Most used Apps

As our company builts free Android and iOS mobile apps and company's main sourse of revenue is from in apps ads, so this project is about finding apps which are most profitable in Google store and App store. This would help our compnay to understand what type of apps are likely to attract more users. Our goal is to build a minimal Android app and launch it on Google Store. If the app have positive reponse we will make its ios version with some improvements to launch it on Apple Store so we need to find the app profile which fits both Apple Store and Google Store as well.

In [1]:

```python
from csv import reader

#Data set from Googleplay


opened_file = open('googleplaystore.csv')
read_file = reader(opened_file)
android = list(read_file)
android_header = android[0]
android = android[1:]


#Data set from Appstore

opened_file = open('AppleStore.csv')
read_file = reader(opened_file)
ios = list(read_file)
ios_header = ios[0]
ios = ios[1:]
```

We will write a function name as 'explore_data' to explore data in more readable way. It will also print number of rows and columns.

In [2]:

```python
def explore_data(dataset, start, end, rows_and_columns=False):
    dataset_slice = dataset[start:end]
    for row in dataset_slice:
        print(row)
        print('\n')
    if rows_and_columns:
        print('Number of rows: ', len(dataset))
        print('Number of columns: ', len(dataset[0]))

print(android_header)
print('\n')
explore_data(android, 0, 3, True)
```

```
['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type', 'Price', 'Content Rating', 'G
enres', 'Last Updated', 'Current Ver', 'Android Ver']


['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']


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

File showing data of Google Play Store have 10841 rows and 13 columns. The important columns for analysis are 'App', 'Category', 'Reviews', 'Installs', 'Type', 'Price', and 'Genres'.

Reviews, Installs, Type, Price, and Genres.

```
print(ios_header)
print('\n')
explore_data(ios, 0, 3, True)
```

```
['id', 'track_name', 'size_bytes', 'currency', 'price', 'rating_count_tot', 'rating_count_ver', 'u
ser_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', '1
2+', '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
```

File showing data of App Store have 7197 rows and 16 columns. The important columns are: 'track_name', 'currency', 'price', 'rating_count_tot', 'rating_count_ver', and 'prime_genre'.

```
print(android[10472])
print('\n')
print(android_header)
print('\n')
print(android[0])
```

```
['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', 'Price', 'Content Rating', 'G
enres', 'Last Updated', 'Current Ver', 'Android Ver']


['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']
```

Before analysing the data, it is important to clean the data. We will remove or correct any inaccuate data. We will also detect duplicate data and will remove it. As we can see there is a error for row 10472 as it is giving 19 for rating so we will remove this row.

```
print(len(android))
del android[10472]
print(len(android))
```

```
10841
10840
```

```
for app in android:
    name = app[0]
    if name == 'Instagram':
        print(app)
```

```
['Instagram', 'SOCIAL', '4.5', '66577313', 'Varies with device', '1,000,000,000+', 'Free', '0', 'T
```

```
een', 'Social', 'July 31, 2018', 'Varies with device', 'Varies with device']
['Instagram', 'SOCIAL', '4.5', '66577446', 'Varies with device', '1,000,000,000+', 'Free', '0', 'T
een', 'Social', 'July 31, 2018', 'Varies with device', 'Varies with device']
['Instagram', 'SOCIAL', '4.5', '66577313', 'Varies with device', '1,000,000,000+', 'Free', '0', 'T
een', 'Social', 'July 31, 2018', 'Varies with device', 'Varies with device']
['Instagram', 'SOCIAL', '4.5', '66509917', 'Varies with device', '1,000,000,000+', 'Free', '0', 'T
een', 'Social', 'July 31, 2018', 'Varies with device', 'Varies with device']
```

Google play data set have duplicate entries that can be seen above. App name 'Instagram' has multiple entries.

In [7]:

```python
same_app = []
unique_app = []

for app in android:
    name = app[0]
    if name in unique_app:
        same_app.append(name)
    else:
        unique_app.append(name)
print('Number of same apps ', len(same_app))
print('\n')
print('Example of same apps \n', same_app[:15])
```

```
Number of same apps  1181


Example of same apps
 ['Quick PDF Scanner + OCR FREE', 'Box', 'Google My Business', 'ZOOM Cloud Meetings', 'join.me - S
imple Meetings', 'Box', 'Zenefits', 'Google Ads', 'Google My Business', 'Slack', 'FreshBooks
Classic', 'Insightly CRM', 'QuickBooks Accounting: Invoicing & Expenses', 'HipChat - Chat Built fo
r Teams', 'Xero Accounting Software']
```

After knowing that we have 1181 duplicate apps and we want to remove it because as it is a deformaty in data and we want our data to be precise and accurate as possible. As you can see 'Instagram' app have different reviews which means the highest reviews are for the latest data so we want this row for analysis.

In [8]:

```python
max_reviews = {}

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

    if name in max_reviews and max_reviews[name] < n_reviews:
        max_reviews[name] = n_reviews

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

In [9]:

```python
print('Expected lenght = ', len(android)-1181)
print('Actual length = ', len(max_reviews))
```

```
Expected lenght =  9659
Actual length =  9659
```

As indicated that we have 1181 cases where apps have multiple entries so the length of our dictionary should be equal to difference between length of data minus 1181.

In [10]:

```python
clean_android = []
already_added = []

for app in android:
```

```
    name = app[0]
    n_reviews = float(app[3])

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

Now its time to explore data and make sure we have 9656 rows.

In [11]:

```
explore_data(clean_android, 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', 'T
een', 'Art & Design', 'June 8, 2018', 'Varies with device', '4.2 and up']


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

As our company wants to make an English app so we do not need apps which are not in English so we would be removing those apps.

In [12]:

```
def english(string):
    non_ascii = 0
    for character in string:
        if ord(character) > 127:
            non_ascii += 1

    if non_ascii > 3:
        return False

    else:
        return True
print(english('Docs To Go™ Free Office Suite'))
print(english('Instachat 😜'))
print(english('爱奇艺PPS -《欢乐颂2》电视剧热播'))
print(english('Instagram'))
```

```
True
True
False
True
```

In [13]:

```
english_android = []
english_ios = []

for app in clean_android:
    name = app[0]
    if english(name):
        english_android.append(app)

for app in ios:
    name = app[1]
    if english(name):
        english_ios.append(app)


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

```
print('\n')
explore_data(english_ios, 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', 'T
een', 'Art & Design', 'June 8, 2018', 'Varies with device', '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', '1
2+', '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

In [14]:

```python
android_new = []
ios_new = []

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

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

print(len(android_new))
print(len(ios_new))
```

8864
3222

Till here we have clean the data by removing duplicates, removing apps in language other then English and removing inaccurate data. We have also isolated free apps into andoid_final and ios_final.

As our goal is to build a minimal Android app and launch it on Google Store. If the app have positive reponse we will make its ios version with some improvements to launch it on Apple Store so we need to find the app profile which fits both Apple Store and Google Store as well. To do that we should have a frequency table to know number of apps in each genre.

In [15]:

```python
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_percentage = {}
    for key in table:
        percentage = (table[key]/total)*100
        table_percentage[key] = percentage
    return table_percentage


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])
```

```
display_table(ios_new, -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
```

Limitation of this data : These apps are only English free apps.

On Apple Store most common genre are 'Games' (more then a half) 58%, Entertainment about 8% followed by Photo & Video about 5%. Educational apps are only 3.7% only and Social Networking apps are 3.3%.

The General impression is that App Store is dominated for fun (Games, Entertainment photos and video) but apps having practical purpose (education, lifestyle) are rare.

```
display_table(android_new, 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
```

The pattern is different of Google Play Store from App Store as Google Play Store has more practical purpose apps (Family about 19%, Tools 8.5%, Business 4.6%, Lifestyle 3,9%, health and fitness 3%). However, if we investigate this further, we can see that the family category (which accounts for almost 19% of the apps) means mostly games for kids.

In [18]:

```
display_table(android_new, -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  :  0.4287003610108303
Trivia  :  0.41741877256317694
```

```
Educational;Education  :  0.39485559556787
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
```

'Genre' has more catagory then 'Catagory' column but as we are looking at big picture so we will work with 'Catagory' column.

