

googleplaystore Analysis

Import library

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib import style
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline
```

To read the file

In [2]:

```
p=pd.read_excel(r'C:\LINGESH\Data\googleplaystore(edited with excel).xls')
```

to view all data

In [3]:



p

Out[3]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Co R
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND DESIGN	4.1	159	19	10000	Free	0.00	Eve
1	Coloring book moana	ART_AND DESIGN	3.9	967	14	500000	Free	0.00	Eve
2	U Launcher Lite â€“ FREE Live Cool Themes, Hide...	ART_AND DESIGN	4.7	87510	8.7	5000000	Free	0.00	Eve
3	Sketch - Draw & Paint	ART_AND DESIGN	4.5	215644	25	50000000	Free	0.00	
4	Pixel Draw - Number Art Coloring Book	ART_AND DESIGN	4.3	967	2.8	100000	Free	0.00	Eve
...
10834	FP VoiceBot	FAMILY	NaN	17	157k	100	Paid	0.99	M
10835	Fast Tract Diet	HEALTH_AND_FITNESS	4.4	35	2.4	1000	Paid	7.99	Eve
10836	Trine 2: Complete Story	GAME	3.8	252	11	10000	Paid	16.99	
10837	sugar, sugar	FAMILY	4.2	1405	9.5	10000	Paid	1.20	Eve
10838	Word Search Tab 1 FR	FAMILY	NaN	0	1020k	50	Paid	1.04	Eve

10839 rows × 13 columns

In [6]:



p1=p.head(500)

In [7]:



```
p1
```

Out[7]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genre
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND DESIGN	4.1	159	19	10000	Free	0.0	Everyone	Art & Design
1	Coloring book moana	ART_AND DESIGN	3.9	967	14	500000	Free	0.0	Everyone	Art Design;Preteen;Play
2	FREE Live Cool U Launcher Lite	ART_AND DESIGN	4.7	87510	8.7	5000000	Free	0.0	Everyone	Art & Design

In [8]:



```
p2=p.tail(10)
```

In [9]:

p2

Out[9]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating
10829	Circle Colors Pack-FN Theme	PERSONALIZATION	4.2	6	89k	50	Paid	0.99	Everyone
10830	Solitaire+	GAME	4.6	11235	0	100000	Paid	2.99	Everyone
10831	Fruit Ninja Classic	GAME	4.3	85468	36	1000000	Paid	0.99	Everyone
10832	FO Bixby	PERSONALIZATION	5.0	5	861k	100	Paid	0.99	Everyone
10833	Mu.F.O.	GAME	5.0	2	16	1	Paid	0.99	Everyone
10834	FP VoiceBot	FAMILY	NaN	17	157k	100	Paid	0.99	Mat 1
10835	Fast Tract Diet	HEALTH_AND_FITNESS	4.4	35	2.4	1000	Paid	7.99	Everyone
10836	Trine 2: Complete Story	GAME	3.8	252	11	10000	Paid	16.99	Teen
10837	sugar, sugar	FAMILY	4.2	1405	9.5	10000	Paid	1.20	Everyone
10838	Word Search Tab 1 FR	FAMILY	NaN	0	1020k	50	Paid	1.04	Everyone

To change the column name

In [10]:

```
p.rename(columns={'Price':'APP Price'}, inplace=True)
```

In [11]:



p

Out[11]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Co R
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND DESIGN	4.1	159	19	10000	Free	0.00	Eve
1	Coloring book moana	ART_AND DESIGN	3.9	967	14	500000	Free	0.00	Eve
2	U Launcher Lite â€“ FREE Live Cool Themes, Hide...	ART_AND DESIGN	4.7	87510	8.7	5000000	Free	0.00	Eve
3	Sketch - Draw & Paint	ART_AND DESIGN	4.5	215644	25	50000000	Free	0.00	
4	Pixel Draw - Number Art Coloring Book	ART_AND DESIGN	4.3	967	2.8	100000	Free	0.00	Eve
...
10834	FP VoiceBot	FAMILY	NaN	17	157k	100	Paid	0.99	M
10835	Fast Tract Diet	HEALTH_AND_FITNESS	4.4	35	2.4	1000	Paid	7.99	Eve
10836	Trine 2: Complete Story	GAME	3.8	252	11	10000	Paid	16.99	
10837	sugar, sugar	FAMILY	4.2	1405	9.5	10000	Paid	1.20	Eve
10838	Word Search Tab 1 FR	FAMILY	NaN	0	1020k	50	Paid	1.04	Eve

10839 rows × 13 columns

To print top 5 rows

In [12]:



```
p.head(5)
```

Out[12]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND DESIGN	4.1	159	19	10000	Free	0.0	Everyone
1	Coloring book moana	ART_AND DESIGN	3.9	967	14	500000	Free	0.0	Everyone
2	U Launcher Lite â€“ FREE Live Cool Themes, Hide...	ART_AND DESIGN	4.7	87510	8.7	5000000	Free	0.0	Everyone
3	Sketch - Draw & Paint	ART_AND DESIGN	4.5	215644	25	50000000	Free	0.0	Teen
4	Pixel Draw - Number Art Coloring Book	ART_AND DESIGN	4.3	967	2.8	100000	Free	0.0	Everyone

To print bottom 5 rows

In [13]:



```
p.tail(5)
```

Out[13]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating
10834	FP VoiceBot	FAMILY	NaN	17	157k	100	Paid	0.99	Mature 17+
10835	Fast Tract Diet	HEALTH_AND_FITNESS	4.4	35	2.4	1000	Paid	7.99	Everyone
10836	Trine 2: Complete Story	GAME	3.8	252	11	10000	Paid	16.99	Teen
10837	sugar, sugar	FAMILY	4.2	1405	9.5	10000	Paid	1.20	Everyone
10838	Word Search Tab 1 FR	FAMILY	NaN	0	1020k	50	Paid	1.04	Everyone

To check how many columns and rows

In [14]:



```
p.shape
```

Out[14]:

```
(10839, 13)
```

to check the data types of attributes

In [15]:



```
p.dtypes
```

Out[15]:

```
App          object
Category     object
Rating       float64
Reviews      int64
Size         object
Installs     int64
Type         object
APP Price    float64
Content Rating  object
Genres        object
Last Updated   object
Current Ver    object
Android Ver    object
dtype: object
```

To print the number of dimensions

In [16]:

```
p.ndim
```

Out[16]:

```
2
```

To print the index

In [17]:

```
p.index
```

Out[17]:

```
RangeIndex(start=0, stop=10839, step=1)
```

To check the size of the data

In [18]:

```
p.size
```

Out[18]:

```
140907
```

To check the whether data type is empty or not

In [19]:

```
p.empty
```

Out[19]:

```
False
```

Print the column names

In [20]:



```
p.columns
```

Out[20]:

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

print the information of dataset

In [21]:



```
p.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10839 entries, 0 to 10838
Data columns (total 13 columns):
 #   Column            Non-Null Count  Dtype  
 ---  -- 
 0   App               10838 non-null   object  
 1   Category          10839 non-null   object  
 2   Rating             9366 non-null   float64 
 3   Reviews            10839 non-null   int64   
 4   Size               10839 non-null   object  
 5   Installs           10839 non-null   int64   
 6   Type               10839 non-null   object  
 7   APP Price          10839 non-null   float64 
 8   Content Rating    10839 non-null   object  
 9   Genres              10839 non-null   object  
 10  Last Updated       10839 non-null   object  
 11  Current Ver        10831 non-null   object  
 12  Android Ver        10837 non-null   object  
dtypes: float64(2), int64(2), object(9)
memory usage: 1.1+ MB
```

Print statistical information

In [22]:



```
p.describe()
```

Out[22]:

	Rating	Reviews	Installs	APP Price
count	9366.000000	1.083900e+04	1.083900e+04	10839.000000
mean	4.191757	4.441939e+05	1.546577e+07	1.027463
std	0.515219	2.927893e+06	8.503315e+07	15.950436
min	1.000000	0.000000e+00	0.000000e+00	0.000000
25%	4.000000	3.800000e+01	3.000000e+03	0.000000
50%	4.300000	2.094000e+03	1.000000e+05	0.000000
75%	4.500000	5.478300e+04	5.000000e+06	0.000000
max	5.000000	7.815831e+07	1.000000e+09	400.000000

Print all statistical information

In [23]:



```
p.describe(include='all')
```

Out[23]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP
count	10838	10839	9366.000000	1.083900e+04	10839.0	1.083900e+04	10839	10839.0
unique	9657	33	NaN	NaN	460.0	NaN	2	NaN
top	ROBLOX	FAMILY	NaN	NaN	0.0	NaN	Free	NaN
freq	9	1971	NaN	NaN	1694.0	NaN	10039	NaN
mean	NaN	NaN	4.191757	4.441939e+05	NaN	1.546577e+07	NaN	1.027463
std	NaN	NaN	0.515219	2.927893e+06	NaN	8.503315e+07	NaN	15.950436
min	NaN	NaN	1.000000	0.000000e+00	NaN	0.000000e+00	NaN	0.000000
25%	NaN	NaN	4.000000	3.800000e+01	NaN	3.000000e+03	NaN	0.000000
50%	NaN	NaN	4.300000	2.094000e+03	NaN	1.000000e+05	NaN	0.000000
75%	NaN	NaN	4.500000	5.478300e+04	NaN	5.000000e+06	NaN	0.000000
max	NaN	NaN	5.000000	7.815831e+07	NaN	1.000000e+09	NaN	400.000000

To check is there any missing value

In [24]:



```
p.isna()
```

Out[24]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating	Genres	Last Updated
0	False	False	False	False	False	False	False	False	False	False	Fal
1	False	False	False	False	False	False	False	False	False	False	Fal
2	False	False	False	False	False	False	False	False	False	False	Fal
3	False	False	False	False	False	False	False	False	False	False	Fal
4	False	False	False	False	False	False	False	False	False	False	Fal
...
10834	False	False	True	False	False	False	False	False	False	False	Fal
10835	False	False	False	False	False	False	False	False	False	False	Fal
10836	False	False	False	False	False	False	False	False	False	False	Fal
10837	False	False	False	False	False	False	False	False	False	False	Fal
10838	False	False	True	False	False	False	False	False	False	False	Fal

10839 rows × 13 columns

sum of missing value

In [25]:



```
p.isna().sum()
```

Out[25]:

```
App                  1
Category              0
Rating                1473
Reviews               0
Size                  0
Installs               0
Type                  0
APP Price              0
Content Rating          0
Genres                 0
Last Updated            0
Current Ver              8
Android Ver              2
dtype: int64
```

Check the duplicated values

In [26]:



```
p.duplicated()
```

Out[26]:

```
0      False
1      False
2      False
3      False
4      False
...
10834    False
10835    False
10836    False
10837    False
10838    False
Length: 10839, dtype: bool
```

Sum of duplicated value

In [27]:



```
p.duplicated().sum()
```

Out[27]:

```
483
```

number of unique values in dataset

In [28]:



```
p.nunique()
```

Out[28]:

```
App           9657
Category       33
Rating         39
Reviews        6001
Size          460
Installs       20
Type           2
APP Price      92
Content Rating  6
Genres         119
Last Updated   1377
Current Ver     2783
Android Ver     33
dtype: int64
```

Print unique genres in data set

In [29]:



```
p.Genres.unique()
```

Out[29]:

```
array(['Art & Design', 'Art & Design;Pretend Play',
       'Art & Design;Creativity', 'Art & Design;Action & Adventure',
       'Auto & Vehicles', 'Beauty', 'Books & Reference', 'Business',
       'Comics', 'Comics;Creativity', 'Communication', 'Dating',
       'Education;Education', 'Education', 'Education;Creativity',
       'Education;Music & Video', 'Education;Action & Adventure',
       'Education;Pretend Play', 'Education;Brain Games', 'Entertainmen
t',
       'Entertainment;Music & Video', 'Entertainment;Brain Games',
       'Entertainment;Creativity', 'Events', 'Finance', 'Food & Drink',
       'Health & Fitness', 'House & Home', 'Libraries & Demo',
       'Lifestyle', 'Lifestyle;Pretend Play',
       'Adventure;Action & Adventure', 'Arcade', 'Casual', 'Card',
       'Casual;Pretend Play', 'Action', 'Strategy', 'Puzzle', 'Sports',
       'Music', 'Word', 'Racing', 'Casual;Creativity',
       'Casual;Action & Adventure', 'Simulation', 'Adventure', 'Board',
       'Trivia', 'Role Playing', 'Action;Action & Adventure',
```

Print unique Category in data set

In [30]:



```
p.Category.unique()
```

Out[30]:

```
array(['ART_AND DESIGN', 'AUTO_AND VEHICLES', 'BEAUTY',
       'BOOKS_AND_REFERENCE', 'BUSINESS', 'COMICS', 'COMMUNICATION',
       'DATING', 'EDUCATION', 'ENTERTAINMENT', 'EVENTS', 'FINANCE',
       'FOOD_AND_DRINK', 'HEALTH_AND_FITNESS', 'HOUSE_AND_HOME',
       'LIBRARIES_AND_DEMO', 'LIFESTYLE', 'GAME', 'FAMILY', 'MEDICAL',
       'SOCIAL', 'SHOPPING', 'PHOTOGRAPHY', 'SPORTS', 'TRAVEL_AND_LOCAL',
       'TOOLS', 'PERSONALIZATION', 'PRODUCTIVITY', 'PARENTING', 'WEATHER',
       'VIDEO_PLAYERS', 'NEWS_AND_MAGAZINES', 'MAPS_AND_NAVIGATION'],
      dtype=object)
```

To count the Category names

In [31]:

N

```
p.Category.value_counts()
```

Out[31]:

FAMILY	1971
GAME	1144
TOOLS	843
MEDICAL	463
BUSINESS	460
PRODUCTIVITY	424
PERSONALIZATION	392
COMMUNICATION	387
SPORTS	384
LIFESTYLE	382
FINANCE	366
HEALTH_AND_FITNESS	341
PHOTOGRAPHY	335
SOCIAL	295
NEWS_AND_MAGAZINES	283
SHOPPING	260
TRAVEL_AND_LOCAL	258

In [32]:



```
p.Rating.value_counts()
```

Out[32]:

```
4.4    1109
4.3    1076
4.5    1038
4.2    952
4.6    823
4.1    708
4.0    568
4.7    499
3.9    386
3.8    303
5.0    274
3.7    239
4.8    234
3.6    174
3.5    163
3.4    128
3.3    102
4.9    87
3.0    83
3.1    69
3.2    64
2.9    45
2.8    42
2.7    25
2.6    25
2.5    21
2.3    20
2.4    19
1.0    16
2.2    14
1.9    13
2.0    12
2.1    8
1.8    8
1.7    8
1.6    4
1.5    3
1.4    3
1.2    1
Name: Rating, dtype: int64
```



Top 10 highest paid app

In [33]:



```
m_paid_apps=p.sort_values('APP Price',ascending=False).head(10)
```



In [34]:



```
m_paid_apps
```

Out[34]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating
10254	I'm Rich - Trump Edition	LIFESTYLE	3.6	275	7.3	10000	Paid	400.00	Everyone
10253	ðŸ'Ž I'm rich	LIFESTYLE	3.8	718	26	10000	Paid	399.99	Everyone
10393	I am Rich	FINANCE	4.3	180	3.8	5000	Paid	399.99	Everyone
10389	I Am Rich Pro	FAMILY	4.4	201	2.7	5000	Paid	399.99	Everyone
10386	I am rich(premium)	FINANCE	3.5	472	965k	5000	Paid	399.99	Everyone
10385	I am Rich!	FINANCE	3.8	93	22	1000	Paid	399.99	Everyone
10390	I am rich (Most expensive app)	FINANCE	4.1	129	2.7	1000	Paid	399.99	Teen
10381	I am Rich Plus	FAMILY	4.0	856	8.7	10000	Paid	399.99	Everyone
10380	I am rich	LIFESTYLE	3.8	3547	1.8	100000	Paid	399.99	Everyone
10383	I Am Rich Premium	FINANCE	4.1	1867	4.7	50000	Paid	399.99	Everyone

To check free apps

In [35]:



```
p.freeapps=p[p['Type']=='Free']
```

In [36]:



```
p.freeapps
```

Out[36]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND DESIGN	4.1	159	19	10000	Free	0.0	Everyone
1	Coloring book moana	ART_AND DESIGN	3.9	967	14	500000	Free	0.0	Everyone
2	U Launcher Lite	ART_AND DESIGN	4.7	87510	8.7	5000000	Free	0.0	Everyone

In [37]:



```
p.paidapps=p[p['Type']=='Paid']
```

In [38]:



```
p.paidapps
```

Out[38]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating
10039	TurboScan: scan documents and receipts in PDF	BUSINESS	4.7	11442	6.8	100000	Paid	4.99	Everyone
10040	Tiny Scanner Pro: PDF Doc Scan	BUSINESS	4.8	10295	39	100000	Paid	4.99	Everyone
10041	TurboScan: scan documents and	BUSINESS	4.7	11442	6.8	100000	Paid	4.99	Everyone

To find most rated apps

In [39]:



```
m_rated_apps=p.sort_values('Rating', ascending=False).head(5)
```

In [40]:

```
m_rated_apps
```

Out[40]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating	Genre
7795	D-H Pharmacy	LIFESTYLE	5.0	6	2.3	1000	Free	0.00	Everyone	Lifestyle
9826	Trinity Church Deltona, FL	LIFESTYLE	5.0	33	28	500	Free	0.00	Everyone	Lifestyle
10618	Morse Player	FAMILY	5.0	12	2.4	100	Paid	1.99	Everyone	Education
8788	EK Bailey Preaching Conference	EVENTS	5.0	3	30	500	Free	0.00	Everyone	Events
9873	FN	BUSINESS	5.0	14	3.3	50	Free	0.00	Everyone	Business

To find the most under rated apps

In [41]:

```
u_rated_apps=p.sort_values('Rating').head(5)
```

In [42]:



```
u_rated_apps
```

Out[42]:

		App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating	Genres
9809	Lottery Ticket Checker - Florida Results & Lotto	TOOLS	1.0	3	41	500	Free	0.00	Everyone	Tools	
10449	Truck Driving Test Class 3 BC	FAMILY	1.0	1	2	50	Paid	1.49	Everyone	Education	
8163	DS Creator 2.0	TOOLS	1.0	2	4.4	500	Free	0.00	Everyone	Tools	
4825	Clarksburg AH	MEDICAL	1.0	1	28	50	Free	0.00	Everyone	Medical	
10219	Speech Therapy: F	FAMILY	1.0	1	16	10	Paid	2.99	Everyone	Education	

top 10 rated free apps

In [43]:



```
p.topratedfreeapps=p.freeapps.sort_values('Rating',ascending=False).head(10)
```

In [44]:



```
p.topratedfreeapps
```

Out[44]:

		App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating	
4930	AJ Rafael Music Lessons	FAMILY	5.0	1	8.2	10	Free	0.0	Everyone	Enter	
628	Online Girls Chat Group	DATING	5.0	5	5	100	Free	0.0	Mature 17+		
7514	Cy-Fair VFD EMS Protocols	MEDICAL	5.0	2	20	100	Free	0.0	Everyone 10+		
7516	Cy-Fair Christian Church	LIFESTYLE	5.0	2	9.3	100	Free	0.0	Everyone		
7553	CZ-Help	BOOKS AND REFERENCE	5.0	2	1.4	5	Free	0.0	Everyone		

Check the co relationship between data

In [45]:

```
p.corr()
```

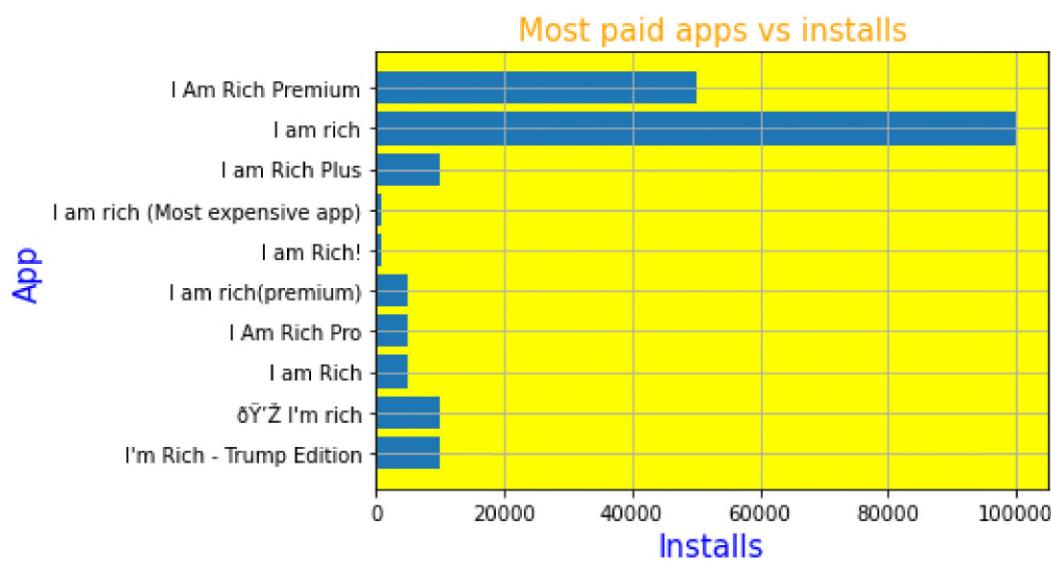
Out[45]:

	Rating	Reviews	Installs	APP Price
Rating	1.000000	0.068141	0.051355	-0.021903
Reviews	0.068141	1.000000	0.643121	-0.009668
Installs	0.051355	0.643121	1.000000	-0.011691
APP Price	-0.021903	-0.009668	-0.011691	1.000000

Data visualization

In [46]:

```
plt.barh(m_paid_apps['App'],m_paid_apps['Installs'])
plt.title('Most paid apps vs installs',fontsize=15,color='orange')
ax=plt.axes()
ax.set_facecolor('Yellow')
plt.xlabel('Installs',fontsize=15,color='blue')
plt.ylabel('App',fontsize=15,color='blue')
plt.grid(True)
plt.show()
```

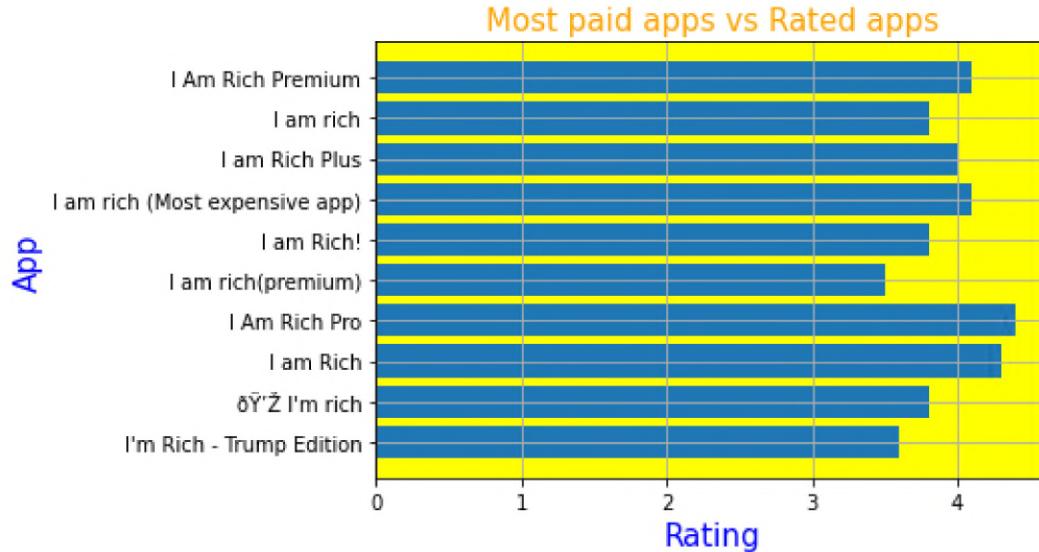


Most paid apps vs Rated Apps

In [47]:

▶

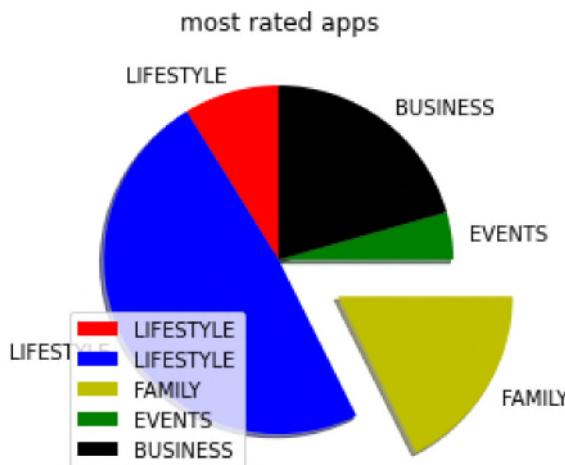
```
plt.barh(m_paid_apps['App'],m_paid_apps['Rating'])
plt.title('Most paid apps vs Rated apps',fontsize=15,color='orange')
ax=plt.axes()
ax.set_facecolor('Yellow')
plt.xlabel('Rating',fontsize=15,color='blue')
plt.ylabel('App',fontsize=15,color='blue')
plt.grid(True)
plt.show()
```



Reviews and category of most rated apps

In [48]:

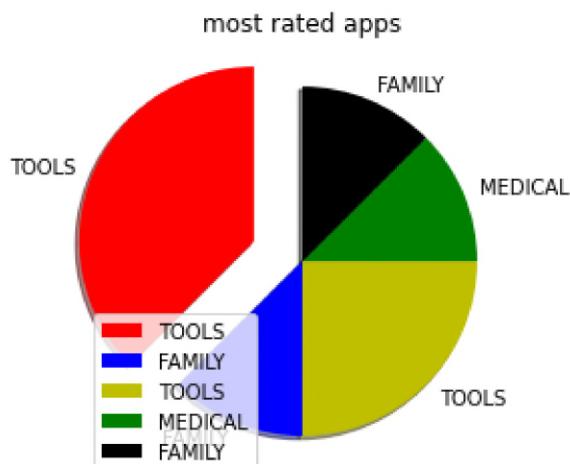
```
y=m_rated_apps.Reviews  
l=m_rated_apps.Category  
c=['r','b','y','g','k']  
  
plt.pie(y,labels=l,startangle=90,shadow=True,colors=c,explode=(0,0,0.4,0,0))  
plt.legend(loc='lower left')  
plt.title('most rated apps')  
plt.show()
```



Review the under rated apps

In [49]:

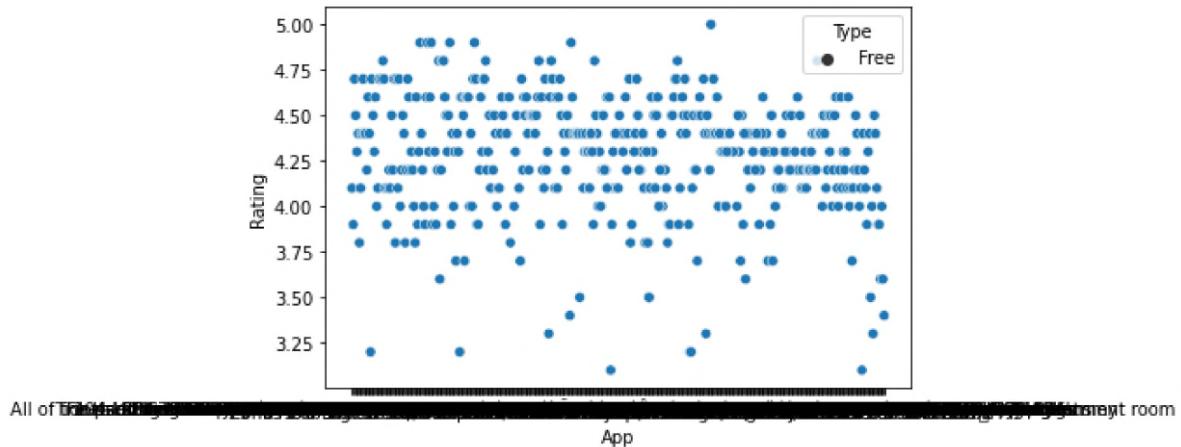
```
y=u_rated_apps.Reviews  
l=u_rated_apps.Category  
c=['r','b','y','g','k']  
  
plt.pie(y,labels=l,startangle=90,shadow=True,colors=c,explode=(0.3,0,0,0,0))  
plt.legend(loc='lower left')  
plt.title('most rated apps')  
plt.show()
```



Top 100 apps in dataset

In [50]:

```
sns.scatterplot(x='App',y='Rating',style='Type',data=p1)
plt.figure(figsize=(16,80))
plt.show()
```

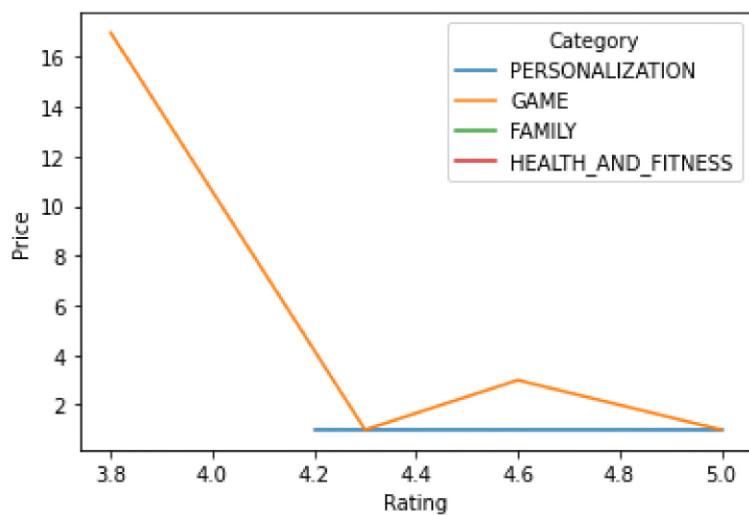


<Figure size 1152x5760 with 0 Axes>

Rating vs App Price

In [52]:

```
sns.lineplot(x='Rating',y='Price',hue='Category',data=p2)
plt.show()
```



In [51]:

```
m_rated_cat=p[p['Category']=='BUSINESS']
m_rated_cat
```

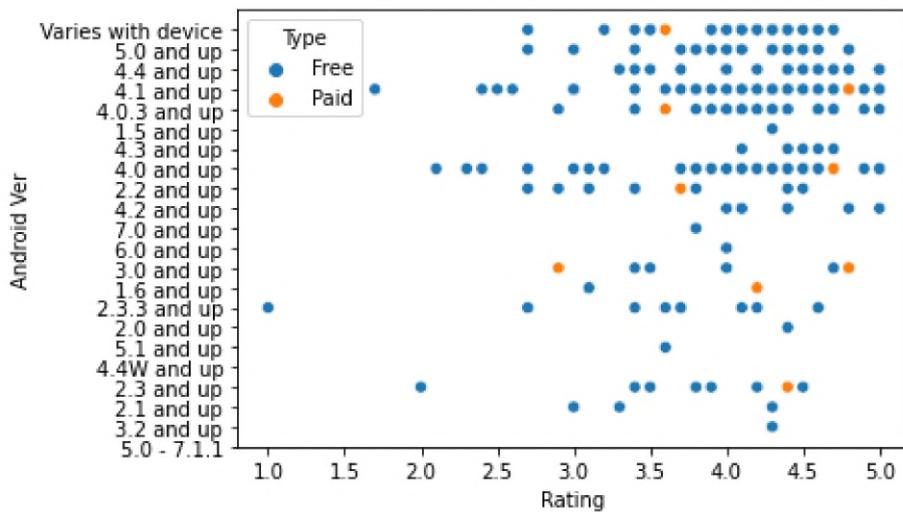
Out[51]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating	Genres	Last Updated
187	Visual Voicemail by MetroPCS	BUSINESS	4.1	16129	0	10000000	Free	0.00	Everyone	Business	July 30, 2018
188	Indeed Job Search	BUSINESS	4.3	674730	0	50000000	Free	0.00	Everyone	Business	May 21, 2018
189	Uber Driver	BUSINESS	4.4	1254730	0	10000000	Free	0.00	Everyone	Business	August 3, 2018
190	ADP Mobile Solutions	BUSINESS	4.3	85185	29	5000000	Free	0.00	Everyone	Business	July 17, 2018

Scatter plot graph Rating vs android version for BUISNESS category

In [53]:

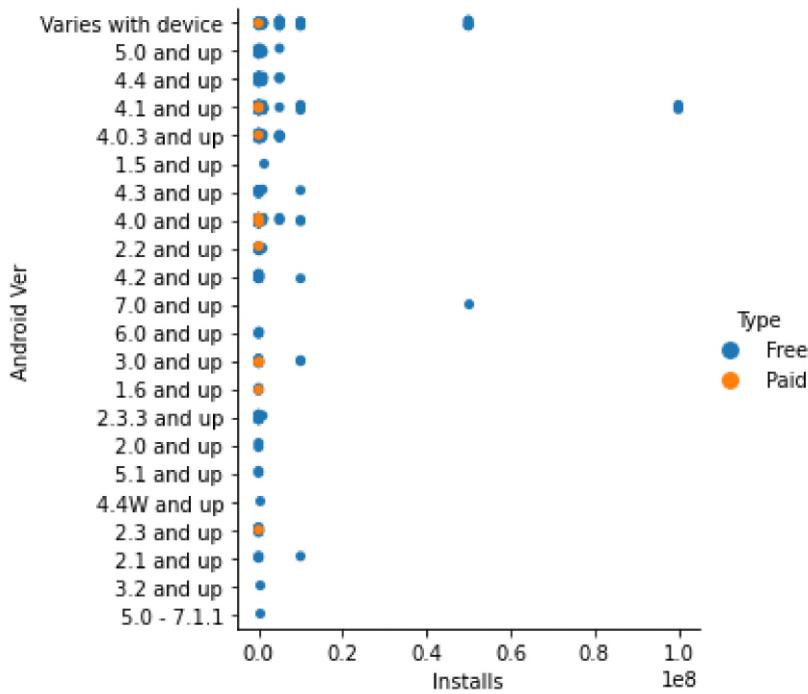
```
sns.scatterplot(x='Rating',y='Android Ver',hue='Type',data=m_rated_cat)
plt.show()
```



cat plot graph Installs vs android version for BUISNESS category

In [54]:

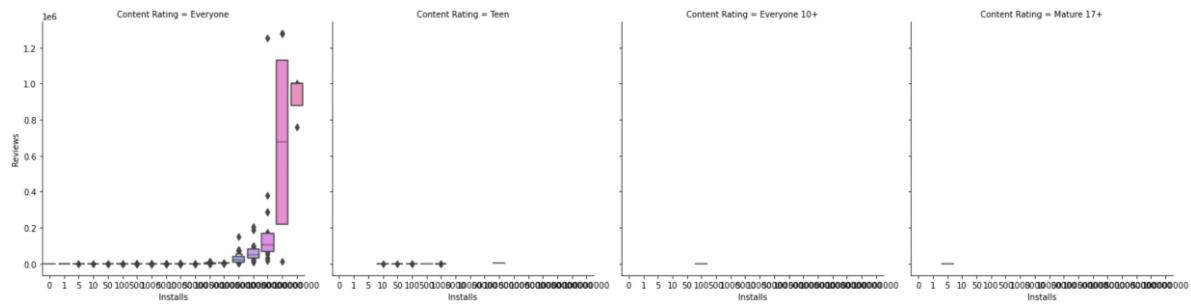
```
sns.catplot(x='Installs',y='Android Ver',hue='Type',data=m_rated_cat)
plt.show()
```



cat plot graph Installs vs Reviews for BUISNESS category

In [55]:

```
sns.catplot(x='Installs',y='Reviews',col='Content Rating',kind='boxen',data=m_rated_cat)
plt.figure(figsize=(16,180))
plt.show()
```

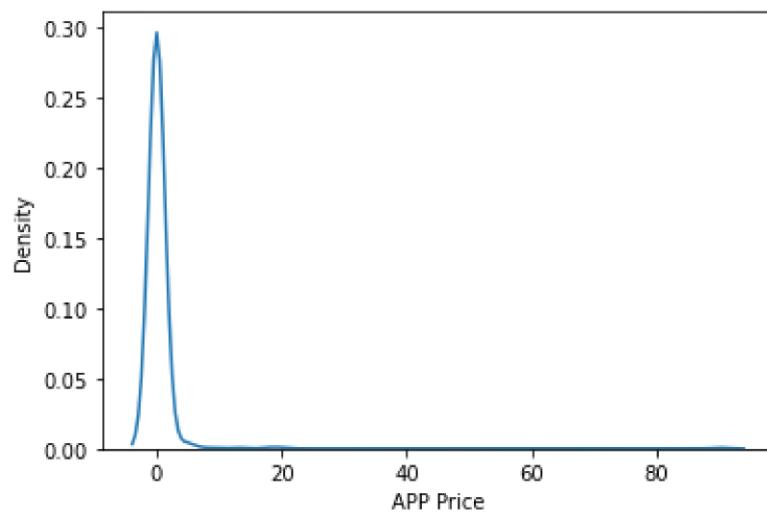


<Figure size 1152x12960 with 0 Axes>

kdeplot graph App price for BUISNESS category

In [56]:

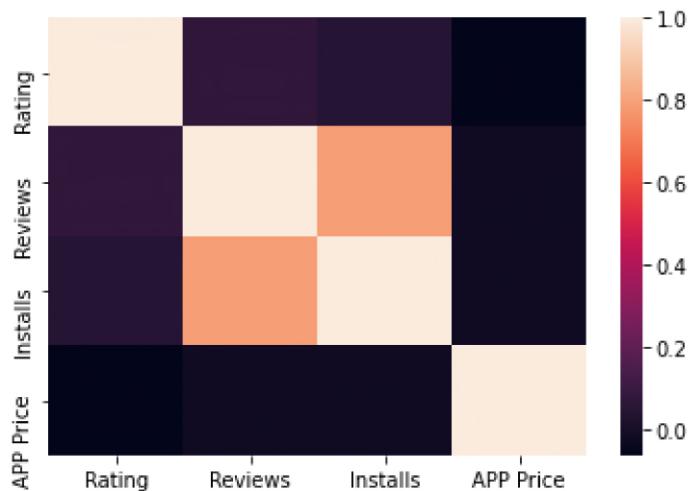
```
sns.kdeplot(m_rated_cat['APP Price'])
plt.show()
```



Heat graph for BUISNESS category

In [57]:

```
sns.heatmap(m_rated_cat.corr())
plt.show()
```



In [59]:

```
sns.heatmap(m_rated_cat.corr(), cmap='RdBu', vmin=-1, vmax=1)
plt.show()
```



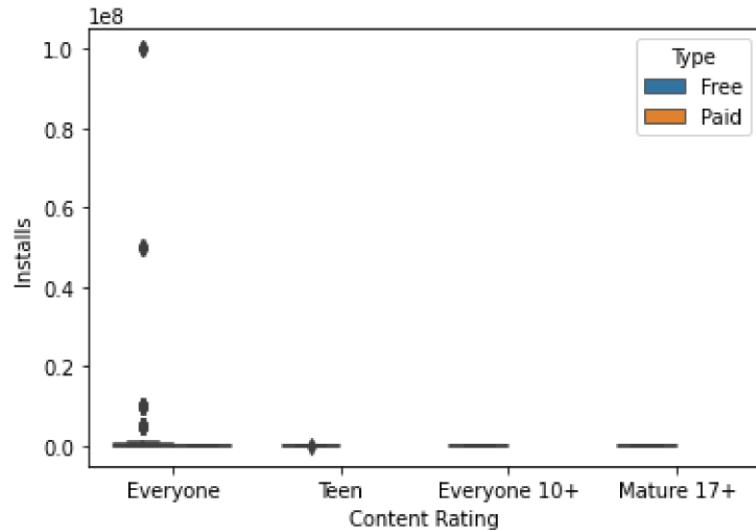
In [66]:

```
sns.heatmap(m_rated_cat.corr(), cmap='RdBu', annot=True)
plt.show()
```



In [62]:

```
sns.boxplot(x='Content Rating',y='Installs',hue='Type',width=0.7,data=m_rated_cat)  
plt.show()
```



In [73]:

```
p.entertainment=p[p['Genres']=='Entertainment']
```

In [78]:

```
p1.entertainment=p.entertainment.sort_values('Installs',ascending=False).head(20)
```

In [79]:

```
p1.entertainment
```

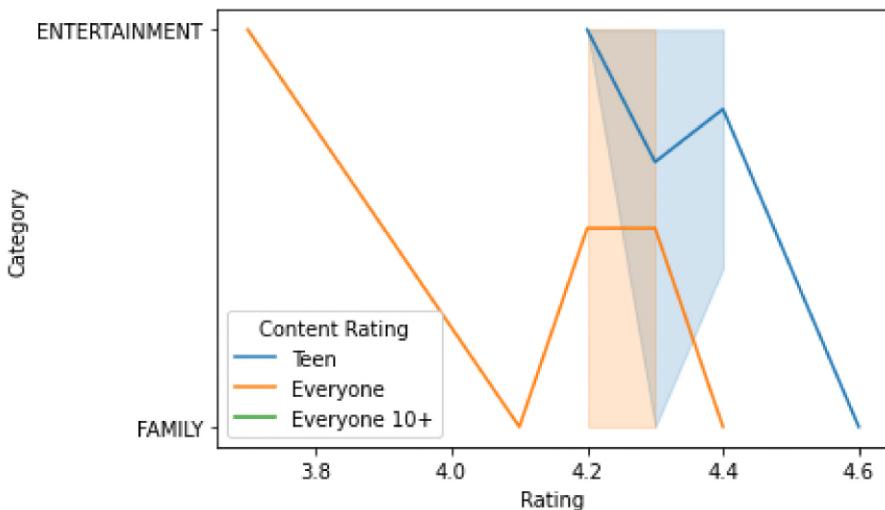
Out[79]:

	App	Category	Rating	Reviews	Size	Installs	Type	APP Price	Content Rating	Genres
849	Google Play Games	ENTERTAINMENT	4.3	7165362	0	10000000000	Free	0.0	Teen	Entertainmen
5462	Google Play Games	FAMILY	4.3	7168735	0	10000000000	Free	0.0	Teen	Entertainmen
839	Netflix	ENTERTAINMENT	4.4	5456208	0	100000000	Free	0.0	Teen	Entertainmen
858	Talking Angela	ENTERTAINMENT	3.7	1828284	52	1000000000	Free	0.0	Everyone	Entertainmen
615	Allo	ENTERTAINMENT	4.1	5456208	0	1000000000	Free	0.0	Teen	Entertainmen

Rating vs category top 20 most entertaing apps

In [80]:

```
sns.lineplot(x='Rating',y='Category',hue='Content Rating',data=p1.entertainment)
plt.show()
```



In [4]:

```
p3=p.head(300)
```

In [5]:



p3

Out[5]:

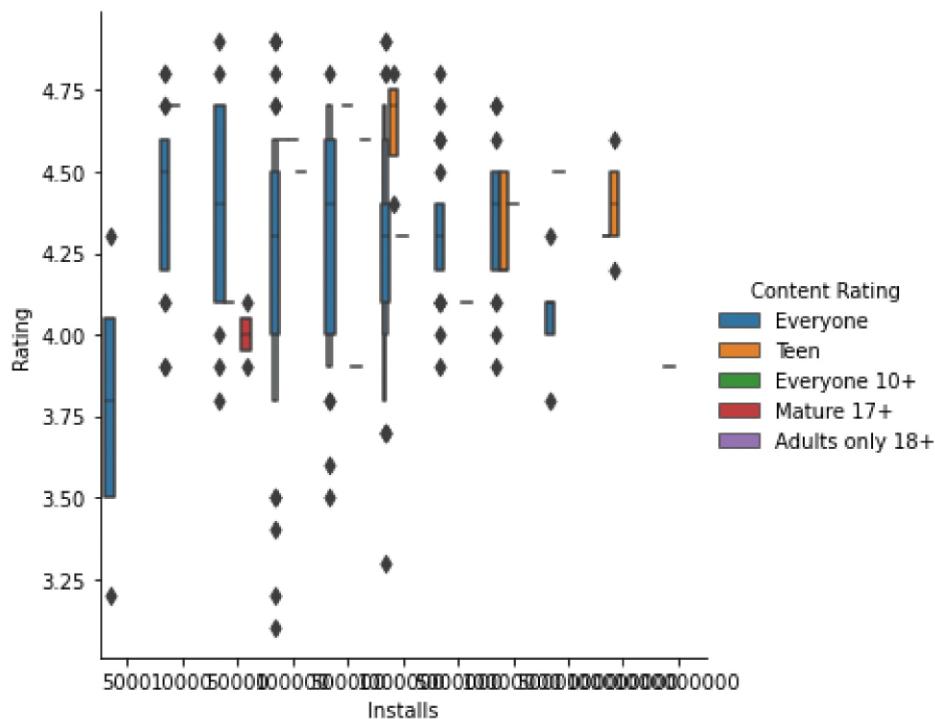
	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genl
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND DESIGN	4.1	159	19	10000	Free	0.0	Everyone	Art & Des
1	Coloring book moana	ART_AND DESIGN	3.9	967	14	500000	Free	0.0	Everyone	Art Design;Prete P
2	U Launcher Lite â€“ FREE Live Cool Themes,	ART_AND DESIGN	4.7	87510	8.7	5000000	Free	0.0	Everyone	Art & Des

Rating vs installs of 1st 300 apps

In [89]:



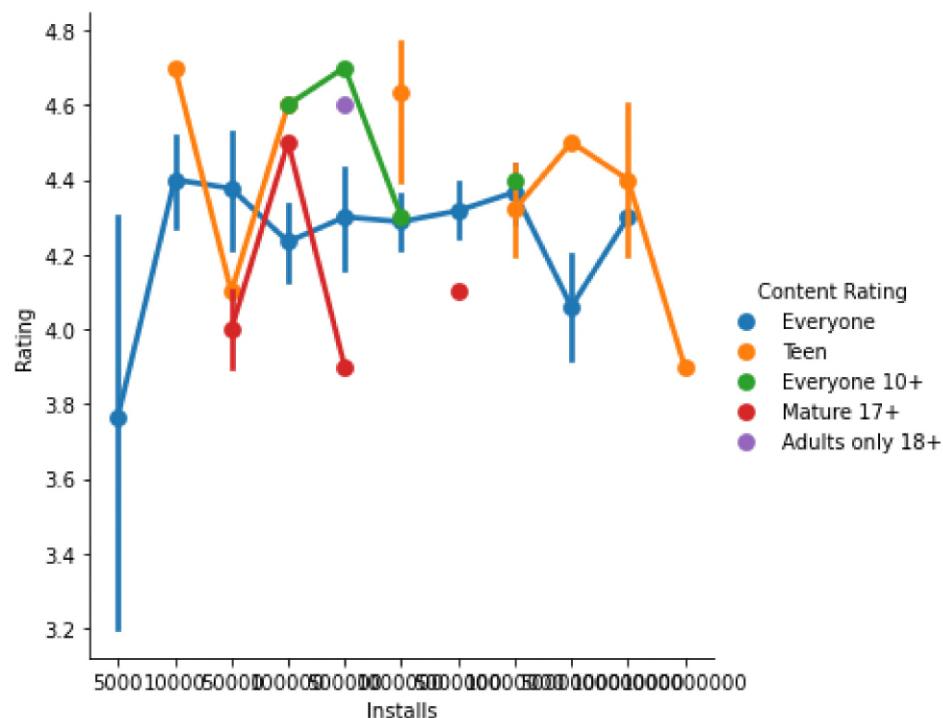
```
sns.catplot(x='Installs',y='Rating',hue='Content Rating',kind='boxen',data=p3)  
plt.show()
```



<Figure size 11520x7200 with 0 Axes>

In [93]:

```
sns.catplot(x='Installs',y='Rating',hue='Content Rating',kind='point',data=p3)
plt.show()
```

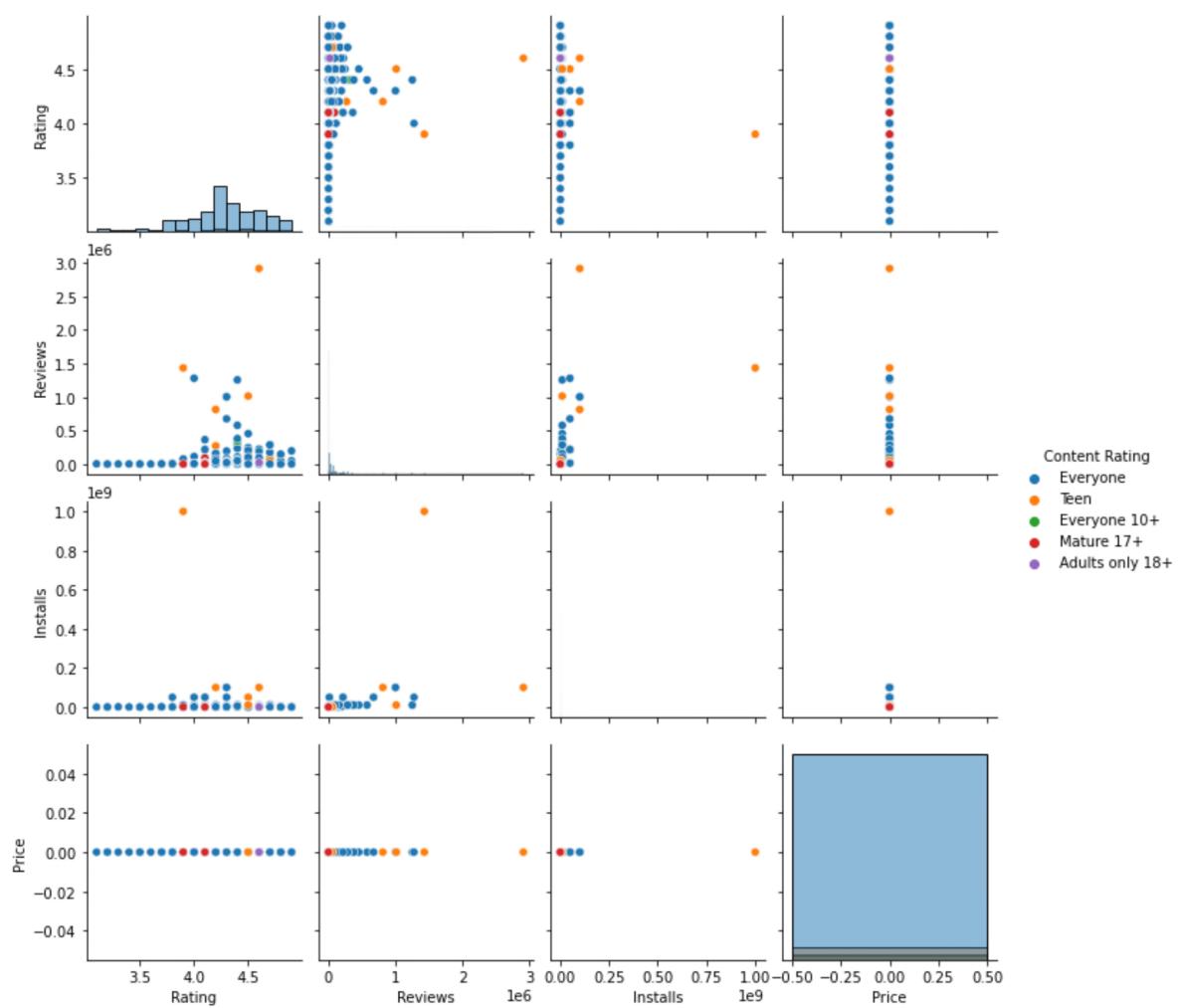


Using pair plot p3

In [6]:

▶

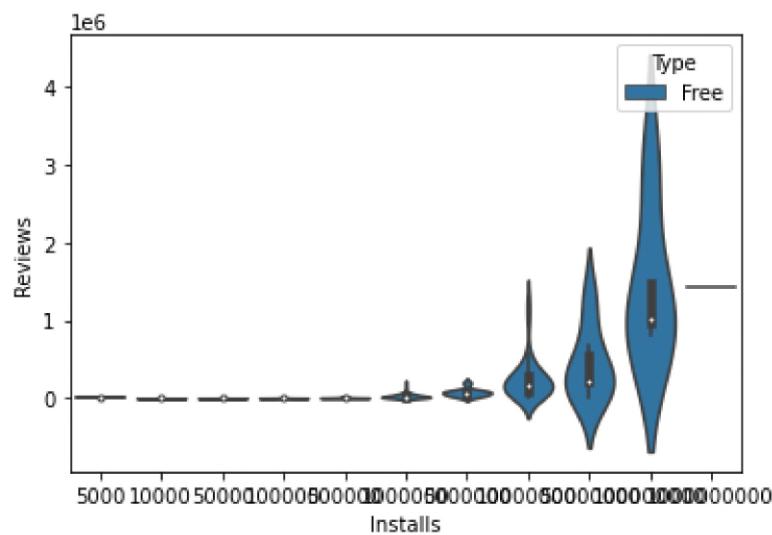
```
sns.pairplot(p3,diag_kind='hist',hue='Content Rating')
plt.show()
```



Using violin plot Install vs Review

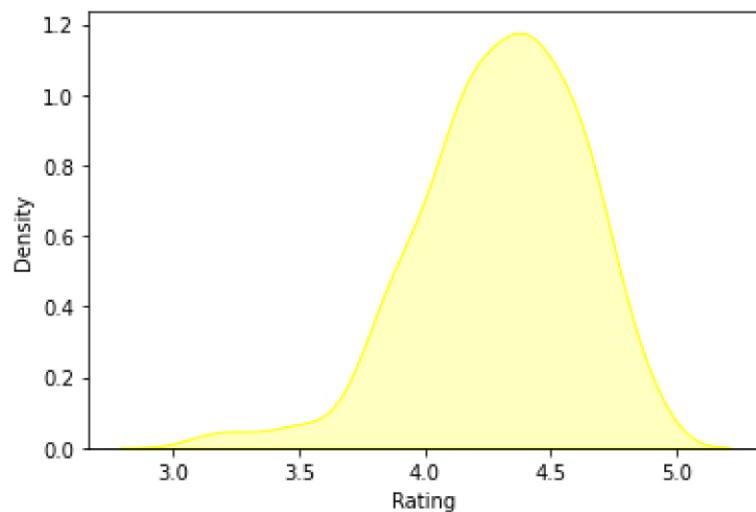
In [10]:

```
sns.violinplot(x='Installs',y='Reviews',hue='Type',data=p3)  
plt.show()
```



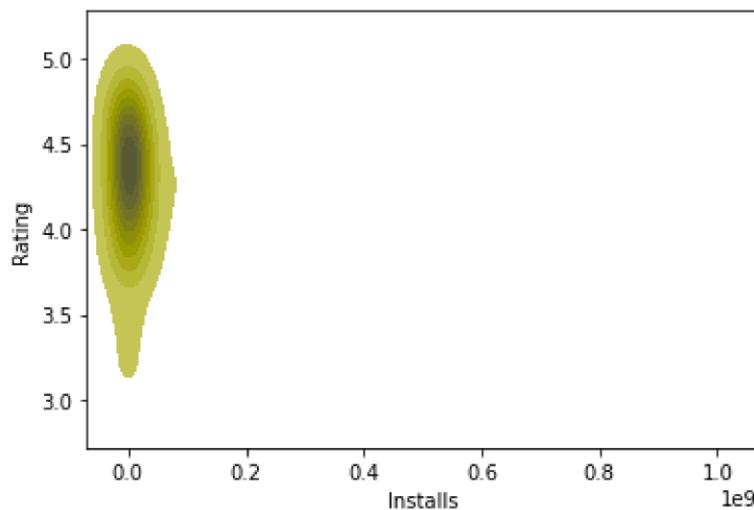
In [13]:

```
sns.kdeplot(p3['Rating'],color='yellow',shade=True)  
plt.show()
```



In [18]:

```
sns.kdeplot(p3['Installs'],p3['Rating'],color='yellow',shade=True)  
plt.show()
```



In []:



In []:



In []:



In []:



In []:



In []:



In []:



In []:



In []:

