

Capstone Project

Play Store App Review Analysis

Team

G. V. Kapeesh Varma
Sridhar Nagar

CONTENTS

Sr. No.	Topic
1)	PROBLEM STATEMENT
2)	APPROACH
3)	DATA SUMMARY
4)	DATA VISUALIZATIONS
5)	CONCLUSIONS

PROBLEM STATEMENT

Google Play was launched on March 6, 2012, bringing together Android Market, and the Google under one brand, marking a shift in Google's digital distribution strategy. The Google Play store had over 108.5 billion app downloads in 2020.

The Play Store Apps data has enormous potential to drive app-making businesses to success. Actionable insights can be drawn for developers to work on and capture the Android market. The aim of this project is to explore and analyze this data to discover key factors responsible for app engagement and success.

APPROACH

In this project, we analyzed thousands of applications on Google Play Store and their corresponding reviews to understand the Android App market. The project has been divided into 2 parts: Apps analysis and Sentiment Analysis of User reviews. Initially, we imported, cleaned and analyzed various features of the Apps dataset through data exploration and visualizations.

In the next part, we loaded and cleaned the User Reviews dataset. The two datasets are then merged to visualize the composition of the total reviews and the sentiment polarity distribution.

DATA SUMMARY

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10841 entries, 0 to 10840
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   App                    10841 non-null  object
1   Category               10841 non-null  object
