Capstone Project - I
EDA : Play Store App Review

Presented by:
Sandeep K

Let's build a Successful App-development Business

Content

Introduction

- ✓ Mobile App Market is set to **grow 20%** by 2023
- ✓ Android Apps **comprise 90% of the Mobile App Market**
- ✓ Android is the most popular operating system in the world, with over **2.5 billion active users spanning over 190 countries.**
- ✓ Google Play was **launched on March 6, 2012**, bringing together Android Market marking a shift in Google's digital distribution strategy .
- ✓ Android is the dominant mobile operating system today more than **85% of all mobile devices running Google's OS.** The Google Play Store is the largest and most popular Android app store.
- ✓ There are more than **3.04 million apps found on Google Play Store**
- ✓ The Play Store apps data has enormous potential to drive app-making businesses to success.

Objective

AI

- ✓ Analyze and understand the important factors , which affects success of a Mobile App Business

- ✓ Identify successful App genres / categories to understand their overall performance and KPIs which lead them to Success

- ✓ Investigate various dependent variables and impact on each other, which may affect business significantly

- ✓ Understand the trends of some important key factors which makes App Business Successful

- ✓ Analyze and understand App market, customer segments, their sentiments, behavior pattern and subjectivity of sentiments

- ✓ Finally, draw appropriate conclusions, and

1. Provide recommendations to App Making Management Team
2. Technical Insights to App Development & Quality Team

- ✓ Which will help App Makers (management) and Developers to recognize the gap and meet customer expectations.

Problem Statement

AI

I had received a mandate from **ABC Co. Ltd**. That, company wants to build a successful Mobile App, which should be accepted by a **huge customer base** of **Play Store** and create a **positive Brand Image** among customers. It should be one of the **Successful App on Play Store.**

Deliverables :

1. Analyze and identify key factors responsible for app engagement and success.
2. Submit recommendations to select and plan development of a Successful App
3. Suggest Actionable Insights to Developers to build a competitive App in best possible manner

My Approach :

1. Import and clean Play Store App Database (App wise data and Sentiment Survey Data)
2. Give appropriate treatment to missing values and prepare Data Frames to analyze
3. Understand data features and explore important data variables
4. Analyze and understand Business Models for App making Business
5. Brief overview of Data Variables and their Relationship with each other
6. Identify Top 10 Performing : Genres | Categories | Apps
7. Analyze important factors responsible for App engagement and success as below,
• Impact of Ratings on Installs
• Content Rating wise Apps on play store (Free App & Paid App)
• Identify Top 3 Content Rating segments contributing most to the success of an App

Problem Statement

AI

- Top 3 Content rating segments giving best Ratings
- Prepare Annexures of Top 10 Apps preferred by each Top 3 Content Rating Segment
- Identify most preferred App Size in general and Content Rating segment wise
- Identify Top 5 Heavy (big size) App Categories

8. <u>Correlation Analysis of all Quantitative Variables & their impact as below,</u>
- Impact of App Updates on Size of the App
- Impact of Price on Size of the App
- App price trend analysis
- Identify Top 5 most accepted Android & Current Versions
- Analyze impact of Android & Current Versions on Reviews / Size / Price
- Investigate Rating / Reviews / Size impact on each other
- Investigate Price / Rating / Last Update impact on each other

9. <u>Correlation Analysis of Sentiment data variables & their impact as below,</u>
- investigate Distribution of Sentiment Polarity and Subjectivity
- Content rating wise sentiment analysis
- Analyze Sentiment relation with Size / Install / rating / last Update
- Identify Top 10 – Most positive sentiment Genres | Categories | Apps

10. <u>Understand relation of Sentiment Polarity & Subjectivity</u>
11. <u>Understand overall Translated Reviews and Sentiment through - Word Cloud</u>
12. <u>Conclusion</u>
- Provide recommendations to App Making Management Team
- Technical Insights to App Development & Quality Team

Data Pipeline

Data Processing – 1: In this first part we have removed unnecessary features. Also handled some missing and inconsistent values. Since there were many records with all null values and inconsistent data.

Data Processing – 2: In this first part I have manually gone through all features selected from above Part – 1, and encoded the categorical features, changed the columns containing date-time values.

EDA : In this segment, I have performed exploratory data analysis of various important factors which affects overall App experience, technology and customer sentiments.

After analyzing these factors, I have presented the same in the form of Visualization, which helps us to understand trend and relations of all variables.

Then same has been concluded in the form of,
1. Recommendations to App Making Management Team
2. Technical Insights to App Development & Quality Team

Description of Dataset - 1

AI

Columns Description : <u>Play Store App Data</u>

1. App :   This column consist of the names of the Mobile Apps.
2. Category :   This column consist of the category of the Mobile Apps.
3. Rating :   This column consist of the user ratings of the Mobile Apps.
4. Reviews :   This column consist of the user reviews of the Mobile Apps.
5. Size :   This column consist of the size of the Mobile Apps in KB / MB.
6. Install :   This column consist of the number of installations of the Mobile Apps.
7. Type :   This column consist of the subscription type (free/paid)of the Mobile Apps.
8. Price :   This column consist of the subscription cost of the Mobile Apps.
9. Content Rating: This column consist of the user group category of the Mobile Apps.
10. Genres: This column consist of the Genre type name of the Mobile Apps.
11. Last Updated: This column consist of the Last date of app update of the App.
12. Current Ver: This column consist of the name of the current version of the Mobile App.
13. Android Ver: This column consist of the name of the android version for the Mobile App.

Description of Dataset - 2

AI

Columns Description : <u>Reviews Data</u>

App :   This column consist of the names of the Mobile Apps.

Translated_Review :   This column consist of the system translated user reviews of the Mobile Apps.

Sentiment :   This column consist of the sentiment type (Positive / Negative / Neutral) of the user review for Mobile App.

Sentiment Polarity: This column consist of the category of Sentiment Polarity for the user review Sentiment polarity defines the orientation of the expressed sentiment. It defines the extent of the sentiment of the text in numerical data.

Sentiment Subjectivity :   This column consist of a measure of the intensity of the sentiment one expressed through their text. It refers to the text that contains text which is usually expressed by a human having typical moods, emotions, and feelings.

Data Cleaning

1. Created copies of Main datasets and started working on the same, to avoid disturbance in original data frame.

2. Checked unique / null values and Datatypes of all columns

3. Identified and removed duplicate records

4. Checked and updated Missing values of Play Store dataset & Reviews Dataset

5. Handled missing values with the help of Mean & Median values of dataset

6. Converted datatypes into appropriate formats

- Some numerical datasets were having objective datatype – converted these datatypes to 'int' & 'float'

- Corrected Date data format as Date-Time format

6. Column "Size" was having different size units, converted these sizes in an uniform unit (KB)

7. Column 'Price' were having '$' sign, removed the same and changed column datatype to 'float'

8. Corrected Column names by removing spaces and replacing them with underscore

AI

Data Preparation

- **Loading the data sets:** Two datasets, First Play store app dataset and User Reviews dataset.
- **Import Libraries:** NumPy, Pandas, Seaborn, Matplotlib & Word Cloud
- **Data cleaning:** Null values, Finding and removing Outliers, Removing duplicate data.
- **Data Imputation:** Filling the missing categorical values with mode and numerical values with median. Conversion of price, installs, reviews into numerical values.
- **Exploratory Data Analysis:** Analyzing the data sets to summarize their main characteristics using statistical graphics and data visualizations method.

Get...Set....Go....

Let's understand :
App Business Models

Finalizing **Business Model** is one of the important cornerstone for any Business, hence before proceeding further with App Development and launching Strategy, one need to finalize the appropriate Mobile App Business model (Free / Paid) for the desired success of the App.

There are two types of App Development Business Models.

1. Free Apps
2. Paid Apps

Let's follow below steps to finalize the same,

1. First Analyze available statistical data for both the Apps (Free & Paid).
2. Understand their Pro-Cons with the help of stats
3. Identify critical KPIs and performance of existing successful Apps

Analysis of Free & Paid App

After having a brief quantitative analysis for Mobile App Business, we can easy understand that,

1. The Market share of Free Apps and Paid Apps is drastically different
2. Contribution of Customer segments are also different in both
3. Sentiment Polarity & subjectivity is not same for the both
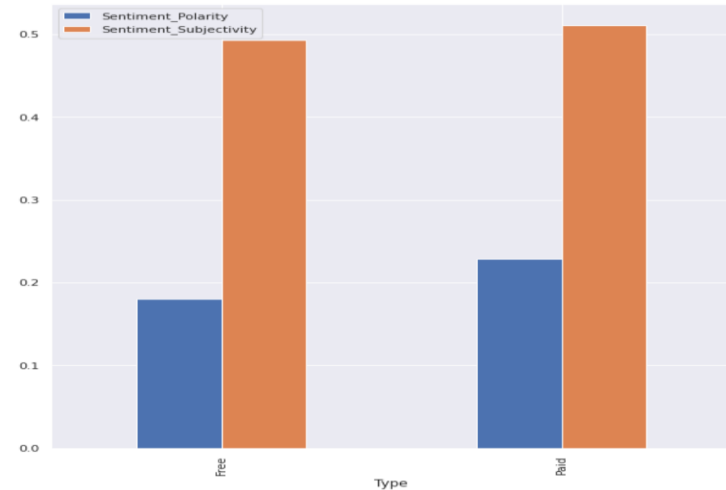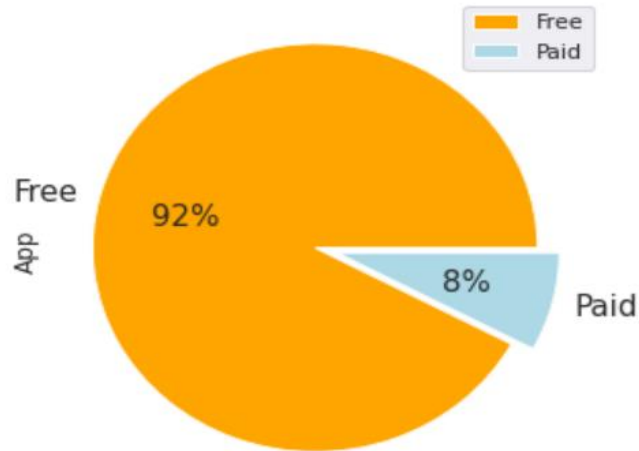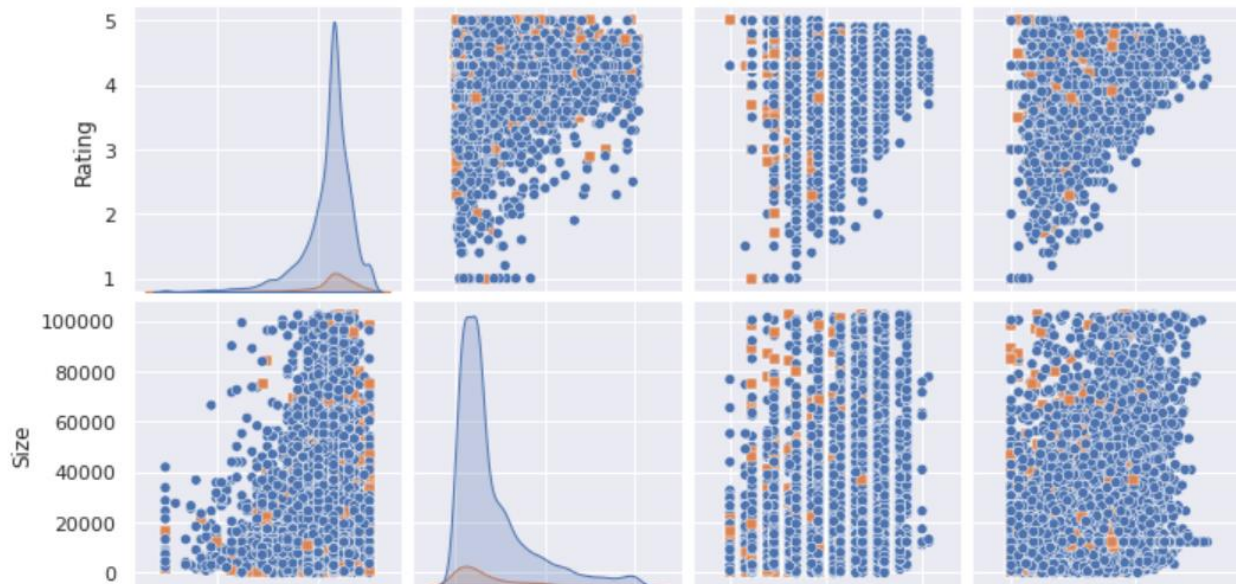
Let's understand :
Success factor relations

To have a glance on the relationship of various factors, we can have a glance on given 'Pairwise Plot' for below pairs,

1. Rating Vs Rating
2. Rating Vs Size
3. Rating Vs Installs
4. Rating Vs Reviews

5. Size vs rating
6. Size Vs Size
7. Size Vs Installs
8. Size vs Reviews

Pairwise Plot - Rating, Size, Installs, Reviews

Let's understand :
Success factor relations

To have a glance on the relationship of various factors, we can have a glance on given 'Pairwise Plot' for below pairs,

1. Rating Vs Installs
2. Size Vs Installs
3. Installs Vs Installs
4. Reviews Vs Installs

5. Reviews Vs Rating
6. Reviews Vs Size
7. Reviews Vs Installs
8. Reviews Vs Reviews
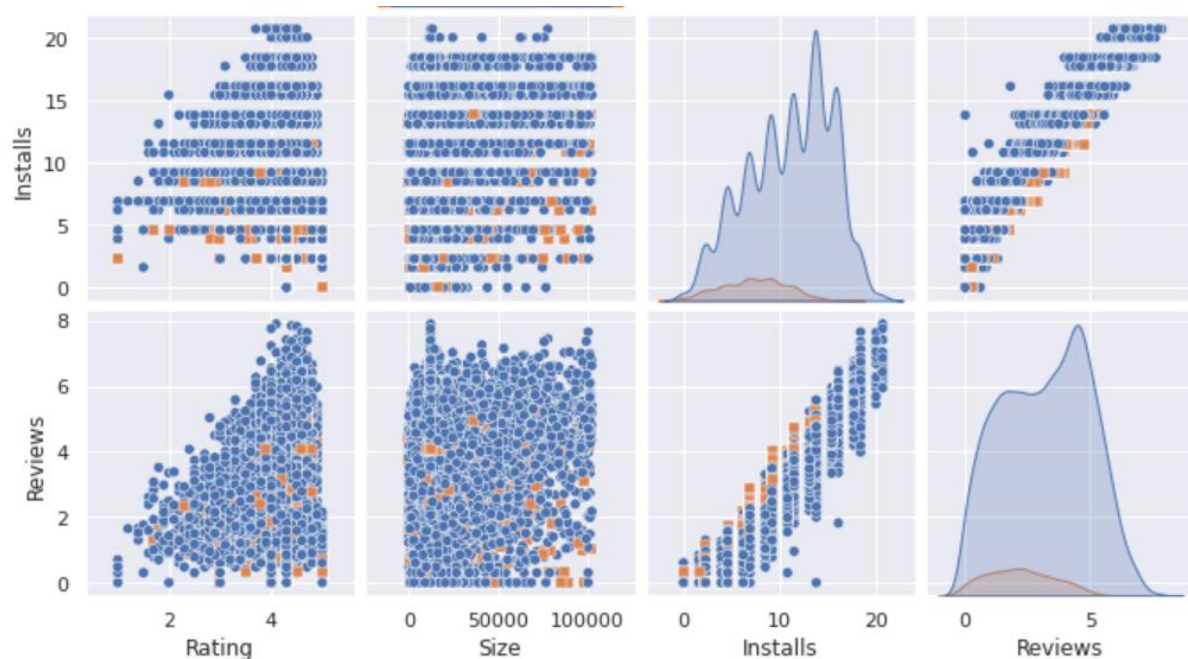
Free Apps

From **avg. installs** we can understand — Free App **attracts huge downloads**

From **avg. reviews** we can understand — Free App **acquires more views & reviews**

As a business model, Free App allows **multiple revenue streams** (Ads, in-app purchases etc.)

