

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

## 1. Load the data file using pandas.

```
In [2]: data = pd.read_csv('googleplaystore.csv')
data.head()
```

```
Out[2]:
```

	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 moana	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	Pixol 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

## 2. Check for null values

```
In [3]: data.isnull().sum()
```

```
Out[3]:
```

App	0
Category	0
Rating	1474
Reviews	0
Size	0
Installs	0
Type	1
Price	0
Content Rating	1
Genres	8
Last Updated	8
Current Ver	8
Android Ver	3
dtype: int64	

## 3. Drop records with nulls

```
In [4]: data = data.dropna()
```

## 4. Variables seem to have incorrect type and inconsistent formatting.

```
In [5]: data = data[data['Size'].str.contains('Varies with device') == False]
data.reset_index()
```

```
Out[5]:
```

	index	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	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	1	Coloring book moana	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	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	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	4	Pixol 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

```
In [6]: data['Size'] = data['Size'].str.extract('(.*?)')
data['Size'] = data['Size'].str.extract('(.*?)')
data['Size'] = pd.to_numeric(data['Size'], errors='coerce')
data_K['Size'] = data_K['Size'].astype(int)
```

C:\Users\VFIR\AppData\Local\Temp\ipykernel\_14730\1724201669.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

C:\Users\VFIR\AppData\Local\Temp\ipykernel\_14730\1724201669.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

data\_K['Size'] = data\_K['Size'].astype(int)

```
Out[6]:
```

	index	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19000	10000+	Free	0	Everyone	Art & Design	January 7, 2018	1.0.0	4.0.3 and up
1	1	Coloring book moana	ART_AND_DESIGN	3.9	967	14000	500000+	Free	0	Everyone	Art & Design/Pretend Play	January 15, 2018	2.0.0	4.0.3 and up
2	2	U Launcher Lite - FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	8000	5000000+	Free	0	Everyone	Art & Design/Pretend Play	January 15, 2018	1.2.4	4.0.3 and up
3	3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25000	50,000,000+	Free	0	Teen	Art & Design	June 8, 2018	Varies with device	4.2 and up
4	4	Pixol Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967	2000	100,000+	Free	0	Everyone	Art & Design/Creativity	June 20, 2018	1.1	4.4 and up

```
In [7]: data_K['Size'] = data_K['Size'].str.extract('(.*?)')
data_K['Size'] = data_K['Size'].str.extract('(.*?)')
data_K['Size'] = pd.to_numeric(data_K['Size'], errors='coerce')
data_K
```

C:\Users\VFIR\AppData\Local\Temp\ipykernel\_14730\174997127.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

data\_K['Size'] = data\_K['Size'].str.extract('(.\*?)')
data\_K['Size'] = data\_K['Size'].str.extract('(.\*?)')
data\_K['Size'] = pd.to\_numeric(data\_K['Size'], errors='coerce')
data\_K

C:\Users\VFIR\AppData\Local\Temp\ipykernel\_14730\174997127.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

data\_K['Size'] = data\_K['Size'].astype(int)

```
Out[7]:
```

	index	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
58	58	Restart Navigator	AUTO_AND_VEHICLES	4.0	1403	201	100,000+	Free	0	Everyone	Auto & Vehicles	September 26, 2014	1.0.1	2.2 and up
209	209	Plagin'ADT v6.0	BUSINESS	3.1	4034	23	100,000+	Free	0	Everyone	Business	September 11, 2015	3.0.1.11 (Build 311)	2.2 and up
384	384	Hingsou Daler - Call Phons	COMMUNICATION	4.0	122498	79	10,000,000+	Free	0	Everyone	Communication	September 2, 2015	0.1.10944346	4.0.3 and up
495	495	Cake! +	COMMUNICATION	4.0	9498	118	1,000,000+	Free	0	Everyone	Communication	June 7, 2016	5.26.0	2.3 and up
468	468	GD iBook	COMMUNICATION	4.2	124346	695	10,000,000+	Free	0	Everyone	Communication	July 9, 2014	2.8	2.3 and up

```
In [8]: data_K['Size'] = data_K['Size'].str.extract('(.*?)')
data_K['Size'] = data_K['Size'].str.extract('(.*?)')
data_K['Size'] = pd.to_numeric(data_K['Size'], errors='coerce')
data_K
```

C:\Users\VFIR\AppData\Local\Temp\ipykernel\_14730\174997127.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

data\_K['Size'] = data\_K['Size'].str.extract('(.\*?)')
data\_K['Size'] = data\_K['Size'].str.extract('(.\*?)')
data\_K['Size'] = pd.to\_numeric(data\_K['Size'], errors='coerce')
data\_K

C:\Users\VFIR\AppData\Local\Temp\ipykernel\_14730\174997127.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

data\_K['Size'] = data\_K['Size'].astype(int)

```
Out[8]:
```

	index	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
10390	10390	News Mnenarrat	NEWS_AND_MAGAZINES	3.8	881	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.5	1.6 and up
10064	10064	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10065	10065	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10067	10067	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10068	10068	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10069	10069	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10070	10070	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10071	10071	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10072	10072	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10073	10073	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10074	10074	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10075	10075	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10076	10076	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10077	10077	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10078	10078	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10079	10079	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10080	10080	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10081	10081	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10082	10082	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10083	10083	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10084	10084	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10085	10085	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10086	10086	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10087	10087	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10088	10088	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10089	10089	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10090	10090	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10091	10091	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10092	10092	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10093	10093	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10094	10094	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10095	10095	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10096	10096	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10097	10097	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10098	10098	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10099	10099	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10100	10100	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10101	10101	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10102	10102	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10103	10103	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10104	10104	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10105	10105	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10106	10106	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10107	10107	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10108	10108	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10109	10109	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10110	10110	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10111	10111	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014	1.0.0	4.1 and up
10112	10112	News Mnenarrat	NEWS_AND_MAGAZINES	4.0	7	2000	100,000+	Free	0	Everyone	News & Magazines	January 20, 2014		