

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

## Importing File

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

```
In [3]: play
```

Out[3]:

		Unnamed: 0	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size	Last_Update
0	0	Women Workout at Home - Female Fitness		50000000.0	4.8	Health & Fitness	474119.0	17.0	September 18, 2021
1	1	cult.fit Fitness, Meditation, Healthy food, Do...		5000000.0	4.6	Health & Fitness	97727.0	0.0	September 24, 2021
2	2	Google Fit: Activity Tracking		1000000000.0	3.9	Health & Fitness	469685.0	0.0	September 23, 2021
3	3	30 Day Fitness Challenge - Workout at Home		10000000.0	4.8	Health & Fitness	598899.0	21.0	August 20, 2021
4	4	FitOn - Free Fitness Workouts & Personalized P...		5000000.0	4.6	Health & Fitness	50594.0	50.0	September 5, 2021
...	...	...		...	...	...	...	...	...
307	308	Feisar Watch Face		100000.0	4.4	Watch faces	3349.0	0.0	September 26, 2021
308	309	Marine Commander Watch Face for WearOS		500000.0	4.3	Watch faces	7752.0	0.0	January 29, 2021
309	310	Cronosurf Wave watch		1000000.0	4.3	Watch faces	5358.0	0.0	September 18, 2021
310	312	O-Xyde Watch Face		100000.0	4.3	Watch faces	3158.0	0.0	September 26, 2021
311	313	Ultra Watch Face		500000.0	3.6	Watch faces	2670.0	0.0	September 18, 2021

312 rows × 11 columns

# Dropping Null Values

```
In [4]: play.drop('Unnamed: 0',axis=1,inplace=True) # Dropping Null Values
```

## Replacing Null Values with zero

```
In [5]: play['Current_Version'] = play['Current_Version'].apply(lambda x : str(x).replace
play['Required_Android'] = play['Required_Android'].apply(lambda x : str(x).repla
```

## Dropping irrelevant columns

```
In [6]: # playstore.rename(columns={'Size':'Size_in_MegaBytes'})
```

```
play.drop('Current_Version',axis=1,inplace=True)
play.drop('Required_Android',axis=1,inplace=True)
```

## Renaming columns

```
In [7]: play.rename(columns={'Size':'Size_in_MegaBytes'},inplace=True)
```

```
In [8]: play
```

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
0	Women Workout at Home - Female Fitness	500000000.0	4.8	Health & Fitness	474119.0	17.0	Septe 18,
1	cult.fit Fitness, Meditation, Healthy food, Do...	5000000.0	4.6	Health & Fitness	97727.0	0.0	Septe 24,
2	Google Fit: Activity Tracking	1000000000.0	3.9	Health & Fitness	469685.0	0.0	Septe 23,
3	30 Day Fitness Challenge - Workout at	100000000.0	4.8	Health & Fitness	598899.0	21.0	Augu

```
In [9]: play.Size_in_MegaBytes.median()
```

```
Out[9]: 17.0
```

## Changing datatype

```
In [11]: play['Size_in_MegaBytes'] = play['Size_in_MegaBytes'].astype(float)
```

In [12]: play

Out[12]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
0	Women Workout at Home - Female Fitness	500000000.0	4.8	Health & Fitness	474119.0	17.0	September 18, 2021
1	cult.fit Fitness, Meditation, Healthy food, Do...	5000000.0	4.6	Health & Fitness	97727.0	0.0	September 24, 2021
2	Google Fit: Activity Tracking	1000000000.0	3.9	Health & Fitness	469685.0	0.0	September 23, 2021
3	30 Day Fitness Challenge - Workout at Home	100000000.0	4.8	Health & Fitness	598899.0	21.0	August 20, 2021
4	FitOn - Free Fitness Workouts & Personalized P...	5000000.0	4.6	Health & Fitness	50594.0	50.0	September 1, 2021
...	...	...	...	...	...	...	...
307	Feisar Watch Face	100000.0	4.4	Watch faces	3349.0	0.0	September 26, 2021
308	Marine Commander Watch Face for WearOS	500000.0	4.3	Watch faces	7752.0	0.0	January 28, 2021
309	Cronosurf Wave watch	1000000.0	4.3	Watch faces	5358.0	0.0	September 18, 2021
310	O-Xyde Watch Face	100000.0	4.3	Watch faces	3158.0	0.0	September 26, 2021
311	Ultra Watch Face	500000.0	3.6	Watch faces	2670.0	0.0	September 18, 2021

312 rows × 8 columns



In [13]: `play.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 312 entries, 0 to 311
Data columns (total 8 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   App_Names        312 non-null    object  
 1   No.Of_Installs   312 non-null    float64 
 2   Rating           312 non-null    float64 
 3   Category         312 non-null    object  
 4   No.Of_Review     312 non-null    float64 
 5   Size_in_MegaBytes 312 non-null    float64 
 6   Last_Update       312 non-null    object  
 7   Developer         312 non-null    object  
dtypes: float64(4), object(4)
memory usage: 19.6+ KB
```

In [14]: `play.shape`

Out[14]: (312, 8)

In [15]: `play.isnull().sum()`

```
App_Names      0
No.Of_Installs 0
Rating         0
Category       0
No.Of_Review   0
Size_in_MegaBytes 0
Last_Update    0
Developer      0
dtype: int64
```

## Converting the datatype of `Last_Update` column with datetime datatype

In [16]: `play['Last_Update'] = pd.to_datetime(play['Last_Update'])`

In [17]: `play.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 312 entries, 0 to 311
Data columns (total 8 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   App_Names        312 non-null    object  
 1   No_Of_Installs   312 non-null    float64 
 2   Rating           312 non-null    float64 
 3   Category         312 non-null    object  
 4   No_Of_Review     312 non-null    float64 
 5   Size_in_MegaBytes 312 non-null    float64 
 6   Last_Update      312 non-null    datetime64[ns]
 7   Developer        312 non-null    object  
dtypes: datetime64[ns](1), float64(4), object(3)
memory usage: 19.6+ KB
```

In [18]: `play['Size_in_MegaBytes'].replace(0.0, 17.0, inplace=True)`

## Required DataSet After Cleaning

In [19]: `play`

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Up
0	Women Workout at Home - Female Fitness	500000000.0	4.8	Health & Fitness	474119.0	17.0	2021-0
1	cult.fit Fitness, Meditation, Healthy food, Do...	5000000.0	4.6	Health & Fitness	97727.0	17.0	2021-0
2	Google Fit: Activity Tracking	1000000000.0	3.9	Health & Fitness	469685.0	17.0	2021-0
3	30 Day Fitness Challenge - Workout at	100000000.0	4.8	Health & Fitness	598899.0	21.0	2021-0

In [20]: `play.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 312 entries, 0 to 311
Data columns (total 8 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   App_Names        312 non-null    object  
 1   No_Of_Installs   312 non-null    float64 
 2   Rating           312 non-null    float64 
 3   Category          312 non-null    object  
 4   No_Of_Review     312 non-null    float64 
 5   Size_in_MegaBytes 312 non-null    float64 
 6   Last_Update       312 non-null    datetime64[ns]
 7   Developer         312 non-null    object  
dtypes: datetime64[ns](1), float64(4), object(3)
memory usage: 19.6+ KB
```

## Data Visualization

In [21]: `import seaborn as sns  
import matplotlib.pyplot as plt`

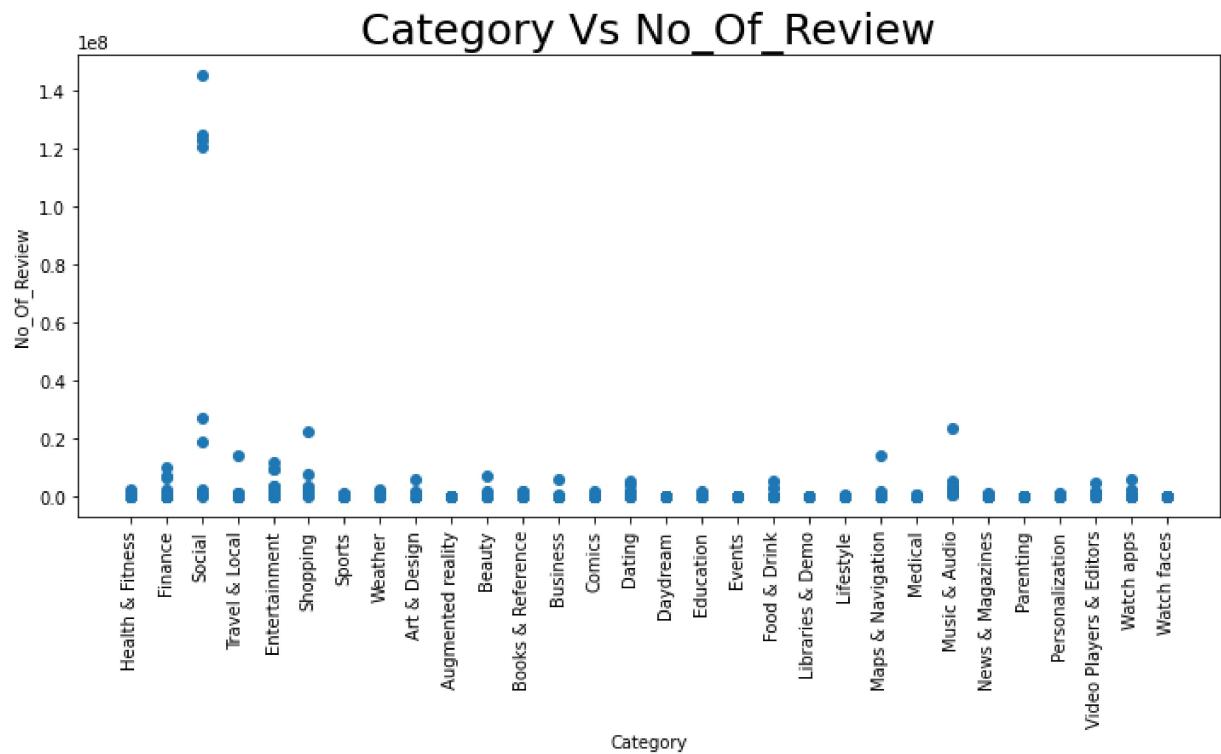
## Review

**The Category 'Social' is most reviewed Category**

**Sports and Daydream is least reviewed Category**

```
In [22]: plt.figure(figsize = (12,5))
```

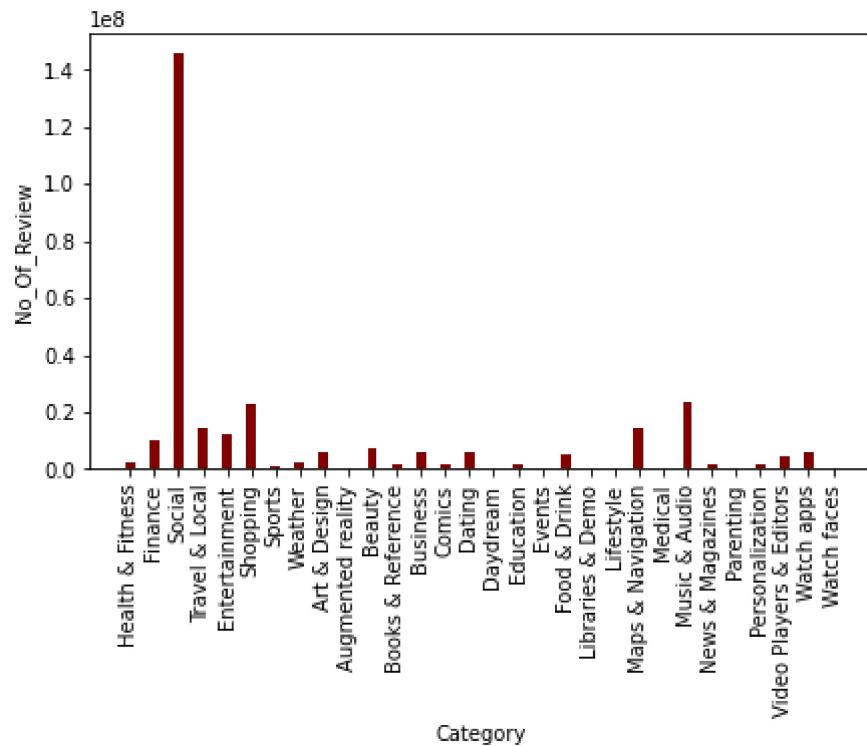
```
plt.scatter(play['Category'],play['No_Of_Review'])
plt.xticks(rotation = 90)
plt.title('Category Vs No_Of_Review',size = 25 )
plt.xlabel('Category')
plt.ylabel('No_Of_Review')
plt.show()
```