From **average ratings** we can understand that - Free app may see lower engagement

As a business model — There is an **uncertainty about revenue generation** as there is no fixed revenue stream

As per above Stat : **92% Apps are free**, its **difficult to stand out** and become successful

Paid Apps

From **avg. ratings** we can understand that - Paid apps must be having better features than Free Apps

Due to paid subscription, there is a **guarantee of revenue**

As per above Stat : **8% Apps are paid, competition is less** than Free Apps

From **avg. installs** we can understand that - **Unable to attract huge downloads**

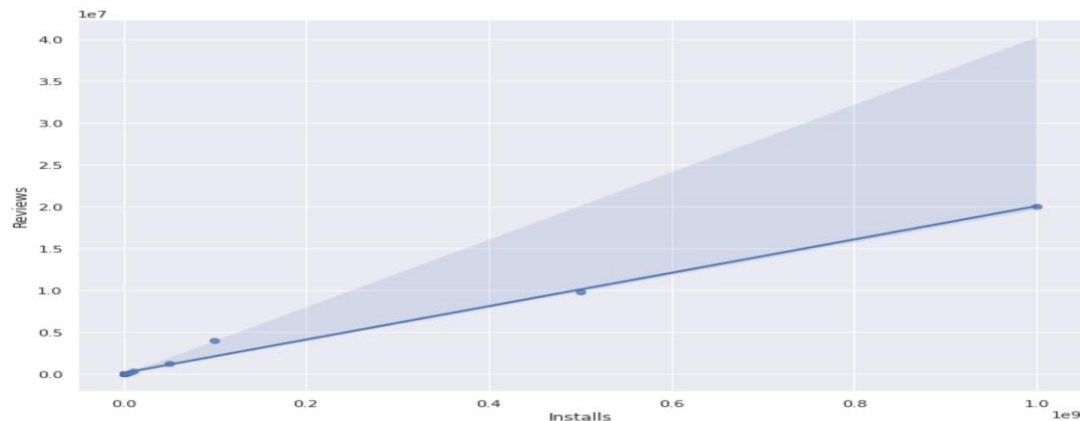Need to provide **better features than Free app**, which may require more investment than Free Apps

May **not have multiple Revenue Streams** - as it is paid app and mostly paid customer don't want interruptions like Ads.

Installs Vs Reviews

Inference 1 – Free App

1. For free apps - reviews are directly proportional to installs
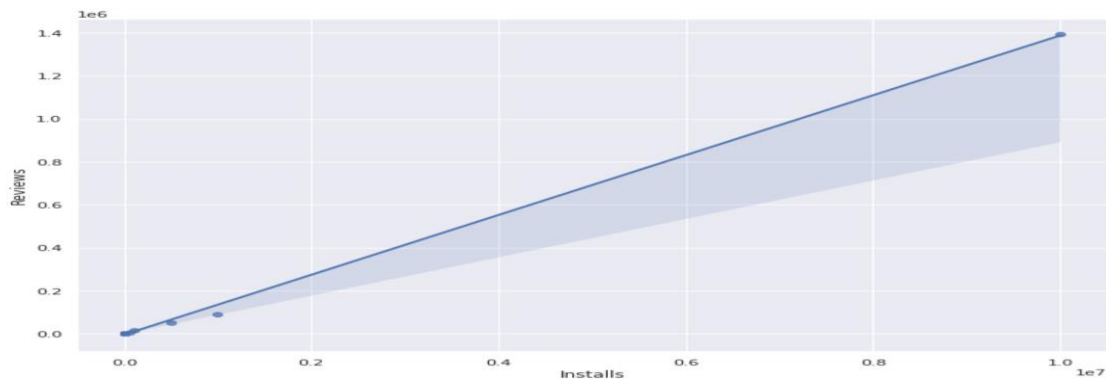
2. As installs increases reviews are also increasing

Inference 2 – Paid App

1. For Paid Apps also, reviews are directly proportional to installs

2. As installs increases reviews are also increasing

FREE App Plot

PAID App Plot

Installs Vs Rating

Inference 1 – Free App

1. There is positive increment in Ratings with respect to Installs for Free Apps
2. Rate of increase is slow but it's crucial as it affects performance of the App significantly.

Inference 2 – Paid App

1. As per graph, there is a good positive increment in Ratings with respect to Installs for Paid Apps
2. Rate of increase of rating is higher than the increment rate of ratings in Free Apps.
3. Paid apps are getting good rating faster because they have a dedicated customer base which reviews Paid Apps faster than Free Apps
4. Quality of service of Paid apps helping them to gain higher rate of incremental rating.

FREE App Chart
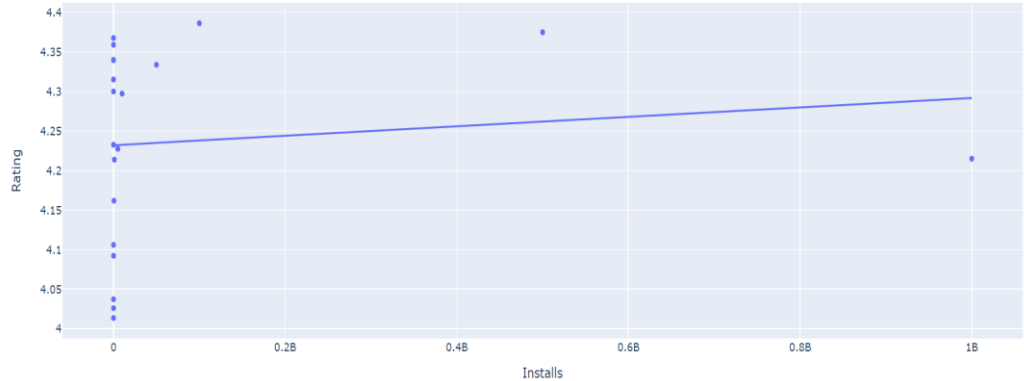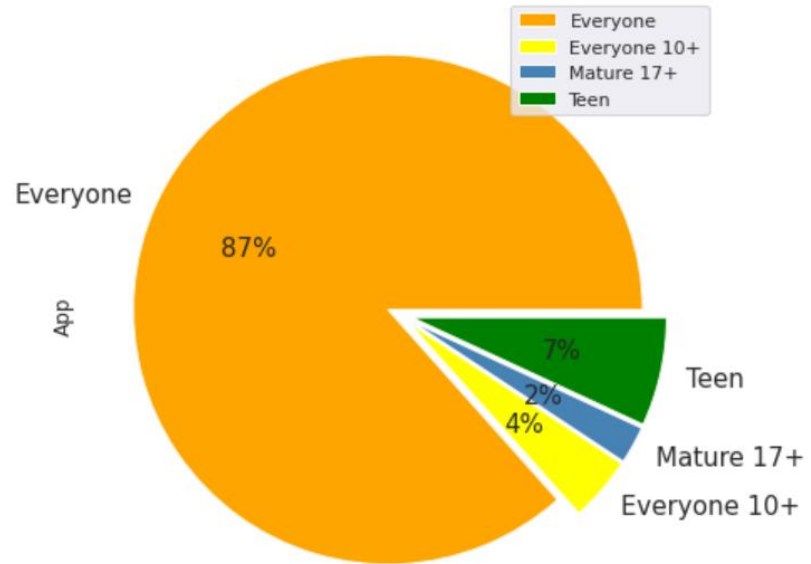
**PAID App Chart**
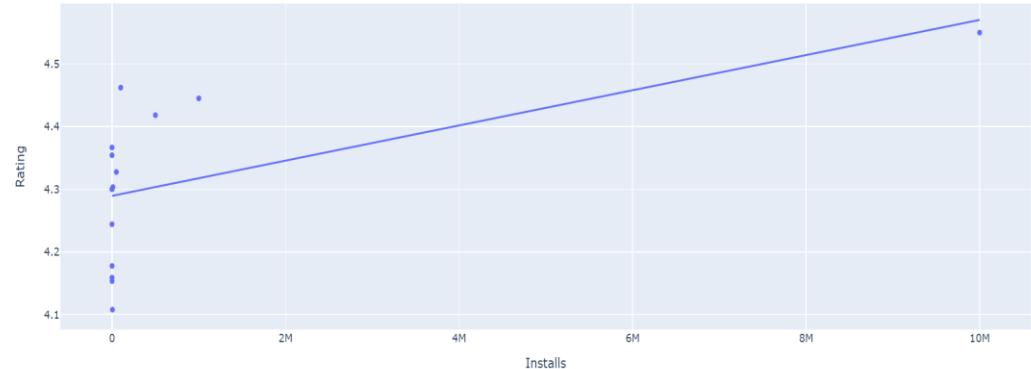
Content Rating Analysis

INFERENCE −

1.In Paid Apps segment 'Everyone' category apps contributing to 87% of the App Market. Which is bigger than Free Apps. Hence we can expect very high competition in this segment.

2.Second biggest category is Teen category apps, which contributes to 7%, we can expect a moderate competition here as well.

3.In Everyone 10+, Mature 17+ category we can expect less competition as no of Apps are around 3-4 % of the market.

Content Rating Analysis
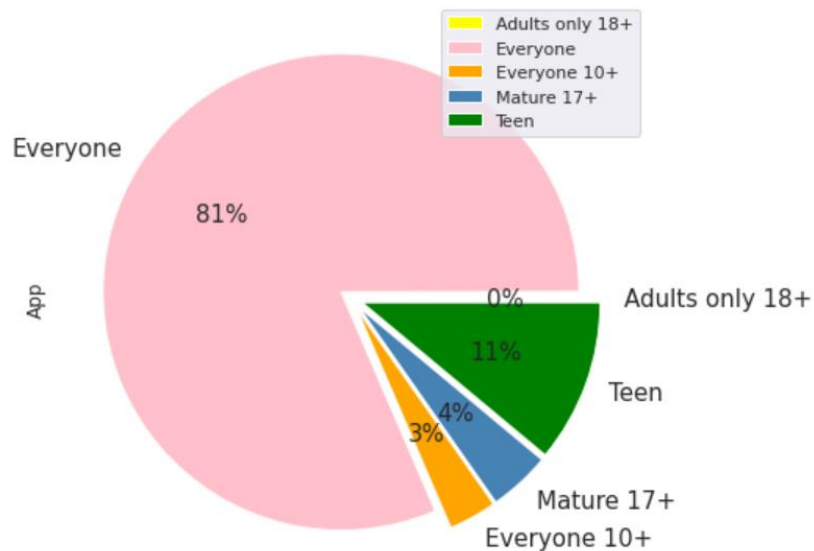
INFERENCE –

1.In Free Apps segment 'Everyone' category apps contributing to 81% of the App Market. Hence we can expect very high competition in this segment.

2.Teen category app contributes to 11% of Free App market. This is the second biggest category, hence we can expect a moderate competition here as well.

3.In Adult 18+, Mature 17_ and Everyone 10+ category we can expect less competition as no of Apps are around 3-4 % of the market.

Top 3 Content Rating Segments

Best Ratings

96.5% contribution to overall Free App review

Content_Rating	Rating
Everyone	4.443094
Everyone 10+	4.444000
Teen	4.427742
Mature 17+	4.369444

91.5% contribution to overall paid App review

Content_Rating	Rating
Everyone	4.565306
Teen	4.521429
Everyone 10+	4.500000
Mature 17+	4.500000

Top 3 - Content Rating Segments – Free Apps

Top Installs | Ratings | Reviews

Content_Rating	Rating	Reviews	Size	Installs
Everyone	4.443094	1.962228e+06	25415.329650	6.976059e+07
Everyone 10+	4.444000	3.711211e+06	51507.200000	6.920000e+07
Teen	4.427742	2.384168e+06	38573.741935	6.716129e+07
Mature 17+	4.369444	1.773997e+06	28407.444444	4.777778e+07

Top 3 - Content Rating Segments – Paid Apps

Best Installs | Ratings | Reviews

Content_Rating	Rating	Reviews	Size	Installs
Mature 17+	4.500000	378627.000000	28160.000000	5.500000e+06
Everyone 10+	4.500000	338314.000000	43904.000000	1.725000e+06
Everyone	4.565306	39785.040816	24360.040816	3.775510e+05
Teen	4.521429	32151.285714	44609.714286	3.214286e+05

Top 3 Content Rating Segment

AI

Contribution % in App Market

Free App - Below 3 Segments contributing at the extent of	Paid App - Below 3 Segments contributing at the extent of

96.5%

91.5%

Free App	Paid App
1. Teen	1. Mature 17+
2. Everyone 10+	2. Everyone 10+
3. Everyone	3. Everyone

Some IMP Q & A

Annexure 3 – Top 10 Free Apps – Preferred by "Everyone" Segment

App	Rating	Reviews	Installs
Google Photos	4.5	10847682.0	1.000000e+09
WhatsApp Messenger	4.4	69109672.0	1.000000e+09
Google	4.4	8021623.0	1.000000e+09
Google Drive	4.4	2728941.0	1.000000e+09
Google Chrome: Fast & Secure	4.3	9642112.0	1.000000e+09
Maps - Navigate & Explore	4.3	9231613.0	1.000000e+09
Gmail	4.3	4604324.0	1.000000e+09
Google Street View	4.2	2129707.0	1.000000e+09
Skype - free IM & video calls	4.1	10484169.0	1.000000e+09
Messenger – Text and Video Chat for Free	4.0	56642847.0	1.000000e+09

