Play Store Data Analysis

Problem Statement: In this project I leveraged a dataset of my choice, which I had previously analyzed during an Exploratory Data Analysis (EDA) project in Python. It was Play Store Data Analysis, which involves a dataset of various apps.

I have designed an interactive and informative Power BI dashboard using the dataset that was used during EDA. The dashboard effectively conveyed insights, trends, and key findings from the data, making it accessible and understandable for a diverse audience.

Dataset Selection: In this project, I am going to analyze Play Store dataset. This dataset contains information of apps available in google play store.

Why Google PlayStore data: Google Play was launched on March 6,2012, bringing together Android Market making a shift in Google’s digital distribution strategy. There are more than 5 million apps found on Google Play Store. Android is the most popular operating system in the world. Lot of developers and designers work on it to make an app successful on the Play Store.

Google Play Statistics: Over 3 billion active users spanning over 190 countries.

# Dataset details:

* Dataset name: Play store data
* Description: The dataset contains information of apps available in Google Play Store, their specifications, and details about their parameters like price, reviews, category, genre, etc.

# Key Attributes:

* App: Name of the application.
* Category: Category pf the application.
* Rating: Rating given to the application.
* Reviews: Number of reviews given to the application.
* Size: Size of the application.
* Installs: Number of downloads of the application.
* Type: Free or paid.
* Price: Price of the application if it is paid.
* Content Rating: It is age appropriate or not.
* Genres: Type of genre the application belongs to.
* Last Updated: When the last time the application is updated.
* Current Ver: Current version of the application’
* Android version: Minimum android version required to run the application.

# Problem areas to Explore:

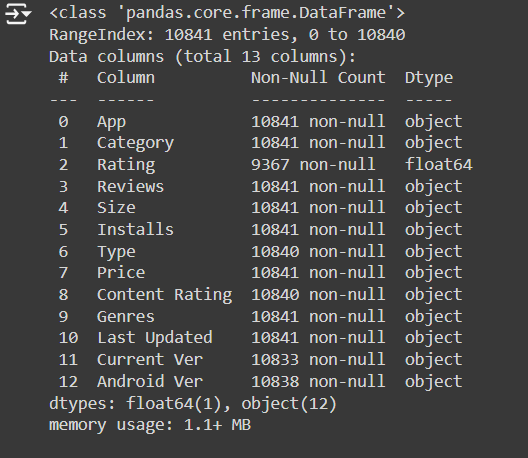
* Which are the popular apps, their category, genre, etc.
* What is the percent share of different apps by type, category, etc.
* How the pricing is varying with different parameters.
* What are the different correlations between type of parameters.
* Which category of content is found more?

# The objective of this project is to explore and analyze the data to discover key findings pertaining to play store dataset.

1. Studying and understanding the dataset.
2. Clean the data.
3. Prepare the data for processing and analysis.
4. Analyze the data deeply and extract insights.

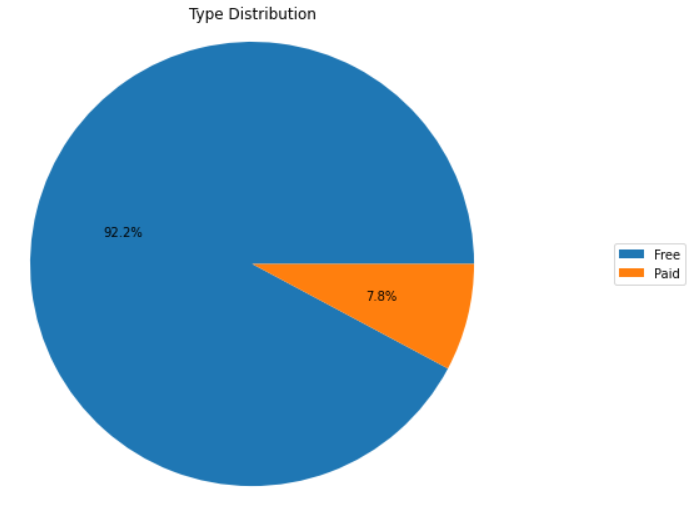
# Introduction of the dataset:

The dataset has 10841 observations in it with 13 columns and it is a mix between categorical and numeric values.