```
In [23]: plt.figure(figsize =(7,4))
plt.bar(play['Category'],play['No_Of_Review'],width = 0.4,color ='maroon')
plt.xticks(rotation = 90)

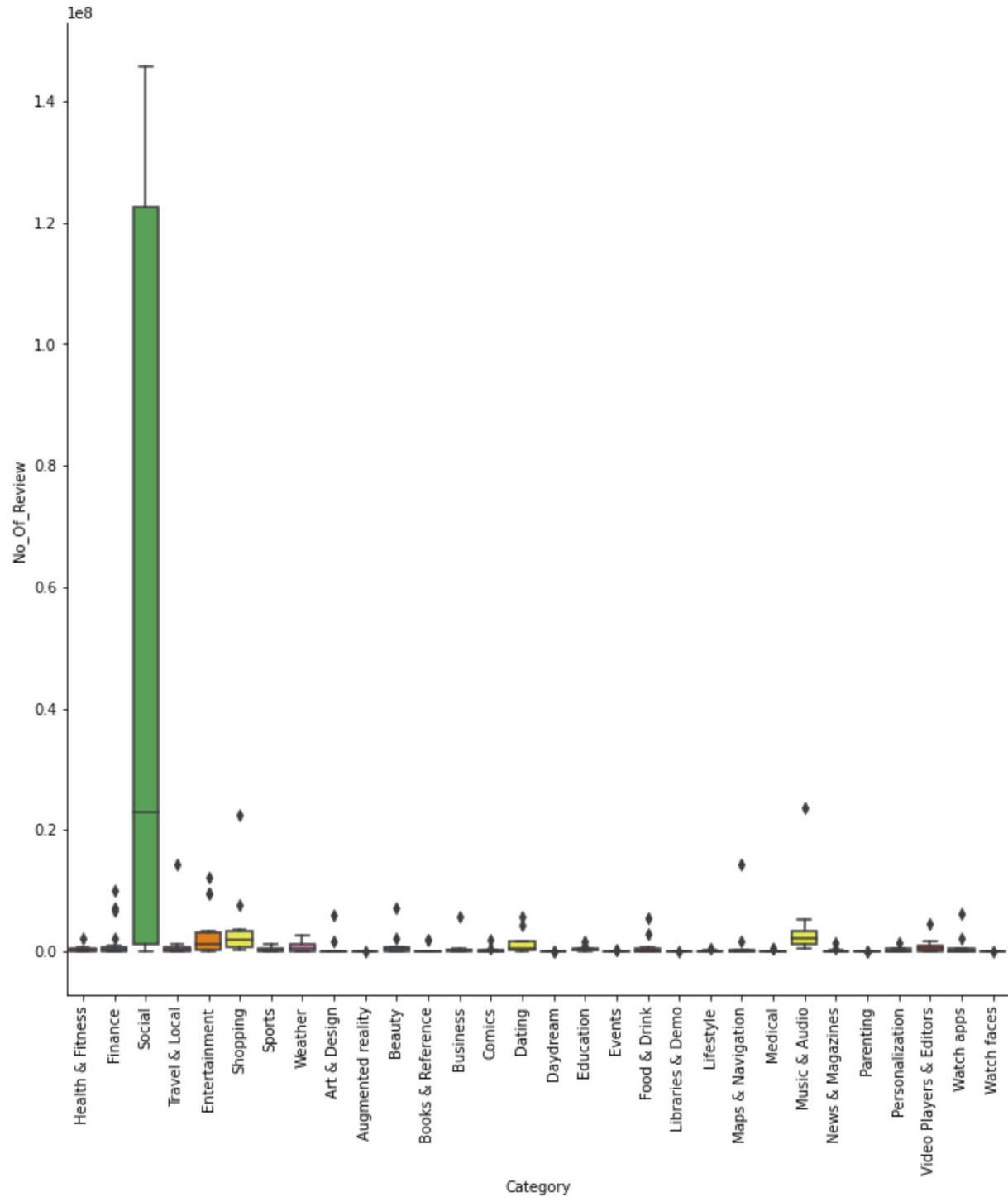
plt.xlabel('Category')
plt.ylabel('No_Of_Review')
plt.show()

# The category 'Social' is most reviewed application
```



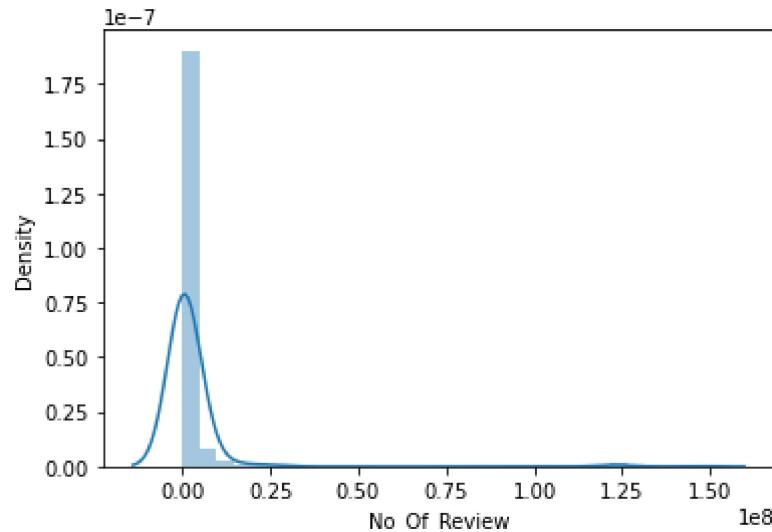
```
In [24]: plt.figure(figsize = (7,4))
g = sns.catplot( x="Category", y='No_Of_Review' ,data = play , kind="box", height=10)
g.set_xticklabels(rotation=90)
g.set(xticks=range(0,30))
plt.show()
```

<Figure size 504x288 with 0 Axes>



```
In [25]: sns.distplot(play['No_of_Review'], bins=30);
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
warnings.warn(msg, FutureWarning)
```



**Social is most reviewed category**

**Augmented Reality, Daydream, lifestyle, Mediacl and watch faces are least reviewed category**

**Whatsapp messenger most revied app**

```
In [26]: play[play['No_Of_Review']==145593486.0]
```

Out[26]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
24	WhatsApp Messenger	5.000000e+09	4.1	Social	145593486.0	17.0	2021-09-14

## All Sport And Battery Daydream is least reviewed app

```
In [27]: play[play['No_Of_Review']==6.0]
```

Out[27]:

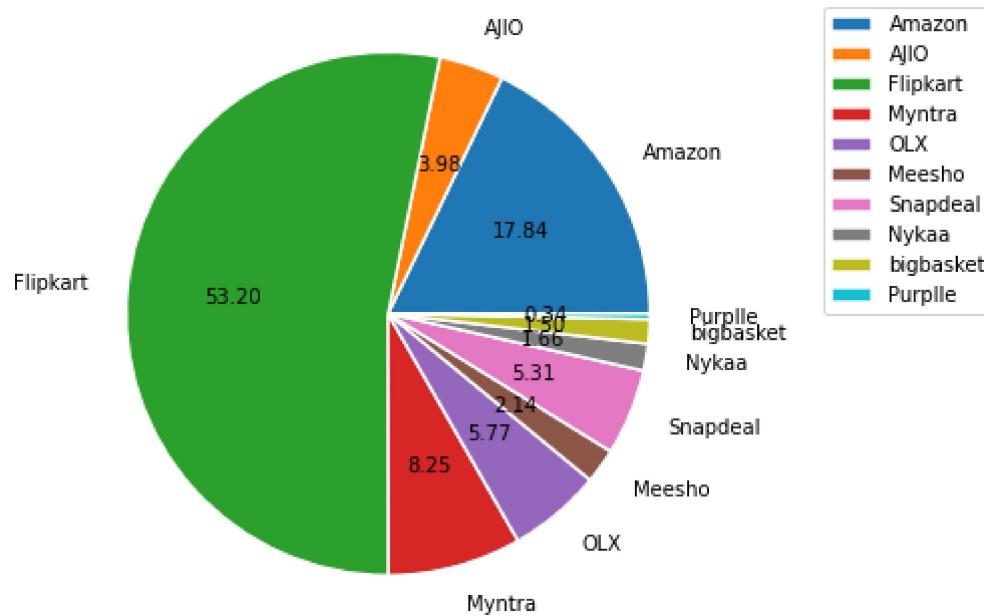
	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updater
60	All Sports	1000.0	4.3	Sports	6.0	4.8	2021-07-01
158	Battery Daydream	1000.0	4.5	Daydream	6.0	1.8	2020-08-27

## Shopping Category

```
In [28]: labels = ['Amazon', 'AJIO', 'Flipkart', 'Myntra', 'OLX',
               'Meesho', 'Snapdeal', 'Nykaa', 'bigbasket', 'Purplle']

plt.pie(play[play['Category']=='Shopping'].loc[50:59, 'No_Of_Review'], labels = labels,
        wedgeprops = { 'linewidth' :1.5, 'edgecolor' : 'white' }, radius = 1.5, autopct = '%.2f')
plt.legend(labels, loc=(1.5,0.5))

plt.show();
```



## Flipkart has maximum review in shopping

```
In [29]: play[play['No_Of_Review']==22445445.0]
```

Out[29]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
52	Flipkart Online Shopping App	1000000000.0	4.3	Shopping	22445445.0	97.0	2021-08-18

```
In [30]: play[play['Category']=='Shopping']['No_Of_Review'].min()
```

Out[30]: 144828.0

## Least Reviewed app in shopping category

```
In [31]: play[play['No_Of_Review']==144828.0].loc[59:59,:]
```

```
Out[31]:
```

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
59	Purple: Beauty Shopping App. Buy Cosmetics On...	10000000.0	4.5	Shopping	144828.0	17.0	2021-09-23



```
In [32]: play[play['Category']=='Music & Audio']['No_Of_Review'].max()
```

```
Out[32]: 23685909.0
```

## MUSIC

```
In [33]: play[play['Category']=='Music & Audio'].loc[246:255,'No_Of_Review']
```

```
Out[33]:
```

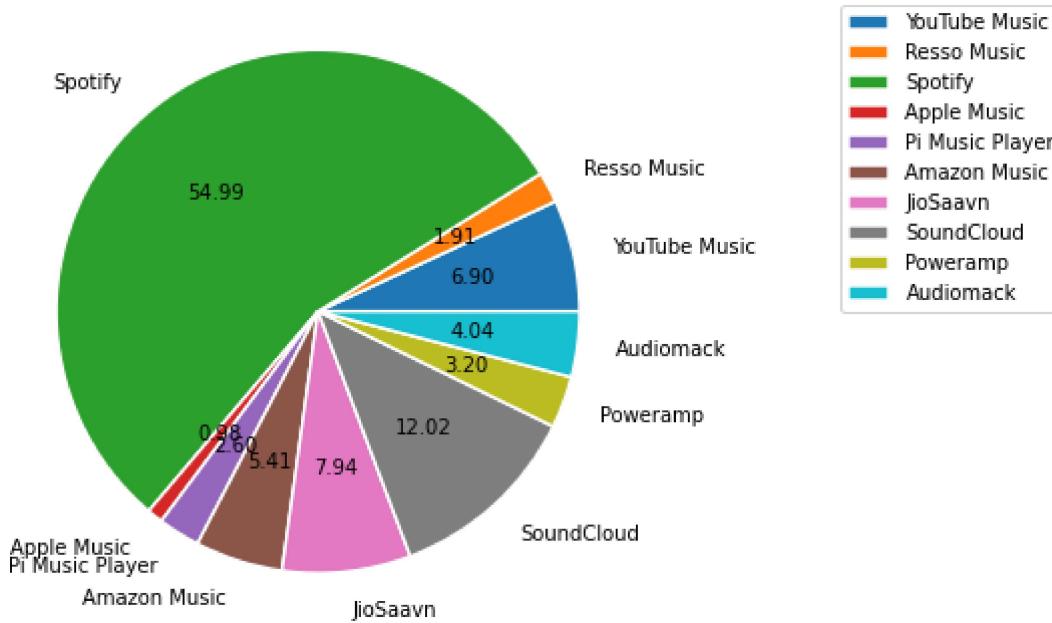
246	2973674.0
247	820777.0
248	23685909.0
249	423669.0
250	1118643.0
251	2332186.0
252	3421173.0
253	5177074.0
254	1377946.0
255	1741054.0

Name: No\_Of\_Review, dtype: float64