Annexure 4 – Top 10 Free Apps – Preferred by "Everyone 10+" Segment

App	Rating	Reviews	Installs
Subway Surfers	4.5	27711703.0	1.000000e+09
Flipboard: News For Our Time	4.4	1284018.0	5.000000e+08
Clash of Clans	4.6	44881447.0	1.000000e+08
Clash Royale	4.6	23125280.0	1.000000e+08
Shadow Fight 2	4.6	10981850.0	1.000000e+08
Minion Rush: Despicable Me Official Game	4.5	10214862.0	1.000000e+08
ROBLOX	4.5	4443407.0	1.000000e+08
Zombie Tsunami	4.4	4918776.0	1.000000e+08
Jetpack Joyride	4.4	4637439.0	1.000000e+08
Plants vs. Zombies FREE	4.4	4064868.0	1.000000e+08

Annexure 5 – Top 10 Free Apps – Preferred by "Teen" Segment

App	Rating	Reviews	Installs
Instagram	4.5	66509917.0	1.000000e+09
YouTube	4.3	25623548.0	1.000000e+09
Google Play Games	4.3	7168735.0	1.000000e+09
Google+	4.2	4828372.0	1.000000e+09
Facebook	4.1	78128208.0	1.000000e+09
Google Play Books	3.9	1433233.0	1.000000e+09
Google News	3.9	878065.0	1.000000e+09
Google Play Movies & TV	3.7	906384.0	1.000000e+09
UC Browser - Fast Download Private & Secure	4.5	17712922.0	5.000000e+08
Facebook Lite	4.3	8595964.0	5.000000e+08

Annexure 6 – Top 10 Paid Apps – Preferred by "Mature 17+" Segment

App	Rating	Reviews	Installs
Hitman Sniper	4.6	408292.0	10000000.0
Grand Theft Auto: San Andreas	4.4	348962.0	1000000.0
Meme Generator	4.6	3771.0	100000.0
XCOM®: Enemy Within	4.2	13752.0	100000.0
Mental Hospital V	4.2	2460.0	50000.0
Essential Anatomy 3	4.1	1533.0	50000.0
Moco+ - Chat, Meet People	4.2	1546.0	10000.0
H*nest Meditation	4.9	145.0	5000.0
Q Alerts: QAnon Drop Notifications, Research +++	4.7	143.0	5000.0
myGrow	4.6	84.0	1000.0

Annexure 7 – Top 10 Paid Apps – Preferred by "Everyone 10+" Segment

App	Rating	Reviews	Installs
Minecraft	4.5	2375336.0	10000000.0
Stickman Legends: Shadow Wars	4.4	38419.0	1000000.0
Card Wars - Adventure Time	4.3	129603.0	1000000.0
Infinity Dungeon VIP	4.3	21804.0	1000000.0
Hitman GO	4.6	84114.0	500000.0
Cytus II	4.7	16851.0	100000.0
Riptide GP: Renegade	4.7	4328.0	100000.0
Teeny Titans - Teen Titans Go!	4.6	20463.0	100000.0
Clue	4.6	19922.0	100000.0
Guns'n'Glory WW2 Premium	4.5	4656.0	100000.0

Annexure 8 – Top 10 Paid Apps – Preferred by "Everyone" Segment

App	Rating	Reviews	Installs
Where's My Water?	4.7	188740.0	1000000.0
Bloons TD 5	4.6	190086.0	1000000.0
DraStic DS Emulator	4.6	87766.0	1000000.0
Cut the Rope GOLD	4.6	61264.0	1000000.0
Tasker	4.6	43045.0	1000000.0
Threema	4.5	51110.0	1000000.0
Lost Journey (Dreamsky)	4.5	32344.0	1000000.0
Sleep as Android Unlock	4.5	23966.0	1000000.0
True Skate	4.4	129409.0	1000000.0
Facetune - For Free	4.4	49553.0	1000000.0

Identify Most Preferred App Size – for Free Apps

INFERENCE -

1. As per stat given beside, Avg. App Size (mostly accepted size) preferred by people for "Free Apps" is 29 MB

2. Median size of the "Free App" is 12 MB (more 50% apps have 12+ MB App Size)

3. We can consider 3rd Quartile size value as well to have more precise App size range

4. But Free apps does not provide that high quality content or additional customized features, hence no need to consider 3rd quartile size as higher side of App size range

From above 2 observations, we can conclude that to make a Successful and most acceptable Free App, it's size should be in the range of 12 to 30 MB

```
count         784.000000
mean        29818.216837
std         26343.315824
min          1228.000000
25%         12288.000000
50%         12288.000000
75%         46080.000000
max        102400.000000
Name: Size, dtype: float64
```

Identify Most Preferred App Size – for Paid Apps

INFERENCE :

1. As per stat given beside, Avg. App Size (mostly accepted size) preferred by people for "Paid Apps" is 30 MB

2. Median size of the "Paid App" is 21 MB (more 50% apps have 21+ MB App Size)

3. Paid Apps consist of higher quality and features in the App. Due to paid model customer expect high quality and more features in Paid Apps. Due to these features and quality, App size is on higher side as compared to Free Apps.

4. As per above observations, successful paid app size should be between range of 23 to 30 MB

```
count        73.000000
mean      30489.452055
std       27352.354761
min          50.000000
25%       12288.000000
50%       21504.000000
75%       44032.000000
max      102400.000000
Name: Size, dtype: float64
```

1. But because of additional features and quality, we need to assume higher side of Paid App size range.

2. Here we should consider 3rd quartile size of paid App, which is 43 MB to recommend appropriate most preferred size to make a successful Paid App, with desired features and quality

3. From above observations, we can conclude that to make a Successful and most acceptable Paid App, it's size should be in the range of 23 to 45 MB

Content rating wise preferred App Size

AI

INFERENCE – Free Apps

1. As per table 1 – Everyone 10+ category prefer heavy apps having size up to 52 MB (Free Apps)
2. Teen category prefer app size from up to 39 MB
3. Mature 17+ and Everyone category prefer app size from 25 to 28 MB
4. We can refer these ranges for more precise App size Content Rating Category wise

Content_Rating	Size
Everyone 10+	51507.200000
Teen	38573.741935
Mature 17+	28407.444444
Everyone	25415.329650

INFERENCE – Paid Apps

1. As per table 1 – Teen & Everyone 10+ category prefer up to 44 MB Free Apps
2. Mature 17+ and Everyone category prefer app size from 24 to 28 MB
3. We can refer these ranges for more precise App size Content Rating Category wise

Content_Rating	Size
Teen	44609.714286
Everyone 10+	43904.000000
Mature 17+	28160.000000
Everyone	24360.040816

Top 5 – Heavy Apps

AI

INFERENCE –

1. Most heavy app falls under Parenting category. This is due to huge images, videos and chat data.
2. Next heavy category is Family and Games. Due video streaming on both apps, these apps are heavy
3. Next comes Finance & Sports, due to high speed data transfer and video streaming these apps are getting heavy
4. If App makers are planning to build any App in this category, they can assume our Free App range up to 45 MB which was 3rd quartile of our size data.

Category	Size
PARENTING	63488.000000
FAMILY	47610.154472
GAME	46954.540541
FINANCE	42170.181818
SPORTS	34443.636364

INFERENCE –

1. Most heavy app falls under Sports category. This is due to high streaming videos with HD quality, because these apps are paid
2. Next heavy category is Family and Games. Due video streaming on both apps, these apps are heavy
3. Next comes Business & Photography, due to high speed data transfer and image processing as well as storage these apps are getting heavy
4. If App makers are planning to build any App in this category, they can assume our Free App range up to 43 MB which was 3rd quartile of our size data.

Category	Size
SPORTS	74752.000000
GAME	48038.375000
FAMILY	30760.900000
BUSINESS	23449.500000
PHOTOGRAPHY	20684.333333

Free Apps

Correlation of quantitative Variables

INFERENCE -

1. For Free App category, heatmap given beside shows strong positive correlation between Installs & Reviews i.e. 0.78

2. Size & Installs and Rating & Installs shows significant negative correlation

3. Lowest negative correlation is between rating and Size but it's significant i.e. -0.36

4. Need to further investigate these relations to understand exact trend of these variables and its long term impact on App performance

Paid Apps

Correlation of quantitative Variables

INFERENCE -

1. For Paid App category, heatmap given beside shows strong positive correlation between Install & Reviews i.e. 0.96

2. Price, Size & rating shows negative correlation with respect to all variables

3. Lowest negative correlation is between Size & rating and Size and installs, they must be stable variables for paid apps

4. Let's further investigate these relations to understand exact trend of these variables and its long term impact on App performance
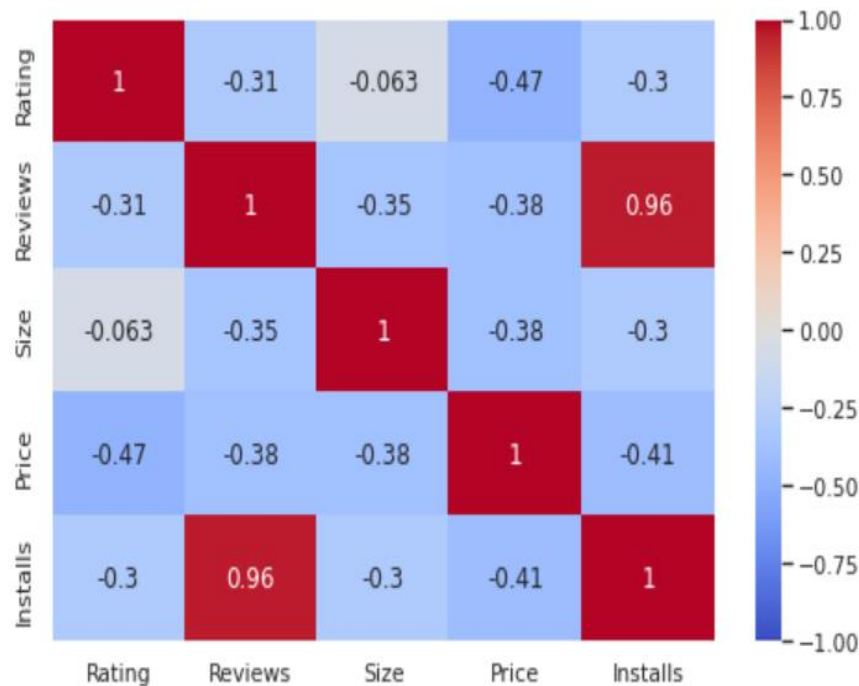
Some IMP Q & A

Annexure 1 – Top 10 - Free

AI

Top 10 Genres

	Genres
0	Communication
1	Travel & Local
2	Productivity
3	Social
4	Education;Education
5	Adventure;Action & Adventure
6	News & Magazines
7	Tools
8	Video Players & Editors
9	Casual

Top 8 categories – Within Top 10 Genres

	Category
0	COMMUNICATION
1	TRAVEL_AND_LOCAL
2	PRODUCTIVITY
3	SOCIAL
4	FAMILY
5	NEWS_AND_MAGAZINES
6	TOOLS
7	VIDEO_PLAYERS
8	GAME

Top 9 Apps – Within Top 8 Categories

	App
0	Instagram
1	WhatsApp Messenger
2	Google
3	Google Drive
4	YouTube
5	Google Chrome: Fast & Secure
6	Maps - Navigate & Explore
7	Gmail
8	Google+
9	Google Street View

Annexure 2 – Top 10 - Paid

AI

Top 20 Genres

Genres
Arcade;Action & Adventure
Action
Communication
Lifestyle
Sports
Card;Action & Adventure
Arcade
Strategy
Puzzle;Brain Games
Photography
Education;Pretend Play
Sports;Action & Adventure
Adventure
Personalization
Tools
Productivity
Puzzle
Role Playing
Weather
Business

Top 10 categories – Within Top 20 Genres

Category
GAME
COMMUNICATION
LIFESTYLE
SPORTS
FAMILY
PHOTOGRAPHY
PERSONALIZATION
TOOLS
PRODUCTIVITY
WEATHER

Top 10 Apps – Within Top 10 Categories

App
Hitman Sniper
Minecraft
Where's My Water?
Bloons TD 5
Five Nights at Freddy's
DraStic DS Emulator
Cut the Rope GOLD
Tasker
Threema
Lost Journey (Dreamsky)

Impact of Updates on Size

INFERENCE –

1.In Free apps, as graph shows the Size of the App is increasing with the new updates in Apps

2.New updates helps to enhance new features in app, quality of the visuals and processing speed of the data, due to these features Apps must be getting heavy
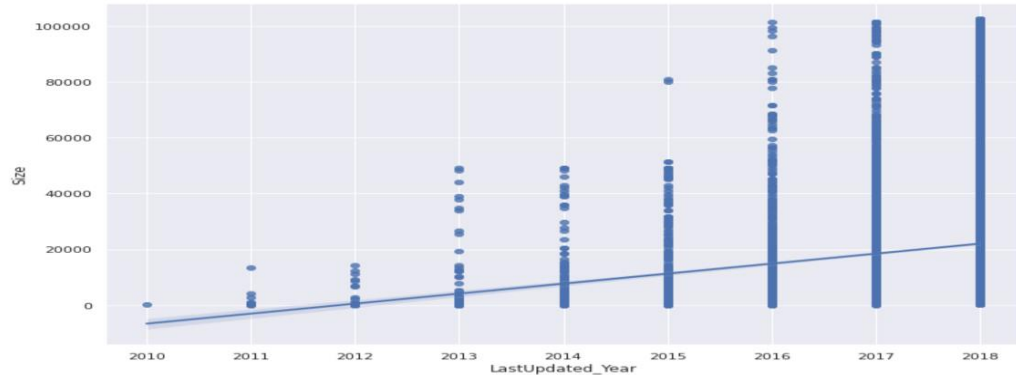
FREE App Chart

INFERENCE –

1.In Paid apps also, as graph shows the Size of the App is increasing with the new updates in Apps

2.New updates helps to enhance new features in app, quality of the visuals and processing speed of the data, due to these features Apps must be getting heavy
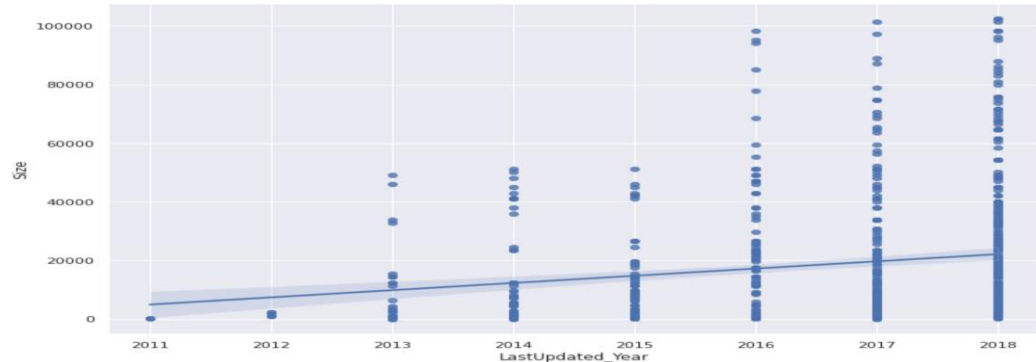
PAID App Chart

Price Vs Size

INFERENCE –

1. Here we can see that Most of the apps having app size up to 30 MB have been saturated between range of Rs 10 – Rs 20 subscription cost

2. Here App makers can be advised, if they are going for Paid App Development they need to keep the price of App should be between 10 – 20 Rs to attract more customers and enhance their App business.

