

# Data Visualization on Google Play Store Apps

Dataset has records for each application present in google play store till August 2018. Each record provides information about the app like app size, ratings for the app, number of installs, category it belongs and more; we have total of 13 variables for each of the application.

**Source:** <https://www.kaggle.com/lava18/google-play-store-apps>

## Visualization Workflow

**1.Agenda:** To provide the users with relatable information about the android market of google play store and give insights for developers on areas which needs improvement.

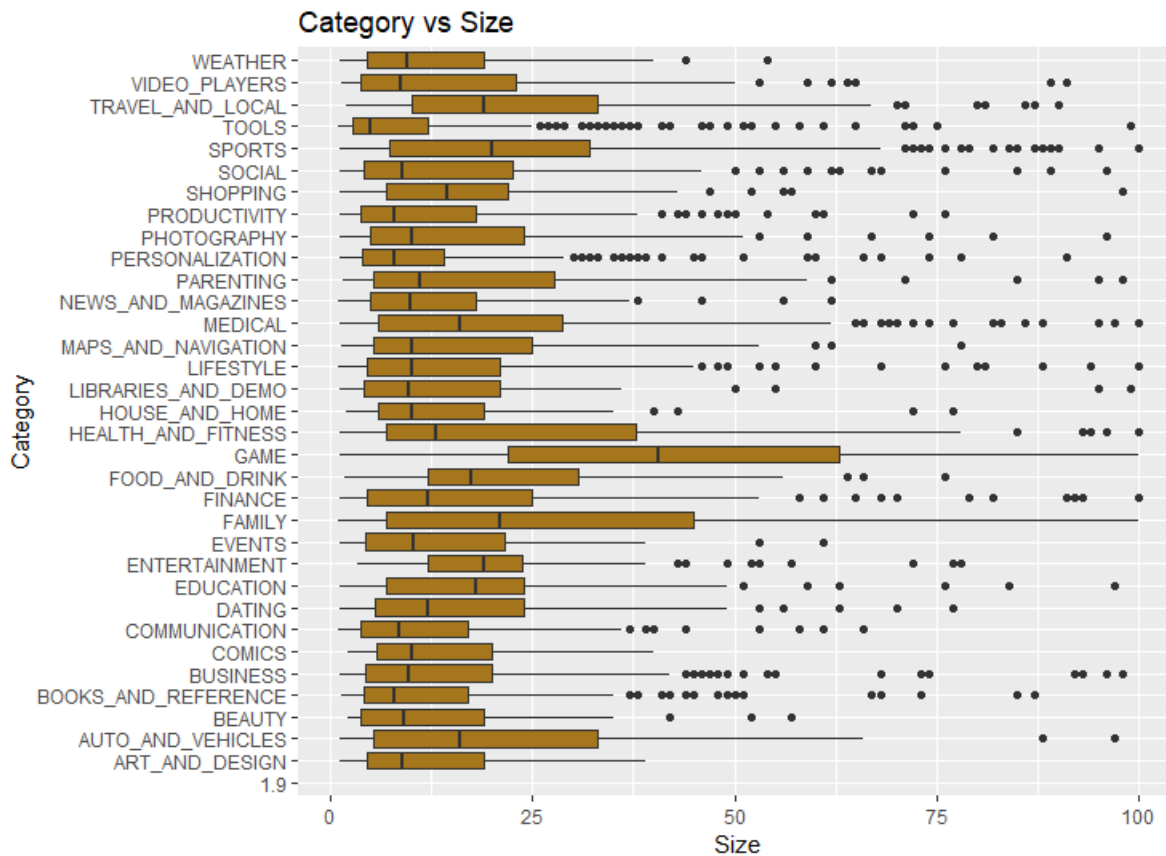
**2.Refining Data:** Necessary formatting has been done to the data set, which helped us in smooth plotting in R .

**3.Editorial thinking:** We want the audience to benefit by knowing the information about each app and which is doing better, and which is not. Additionally, the developers may also focus on the apps that need improvement based on the reviews and ratings on the apps. To achieve the same, we have considered variables relationship and depicted the correlation.