```
In [34]: labels = ['YouTube Music','Resso Music','Spotify','Apple Music','Pi Music Player','Amazon Music','JioSaavn','SoundCloud','Poweramp','Audiomack']

plt.pie(play[play['Category']=='Music & Audio'].loc[246:255,'No_Of_Review'],labels=labels,wedgeprops = { 'linewidth' :1.5, 'edgecolor' : 'white' },radius = 1.5,autopct = '%.2f')
plt.legend(labels, loc=(1.7,0.5))

plt.show();
```



## Spotify is most reviewed Music App

```
In [35]: play[play['No_Of_Review']==23685909.0]
```

```
Out[35]:
```

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
248	Spotify: Music and Podcasts	1.000000e+09	4.4	Music & Audio	23685909.0	17.0	2021-09-2

```
In [36]: play[play['Category']=='Music & Audio']['No_Of_Review'].min()
```

```
Out[36]: 423669.0
```

## Apple Music is least Reviewed music app

In [37]: `play[play['No_Of_Review'] == 423669.0]`

Out[37]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
249	Apple Music	500000000.0	3.6	Music & Audio	423669.0	17.0	2021-07-2

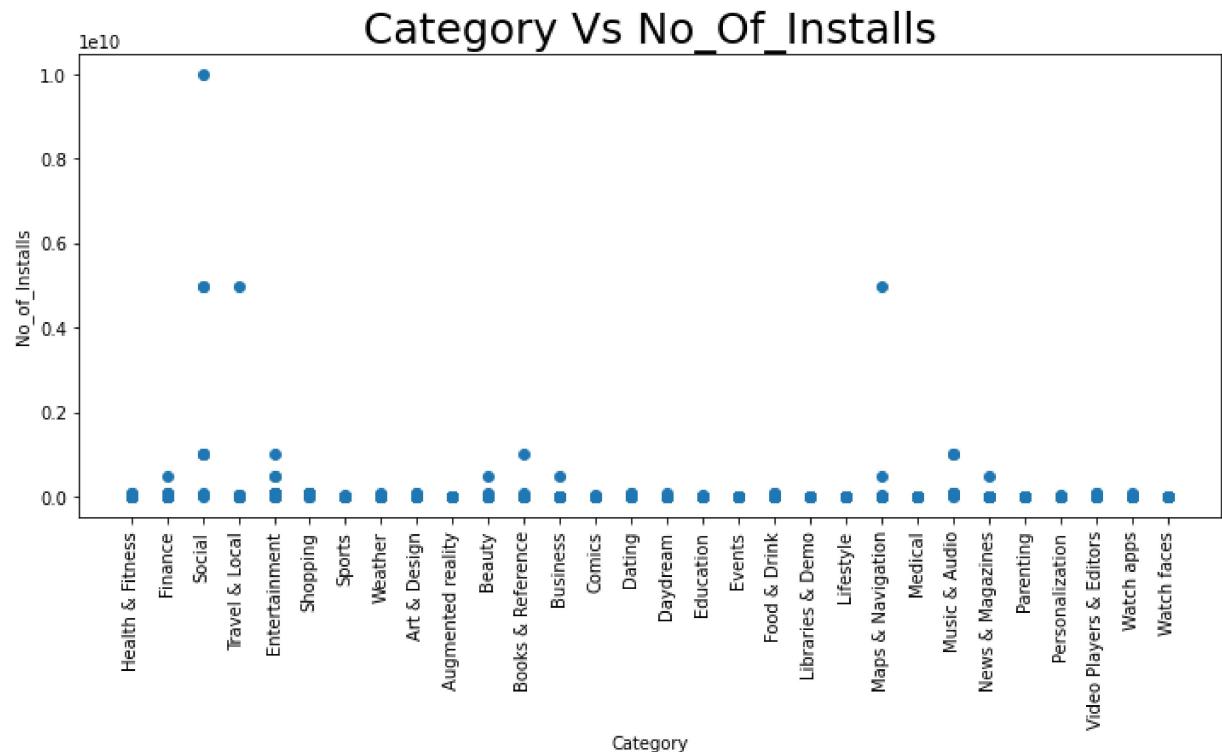
## Number Of Installs

### Social Category are installed the most

In [38]: `plt.figure(figsize = (12,5))`

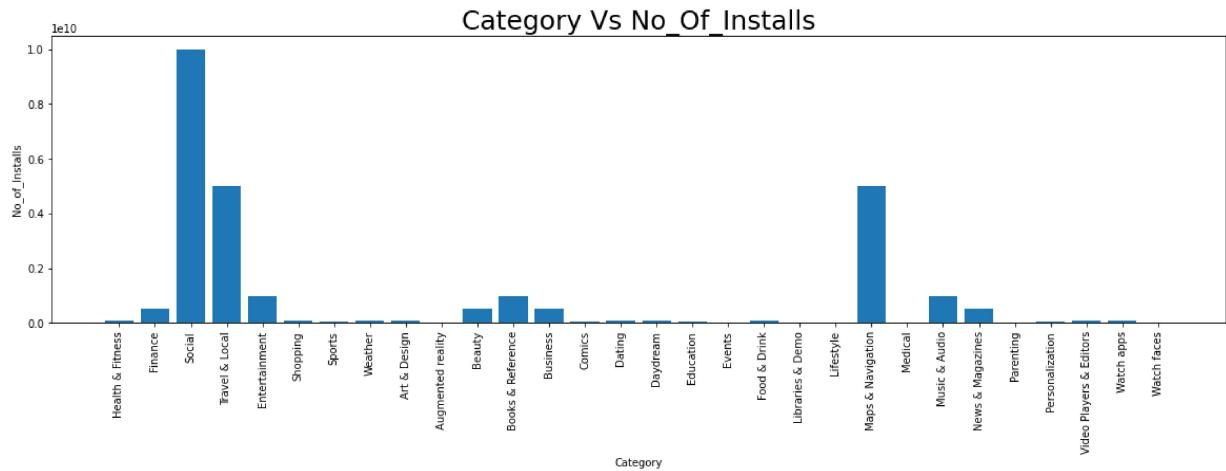
```
plt.scatter(play['Category'], play['No_Of_Installs'])
plt.xticks(rotation = 90)
plt.title('Category Vs No_Of_Installs', size = 25 )
plt.xlabel('Category')
plt.ylabel('No_of_Installs')
plt.show()
```

# Social apps are most installed apps



```
In [39]: plt.figure(figsize =(20,5))
plt.bar(play['Category'],play['No_Of_Installs'])
plt.xticks(rotation = 90)
plt.title('Category Vs No_Of_Installs',size = 25 )
plt.xlabel('Category')
plt.ylabel('No_of_Installs')
plt.show()
```

# social apps are installed th most

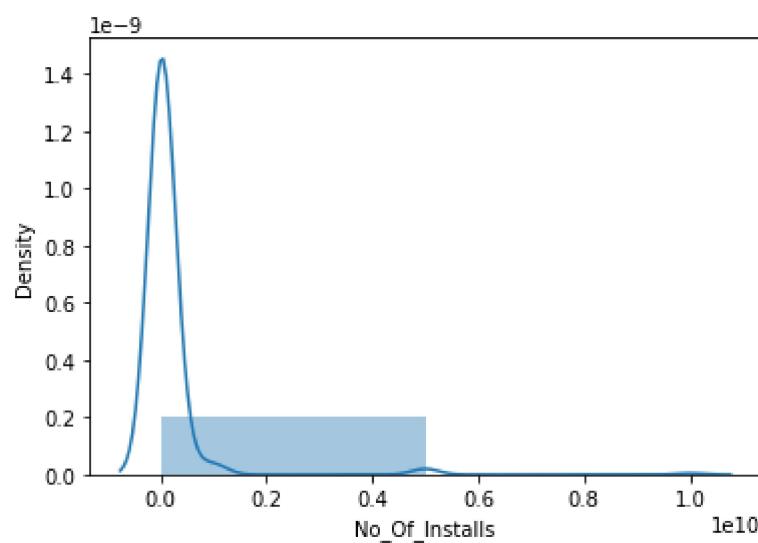


```
In [40]: plt.figure(figsize = (7,4))
g = sns.catplot( x="Category", y='No_Of_Installs' ,data = play , kind="box", height=4)
g.set_xticklabels(rotation=90)
g.set(xticks=range(0,30))
plt.show()
```

<Figure size 504x288 with 0 Axes>

In [41]: `sns.distplot(play['No_of_Installs'], bins=2);`

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```



In [42]: `play[play['Category']=='Social']['No_of_Installs'].max()`

Out[42]: 10000000000.0

## YouTube is most installed Social app

In [43]: `play[play['No_of_Installs']==10000000000.0]`

	App_Names	No_of_Installs	Rating	Category	No_of_Review	Size_in_MegaBytes	Last_Update
23	YouTube	1.000000e+10	4.3	Social	123129731.0	17.0	2021-09-24

◀ ▶

In [44]: `play['No_of_Installs'].min()`

Out[44]: 1000.0

## These are the least Installed apps

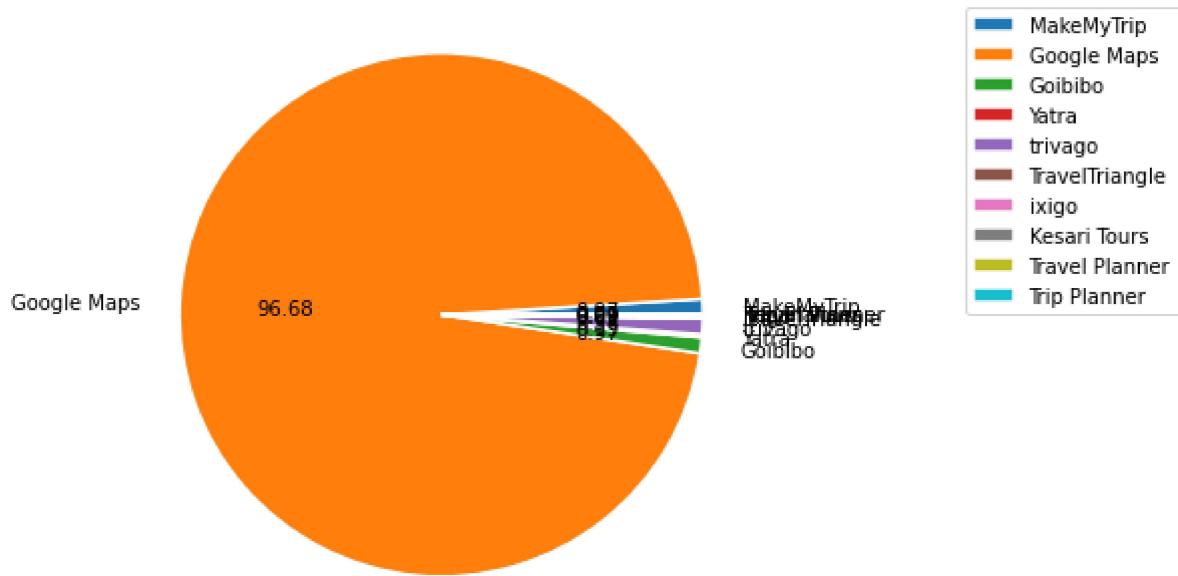
In [45]: `play[play['No_of_Installs']==1000.0]`

Out[45]:

	App_Names	No_of_Installs	Rating	Category	No_of_Review	Size_in_MegaBytes	Last_Updated
60	All Sports	1000.0	4.3	Sports	6.0	4.8	2021-07-10
95	Halo AR - Augmented Reality, VR & 3D Lens Crea...	1000.0	3.6	Augmented reality	10.0	64.0	2021-09-01
158	Battery Daydream	1000.0	4.5	Daydream	6.0	1.8	2020-08-10
189	Eventtus: Virtual Events	1000.0	3.9	Events	9.0	46.0	2021-06-10
214	DDC Schedule	1000.0	4.2	Libraries & Demo	18.0	8.6	2021-07-10
215	Material Gallery - Android Source Code	1000.0	4.7	Libraries & Demo	100.0	6.4	2021-06-10
262	Nuzzera - Experience News & Magazines	1000.0	4.7	News & Magazines	79.0	2.9	2020-06-10
265	News & Magazines	1000.0	4.5	News & Magazines	22.0	3.7	2017-10-10

## Travel & Local

```
In [46]: labels = ['MakeMyTrip', 'Google Maps', 'Goibibo', 'Yatra', 'trivago',  
    'TravelTriangle', 'ixigo', 'Kesari Tours', 'Travel Planner', 'Trip Planner'  
  
plt.pie(play[play['Category']=='Travel & Local'].loc[30:39,'No_of_Installs'],labels=labels,wedgeprops = { 'linewidth' :1.5, 'edgecolor' : 'white' },radius = 1.5,auto_startAngle=True)  
plt.legend(labels, loc=(1.7,0.5))  
plt.show();
```



```
In [47]: play[play['Category']=='Travel & Local']['No_of_Installs'].max()
```

Out[47]: 50000000000.0

## Google Maps is most installed Travel and local app

```
In [48]: play[play['No_of_Installs']==5000000000.0].loc[31:31,:]
```

	App_Names	No_of_Installs	Rating	Category	No_of_Review	Size_in_MegaBytes	Last_Update
31	Google Maps	5.000000e+09	3.8	Travel & Local	14215173.0	17.0	2021-09-22

```
In [49]: play[play['Category']=='Travel & Local']['No_of_Installs'].min()
```

Out[49]: 100000.0

## Kesari tour and trip planner is least installed travel and local apps

```
In [50]: play[play['No_of_Installs']==100000.0].loc[37:37,:]
```

Out[50]:

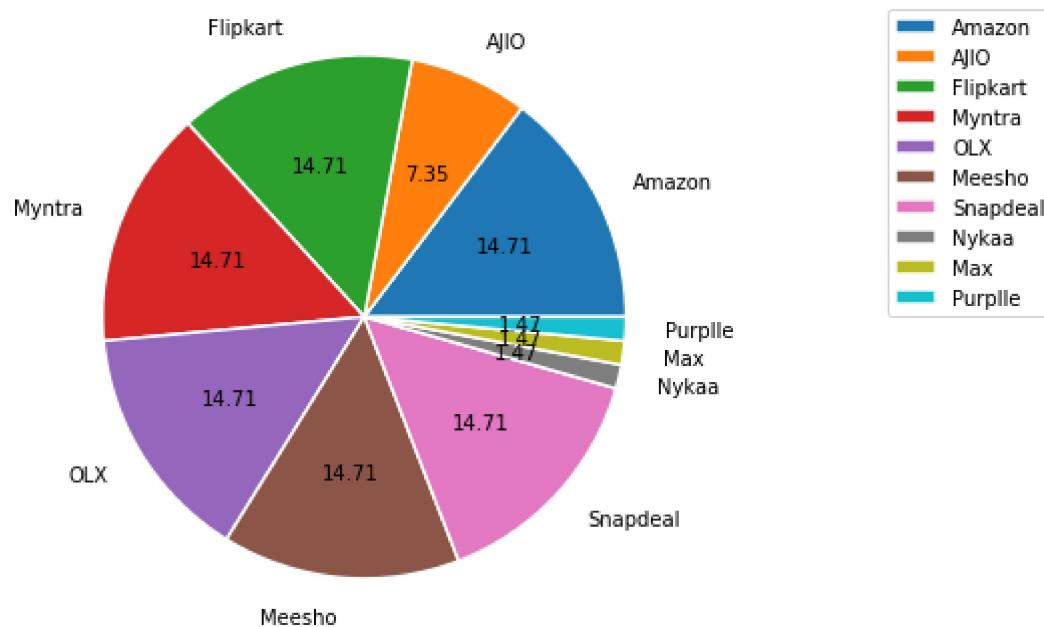
	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
37	Kesari Tours	1000000.0	3.6	Travel & Local	1124.0	28.0	2021-02-11

## Shopping

```
In [51]: labels = ['Amazon', 'AJIO', 'Flipkart', 'Myntra', 'OLX',
               'Meesho', 'Snapdeal', 'Nykaa', 'Max ', 'Purplle']

plt.pie(play[play['Category']=='Shopping'].loc[50:59,'No_of_Installs'], labels = labels,
        wedgeprops = { 'linewidth' : 1.5, 'edgecolor' : 'white' }, radius = 1.5, autopct='%.2f')
plt.legend(labels, loc=(1.7,0.5))

plt.show();
```



```
In [52]: play[play['Category']=='Shopping']['No_of_Installs'].max()
```

Out[52]: 1000000000.0

**OLX , Mesho, Snapdeal is most installed Shopping App**

In [53]: `play[play['No_Of_Installs']==100000000.0].loc[50:60,:]`

Out[53]:

		App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
50		Amazon Shopping, UPI, Money Transfer, Bill Pay...	100000000.0	4.3	Shopping	7525721.0	17.0	2021-09-10
52		Flipkart Online Shopping App	100000000.0	4.3	Shopping	22445445.0	97.0	2021-08-18
53		Myntra Online Shopping App - Shop Fashion & more	100000000.0	4.3	Shopping	3482732.0	17.0	2021-09-25
54		OLX: Buy & Sell Near You with Online Classifieds	100000000.0	4.2	Shopping	2434758.0	22.0	2021-09-21
55		Meesho: Online Shopping App	100000000.0	4.4	Shopping	903862.0	16.0	2021-09-21
56		Snapdeal Shopping App -Free Delivery on all or...	100000000.0	4.3	Shopping	2241650.0	30.0	2021-09-16



In [54]: `play[play['Category']=='Shopping']['No_Of_Installs'].min()`

Out[54]: 10000000.0

In [55]: `play[play['Category']=='Shopping']`

Out[55]:

		App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
50	Amazon Shopping, UPI, Money Transfer, Bill Pay...		1000000000.0	4.3	Shopping	7525721.0	17.0	2021-09-10
51	AJIO Online Shopping - Handpicked Curated Fashion		50000000.0	4.3	Shopping	1681335.0	19.0	2021-09-20
52	Flipkart Online Shopping App		1000000000.0	4.3	Shopping	22445445.0	97.0	2021-08-18
53	Myntra Online Shopping App - Shop Fashion & more		1000000000.0	4.3	Shopping	3482732.0	17.0	2021-09-25
54	OLX: Buy & Sell Near You with Online Classifieds		1000000000.0	4.2	Shopping	2434758.0	22.0	2021-09-21
55	Meesho: Online Shopping App		1000000000.0	4.4	Shopping	903862.0	16.0	2021-09-21
56	Snapdeal Shopping App -Free Delivery on all or...		1000000000.0	4.3	Shopping	2241650.0	30.0	2021-09-16
57	Nykaa: Beauty Shopping App. Buy Makeup & Cosme...		10000000.0	4.6	Shopping	699145.0	70.0	2021-09-22
58	bigbasket- Online Grocery Shopping, Home Delivery		10000000.0	4.4	Shopping	632870.0	20.0	2021-09-09
59	Purple: Beauty Shopping App. Buy Cosmetics On...		10000000.0	4.5	Shopping	144828.0	17.0	2021-09-23



## Nykaa , bigbasket , Purplle are least installed apps

In [56]: `play[play['No.Of.Installs']==10000000.0].loc[50:59]`

Out[56]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
57	Nykaa: Beauty Shopping App. Buy Makeup & Cosme...	10000000.0	4.6	Shopping	699145.0	70.0	2021-09-22
58	bigbasket- Online Grocery Shopping, Home Delivery	10000000.0	4.4	Shopping	632870.0	20.0	2021-09-09
59	Purplle: Beauty Shopping App. Buy Cosmetics On...	10000000.0	4.5	Shopping	144828.0	17.0	2021-09-23



## Developer

There Are 13 Apps with Google LLC Developer

In [57]: `play[play['Developer']=='Google LLC']`

Out[57]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
2	Google Fit: Activity Tracking	1.000000e+08	3.9	Health & Fitness	469685.0	17.0	2021-09-1
11	Google Pay: Save, Pay, Manage	5.000000e+08	4.2	Finance	6743758.0	17.0	2021-09-1
23	YouTube	1.000000e+10	4.3	Social	123129731.0	17.0	2021-09-1
31	Google Maps	5.000000e+09	3.8	Travel & Local	14215173.0	17.0	2021-09-1
111	Google Play Books & Audiobooks	1.000000e+09	4.1	Books & Reference	1845020.0	17.0	2021-09-1
120	Google My Business	1.000000e+07	4.3	Business	261634.0	17.0	2021-09-1
150	Daydream	1.000000e+07	3.4	Daydream	5090.0	35.0	2019-09-1
151	Google VR Services	1.000000e+08	3.9	Daydream	42519.0	17.0	2019-09-1
153	YouTube VR	1.000000e+07	3.7	Daydream	4127.0	24.0	2020-10-1
155	Daydream Keyboard	1.000000e+07	3.9	Daydream	2347.0	46.0	2019-09-1
228	Google Maps	5.000000e+09	3.8	Maps & Navigation	14215248.0	17.0	2021-09-1
235	Google Maps Go	5.000000e+08	4.3	Maps & Navigation	179772.0	2.3	2021-02-1
246	YouTube Music	1.000000e+09	4.3	Music & Audio	2973674.0	17.0	2021-09-1



In [58]: `a = play[play['Developer']=='Google LLC']  
print(f"Number of Google Developed Apps are {len(a)}")`

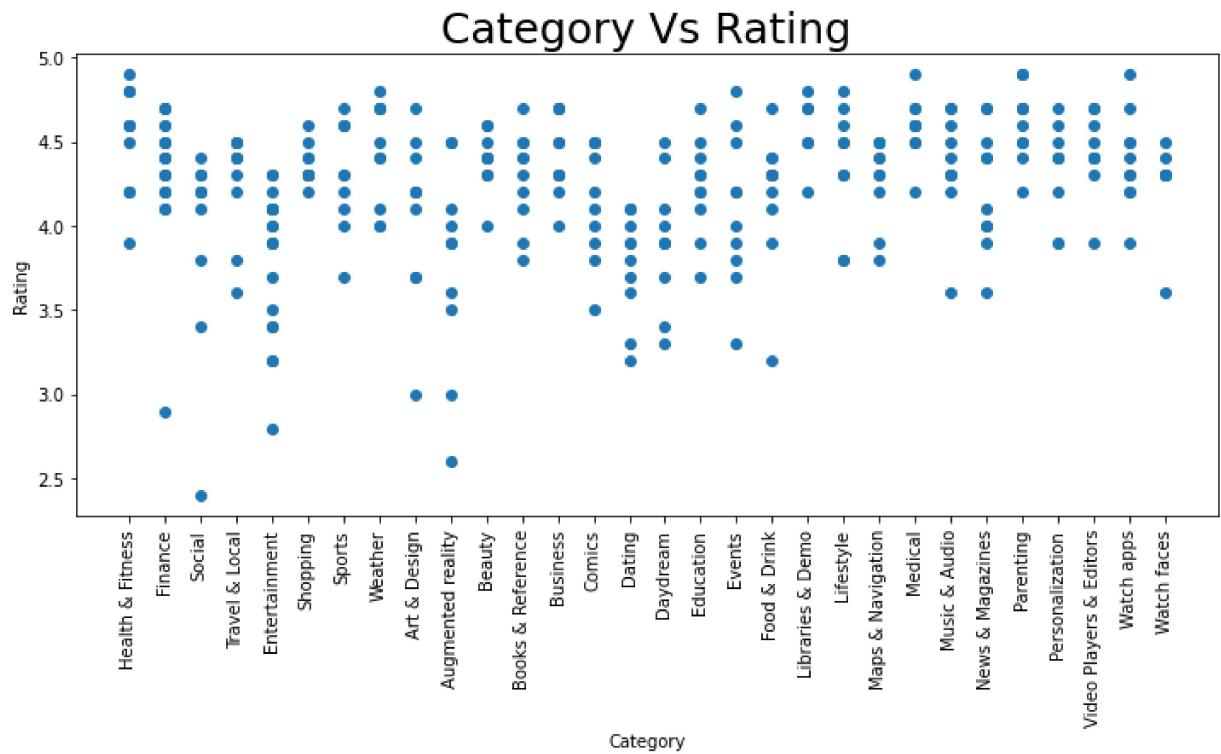
Number of Google Developed Apps are 13

## Rating

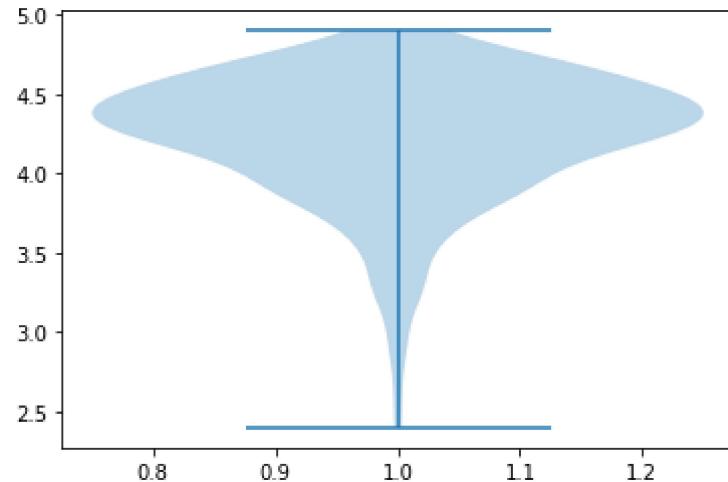
**Lowest rating is given to social media category whereas highest rating is almost same for all category not much difference**