3. If App cost is high App makers can explore another channels of revenue through same app like in-app purchases etc.
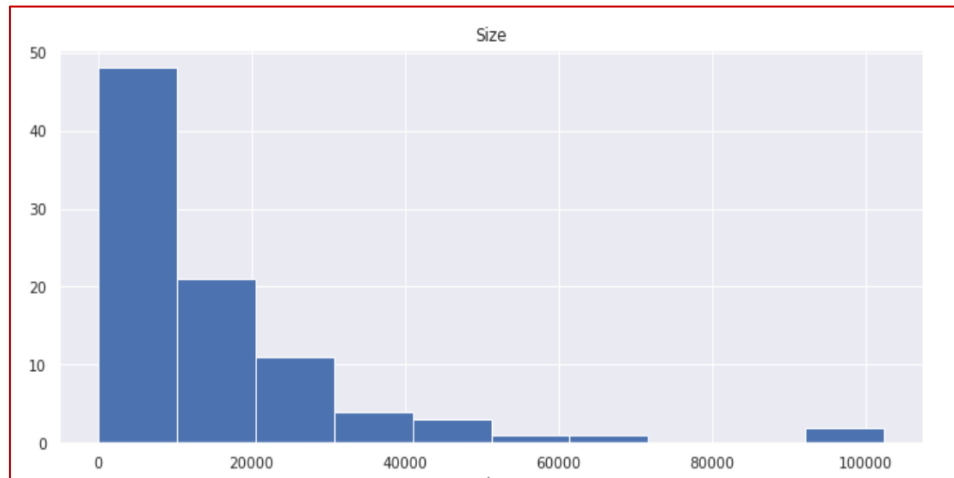
Avg. Size of Paid App Categories –

Top 10 Genres – Avg. App Size is – 28 MB
Top 10 Categories – Avg. App Size is – 24 MB
Top 10 Apps – Avg. App Size is – 37 MB

Price Trend Analysis

INFERNCE –

1. As per price trend reflecting in chart given beside, most of the Apps in all the categories are trying to maintaining their price in the range of 10 - 20 Rs.

2. And as we have seen earlier, it is the most acceptable price among customers for Paid Apps

3. As seen earlier, If App cost is high due to content quality / image / video quality / storage / any other features, App makers can explore another channels of revenue through same app like in-app purchases, different temporary subscription schemes etc.

App Pricing Trend across Categories
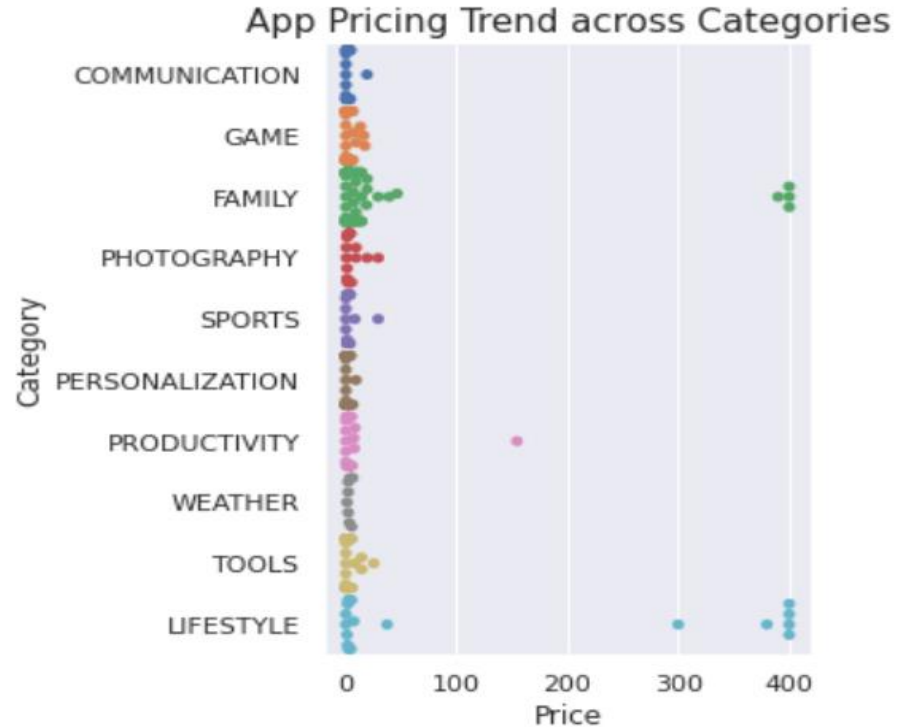
Top 5 – Android Versions

Below top 5 - Android versions are most accepted android versions by customers
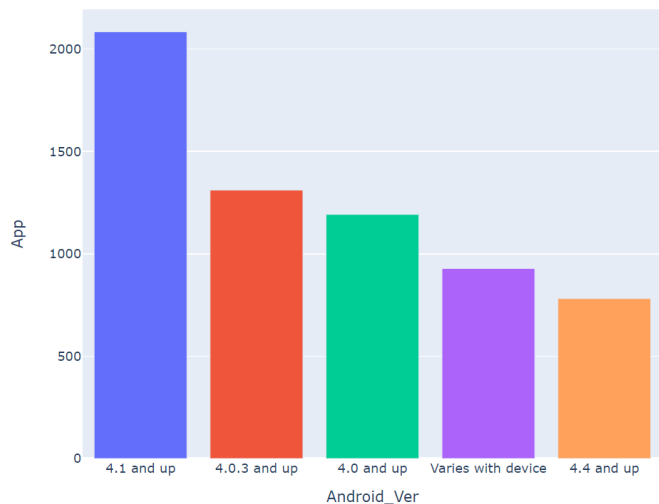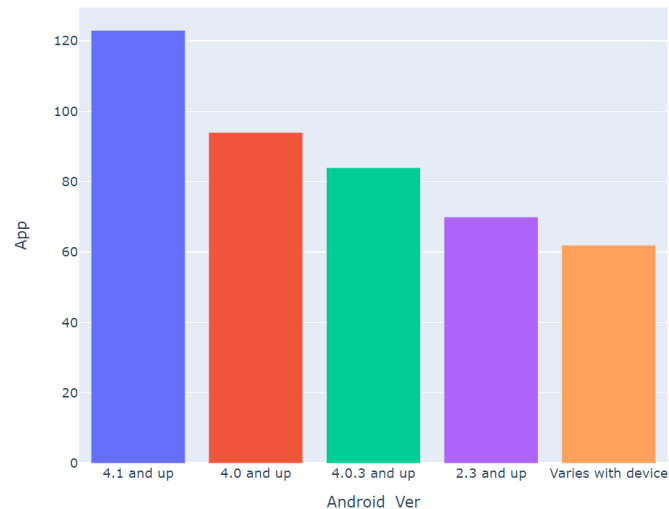
FREE App Chart

PAID App Chart

Top 5 – Current Versions

Below top 5 - Android versions are most accepted android versions by customers
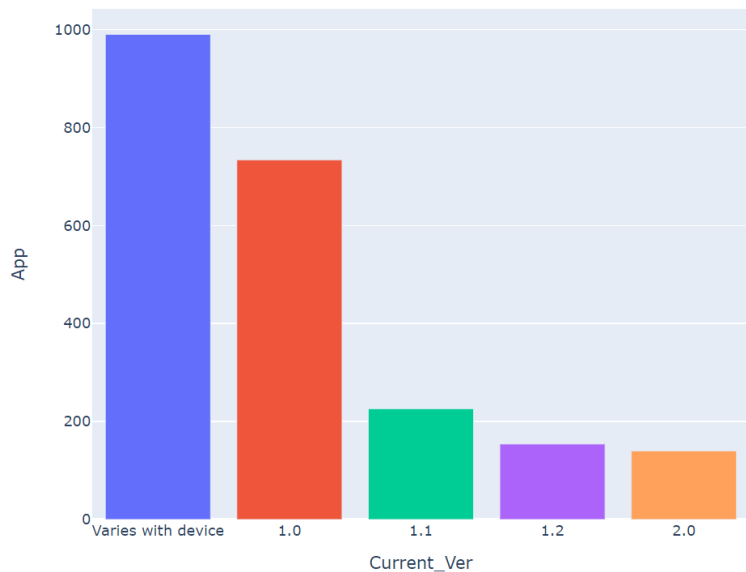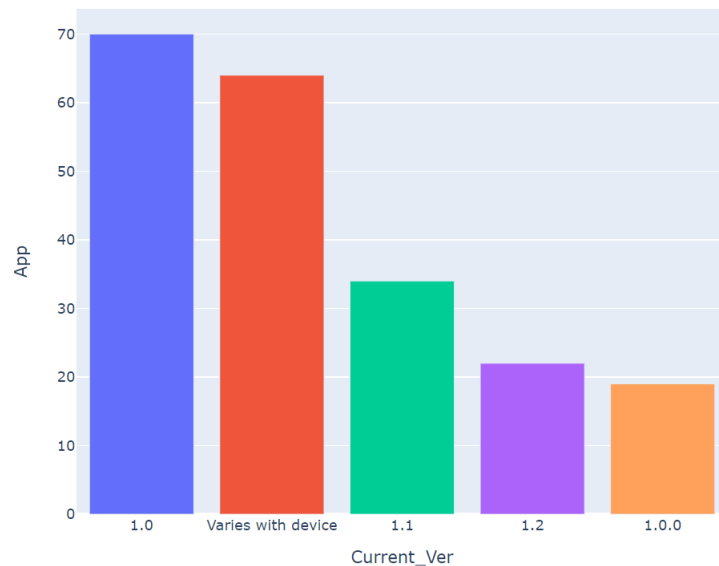
FREE App Chart

PAID App Chart

Current Version Analysis

INFERENCE –

In Paid Apps, as per chart below, we can say that, No relation can be established between Current version and Size of the App.
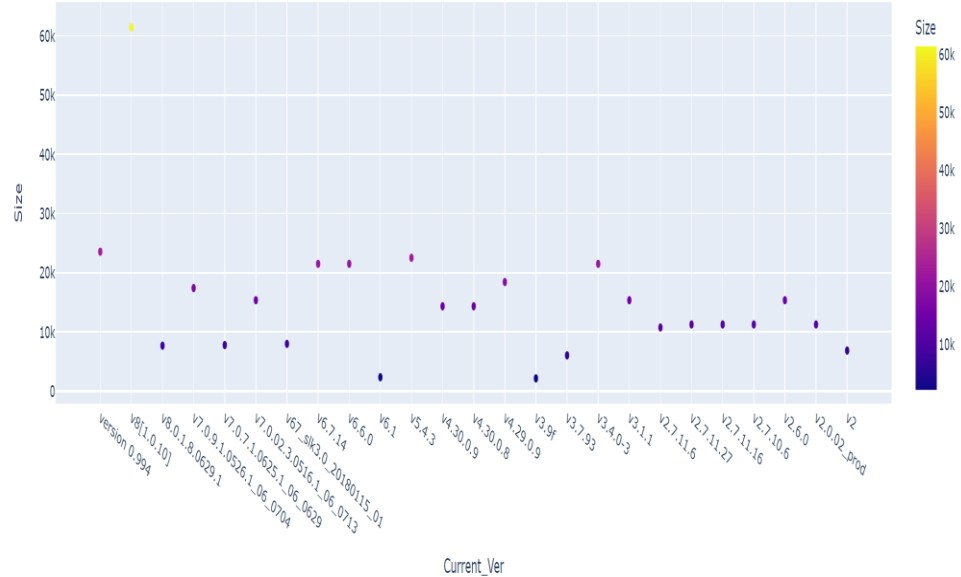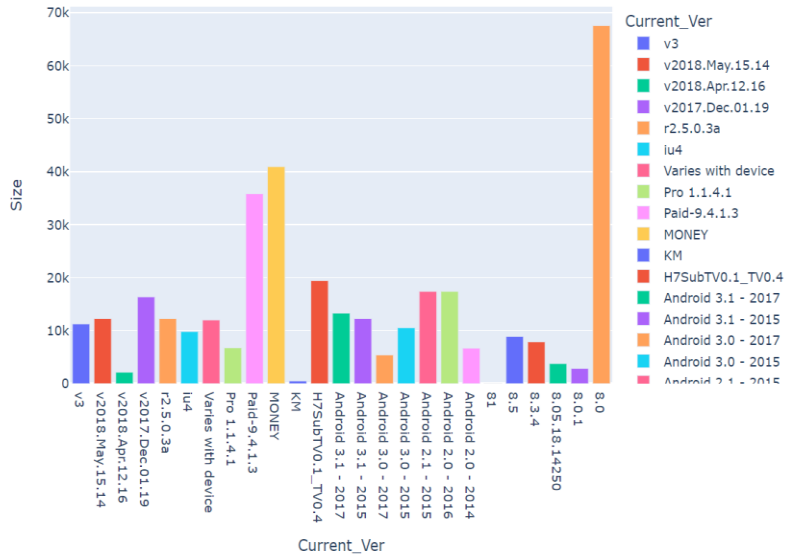
INFERENCE –

In Free Apps also, as per chart above, we can say that, No relation can be established between Current version and Size of the App.

Android Version Vs Rating

INFERENCE –

In Paid Apps, as per chart below, Android versions of Series 4.0 to Series 4.4 are showing good ratings and customers positive acceptance
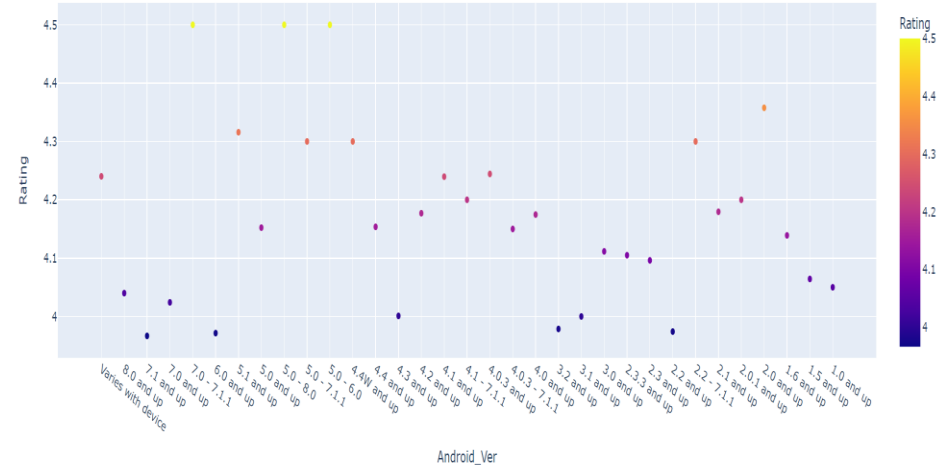
INFERENCE –

In Free Apps, as per chart above, Android versions of Series 4.0 and above are showing good ratings and customers positive acceptance

Android Version Vs Reviews

INFERENCE –

In Paid Apps also, Android versions of Series 4.0 to Series 4.4 and Series 2.3 to Series 3.0 are showing highest mean reviews.
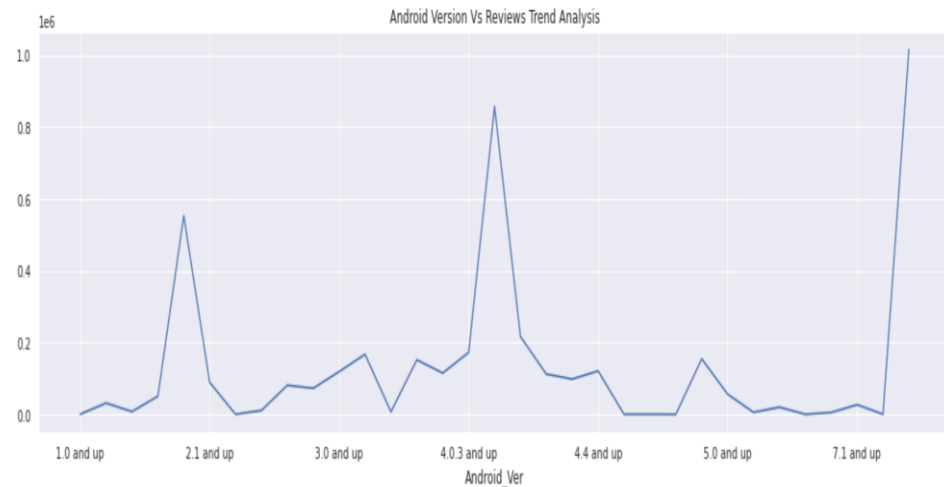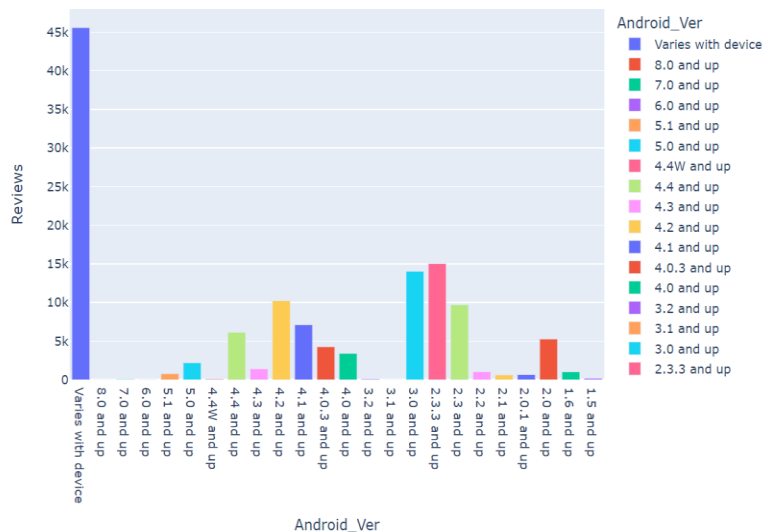
INFERENCE –

In Free Apps, Android versions of Series 4.0 to Series 4.4 are showing highest mean reviews.

Android Version Vs Size

INFERENCE –

In Paid Apps also, we can say Size increment trend with latest Android versions.

Size of app increasing with latest Android Versions
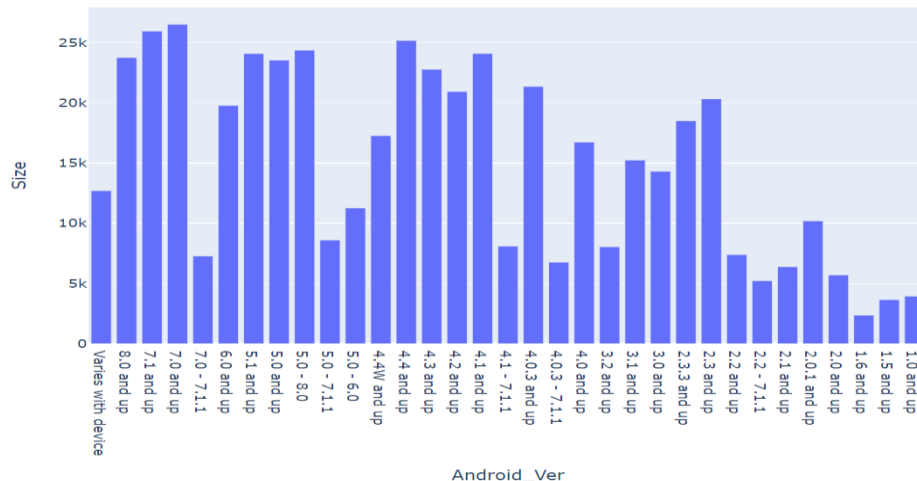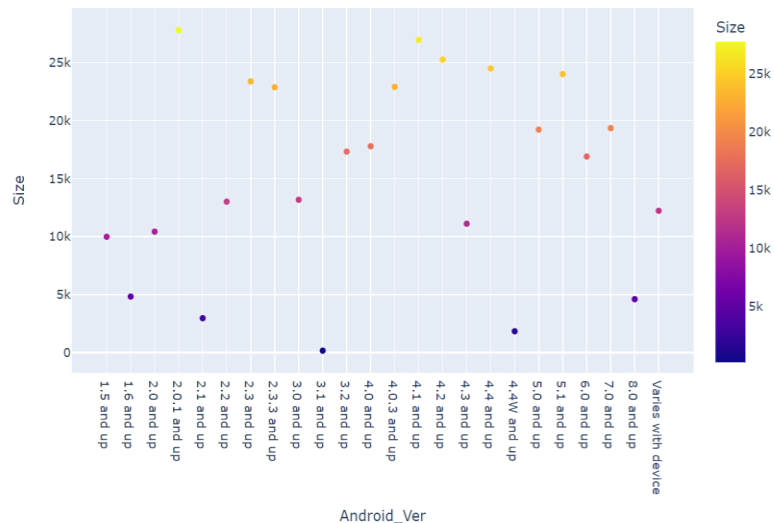
INFERENCE –

In Free Apps, as per above graph, we can easily make out the trend of increase in size with latest Android versions.

Size of app increasing with latest Android Versions

Android Version Vs Price

Paid App – Android Version Vs Price

INFERENCE –

1. In Paid Apps, Android Version Series of 4.0 to 4.4 are mostly accepted version

2. As we can check in graph most of the prices are concentrated around Android version 4.0 to 4.4

3. And these Android versions are having average subscription cost of around Rs 10 - 20/-
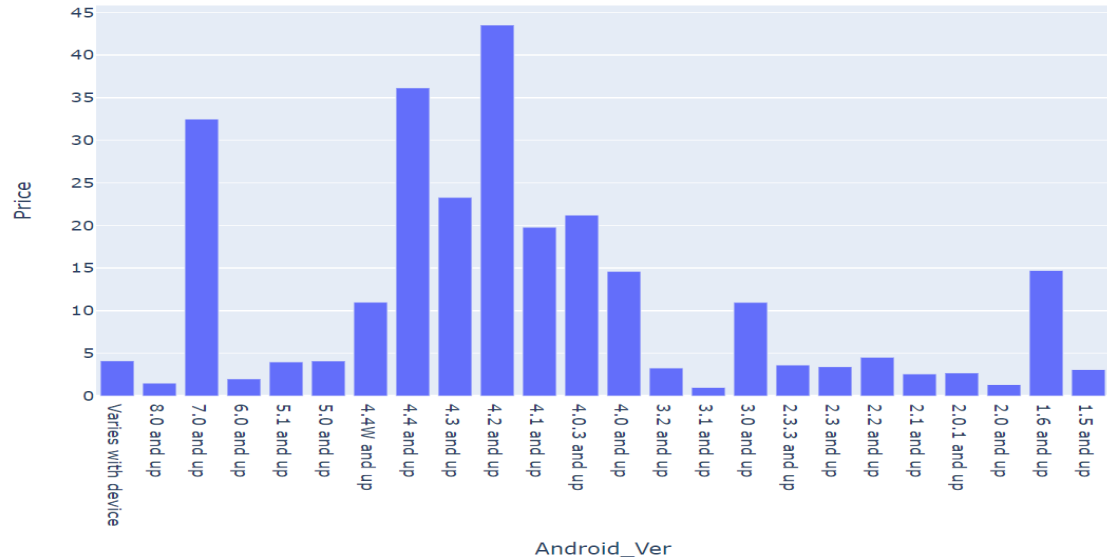
Reviews Vs Rating

INFERENCE –

As per below chart, in successful Free Apps, Ratings are showing positive incremental trends with increasing reviews.

This might be due to people reacting positively towards services provided by the Apps in terms of Technology, Size, quality etc.

And these positive reviews are driving rating higher with increase in such positive feedbacks.
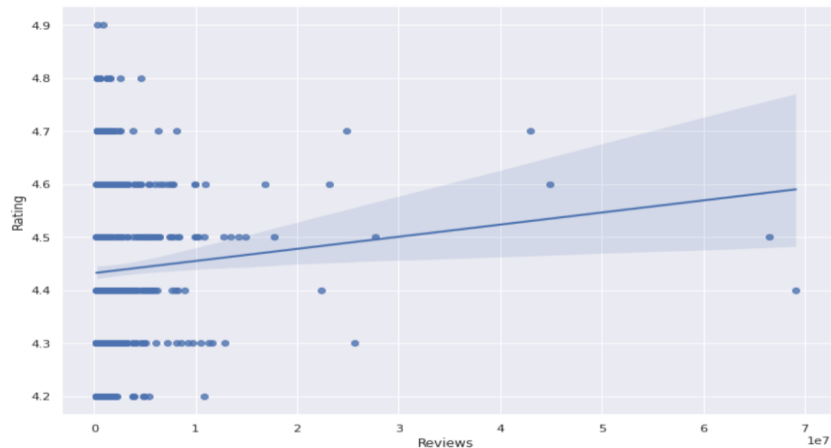
INFERENCE –

In Successful Paid Apps, Ratings are showing slightly negative incremental trends with increasing reviews.

This might be due to people who are paying for App, after checking earlier Ratings, expects more when they subscribe paid apps.

If they get slight dissatisfaction their views are getting slightly negative and hence maintaining or declining their feedback.
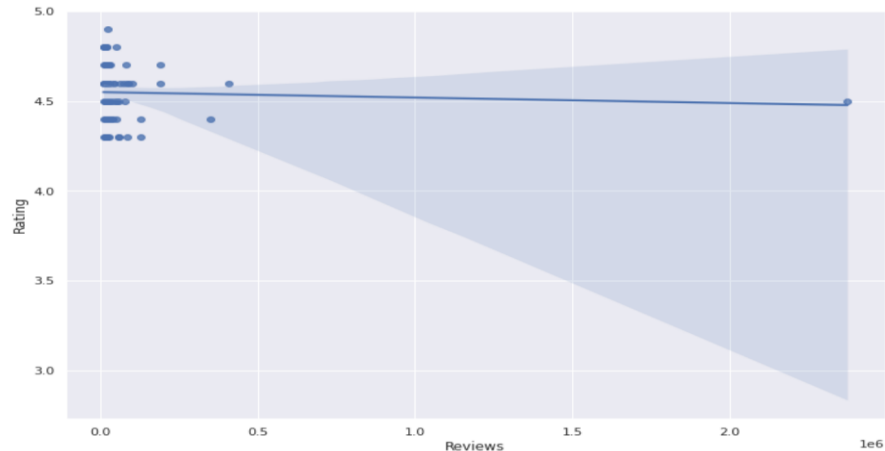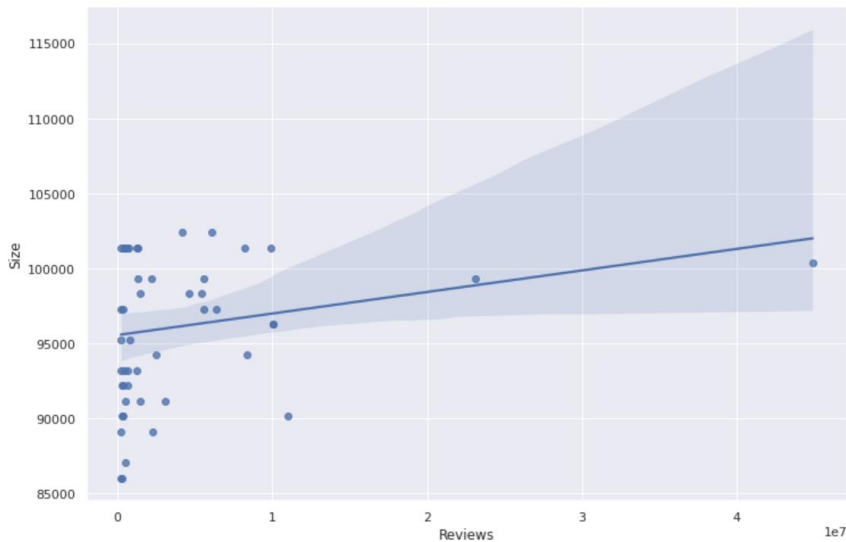
Reviews Vs Size

INFERENCE −

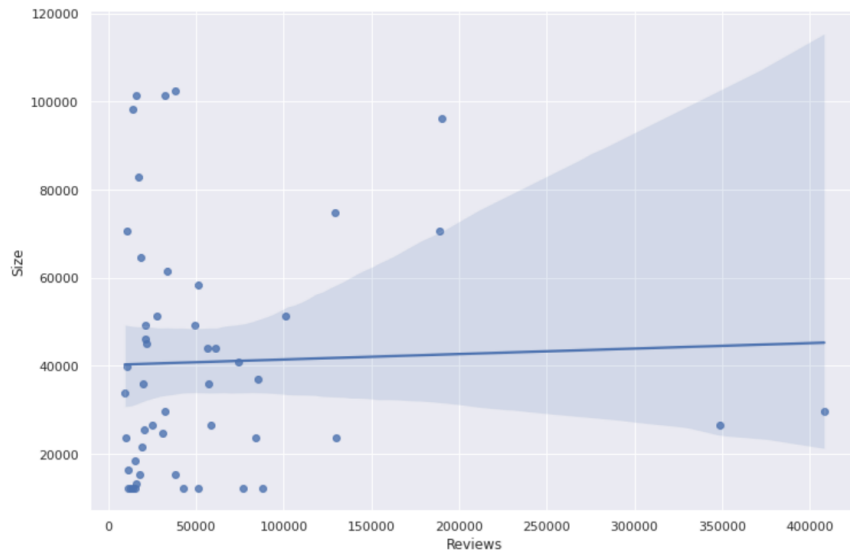In Free Apps, with increase in Size reviews are showing positive incremental trend.

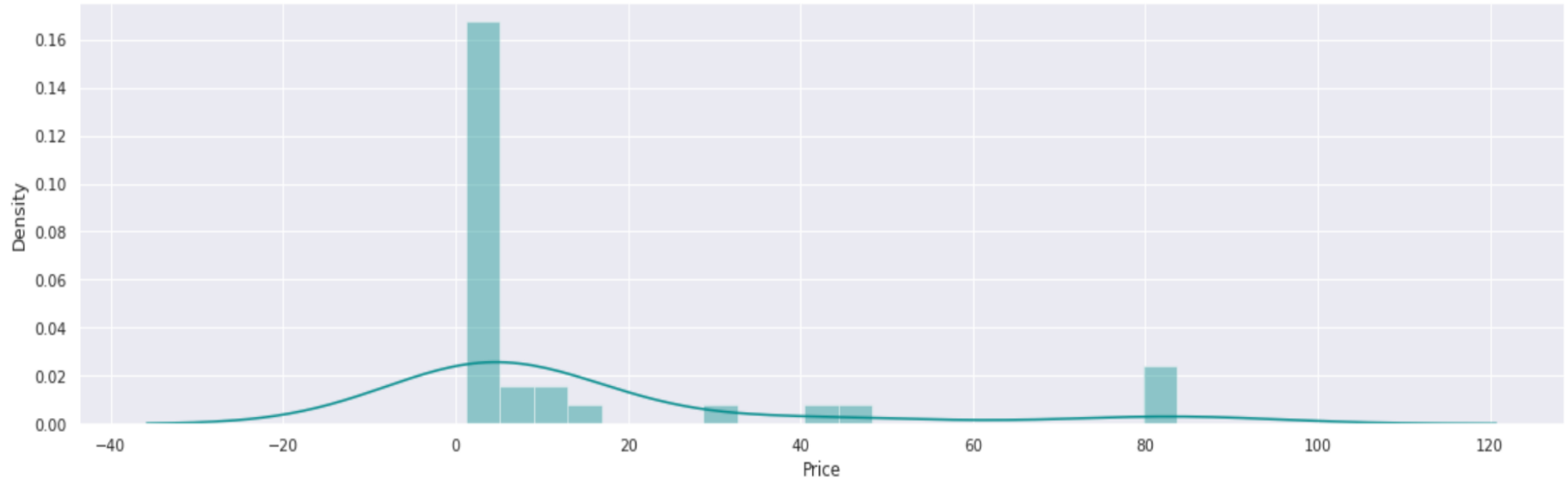Means, people might be getting better service after increase in size of the App.

INFERENCE −

In Paid Apps, with increase in Size of the app, reviews are not getting changed, they are stable.

Hence, Paid app customers are not that sensitive towards increase in the size of the App.
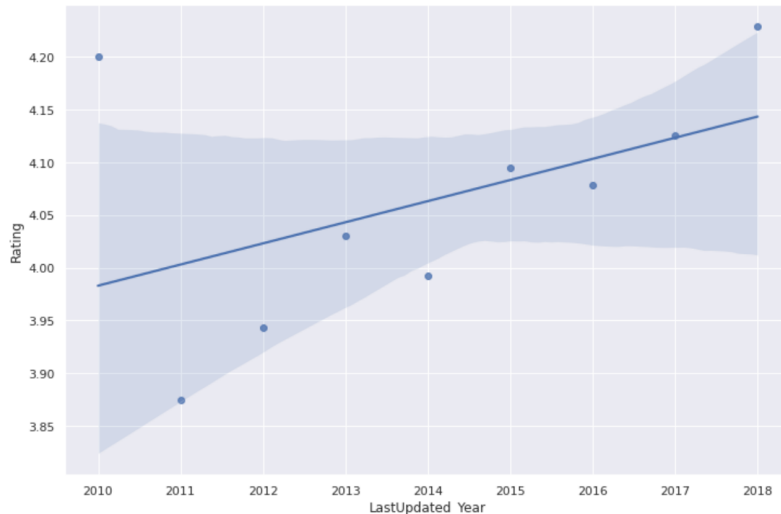
Price Vs Rating – Paid Apps

INFERENCE – People have highest acceptance of price, in the range of Rs 10 – 20/-

Last Update Vs Rating

INFERENCE –

In Free Apps, with Latest updates of Apps their rating is also increasing.
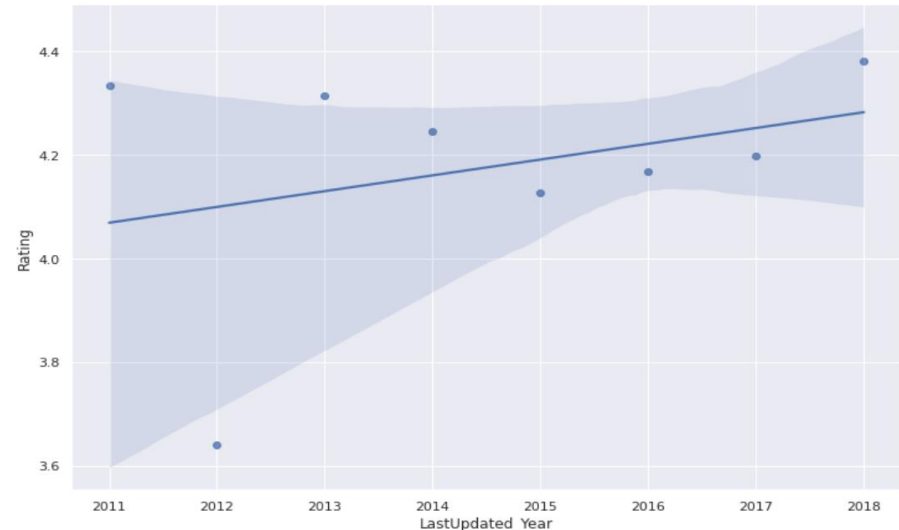Means, people might be getting better service after getting latest updates in the App and they are positive towards such updates.

INFERENCE –

In Paid Apps, with Latest updates of Apps their rating is also increasing.

Means, people might be getting better service after getting latest updates in the App and they are positive towards such updates.

Sentiment – Correlation Analysis

INFERENCE :

1.Here we can see strong correlation of Installs & Rating.

2.Other parameters are not showing any significant correlation with each other.

3.As per some normal incremental correlation coefficient no. like 0.11 (Last Update & Rating), 0.18 (Last Update & Size) etc. may show us some slight trends over a period.

4.We will analyse the same in detail to understand their impact on App performance.

Sentiment Subjectivity Distribution
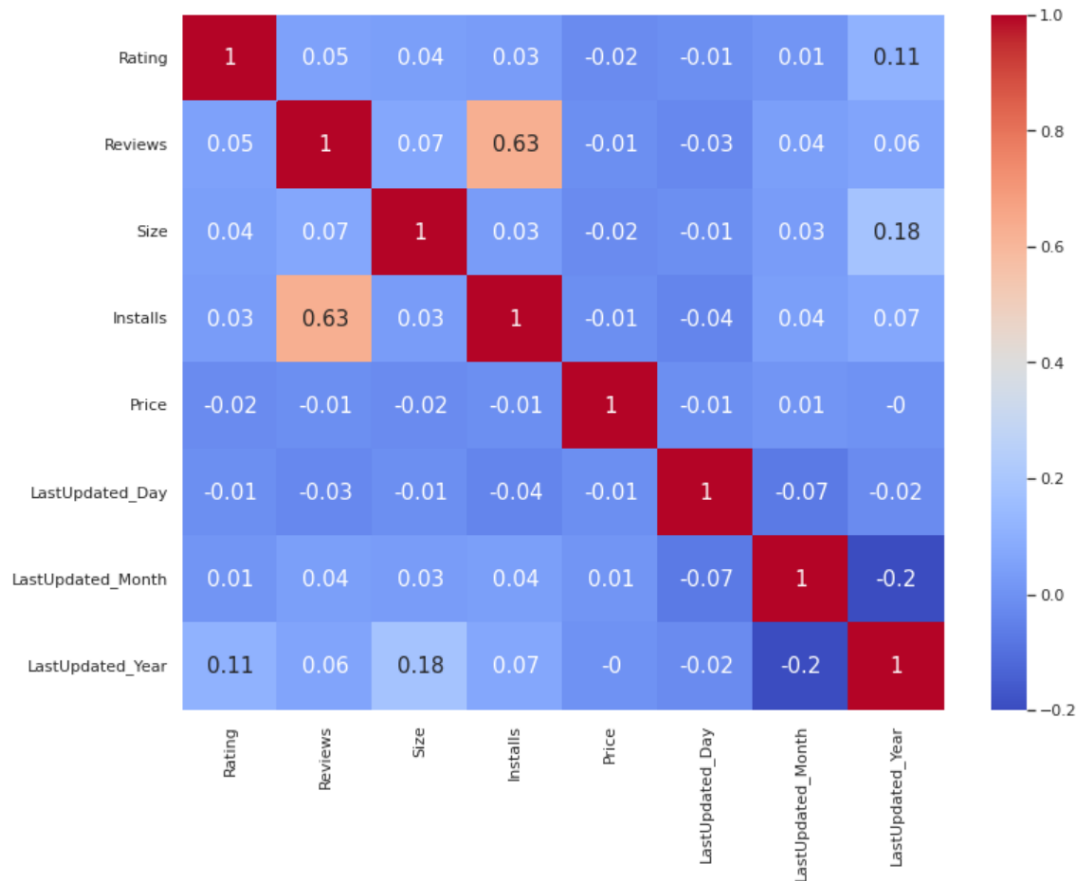
Sentiment Subjectivity Distribution
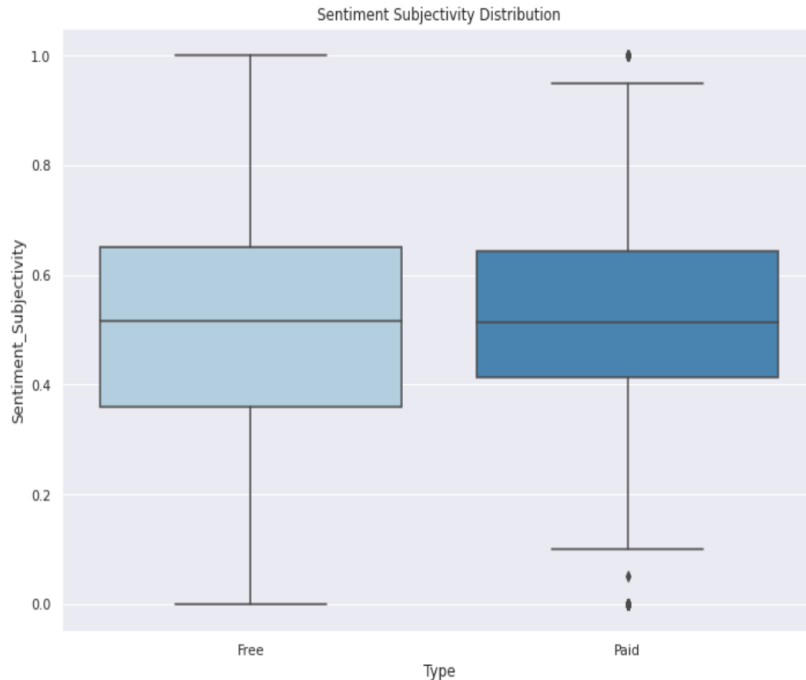
INFERENCE –

1. Here this Sentiment Subjectivity Box Plot also giving us same sentiment trend like polarity

2. Here also, **we can confirm that customers sentiment subjectivity i.e. intensity of emotions while expressing their views is stable and better in terms of Paid Apps.** That might be due to Paid Apps are providing good services than the Free Apps.

3. We can also see that Free Apps sentiment subjectivity is slightly negatively skewed as compared to Paid Apps

Sentiment Polarity Distribution

INFERENCE –

1.Here this Box Plot gives us five important numbers related to relation of Sentiment Polarity in Free & Paid Apps

2.Free apps have higher sentiment polarity parameters with respect to Paid Apps (Median, Max, Inter Quartile Ranges etc.)

