

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

```
#READ DATA
```

```
df_data=pd.read_csv('https://raw.githubusercontent.com/DivyjaThakur24/GoogleAppRating-1')
```

```
df_data.head()
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	User Ver
--	-----	----------	--------	---------	------	----------	------	-------	----------------	--------	----------

0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10,000+	Free	0	Everyone	Art & Design	1
---	--	----------------	-----	-----	-----	---------	------	---	----------	--------------	---

1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free	0	Everyone	Design/Pretend Play	1
---	---------------------	----------------	-----	-----	-----	----------	------	---	----------	---------------------	---

2	Lite - FREE Live Cool Themes, Hide...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	1
---	---------------------------------------	----------------	-----	-------	------	------------	------	---	----------	--------------	---

3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	1
---	-----------------------	----------------	-----	--------	-----	-------------	------	---	------	--------------	---

4	Pixel Draw - Number Coloring Book	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Design/Creativity	1
---	-----------------------------------	----------------	-----	-----	------	----------	------	---	----------	-------------------------	---

```
df_data.head(5)
```

```
pandas.core.frame.DataFrame
```

```
df_data.shape()
```

```
df_data.head(13)
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	User Ver
--	-----	----------	--------	---------	------	----------	------	-------	----------------	--------	----------

0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10,000+	Free	0	Everyone	Art & Design	1
---	--	----------------	-----	-----	-----	---------	------	---	----------	--------------	---

1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free	0	Everyone	Design/Pretend Play	1
---	---------------------	----------------	-----	-----	-----	----------	------	---	----------	---------------------	---

2	Lite - FREE Live Cool Themes, Hide...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	1
---	---------------------------------------	----------------	-----	-------	------	------------	------	---	----------	--------------	---

3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	1
---	-----------------------	----------------	-----	--------	-----	-------------	------	---	------	--------------	---

4	Pixel Draw - Number Coloring Book	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Design/Creativity	1
---	-----------------------------------	----------------	-----	-----	------	----------	------	---	----------	-------------------------	---

5	Paper flowers instructions	ART_AND_DESIGN	4.4	167	5.6M	50,000+	Free	0	Everyone	Art & Design	1
---	----------------------------	----------------	-----	-----	------	---------	------	---	----------	--------------	---

6	Smoke Effect Maker - Smoke Editor	ART_AND_DESIGN	3.8	178	19M	50,000+	Free	0	Everyone	Art & Design	1
---	-----------------------------------	----------------	-----	-----	-----	---------	------	---	----------	--------------	---

7	Infinite Painter	ART_AND_DESIGN	4.1	36815	29M	1,000,000+	Free	0	Everyone	Art & Design	1
---	------------------	----------------	-----	-------	-----	------------	------	---	----------	--------------	---

8	Garden Coloring Book	ART_AND_DESIGN	4.4	13791	33M	1,000,000+	Free	0	Everyone	Art & Design	1
---	----------------------	----------------	-----	-------	-----	------------	------	---	----------	--------------	---

9	Kids Paint Free - Drawing Fun	ART_AND_DESIGN	4.7	121	3.1M	10,000+	Free	0	Everyone	Art & Design/Creativity	1
---	-------------------------------	----------------	-----	-----	------	---------	------	---	----------	-------------------------	---

10	Text on Photo - Fontee	ART_AND_DESIGN	4.4	13880	28M	1,000,000+	Free	0	Everyone	Art & Design	1
----	------------------------	----------------	-----	-------	-----	------------	------	---	----------	--------------	---

11	Name Art Photo Editor - Focus Filters	ART_AND_DESIGN	4.4	8788	12M	1,000,000+	Free	0	Everyone	Art & Design	1
----	---------------------------------------	----------------	-----	------	-----	------------	------	---	----------	--------------	---

12	Tattoo Name On My Photo Editor	ART_AND_DESIGN	4.2	44829	20M	10,000,000+	Free	0	Teen	Art & Design	1
----	--------------------------------	----------------	-----	-------	-----	-------------	------	---	------	--------------	---

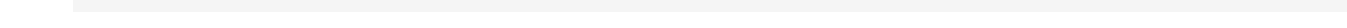
```
df_data.shape
```

```
df_data.describe()
```

	Rating
--	--------

count	936700000
mean	4.193338
std	0.537431
min	1.000000
25%	4.000000
50%	4.300000
75%	4.500000
max	19.000000

```
df_data.boxplot()
```



```
df_data.hist()
```



```
df_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 9367 non-null float64  
2 Reviews 10841 non-null object  
3 Size 10841 non-null object  
4 Installs 10841 non-null object  
5 Type 10841 non-null object  
6 Price 10841 non-null object  
7 Content Rating 10840 non-null object  
8 Genres 10841 non-null object  
9 Last Updated 10841 non-null object  
10 Current Ver 10833 non-null object  
11 Android Ver 10838 non-null object  
memory usage: 1.1+ MB

```
#DATA CLEANING
```

```
df_data.isnull()
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
--	-----	----------	--------	---------	------	----------	------	-------	----------------	--------	--------------	-------------	-------------

0	False	False	False	False	False	False	False	False	False	False	False	False	False
---	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

1	False	False	False	False	False	False	False	False	False	False	False	False	False
---	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

2	False	False	False	False	False	False	False	False	False	False	False	False	False
---	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

3	False	False	False	False	False	False	False	False	False	False	False	False	False
---	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

4	False	False	False	False	False	False	False	False	False	False	False	False	False
---	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

...	...	...	...	...	...	...	...	...	...	...	...	...	...
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

10836	False	False	False	False	False	False	False	False	False	False	False	False	False
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

10837	False	False	False	False	False	False	False	False	False	False	False	False	False
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

10838	False	False	True	False	False	False	False	False	False	False	False	False	False
-------	-------	-------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

10839	False	False	False	False	False	False	False	False	False	False	False	False	False
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

10840	False	False	False	False	False	False	False	False	False	False	False	False	False
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

10841 rows x 13 columns

```
df_data.isnull().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
dtype: int64	

```
df_data[df_data.Rating > 5]
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated
--	-----	----------	--------	---------	------	----------	------	-------	----------------	--------	--------------