```
In [59]: plt.figure(figsize = (12,5))

plt.scatter(play['Category'],play['Rating'])
plt.xticks(rotation = 90)
plt.title('Category Vs Rating',size = 25 )
plt.xlabel('Category')
plt.ylabel('Rating')
plt.show()
# Lowest rating is given to social media category whereas highest rating is aln
```

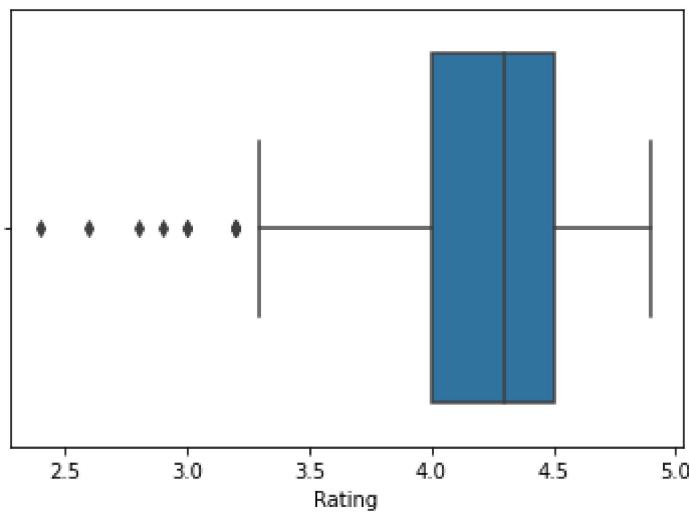


```
In [60]: plt.violinplot(play['Rating'])
plt.show()
```

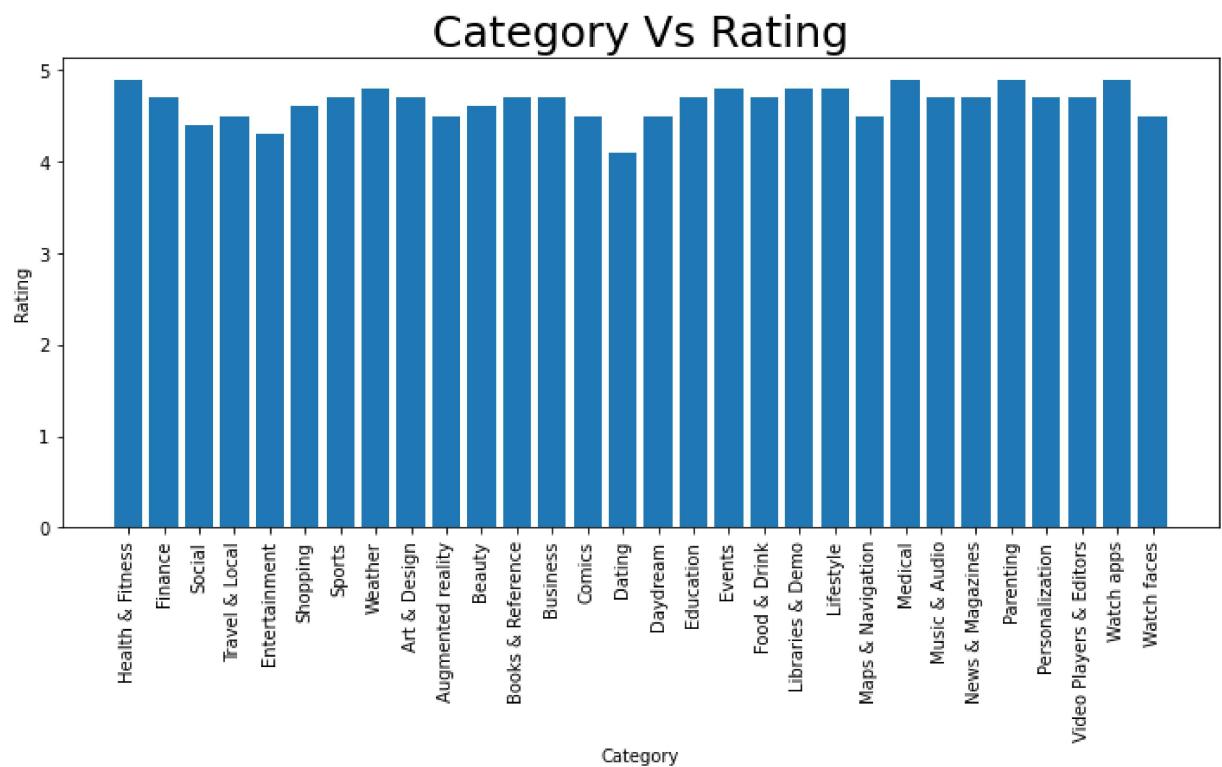


```
In [61]: sns.boxplot(play["Rating"])
plt.show()
```

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.  
warnings.warn(

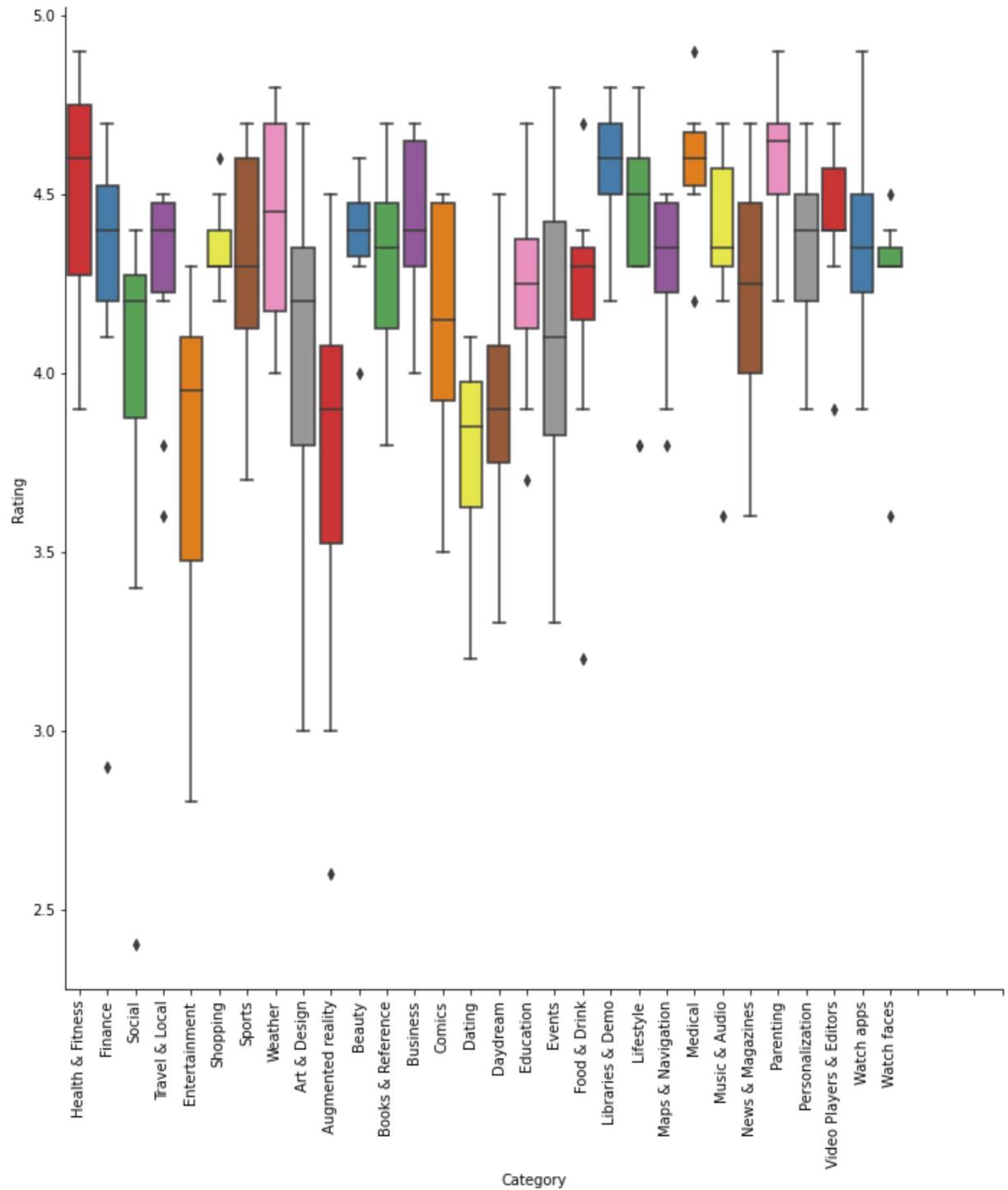


```
In [62]: plt.figure(figsize =(12,5))
plt.bar(play['Category'],play['Rating'])
plt.xticks(rotation = 90)
plt.title('Category Vs Rating',size = 25 )
plt.xlabel('Category')
plt.ylabel('Rating')
plt.show()
```



```
In [63]: plt.figure(figsize = (7,2))
g = sns.catplot( x="Category", y='Rating' ,data = play , kind="box", height = 10,
g.set_xticklabels(rotation=90)
g.set( xticks=range(0,34))
plt.show()
```

&lt;Figure size 504x144 with 0 Axes&gt;



## Social is least reviewed category

facebook is least reviewed app

```
In [64]: play[play['Rating']==2.4]
```

Out[64]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
21	Facebook	5.000000e+09	2.4	Social	121031089.0	17.0	2021-09-22

```
In [65]: play[play['Rating']<3]
```

Out[65]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_U
21	Facebook	5.000000e+09	2.4	Social	121031089.0	17.0	2021-
97	Genesis Augmented Reality	1.000000e+04	2.6	Augmented reality	332.0	35.0	2018-
172	Eros Now - Movies, Originals, Music & TV Shows	1.000000e+07	2.8	Entertainment	119135.0	18.0	2021-
193	Spendee - Budget and Expense Tracker & Planner	1.000000e+06	2.9	Finance	30260.0	63.0	2021-

In [66]: `play['Rating'].max()`

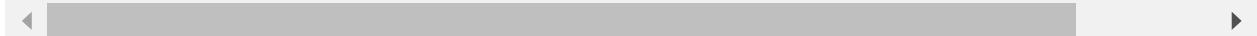
Out[66]: 4.9

## There are five apps with highest rating

In [67]: `play[play["Rating"]==4.9]`

Out[67]:

		App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
	8	Home Workout - No Equipment	1000000000.0	4.9	Health & Fitness	2183906.0	17.0	2021-08-0
	244	MDCalc Medical Calculator	500000.0	4.9	Medical	7517.0	62.0	2021-08-1
	267	The Happy Child - Parenting App	100000.0	4.9	Parenting	8770.0	17.0	2021-06-2
	275	Positive Parenting Solutions	50000.0	4.9	Parenting	640.0	26.0	2021-08-1
	295	Period Tracker - Period Calendar Ovulation Tra...	1000000000.0	4.9	Watch apps	6151543.0	17.0	2021-08-1



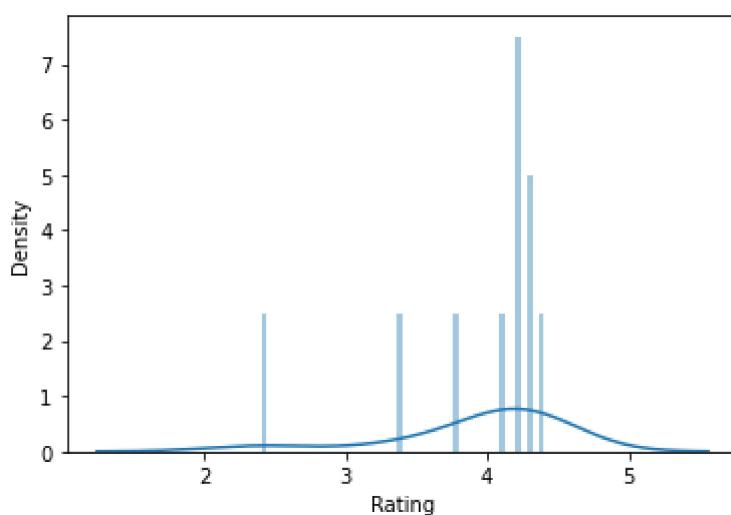
```
In [68]: sns.distplot(play[play['Category']=='Social']['Rating'],bins=50);
```

```
a = play[play['Category']=='Social']['Rating'].min()  
b = play[play['Category']=='Social']['Rating'].max()  
print(a)  
print(b)
```

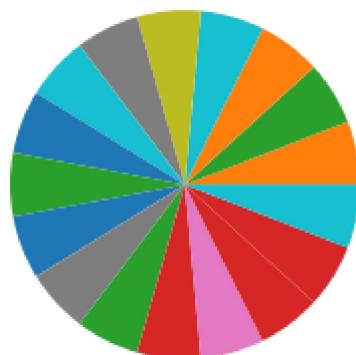
```
2.4  
4.4
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histogram s).
```

```
warnings.warn(msg, FutureWarning)
```



```
In [69]: plt.pie(play['Rating']<3.5)  
plt.show()  
# these are the 18 apps
```



# Reddit is most rated Social app

In [70]: `play[play['Rating']==4.4].loc[25:25,:]`

Out[70]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
25	Reddit	500000000.0	4.4	Social	2247394.0	17.0	2021-09-21

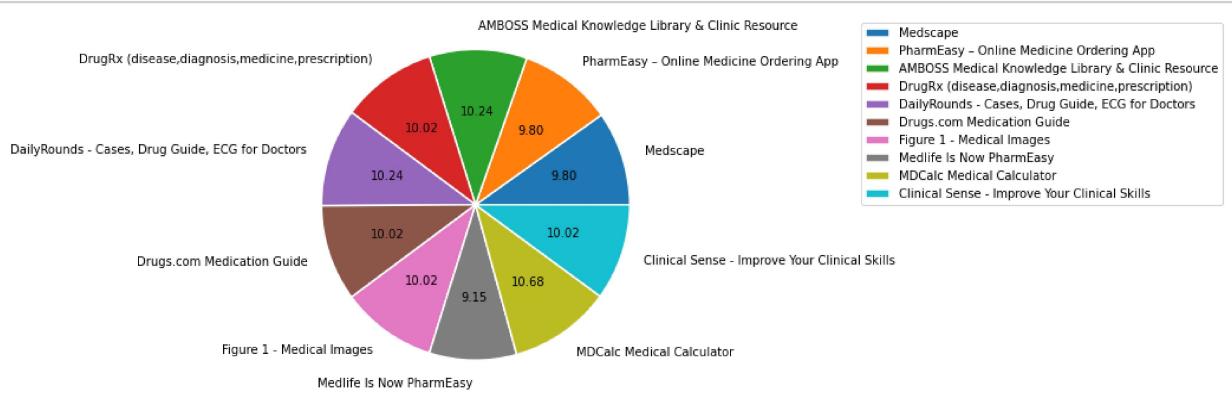
## Medical Category

In [71]: `play[play['Category']=='Medical']  
# ["App_Names"].unique()  
# .Loc[50:, 'No_Of_Review']`

Out[71]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_ir
236	Medscape	5000000.0	4.5	Medical	60521.0	
237	PharmEasy – Online Medicine Ordering App	10000000.0	4.5	Medical	360071.0	
238	AMBOSS Medical Knowledge Library & Clinic Reso...	100000.0	4.7	Medical	3235.0	
239	DrugRx (disease,diagnosis,medicine,prescription)	50000.0	4.6	Medical	266.0	
240	DailyRounds - Cases, Drug Guide, ECG for Doctors	100000.0	4.7	Medical	16298.0	
241	Drugs.com Medication Guide	1000000.0	4.6	Medical	25843.0	
242	Figure 1 - Medical Images	1000000.0	4.6	Medical	13536.0	
243	Medlife Is Now PharmEasy	10000000.0	4.2	Medical	362672.0	
244	MDCalc Medical Calculator	500000.0	4.9	Medical	7517.0	
245	Clinical Sense - Improve Your Clinical Skills	1000000.0	4.6	Medical	11959.0	

```
In [72]: labels = ['Medscape', 'PharmEasy - Online Medicine Ordering App',  
               'AMBOSS Medical Knowledge Library & Clinic Resource',  
               'DrugRx (disease,diagnosis,medicine,prescription)',  
               'DailyRounds - Cases, Drug Guide, ECG for Doctors',  
               'Drugs.com Medication Guide', 'Figure 1 - Medical Images',  
               'Medlife Is Now PharmEasy', 'MDCalc Medical Calculator',  
               'Clinical Sense - Improve Your Clinical Skills']  
  
plt.pie(play[play['Category']=='Medical'].loc[236:245,'Rating'],labels = labels,]  
        wedgeprops = { 'linewidth' :1.5, 'edgecolor' : 'white' },radius = 1.5,auto  
plt.legend(labels, loc=(2,0.5))  
plt.show();
```



**MDCalc Medical Calculator is most rated medical app**

In [73]: `play[play['Rating']==4.9]`

Out[73]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
8	Home Workout - No Equipment	1000000000.0	4.9	Health & Fitness	2183906.0	17.0	2021-08-0
244	MDCalc Medical Calculator	500000.0	4.9	Medical	7517.0	62.0	2021-08-1
267	The Happy Child - Parenting App	100000.0	4.9	Parenting	8770.0	17.0	2021-06-2
275	Positive Parenting Solutions	50000.0	4.9	Parenting	640.0	26.0	2021-08-1
295	Period Tracker - Period Calendar Ovulation Tra...	1000000000.0	4.9	Watch apps	6151543.0	17.0	2021-08-1

In [74]: `play[play['Category']=='Medical']['Rating'].min()`

Out[74]: 4.2

## Medlife Is Now PharmEasy is least rated medical app

In [75]: `play[play['Rating']==4.2].loc[243:243,:]`

Out[75]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
243	Medlife Is Now PharmEasy	10000000.0	4.2	Medical	362672.0	17.0	2021-05-0

## Health and fitness

```
In [76]: play[play['Category']=='Health & Fitness']
# ["App_Names"].unique()
# .Loc[50:, 'No_Of_Review']
```

Out[76]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
0	Women Workout at Home - Female Fitness	50000000.0	4.8	Health & Fitness	474119.0	17.0	2021-09-18
1	cult.fit Fitness, Meditation, Healthy food, Do...	5000000.0	4.6	Health & Fitness	97727.0	17.0	2021-09-24
2	Google Fit: Activity Tracking	100000000.0	3.9	Health & Fitness	469685.0	17.0	2021-09-23
3	30 Day Fitness Challenge - Workout at Home	10000000.0	4.8	Health & Fitness	598899.0	21.0	2021-08-20
4	FitOn - Free Fitness Workouts & Personalized P...	5000000.0	4.6	Health & Fitness	50594.0	50.0	2021-09-05
5	Health Diet Foods Fitness Help	500000.0	4.6	Health & Fitness	7306.0	29.0	2021-02-24
6	Calorie Counter, Diet Plan, Dietitians, Trainers	10000000.0	4.5	Health & Fitness	283458.0	45.0	2021-09-19
7	FITTR: Fat-loss plan, workout & personal training	1000000.0	4.2	Health & Fitness	17223.0	85.0	2021-09-24
8	Home Workout - No Equipment	100000000.0	4.9	Health & Fitness	2183906.0	17.0	2021-08-03
9	HealthKart: Health & Bodybuilding Supplements App	1000000.0	4.2	Health & Fitness	51124.0	34.0	2021-09-25

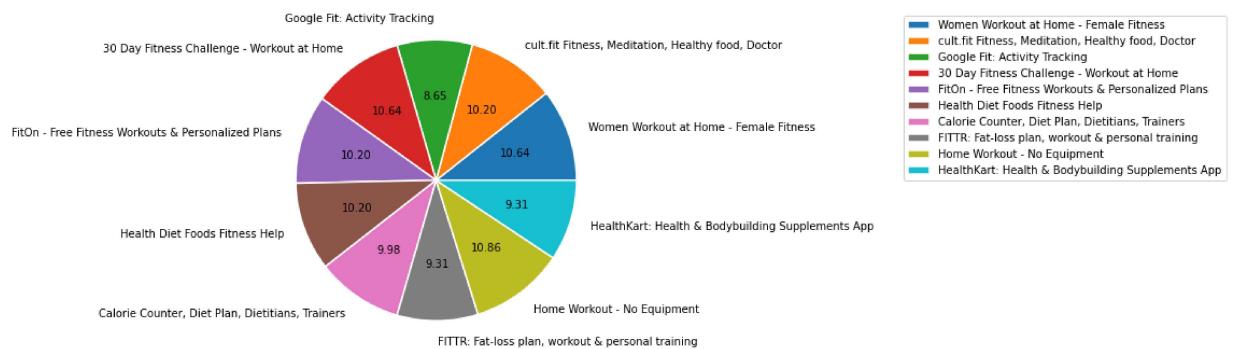


```
In [77]: labels = ['Women Workout at Home - Female Fitness',
   'cult.fit Fitness, Meditation, Healthy food, Doctor',
   'Google Fit: Activity Tracking',
   '30 Day Fitness Challenge - Workout at Home',
   'FitOn - Free Fitness Workouts & Personalized Plans',
   'Health Diet Foods Fitness Help',
   'Calorie Counter, Diet Plan, Dietitians, Trainers',
   'FITTR: Fat-loss plan, workout & personal training',
   'Home Workout - No Equipment',
   'HealthKart: Health & Bodybuilding Supplements App']

plt.pie(play[play['Category']=='Health & Fitness'].loc[0:9,'Rating'],labels = labels,
       wedgeprops = { 'linewidth' :1.5, 'edgecolor' : 'white' },radius = 1.5,auto_startAngle=90)

plt.legend(labels, loc=(2.5,0.5))

plt.show();
```



## Google Fit: Activity Tracking Least Rated Health And Fitness App

