

googleplaystore-analysis

November 16, 2024

```
[3]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[5]: g_data=pd.read_csv('E:\Sql query\googleplaystore.csv')
g_data
```

```
[5]:
```

	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]

[]:

1- Display top 10 rows

[16]: `g_data.head(10)`

```
[16]:
```

	App	Category	Rating	\
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	
1	Coloring book moana	ART_AND_DESIGN	3.9	
2	U Launcher Lite - FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	
3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	
5	Paper flowers instructions	ART_AND_DESIGN	4.4	
6	Smoke Effect Photo Maker - Smoke Editor	ART_AND_DESIGN	3.8	
7	Infinite Painter	ART_AND_DESIGN	4.1	
8	Garden Coloring Book	ART_AND_DESIGN	4.4	
9	Kids Paint Free - Drawing Fun	ART_AND_DESIGN	4.7	

	Reviews	Size	Installs	Type	Price	Content	Rating	\
0	159	19M	10,000+	Free	0	Everyone		
1	967	14M	500,000+	Free	0	Everyone		
2	87510	8.7M	5,000,000+	Free	0	Everyone		
3	215644	25M	50,000,000+	Free	0	Teen		

4	967	2.8M	100,000+	Free	0	Everyone
5	167	5.6M	50,000+	Free	0	Everyone
6	178	19M	50,000+	Free	0	Everyone
7	36815	29M	1,000,000+	Free	0	Everyone
8	13791	33M	1,000,000+	Free	0	Everyone
9	121	3.1M	10,000+	Free	0	Everyone

	Genres	Last Updated	Current Ver \
0	Art & Design	January 7, 2018	1.0.0
1	Art & Design;Pretend Play	January 15, 2018	2.0.0
2	Art & Design	August 1, 2018	1.2.4
3	Art & Design	June 8, 2018	Varies with device
4	Art & Design;Creativity	June 20, 2018	1.1
5	Art & Design	March 26, 2017	1.0
6	Art & Design	April 26, 2018	1.1
7	Art & Design	June 14, 2018	6.1.61.1
8	Art & Design	September 20, 2017	2.9.2
9	Art & Design;Creativity	July 3, 2018	2.8

	Android Ver
0	4.0.3 and up
1	4.0.3 and up
2	4.0.3 and up
3	4.2 and up
4	4.4 and up
5	2.3 and up
6	4.0.3 and up
7	4.2 and up
8	3.0 and up
9	4.0.3 and up

[]:

2- Display least 10 rows

[19]: `g_data.tail(10)`

[19]:

	App	Category \
10831	payermonstationnement.fr	MAPS_AND_NAVIGATION
10832	FR Tides	WEATHER
10833	Chemin (fr)	BOOKS_AND_REFERENCE
10834	FR Calculator	FAMILY
10835	FR Forms	BUSINESS
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	\
10831	NaN	38	9.8M	5,000+	Free	0	
10832	3.8	1195	582k	100,000+	Free	0	
10833	4.8	44	619k	1,000+	Free	0	
10834	4.0	7	2.6M	500+	Free	0	
10835	NaN	0	9.6M	10+	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	\
10831	Everyone	Maps & Navigation	June 13, 2018	
10832	Everyone	Weather	February 16, 2014	
10833	Everyone	Books & Reference	March 23, 2014	
10834	Everyone	Education	June 18, 2017	
10835	Everyone	Business	September 29, 2016	
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
10831	2.0.148.0	4.0 and up
10832	6.0	2.1 and up
10833	0.8	2.2 and up
10834	1.0.0	4.1 and up
10835	1.1.5	4.0 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

[]:

3- Find the shape of dataset also no.of rows and columns

[22]: g_data.shape

[22]: (10841, 13)

```
[26]: print("No of Rows :", g_data.shape[0])
      print("No of columns :", g_data.shape[1])
```

```
No of Rows : 10841
No of columns : 13
```

```
[ ]:
```

4- Get information of dataset

```
[29]: g_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
```

```
[ ]:
```

5- Get overall Statistics

```
[34]: g_data.describe()  # data set has only one number column .
```

```
[34]:
```

	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

```
[36]: g_data.describe(include='all')
```

```
[36]:
```

	App	Category	Rating	Reviews		Size	Installs \
count	10841	10841	9367.000000	10841		10841	10841
unique	9660	34	NaN	6002		462	22
top	ROBLOX	FAMILY	NaN	0	Varies with device	1,000,000+	
freq	9	1972	NaN	596		1695	1579
mean	NaN	NaN	4.193338	NaN		NaN	NaN
std	NaN	NaN	0.537431	NaN		NaN	NaN
min	NaN	NaN	1.000000	NaN		NaN	NaN
25%	NaN	NaN	4.000000	NaN		NaN	NaN
50%	NaN	NaN	4.300000	NaN		NaN	NaN
75%	NaN	NaN	4.500000	NaN		NaN	NaN
max	NaN	NaN	19.000000	NaN		NaN	NaN

	Type	Price	Content	Rating	Genres	Last Updated \
count	10840	10841		10840	10841	10841
unique	3	93		6	120	1378
top	Free	0	Everyone	Tools		August 3, 2018
freq	10039	10040		8714	842	326
mean	NaN	NaN		NaN	NaN	NaN
std	NaN	NaN		NaN	NaN	NaN
min	NaN	NaN		NaN	NaN	NaN
25%	NaN	NaN		NaN	NaN	NaN
50%	NaN	NaN		NaN	NaN	NaN
75%	NaN	NaN		NaN	NaN	NaN
max	NaN	NaN		NaN	NaN	NaN

	Current Ver	Android Ver
count	10833	10838
unique	2832	33
top	Varies with device	4.1 and up
freq	1459	2451
mean	NaN	NaN
std	NaN	NaN
min	NaN	NaN
25%	NaN	NaN
50%	NaN	NaN
75%	NaN	NaN
max	NaN	NaN

```
[ ]:
```

6- total no of App titles contains Astrology

```
[39]: g_data.columns
```

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

```
[41]: g_data['App'].str.contains('Astrology', case=False)
```

```
[41]: 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
```

```
[43]: g_data[g_data['App'].str.contains('Astrology', case=False)]
```

```
[43]:
```

	App	Category	Rating	\
1570	Horoscopes - Daily Zodiac Horoscope and Astrology	LIFESTYLE	4.6	
1592	Astrology - Min Thein Kha BayDin	LIFESTYLE	4.7	
10840	iHoroscope - 2018 Daily Horoscope & Astrology	LIFESTYLE	4.5	

	Reviews	Size	Installs	Type	Price	Content Rating	Genres	\
1570	161143	11M	10,000,000+	Free	0	Everyone 10+	Lifestyle	
1592	2225	15M	100,000+	Free	0	Everyone	Lifestyle	
10840	398307	19M	10,000,000+	Free	0	Everyone	Lifestyle	

	Last Updated	Current Ver	Android Ver
1570	June 25, 2018	5.2.4(881)	4.0.3 and up
1592	July 26, 2018	4.2.1	4.0.3 and up
10840	July 25, 2018	Varies with device	Varies with device

```
[49]: len(g_data[g_data['App'].str.contains('Astrology', case=False)])
```

```
[49]: 3
```

```
[ ]:
```

7- Find avg App rating

```
[51]: g_data.columns
```

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

```
[57]: g_data['Rating'].mean()
```

```
[57]: 4.193338315362443
```

```
[ ]:
```

8- Find total no of unique Category

```
[61]: g_data['Category'].nunique()
```

```
[61]: 34
```

```
[ ]:
```

9- Which category getting the highest avg rating

```
[69]: g_data.groupby('Category')['Rating'].mean().sort_values(ascending=False)
```

```
[69]: 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

```

```
[ ]:
```

10- Find total no of Apps Having 5 Star Rating

```
[74]: g_data.columns
```

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

```
[78]: len(g_data[g_data['Rating']==5.0])
```

```
[78]: 274
```

```
[ ]:
```

11- Find Avg of Reviews

```
[92]: g_data.columns
```

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

```
[94]: g_data[g_data['Reviews']=='3.0M']
```

```
[94]:
```

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

```
[96]: g_data['Reviews']=g_data['Reviews'].replace('3.0M',3.0)
```

```
[100]: g_data['Reviews']=g_data['Reviews'].astype('float')
```

```
[102]: g_data['Reviews'].mean()
```

```
[102]: 444111.9265750392
```

```
[ ]:
```

12- Find total no of free and paid Apps.

```
[105]: g_data.columns
```

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

```
[113]: g_data['Type'].value_counts()
```

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

```
[ ]:
```

13- find maximum Reviews

```
[133]: g_data['Reviews'].max()
```

```
[133]: 78158306.0
```

```
[ ]:
```

14- Display top 5 apps having highest reviews.

```
[138]: index=g_data['Reviews'].sort_values(ascending=False).head().index  
g_data.iloc[index]['App']
```

```
[138]: 2544          Facebook  
3943          Facebook  
381    WhatsApp Messenger  
336    WhatsApp Messenger  
3904    WhatsApp Messenger  
Name: App, dtype: object
```

```
[ ]:
```

15- Find avg Rating of free and paid App.

```
[144]: g_data.groupby('Type')['Rating'].mean()
```

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

```
[ ]:
```

16- Display top app having max install

```
[148]: g_data.columns
```

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

```
[ ]: g_data['Installs'].str.replace(',', '')
g_data['Installs'].str.replace('+', '')
g_data['Installs'].str.replace('free', '0')
g_data['Installs'].astype('int')
```

```
[164]: g_data['Installs']=g_data['Installs'].str.replace(',', '')
```

```
[166]: g_data['Installs']=g_data['Installs'].str.replace('+', '')
```

```
[174]: g_data['Installs']=g_data['Installs'].str.replace('Free', '0')
```

```
[176]: g_data['Installs'].astype('int')
```

```
[176]: 0      10000
1     500000
2     5000000
3     50000000
4     100000
...
10836      5000
10837       100
10838      1000
10839      1000
10840    10000000
Name: Installs, Length: 10841, dtype: int32
```

```
[178]: index1=g_data['Installs'].sort_values(ascending=False).head().index
```

```
[180]: g_data.iloc[index]['App']
```

```
[180]: 2544          Facebook
      3943          Facebook
      381  WhatsApp Messenger
      336  WhatsApp Messenger
      3904  WhatsApp Messenger
      Name: App, dtype: object
```

```
[ ]:
```