

# Google Play Store Installs Report and Analysis

## Introduction

This report analyzes data collected on 10841 apps in the Google Play store. For this user story, I am reporting on whether a possible correlation can be seen when comparing installs to content rating.

## Body

### Data

This report was written using MyJupyterNotebooks. First I imported Pandas, Matplotlib, Seaborn, and Numpy to analyze and write reports on the data. Then I imported the Google Play app data via .csv spreadsheet. Next, I requested the number of rows and columns (10841 and 13), the data types of each column (all were strings or "object"), and the column names ('App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type', 'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver', and 'Android Ver').

### Method

Since I only need category, content rating, and installs for this analysis, I dropped the other columns. Next, I needed to reformat the Installs category so that it could be read as a float or decimal number; to do this, I removed the symbols '+', ',', and ' ' from the column. Next I converted the column to a float. Finally, I changed the column name of "Content Rating" to "Rating" as the space in the middle of the column was causing a formatting issue.

### Results

After creating a histogram and finding the mean and sum of the data, I found that apps in the "Everyone" content rating category had the most installs at 3.988306e+20 total installs. This number is approximated due to the fact that the original dataset rounded numbers in the Installs column (e.g. if one app had 10,827 installs, it is represented as 10,000+).

## Analysis

Histogram showing count of all installs for all apps in each content rating:

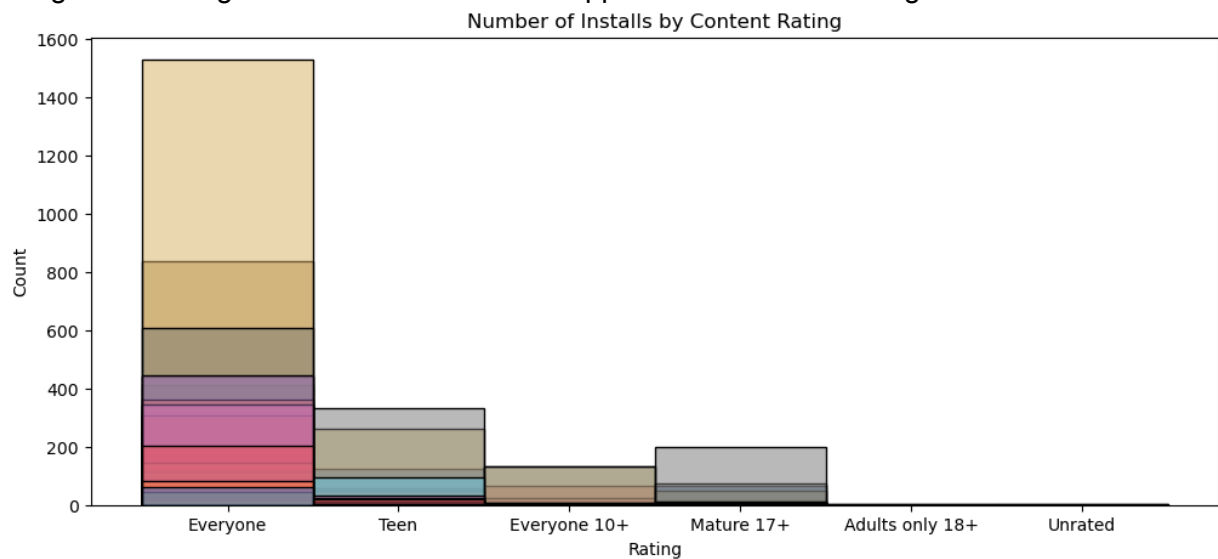
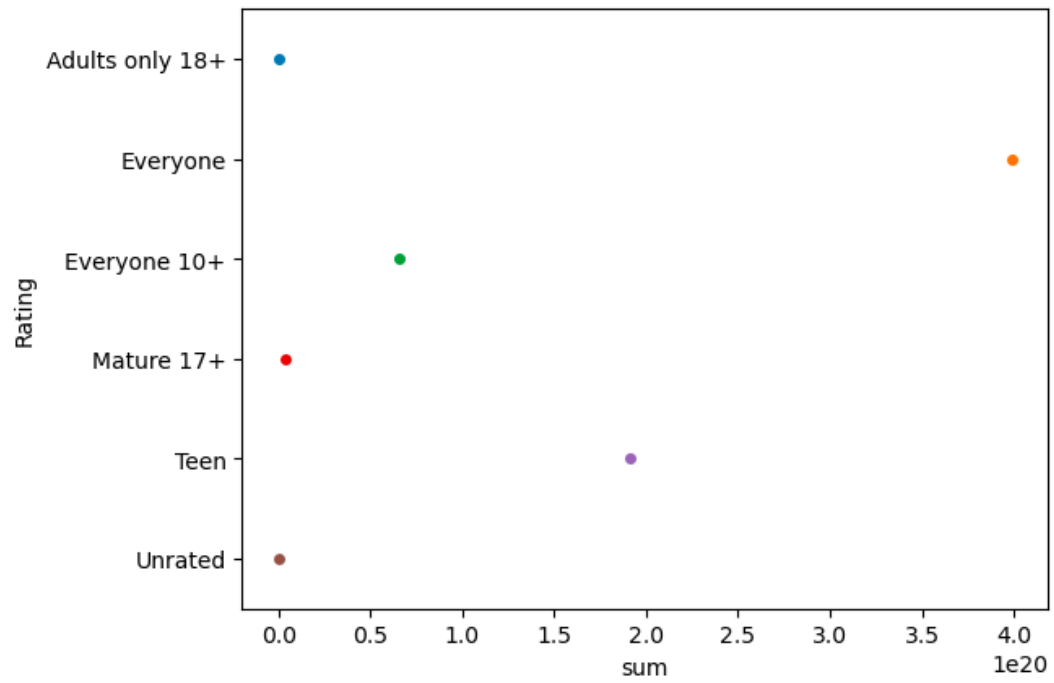


