments.
- Paid apps have low negative and neutral sentiments compare to free.

YAY!! The story Unfolds here...

CONCLUSION:

✓ Facebook has the maximum number of reviews.

✓ There are total 15 apps which are uninstalled.

✓ There are total 20 apps having 1 B installs.

✓ Games are the most downloaded category followed by communication and tools.

✓ Most of the apps are from content rating as everyone equivalent to 81.8%.

✓ Content rating as Everyone has the maximum no. of installs.

✓ Minecraft application generates the maximum revenue.

✓ Most of the apps are for family followed by games and tools.

✓ 92% of the apps on play store are free while 8% are paid apps.

CONCLUSION:

- ✓ Average rating of apps on play store is 4.2
- ✓ Most of the apps are of size between 1Kb to 20Mb.
- ✓ Education category has maximum average rating as 4.3
- ✓ Tools are the most famous genres having 826 apps on play store.
- ✓ Free apps has varying ratings while paid apps generally have high ratings.
- ✓ Mostly apps are updated in the year 2017,2018,2019.
- ✓ 'Bowmaster' app has the maximum translated reviews as 312.
- ✓ Sentiment subjectivity mostly lies between 0.4 to 0.8
- ✓ There are almost 64% positive Reviews given to apps ,i.e., most of the apps have positive effect on users.