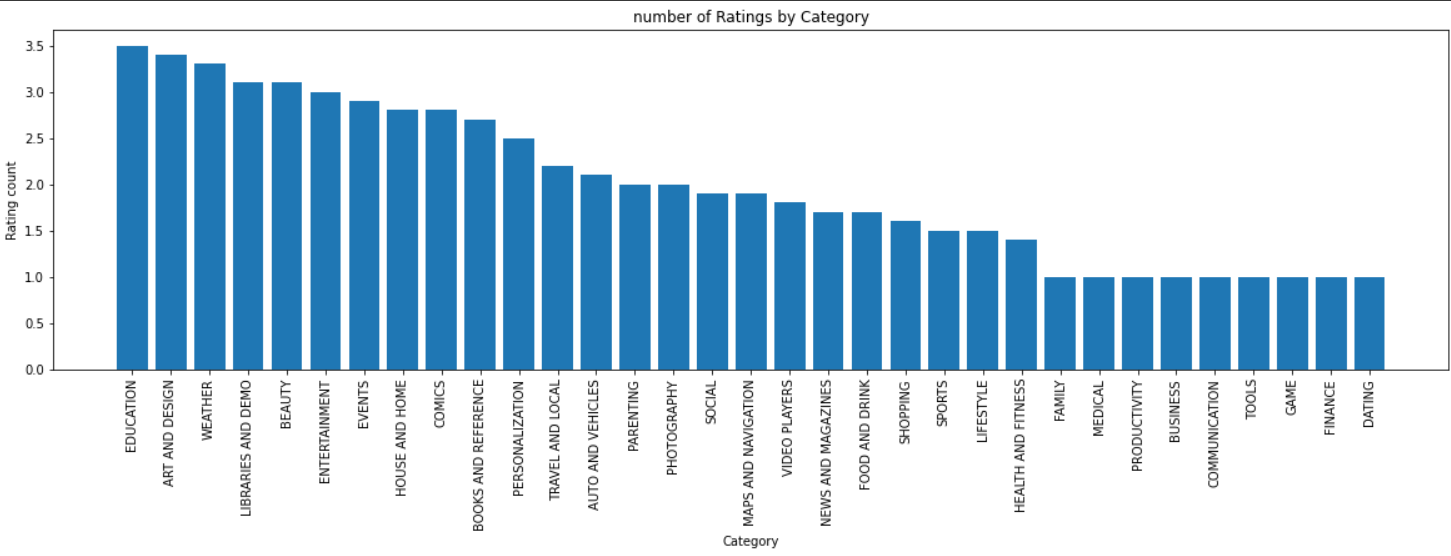
Visualization

1.No of apps free or paid)



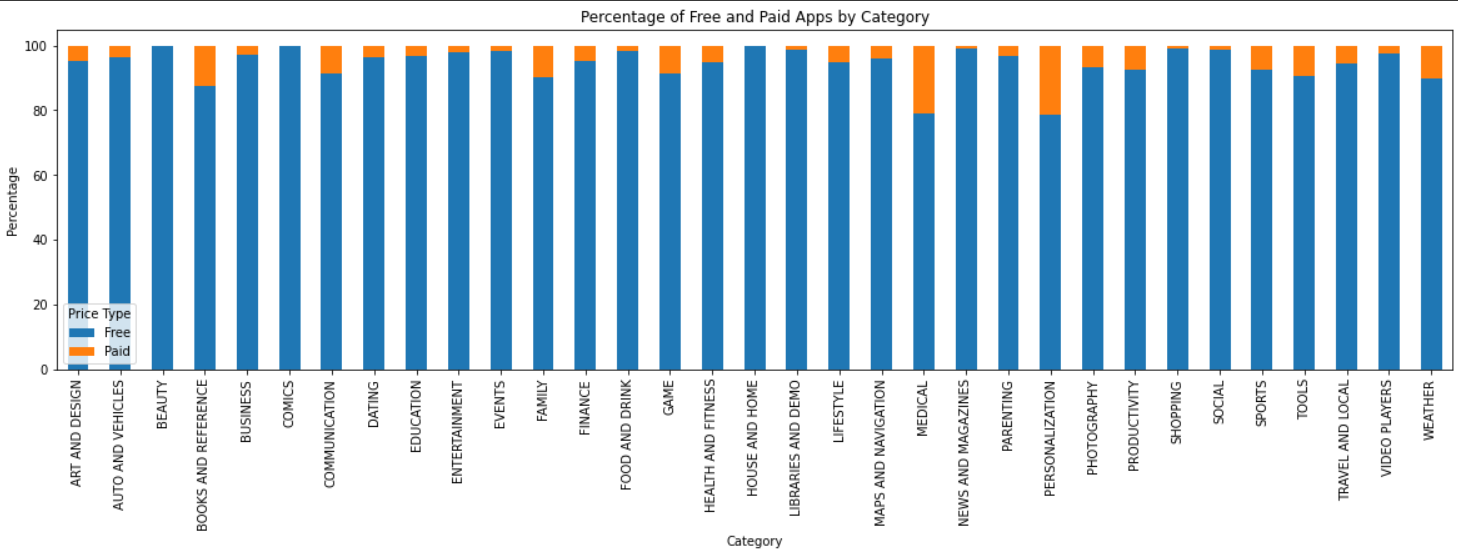
A pie chart shares the distribution of apps by their type (free or paid) very clearly. We see that the total number of apps which are paid are 7.8% and rest are free.

2.Rating Count of each Category

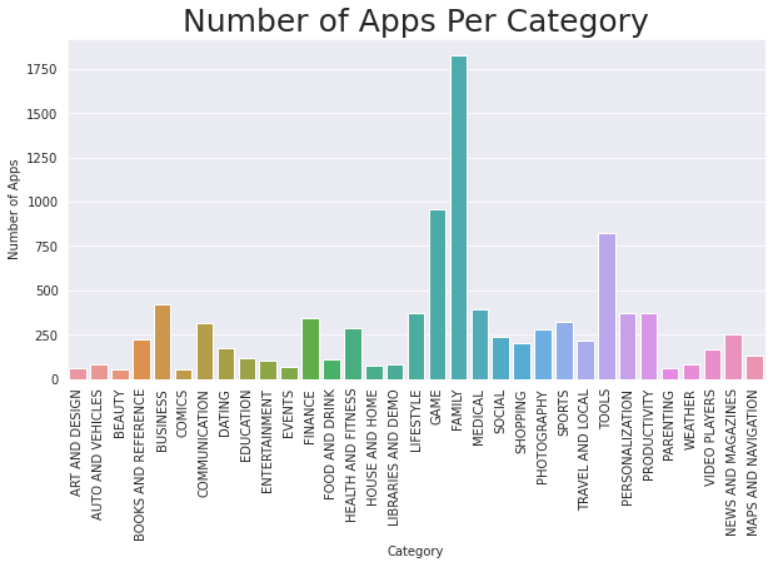


Bar chart would be best suited for the visualization to understand which category has lowest ratings and which has highest. We see that there are apps in the category that has lowest rating and through this we can identify the category in which we can work on by exceptionally making a better app.