3.Here we understand that customers sentiment are stable and better in terms of Paid Apps. That might be due to Paid Apps are providing good services than the Free Apps

4.Here we can see that Free Apps sentiment polarity is slightly negatively skewed as compared to Paid Apps


Sentiment Polarity Distribution
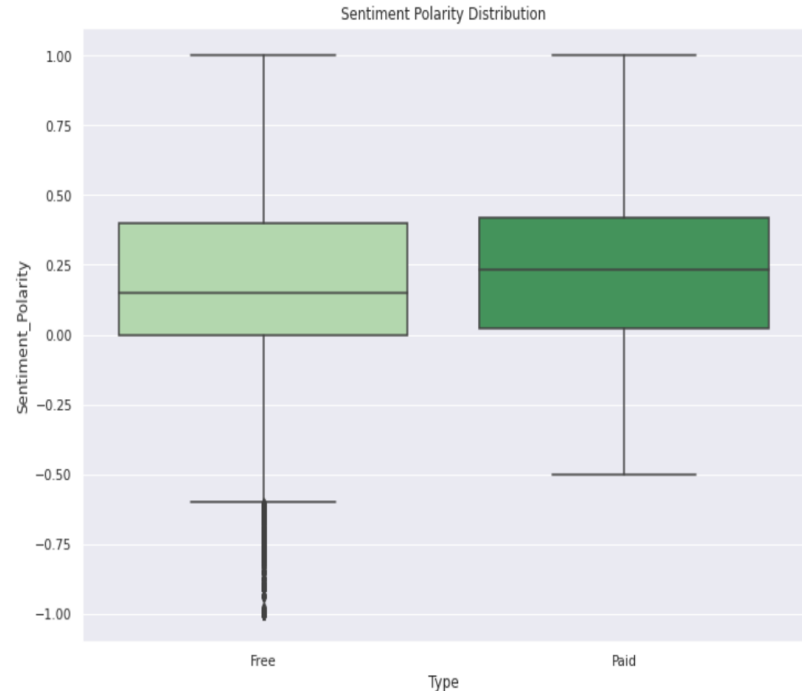
Overall Positivity & Subjectivity
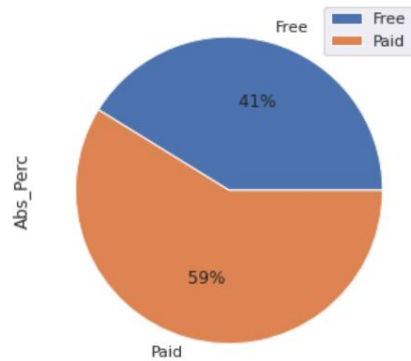
INFERENCE -
1.In this pie chart we can see that as a whole people are positive about Free & Paid both the apps

2.People are slightly conservative while giving their reviews about Apps

3.We can also see that overall positivity of Paid Apps is more that Free Apps. This means people are finding Paid App service better than Free Apps
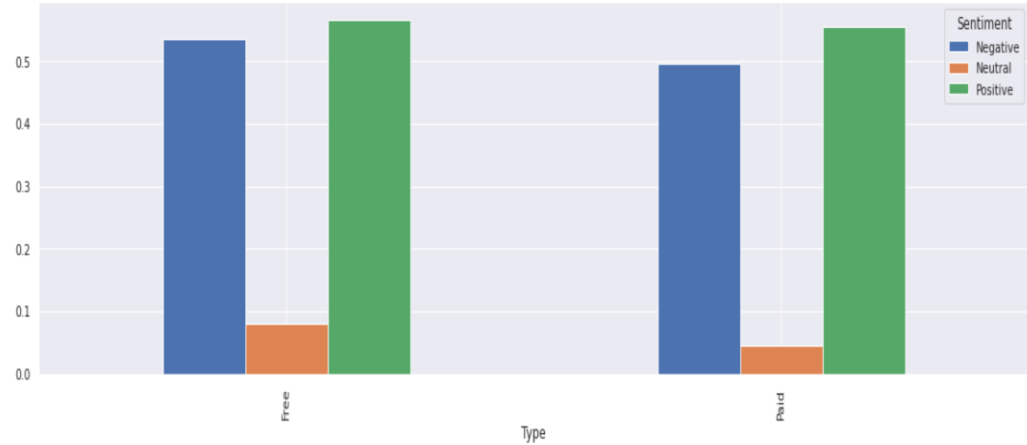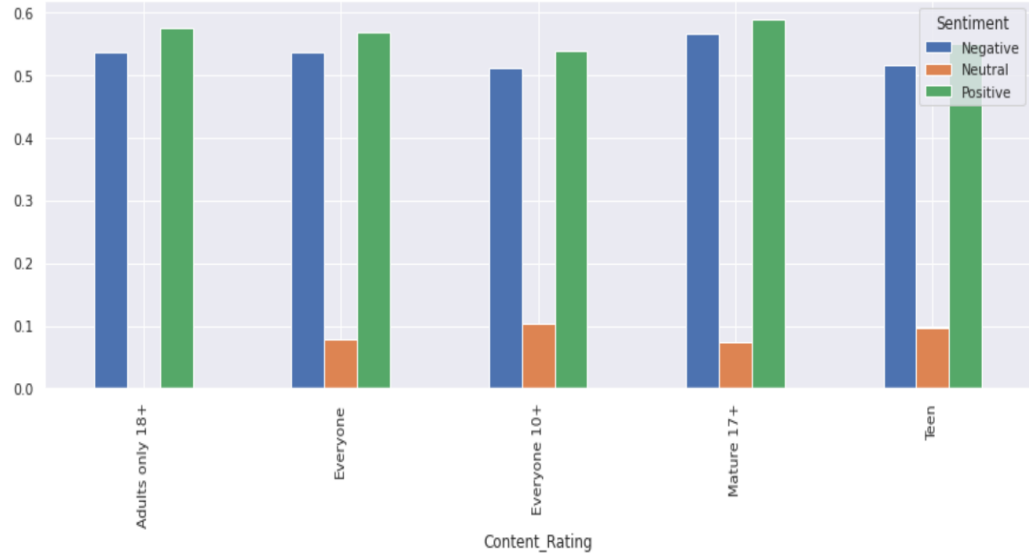
INFERENCE –

1.In above chart we can see positive bar is higher than negative bar, people are overall positive about Free & Paid Apps

2.Very few people are neutral, most of the people are vocal about their experiences

3.Overall Positive subjectivity is more

Content Rating wise Sentiments
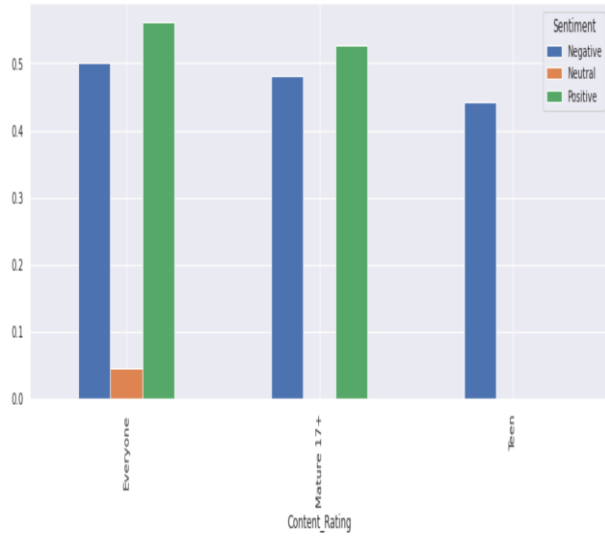
PAID App Plot

FREE App Plot

Some IMP Q & A

Annexure 9 – Top 10 – Positive Sentiment – Free Apps

AI

Top 10 Genres

Genres
Communication
Photography
Tools
Casual
Productivity
Social
Education;Education
Books & Reference
Strategy
Travel & Local

Top 9 categories – Within Top 10 Genres

Category
COMMUNICATION
PHOTOGRAPHY
TOOLS
FAMILY
PRODUCTIVITY
SOCIAL
BOOKS_AND_REFERENCE
TRAVEL_AND_LOCAL
GAME

Top 10 Apps – Within Top 10 Categories

App
Google Photos
Google
Gmail
Google Duo - High Quality Video Calls
Candy Crush Saga
Google Translate
Dropbox
Facebook Lite
Duolingo: Learn Languages Free
Bible

Annexure 10 – Top 10 – Positive Sentiment – Paid Apps

AI

Top 10 Genres

Genres
Photography
Casual
Education;Education
Strategy
Racing
Tools
Productivity
Communication
Weather
Books & Reference

Top 8 categories – Within Top 10 Genres

Category
PHOTOGRAPHY
FAMILY
TOOLS
PRODUCTIVITY
COMMUNICATION
GAME
WEATHER
BOOKS_AND_REFERENCE

Top 10 Apps – Within Top 10 Categories

| App |
| --- |
| Google Photos |
| Google |
| Google Duo - High Quality Video Calls |
| Candy Crush Saga |
| Google Translate |
| Dropbox |
| Duolingo: Learn Languages Free |
| Bible |
| Clash of Clans |
| Clash Royale |

Impact of Size on Sentiments

INFERENCE –

1.For Free Apps, increase in size does not have significant impact on subjectivity

2.But Sentiment Polarity is showing declining trend with increase in size
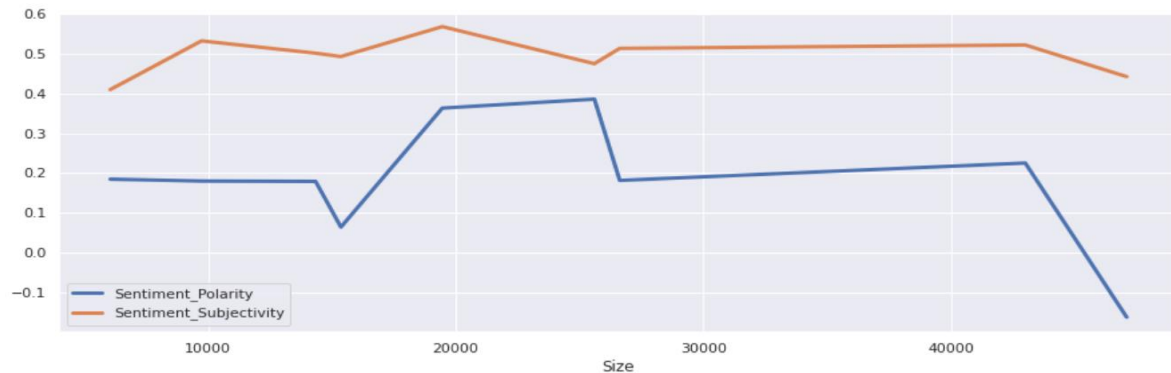
FREE App Chart
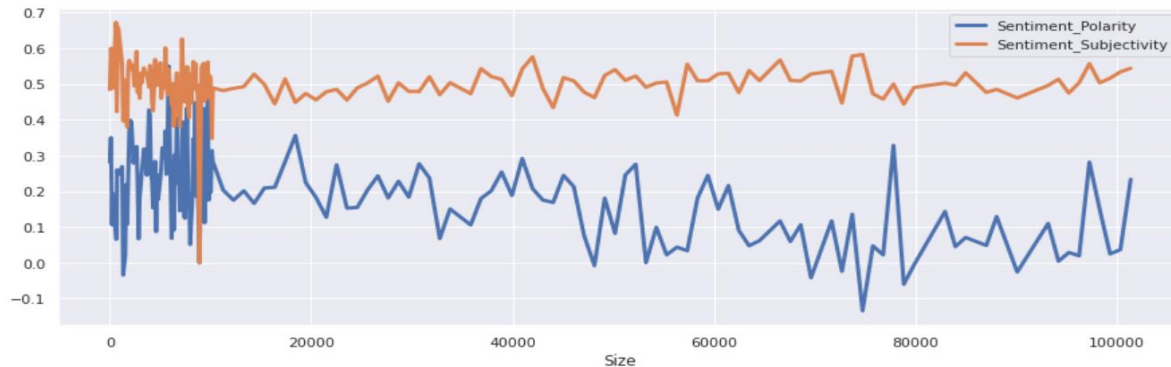
INFERENCE –

1.For Paid Apps also, increase in size does not have significant impact on subjectivity

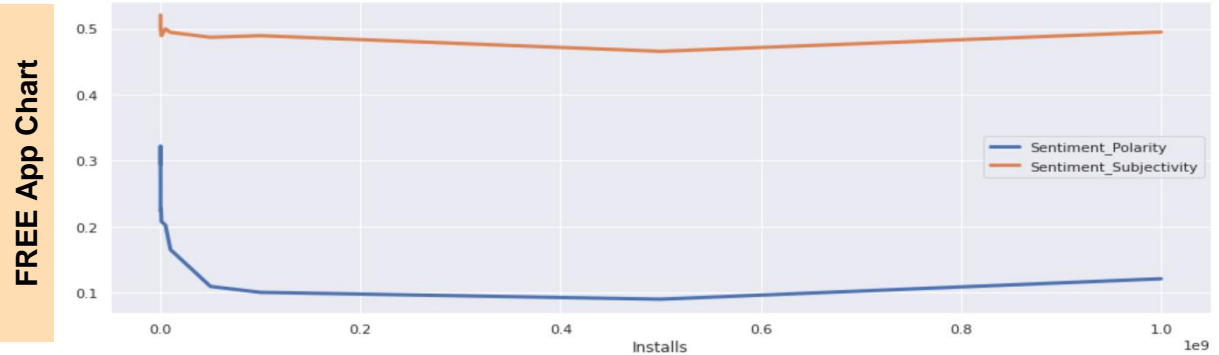2.But Sentiment Polarity is showing declining trend with increase in size

PAID App Chart

# Impact of Installs on Sentiments

INFERENCE –

1.For Free Apps, increasing installs does not show any significant positive / negative trend for both sentiment polarity as well as sentiment subjectivity.

FREE App Chart

INFERENCE –

1.For Paid Apps, increasing installs are showing slight positive trend for sentimen polarity.

2.Subjectivity is stable while increasing installs
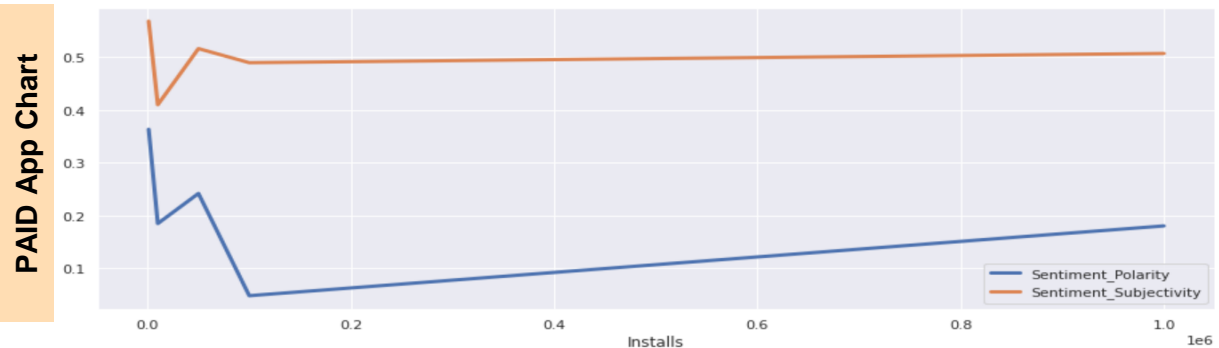
PAID App Chart

Rating Vs Sentiments

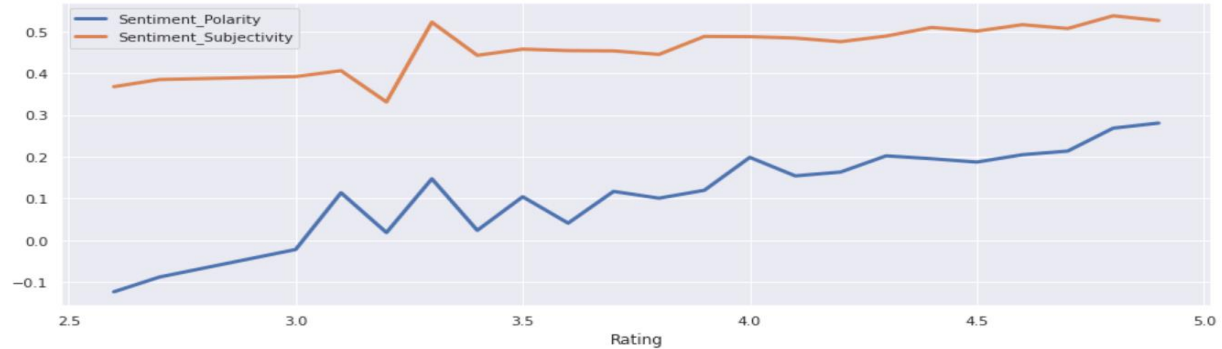**INFERENCE –**

1.For Free Apps, in chart we can see rating is highly correlated with rating.

2.Sentiment polarity shows higher growth rate than sentiment subjectivity

FREE App Chart

INFERENCE –

1.For Paid Apps also, increasing rating are showing positive trend for sentiment polarity & subjectivity both.

2.Rate of polarity increase is higher than subjectivity

PAID App Chart
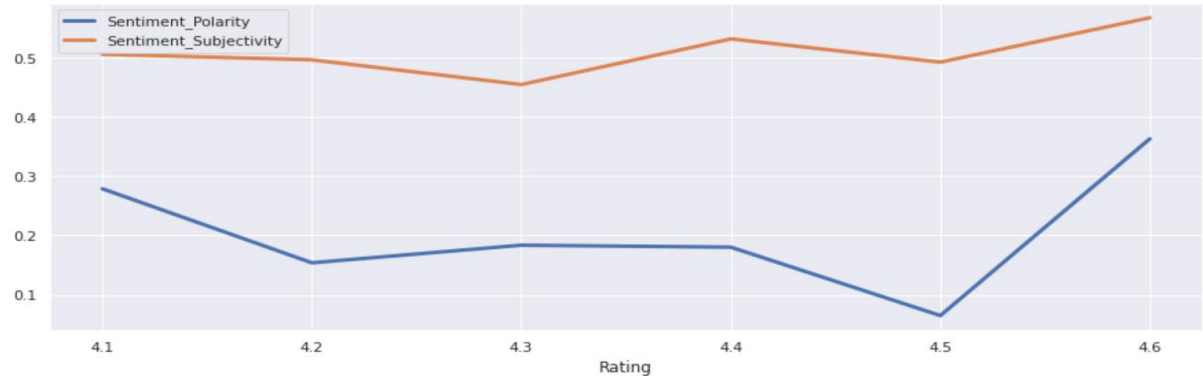
Impact of Update on Sentiments

INFERENCE –

1. For Paid Apps, recent last updates are showing positive trend for sentiment polarity.

2. Sentiment Subjectivity looks stable

PAID App Chart
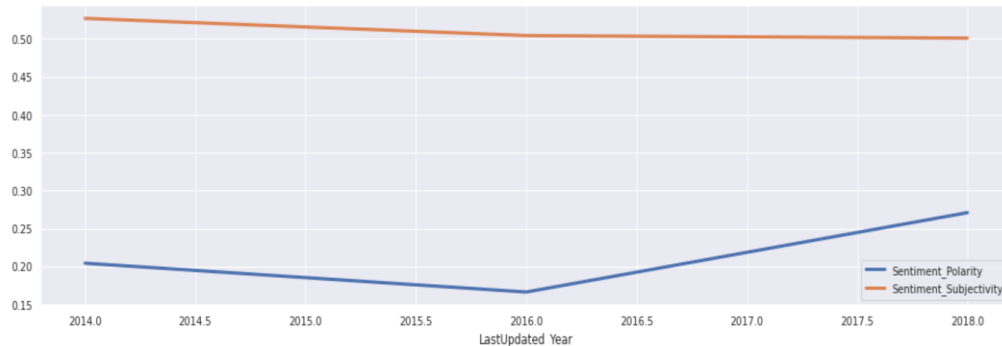
INFERENCE –

1. For Free Apps, recent last updates are not showing significant positive trend for both sentiment polarity as well as Sentiment Subjectivity

2. Polarity & Subjectivity both looks stable

FREE App Chart
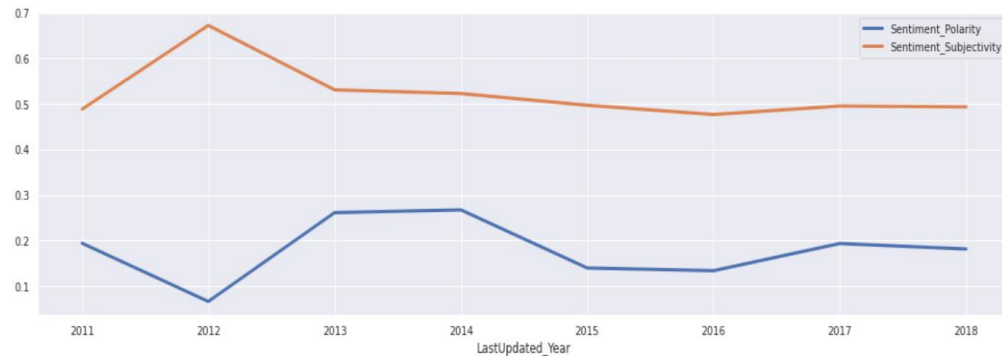
Sentiment Polarity Vs Subjectivity

INFERENCE –

1.Here we can see that, as polarity is going towards two extremes sentiment subjectivity is increasing.

2.People having their views very intense and emotional for highly negative or highly positive reviews

3.Polarity is moving towards high subjectivity to positive as well as negative end.

Word Cloud Analysis – Free Apps

Word Clouds are visual displays of text data – simple text analysis. Word Clouds **display the most prominent or frequent words in a body of text.** We can make out some important information from word cloud and can understand reviewer's sentiments / concerns in brief at certain extent.

INFERENCE –

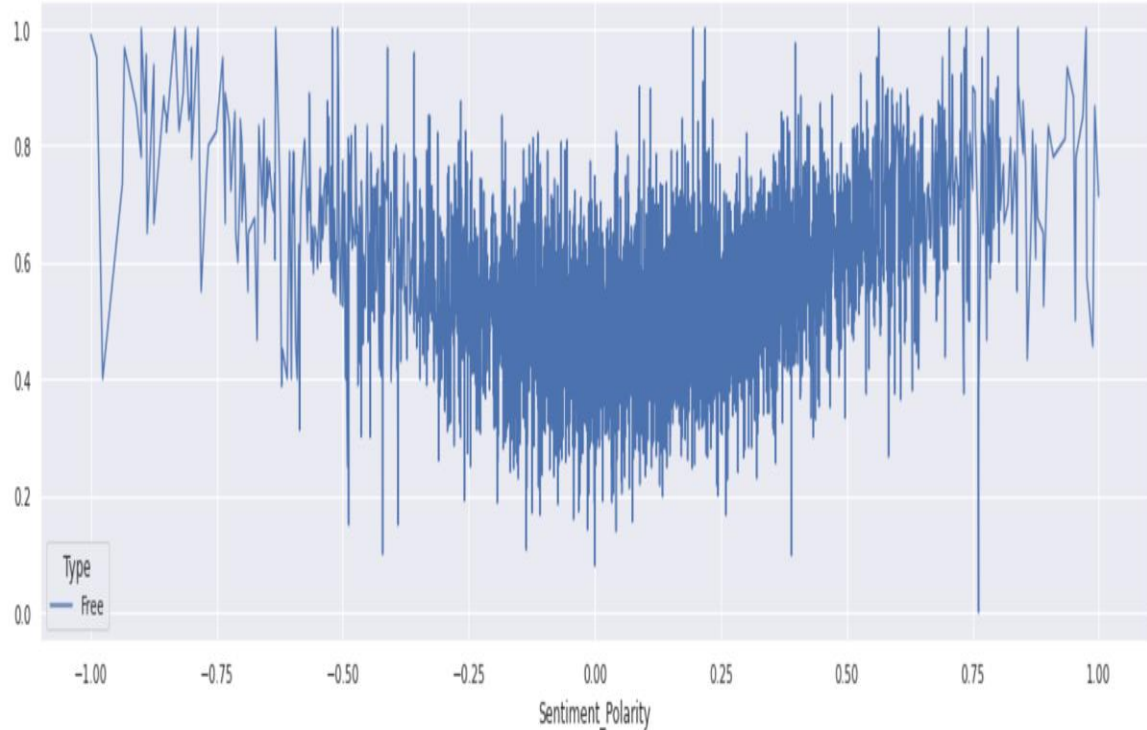1.Game – It's an prominent word here, it means many people who are Using Gaming Apps are writing reviews and expressing their views frequently. (some similar words like level, play etc gives same indication)

2.Great / good / Love / best / nice / thank – All these words are showing positive vibes about reviews and app features.

3.Update / account / work / still – Such words indicating concerns of people. So need to focus on such feedbacks.

4.Start slow / another / note / bad device – These words also represent serious concerns or feedback from customers. Need to look after these reviews as well.

Word Cloud Analysis – Paid Apps

Word Clouds are visual displays of text data – simple text analysis. Word Clouds **display the most prominent or frequent words in a body of text.** We can make out some important information from word cloud and can understand reviewer's sentiments / concerns in brief at certain extent.

INFERENCE –

1.Great / good – These kind of prominent words are showing positive sentiments about paid apps.

2.Game – many paid apps are Gaming apps, hence this word is also prominent here

3.Love – These kinds of word indicates, that there are many dating apps in paid apps type and people are frequently using this word.

4. Diabetes / tablet / calorie – These words indicates that there are many health related apps in paid apps type.

1.Update / Please / Problem / Version / Change / User / Issue - Such words indicating concerns of people. They might be wanted App makers to update their version, fix some bugs etc. So need to focus on such feedbacks.

Conclusion

AI

This Analysis has been concluded in two parts **RECOMMENDATIONS | ACTIONABLE INSIGHTS**

Conclusion & Recommendations (App makers Management Team)

1. In order to start planning or strategizing launch of any New App, you need to finalise a Business Model first. Either Free App model or Paid App Model

2. Every model has their own Pros & Cons (refer slide 17 & 18), hence you need to check the same, that which model suits to your budget, resources and skill sets.

3. If you want to explore both the Business models and want to pursue their cost benefit analysis, then you have to choose an App category which works best as per your capabilities.

4. Here we would like to suggest some important genres and categories which may help you to build and launch a profitable App

5. In Free App category – Most successful categories are as below

- Communication | Travel | Productivity | Social | Family

- Communication category consist of popular messenger apps like Whatsapp

- Social category consist of apps like Instagram

Annexure – 1 of our report will help you with all these details

You can select desired category and app to start your Free App planning accordingly.

Conclusion

6. If you want to initiate a Paid Type of App then you can start with below top Paid App categories

• Gaming Apps | Communication | Lifestyle | Sports | Family

Annexure – 2 of our report will help you with all these details

You can select desired category and app to start your Free App planning

7. Also you can it by your Target Audience category, there are five different customer segments in Free & paid App category,

• Everyone | Everyone 10+ | Teen | Mature 17+ | Adult 18+

In these categories Everyone & Everyone 10+ categories are in top three customer bases of both Free and Paid Apps

You can target these categories because,

8. These 2 categories contributing at the extent of around 66% and 50% respectively to the success of Free and paid apps

9. Mature 17+ category alone contributing @ 50% to the success of paid apps

10. In Annexure 3 to 8, we have given most preferred Apps list by each of these categories separately. You can refer the same for finalizing an app type.

11. We also need to consider a huge competition in theses three categories, more than 90% apps on play store are serving these three categories. Hence competition will be tough. Our App needs to have an uniqueness in it's services to become successful.

Conclusion

12. In Free app category people are preferring an app size of **12 – 30 MB** and in paid App category people are preferring an avg. app size between **23 to 45 MB**

13. You can also use **varies with devise** app versions as people have positive acceptance towards that as well.

14. App needs to get **updated frequently** with **latest technology**

15. Periodical updates affects user experience, reviews and rating of app.

16. Due to additional features / quality of user interface / storage these things are making paid apps slightly heavier than free apps

17. If we go in details content rating wise also we can select appropriate size for our app.

18. In terms of technology people are preferring Android Version apps having version series 4.0 and above. These updated versions are mostly appreciated and liked by people.

19. Certain categories like Sport / family / Game includes video streaming and storage hence these apps are most heavier apps in terms of size on the play store

20. As we investigated correlation of various important factors which affects performance or success of apps, we found below observations

- Size and Ratings in Free App category are negatively correlated with each other.

- Installs and Reviews are strongly correlated with each other. Reviews are mostly positive.

- As a whole and on an average People are showing overall positivity in reviews.

Conclusion

- **Installs and rating** are slightly positively correlated. Slow rate of rating growth with increasing installs.

- These positive **reviews are pushing installations** of app

- But as **installations are increasing**, as people have already gone through positive reviews they expect higher than those reviews.

- And Free apps are not satisfying their expectation in terms of features and hence people are getting slightly negative and showing a **low rate of growth** in rating with increasing installs.

- After an installation **threshold of 0.5 Billion** rating rate slows down, Need to review customer sentiment at this at 0.5 B installs of app.

- In paid apps also we saw a **positive correlation** of Installs and reviews.

21. Hence for success of an app, reviews are very important, they are driving success of the app. You need to develop a mechanism in app which will frequently ask for reviews and **gain maximum reviews.**

22. In paid apps **Size and Price** are adversely affecting apps success.

23. Make sure to have App **pricing between Rs 10 -20/-** Maximum people are tend to pay this amount on an average. Because of this you can gain maximum customers and to generate more revenue you can explore another methods like in app purchases, temp paid services, temp subscription etc

24. While planning operations of the App, you need to have mechanism to address people queries or negative reviews.

25. Sentiment Subjectivity is high within Mature 17+ , Teen and Everyone 10+ categories.

So while planning of these category apps need to take this feedback redressal mechanism

Conclusion – Technical Recommendations

To App Development Team	

Technical Recommendations	Details
Technology	Android Version - Series 4.0 and above or Android version - Varies with Devise versions
Size	Free App - between 12 - 30 MB | Paid App - 23 - 45 MB | Heavy App - Max. 60 MB for paid Apps and 40 MB for Free Apps
Review Mechanism	A pop-up mechanism for review after every long activity on App Periodical review mechanism for people who gives frequent review Negative feedback tracing and tracking of redressal - Mechanism / Feature
User Interface	Free App - Can be of moderate technology | Paid App - Should be HD technology for better user experience
App Updates	App updation should be done periodically with latest technology
App Experience	Our Annexures consist of some top play store apps, study their technical specification to get an edge over other competitors
Annexures	Annexure 1 - Top 10 Genres | Catgories | Apps (Free Apps) Annexure 2 - Top 10 Genres | Catgories | Apps (Paid Apps) Annexure 3 - Top 10 Apps (Everyone Segment - Free apps) Annexure 4 - Top 10 Apps (Everyone 10+ Segment - Free apps) Annexure 5 - Top 10 Apps (Teen Segment - Free apps) Annexure 6 - Top 10 Apps (Mature 17+ Segment - Paid apps) Annexure 7 - Top 10 Apps (Everyone 10+ Segment - Paid apps) Annexure 8 - Top 10 Apps (Everyone Segment - Paid apps) Annexure 9 - Top 10 Apps (Most Positive Sentiments - Free apps) Annexure 10 - Top 10 Apps (Most Positive Sentiments - Paid apps)

Thank You !!