```
In [78]: play[play['App_Names']=='Google Fit: Activity Tracking']
```

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
2	Google Fit: Activity Tracking	1000000000.0	3.9	Health & Fitness	469685.0	17.0	2021-09-23

## Home Workout - No Equipment Most Rated Health And Fitness App

```
In [79]: play[play['App_Names']=='Home Workout - No Equipment']
```

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
8	Home Workout - No Equipment	1000000000.0	4.9	Health & Fitness	2183906.0	17.0	2021-08-03

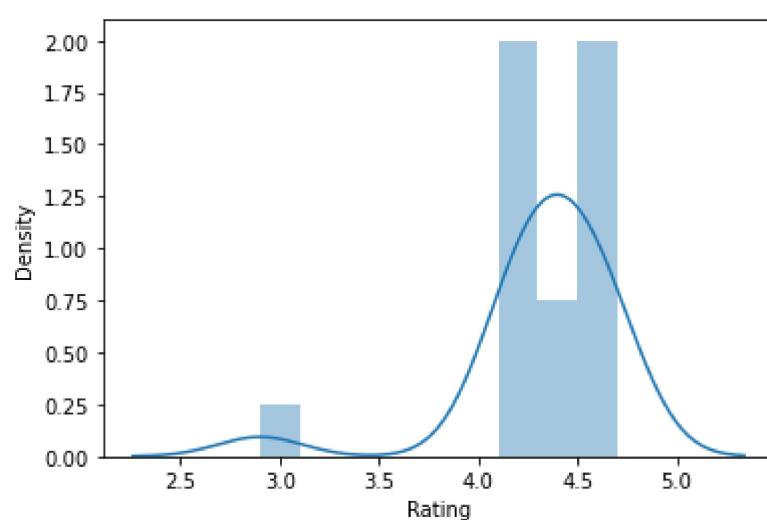
## Finance

```
In [80]: sns.distplot(play[play['Category']=='Finance']['Rating'],bins=9);
a = play[play['Category']=='Finance']['Rating'].min()
b = play[play['Category']=='Finance']['Rating'].max()
print(a)
print(b)
```

2.9  
4.7

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)



**Spendee - Budget and Expense Tracker & Planner is Least Rated Finance App**

In [81]: `play[play['App_Names']=='Spendee - Budget and Expense Tracker & Planner']`

Out[81]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
193	Spendee - Budget and Expense Tracker & Planner	10000000.0	2.9	Finance	30260.0	63.0	2021-08-1

## CRED is most rated finance app

In [82]: `play[play['Rating']==4.7].loc[12:12,:]`

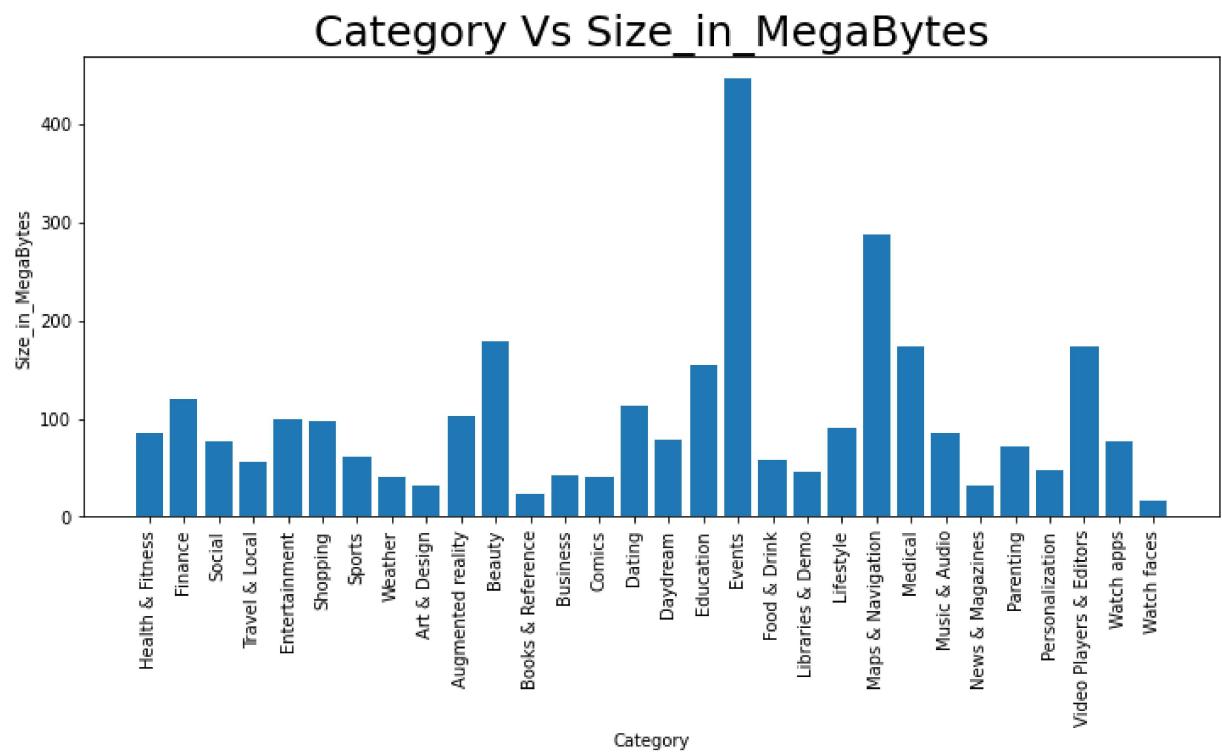
Out[82]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Update
12	CRED: Credit Card Bills, Credit Score & Pay Rent	10000000.0	4.7	Finance	674015.0	49.0	2021-09-10

## Size In MegaBytes

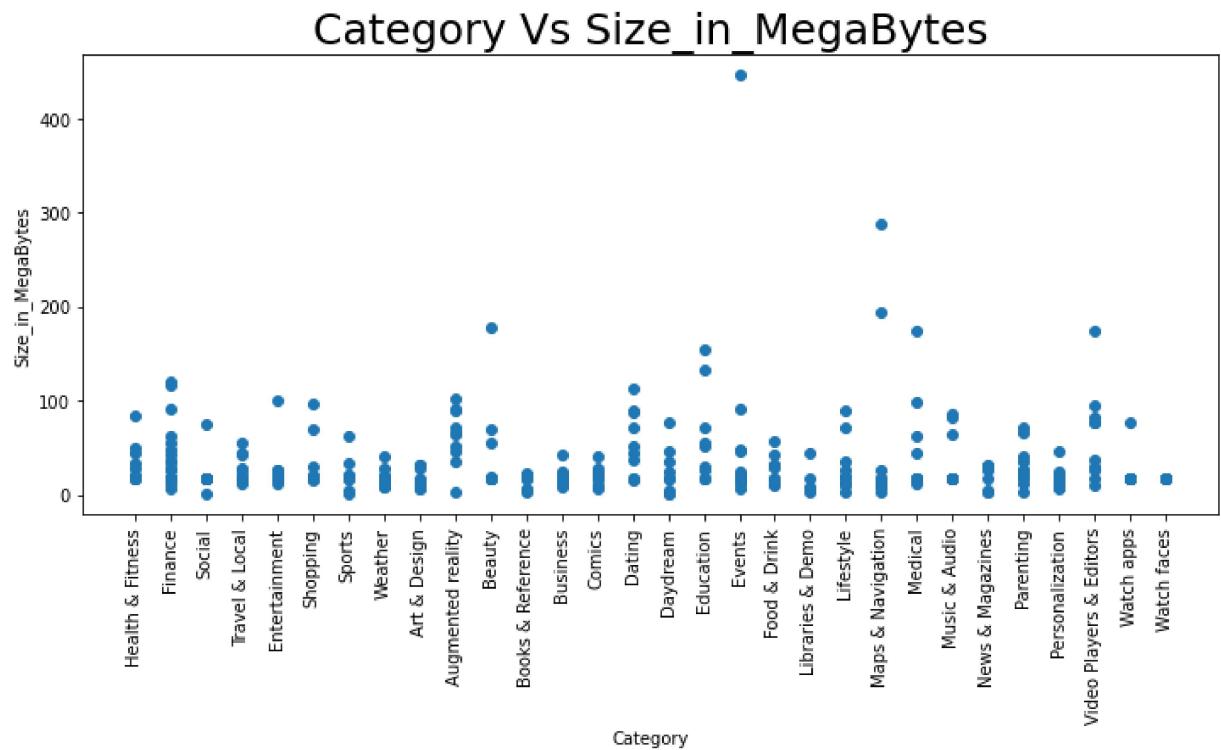
The app size of Event Category is most

```
In [83]: plt.figure(figsize =(12,5))
plt.bar(play['Category'],play['Size_in_MegaBytes'])
plt.xticks(rotation = 90)
plt.title('Category Vs Size_in_MegaBytes',size = 25 )
plt.xlabel('Category')
plt.ylabel('Size_in_MegaBytes')
plt.show()
```



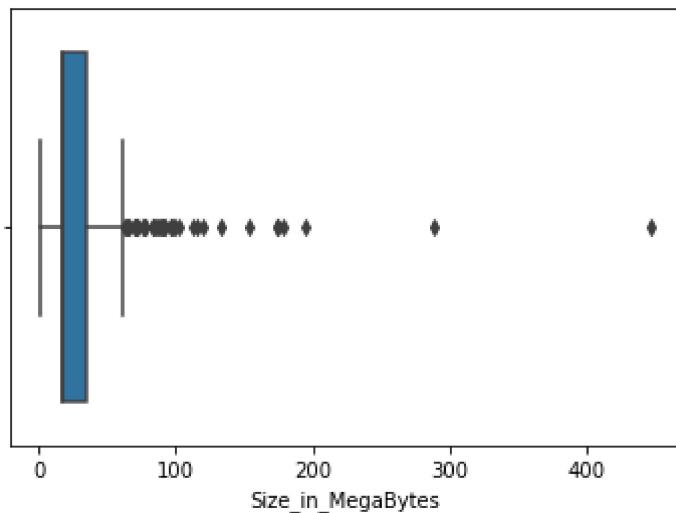
```
In [84]: plt.figure(figsize = (12,5))
```

```
plt.scatter(play['Category'],play['Size_in_MegaBytes'])
plt.xticks(rotation = 90)
plt.title('Category Vs Size_in_MegaBytes',size = 25 )
plt.xlabel('Category')
plt.ylabel('Size_in_MegaBytes')
plt.show()
# the size of Event app is largest
```



In [85]: `sns.boxplot(play["Size_in_MegaBytes"])  
plt.show()`

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.  
warnings.warn(
```



## Whova - Event & Conference App has largest size

In [86]: `play[play['App_Names']=='Whova - Event & Conference App']`

Out[86]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
186	Whova - Event & Conference App	5000000.0	4.8	Events	14805.0	447.0	2021-09-1

In [ ]:

In [87]: `play['Size_in_MegaBytes'].min()`

Out[87]: 1.2

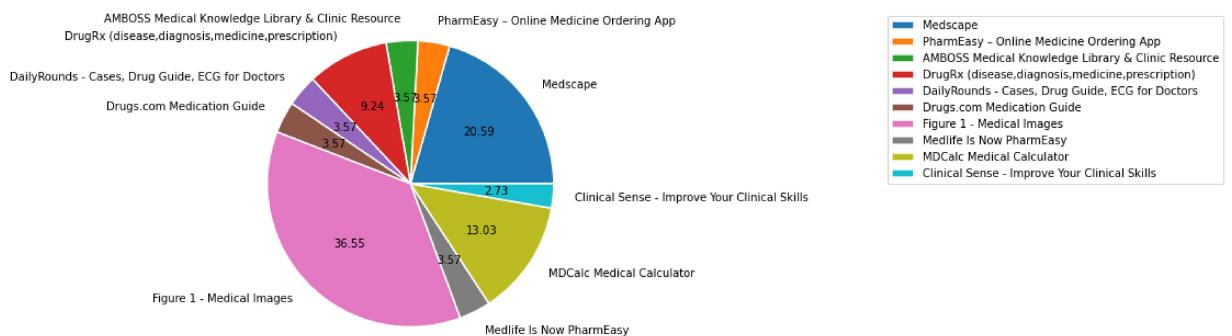
## Lucid - DayDream Screensaver of Daydream category has smallest size

In [88]: `play[play['Size_in_MegaBytes']==1.2]`

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
152	Lucid - DayDream Screensaver	1000000.0	3.9	Daydream	3243.0	1.2	2020-08-01

## Medical

```
In [89]: labels = ['Medscape', 'PharmEasy - Online Medicine Ordering App',  
              'AMBOSS Medical Knowledge Library & Clinic Resource',  
              'DrugRx (disease,diagnosis,medicine,prescription)',  
              'DailyRounds - Cases, Drug Guide, ECG for Doctors',  
              'Drugs.com Medication Guide', 'Figure 1 - Medical Images',  
              'Medlife Is Now PharmEasy', 'MDCalc Medical Calculator',  
              'Clinical Sense - Improve Your Clinical Skills']  
  
plt.pie(play[play['Category']=='Medical'].loc[236:245,'Size_in_MegaBytes'],labels=labels,wedgeprops = { 'linewidth' :1.5, 'edgecolor' : 'white' },radius = 1.5, autopct='%1.1f')  
plt.legend(labels, loc=(2.5,0.5))  
  
plt.show();
```



In [90]: `play[play['Category']=='Medical']['Size_in_MegaBytes'].min()`

Out[90]: 13.0

## Clinical Sense - Improve Your Clinical Skills is of smallest size in Medical Category

In [91]: `play[play['Size_in_MegaBytes']==13.0]`

Out[91]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
123	PhonePe Business	10000000.0	4.3	Business	218718.0	13.0	2021
178	JioCinema: Movies TV Originals	50000000.0	3.9	Entertainment	859918.0	13.0	2021
209	foodpanda: Fastest food delivery, amazing offers	10000000.0	3.2	Food & Drink	189695.0	13.0	2019
245	Clinical Sense - Improve Your Clinical Skills	1000000.0	4.6	Medical	11959.0	13.0	2020
273	Parenting tips, babycare, baby health & baby food	100000.0	4.4	Parenting	1120.0	13.0	2021
281	Launcher Plus One	1000000.0	4.5	Personalization	64650.0	13.0	2021

◀ ▶

In [92]: `play[play['Category']=='Medical']['Size_in_MegaBytes'].max()`

Out[92]: 174.0

## Figure 1 - Medical Images is of largest size in medical category

In [93]: `play[play['Size_in_MegaBytes']==174.0]`

Out[93]:

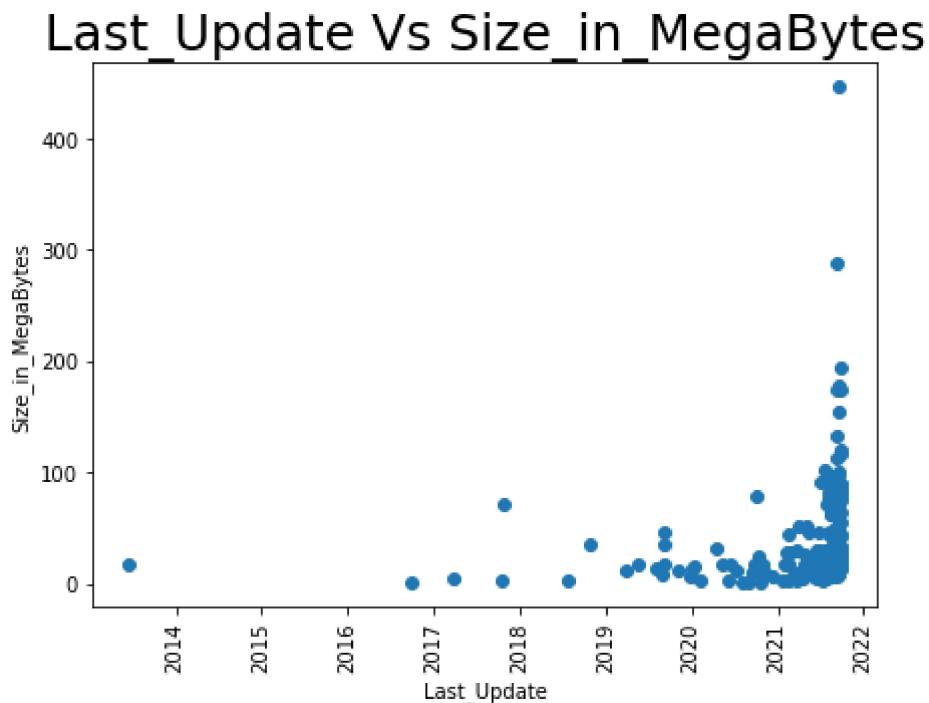
	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
242	Figure 1 - Medical Images	1000000.0	4.6	Medical	13536.0	174.0	2021-09-1
287	VN Video Editor Maker VlogNow	50000000.0	4.6	Video Players & Editors	957524.0	174.0	2021-09-2

◀ ▶

# With increase in updates the size of app got increased

```
In [94]: plt.figure(figsize = (7,5))

plt.scatter(play['Last_Update'],play['Size_in_MegaBytes'])
plt.xticks(rotation = 90)
plt.title('Last_Update Vs Size_in_MegaBytes',size = 25 )
plt.xlabel('Last_Update')
plt.ylabel('Size_in_MegaBytes')
plt.show()
# With increase in updates the size of app got increased
```



## Developer

In [95]: `play[play['Developer']=='Google LLC']`

Out[95]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
2	Google Fit: Activity Tracking	1.000000e+08	3.9	Health & Fitness	469685.0	17.0	2021-09-
11	Google Pay: Save, Pay, Manage	5.000000e+08	4.2	Finance	6743758.0	17.0	2021-09-
23	YouTube	1.000000e+10	4.3	Social	123129731.0	17.0	2021-09-
31	Google Maps	5.000000e+09	3.8	Travel & Local	14215173.0	17.0	2021-09-
111	Google Play Books & Audiobooks	1.000000e+09	4.1	Books & Reference	1845020.0	17.0	2021-09-
120	Google My Business	1.000000e+07	4.3	Business	261634.0	17.0	2021-09-
150	Daydream	1.000000e+07	3.4	Daydream	5090.0	35.0	2019-09-
151	Google VR Services	1.000000e+08	3.9	Daydream	42519.0	17.0	2019-09-
153	YouTube VR	1.000000e+07	3.7	Daydream	4127.0	24.0	2020-10-
155	Daydream Keyboard	1.000000e+07	3.9	Daydream	2347.0	46.0	2019-09-
228	Google Maps	5.000000e+09	3.8	Maps & Navigation	14215248.0	17.0	2021-09-
235	Google Maps Go	5.000000e+08	4.3	Maps & Navigation	179772.0	2.3	2021-02-
246	YouTube Music	1.000000e+09	4.3	Music & Audio	2973674.0	17.0	2021-09-

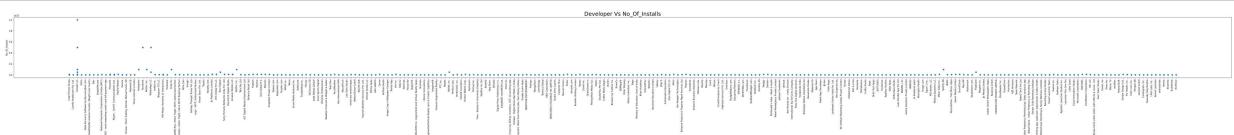
In [96]: `a = play[play['Developer']=='Google LLC']  
print(f"Number of Google Developed Apps are {len(a)}")`

Number of Google Developed Apps are 13

## The most developed app is of Google LLC

```
In [97]: plt.figure(figsize = (100,5))

plt.scatter(play['Developer'],play['No_Of_Installs'])
plt.xticks(rotation = 90)
plt.title('Developer Vs No_Of_Installs',size = 25 )
plt.xlabel('Developer')
plt.ylabel('No_Of_Installs')
plt.show()
# The most developed app is of Google LLC
```



## Last Update

```
In [98]: play['Last_Update'].max()

Out[98]: Timestamp('2021-09-27 00:00:00')

In [99]: play['Last_Update'].min()

Out[99]: Timestamp('2013-06-13 00:00:00')
```

## Daydream Launcher app was Updated far before

```
In [100]: play[play['Last_Update']==('2013-06-13 00:00:00')]

Out[100]:
```

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Upda
156	Daydream Launcher	100000.0	4.0	Daydream	408.0	17.0	2013-06-

## There are 13 Recently Updated Apps

```
In [101]: play[play['Last_Update']==('2021-09-23 00:00:00')]
```

```
# there are 13 most recently updated apps
```

Out[101]:

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
2	Google Fit: Activity Tracking	1.000000e+08	3.9	Health & Fitness	469685.0	17.0	2021-09-23 00:00:00
11	Google Pay: Save, Pay, Manage	5.000000e+08	4.2	Finance	6743758.0	17.0	2021-09-23 00:00:00
13	PhonePe: UPI, Recharge, Investment, Insurance	1.000000e+08	4.3	Finance	7132435.0	17.0	2021-09-23 00:00:00
15	PayPal Mobile Cash: Send and Request Money Fast	1.000000e+08	4.3	Finance	2174584.0	116.0	2021-09-23 00:00:00
17	Groww - Stocks, Demat, Mutual Fund, SIP	1.000000e+07	4.3	Finance	355781.0	17.0	2021-09-23 00:00:00
19	ETMONEY: Mutual Funds, SIP, Portfolio Tracking...	5.000000e+06	4.5	Finance	123002.0	27.0	2021-09-23 00:00:00
27	MX TakaTak Short Video App   Made in India for...	1.000000e+08	4.2	Social	760820.0	17.0	2021-09-23 00:00:00
41	ZEE5: Movies, TV Shows, Web Series, News	1.000000e+08	3.4	Entertainment	1147298.0	17.0	2021-09-23 00:00:00
45	Amazon Prime Video	1.000000e+08	4.3	Entertainment	3004249.0	17.0	2021-09-23 00:00:00
46	Netflix	1.000000e+09	4.3	Entertainment	12101864.0	17.0	2021-09-23 00:00:00
59	Purple: Beauty Shopping App. Buy Cosmetics Online...	1.000000e+07	4.5	Shopping	144828.0	17.0	2021-09-23 00:00:00
83	Canva: Graphic Design, Video Collage, Logo Maker	1.000000e+08	4.7	Art & Design	5971090.0	32.0	2021-09-23 00:00:00

	App_Names	No_Of_Installs	Rating	Category	No_Of_Review	Size_in_MegaBytes	Last_Updated
101	Purple: Beauty Shopping App. Buy Cosmetics On...	1.000000e+07	4.5	Beauty	144828.0	17.0	2021-1
148	Woo - The Dating App Women Love	1.000000e+07	4.1	Dating	107227.0	45.0	2021-1
161	DIKSHA - Platform for School Education	1.000000e+07	4.3	Education	353659.0	17.0	2021-1
173	ZEE5: Movies, TV Shows, Web Series, News	1.000000e+08	3.4	Entertainment	1147308.0	17.0	2021-1
203	EatSure - Order Food   Food court on an app	5.000000e+06	4.3	Food & Drink	131790.0	18.0	2021-1
224	Fabulous: Daily Routine Planner	1.000000e+07	4.6	Lifestyle	426087.0	90.0	2021-1
246	YouTube Music	1.000000e+09	4.3	Music & Audio	2973674.0	17.0	2021-1
253	SoundCloud - Play Music, Podcasts & New Songs	1.000000e+08	4.7	Music & Audio	5177074.0	83.0	2021-1
263	Vikatan News App: Magazine & Latest News Publi...	1.000000e+06	4.7	News & Magazines	26235.0	31.0	2021-1
287	VN Video Editor Maker VlogNow	5.000000e+07	4.6	Video Players & Editors	957524.0	174.0	2021-1
288	Tempo - Music Video Maker	1.000000e+07	4.4	Video Players & Editors	271644.0	78.0	2021-1
297	Calm - Meditate, Sleep, Relax	1.000000e+07	4.3	Watch apps	436097.0	17.0	2021-1

In [ ]: