

Screenshots of App Rating Project

Output - 1

df

	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 m oar a	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
...
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 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

10841 rows x 13 columns

Output - 2

```
df.isna().sum()
App                0
Category           0
Rating            1474
Reviews            0
Size               0
Installs           0
Type               1
Price              0
Content Rating     1
Genres             0
Last Updated       0
Current Ver        8
Android Ver        3
dt type: int64

df.dropna(inplace=True)

df.isna().sum()
App                0
Category           0
Rating             0
Reviews            0
Size               0
Installs           0
Type               0
Price              0
Content Rating     0
Genres             0
Last Updated       0
Current Ver        0
Android Ver        0
..                .
```

Output - 3

	Genres	Last Updated	Current Ver	Android Ver	size
	Art & Design	January 7, 2018	1.0.0	4.0.3 and up	19000.0
	Art & Design; Pretend Play	January 15, 2018	2.0.0	4.0.3 and up	14000.0
	Art & Design	August 1, 2018	1.2.4	4.0.3 and up	8000.0
	Art & Design	June 8, 2018	Varies with device	4.2 and up	25000.0
	Art & Design; Creativity	June 20, 2018	1.1	4.4 and up	2000.0

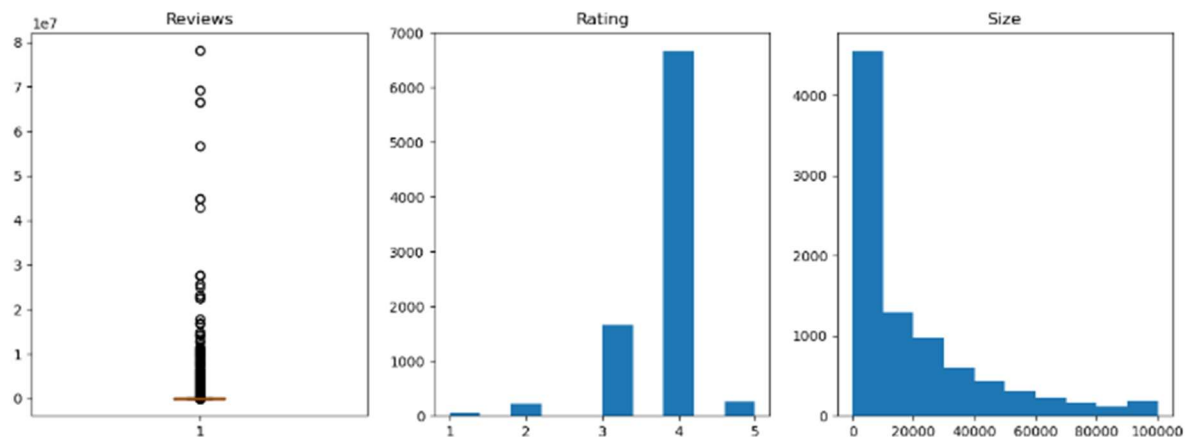
	Education	June 18, 2017	1.0.0	4.1 and up	2000.0
	Education	July 25, 2017	1.48	4.1 and up	53000.0
	Education	July 6, 2018	1.0	4.1 and up	3000.0
	Books & Reference	January 19, 2015	Varies with device	Varies with device	0.0
	Lifestyle	July 25, 2018	Varies with device	Varies with device	19000.0

Output - 4

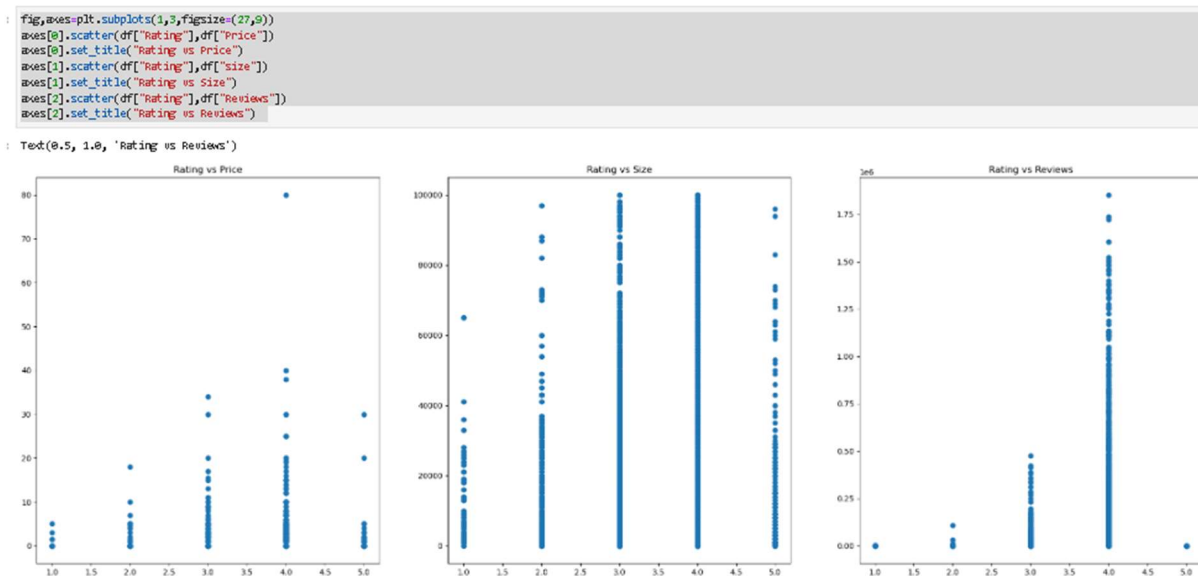
	App	Category	Rating	Reviews	Installs
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	10000
1	Coloring book mcana	ART_AND_DESIGN	3.9	967	500000
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	5000000
3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	50000000
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967	100000


```
df["Price"] = df["Price"].astype(float)
df.info()
<class 'pandas.core.frame.DataFrame'>
Int64Index: 8886 entries, 0 to 10840
Data columns (total 13 columns):
#   Column              Non-Null Count  Dtype
---  -
0   App                  8886 non-null   object
1   Category              8886 non-null   object
2   Rating                8886 non-null   int32
3   Reviews               8886 non-null   int32
4   Installs              8886 non-null   int32
5   Type                  8886 non-null   object
6   Price                 8886 non-null   float64
7   Content Rating        8886 non-null   object
8   Genres                8886 non-null   object
9   Last Updated          8886 non-null   object
10  Current Ver           8886 non-null   object
11  Android Ver           8886 non-null   object
12  size                  8886 non-null   float64
dtypes: float64(2), int32(3), object(8)
memory usage: 867.8+ KB
```

Output - 5



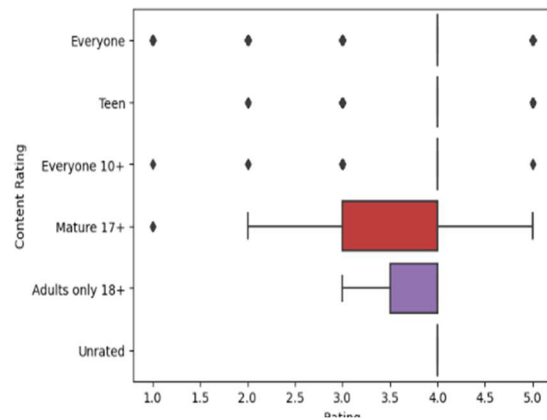
Output - 6



Output - 7

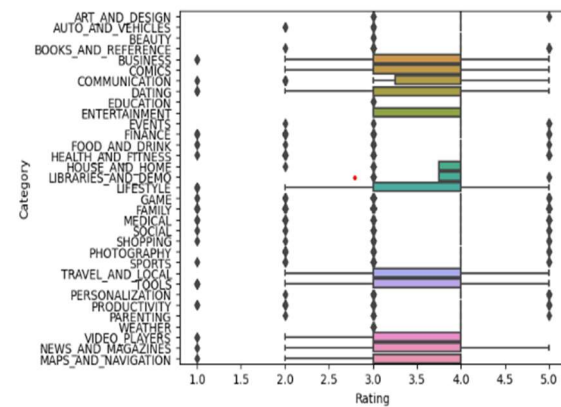
```
import seaborn as sns
sns.boxplot(x="Rating", y="Content_Rating", data=df)
```

```
<seaborn.axis.FacetGrid>
```



```
sns.boxplot(x="Rating", y="Category", data=df)
```

```
<seaborn.axis.FacetGrid>
```



Output - 8

```
inp1["Reviews"] = np.log(inp1["Reviews"])
```

```
inp1["Installs"] = np.log(inp1["Installs"])
```

```
inp1.head()
```

	App	Category	Rating	Reviews	Installs
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4	5.068904	9.210340
1	Coloring book mcana	ART_AND_DESIGN	3	6.874198	13.122363
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4	11.379508	15.424948
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4	6.874198	11.512925
5	Paper flowers instructions	ART_AND_DESIGN	4	5.117994	10.819778

Output - 9

```
inp2 = pd.get_dummies(inp1, columns=["Category", "Content_Rating", "Genres", "Type"])
```

```
inp2.head()
```

	Reviews	Installs	Price	size	Category_ART_AND_DESIGN	Category_AUTO_AND_VEHICLES	Category_BEAUTY	Category_BOOKS_AND_REFERENCE	Category_BUSINESS	...	Genres_Travel & local
4	5.068904	9.210340	0.0	19000.0	1	0	0	0	0	...	0
3	6.874198	13.122363	0.0	14000.0	1	0	0	0	0	...	0
4	11.379508	15.424948	0.0	8000.0	1	0	0	0	0	...	0
4	6.874198	11.512925	0.0	2000.0	1	0	0	0	0	...	0
4	5.117994	10.819778	0.0	5000.0	1	0	0	0	0	...	0

Output - 10

x_train

```
array([[ 9.12841566, 13.12236338,  0.         , ...,  0.         ,
         1.         ,  0.         ],
       [ 8.85337967, 13.81551056,  0.         , ...,  0.         ,
         1.         ,  0.         ],
       [12.58716564, 15.42494847,  0.         , ...,  0.         ,
         1.         ,  0.         ],
       ...,
       [13.44466135, 16.11889565,  0.         , ...,  0.         ,
         1.         ,  0.         ],
       [ 5.73889978, 10.81977828,  0.         , ...,  0.         ,
         1.         ,  0.         ],
       [10.63888853, 13.81551056,  0.         , ...,  0.         ,
         1.         ,  0.         ]])
```

y_train

```
array([4, 4, 4, ..., 4, 4, 4])
```

Output - 11

```
from sklearn.linear_model import LinearRegression
reg=LinearRegression()
reg.fit(x_train,y_train)
```

LinearRegression()

```
from sklearn.metrics import r2_score
y_pred=reg.predict(x_test)
r_sq_test=r2_score(y_test,y_pred)
y_pred_train=reg.predict(x_train)
r_sq_train=r2_score(y_train,y_pred_train)
print("R square of train set : ",r_sq_train)
print("R square of test set : ",r_sq_test)
```

```
R square of train set :  0.11282177383868853
R square of test set :  0.08616438378931161
```

Output - 12

```
y_pred_table=pd.DataFrame({"Actual Rating":y_test,"Predicted Rating":y_pred})
```

y_pred_table

	Actual Rating	Predicted Rating
0	3	3.672290
1	4	3.939140
2	4	4.210368
3	4	4.138533
4	5	3.464796
...
2425	4	3.842490
2426	4	3.803011
2427	4	3.599883
2428	4	3.848690
2429	3	3.593998

2430 rows x 2 columns

```
plt.scatter(y_test,y_pred,color="blue")
plt.xlabel("Actual Rating")
plt.ylabel("Predicted Rating")
plt.title("Actual vs Predicted Rating")
plt.show()
```