**4.Design Solution:** We have come up with 3 static, 3 animated and 2 interactive charts to narrate the story of GooglePlayStoreApps

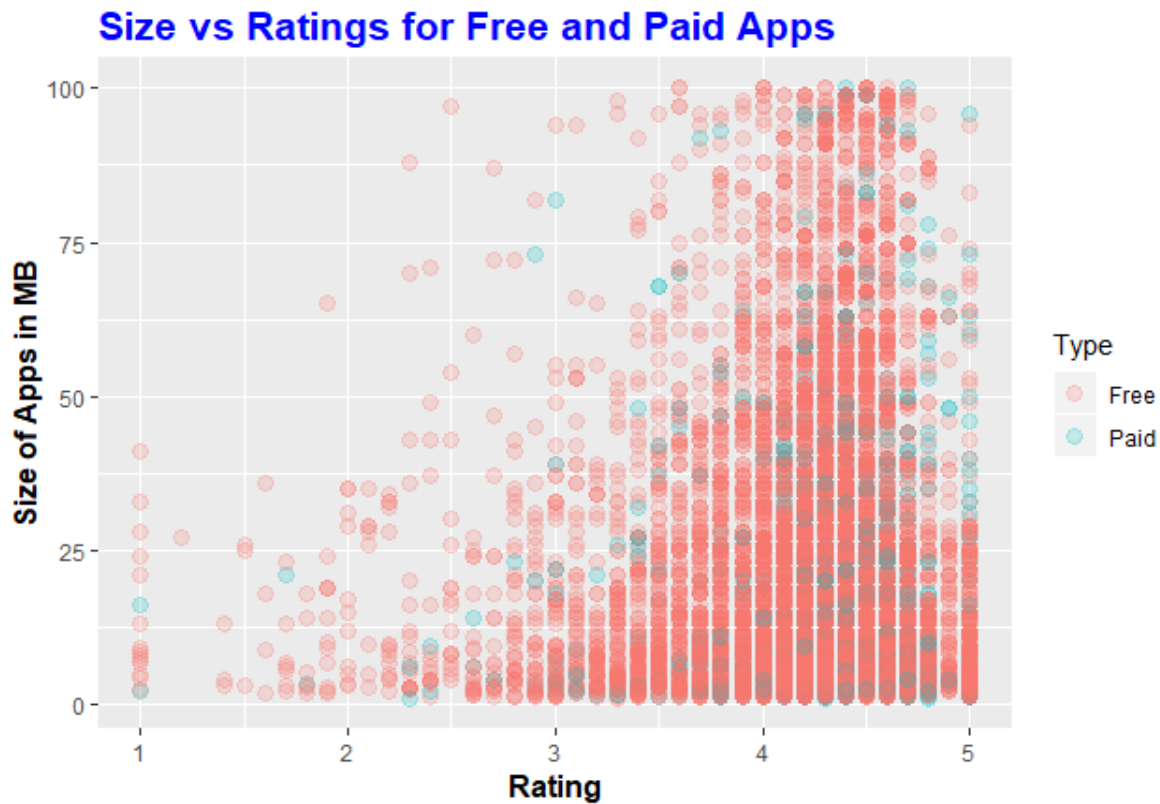
## Static Charts:

### 1.Category vs Size



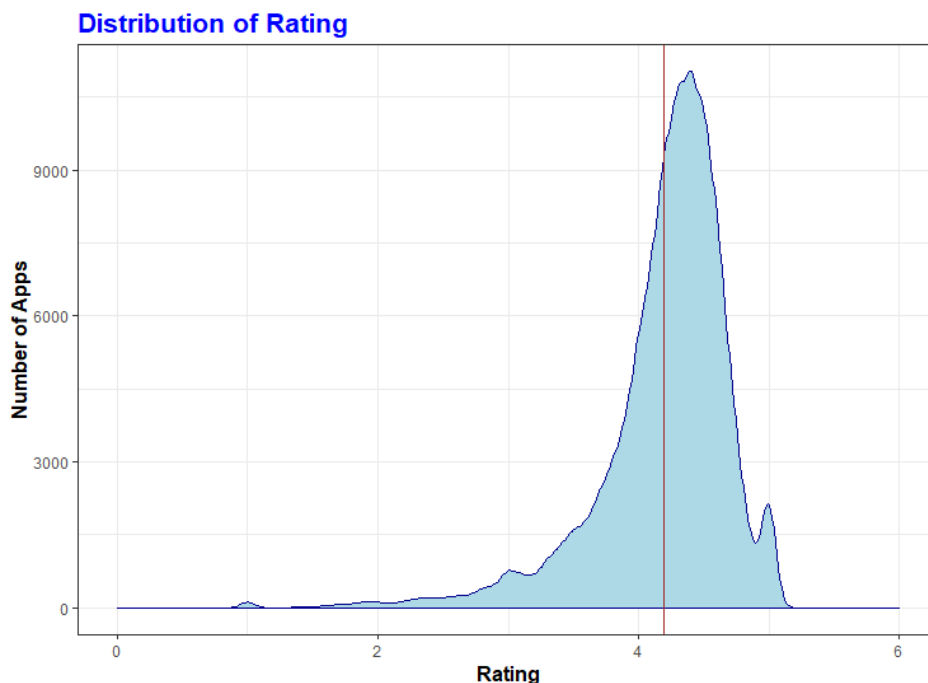
- ❖ This chart talks about the categories and the sizes associated with each category. The boxplot tells us the mean of each category's size and other related statistics. By observing we get to know that the category Game requires more memory followed by the category Family