Table showing the sum and mean of installs in each content rating:

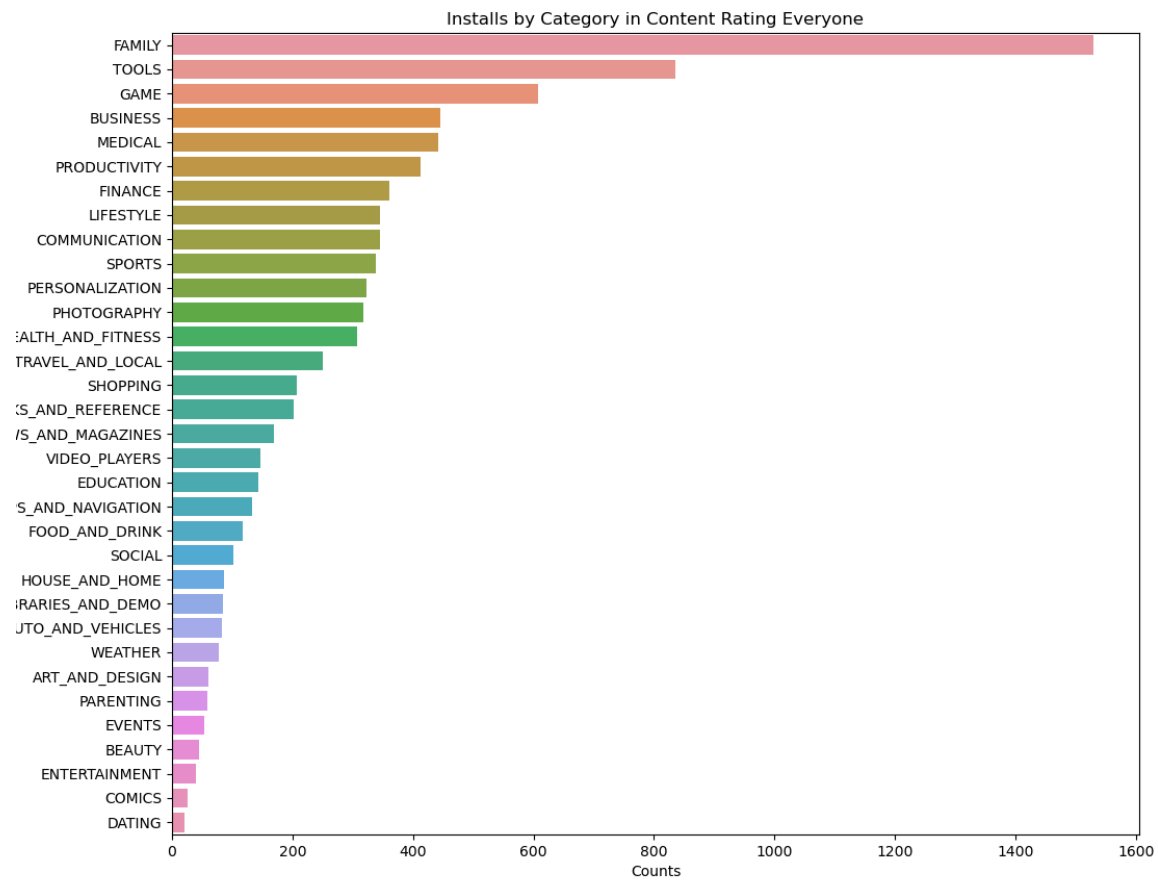
```
sumTotal=data.groupby('Rating')['Installs'].agg(['sum', 'mean'])
sumTotal
```

	sum	mean
Rating		
Adults only 18+	1.100000e+13	3.666667e+12
Everyone	3.988306e+20	4.576895e+16
Everyone 10+	6.555855e+19	1.583540e+17
Mature 17+	3.947469e+18	7.910760e+15
Teen	1.913141e+20	1.583726e+17
Unrated	5.000500e+09	2.500250e+09

Scatterplot showing sum of all installs in each content rating:



Bar chart showing installs by category in the Everyone content rating:



## Conclusion

Out of 10,841 total apps analyzed, apps rated “Everyone” had the highest total installs. It belongs to 33 categories, and here is the count breakdown by category:

Category	Counts
FAMILY	1529
TOOLS	835
GAME	608
BUSINESS	445
MEDICAL	442
PRODUCTIVITY	413
FINANCE	361
LIFESTYLE	345
COMMUNICATION	345
SPORTS	338
PERSONALIZATION	322
PHOTOGRAPHY	318
HEALTH_AND_FITNESS	307
TRAVEL_AND_LOCAL	251
SHOPPING	208
BOOKS_AND_REFERENCE	202
NEWS_AND_MAGAZINES	169
VIDEO_PLAYERS	146
EDUCATION	144
MAPS_AND_NAVIGATION	133
FOOD_AND_DRINK	117
SOCIAL	101
HOUSE_AND_HOME	86
LIBRARIES_AND_DEMO	85
AUTO_AND_VEHICLES	83
WEATHER	78
ART_AND_DESIGN	61
PARENTING	58
EVENTS	53
BEAUTY	45
ENTERTAINMENT	39
COMICS	26
DATING	21