10472	Life Made Wi-Fi Touchscreen Card	PRODUCTIVITY	1.9	190	3.0M	1,000+	Free	0	Everyone	NaN	February 11, 2018
-------	----------------------------------	--------------	-----	-----	------	--------	------	---	----------	-----	-------------------

```
df_data.drop([10472],inplace=True)
```

```
df_data[10450:10478]
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres
--	-----	----------	--------	---------	------	----------	------	-------	----------------	--------

10450	Police Field Interview Fi Card	PRODUCTIVITY	4.5	112	2.4M	1,000+	Paid	\$7.99	Everyone	Productivity
-------	--------------------------------	--------------	-----	-----	------	--------	------	--------	----------	--------------

10451	LG Hi-Fi Plus Manager	TOOLS	3.5	154	3.4M	10,000+	Free	0	Everyone	Tools
-------	-----------------------	-------	-----	-----	------	---------	------	---	----------	-------

10452	WiFi - Free Fast WiFi Connect & Find Wi-Fi Map	COMMUNICATION	4.0	13469	5.1M	1,000,000+	Free	0	Everyone	Communication
-------	--	---------------	-----	-------	------	------------	------	---	----------	---------------

10453	Talkie Pro - Wi-Fi Calling, Chats, File Sharing	COMMUNICATION	4.5	201	Varies with device	1,000+	Paid	\$2.99	Everyone	Communication
-------	---	---------------	-----	-----	--------------------	--------	------	--------	----------	---------------

10454	Safe Wi-Fi Pro	PRODUCTIVITY	2.9	32	9.7M	1,000+	Free	0	Everyone	Productivity
-------	----------------	--------------	-----	----	------	--------	------	---	----------	--------------

10455	Root! Pro - Wi-Fi Unlock Key	TOOLS	3.9	6736	72k	500,000+	Free	0	Everyone	Tools
-------	------------------------------	-------	-----	------	-----	----------	------	---	----------	-------

10456	Sat-Fi Monitor Pro - analyzer of Wi-Fi networks	TOOLS	4.6	85	2.4M	1,000+	Paid	\$2.99	Everyone	Tools
-------	---	-------	-----	----	------	--------	------	--------	----------	-------

10458	Candy simply-Fi	LIFESTYLE	2.1	2390	35M	100,000+	Free	0	Everyone	Lifestyle
-------	-----------------	-----------	-----	------	-----	----------	------	---	----------	-----------

10459	SCI-Fi UI	FAMILY	4.7	15	3.9M	100+	Paid	\$1.99	Everyone	Entertainment
-------	-----------	--------	-----	----	------	------	------	--------	----------	---------------

10460	Rabbit Wi-Fi settings shortcut	TOOLS	4.5	142	26k	5,000+	Paid	\$1.00	Everyone	Tools
-------	--------------------------------	-------	-----	-----	-----	--------	------	--------	----------	-------

10461	Wi-Fi Master	TOOLS	4.1	479	29k	50,000+	Free	0	Everyone	Tools
-------	--------------	-------	-----	-----	-----	---------	------	---	----------	-------

10462	Wi-Fi Micro Fi	FAMILY	4.5	3673	5.8M	500,000+	Free	0	Everyone	Tools
-------	----------------	--------	-----	------	------	----------	------	---	----------	-------

10463	Portable Wi-Fi hotspot Free-Fi	FAMILY	3.7	39	2.8M	5,000+	Free	0	Everyone	Education
-------	--------------------------------	--------	-----	----	------	--------	------	---	----------	-----------

10464	Portable Wi-Fi hotspot Free-Fi	FAMILY	2.9	691	5.8M	100,000+	Free	0	Everyone	Strategy
-------	--------------------------------	--------	-----	-----	------	----------	------	---	----------	----------

10465	Portable Wi-Fi hotspot Free-Fi	COMMUNICATION	4.0	1711	2.1M	100,000+	Free	0	Everyone	Communication
-------	--------------------------------	---------------	-----	------	------	----------	------	---	----------	---------------

10466	Portable Wi-Fi hotspot Free-Fi	PRODUCTIVITY	4.2	115	1.4M	10,000+	Free	0	Everyone	Productivity
-------	--------------------------------	--------------	-----	-----	------	---------	------	---	----------	--------------

10467	FI CFL Networks	FINANCE	3.7	112	3.9M	10,000+	Free	0	Everyone	Finance
-------	-----------------	---------	-----	-----	------	---------	------	---	----------	---------

10468	Tassa.fi Finland TownWiFi	LIFESTYLE	3.6	346	7.5M	50,000+	Free	0	Everyone	Lifestyle
-------	---------------------------	-----------	-----	-----	------	---------	------	---	----------	-----------

10469	Wi-Fi Everywhere	COMMUNICATION	3.9	2372	58M	500,000+	Free	0	Everyone	Communication
-------	------------------	---------------	-----	------	-----	----------	------	---	----------	---------------

10470	Jazz Wi-Fi	COMMUNICATION	3.4	49	4.0M	10,000+	Free	0	Everyone	Communication
-------	------------	---------------	-----	----	------	---------	------	---	----------	---------------

10471	Xposed Wi-Fi-Pwd osmino	PERSONALIZATION	3.5	1042	404k	100,000+	Free	0	Everyone	Personalization
-------	-------------------------	-----------------	-----	------	------	----------	------	---	----------	-----------------

10473	osmino Wi-Fi free	TOOLS	4.2	134203	4.1M	10,000,000+	Free	0	Everyone	Tools
-------	-------------------	-------	-----	--------	------	-------------	------	---	----------	-------

10474	Sat-Fi Voice	COMMUNICATION	3.4	37	14M	1,000+	Free	0	Everyone	Communication
-------	--------------	---------------	-----	----	-----	--------	------	---	----------	---------------

10475	Wi-Fi Visualizer	TOOLS	3.9	132	2.6M	50,000+	Free	0	Everyone	Tools
-------	------------------	-------	-----	-----	------	---------	------	---	----------	-------

10476	Lennox iCoo	LIFESTYLE	3.0	552	7.6M	50,000+	Free	0	Everyone	Lifestyle
-------	-------------	-----------	-----	-----	------	---------	------	---	----------	-----------

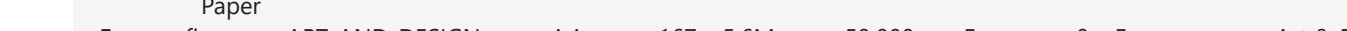
10477	Sci-Fi Sounds and Ringtones	PERSONALIZATION	3.6	128	11M	10,000+	Free	0	Everyone	Personalization
-------	-----------------------------	-----------------	-----	-----	-----	---------	------	---	----------	-----------------

10478	Sci-Fi Sounds	FAMILY	3.2	4	8.0M	1,000+	Free	0	Everyone	Entertainment
-------	---------------	--------	-----	---	------	--------	------	---	----------	---------------