## 2. Distribution of Ratings wrt Size of Apps



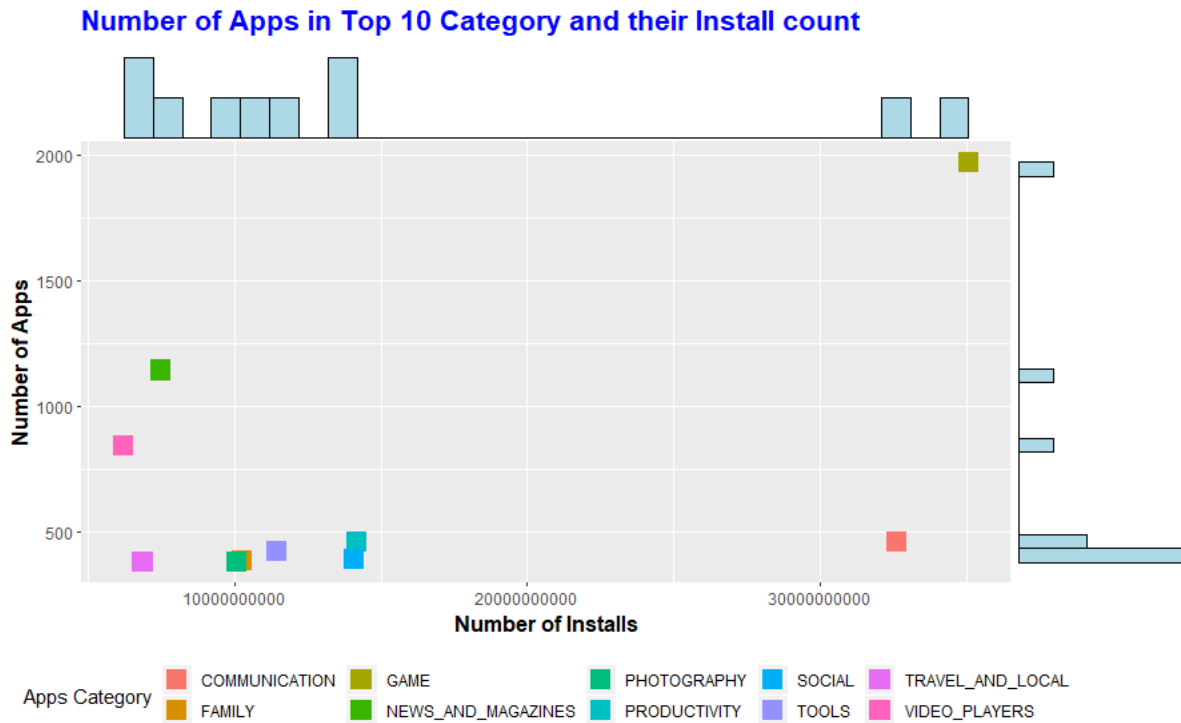
- ❖ This scatter plot talks about the distribution of ratings with respect to size of the apps. The color variation is shown to distinguish between free and paid apps. The observation made from this plot is that paid apps have better review than free apps and apps < size 65MB have better review.

## 3. Distribution of Ratings



- ❖ This graph shows the distribution of ratings given by user to apps in general. The average rating according to the dataset is 4.19 which is relatively high.

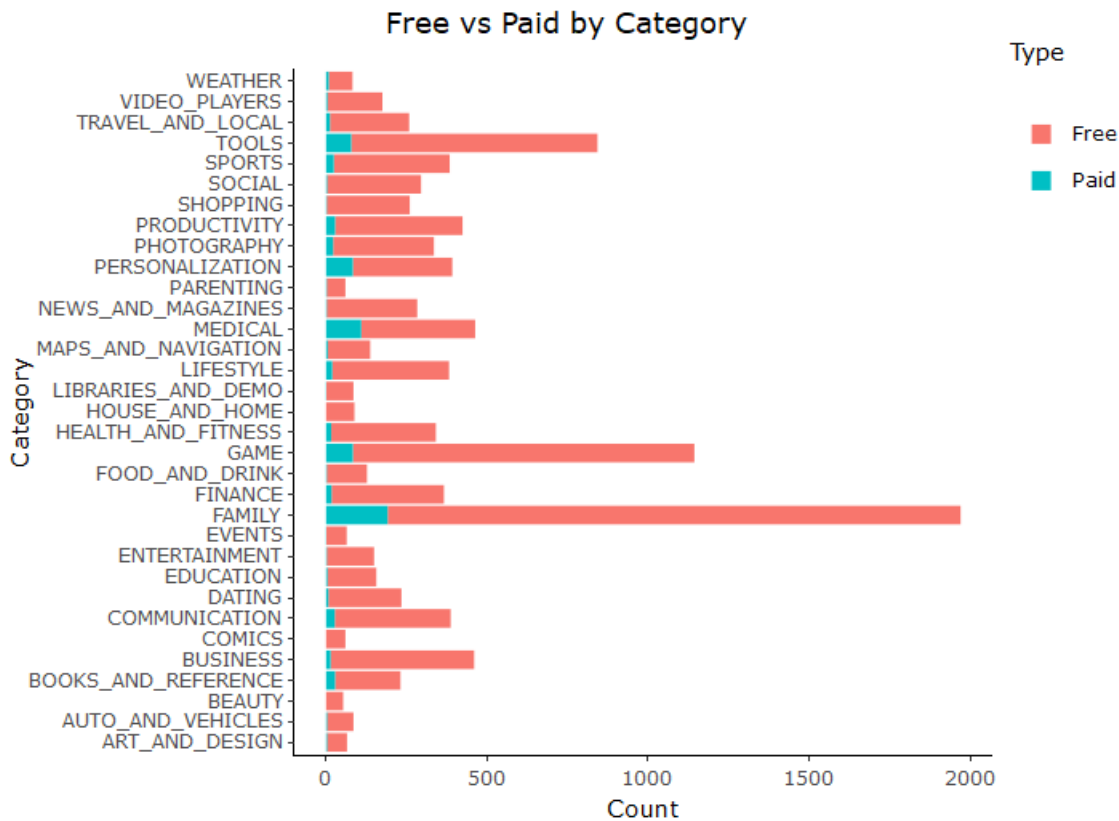
#### 4. Top 10 Category Apps Count and Downloads



- ❖ The categories with more downloads such as Communication and Productivity have less number of apps compared to categories with less less downloads (News and Magazines and video players). Developers can target categories which have more installs, for new apps development.

