In [1]:
```
# Data Analysis on Google Play Store
# Get TheDataFrame Data From Google Play Store Data Set
import numpy as np
import pandas as pd
```

In [2]:
```
df = pd.read_csv("C:\\Users\\user\\Desktop\\Pandas\\4.Case study and projects\\Project-1\\1.googleplaystore.csv")
print(df)
```

```
                                                 App             Category  \
0         Photo Editor & Candy Camera & Grid & ScrapBook       ART_AND_DESIGN
1                                    Coloring book moana       ART_AND_DESIGN
2      U Launcher Lite – FREE Live Cool Themes, Hide ...       ART_AND_DESIGN
3                                    Sketch - Draw & Paint       ART_AND_DESIGN
4                  Pixel Draw - Number Art Coloring Book       ART_AND_DESIGN
...                                                  ...                  ...
10836                                  Sya9a Maroc - FR               FAMILY
10837                  Fr. Mike Schmitz Audio Teachings               FAMILY
10838                           Parkinson Exercices FR              MEDICAL
10839                     The SCP Foundation DB fr nn5n  BOOKS_AND_REFERENCE
10840     iHoroscope - 2018 Daily Horoscope & Astrology            LIFESTYLE

       Rating Reviews                Size     Installs  Type Price  \
0         4.1     159                19M      10,000+  Free     0
1         3.9     967                14M     500,000+  Free     0
2         4.7   87510               8.7M   5,000,000+  Free     0
3         4.5  215644                25M  50,000,000+  Free     0
4         4.3     967               2.8M     100,000+  Free     0
...       ...     ...                ...          ...   ...   ...
10836     4.5      38                53M       5,000+  Free     0
10837     5.0       4               3.6M         100+  Free     0
10838     NaN       3               9.5M       1,000+  Free     0
10839     4.5     114  Varies with device       1,000+  Free     0
10840     4.5  398307                19M  10,000,000+  Free     0

      Content Rating                     Genres      Last Updated  \
0           Everyone               Art & Design   January 7, 2018
1           Everyone  Art & Design;Pretend Play  January 15, 2018
2           Everyone               Art & Design    August 1, 2018
3               Teen               Art & Design      June 8, 2018
4           Everyone    Art & Design;Creativity     June 20, 2018
...              ...                        ...               ...
10836       Everyone                  Education     July 25, 2017
10837       Everyone                  Education      July 6, 2018
10838       Everyone                    Medical  January 20, 2017
10839     Mature 17+          Books & Reference  January 19, 2015
10840       Everyone                  Lifestyle     July 25, 2018
```

```
             Current Ver         Android Ver
0                  1.0.0        4.0.3 and up
1                  2.0.0        4.0.3 and up
2                  1.2.4        4.0.3 and up
3      Varies with device          4.2 and up
4                    1.1          4.4 and up
...                  ...                 ...
10836               1.48          4.1 and up
10837                1.0          4.1 and up
10838                1.0          2.2 and up
10839  Varies with device  Varies with device
10840  Varies with device  Varies with device

[10841 rows x 13 columns]
```

In [3]:
```
pd.set_option("Display.max_columns",None)
df=pd.read_csv("C:\\Users\\user\\Desktop\\Pandas\\4.Case study and projects\\Project-1\\1.googleplaystore.csv")
print(df)
```

```
                                                 App             Category  \
0         Photo Editor & Candy Camera & Grid & ScrapBook       ART_AND_DESIGN
1                                    Coloring book moana       ART_AND_DESIGN
2      U Launcher Lite – FREE Live Cool Themes, Hide ...       ART_AND_DESIGN
3                                    Sketch - Draw & Paint       ART_AND_DESIGN
4                  Pixel Draw - Number Art Coloring Book       ART_AND_DESIGN
...                                                  ...                  ...
10836                                   Sya9a Maroc - FR              FAMILY
10837                  Fr. Mike Schmitz Audio Teachings               FAMILY
10838                           Parkinson Exercices FR               MEDICAL
10839                     The SCP Foundation DB fr nn5n  BOOKS_AND_REFERENCE
10840      iHoroscope - 2018 Daily Horoscope & Astrology             LIFESTYLE

       Rating Reviews                Size     Installs  Type Price  \
0         4.1     159                 19M      10,000+  Free     0
1         3.9     967                 14M     500,000+  Free     0
2         4.7   87510                8.7M   5,000,000+  Free     0
3         4.5  215644                 25M  50,000,000+  Free     0
4         4.3     967                2.8M     100,000+  Free     0
...       ...     ...                 ...          ...   ...   ...
10836     4.5      38                 53M       5,000+  Free     0
10837     5.0       4                3.6M         100+  Free     0
10838     NaN       3                9.5M       1,000+  Free     0
10839     4.5     114  Varies with device       1,000+  Free     0
10840     4.5  398307                 19M  10,000,000+  Free     0

      Content Rating                     Genres      Last Updated  \
0           Everyone               Art & Design   January 7, 2018
1           Everyone  Art & Design;Pretend Play  January 15, 2018
2           Everyone               Art & Design    August 1, 2018
```

```
3              Teen                Art & Design      June 8, 2018
4          Everyone    Art & Design;Creativity     June 20, 2018
...             ...                       ...               ...
10836      Everyone                 Education     July 25, 2017
10837      Everyone                 Education      July 6, 2018
10838      Everyone                   Medical  January 20, 2017
10839    Mature 17+          Books & Reference  January 19, 2015
10840      Everyone                 Lifestyle     July 25, 2018

              Current Ver         Android Ver
0                   1.0.0        4.0.3 and up
1                   2.0.0        4.0.3 and up
2                   1.2.4        4.0.3 and up
3      Varies with device          4.2 and up
4                     1.1          4.4 and up
...                   ...                 ...
10836                1.48          4.1 and up
10837                 1.0          4.1 and up
10838                 1.0          2.2 and up
10839  Varies with device  Varies with device
10840  Varies with device  Varies with device

[10841 rows x 13 columns]
```

In [4]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10841 entries, 0 to 10840
Data columns (total 13 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   App             10841 non-null  object 
 1   Category        10841 non-null  object 
 2   Rating          9367 non-null   float64
 3   Reviews         10841 non-null  object 
 4   Size            10841 non-null  object 
 5   Installs        10841 non-null  object 
 6   Type            10840 non-null  object 
 7   Price           10841 non-null  object 
 8   Content Rating  10840 non-null  object 
 9   Genres          10841 non-null  object 
 10  Last Updated    10841 non-null  object 
 11  Current Ver     10833 non-null  object 
 12  Android Ver     10838 non-null  object 
dtypes: float64(1), object(12)
memory usage: 1.1+ MB
```

In [5]:
```
# 1. Display Top 5 Rows of The Dataset
# Default Value is 5
```

```
df.head()
```

Out[5]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	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
1	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
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	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
3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	June 8, 2018	Varies with device	4.2 and up
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Design;Creativity	June 20, 2018	1.1	4.4 and up

In [6]:
```
# 1. Display Top 3 Rows of The Dataset
df.head(3)
```

Out[6]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	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
1	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
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	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

In [7]:
```
# 2. Check the Last 7 Rows of The Dataset
df.tail(7)
```

Out[7]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
10834	FR Calculator	FAMILY	4.0	7	2.6M	500+	Free	0	Everyone	Education	June 18, 2017	1.0.0	4.1 and up
10835	FR Forms	BUSINESS	NaN	0	9.6M	10+	Free	0	Everyone	Business	September 29, 2016	1.1.5	4.0 and up

10836	Sya9a Maroc - FR	FAMILY	4.5	38	53M	5,000+	Free	0	Everyone	Education	July 25, 2017	1.48	4.1 and up
10837	Fr. Mike Schmitz Audio Teachings	FAMILY	5.0	4	3.6M	100+	Free	0	Everyone	Education	July 6, 2018	1.0	4.1 and up
10838	Parkinson Exercices FR	MEDICAL	NaN	3	9.5M	1,000+	Free	0	Everyone	Medical	January 20, 2017	1.0	2.2 and up
10839	The SCP Foundation DB fr nn5n	BOOKS_AND_REFERENCE	4.5	114	Varies with device	1,000+	Free	0	Mature 17+	Books & Reference	January 19, 2015	Varies with device	Varies with device
10840	iHoroscope - 2018 Daily Horoscope & Astrology	LIFESTYLE	4.5	398307	19M	10,000,000+	Free	0	Everyone	Lifestyle	July 25, 2018	Varies with device	Varies with device

In [8]:
```
# 2. Check the Last 5 Rows of The Dataset
# Default Value is 5
df.tail()
```

Out[8]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
10836	Sya9a Maroc - FR	FAMILY	4.5	38	53M	5,000+	Free	0	Everyone	Education	July 25, 2017	1.48	4.1 and up
10837	Fr. Mike Schmitz Audio Teachings	FAMILY	5.0	4	3.6M	100+	Free	0	Everyone	Education	July 6, 2018	1.0	4.1 and up
10838	Parkinson Exercices FR	MEDICAL	NaN	3	9.5M	1,000+	Free	0	Everyone	Medical	January 20, 2017	1.0	2.2 and up
10839	The SCP Foundation DB fr nn5n	BOOKS_AND_REFERENCE	4.5	114	Varies with device	1,000+	Free	0	Mature 17+	Books & Reference	January 19, 2015	Varies with device	Varies with device
10840	iHoroscope - 2018 Daily Horoscope & Astrology	LIFESTYLE	4.5	398307	19M	10,000,000+	Free	0	Everyone	Lifestyle	July 25, 2018	Varies with device	Varies with device

In [9]:
```
# 3. Find Shape of Our Dataset (Number of Rows & Number of Columns)
df.shape
```

Out[9]:
```
(10841, 13)
```

In [10]:
```
#3. Find the Number of Apps in Google play store and columns( shape)
df.shape
```

```
print("Number of Apps=",df.shape[0])
print("Number of Cols=",df.shape[1])
```

```
Number of Apps= 10841
Number of Cols= 13
```

In [11]:
```
# 4. Display The Columns and All The Apps
df.columns
```

Out[11]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [12]:
```
for i in df.columns:
    print(i)
```

```
App
Category
Rating
Reviews
Size
Installs
Type
Price
Content Rating
Genres
Last Updated
Current Ver
Android Ver
```

In [13]:
```
# 5.Get the information abount our Dataset Like Total Number of Rows , \
# Total Number of Columns,Data types of Each column And Memory Requirement
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10841 entries, 0 to 10840
Data columns (total 13 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   App             10841 non-null  object 
 1   Category        10841 non-null  object 
 2   Rating          9367 non-null   float64
 3   Reviews         10841 non-null  object 
 4   Size            10841 non-null  object 
 5   Installs        10841 non-null  object 
 6   Type            10840 non-null  object 
 7   Price           10841 non-null  object 
 8   Content Rating  10840 non-null  object 
 9   Genres          10841 non-null  object 
```

```
 10  Last Updated    10841 non-null  object
 11  Current Ver     10833 non-null  object
 12  Android Ver     10838 non-null  object
dtypes: float64(1), object(12)
memory usage: 1.1+ MB
```

In [14]:
```
#6. Get the Statistical Information about Google Play Store Apps(Called Descriptive statistics)
df.describe() # Gives Descriptive statistical Information only on Numerical Columns
```

Out[14]:

	Rating
count	9367.000000
mean	4.193338
std	0.537431
min	1.000000
25%	4.000000
50%	4.300000
75%	4.500000
max	19.000000

In [15]:
```
#6. Get the Statistical Information about Google Play Store Apps(Called Descriptive statistics)
df.describe(include='all') # Gives Descriptive statistical Information of all Columns
```

Out[15]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
count	10841	10841	9367.000000	10841	10841	10841	10840	10841	10840	10841	10841	10833	10838
unique	9660	34	NaN	6002	462	22	3	93	6	120	1378	2832	33
top	ROBLOX	FAMILY	NaN	0	Varies with device	1,000,000+	Free	0	Everyone	Tools	August 3, 2018	Varies with device	4.1 and up
freq	9	1972	NaN	596	1695	1579	10039	10040	8714	842	326	1459	2451
mean	NaN	NaN	4.193338	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
std	NaN	NaN	0.537431	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
min	NaN	NaN	1.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
25%	NaN	NaN	4.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
50%	NaN	NaN	4.300000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75%	NaN	NaN	4.500000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
max	NaN	NaN	19.000000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [16]:
```
#7.Find the Name of Apps contains Astrology
#step-0  Gives all column names
df.columns
```

Out[16]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [17]:
```
# step-1: First find app names
df["App"] # displays all apps names values from apps column
```

Out[17]:
```
0          Photo Editor & Candy Camera & Grid & ScrapBook
1                                      Coloring book moana
2        U Launcher Lite – FREE Live Cool Themes, Hide ...
3                                    Sketch - Draw & Paint
4                     Pixel Draw - Number Art Coloring Book
                               ...
10836                                    Sya9a Maroc - FR
10837                     Fr. Mike Schmitz Audio Teachings
10838                              Parkinson Exercices FR
10839                        The SCP Foundation DB fr nn5n
10840         iHoroscope - 2018 Daily Horoscope & Astrology
Name: App, Length: 10841, dtype: object
```

In [18]:
```
# step-2: Find the app names contains Astrology word
df["App"].str.contains("Astrology") # Gives Boolean array
```

Out[18]:
```
0        False
1        False
2        False
3        False
4        False
         ...
10836    False
10837    False
10838    False
10839    False
10840     True
Name: App, Length: 10841, dtype: bool
```

In [19]:
```
# step-3: pass boolean array to DataFrame object
df[df["App"].str.contains("Astrology")] # Case sensitive
```

Out[19]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
1570	Horoscopes – Daily Zodiac	LIFESTYLE	4.6	161143	11M	10,000,000+	Free	0	Everyone	Lifestyle	June 25,	5.2.4(881)	4.0.3 and up

	Horoscope and Astrology								10+		2018		
1592	□□□□□ Astrology - Min Thein Kha BayDin	LIFESTYLE	4.7	2225	15M	100,000+	Free	0	Everyone	Lifestyle	July 26, 2018	4.2.1	4.0.3 and up
10840	iHoroscope - 2018 Daily Horoscope & Astrology	LIFESTYLE	4.5	398307	19M	10,000,000+	Free	0	Everyone	Lifestyle	July 25, 2018	Varies with device	Varies with device

In [20]:
```
#  OR
df[df["App"].str.contains("Astrology",case=False)]      # Ingore the Case sensitive
```

Out[20]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
1570	Horoscopes – Daily Zodiac Horoscope and Astrology	LIFESTYLE	4.6	161143	11M	10,000,000+	Free	0	Everyone 10+	Lifestyle	June 25, 2018	5.2.4(881)	4.0.3 and up
1592	□□□□□ Astrology - Min Thein Kha BayDin	LIFESTYLE	4.7	2225	15M	100,000+	Free	0	Everyone	Lifestyle	July 26, 2018	4.2.1	4.0.3 and up
10840	iHoroscope - 2018 Daily Horoscope & Astrology	LIFESTYLE	4.5	398307	19M	10,000,000+	Free	0	Everyone	Lifestyle	July 25, 2018	Varies with device	Varies with device

In [21]:
```
# 8.Find the total number of unique Category
df['Category']
```

Out[21]:
```
0              ART_AND_DESIGN
1              ART_AND_DESIGN
2              ART_AND_DESIGN
3              ART_AND_DESIGN
4              ART_AND_DESIGN
                  ...
10836                  FAMILY
10837                  FAMILY
10838                 MEDICAL
10839    BOOKS_AND_REFERENCE
10840               LIFESTYLE
Name: Category, Length: 10841, dtype: object
```

In [22]:
```
# 8.Find the total number of Unique Category
uv=df['Category'].nunique()
print("Total number of Unique Category= ",uv)
```

```
Total number of Unique Category=  34
```

In [23]:
```
# 9. Find which Category getting Highest Average Rating
for gn,ginfo in df.groupby("Category"):
    print("Group Name:",gn)
    print("Gropu Information")
    print(ginfo)
```

```
...        ...                 ...                 ...                 ...
10612  Weather     April 12, 2018                 1.0          5.0 and up
10625  Weather     April 13, 2018                 1.0          5.0 and up
10627  Weather      June 21, 2018           4.6.1404          4.4 and up
10713  Weather      July 24, 2018  Varies with device  Varies with device
10832  Weather  February 16, 2014                 6.0          2.1 and up

[82 rows x 13 columns]
```

In [24]:

```
# 9. Find which Category getting Highest Average Rating
for x in df.groupby("Category")["Rating"]:
    print(x)
```

```
('1.9', 10472    19.0
Name: Rating, dtype: float64)
('ART_AND_DESIGN', 0       4.1
1       3.9
2       4.7
3       4.5
4       4.3
       ...
7174    NaN
8679    4.2
8712    4.3
8871    NaN
8888    5.0
Name: Rating, Length: 65, dtype: float64)
('AUTO_AND_VEHICLES', 49       4.2
50       4.0
51       3.8
52       4.6
53       3.9
        ...
10137    2.1
10345    3.9
10486    4.6
10764    NaN
10811    NaN
Name: Rating, Length: 85, dtype: float64)
('BEAUTY', 98       4.7
99       4.9
100      4.7
101      3.9
102      3.9
103      4.2
104      4.6
105      4.3
106      4.7
107      4.7
```

```
3667     4.2
3668     4.3
3669     3.6
        ...
10237    4.2
10318    4.4
10519    NaN
10647    3.9
10823    NaN
Name: Rating, Length: 175, dtype: float64)
('WEATHER', 3626     4.4
3627     4.8
3628     4.4
3629     4.8
3630     4.5
        ...
10612    NaN
10625    NaN
10627    NaN
10713    4.4
10832    3.8
Name: Rating, Length: 82, dtype: float64)
```

In [25]:
```
df.groupby("Category")["Rating"].mean()
```

Out[25]:
```
Category
1.9                    19.000000
ART_AND_DESIGN          4.358065
AUTO_AND_VEHICLES       4.190411
BEAUTY                  4.278571
BOOKS_AND_REFERENCE     4.346067
BUSINESS                4.121452
COMICS                  4.155172
COMMUNICATION           4.158537
DATING                  3.970769
EDUCATION               4.389032
ENTERTAINMENT           4.126174
EVENTS                  4.435556
FAMILY                  4.192272
FINANCE                 4.131889
FOOD_AND_DRINK          4.166972
GAME                    4.286326
HEALTH_AND_FITNESS      4.277104
HOUSE_AND_HOME          4.197368
LIBRARIES_AND_DEMO      4.178462
LIFESTYLE               4.094904
MAPS_AND_NAVIGATION     4.051613
MEDICAL                 4.189143
NEWS_AND_MAGAZINES      4.132189
```

```
PARENTING                4.300000
PERSONALIZATION          4.335987
PHOTOGRAPHY              4.192114
PRODUCTIVITY             4.211396
SHOPPING                 4.259664
SOCIAL                   4.255598
SPORTS                   4.223511
TOOLS                    4.047411
TRAVEL_AND_LOCAL         4.109292
VIDEO_PLAYERS            4.063750
WEATHER                  4.244000
Name: Rating, dtype: float64
```

In [26]:
```
df.groupby("Category")["Rating"].mean().sort_values(ascending=False)
```

Out[26]:
```
Category
1.9                    19.000000
EVENTS                  4.435556
EDUCATION               4.389032
ART_AND_DESIGN          4.358065
BOOKS_AND_REFERENCE     4.346067
PERSONALIZATION         4.335987
PARENTING               4.300000
GAME                    4.286326
BEAUTY                  4.278571
HEALTH_AND_FITNESS      4.277104
SHOPPING                4.259664
SOCIAL                  4.255598
WEATHER                 4.244000
SPORTS                  4.223511
PRODUCTIVITY            4.211396
HOUSE_AND_HOME          4.197368
FAMILY                  4.192272
PHOTOGRAPHY             4.192114
AUTO_AND_VEHICLES       4.190411
MEDICAL                 4.189143
LIBRARIES_AND_DEMO      4.178462
FOOD_AND_DRINK          4.166972
COMMUNICATION           4.158537
COMICS                  4.155172
NEWS_AND_MAGAZINES      4.132189
FINANCE                 4.131889
ENTERTAINMENT           4.126174
BUSINESS                4.121452
TRAVEL_AND_LOCAL        4.109292
LIFESTYLE               4.094904
VIDEO_PLAYERS           4.063750
MAPS_AND_NAVIGATION     4.051613
TOOLS                   4.047411
```

```
DATING                 3.970769
Name: Rating, dtype: float64
```

In [27]:
```
# 9. Find which Category getting Highest Average Rating
df.groupby("Category")["Rating"].mean().sort_values(ascending=False).head(1)
```

Out[27]:
```
Category
1.9    19.0
Name: Rating, dtype: float64
```

In [28]:
```
# 9. Find which Category getting Highest Average Rating
df.groupby("Category")["Rating"].mean().sort_values(ascending=False).tail(1)
```

Out[28]:
```
Category
DATING    3.970769
Name: Rating, dtype: float64
```

In [29]:
```
# 10.Find the total number App having 5 star Rating Apps
df["Rating"]
```

Out[29]:
```
0        4.1
1        3.9
2        4.7
3        4.5
4        4.3
        ...
10836    4.5
10837    5.0
10838    NaN
10839    4.5
10840    4.5
Name: Rating, Length: 10841, dtype: float64
```

In [30]:
```
df["Rating"]==5
```

Out[30]:
```
0        False
1        False
2        False
3        False
4        False
         ...
10836    False
10837     True
10838    False
10839    False
10840    False
Name: Rating, Length: 10841, dtype: bool
```

In [31]:
```
# 10.Find the total number App having 5 star Rating Apps
star=len(df[df["Rating"]==5])
```

```
print("Total Number of App Having 5 star Rating Apps=",star)
```

```
Total Number of App Having 5 star Rating Apps= 274
```

In [32]:
```
# 11.Find the Average Values of Reviews
df["Reviews"].dtype
```

Out[32]:
```
dtype('O')
```

In [33]:
```
df[df["Reviews"]=="3.0M"]
```

Out[33]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
10472	Life Made WI-Fi Touchscreen Photo Frame	1.9	19.0	3.0M	1,000+	Free	0	Everyone	NaN	February 11, 2018	1.0.19	4.0 and up	NaN

In [34]:
```
df["Reviews"]=df["Reviews"].replace("3.0M",3.0)
```

In [35]:
```
df[df["Reviews"]==3.0]
```

Out[35]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
10472	Life Made WI-Fi Touchscreen Photo Frame	1.9	19.0	3.0	1,000+	Free	0	Everyone	NaN	February 11, 2018	1.0.19	4.0 and up	NaN

In [36]:
```
# 11. Find the Average Values of Reviews
df["Reviews"].dtype
```

Out[36]:
```
dtype('O')
```

In [37]:
```
df["Reviews"]=df["Reviews"].astype("float")
```

In [38]:
```
df["Reviews"].dtype
```

Out[38]:
```
dtype('float64')
```

In [39]:
```
Av=df["Reviews"].mean()
print("The average Values of Reviews=",Av)
```

```
The average Values of Reviews= 444111.9265750392
```

In [40]:
```
pd.set_option("Display.Max_Columns",None)
df=pd.read_csv("C:\\Users\\user\\Desktop\\Pandas\\4.Case study and projects\\Project-1\\1.googleplaystore.csv")
print(df)
```

```
10838                1.0         2.2 and up
10839  Varies with device  Varies with device
10840  Varies with device  Varies with device

[10841 rows x 13 columns]
```

In [41]:
```
# 12. Find Total Number of Apps Free and Paid Apps
df[df["Type"]=="Paid"]
```

Out[41]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
234	TurboScan: scan documents and receipts in PDF	BUSINESS	4.7	11442	6.8M	100,000+	Paid	$4.99	Everyone	Business	March 25, 2018	1.5.2	4.0 and up
235	Tiny Scanner Pro: PDF Doc Scan	BUSINESS	4.8	10295	39M	100,000+	Paid	$4.99	Everyone	Business	April 11, 2017	3.4.6	3.0 and up
290	TurboScan: scan documents and receipts in PDF	BUSINESS	4.7	11442	6.8M	100,000+	Paid	$4.99	Everyone	Business	March 25, 2018	1.5.2	4.0 and up
291	Tiny Scanner Pro: PDF Doc Scan	BUSINESS	4.8	10295	39M	100,000+	Paid	$4.99	Everyone	Business	April 11, 2017	3.4.6	3.0 and up
427	Puffin Browser Pro	COMMUNICATION	4.0	18247	Varies with device	100,000+	Paid	$3.99	Everyone	Communication	July 5, 2018	7.5.3.20547	4.1 and up
...	...	...	...	...	...	...	...	...	...	...	...	...	...
10735	FP VoiceBot	FAMILY	NaN	17	157k	100+	Paid	$0.99	Mature 17+	Entertainment	November 25, 2015	1.2	2.1 and up
10760	Fast Tract Diet	HEALTH_AND_FITNESS	4.4	35	2.4M	1,000+	Paid	$7.99	Everyone	Health & Fitness	August 8, 2018	1.9.3	4.2 and up
10782	Trine 2: Complete Story	GAME	3.8	252	11M	10,000+	Paid	$16.99	Teen	Action	February 27, 2015	2.22	5.0 and up
10785	sugar, sugar	FAMILY	4.2	1405	9.5M	10,000+	Paid	$1.20	Everyone	Puzzle	June 5, 2018	2.7	2.3 and up
10798	Word Search Tab 1 FR	FAMILY	NaN	0	1020k	50+	Paid	$1.04	Everyone	Puzzle	February 6, 2012	1.1	3.0 and up

800 rows × 13 columns

In [42]:
```
# 12. Find Total Number of Apps Free and Paid Apps
df["Type"].value_counts()
```

Out[42]:
```
Type
Free    10039
Paid      800
0           1
Name: count, dtype: int64
```

In [43]:
```
# 13.Find Which App Has Maximum Review
df["Reviews"]
```

Out[43]:
```
0           159
1           967
2         87510
3        215644
4           967
          ...
10836        38
10837         4
10838         3
10839       114
10840    398307
Name: Reviews, Length: 10841, dtype: object
```

In [44]:
```
df["Reviews"].max()
```

Out[44]:
```
'9992'
```

In [45]:
```
df[df["Reviews"].max()==df["Reviews"]]
```

Out[45]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
2989	GollerCepte Live Score	SPORTS	4.2	9992	31M	1,000,000+	Free	0	Everyone	Sports	May 23, 2018	6.5	4.1 and up

In [46]:
```
df[df["Reviews"].max()==df["Reviews"]]["App"]
```

Out[46]:
```
2989    GollerCepte Live Score
Name: App, dtype: object
```

In [47]:
```
# 14. Find the Top 5 App Names having highest Review
indices=df["Reviews"].sort_values(ascending=False).head().index
df.iloc[indices]["App"]
```

Out[47]:
```
2989              GollerCepte Live Score
4970        Ad Block REMOVER - NEED ROOT
2723                 SnipSnap Coupon App
2705                 SnipSnap Coupon App
```

```
3079    US Open Tennis Championships 2018
Name: App, dtype: object
```

In [48]:
```
# OR
df.loc[indices]["App"]
```

Out[48]:
```
2989               GollerCepte Live Score
4970         Ad Block REMOVER - NEED ROOT
2723                   SnipSnap Coupon App
2705                   SnipSnap Coupon App
3079    US Open Tennis Championships 2018
Name: App, dtype: object
```

In [49]:
```
# 15.Find The Average Rating of free and Paid Apps
df.groupby("Type")["Rating"].mean()
```

Out[49]:
```
Type
0       19.000000
Free     4.186203
Paid     4.266615
Name: Rating, dtype: float64
```

In [50]:
```
# 16. Find Top 5 Apps Having Maximum Installs
df["newInstalls"]=df["Installs"].str.replace(",","")
df["newInstalls"]
```

Out[50]:
```
0          10000+
1         500000+
2        5000000+
3       50000000+
4         100000+
           ...
10836       5000+
10837        100+
10838       1000+
10839       1000+
10840    10000000+
Name: newInstalls, Length: 10841, dtype: object
```

In [51]:
```
df["newInstalls"].str.replace("+","")
```

Out[51]:
```
0          10000
1         500000
2        5000000
3       50000000
4         100000
           ...
10836       5000
10837        100
10838       1000
```

```
10839        1000
10840    10000000
Name: newInstalls, Length: 10841, dtype: object
```

In [52]:
```
df["newInstalls"]=df["newInstalls"].str.replace("+","")
df["newInstalls"]
```

Out[52]:
```
0           10000
1          500000
2         5000000
3        50000000
4          100000
           ...
10836        5000
10837         100
10838        1000
10839        1000
10840    10000000
Name: newInstalls, Length: 10841, dtype: object
```

In [53]:
```
df["newInstalls"].dtype
```

Out[53]:
```
dtype('O')
```

In [54]:
```
df.loc[df["newInstalls"]=="Free"]
```

Out[54]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver	newInstalls
10472	Life Made WI-Fi Touchscreen Photo Frame	1.9	19.0	3.0M	1,000+	Free	0	Everyone	NaN	February 11, 2018	1.0.19	4.0 and up	NaN	Free

In [55]:
```
df["newInstalls"]=df["newInstalls"].str.replace("Free","0")
```

In [56]:
```
df["newInstalls"].astype(int)
```

Out[56]:
```
0           10000
1          500000
2         5000000
3        50000000
4          100000
           ...
10836        5000
10837         100
10838        1000
10839        1000
```

```
10840    10000000
Name: newInstalls, Length: 10841, dtype: int32
```

In [57]:
```
df["newInstalls"].sort_values().tail()
```

Out[57]:
```
3796    500000000
1702    500000000
1885    500000000
1662    500000000
385     500000000
Name: newInstalls, dtype: object
```

In [58]:
```
indices=df["newInstalls"].sort_values().tail().index
print(indices)
```

```
Index([3796, 1702, 1885, 1662, 385], dtype='int64')
```

In [59]:
```
df.iloc[indices]["App"]
```

Out[59]:
```
3796            Twitter
1702                Pou
1885                Pou
1662                Pou
385     Viber Messenger
Name: App, dtype: object
```

In [60]:
```
df.loc[indices][["App","Installs"]]
```

Out[60]:

	App	Installs
3796	Twitter	500,000,000+
1702	Pou	500,000,000+
1885	Pou	500,000,000+
1662	Pou	500,000,000+
385	Viber Messenger	500,000,000+