Now we have to know which genre is used most by the users. We will calculate installs for each genre app. In the file of Google play

store we have a column name as 'Install' which tells us number of installs but for Apple Store we do not have that column so we will use 'rating_count_tot' column to calculate installs for each genre app.

```python
genres_ios = freq_table(ios_new, -5)
for genre in genres_ios:
    total = 0
    len_genre = 0
    for app in ios_new:
        genre_app = app[-5]
        if genre_app == genre:
            n_ratings = float(app[5])
            total += n_ratings
            len_genre += 1
    avg_n_ratings = total/len_genre
    print(genre, ' : ', avg_n_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
```

As you can see from this data, Navigation apps has most users but this is influenced by Waze or Google maps.

```python
for app in ios_new:
    if app[-5] == 'Navigation':
        print(app[1], ' : ', app[5])
```

```
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
```

This same pattern applies for Social Networking genre as well where few giant companies like Facebook, Skype, Pininterest has influenced that average. We are looking for most popular genre overall.

```python
for app in ios_new:
    if app[-5] == 'Reference':
        print(app[1], ' : ', app[5])
```

```
Bible  :  985920
Dictionary.com Dictionary & Thesaurus  :  200047
Dictionary.com Dictionary & Thesaurus for iPad  :  54175
Google Translate  :  26786
```

```
Muslim Pro: Ramadan 2017 Prayer Times, Azan, Quran  :  18418
New Furniture Mods - Pocket Wiki & Game Tools for Minecraft PC Edition  :  17588
Merriam-Webster Dictionary  :  16849
Night Sky  :  12122
City Maps for Minecraft PE - The Best Maps for Minecraft Pocket Edition (MCPE)  :  8535
LUCKY BLOCK MOD ™ for Minecraft PC Edition - The Best Pocket Wiki & Mods Installer Tools  :  4693
GUNS MODS for Minecraft PC Edition - Mods Tools  :  1497
Guides for Pokémon GO - Pokemon GO News and Cheats  :  826
WWDC  :  762
Horror Maps for Minecraft PE - Download The Scariest Maps for Minecraft Pocket Edition (MCPE) Free
:  718
VPN Express  :  14
Real Bike Traffic Rider Virtual Reality Glasses  :  8
教えて!goo  :  0
Jishokun-Japanese English Dictionary & Translator  :  0
```

As you can see the Reference genre is also influenced by Bible and Dictionary.com which has skew up the average.

One thing we could do is take another popular book and turn it into an app where we could add different features besides the raw version of the book. This might include quotes of the day from the book, an audio version of the book etc. We could also add a dictionary within the app, so users don't need to quit our app to look up words.

This idea seems to fit well with the fact that the App Store is dominated by for-fun apps. This suggests the market might be a bit saturated with for-fun apps, which means a practical app might have more of a chance to stand out among the huge number of apps on the App Store.

Other genres that seem popular include weather, book, food and drink, or finance. The book genre seem to overlap a bit with the app idea we described above, but the other genres don't seem too interesting to us: Weather apps — people generally don't spend too much time in-app, and the chances of making profit from in-app adds are low. Also, getting reliable live weather data may require us to connect our apps to non-free APIs. Food and drink — examples here include Starbucks, Dunkin' Donuts, McDonald's, etc. So making a popular food and drink app requires actual cooking and a delivery service, which is outside the scope of our company. Finance apps — these apps involve banking, paying bills, money transfer, etc. Building a finance app requires domain knowledge, and we don't want to hire a finance expert just to build an app.

In [22]:
```python
display_table(android_new, 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
```

For Google Play store data we do have number of installs but the issue with this data is that it is not precise because we dont know 1,000,000+ installs is 1000,000 or 2000,000 or 3000,000. But for our mission we dont need that precision so we would take 1000,000+ as 1000,0000 installs.

In [23]:
```python
categories_android = freq_table(android_new, 1)

for category in categories_android:
        total = 0
```

```
        len_category = 0
        for app in android_new:
            category_app = app[1]
            if category_app == category:
                n_installs = app[5]
                n_installs = n_installs.replace(',', '')
                n_installs = n_installs.replace('+', '')
                total += float(n_installs)
                len_category += 1
        avg_n_installs = total/len_category
        print(category, ' : ', avg_n_installs)
```

```
ART_AND_DESIGN  :  1986335.0877192982
AUTO_AND_VEHICLES  :  647317.8170731707
BEAUTY  :  513151.88679245283
BOOKS_AND_REFERENCE  :  8767811.894736841
BUSINESS  :  1712290.1474201474
COMICS  :  817657.2727272727
COMMUNICATION  :  38456119.167247385
DATING  :  854028.8303030303
EDUCATION  :  1833495.145631068
ENTERTAINMENT  :  11640705.88235294
EVENTS  :  253542.22222222222
FINANCE  :  1387692.475609756
FOOD_AND_DRINK  :  1924897.7363636363
HEALTH_AND_FITNESS  :  4188821.9853479853
HOUSE_AND_HOME  :  1331540.5616438356
LIBRARIES_AND_DEMO  :  638503.734939759
LIFESTYLE  :  1437816.2687861272
GAME  :  15588015.603248259
FAMILY  :  3695641.8198090694
MEDICAL  :  120550.61980830671
SOCIAL  :  23253652.127118643
SHOPPING  :  7036877.311557789
PHOTOGRAPHY  :  17840110.40229885
SPORTS  :  3638640.1428571427
TRAVEL_AND_LOCAL  :  13984077.710144928
TOOLS  :  10801391.298666667
PERSONALIZATION  :  5201482.6122448975
PRODUCTIVITY  :  16787331.344927534
PARENTING  :  542603.6206896552
WEATHER  :  5074486.197183099
VIDEO_PLAYERS  :  24727872.452830188
NEWS_AND_MAGAZINES  :  9549178.467741935
MAPS_AND_NAVIGATION  :  4056941.7741935486
```

On average, Communication has most installs 38456119 but this number is skewed because of giant apps like Facebook, Skype, Messenger, Google Chrome that have one over billions of installs.

In [24]:

```
for app in android_new:
    if (app[1] == 'COMMUNICATION') and (app[5] == '1,000,000,000+'
                                    or app[5] == '500,000,000+'
                                    or app[5] == '100,000,000+'):
        print(app[0], ' : ', app[5])
```

```
WhatsApp Messenger  :  1,000,000,000+
imo beta free calls and text  :  100,000,000+
Android Messages  :  100,000,000+
Google Duo - High Quality Video Calls  :  500,000,000+
Messenger – Text and Video Chat for Free  :  1,000,000,000+
imo free video calls and chat  :  500,000,000+
Skype - free IM & video calls  :  1,000,000,000+
Who  :  100,000,000+
GO SMS Pro - Messenger, Free Themes, Emoji  :  100,000,000+
LINE: Free Calls & Messages  :  500,000,000+
Google Chrome: Fast & Secure  :  1,000,000,000+
Firefox Browser fast & private  :  100,000,000+
UC Browser - Fast Download Private & Secure  :  500,000,000+
Gmail  :  1,000,000,000+
Hangouts  :  1,000,000,000+
Messenger Lite: Free Calls & Messages  :  100,000,000+
```

```
Messenger Lite: Free Calls & Messages  :  100,000,000+
Kik  :  100,000,000+
KakaoTalk: Free Calls & Text  :  100,000,000+
Opera Mini - fast web browser  :  100,000,000+
Opera Browser: Fast and Secure  :  100,000,000+
Telegram  :  100,000,000+
Truecaller: Caller ID, SMS spam blocking & Dialer  :  100,000,000+
UC Browser Mini -Tiny Fast Private & Secure  :  100,000,000+
Viber Messenger  :  500,000,000+
WeChat  :  100,000,000+
Yahoo Mail – Stay Organized  :  100,000,000+
BBM - Free Calls & Messages  :  100,000,000+
```

We will remove communiation apps which have more then 100 million installs which will reduce our average almost ten times.

In [25]:

```python
under_100_m = []
for app in android_new:
    n_installs = app[5]
    n_installs = n_installs.replace(',', '')
    n_installs = n_installs.replace('+', '')
    if (app[1] == 'COMMUNICATION') and (float(n_installs) < 100000000):
        under_100_m.append(float(n_installs))

print(sum(under_100_m) / len(under_100_m))
```

```
3603485.3884615386
```

We see the same pattern for the video players category with 24,727,872 installs. The market is dominated by apps like Youtube, Google Play Movies & TV, or MX Player. We can observe the same pattern for social apps (where we have giants like Facebook, Instagram, Google+, etc.), photography apps (Google Photos and other popular photo editors), or productivity apps (Microsoft Word, Dropbox, Google Calendar, Evernote, etc.).

The game genre seems very popular, but we found out this part of the market seems a bit saturated, so we would like to come up with a different app recommendation if possible.

The books and reference genre looks fairly popular as well, with an average number of installs of 8,767,811. It's interesting to explore this in more depth, since we found this genre has some potential to work well on the App Store, and our aim is to recommend an app genre that shows potential for being profitable on both the App Store and Google Play. Let's take a look at some of the apps from this genre and their number of installs:

In [26]:

```python
for app in android_new:
    if app[1] == 'BOOKS_AND_REFERENCE':
        print(app[0], ' : ', app[5])
```

```
E-Book Read - Read Book for free  :  50,000+
Download free book with green book  :  100,000+
Wikipedia  :  10,000,000+
Cool Reader  :  10,000,000+
Free Panda Radio Music  :  100,000+
Book store  :  1,000,000+
FBReader: Favorite Book Reader  :  10,000,000+
English Grammar Complete Handbook  :  500,000+
Free Books - Spirit Fanfiction and Stories  :  1,000,000+
Google Play Books  :  1,000,000,000+
AlReader -any text book reader  :  5,000,000+
Offline English Dictionary  :  100,000+
Offline: English to Tagalog Dictionary  :  500,000+
FamilySearch Tree  :  1,000,000+
Cloud of Books  :  1,000,000+
Recipes of Prophetic Medicine for free  :  500,000+
ReadEra – free ebook reader  :  1,000,000+
Anonymous caller detection  :  10,000+
Ebook Reader  :  5,000,000+
Litnet - E-books  :  100,000+
Read books online  :  5,000,000+
English to Urdu Dictionary  :  500,000+
eBoox: book reader fb2 epub zip  :  1,000,000+
English Persian Dictionary  :  500,000+
```

English Persian Dictionary : 500,000+
Flybook : 500,000+
All Maths Formulas : 1,000,000+
Ancestry : 5,000,000+
HTC Help : 10,000,000+
English translation from Bengali : 100,000+
Pdf Book Download - Read Pdf Book : 100,000+
Free Book Reader : 100,000+
eBoox new: Reader for fb2 epub zip books : 50,000+
Only 30 days in English, the guideline is guaranteed : 500,000+
Moon+ Reader : 10,000,000+
SH-02J Owner's Manual (Android 8.0) : 50,000+
English-Myanmar Dictionary : 1,000,000+
Golden Dictionary (EN-AR) : 1,000,000+
All Language Translator Free : 1,000,000+
Azpen eReader : 500,000+
URBANO V 02 instruction manual : 100,000+
Bible : 100,000,000+
C Programs and Reference : 50,000+
C Offline Tutorial : 1,000+
C Programs Handbook : 50,000+
Amazon Kindle : 100,000,000+
Aab e Hayat Full Novel : 100,000+
Aldiko Book Reader : 10,000,000+
Google I/O 2018 : 500,000+
R Language Reference Guide : 10,000+
Learn R Programming Full : 5,000+
R Programing Offline Tutorial : 1,000+
Guide for R Programming : 5+
Learn R Programming : 10+
R Quick Reference Big Data : 1,000+
V Made : 100,000+
Wattpad 📖 Free Books : 100,000,000+
Dictionary - WordWeb : 5,000,000+
Guide (for X-MEN) : 100,000+
AC Air condition Troubleshoot,Repair,Maintenance : 5,000+
AE Bulletins : 1,000+
Ae Allah na Dai (Rasa) : 10,000+
50000 Free eBooks & Free AudioBooks : 5,000,000+
Ag PhD Field Guide : 10,000+
Ag PhD Deficiencies : 10,000+
Ag PhD Planting Population Calculator : 1,000+
Ag PhD Soybean Diseases : 1,000+
Fertilizer Removal By Crop : 50,000+
A-J Media Vault : 50+
Al-Quran (Free) : 10,000,000+
Al Quran (Tafsir & by Word) : 500,000+
Al Quran Indonesia : 10,000,000+
Al'Quran Bahasa Indonesia : 10,000,000+
Al Quran Al karim : 1,000,000+
Al-Muhaffiz : 50,000+
Al Quran : EAlim - Translations & MP3 Offline : 5,000,000+
Al-Quran 30 Juz free copies : 500,000+
Koran Read &MP3 30 Juz Offline : 1,000,000+
Hafizi Quran 15 lines per page : 1,000,000+
Quran for Android : 10,000,000+
Surah Al-Waqiah : 100,000+
Hisnul Al Muslim - Hisn Invocations & Adhkaar : 100,000+
Satellite AR : 1,000,000+
Audiobooks from Audible : 100,000,000+
Kinot & Eichah for Tisha B'Av : 10,000+
AW Tozer Devotionals - Daily : 5,000+
Tozer Devotional -Series 1 : 1,000+
The Pursuit of God : 1,000+
AY Sing : 5,000+
Ay Hasnain k Nana Milad Naat : 10,000+
Ay Mohabbat Teri Khatir Novel : 10,000+
Arizona Statutes, ARS (AZ Law) : 1,000+
Oxford A-Z of English Usage : 1,000,000+
BD Fishpedia : 1,000+
BD All Sim Offer : 10,000+
Youboox - Livres, BD et magazines : 500,000+
B&H Kids AR : 10,000+
B y H Niños ES : 5,000+
Dictionary.com: Find Definitions for English Words : 10,000,000+
English Dictionary - Offline : 10,000,000+
Bible KJV : 5,000,000+
Borneo Bible, BM Bible : 10,000+