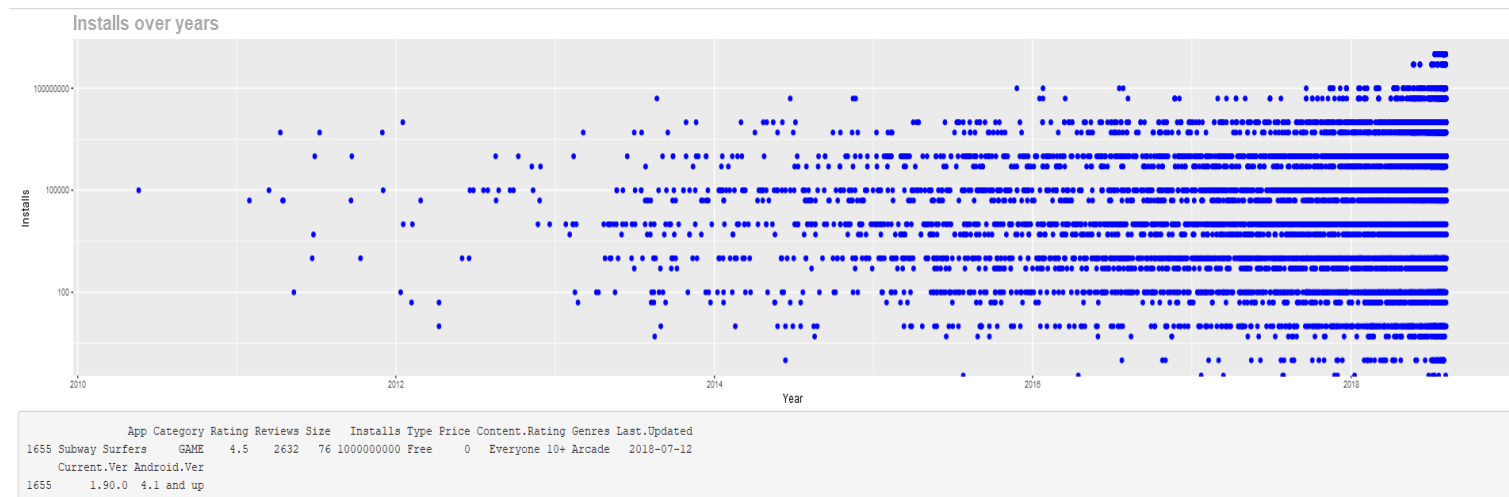
# Interactive Charts

## 1.Free vs Paid by Category



- ❖ This stacked bar graph shows the count of free apps vs paid apps for each category. The insights for this graph are that the category Family has the highest no.of free apps and Beauty has less no.of free apps. Beauty and House\_and\_Home are the only categories with no paid apps. Family has highest paid apps.

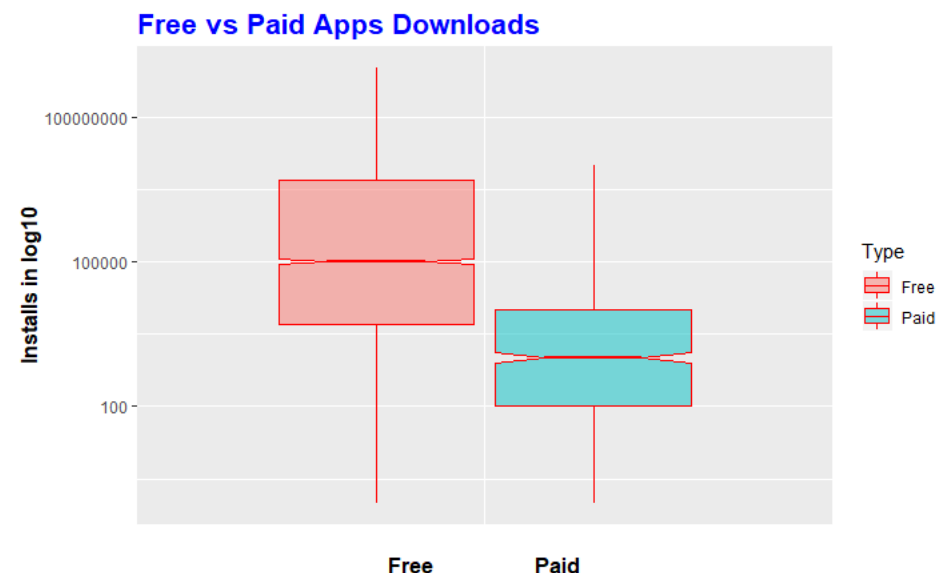
## 2.Installs vs Last.Updated



- ❖ This plot tells us about the Installations that happened over the years(2010-2018). Since it's a lot of data to plot, hover interaction has been implemented. When the cursor is hovered on a point it gives us the entire row details corresponding to that datapoint. In the snapshot we can see that the details of the app subway surfer has been retrieved. This chart has been taken into consideration because it would be easy for the user to just hover on the datapoint to know the entire information about that particular app in one single chart rather than plotting multiple charts for it.

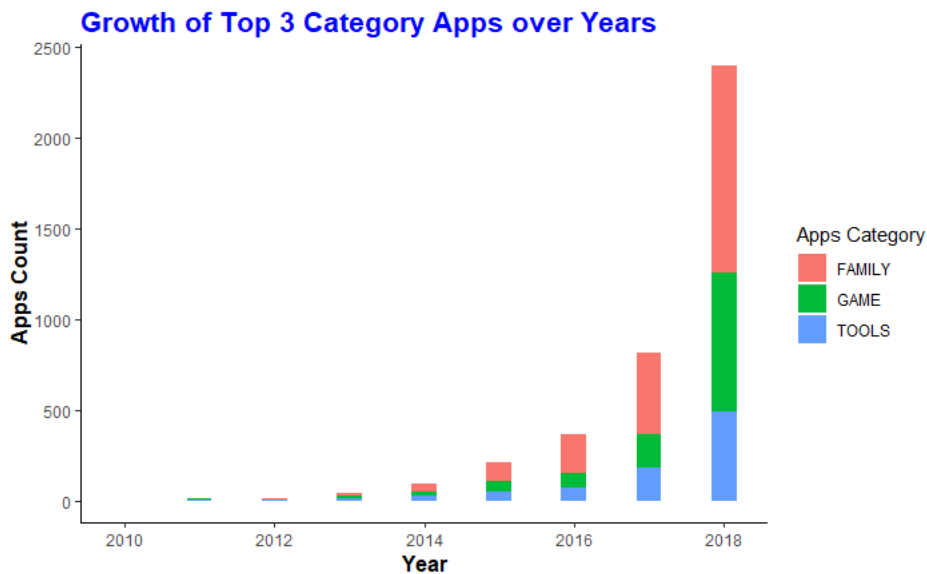
## 3.Animated Charts

### 1.Free vs Paid Apps Downloads



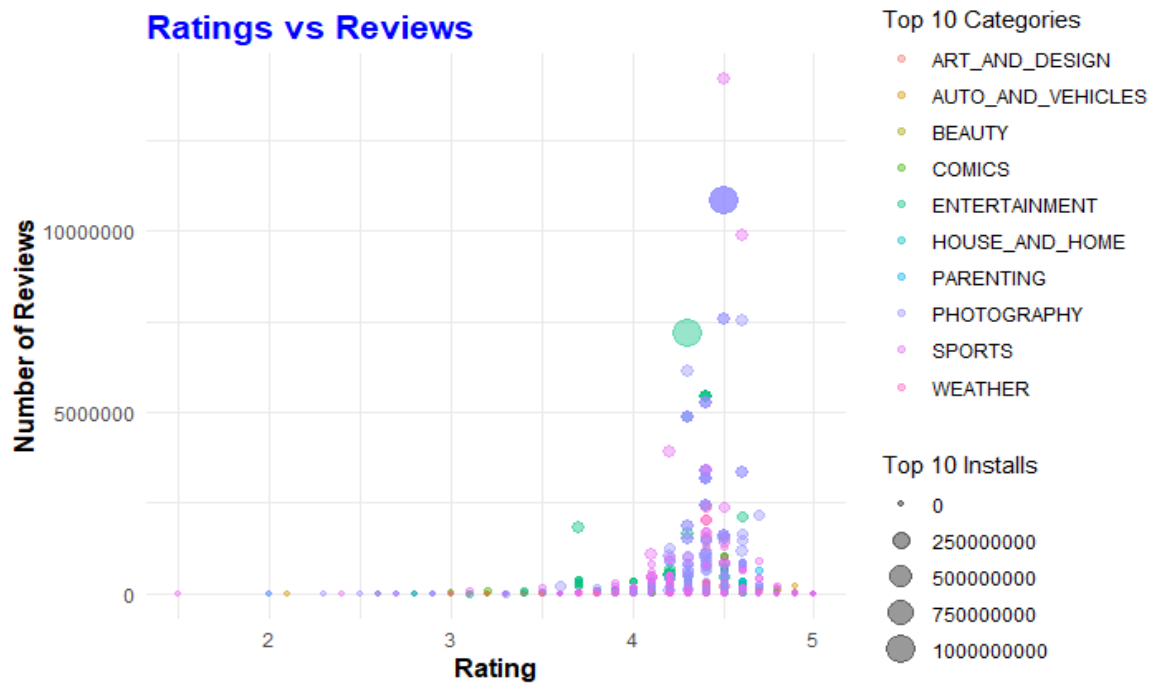
- ❖ This animated boxplot shows the variation between free and paid apps installed. The y-axis scale has been adjusted to support the graph. The animation shows the difference between the free and paid apps corresponding to the #installs. As we can see, the paid apps are downloaded less and even their means have more difference.

## 2. Growth of Top 3 Categories over Years



- ❖ This animated bar graph shows the progress of top 3 categories, based on apps count, over the years. The transition includes the rise and fall of bar graphs. The growth is exponential as expected.

## 3. Ratings vs Reviews for Top 10 Categories



- ❖ This bubble graph shows the transition of ratings vs reviews for top 10 categories. Apps with high # of installs have more reviews and which in turn can influence the customers to install the application.