In [1]:
```
import pandas as pd
```

In [2]:
```
data = pd.read_csv('googleplaystore.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	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 [4]:
```
data.tail(3)
```

Out[4]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
10838	Parkinson Exercises 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 [5]:
```
data.shape
```

Out[5]:
```
(10841, 13)
```

In [6]:
```
print('Number of Rows', data.shape[0])
print('Number of Columns', data.shape[1])
```

```
Number of Rows 10841
Number of Columns 13
```

In [7]:
```
data.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 [11]:
```
data.describe(include='all')
```

Out[11]:

	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

Total Number of App Titles Contain Astrology

In [12]:
```
data.columns
```

Out[12]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [24]:
```
data[data['App'].str.contains('Astrology', case=False)]
```

Out[24]:

	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	101143	11M	10,000,000+	Free	0	Everyone 10+	Lifestyle	June 25, 2018	5.2.A0351	4.0.3 and up
1582	Astrology - Min Them Kha Baghn	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

Total Number of Astrology Titles

In [25]:
```
len(data[data['App'].str.contains('Astrology', case=False)])
```

Out[25]:
```
3
```

Find Average App Rating

In [21]:
```
data['Rating'].mean()
```

Out[21]:
```
4.193338315362448
```

Find Total Number of Unique Category

In [29]:
```
data['Category'].nunique()
```

Out[29]:
```
34
```

Which Category Getting The Highest Average Rating?

In [30]:
```
data.columns
```

Out[30]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [36]:
```
data.groupby('Category')['Rating'].mean().sort_values(ascending=False)
```

Out[36]:
```
Category
1.9                  19.000000
EVENTS               4.435556
EDUCATION            4.389032
ART_AND_DESIGN       4.358065
BOOKS_AND_REFERENCE  4.346067
PERSONALIZATION      4.335087
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
FOOD_AND_DRINK       4.166972
FAMILY               4.191757
PHOTOGRAPHY          4.192114
AUTO_AND_VEHICLES    4.190411
MEDICAL              4.189143
LIBRARIES_AND_DEMO   4.178462
FOOD_AND_DRINK       4.166972
COMMUNICATION        4.158537
HOUSE_AND_HOME       4.150000
NEWS_AND_MAGAZINES   4.132189
FINANCE              4.131889
LIFESTYLE            4.117914
BUSINESS             4.121452
TRAVEL_AND_LOCAL     4.109292
DATING               3.970769
VIDEO_PLAYERS        4.063750
MAPS_AND_NAVIGATION  4.051613
TOOLS                4.047411
DATING               3.970769
Name: Rating, dtype: float64
```

Total Find Number of App having 5 Star Rating

In [37]:
```
data.columns
```

Out[37]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [43]:
```
len(data[data['Rating'] == 5.0])
```

Out[43]:
```
274
```

Find Average Value of Reviews

In [49]:
```
data['Reviews'].dtype
```

Out[49]:
```
dtype('O')
```

In [50]:
```
data['Reviews'].astype('float')
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
<ipython-input-55-dd7f1a04a7b7> in <module>
----> 1 data['Reviews'].astype('float')

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py in astype(self, dtype, copy, errors)
   5875         else:
   5876             # else, only a single dtype is given
-> 5877             new_data = self._mgr.astype(dtype=dtype, copy=copy, errors=errors)
   5878             return self._constructor(new_data).__finalize__(self, method="astype")
   5879

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\internals\managers.py in astype(self, dtype, copy, errors)
    629         self, dtype, copy: bool = False, errors: str = "raise"
    630     ) -> "BlockManager":
--> 631         return self.apply("astype", dtype=dtype, copy=copy, errors=errors)
    632
    633     def convert(

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\internals\managers.py in apply(self, f, align_keys, ignore_failures, **kwargs)
    425                 applied = b.apply(f, **kwargs)
    426             else:
--> 427                 applied = getattr(b, f)(**kwargs)
    428         except (TypeError, NotImplementedError):
    429             if not ignore_failures:

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\internals\blocks.py in astype(self, dtype, copy, errors)
    671             vals1d = values.ravel()
    672             try:
--> 673                 values = astype_nansafe(vals1d, dtype, copy=True)
    674             except (ValueError, TypeError):
    675                 # e.g. astype_nansafe can fail on object-dtype of strings

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\dtypes\cast.py in astype_nansafe(arr, dtype, copy, skipna)
   1099     if copy or is_object_dtype(arr) or is_object_dtype(dtype):
   1100         # Explicit copy, or required since NumPy can't view from / to object.
-> 1101         return arr.astype(dtype, copy=True)
   1102
   1103     return arr.view(dtype)

ValueError: could not convert string to float: '3.0M'
```

In [58]:
```
data[data['Reviews'] == '3.0M']
```

Out[58]:

	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 [59]:
```
data['Reviews'] = data['Reviews'].replace('3.0M', 3.0)
```

In [62]:
```
data['Reviews'] = data['Reviews'].astype('float')
```

In [64]:
```
data['Reviews'].dtype
```

Out[64]:
```
dtype('float64')
```

In [67]:
```
data['Reviews'].mean()
```

Out[67]:
```
444111.9265730392
```

Find Total Number of Free and Paid Apps

In [68]:
```
data.columns
```

Out[68]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [69]:
```
data['Type'].value_counts()
```

Out[69]:
```
Free    10039
Paid      800
0           1
Name: Type, dtype: int64
```

Which App Has Maximum Reviews?

In [70]:
```
data[data['Reviews'].max() == data['Reviews']]['App']
```

Out[70]:
```
2544    Facebook
Name: App, dtype: object
```

Display Top 5 Apps Having Highest Reviews

In [108]:
```
index = data['Reviews'].sort_values(ascending=False).head().index

index
```

Out[108]:
```
Int64Index([2544, 3943, 381, 336, 3984], dtype='int64')
```

In [109]:
```
data.iloc[index]
```

Out[109]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
2544	Facebook	SOCIAL	4.1	78158306.0	Varies with device	1,000,000,000+	Free	0	Teen	Social	August 3, 2018	Varies with device	Varies with device
3943	Facebook	COMMUNICATION	4.1	78158306.0	Varies with device	1,000,000,000+	Free	0	Teen	Social	August 3, 2018	Varies with device	Varies with device
381	WhatsApp Messenger	COMMUNICATION	4.4	69119316.0	Varies with device	1,000,000,000+	Free	0	Everyone	Communication	August 3, 2018	Varies with device	Varies with device
336	WhatsApp Messenger	COMMUNICATION	4.4	69119316.0	Varies with device	1,000,000,000+	Free	0	Everyone	Communication	August 3, 2018	Varies with device	Varies with device
3984	WhatsApp Messenger	COMMUNICATION	4.4	69119316.0	Varies with device	1,000,000,000+	Free	0	Everyone	Communication	August 3, 2018	Varies with device	Varies with device

In [110]:
```
data.iloc[index]['App']
```

Out[110]:
```
2544        Facebook
3943        Facebook
381     WhatsApp Messenger
336     WhatsApp Messenger
3984    WhatsApp Messenger
Name: App, dtype: object
```

Find Average Rating of Free and Paid Apps

In [111]:
```
data.head(1)
```

Out[111]:

	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	159.0	19M	10,000+	Free	0	Everyone	Art & Design	January 7, 2018	1.0.0	4.0.3 and up

In [113]:
```
data.groupby('Type')['Rating'].mean()
```

Out[113]:
```
Type
0       19.000000
Free     4.186203
Paid     4.266615
Name: Rating, dtype: float64
```

Display Top 5 Apps Having Maximum Installs

In [118]:
```
data.columns
```

Out[118]:
```
Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type',
       'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver',
       'Android Ver'],
      dtype='object')
```

In [127]:
```
data['Installs'].dtype
```

Out[127]:
```
dtype('O')
```

In [128]:
```
data['Installs']
```

Out[128]:
```
0            10,000+
1           500,000+
2         5,000,000+
3        50,000,000+
4           100,000+
              ...
10836         5,000+
10837           100+
10838         1,000+
10839           100+
10840    10,000,000+
Name: Installs, Length: 10841, dtype: object
```

In [134]:
```
data['Installs'] = data['Installs'].str.replace(',', '')
```

In [135]:
```
data['Installs']
```

Out[135]:
```
0            10000+
1           500000+
2          5000000+
3         50000000+
4           100000+
              ...
10836         5000+
10837          100+
10838         1000+
10839          100+
10840     10000000+
Name: Installs, Length: 10841, dtype: object
```

In [136]:
```
data['Installs'] = data['Installs'].str.replace('+', '')
```

```
<ipython-input-136-dd2c439fc082>:1: FutureWarning: The default value of regex will change from True to False in a future version. In addition, single character regular expressions will*not* be treated as literal strings when regex=True.
  data['Installs'] = data['Installs'].str.replace('+', '')
```

In [138]:
```
data['Installs'].dtype
```

Out[138]:
```
dtype('O')
```

In [140]:
```
data[data['Installs'] == 'Free']
```

Out[140]:

	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 [144]:
```
data['Installs'] = data['Installs'].str.replace('Free', '0')
```

In [146]:
```
data['Installs'] = data['Installs'].astype('int')
```

In [154]:
```
data['Installs'].dtype
```

Out[154]:
```
dtype('int32')
```

In [160]:
```
index = data['Installs'].sort_values(ascending=False).head().index

index
```

In [165]:
```
data.iloc[index]['App']
```

Out[165]:
```
3936                          Subway Surfers
3943                                Facebook
335    Messenger – Text and Video Chat for Free
3523                            Google Drive
3947                             Google Drive
Name: App, dtype: object
```

In [ ]:
```

```