In [1]:
```
# Group Number: 09
# Group Member: Laxmi Gurung and Neja Gurung
# Project: Using the appstoregames.cs file, write python code to read the file
#        into a Pandas DataFrame analyze the data.
# Date: 11/01/2021


# Import the libraries numpy and pandas for data manipulation.
import pandas as pd
import numpy as np

# Reading the csv file.
appStoreGames = pd.read_csv('appstoregames.csv')
```

In [2]:
```
# 1. Use the appropriate Pandas method to describe the characteristics of your dataset.
appStoreGames.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17007 entries, 0 to 17006
Data columns (total 14 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   ID                           17007 non-null  int64
 1   Name                         17007 non-null  object
 2   Subtitle                     5261 non-null   object
 3   User Rating                  7561 non-null   float64
 4   Count of Ratings             7561 non-null   float64
 5   Price                        16983 non-null  float64
 6   In-app Purchases             7683 non-null   object
 7   Developer                    17007 non-null  object
 8   Age Rating                   17007 non-null  object
 9   Languages                    16947 non-null  object
 10  Size                         17006 non-null  float64
 11  Genre                        17007 non-null  object
 12  Original Release Date        17007 non-null  object
 13  Current Version Release Date  17007 non-null  object
dtypes: float64(4), int64(1), object(9)
memory usage: 1.8+ MB
```

In [3]:
```
# 2. Count and show the number of apps in each genre
appStoreGames['Genre'].value_counts() #a quick way to count the unique values in a sing
```

Out[3]:
```
Games                16286
Education              222
Entertainment          198
Utilities               77
Sports                  60
Reference               32
Stickers                29
Finance                 18
Business                16
Productivity            15
Lifestyle               10
Book                     9
Social Networking        9
News                     7
Health & Fitness         5
Food & Drink             4
Music                    3
```

```
Navigation              3
Medical                 2
Travel                  1
Shopping                1
Name: Genre, dtype: int64
```

In [17]:

```
# 3. Show all statistics of user ratings and size on the same summary table

appStoreGames[['User Rating','Size']].describe()
```

Out[17]:

	User Rating	Size
count	7561.000000	1.700600e+04
mean	4.060905	1.157064e+08
std	0.751428	2.036477e+08
min	1.000000	5.132800e+04
25%	3.500000	2.295014e+07
50%	4.500000	5.676895e+07
75%	4.500000	1.330271e+08
max	5.000000	4.005591e+09

In [25]:

```
# 4. Filter the dataset to count and show all paid apps with ratings under 1.5

# All the paid apps that are greater than stored in paidAppsAboveZero
paidAppsAboveZero = appStoreGames['Price']>0
#appStoreGames[paidAppsAboveZero]

# All the apps which has User Rating below than 1.5
lowerRatings = appStoreGames['User Rating']<1.5

print("The total number of apps that are paid and User Rating is below 1.5 :",
      appStoreGames[lowerRatings & paidAppsAboveZero]['Name'].count())
```

```
The total number of apps that are paid and user rating is below 1.5 : 9
```

In [26]:

```
#Displaying the apps that costs above Zero and has User Rating below 1.5
appStoreGames[lowerRatings & paidAppsAboveZero]
```

Out[26]:

	ID	Name	Subtitle	User Rating	Count of Ratings	Price	In-app Purchases	Developer	Age Rating	L
1957	601301449	NORTH & SOUTH - The Game (Pocket Edition)	NaN	1.0	7.0	1.99	NaN	HandyGames	9+	
2141	633006816	Haunted Island : Mystic of Anarchy Wild Escape 3D	NaN	1.0	14.0	2.99	NaN	Palmacapp	12+	

	ID	Name	Subtitle	User Rating	Count of Ratings	Price	In-app Purchases	Developer	Age Rating	L
3984	884768672	Design CAD - create and edit DWG/DXF/CTM drawi...	NaN	1.0	7.0	9.99	7.99	li xiaolong	4+	
5572	969107262	Venture for iPad	NaN	1.0	5.0	2.99	NaN	Sandy Knoll Software, LLC	4+	
6103	993863316	Escape Mystery Haunted House -Scary Point & Cl...	NaN	1.0	7.0	1.99	NaN	One Connection Media	12+	
6422	1015151892	Nuts! The Battle of the Bulge	NaN	1.0	11.0	4.99	NaN	HexWar Games Ltd	12+	
7385	1062521052	Backgammon Skills	NaN	1.0	6.0	2.99	9.99	Tony Walsh	17+	
11011	1205436748	Harbor Master: Caribbean Merchant	NaN	1.0	8.0	0.99	NaN	Evgeni Petkov	4+	
12235	1258100671	Lightning: D-Day	NaN	1.0	6.0	4.99	NaN	HexWar Games Ltd	9+	

In [15]:

```
#5. Write Python code to calculate which percentage of the apps in the dataset are free
#Do not change Nan values to zero. Print the percentage on a 100 basis with no more tha


# Using '==' operator to extract all the free apps
appsBelowZero = appStoreGames['Price'] == 0
freeApps = appStoreGames[appsBelowZero]['Price'].count() # counting number of free apps
#print(freeApps)

# To count the total number of apps including the Nan values because non-null value in
totalApps = appStoreGames['Name'].count()
#print(totalApps)

freeAppsRate = (freeApps/totalApps)*100

print(f"The percentage of the free apps is {freeAppsRate:.2f}% .")
```

```
The percentage of the free apps is 83.57% .
```