2   Rating                 9367 non-null   float64
3   Reviews                10841 non-null  object
4   Size                   10841 non-null  object
5   Installs               10841 non-null  object
6   Type                   10840 non-null  object
7   Price                  10841 non-null  object
8   Content Rating         10840 non-null  object
9   Genres                 10841 non-null  object
10  Last Updated           10841 non-null  object
11  Current Ver            10833 non-null  object
12  Android Ver            10838 non-null  object
dtypes: float64(1), object(12)
memory usage: 1.1+ MB
None
Total number of apps in the dataset = 10841
```

Apps DataFrame Summary

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated
3891	Create A Superhero HD	FAMILY	3.3	8642	19M	500,000+	Free	0	Teen	Casual	December 30, 2017
6468	BM FlexCheck	AUTO_AND_VEHICLES	3.5	2	34M	100+	Free	0	Everyone	Auto & Vehicles	June 8, 2018
1740	Angry Birds Classic	GAME	4.4	5566805	97M	100,000,000+	Free	0	Everyone	Arcade	May 24, 2018
517	Christian Dating For Free App	DATING	3.9	8722	13M	500,000+	Free	0	Mature 17+	Dating	June 20, 2018
6588	iFORA BP	MEDICAL	3.4	18	45M	1,000+	Free	0	Everyone	Medical	June 5, 2018
4998	AE Basketball	SPORTS	4.0	8100	Varies with device	1,000,000+	Free	0	Everyone	Sports	November 20, 2013
10301	FD Calculator (EMI, SIP, RD & Loan Eligibility)	FINANCE	5.0	104	2.3M	1,000+	Free	0	Everyone	Finance	August 7, 2018
9404	Princess Palace: Royal Pony	FAMILY	4.4	11442	66M	1,000,000+	Free	0	Everyone	Educational;Creativity	June 30, 2016
6022	BD Sports Box & TV	FAMILY	4.2	5	3.3M	500+	Free	0	Everyone	Entertainment	April 13, 2018
4526	Q-See QT View	TOOLS	3.6	7357	19M	100,000+	Free	0	Everyone	Tools	July 12, 2018

Apps DataFrame before Data Wrangling

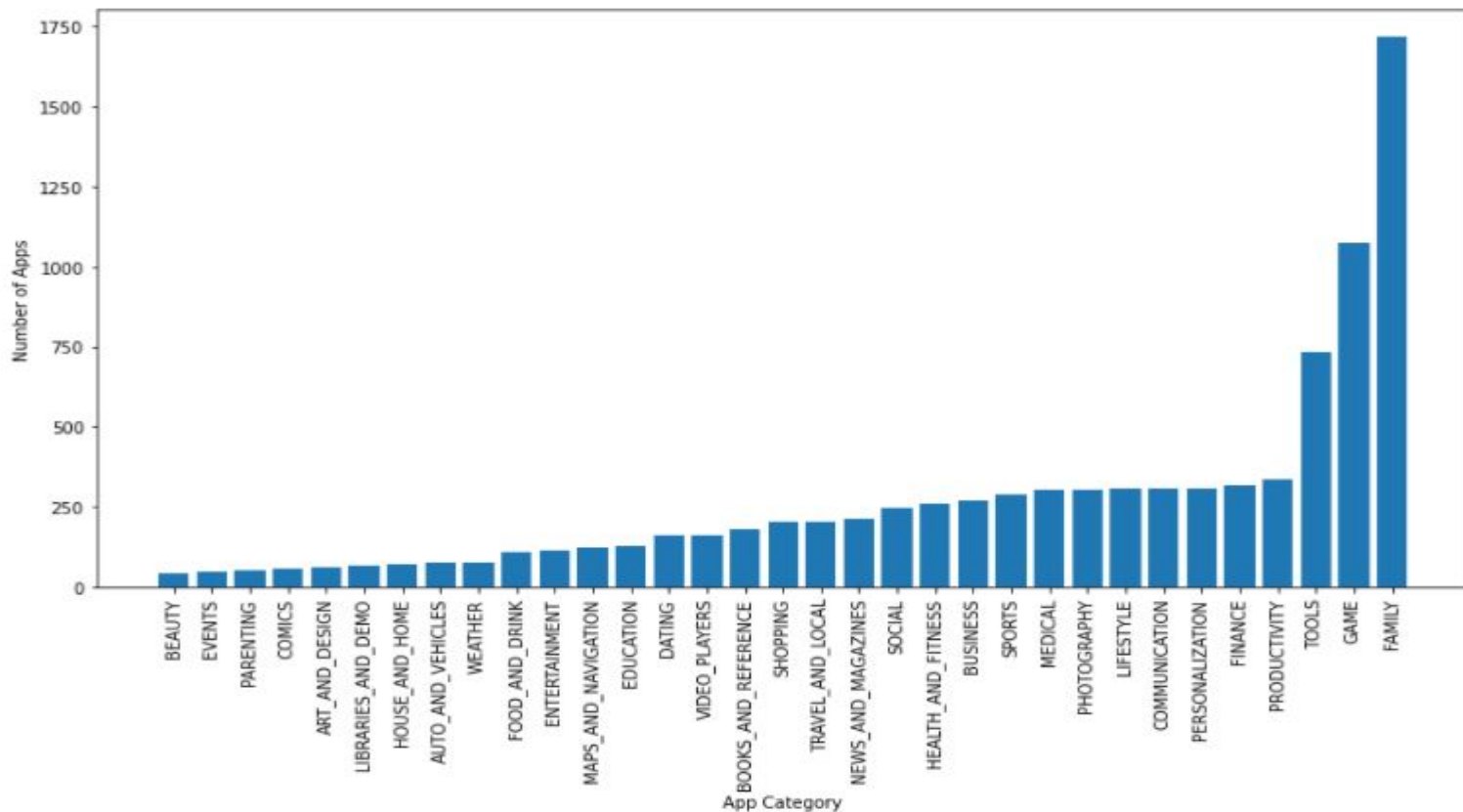
	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated
9066	Multiple Sclerosis Dx & Mgmt.	MEDICAL	4.5	131	60.0	10000	Free	0.0	Everyone	Medical	2017-09-25
6564	Thai Sic Bo	GAME	4.1	14283	4.2	1000000	Free	0.0	Teen	Casino	2017-10-01
10192	Facejjang	PHOTOGRAPHY	4.1	100179	96.0	10000000	Free	0.0	Everyone	Photography	2018-05-17
3232	Google Street View	TRAVEL_AND_LOCAL	4.2	2129707	11.5	1000000000	Free	0.0	Everyone	Travel & Local	2018-08-06
8844	DS Thermometer	WEATHER	3.7	631	3.0	100000	Free	0.0	Everyone	Weather	2015-05-30
902	AMC Theatres	ENTERTAINMENT	4.3	44550	72.0	1000000	Free	0.0	Everyone	Entertainment	2018-07-18
7820	NetClient CS	BUSINESS	3.5	125	4.3	10000	Free	0.0	Everyone	Business	2018-01-05
44	Popsicle Sticks and Similar DIY Craft Ideas	ART_AND_DESIGN	4.2	26	12.0	10000	Free	0.0	Everyone	Art & Design	2018-01-03
7213	Placement Tips for BE (CE/IT)	FAMILY	4.4	7	6.3	500	Free	0.0	Everyone	Education	2016-06-24
6899	BlitzWolf Shutter - BW Shutter	PHOTOGRAPHY	2.9	333	1.8	50000	Free	0.0	Everyone	Photography	2017-11-22

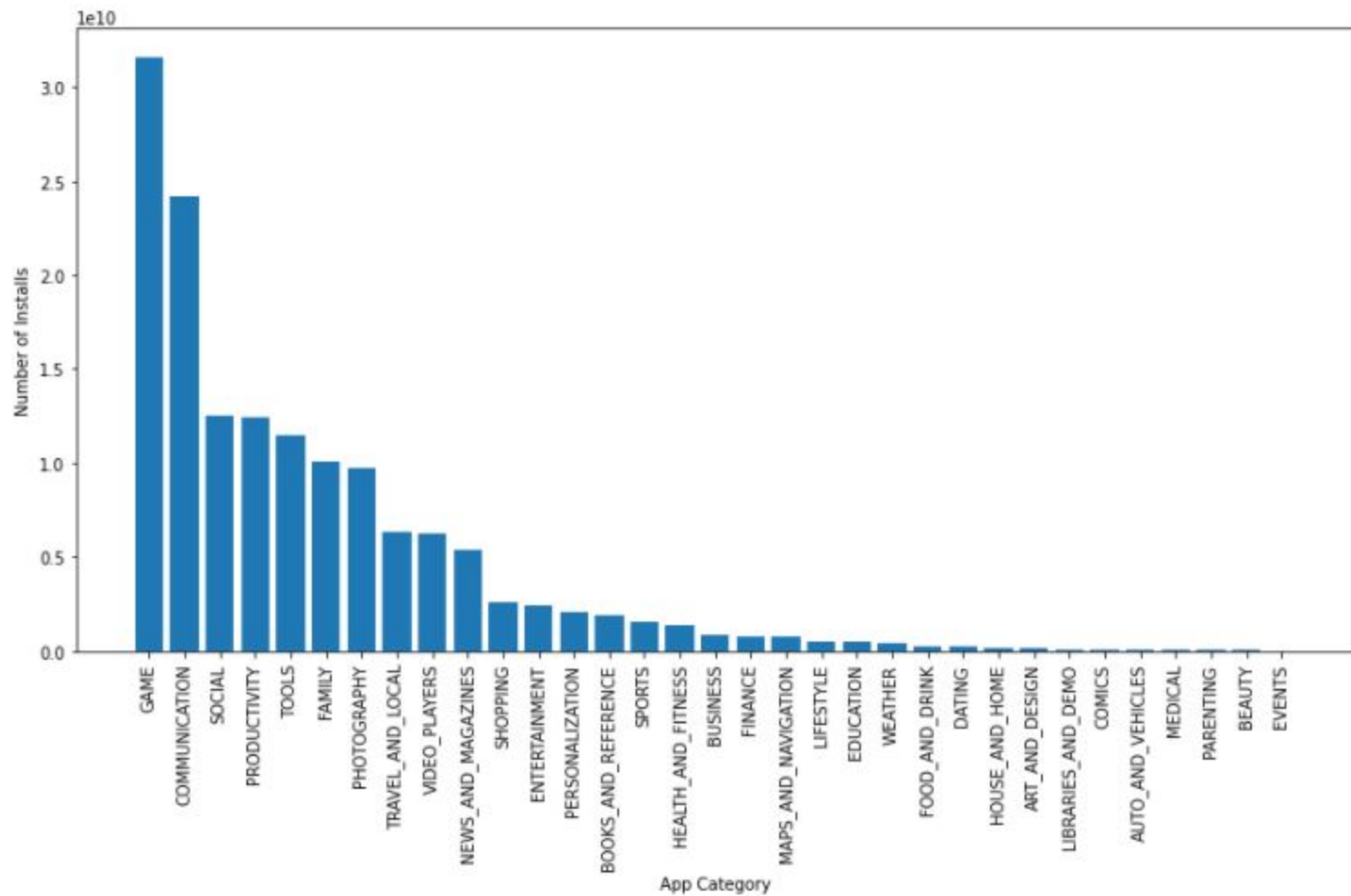
Apps DataFrame after Data Wrangling

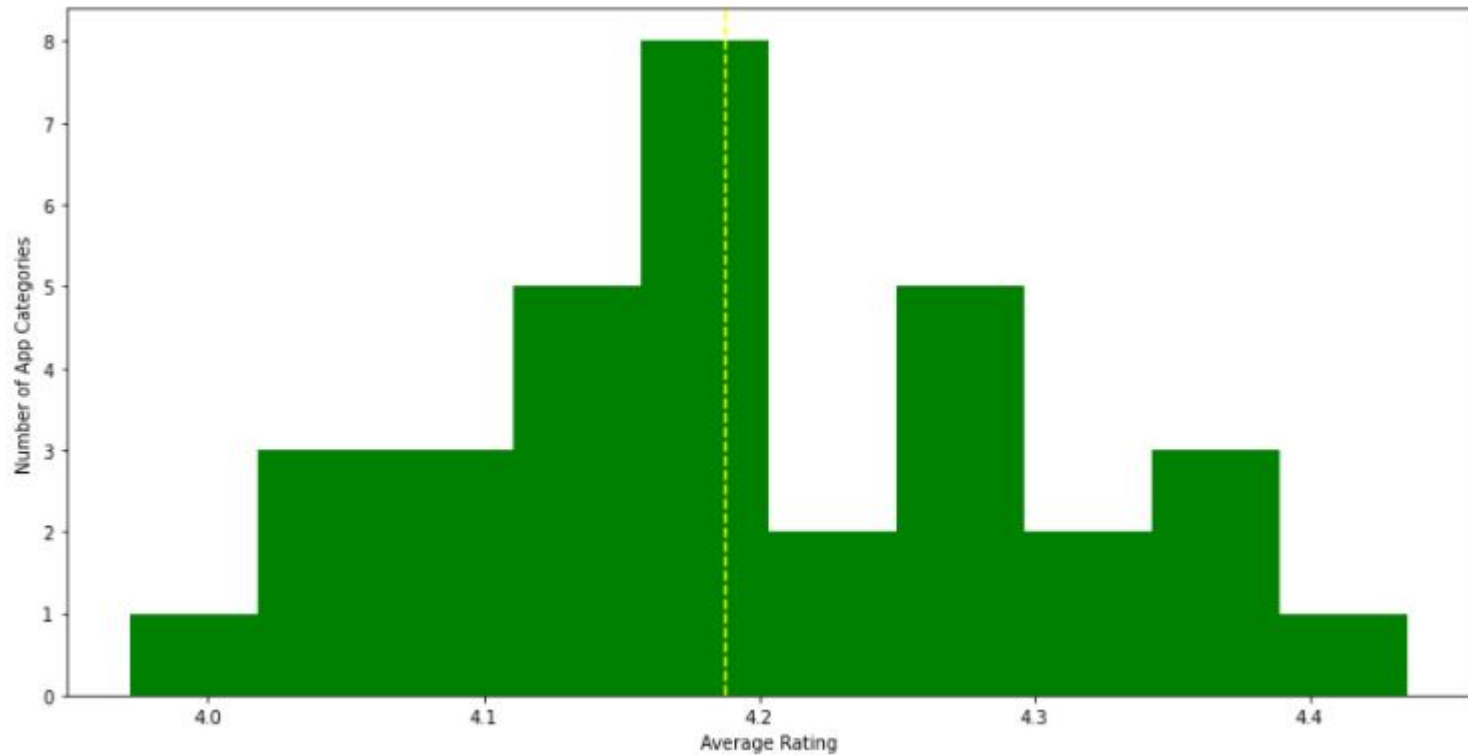
Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver	Translated_Review	Sentiment	Sentiment_Polarity	Sentiment_Subjectivity
_AND_DESIGN	3.9	967	14.0	500000	Free	0.0	Everyone	Design;Pretend Play	2018-01-15	2.0.0	4.0.3 and up	A kid's excessive ads. The types ads allowed a...	Negative	-0.250	1.000000
_AND_DESIGN	3.9	967	14.0	500000	Free	0.0	Everyone	Design;Pretend Play	2018-01-15	2.0.0	4.0.3 and up	It bad >:(Negative	-0.725	0.833333
_AND_DESIGN	3.9	967	14.0	500000	Free	0.0	Everyone	Design;Pretend Play	2018-01-15	2.0.0	4.0.3 and up	like	Neutral	0.000	0.000000
_AND_DESIGN	3.9	967	14.0	500000	Free	0.0	Everyone	Design;Pretend Play	2018-01-15	2.0.0	4.0.3 and up	I love colors inspyering	Positive	0.500	0.600000
_AND_DESIGN	3.9	967	14.0	500000	Free	0.0	Everyone	Design;Pretend Play	2018-01-15	2.0.0	4.0.3 and up	I hate	Negative	-0.800	0.900000

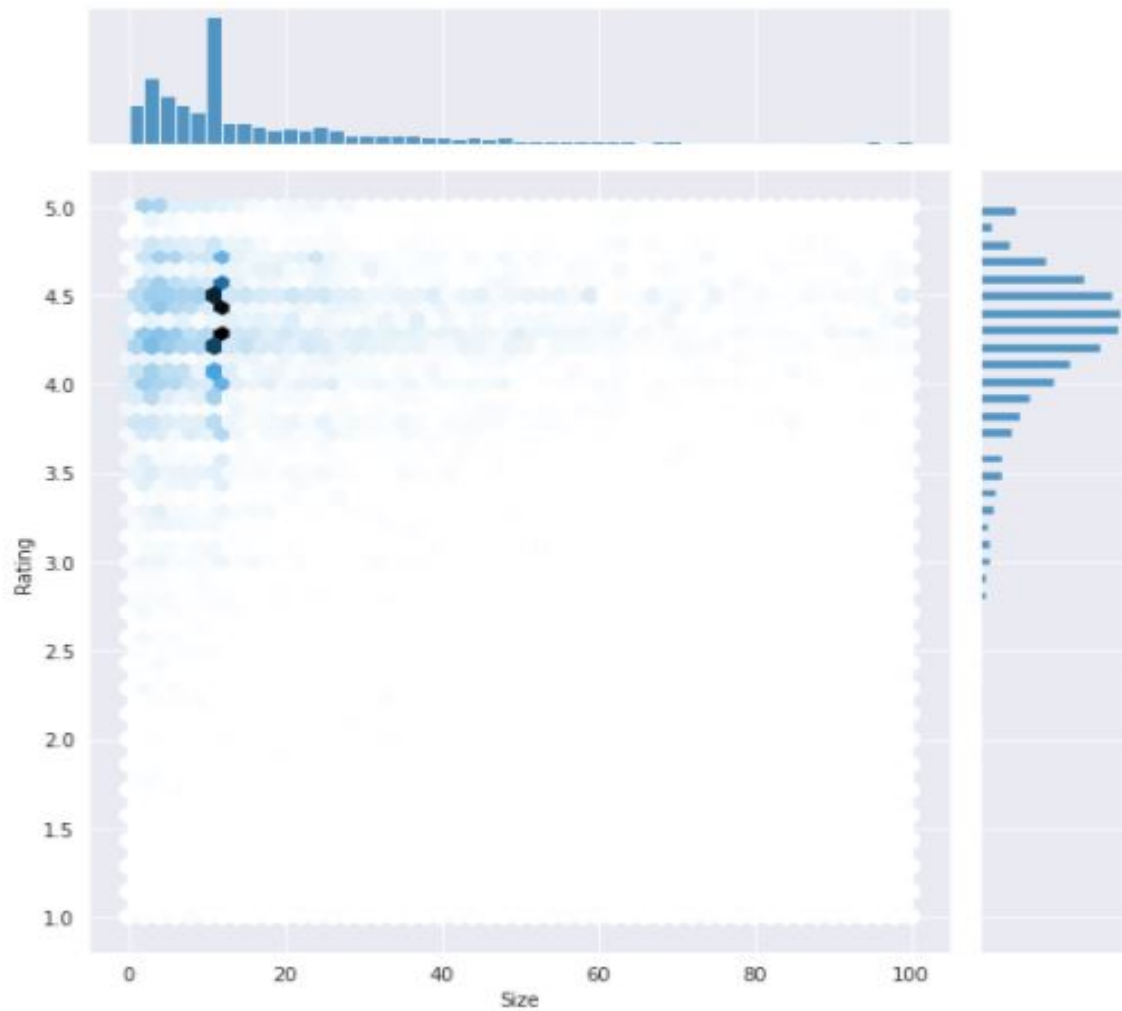
Merged DataFrame (Apps + Reviews)

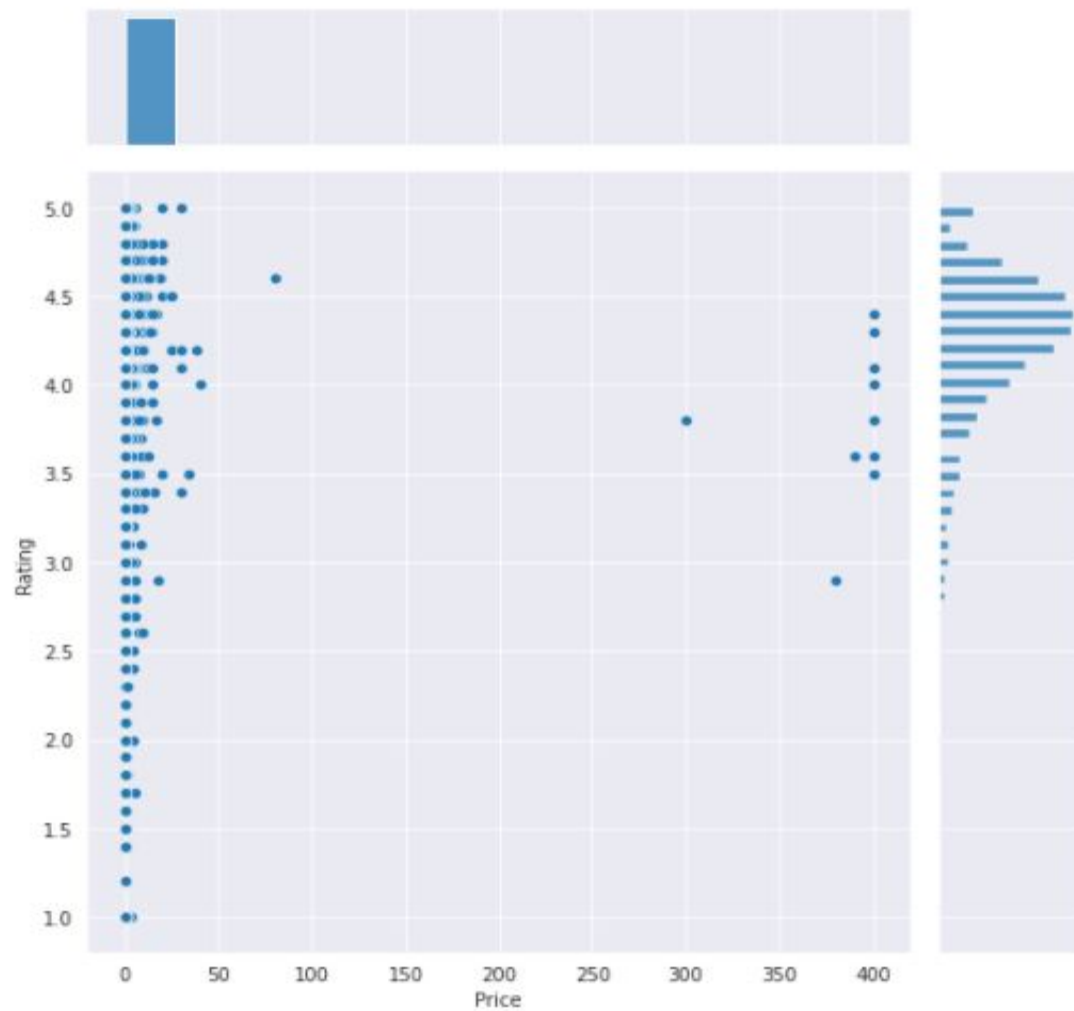
DATA VISUALIZATIONS

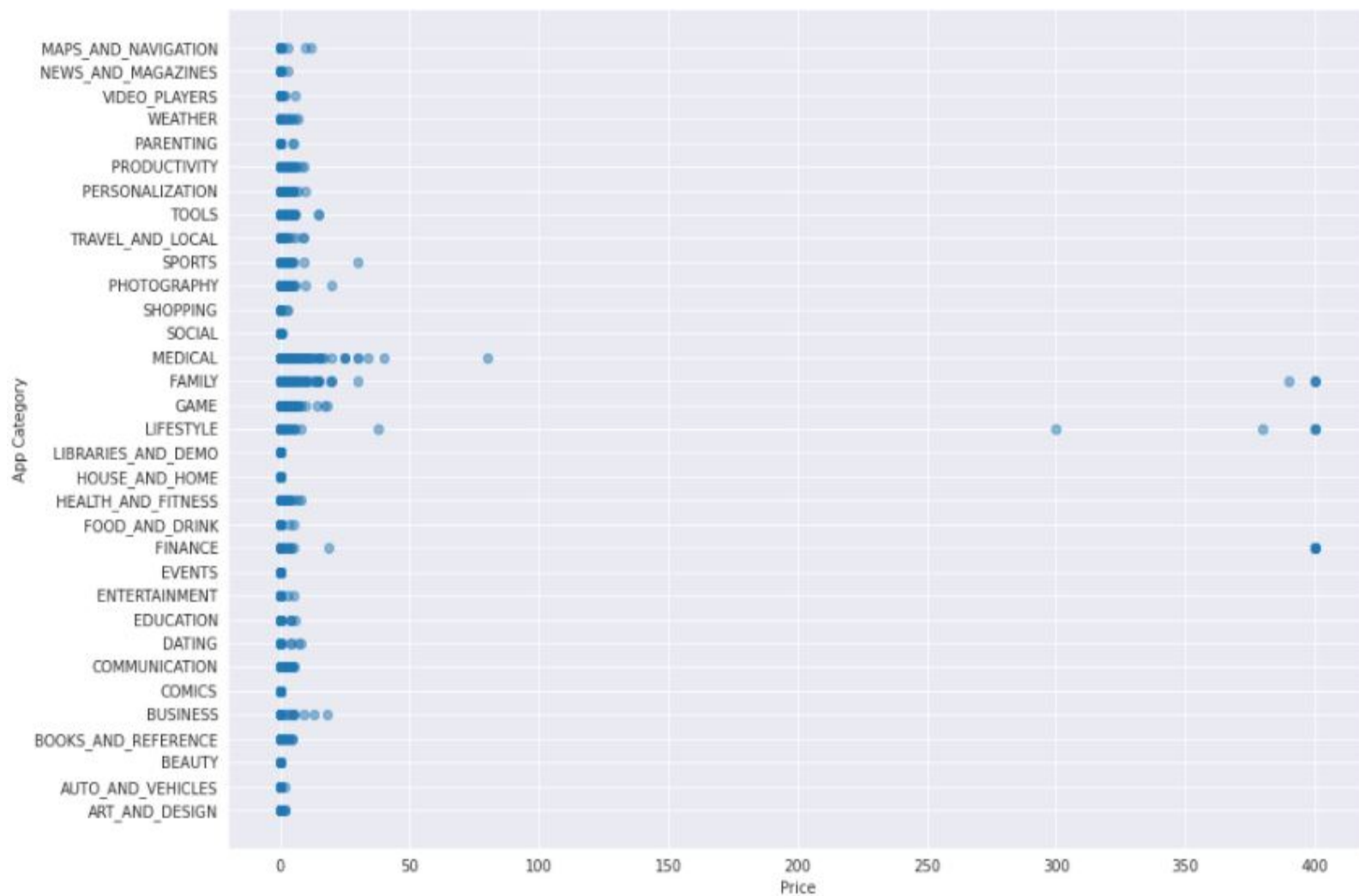


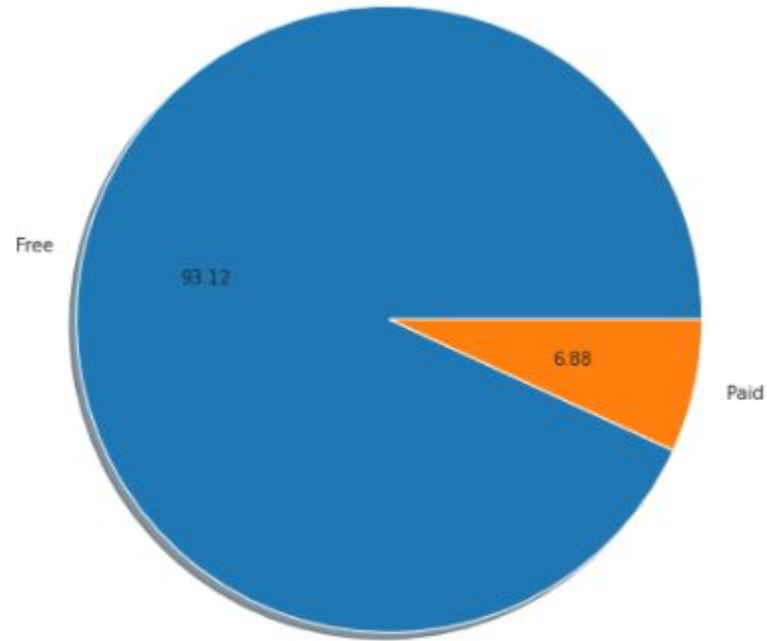






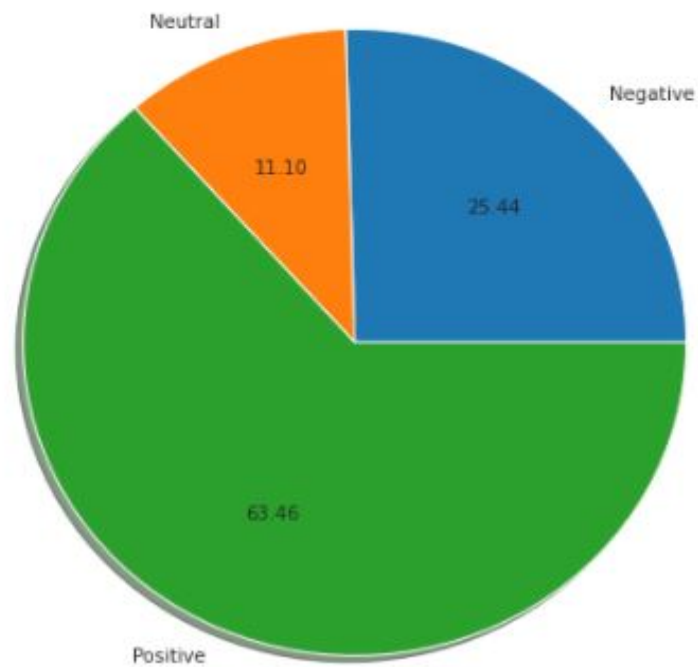




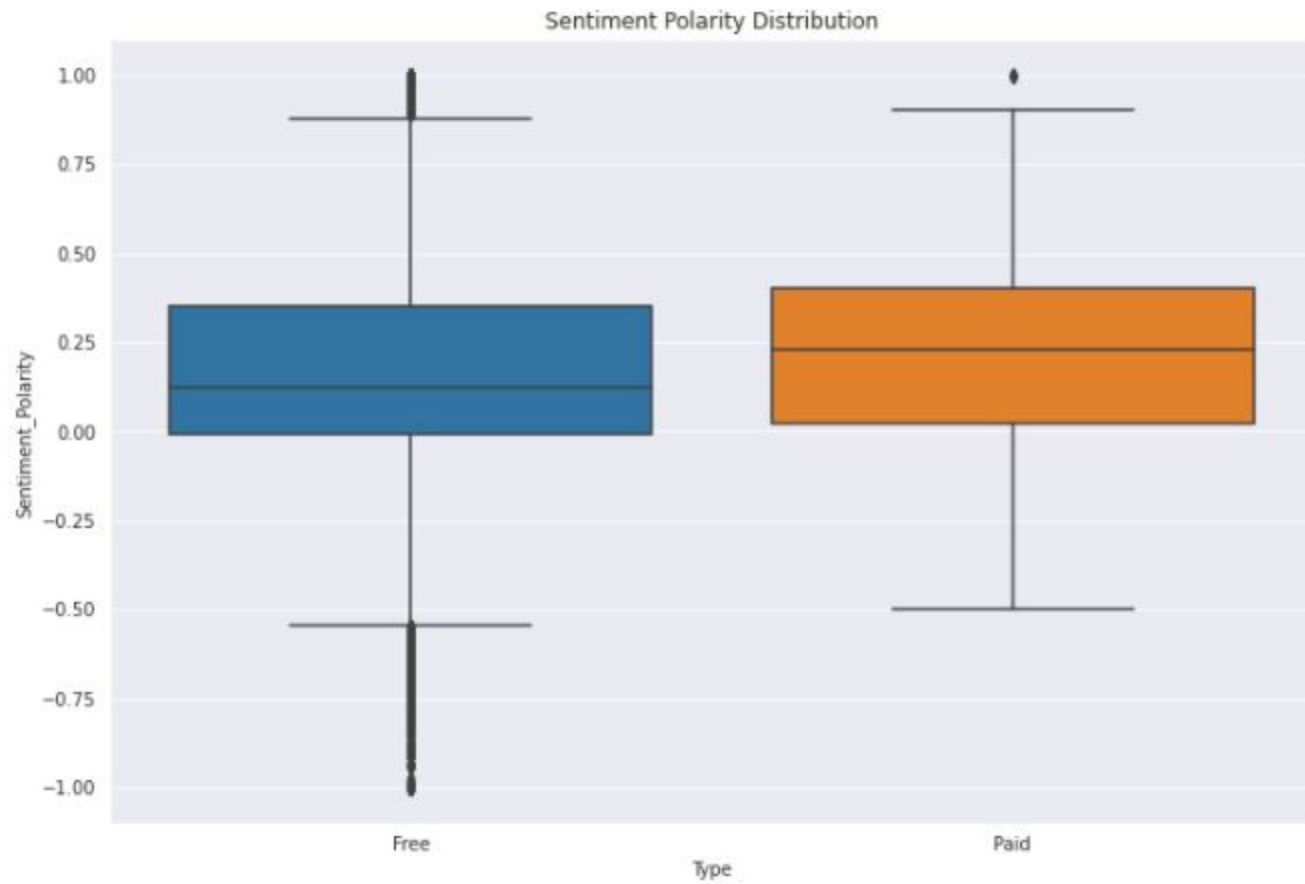


Popularity of Free vs Paid Apps

```
Sentiment  
Positive    37523  
Negative    15041  
Neutral     6560  
dtype: int64
```



Distribution of Sentiments



CONCLUSIONS

1. The most popular App Category is "Game".
2. A large number of Apps fall into "Family" Category i.e., this is the category with highest number of subsequent apps.
3. The total average rating of Play Store Apps is 4.18.
4. The App Categories with least and highest average ratings are "Dating" and "Events" respectively.
5. Free apps are highly popular when compared to Paid apps.
6. Most of the apps receive Positive reviews i.e., about 63.5%. Free apps might have a broader range of reviews than Paid apps, but the median of Sentiment Polarity is higher in Paid apps.