

Executive Summary	Exploratory Data Approach and Data Preparation	Sample Findings from the EDA in Python	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?
-------------------	--	--	--	---	---	---	---

# Analysis on the Mobile Apps in Google Playstore

## Background

The competitive market on smartphones posits a high potential for success in application development. App-making businesses can look at application downloads, reviews, and ratings datasets, to name a few, to realize such potential, especially in the Android market.

## Objectives

This project aims to suggest to developers and application development businesses the next best application. We look at the number of installations, number of reviews, quality of reviews, and average ratings of applications and their corresponding categories/genres to come up with actionable recommendations.

## Method

We have utilized the Google Play Store Apps data we have found in Kaggle. We have performed complete EDA (Exploratory Data Analysis) with this data set (app review and user review with NLP applied) for us to help answer the Key Business Questions for each area we will be sharing. Tableau charts would be limited to the key areas we analyzed best to be shared but more insights can be found in our python notebook.



Executive Summary	Exploratory Data Approach and Data Preparation	Sample Findings from the EDA in Python	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?
-------------------	--	--	--	---	---	---	---

# Data Preparation

## Before Data Preparation

	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

## After Data Preparation

App_ID	App	Category	Rating	Reviews	Installs	Type	Price	Content Rating	Genres	Last Updated	Genre Main	Genre Sub	Size in MB	Installs_Min
0	0 Photo Editor & Candy Camera & Grid & Scrapbook	ART_AND_DESIGN	4.1	159	10,000+	Free	0.0	Everyone	Art & Design	2018-01-07	Art & Design	None	19.0	10000
2	1 Coloring Book Moana	FAMILY	3.9	974	500,000+	Free	0.0	Everyone	Art & Design;Pretend Play	2018-01-15	Art & Design	Pretend Play	14.0	500000
3	2 U Launcher Lite – Free Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	5,000,000+	Free	0.0	Everyone	Art & Design	2018-08-01	Art & Design	None	8.7	5000000
4	3 Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	50,000,000+	Free	0.0	Teen	Art & Design	2018-06-08	Art & Design	None	25.0	50000000

## Initial Observations:

1. There are **MISSING** or **NULL** Values in the dataset.
2. The datatypes of some columns are **INCORRECT**; needs reassignment.
3. Some columns are **IRRELEVANT**, hence needs to be dropped.

## Steps:

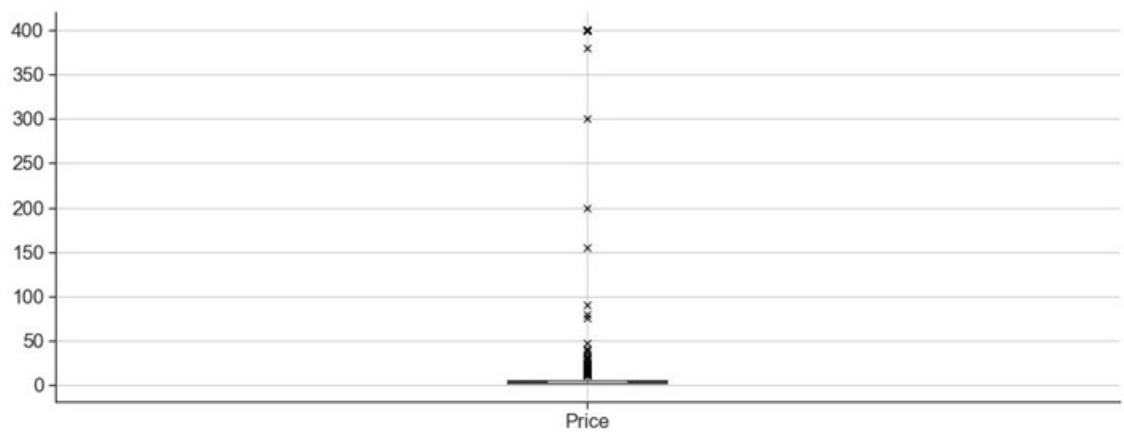
1. Dropped columns **Current Ver** and **Android Ver**
2. Missing Value Imputation on the **1474** entries of Rating column: Imputed values are the **MEAN** of **Category per Type**.
3. Added columns **Genre Main**, **Genre Sub**, **Size in MB**, and **Installs\_Min**
4. Fixed datatypes accordingly. (e.g. Last Updated from string to date format; Revi..



Executive Summary	Exploratory Data Approach and Data Preparation	Sample Findings from the EDA in Python	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?
-------------------	--	--	--	---	---	---	---

# Sample Findings from the EDA

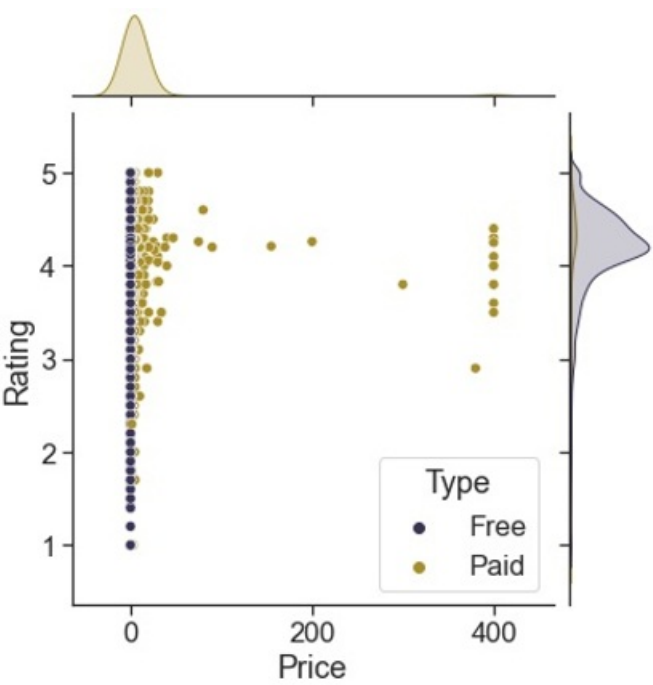
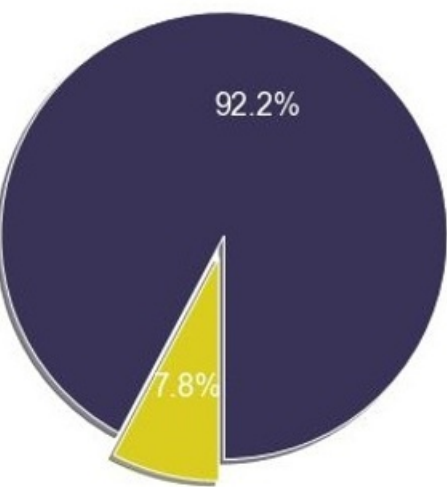
Box Plot



Correlation Chart



Percent of Free App in store

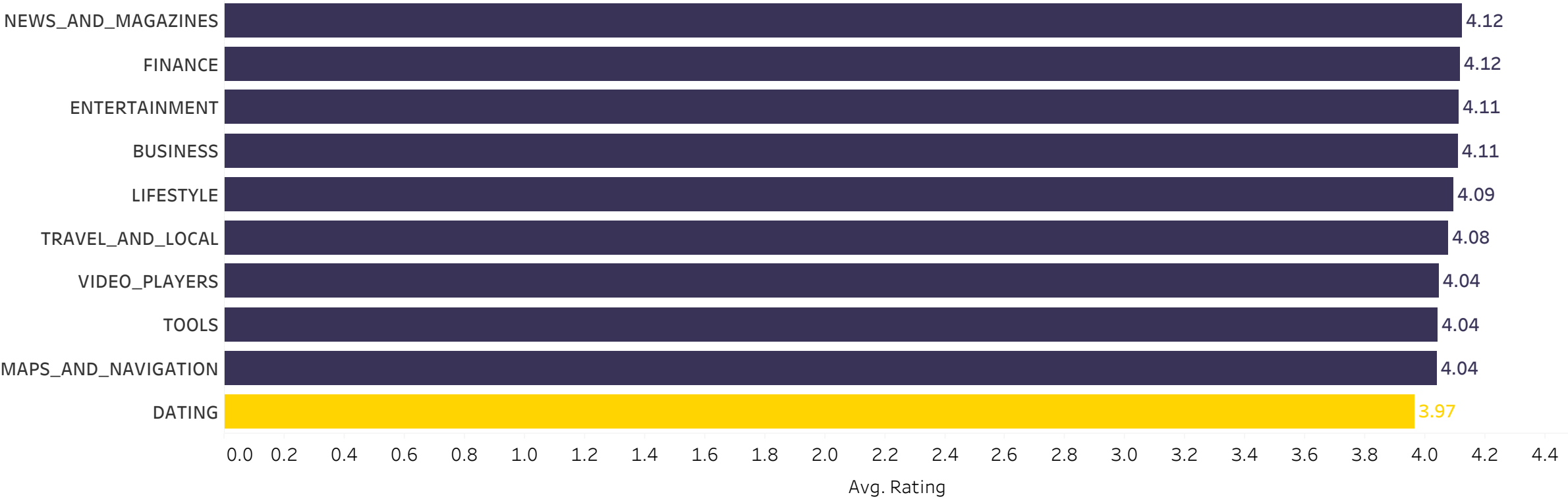


Executive Summary	Exploratory Data Approach and Data Preparation	Sample Findings from the EDA in Python	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?
-------------------	--	--	--	---	---	---	---

# Dating Ranks Lowest Among Categories

The average rating for dating apps is only **3.97**, the only category below the 4.0 rating mark. These are the bottom ten rated categories from Google Play Store.

Top 10 Categories

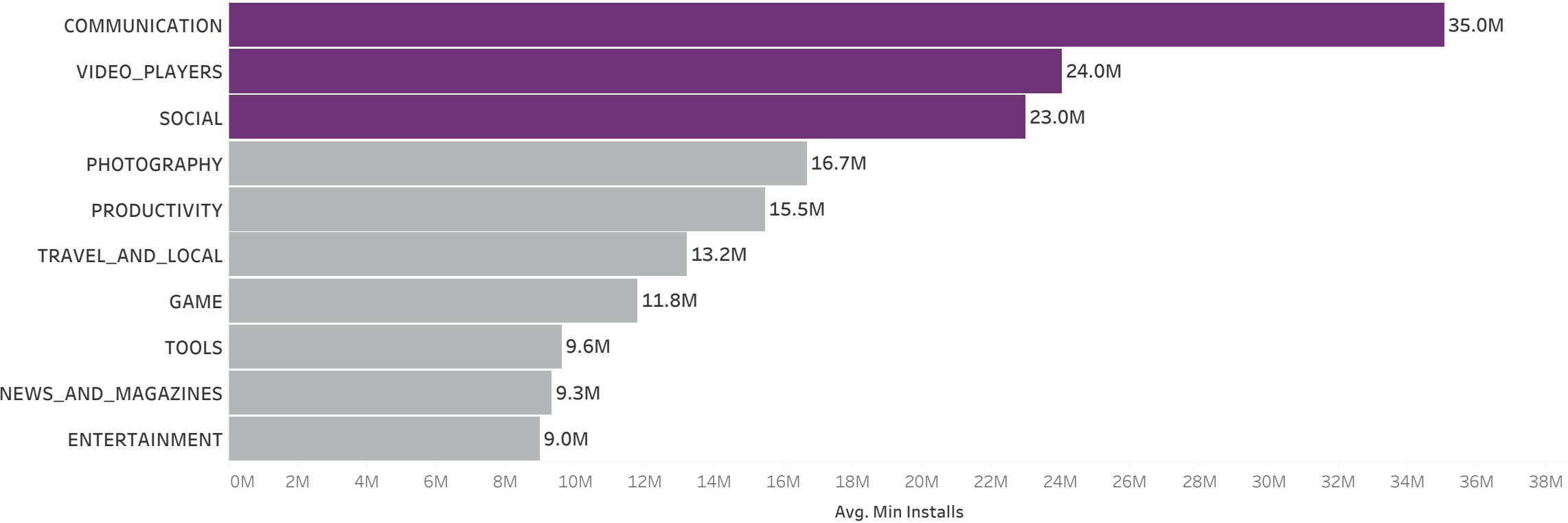


Executive Summary	Exploratory Data Approach and Data Preparation	Sample Findings from the EDA in Python	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?	What do the frequently used words i..
-------------------	--	--	--	---	---	---	---	---------------------------------------

# Entertainment and Social apps are the most popular

**Communication, Video, and Social** apps have the highest number of installs among all categories.  
GMail , FB Messenger , Youtube , and Instagram are some of the popular apps within these categories.

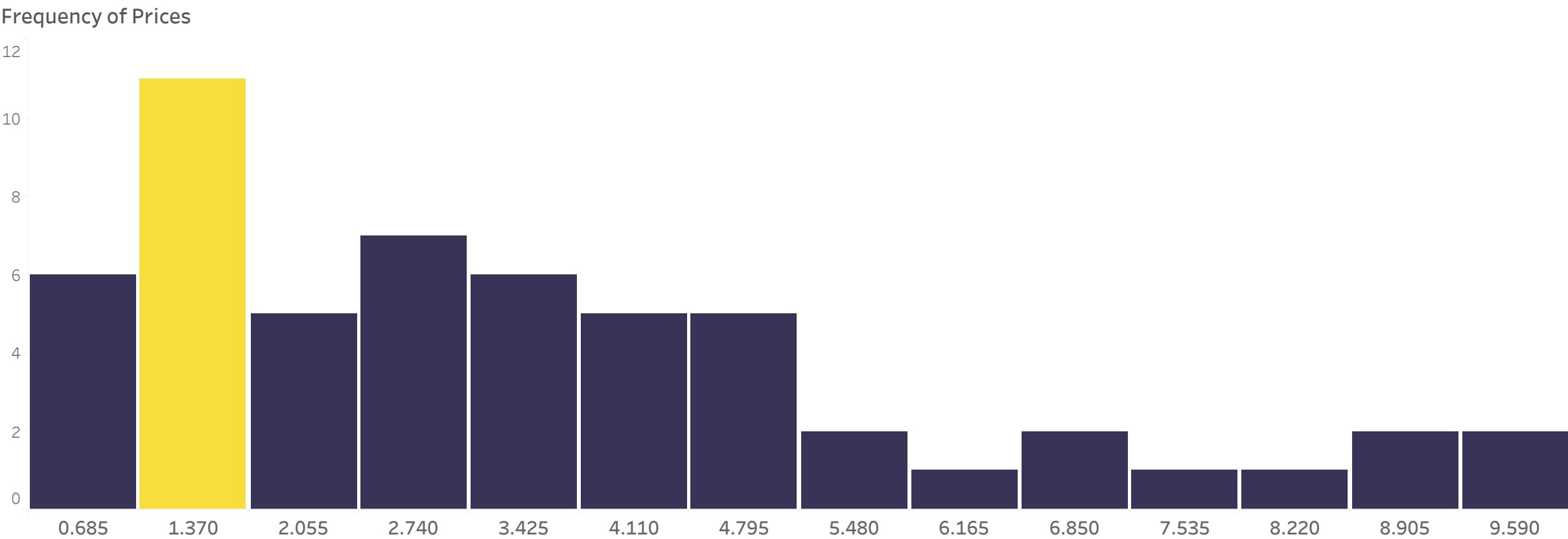
Top 10 Categories



Exploratory Data Approach ..	Sample Findings from the EDA in Python	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?	What do the frequently used words in 'Negative Sentiments' suggest?	What do the frequently used words i..
------------------------------	--	--	---	---	---	---	---	---------------------------------------

# Overall, the profit from each app download is less than \$2

Majority of the **7.8% paid google apps** cost around **\$1- \$2**. The usual price is less than \$10.  
92.2% of the apps across different categories are **free to download** . This should also be considered when coming up with a price strategy.

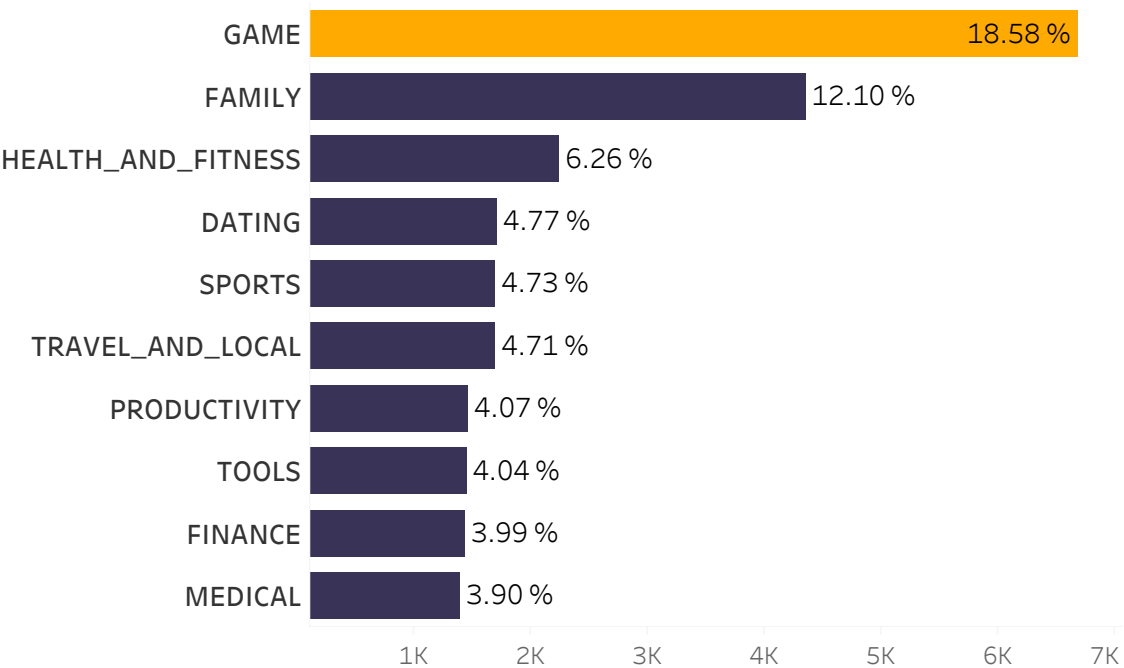


Sample Findings from the E..	What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?	What do the frequently used words in 'Negative Sentiments' suggest?	What do the frequently used words in 'Neutral Sentiments' suggest?	Recommendations
------------------------------	--	---	---	---	---	---	--	-----------------

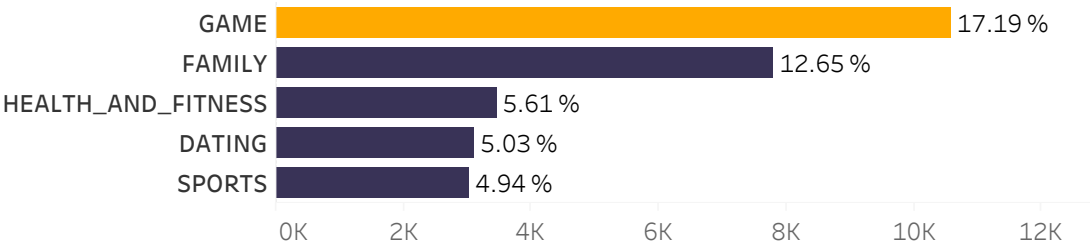
# User Review Distribution By App Category

**18.58%** of all reviews are under Games (Top 3 Category),  
but it has majority of the Negative reviews at **30.08%** and only **17.19%** of the Positive Reviews  
Communication and Social which are the Top 1 and 2 Categories only contributes **2.86%** and **2.06%** of the Total Reviews

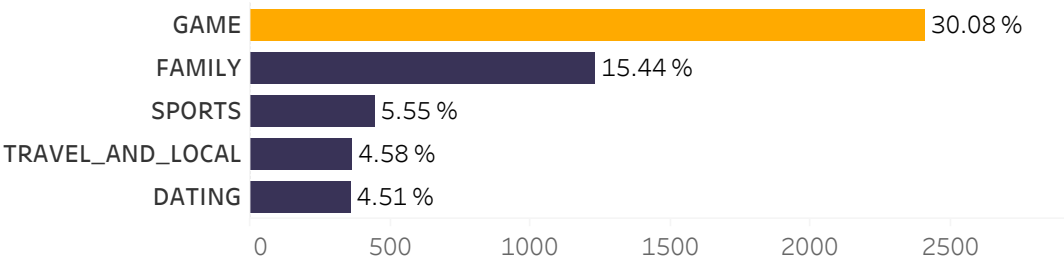
Top 10 Categories by No. of Reviews



Top 5 Categories with Positive Review



Top 5 Categories with Negative Review













What is the lowest rated app category?	What popular categories can app developers and business owners explore?	How much profit can we expect from app downloads?	Which app category gets the most reviews?	What do the frequently used words in 'Positive Sentiments' suggest?	What do the frequently used words in 'Negative Sentiments' suggest?	What do the frequently used words in 'Neutral Sentiments' suggest?	Recommendations
--	---	---	---	---	---	--	-----------------

# Recommendations



## Answers to Key Business Questions

- (1) What is the lowest-rated app category?  
**Dating is the lowest-rated app category, may be due to disappointment related to seeking partners online.**
- (2) What popular categories can developers explore?  
**Communication, Social, and Games are the most popular categories in the Android Market.**
- (3) How much profit can we expect from app downloads?  
**Best price range for paid apps is \$1 to \$2 per download.**
- (4) Which app category gets the most reviews?  
**Games category is the most reviewed category of applications.**

## Recommendations

- (1) On lowest rated app categories:  
**While developers can explore popular categories for their next app development, lowest-rated app categories are still worth noting. While lowest in terms of customer approval, it offers a high potential for a niche market which they can dominate.**
- (2) On most popular app categories:  
**We suggest for developers to explore these categories more, developing a different communication/social/game app that will bring a unique user experience. Such uniqueness will be key for penetrating an already competitive market.**
- (3) On price range:  
**Considering free applications for downloads, in-app purchases can provide customers the option to pay for certain features or functionalities that are not offered on free version of the app. Developers can consider the price range of \$1 to \$2 a pop for these in-app purchases.**
- (4) On most-reviewed category:  
**Should the developers decide to explore game applications category, as inferred from the negative sentiment word cloud, timely update of these apps are crucial for customer satisfaction. Developers can easily strategize around this category with the influx of feedback for game applications.**