Borneo Bible, BM Bible : 10,000+
MOD Black for BM : 100+
BM Box : 1,000+
Anime Mod for BM : 100+
NOOK: Read eBooks & Magazines : 10,000,000+
NOOK Audiobooks : 500,000+
NOOK App for NOOK Devices : 500,000+
Browsery by Barnes & Noble : 5,000+
bp e-store : 1,000+
Brilliant Quotes: Life, Love, Family & Motivation : 1,000,000+
BR Ambedkar Biography & Quotes : 10,000+
BU Alsace : 100+
Catholic La Bu Zo Kam : 500+
Khrifa Hla Bu (Solfa) : 10+
Kristian Hla Bu : 10,000+
SA HLA BU : 1,000+
Learn SAP BW : 500+
Learn SAP BW on HANA : 500+
CA Laws 2018 (California Laws and Codes) : 5,000+
Bootable Methods(USB-CD-DVD) : 10,000+
cloudLibrary : 100,000+
SDA Collegiate Quarterly : 500+
Sabbath School : 100,000+
Cypress College Library : 100+
Stats Royale for Clash Royale : 1,000,000+
GATE 21 years CS Papers(2011-2018 Solved) : 50+
Learn CT Scan Of Head : 5,000+
Easy Cv maker 2018 : 10,000+
How to Write CV : 100,000+
CW Nuclear : 1,000+
CY Spray nozzle : 10+
BibleRead En Cy Zh Yue : 5+
CZ-Help : 5+
Modlitební knížka CZ : 500+
Guide for DB Xenoverse : 10,000+
Guide for DB Xenoverse 2 : 10,000+
Guide for IMS DB : 10+
DC HSEMA : 5,000+
DC Public Library : 1,000+
Painting Lulu DC Super Friends : 1,000+
Dictionary : 10,000,000+
Fix Error Google Playstore : 1,000+
D. H. Lawrence Poems FREE : 1,000+
Bilingual Dictionary Audio App : 5,000+
DM Screen : 10,000+
wikiHow: how to do anything : 1,000,000+
Dr. Doug's Tips : 1,000+
Bible du Semeur-BDS (French) : 50,000+
La citadelle du musulman : 50,000+
DV 2019 Entry Guide : 10,000+
DV 2019 - EDV Photo & Form : 50,000+
DV 2018 Winners Guide : 1,000+
EB Annual Meetings : 1,000+
EC - AP & Telangana : 5,000+
TN Patta Citta & EC : 10,000+
AP Stamps and Registration : 10,000+
CompactiMa EC pH Calibration : 100+
EGW Writings 2 : 100,000+
EGW Writings : 1,000,000+
Bible with EGW Comments : 100,000+
My Little Pony AR Guide : 1,000,000+
SDA Sabbath School Quarterly : 500,000+
Duaa Ek Ibaadat : 5,000+
Spanish English Translator : 10,000,000+
Dictionary - Merriam-Webster : 10,000,000+
JW Library : 10,000,000+
Oxford Dictionary of English : Free : 10,000,000+
English Hindi Dictionary : 10,000,000+
English to Hindi Dictionary : 5,000,000+
EP Research Service : 1,000+
Hymnes et Louanges : 100,000+
EU Charter : 1,000+
EU Data Protection : 1,000+
EU IP Codes : 100+
EW PDF : 5+
BakaReader EX : 100,000+
EZ Quran : 50,000+
FA Part 1 & 2 Past Papers Solved Free - Offline : 5,000+

```
FA Part 1 & 2 Past Papers Solved Free - Offline  :  5,000+
La Fe de Jesus  :  1,000+
La Fe de Jesús  :  500+
Le Fe de Jesus  :  500+
Florida - Pocket Brainbook  :  1,000+
Florida Statutes (FL Code)  :  1,000+
English To Shona Dictionary  :  10,000+
Greek Bible FP (Audio)  :  1,000+
Golden Dictionary (FR-AR)  :  500,000+
Fanfic-FR  :  5,000+
Bulgarian French Dictionary Fr  :  10,000+
Chemin (fr)  :  1,000+
The SCP Foundation DB fr nn5n  :  1,000+
```

There is variety of apps in this genre and still there are some popular apps who are skewing the average

In [27]:

```python
for app in android_new:
    if app[1] == 'BOOKS_AND_REFERENCE' and (app[5] == '1,000,000,000+'
                                            or app[5] == '500,000,000+'
                                            or app[5] == '100,000,000+'):
        print(app[0], ' : ', app[5])
```

```
Google Play Books  :  1,000,000,000+
Bible  :  100,000,000+
Amazon Kindle  :  100,000,000+
Wattpad □ Free Books  :  100,000,000+
Audiobooks from Audible  :  100,000,000+
```

However, it looks like there are only a few very popular apps, so this market still shows potential. Let's try to get some app ideas based on the kind of apps that are somewhere in the middle in terms of popularity (between 1,000,000 and 100,000,000 downloads):

In [28]:

```python
for app in android_new:
    if app[1] == 'BOOKS_AND_REFERENCE' and (app[5] == '1,000,000+'
                                            or app[5] == '5,000,000+'
                                            or app[5] == '10,000,000+'
                                            or app[5] == '50,000,000+'):
        print(app[0], ' : ', app[5])
```

```
Wikipedia  :  10,000,000+
Cool Reader  :  10,000,000+
Book store  :  1,000,000+
FBReader: Favorite Book Reader  :  10,000,000+
Free Books - Spirit Fanfiction and Stories  :  1,000,000+
AlReader -any text book reader  :  5,000,000+
FamilySearch Tree  :  1,000,000+
Cloud of Books  :  1,000,000+
ReadEra – free ebook reader  :  1,000,000+
Ebook Reader  :  5,000,000+
Read books online  :  5,000,000+
eBoox: book reader fb2 epub zip  :  1,000,000+
All Maths Formulas  :  1,000,000+
Ancestry  :  5,000,000+
HTC Help  :  10,000,000+
Moon+ Reader  :  10,000,000+
English-Myanmar Dictionary  :  1,000,000+
Golden Dictionary (EN-AR)  :  1,000,000+
All Language Translator Free  :  1,000,000+
Aldiko Book Reader  :  10,000,000+
Dictionary - WordWeb  :  5,000,000+
50000 Free eBooks & Free AudioBooks  :  5,000,000+
Al-Quran (Free)  :  10,000,000+
Al Quran Indonesia  :  10,000,000+
Al'Quran Bahasa Indonesia  :  10,000,000+
Al Quran Al karim  :  1,000,000+
Al Quran : EAlim - Translations & MP3 Offline  :  5,000,000+
Koran Read &MP3 30 Juz Offline  :  1,000,000+
Hafizi Quran 15 lines per page  :  1,000,000+
Quran for Android  :  10,000,000+
Satellite AR  :  1,000,000+
```

```
Satellite AR  :  1,000,000+
Oxford A-Z of English Usage  :  1,000,000+
Dictionary.com: Find Definitions for English Words  :  10,000,000+
English Dictionary - Offline  :  10,000,000+
Bible KJV  :  5,000,000+
NOOK: Read eBooks & Magazines  :  10,000,000+
Brilliant Quotes: Life, Love, Family & Motivation  :  1,000,000+
Stats Royale for Clash Royale  :  1,000,000+
Dictionary  :  10,000,000+
wikiHow: how to do anything  :  1,000,000+
EGW Writings  :  1,000,000+
My Little Pony AR Guide  :  1,000,000+
Spanish English Translator  :  10,000,000+
Dictionary - Merriam-Webster  :  10,000,000+
JW Library  :  10,000,000+
Oxford Dictionary of English : Free  :  10,000,000+
English Hindi Dictionary  :  10,000,000+
English to Hindi Dictionary  :  5,000,000+
```

Conclusion

We also notice there are quite a few apps built around the book Quran, which suggests that building an app around a popular book can be profitable. It seems that taking a popular book (perhaps a more recent book) and turning it into an app could be profitable for both the Google Play and the App Store markets. We also need to add some special features like daily quotes from the book, an audio version of the book, quizzes on the book, a forum where people can discuss the book, etc.