3. Percentage of Free and Paid Apps by Category

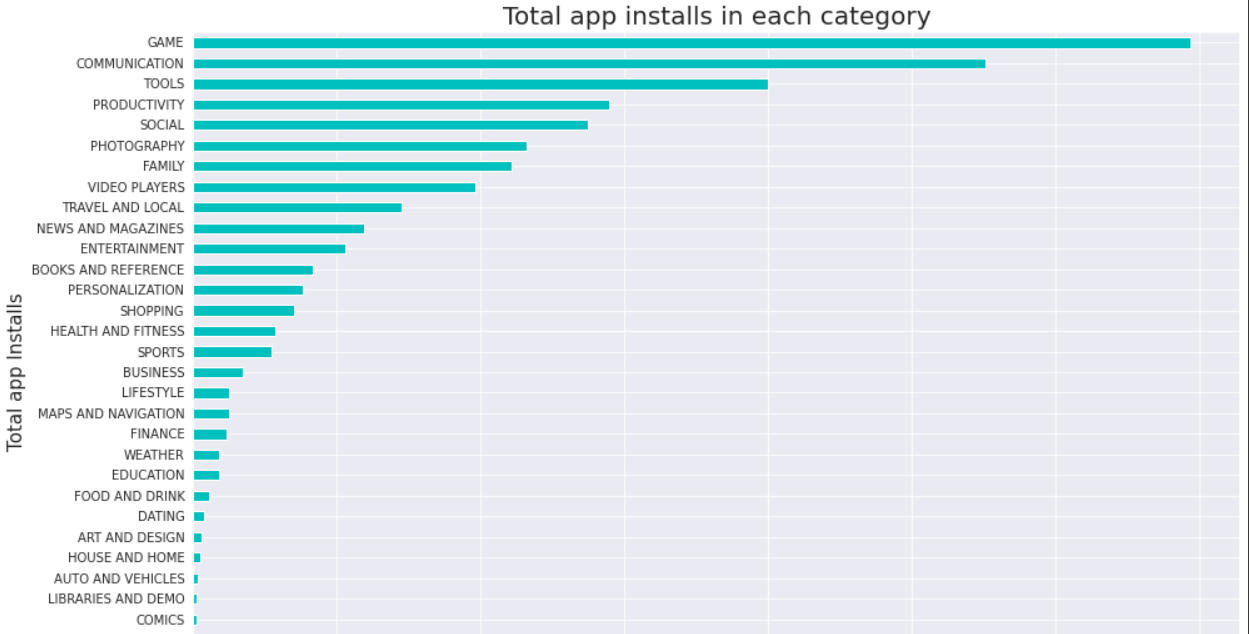
A stacked bar plot shows the percentage for all app categories by type (free and paid). We see that the percentage of paid apps in personalization and Medical are more.

4.Number of Apps per Category



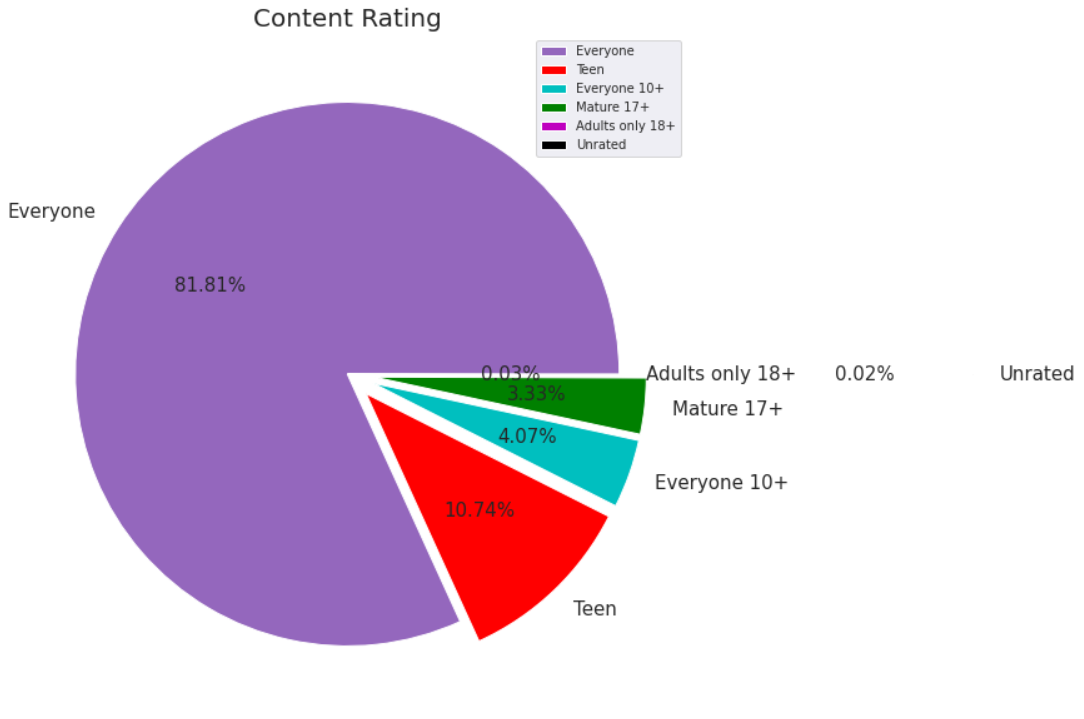
We find that the number of apps in Family category is 18.5% followed by Games with 9.94% and tools app with 8.55%. We see that there are many apps in family, games and tools category, this plot can help us to understand the market size of the category in which we are trying to make our new app.

5.Total App Installs in each Category



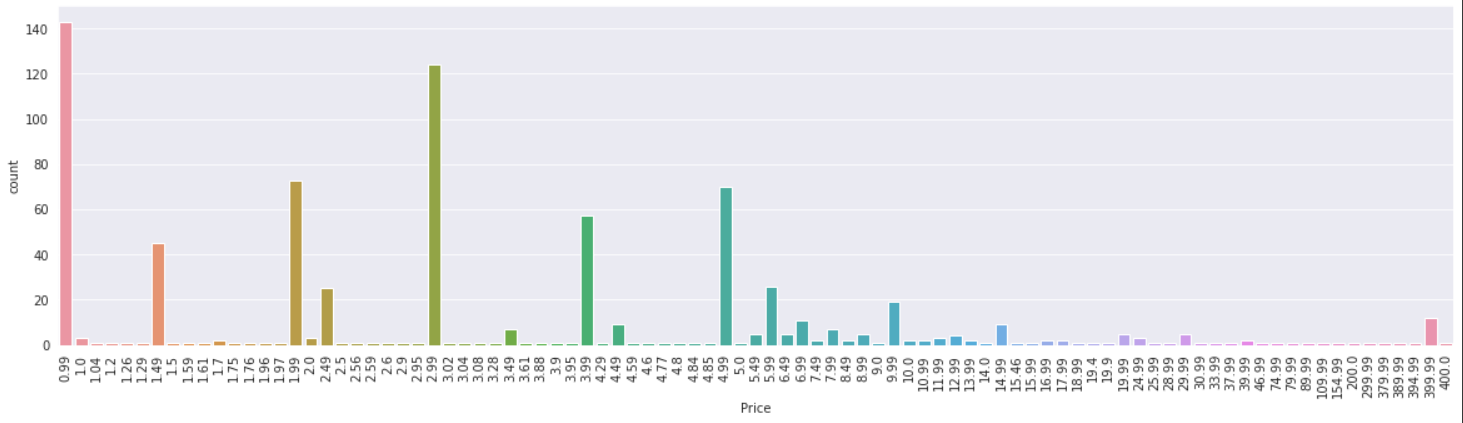
A horizontal bar plot gives us a clear picture of the number of installs with respect to category. We see that the maximum installs are for the gaming category and we followed by communication and tools.

6.Content Rating for Various groups



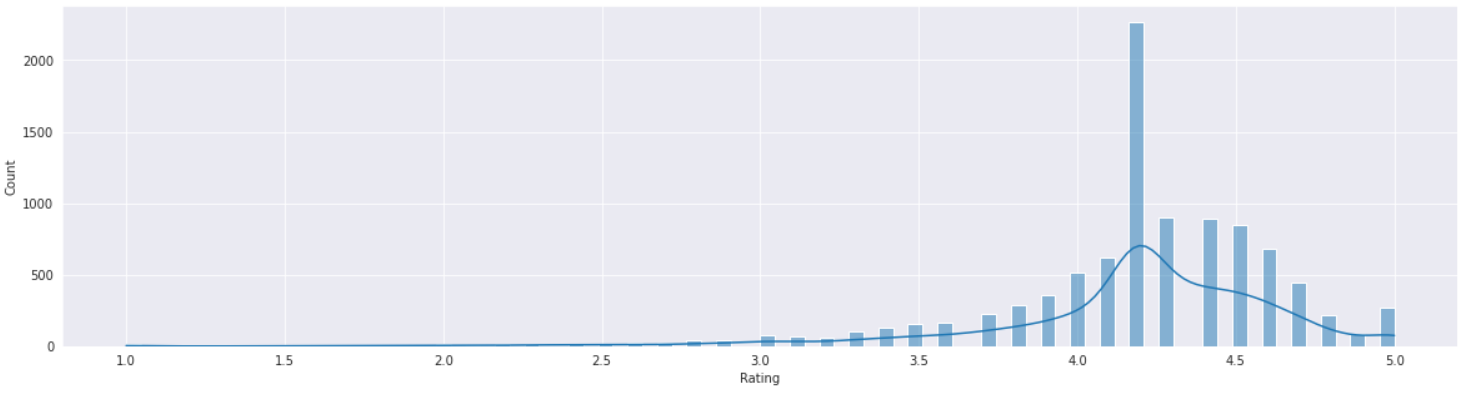
Pie plots are usually best for sharing visually distribution of data. We found that most of the apps are open to be reviewed by Everyone.

7. Paid app counts with respect to price (excluding free apps)



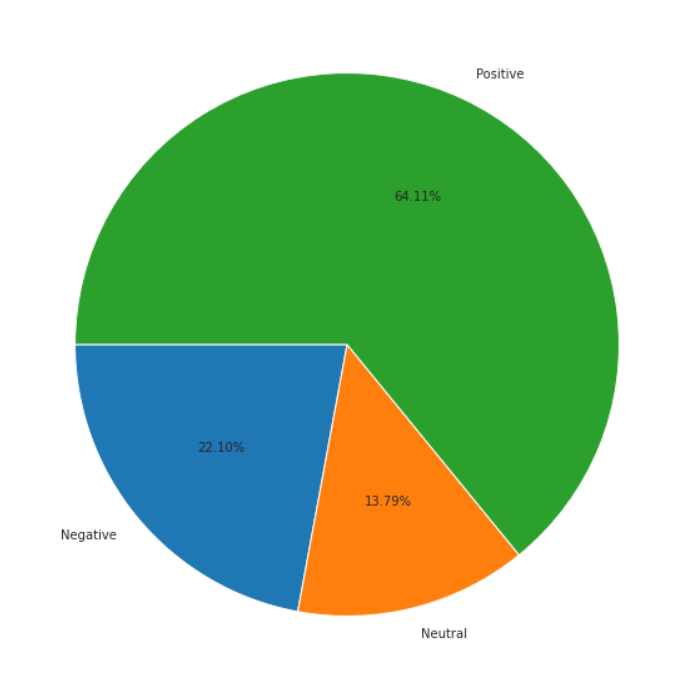
We see that max app belong to price = 0.99 Followed by 2.99, So it will be a good idea for the paid app developer to keep the price in between 0.99 to 2.99.

8.Distribution of ratings across various apps



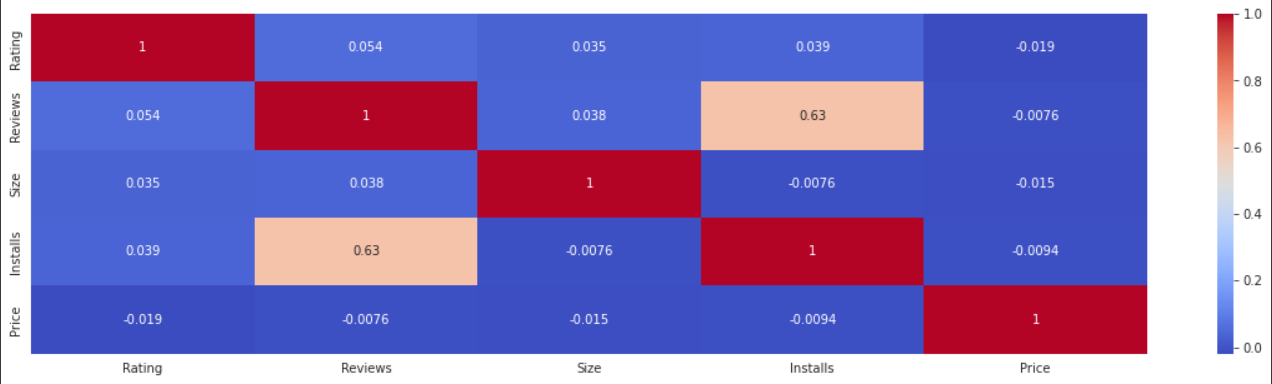
Hist plots are better to analyze distribution of numeric values. We see that most of the apps have a rating of 4.25. Since most of the apps have 4.2 rating, there is an opportunity to have a better app that could fit the market needs.