```
df_data.boxplot()
```



```
df_data.hist()
```



```
#REMOVE COLUMNS THAT ARE 90% EMPTY
```

```
threshold = len(df_data) * 0.1
```

```
df_data.dropna(thresh=threshold, axis=1, inplace=True)
```

```
print(df_data.isnull().sum())
```

App	0
Category	0
Rating	1474
Reviews	0
Size	0
Installs	0
Type	1
Price	0
Content Rating	0
Genres	0
Last Updated	0
Current Ver	8
Android Ver	2
dtype: int64	

```
#DATA IMPUTATION AND MANIPULATION
```

```
#Numerical Values use median
```

```
def impute_median(series):  
    return series.fillna(series.median())
```

```
df_data.isnull().sum()
```

App	0
Category	0
Rating	1474
Reviews	0
Size	0
Installs	0
Type	1
Price	0
Content Rating	0
Genres	0
Last Updated	0
Current Ver	8
Android Ver	2
dtype: int64	

```
df_data.Rating = df_data['Rating'].transform(impute_median)
```

```
df_data.isnull().sum()
```

App	0
Category	0
Rating	1474
Reviews	0
Size	0
Installs	0
Type	1
Price	0
Content Rating	0
Genres	0
Last Updated	0
Current Ver	8
Android Ver	2
dtype: int64	

```
#modes of categorical values
```

```
print(df_data['Type'].mode())  
print(df_data['Current Ver'].mode())  
print(df_data['Android Ver'].mode())
```

0 Free  
dtype: object  
0 Varies with device  
dtype: object  
0 4.1 and up  
dtype: object

```
df_data['Type'].fillna(str(df_data['Type'].mode().values[0]),inplace=True)  
df_data['Current Ver'].fillna(str(df_data['Current Ver'].mode().values[0]),inplace=True)  
df_data['Android Ver'].fillna(str(df_data['Android Ver'].mode().values[0]),inplace=True)
```

```
df_data.isnull().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
dtype: object	

```
## Convert Price, Reviews, Ratings into Numerical Values
```

```
df_data['Price'] = df_data['Price'].apply(lambda x: str(x).replace('$','')) if '$' in str(x) else x  
df_data['Reviews'] = df_data['Reviews'].apply(lambda x: float(x)) if 'k' in str(x) else x  
df_data['Installs'] = df_data['Installs'].apply(lambda x: float(x)) if 'k' in str(x) else x
```

```
df_data['Installs'] = df_data['Installs'].apply(lambda x: str(x).replace('k','')) if 'k' in str(x) else x  
df_data['Installs'] = df_data['Installs'].apply(lambda x: str(x).replace('k','')) if 'k' in str(x) else x  
df_data['Installs'] = df_data['Installs'].apply(lambda x: float(x)) if 'k' in str(x) else x
```

```
df_data.head(10)
```

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres
--	-----	----------	--------	---------	------	----------	------	-------	----------------	--------

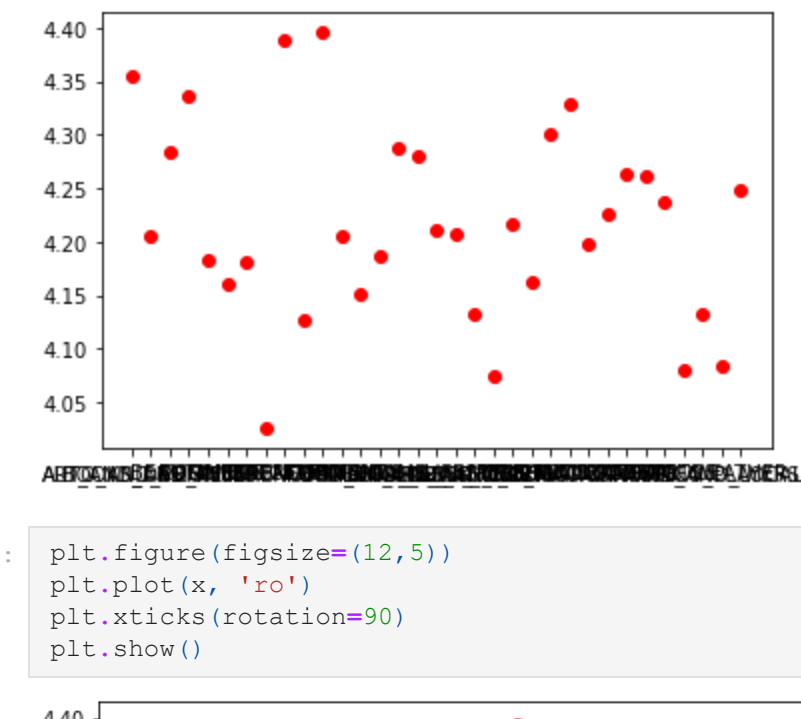
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10000.0	Free	0.0	Everyone	Art & Design
---	--	----------------	-----	-----	-----	---------	------	-----	----------	--------------

1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500000.0	Free	0.0	Everyone	Design/Pretend Play
---	---------------------	----------------	-----	-----	-----	----------	------	-----	----------	---------------------

2	Lite - FREE Live Cool Themes, Hide...	ART_AND_DESIGN	4.7	87510	8.7M	5000000.0	Free	0.0	Everyone	Art & Design
---	---------------------------------------	----------------	-----	-------	------	-----------	------	-----	----------	--------------

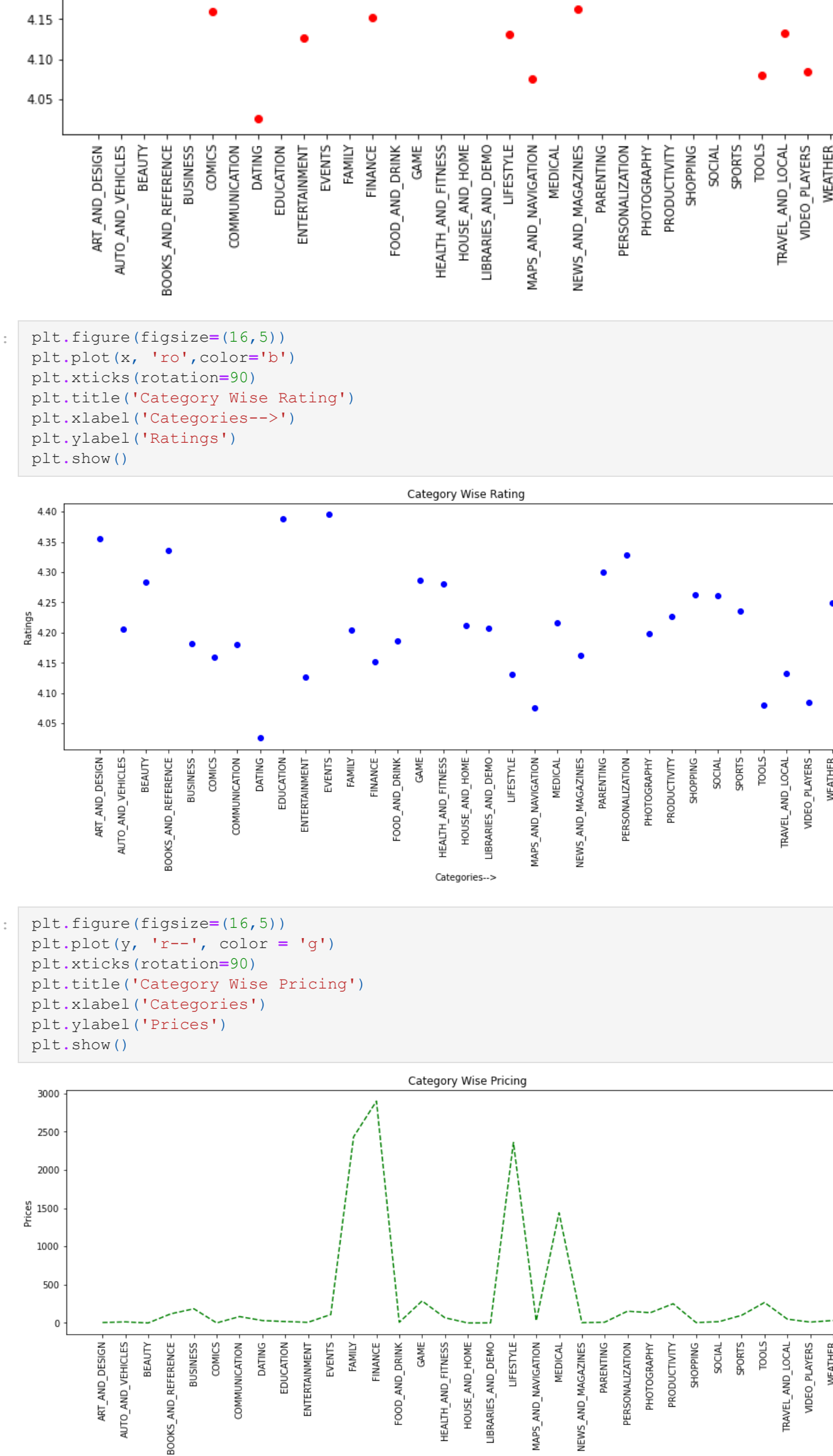
3	Sketch - Draw & Paint</
---	-------------------------





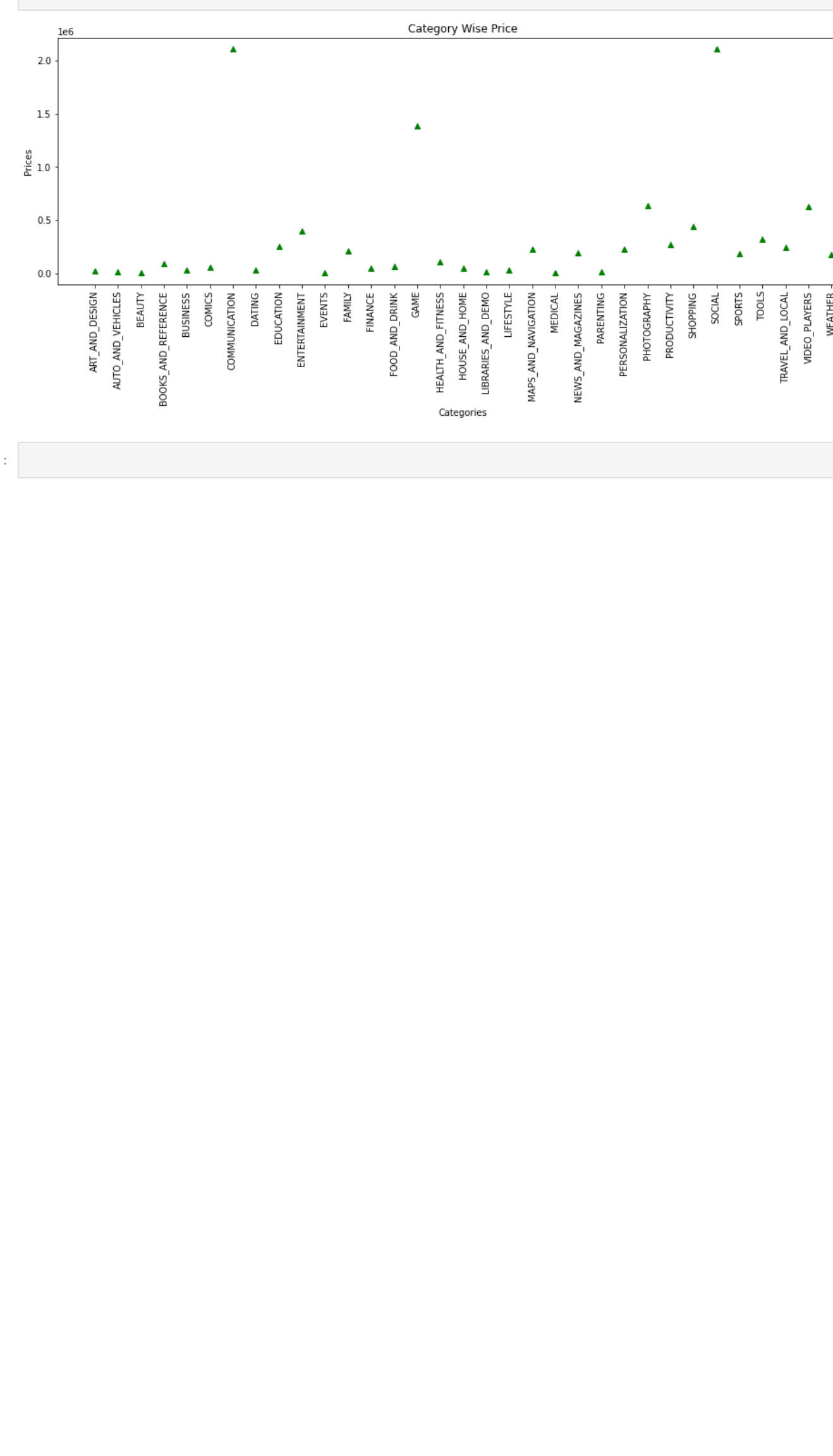
In [50]:

```
plt.figure(figsize=(12,5))
plt.plot(x, 'r-', color='r')
plt.xticks(rotation=90)
plt.show()
```



In [52]:

```
plt.figure(figsize=(16,5))
plt.plot(x, 'r-', color='b')
plt.xticks(rotation=90)
plt.title('Category Wise Rating')
plt.xlabel('Categories-->')
plt.ylabel('Ratings')
plt.show()
```



In [57]:

```
plt.figure(figsize=(16,5))
plt.plot(y, 'r-', color = 'g')
plt.xticks(rotation=90)
plt.title('Category Wise Pricing')
plt.xlabel('Categories')
plt.ylabel('Prices')
plt.show()
```



In [54]:

```
plt.figure(figsize=(16,5))
plt.plot(z, 'g', color = 'g')
plt.xticks(rotation=90)
plt.title('Category Wise Price')
plt.xlabel('Categories')
plt.ylabel('Prices')
plt.show()
```



In [ ]: