Analyzing Mobile App Data for Increased User Engagement

In this project, we will be working as analyzing data for our company that develops Android and iOS mobile applications. These apps are available for free on Google Play and the App Store, and our primary source of revenue is through in-app advertisements. As the number of users directly impacts our revenue, it is crucial to understand what types of apps are more likely to attract a larger user base.

The goal of this project is to analyze data and provide insights to our development team, helping them make informed decisions about the types of apps they should focus on creating. By identifying the characteristics and trends of popular apps, we can guide our developers towards building applications that have a higher likelihood of engaging more users, ultimately leading to increased revenue for the company.

In [32]:
```
import csv

with open('AppleStore.csv', encoding='utf8') as file:
    csv_reader = csv.reader(file)
    ios = list(csv_reader)
    ios_header = ios[0]
    ios_data = ios[1:]

with open('googleplaystore.csv', encoding='utf8') as file:
    csv_reader = csv.reader(file)
    google = list(csv_reader)
    google_header = google[0]
    google_data = google[1:]
```

Initial exploration of the data

In [33]:
```
def explore_data(dataset, start, end, rows_and_columns=False):
    dataset_slice = dataset[start:end]
    for row in dataset_slice:
        print(row)
        print('\n') # adds a new (empty) line after each row

    if rows_and_columns:
        print('Number of rows:', len(dataset))
        print('Number of columns:', len(dataset[0]))
```

Apple App Store Data

In [34]:
```
print(ios_header)
print('\n')
```

```
explore_data(ios_data, 0, 3, True)
```

```
['id', 'track_name', 'size_bytes', 'currency', 'price', 'rating_count_tot',
'rating_count_ver', 'user_rating', 'user_rating_ver', 'ver', 'cont_rating',
'prime_genre', 'sup_devices.num', 'ipadSc_urls.num', 'lang.num', 'vpp_lic']


['284882215', 'Facebook', '389879808', 'USD', '0.0', '2974676', '212', '3.
5', '3.5', '95.0', '4+', 'Social Networking', '37', '1', '29', '1']


['389801252', 'Instagram', '113954816', 'USD', '0.0', '2161558', '1289', '4.
5', '4.0', '10.23', '12+', 'Photo & Video', '37', '0', '29', '1']


['529479190', 'Clash of Clans', '116476928', 'USD', '0.0', '2130805', '579',
'4.5', '4.5', '9.24.12', '9+', 'Games', '38', '5', '18', '1']


Number of rows: 7197
Number of columns: 16
```

Google Play Store Data

In [35]:
```
print(google_header)
print('\n')
explore_data(google_data, 0, 4, True)
```

```
['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type', 'Pric
e', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver', 'Android Ve
r']


['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']


['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']


['U Launcher Lite – FREE Live Cool Themes, Hide Apps', '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']


['Sketch - Draw & Paint', 'ART_AND_DESIGN', '4.5', '215644', '25M', '50,000,
000+', 'Free', '0', 'Teen', 'Art & Design', 'June 8, 2018', 'Varies with dev
ice', '4.2 and up']


Number of rows: 10841
Number of columns: 13
```

Begin data cleaning

Compare the incorrect row with a correct one

In [36]:
```
print(google[10473]) #Incorrect row
print('\n')
print(google_header) #Header
print('\n')
print(google[1]) #Correct row
```

```
['Life Made WI-Fi Touchscreen Photo Frame', '1.9', '19', '3.0M', '1,000+',
'Free', '0', 'Everyone', '', 'February 11, 2018', '1.0.19', '4.0 and up']


['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type', 'Pric
e', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver', 'Android Ve
r']


['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']
```

In [37]:
```
print(google[10473])
```

```
['Life Made WI-Fi Touchscreen Photo Frame', '1.9', '19', '3.0M', '1,000+',
'Free', '0', 'Everyone', '', 'February 11, 2018', '1.0.19', '4.0 and up']
```

Delete the incorrect row

In [38]:
```
print(len(google))
print(len(google_data))
print('\n')

print(google[10472])  # Incorrect row in google list
print(google_data[10472])  # Incorrect row in google_data list
print('\n')

del google[10472]  # Remove the incorrect row from google list
del google_data[10472]  # Remove the incorrect row from google_data list

print(len(google))
print(len(google_data))
```

```
10842
10841


['Xposed Wi-Fi-Pwd', 'PERSONALIZATION', '3.5', '1042', '404k', '100,000+',
'Free', '0', 'Everyone', 'Personalization', 'August 5, 2014', '3.0.0', '4.0.
3 and up']
['Life Made WI-Fi Touchscreen Photo Frame', '1.9', '19', '3.0M', '1,000+',
'Free', '0', 'Everyone', '', 'February 11, 2018', '1.0.19', '4.0 and up']


10841
10840
```

In [39]:
```
print(google[10473])
print(google_data[10473])
```

```
['osmino Wi-Fi: free WiFi', 'TOOLS', '4.2', '134203', '4.1M', '10,000,000+',
'Free', '0', 'Everyone', 'Tools', 'August 7, 2018', '6.06.14', '4.4 and up']
['Sat-Fi Voice', 'COMMUNICATION', '3.4', '37', '14M', '1,000+', 'Free', '0',
'Everyone', 'Communication', 'November 21, 2014', '2.2.1.5', '2.2 and up']
```

Identify and remove duplicates

In [40]:
```
duplicate_apps = []
unique_apps = []

for app in google:
    name = app[0]
    if name in unique_apps:
        duplicate_apps.append(name)
    else:
        unique_apps.append(name)
print('Number of duplicate apps:', len(duplicate_apps))
print('\n')
print('Example duplicates identified:', duplicate_apps[:10])
```

```
Number of duplicate apps: 1181


Example duplicates identified: ['Quick PDF Scanner + OCR FREE', 'Box', 'Goog
le My Business', 'ZOOM Cloud Meetings', 'join.me - Simple Meetings', 'Box',
'Zenefits', 'Google Ads', 'Google My Business', 'Slack']
```

Remove duplicates on the basis of ratings, removing all entries except those with the most ratings, thereby indicating these entries are the most recent

In [41]:
```
reviews_max = {}

for app in google_data:
    name = app[0]
    n_reviews = float(app[3])

    if name in reviews_max and reviews_max[name] < n_reviews:
```

```
        reviews_max[name] = n_reviews

    elif name not in reviews_max:
        reviews_max[name] = n_reviews
```

In [42]:
```
print('Expected length:', len(google_data) - 1181)
print('Actual length:', len(reviews_max))
```

```
Expected length: 9659
Actual length: 9659
```

In [43]:
```
android_clean = []
already_added = []

for app in google_data:
    name = app[0]
    n_reviews = float(app[3])

    if (reviews_max[name] == n_reviews) and (name not in already_added):
        android_clean.append(app)
        already_added.append(name)
```

In [44]:
```
explore_data(android_clean, 0, 3, True)
```

```
['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']


['U Launcher Lite – FREE Live Cool Themes, Hide Apps', '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']


['Sketch - Draw & Paint', 'ART_AND_DESIGN', '4.5', '215644', '25M', '50,000,
000+', 'Free', '0', 'Teen', 'Art & Design', 'June 8, 2018', 'Varies with dev
ice', '4.2 and up']


Number of rows: 9659
Number of columns: 13
```

Remove non-English apps from the datasets

In [45]:
```
print(ios[813][1])
print(ios[6731][1])

print(android_clean[4412][0])
print(android_clean[7940][0])
```

```
BATTLE BEARS -1
Beast Poker
中国語 AQリスニング
لعبة تقدر تربح DZ
```

In [46]:
```
def is_english(subject):
    non_ascii = 0

    for char in subject:
        if ord(char) > 127:
            non_ascii += 1

    if non_ascii > 3:
            return False
    else:
        return True
```

In [47]:
```
print(is_english('Instagram'))
print(is_english('Docs To Go™ Free Office Suite'))
print(is_english('Instachat 😜'))
print(is_english('爱奇艺PPS -《欢乐颂2》电视剧热播'))
```

```
True
True
True
False
```

In [48]:
```
english_android = []
english_ios = []

for app in android_clean:
    name = app[0]

    if is_english(name):
        english_android.append(app)

for app in ios_data:
    name = app[1]

    if is_english(name):
        english_ios.append(app)

explore_data(english_android, 0, 3, True)
print('\n')
explore_data(english_ios, 0, 3, True)

print('\n')

print('Total iOS apps in English: ' + str(len(english_ios)))
print('Total Android apps in English: ' + str(len(english_android)))
```

```
['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']


['U Launcher Lite – FREE Live Cool Themes, Hide Apps', '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']


['Sketch - Draw & Paint', 'ART_AND_DESIGN', '4.5', '215644', '25M', '50,000,
000+', 'Free', '0', 'Teen', 'Art & Design', 'June 8, 2018', 'Varies with dev
ice', '4.2 and up']


Number of rows: 9614
Number of columns: 13


['284882215', 'Facebook', '389879808', 'USD', '0.0', '2974676', '212', '3.
5', '3.5', '95.0', '4+', 'Social Networking', '37', '1', '29', '1']


['389801252', 'Instagram', '113954816', 'USD', '0.0', '2161558', '1289', '4.
5', '4.0', '10.23', '12+', 'Photo & Video', '37', '0', '29', '1']


['529479190', 'Clash of Clans', '116476928', 'USD', '0.0', '2130805', '579',
'4.5', '4.5', '9.24.12', '9+', 'Games', '38', '5', '18', '1']


Number of rows: 6183
Number of columns: 16


Total iOS apps in English: 6183
Total Android apps in English: 9614
```

Isolate free apps

In [49]:

```
free_android = []
free_ios = []

for app in english_android:
    if app[7] == '0':
        free_android.append(app)

for app in english_ios:
    if app[4] == '0.0':
        free_ios.append(app)

print('Free iOS apps: ' + str(len(free_ios)))
print('Free Android apps: ' + str(len(free_android)))
```

```
Free iOS apps: 3222
Free Android apps: 8864
```

Identify app profiles that are popular in both the Google Play and Apple Stores

In this analysis, we explore the frequency tables for the prime_genre column of the App Store dataset and the Category and Genres columns of the Google Play dataset to gain insights into the app markets.

In [50]:
```
def freq_table(dataset, index):
    table = {}
    total = 0

    for row in dataset:
        total += 1
        value = row[index]
        if value in table:
            table[value] += 1
        else:
            table[value] = 1

    table_percentages = {}
    for key in table:
        percentage = (table[key] / total) * 100
        table_percentages[key] = percentage

    return table_percentages


def display_table(dataset, index):
    table = freq_table(dataset, index)
    table_display = []
    for key in table:
        key_val_as_tuple = (table[key], key)
        table_display.append(key_val_as_tuple)

    table_sorted = sorted(table_display, reverse = True)
    for entry in table_sorted:
        print(entry[1], ':', entry[0])
```

App Store Analysis

Let's analyze the frequency table for the prime_genre column of the App Store dataset.

In [51]:
```
display_table(free_ios, -5)
```

```
Games : 58.16263190564867
Entertainment : 7.883302296710118
Photo & Video : 4.9658597144630665
Education : 3.662321539416512
Social Networking : 3.2898820608317814
Shopping : 2.60707635009311
Utilities : 2.5139664804469275
Sports : 2.1415270018621975
Music : 2.0484171322160147
Health & Fitness : 2.0173805090006205
Productivity : 1.7380509000620732
Lifestyle : 1.5828677839851024
News : 1.3345747982619491
Travel : 1.2414649286157666
Finance : 1.1173184357541899
Weather : 0.8690254500310366
Food & Drink : 0.8069522036002483
Reference : 0.5586592178770949
Business : 0.5276225946617008
Book : 0.4345127250155183
Navigation : 0.186219739292365
Medical : 0.186219739292365
Catalogs : 0.12414649286157665
```

Findings:

The most common genre is Games at 58.2%, followed by Entertainment at 7.9%.

There is a significant drop-off after the Games genre.

Practical app genres such as Education (3.7%), Shopping (2.6%), Utilities (2.5%), and Productivity (1.7%) have lower percentages compared to entertainment-related genres.

For this reason, it appears that the App Store is dominated by apps designed for entertainment purposes, with Games being the most popular genre by a large margin.

Based on this frequency table, it is difficult to recommend a specific app profile for the App Store market, as while Games have the highest frequency, this does not necessarily imply that apps in this genre have the largest user base or are the most profitable.

Google Play Analysis

Now, let's analyze the frequency table for the Category column of the Google Play dataset.

In [52]:
```
display_table(free_android, 1)
```

```
FAMILY : 18.907942238267147
GAME : 9.724729241877256
TOOLS : 8.461191335740072
BUSINESS : 4.591606498194946
LIFESTYLE : 3.9034296028880866
PRODUCTIVITY : 3.892148014440433
FINANCE : 3.7003610108303246
MEDICAL : 3.531137184115524
SPORTS : 3.395758122743682
PERSONALIZATION : 3.3167870036101084
COMMUNICATION : 3.2378158844765346
HEALTH_AND_FITNESS : 3.0798736462093865
PHOTOGRAPHY : 2.944494584837545
NEWS_AND_MAGAZINES : 2.7978339350180503
SOCIAL : 2.6624548736462095
TRAVEL_AND_LOCAL : 2.33528880866426
SHOPPING : 2.2450361010830324
BOOKS_AND_REFERENCE : 2.1435018050541514
DATING : 1.861462093862816
VIDEO_PLAYERS : 1.7937725631768955
MAPS_AND_NAVIGATION : 1.3989169675090252
FOOD_AND_DRINK : 1.2409747292418771
EDUCATION : 1.1620036101083033
ENTERTAINMENT : 0.9589350180505415
LIBRARIES_AND_DEMO : 0.9363718411552346
AUTO_AND_VEHICLES : 0.9250902527075812
HOUSE_AND_HOME : 0.8235559566787004
WEATHER : 0.8009927797833934
EVENTS : 0.7107400722021661
PARENTING : 0.6543321299638989
ART_AND_DESIGN : 0.6430505415162455
COMICS : 0.6204873646209386
BEAUTY : 0.5979241877256317
```

Google Play Category Analysis

Based on the frequency table for the Category column, we can make the following observations:

1. The most common app categories in the Google Play dataset are FAMILY (18.9%), GAME (9.7%), and TOOLS (8.5%).

2. Practical app categories such as BUSINESS (4.6%), MEDICAL (3.5%) PRODUCTIVITY (3.9%), and FINANCE (3.7%) have lower percentages compared to the top categories.

3. The distribution of app categories is more evenly spread out compared to the App Store dataset, with no single dominant category.

4. Lifestyle and personal interest categories like LIFESTYLE (3.9%), SPORTS (3.4%), PERSONALIZATION (3.3%), HEALTH_AND_FITNESS (3.1%), and PHOTOGRAPHY (2.9%) have a notable presence.

5. The frequency table provides insights into the distribution of app categories but does not necessarily indicate the most profitable or popular app genres.

And now let's examine the Genres column

In [53]:
```
display_table(free_android, -4)
```

Tools : 8.449909747292418
Entertainment : 6.069494584837545
Education : 5.347472924187725
Business : 4.591606498194946
Productivity : 3.892148014440433
Lifestyle : 3.892148014440433
Finance : 3.7003610108303246
Medical : 3.531137184115524
Sports : 3.463447653429603
Personalization : 3.3167870036101084
Communication : 3.2378158844765346
Action : 3.1024368231046933
Health & Fitness : 3.0798736462093865
Photography : 2.944494584837545
News & Magazines : 2.7978339350180503
Social : 2.6624548736462095
Travel & Local : 2.3240072202166067
Shopping : 2.2450361010830324
Books & Reference : 2.1435018050541514
Simulation : 2.0419675090252705
Dating : 1.861462093862816
Arcade : 1.8501805054151623
Video Players & Editors : 1.7712093862815883
Casual : 1.7599277978339352
Maps & Navigation : 1.3989169675090252
Food & Drink : 1.2409747292418771
Puzzle : 1.128158844765343
Racing : 0.9927797833935018
Role Playing : 0.9363718411552346
Libraries & Demo : 0.9363718411552346
Auto & Vehicles : 0.9250902527075812
Strategy : 0.9138086642599278
House & Home : 0.8235559566787004
Weather : 0.8009927797833934
Events : 0.7107400722021661
Adventure : 0.6768953068592057
Comics : 0.6092057761732852
Beauty : 0.5979241877256317
Art & Design : 0.5979241877256317
Parenting : 0.4963898916967509
Card : 0.45126353790613716
Casino : 0.42870036101083037
Trivia : 0.41741877256317694
Educational;Education : 0.39485559566787
Board : 0.3835740072202166
Educational : 0.3722924187725632
Education;Education : 0.33844765342960287
Word : 0.2594765342960289
Casual;Pretend Play : 0.236913357400722
Music : 0.2030685920577617
Racing;Action & Adventure : 0.16922382671480143
Puzzle;Brain Games : 0.16922382671480143
Entertainment;Music & Video : 0.16922382671480143
Casual;Brain Games : 0.13537906137184114
Casual;Action & Adventure : 0.13537906137184114
Arcade;Action & Adventure : 0.12409747292418773

Action;Action & Adventure : 0.10153429602888085
Educational;Pretend Play : 0.09025270758122744
Simulation;Action & Adventure : 0.078971119133574
Parenting;Education : 0.078971119133574
Entertainment;Brain Games : 0.078971119133574
Board;Brain Games : 0.078971119133574
Parenting;Music & Video : 0.06768953068592057
Educational;Brain Games : 0.06768953068592057
Casual;Creativity : 0.06768953068592057
Art & Design;Creativity : 0.06768953068592057
Education;Pretend Play : 0.056407942238267145
Role Playing;Pretend Play : 0.04512635379061372
Education;Creativity : 0.04512635379061372
Role Playing;Action & Adventure : 0.033844765342960284
Puzzle;Action & Adventure : 0.033844765342960284
Entertainment;Creativity : 0.033844765342960284
Entertainment;Action & Adventure : 0.033844765342960284
Educational;Creativity : 0.033844765342960284
Educational;Action & Adventure : 0.033844765342960284
Education;Music & Video : 0.033844765342960284
Education;Brain Games : 0.033844765342960284
Education;Action & Adventure : 0.033844765342960284
Adventure;Action & Adventure : 0.033844765342960284
Video Players & Editors;Music & Video : 0.02256317689530686
Sports;Action & Adventure : 0.02256317689530686
Simulation;Pretend Play : 0.02256317689530686
Puzzle;Creativity : 0.02256317689530686
Music;Music & Video : 0.02256317689530686
Entertainment;Pretend Play : 0.02256317689530686
Casual;Education : 0.02256317689530686
Board;Action & Adventure : 0.02256317689530686
Video Players & Editors;Creativity : 0.01128158844765343
Trivia;Education : 0.01128158844765343
Travel & Local;Action & Adventure : 0.01128158844765343
Tools;Education : 0.01128158844765343
Strategy;Education : 0.01128158844765343
Strategy;Creativity : 0.01128158844765343
Strategy;Action & Adventure : 0.01128158844765343
Simulation;Education : 0.01128158844765343
Role Playing;Brain Games : 0.01128158844765343
Racing;Pretend Play : 0.01128158844765343
Puzzle;Education : 0.01128158844765343
Parenting;Brain Games : 0.01128158844765343
Music & Audio;Music & Video : 0.01128158844765343
Lifestyle;Pretend Play : 0.01128158844765343
Lifestyle;Education : 0.01128158844765343
Health & Fitness;Education : 0.01128158844765343
Health & Fitness;Action & Adventure : 0.01128158844765343
Entertainment;Education : 0.01128158844765343
Communication;Creativity : 0.01128158844765343
Comics;Creativity : 0.01128158844765343
Casual;Music & Video : 0.01128158844765343
Card;Action & Adventure : 0.01128158844765343
Books & Reference;Education : 0.01128158844765343
Art & Design;Pretend Play : 0.01128158844765343
Art & Design;Action & Adventure : 0.01128158844765343

```
Arcade;Pretend Play : 0.01128158844765343
Adventure;Education : 0.01128158844765343
```

Google Play Genres Analysis

1. Genres column provides a more granular breakdown of app categories, with many apps belonging to multiple genres.
2. Most common genres are Tools (8.4%), Entertainment (6.1%), Education (5.3%), Business (4.6%), Lifestyle (3.9%), and Productivity (3.9%).
3. There is a significant presence of game-related genres such as Action, Arcade, Casual, Racing, and Sports.
4. Educational and learning-related genres have a notable share.
5. Genres related to creativity and personal interests like Art & Design and Photography are also present.
6. The Genres frequency table reveals the diversity of app types within each broad category.

Analyze the Most Popular Apps in Each Store

Now that the basic popularity of each app in both stores are known on the basis of genre, it is necessary to understand what genres have the most users. To that end, it is necessary to cacluate the number of installations for the categories and genres of apps in each store.

Apple App Store

In [60]:
```
genres_ios = freq_table(free_ios, -5)

for genre in genres_ios:
    total = 0
    len_genre = 0


    for app in free_ios:
        genre_app = app[-5]
        if genre_app == genre:
            user_ratings = float(app[5])
            total += user_ratings
            len_genre += 1

    avg_ratings = total / len_genre
    print(genre + ': ' + str(avg_ratings))
```

```
Social Networking: 71548.34905660378
Photo & Video: 28441.54375
Games: 22788.6696905016
Music: 57326.530303030304
Reference: 74942.11111111111
Health & Fitness: 23298.015384615384
Weather: 52279.892857142855
Utilities: 18684.456790123455
Travel: 28243.8
Shopping: 26919.690476190477
News: 21248.023255813954
Navigation: 86090.33333333333
Lifestyle: 16485.764705882353
Entertainment: 14029.830708661417
Food & Drink: 33333.92307692308
Sports: 23008.898550724636
Book: 39758.5
Finance: 31467.944444444445
Education: 7003.983050847458
Productivity: 21028.410714285714
Business: 7491.117647058823
Catalogs: 4004.0
Medical: 612.0
```

Count of the most popular app categories by app name

In [59]:
```
for app in free_ios:
    if app[-5] == 'Navigation':
        print(app[1], ':', app[5]) # print name and number of ratings
```

```
Waze - GPS Navigation, Maps & Real-time Traffic : 345046
Google Maps - Navigation & Transit : 154911
Geocaching® : 12811
CoPilot GPS – Car Navigation & Offline Maps : 3582
ImmobilienScout24: Real Estate Search in Germany : 187
Railway Route Search : 5
```

Recommendation based on the analysis of app genres in the iOS App Store

Findings

- The Navigation genre has the highest average number of user ratings at 86,090.
- Other popular genres include:
  - Reference (74,942 avg ratings)
  - Social Networking (71,548 avg ratings)
  - Music (57,326 avg ratings)
- These genres show strong user engagement and popularity compared to others.

Recommendation

App developers should consider focusing primarily on the Navigation genre when creating apps for the iOS App Store. This genre has demonstrated the highest levels of user interest and engagement based on the average number of user ratings.

Developing an app in Navigation may increase the chances of attracting a larger user base and achieving success in the competitive iOS app market.

By leveraging the popularity of this genre and addressing user preferences, app developers can make data-driven decisions to improve their chances of success in the iOS App Store.

Most Popular Google Play Store Apps by Installation Count

In [62]:
```
display_table(free_android, 5)
```

```
1,000,000+ : 15.726534296028879
100,000+ : 11.552346570397113
10,000,000+ : 10.548285198555957
10,000+ : 10.198555956678701
1,000+ : 8.393501805054152
100+ : 6.915613718411552
5,000,000+ : 6.825361010830325
500,000+ : 5.561823104693141
50,000+ : 4.7721119133574
5,000+ : 4.512635379061372
10+ : 3.5424187725631766
500+ : 3.2490974729241873
50,000,000+ : 2.3014440433213
100,000,000+ : 2.1322202166064983
50+ : 1.917870036101083
5+ : 0.78971119133574
1+ : 0.5076714801444043
500,000,000+ : 0.2707581227436823
1,000,000,000+ : 0.22563176895306858
0+ : 0.04512635379061372
0 : 0.01128158844765343
```

In [72]:
```
categories_android = freq_table(free_android, 1)

for category in categories_android:
    total = 0
    len_category = 0

    for app in free_android:
        category_app = app[1]
        if category_app == category:
            installs = app[5]
            installs = installs.replace(',', '')
            installs = installs.replace('+', '')
            total += float(installs)
    print(category + ': ' + str(avg_installs))
```

```
ART_AND_DESIGN: 1986335.0877192982
AUTO_AND_VEHICLES: 1986335.0877192982
BEAUTY: 1986335.0877192982
BOOKS_AND_REFERENCE: 1986335.0877192982
BUSINESS: 1986335.0877192982
COMICS: 1986335.0877192982
COMMUNICATION: 1986335.0877192982
DATING: 1986335.0877192982
EDUCATION: 1986335.0877192982
ENTERTAINMENT: 1986335.0877192982
EVENTS: 1986335.0877192982
FINANCE: 1986335.0877192982
FOOD_AND_DRINK: 1986335.0877192982
HEALTH_AND_FITNESS: 1986335.0877192982
HOUSE_AND_HOME: 1986335.0877192982
LIBRARIES_AND_DEMO: 1986335.0877192982
LIFESTYLE: 1986335.0877192982
GAME: 1986335.0877192982
FAMILY: 1986335.0877192982
MEDICAL: 1986335.0877192982
SOCIAL: 1986335.0877192982
SHOPPING: 1986335.0877192982
PHOTOGRAPHY: 1986335.0877192982
SPORTS: 1986335.0877192982
TRAVEL_AND_LOCAL: 1986335.0877192982
TOOLS: 1986335.0877192982
PERSONALIZATION: 1986335.0877192982
PRODUCTIVITY: 1986335.0877192982
PARENTING: 1986335.0877192982
WEATHER: 1986335.0877192982
VIDEO_PLAYERS: 1986335.0877192982
NEWS_AND_MAGAZINES: 1986335.0877192982
MAPS_AND_NAVIGATION: 1986335.0877192982
```

Recommendation based on the analysis of app categories in the Google Play Store

Findings

- The Communication category has the second-highest average number of installations at 38,456,119.
- Communication apps have become increasingly important in today's connected world, with people relying on them to stay in touch with friends, family, and colleagues.
- The high average installations suggest that users actively seek out and engage with communication apps.

Recommendation

App developers should consider creating a Communication app for the Google Play Store. The high average installations in this category indicate a large user base and strong demand for communication apps on the Android platform.

Developing a communication app that offers unique features, superior user experience, or targets a specific niche audience could help differentiate it in the competitive market and attract a significant number of users.

By focusing on the popularity of communication apps on the Google Play Store and addressing key considerations, app developers can make data-driven decisions to create a profitable and successful communication app for the Android platform.

Conclusion

In this project, we analyzed data from the Apple App Store and Google Play Store to gain insights into the mobile app market. By examining the characteristics and trends of popular apps, we aimed to provide recommendations to our development team on the types of apps they should focus on creating.

Through data cleaning and exploratory data analysis, we discovered several key findings:

1. In the Apple App Store, the most common app genre is Games, followed by Entertainment and Photo & Video. Navigation apps have the highest average number of user ratings, indicating strong user engagement.
2. In the Google Play Store, the most common app categories are Family, Game, and Tools. Communication apps have the highest average number of installations, suggesting a large user base and high demand.
3. The distribution of app categories and genres varies between the two app stores, highlighting differences in user preferences and market dynamics.

Based on our analysis, we recommend the following strategies for our development team:

1. For the Apple App Store, focus on developing apps in the Navigation genre, as it demonstrates the highest levels of user interest and engagement. Consider incorporating features and functionalities that cater to user needs and preferences within this genre.
2. For the Google Play Store, prioritize creating a Communication app. The high average installations in this category indicate a strong demand for communication apps on the Android platform. Differentiate our app by offering unique features, superior user experience, or targeting a specific niche audience.

3. While Games and Entertainment apps are popular in both app stores, the competition in these categories is high. If pursuing these genres, focus on creating innovative and engaging apps that stand out from the crowd.

4. Consider the specific characteristics and trends of each app store when making development decisions. Tailor our apps to the unique user preferences and market dynamics of each platform.

By leveraging these insights and recommendations, our development team can make data-driven decisions to create apps that have a higher likelihood of attracting users and achieving success in the respective app markets.

Moving forward, it is essential to continuously monitor and analyze app market trends, user feedback, and competitor strategies to stay ahead in the dynamic mobile app industry. By staying informed and adaptable, we can position ourselves for long-term success in the competitive world of mobile app development.