9.Sentiment Analysis



Pie charts are better to demonstrate the proportions of the values. we see that there are many apps that are having positive sentiments respect to various apps

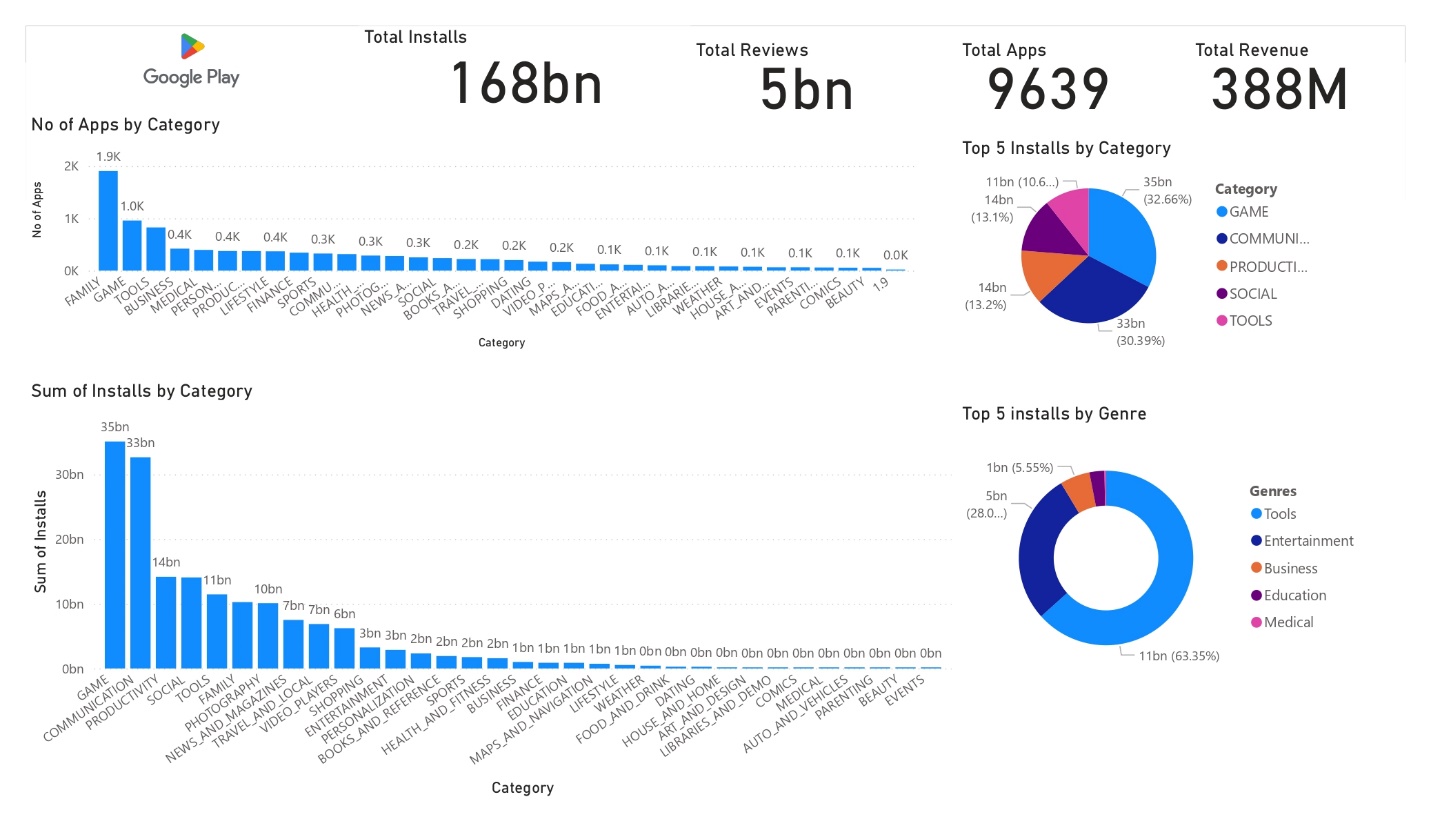
10.Correlation Heatmap



The values in the matrix range from -1 to 1, and represent the strength and direction of the correlation between two variables. A correlation coefficient of -1 indicates a perfect negative correlation, a coefficient of 0 indicates no correlation, and a coefficient of 1 indicates a perfect positive correlation. We see that the correlation between Installs and Price is 0.63 which is good, It means that when there are more reviews on any app the more people tend to install the app. so its a better idea to get reviews on the app.

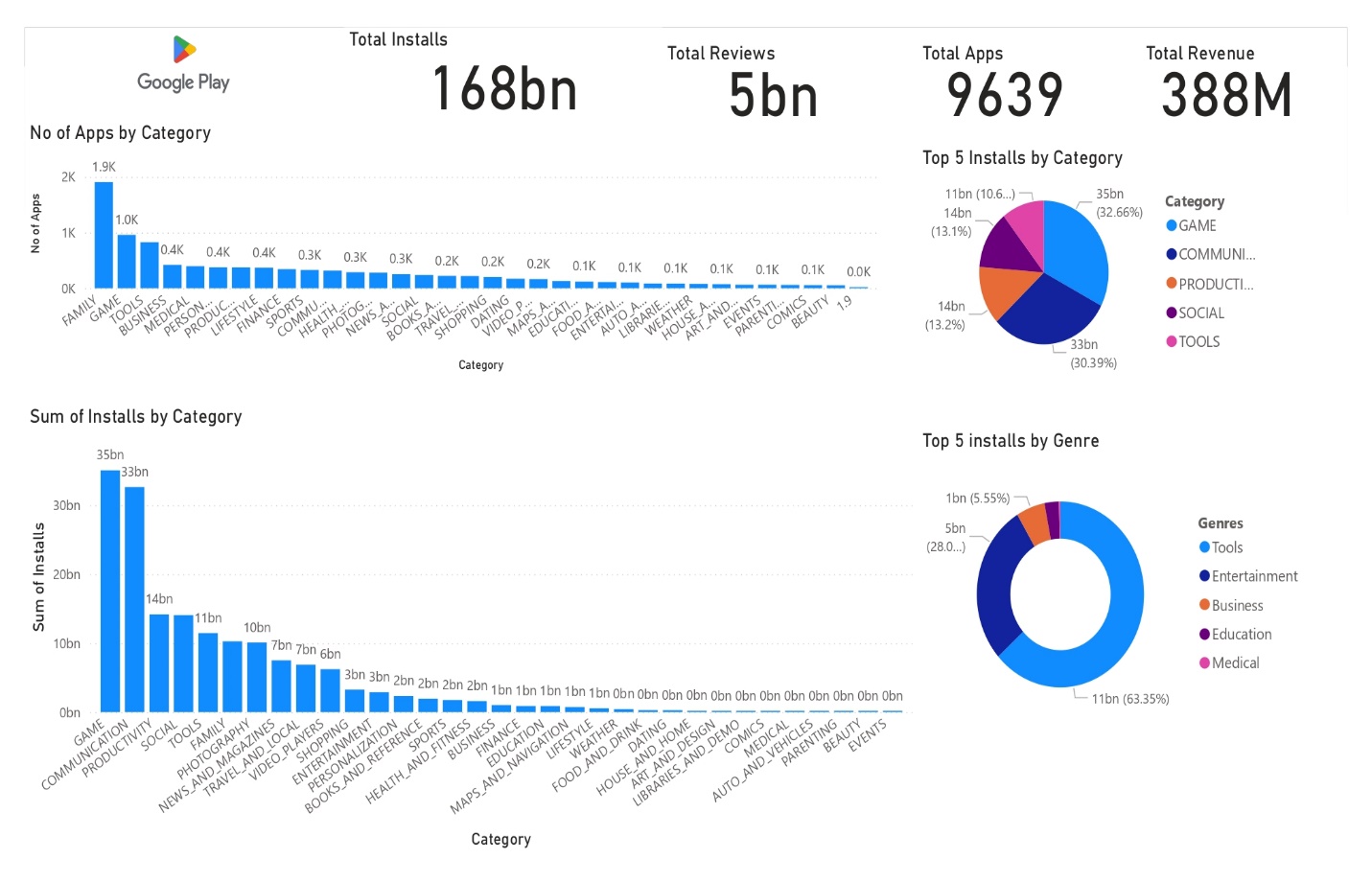
# 

# Dashboard



A dashboard is used to display meaningful insights from a data in the form of visualization, here various graphs are combined into a single sheet to give quick theme of the dataset, instead of going through each and every graph, we can look into a single dashboard and make decisions.

# 



All the individual graphs are combined on a single sheet to give brief and meaningful insights about all of the visualisation, along with some additional data like Total installs, Total Reviews, Total apps, Total revenue.