

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

```
In [2]: data=pd.read_excel(r"C:\Users\User\Downloads\googleplaystore.xlsx")
data
```

Out[2]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
	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 and up
	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 and up
	U Launcher Lite æ™ FREE Live Cool Themes, Hid...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	August 1, 2018	1.2.4	4.0 and up
	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	June 8, 2018	Varies with device	4.0 and up
	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	4.0 and up

	Sya9a Maroc - FR	FAMILY	4.5	38	53M	5,000+	Free	0	Everyone	Education	July 25, 2017	1.48	4.0 and up
	Fr. Mike Schmitz Audio Teachings	FAMILY	5.0	4	3.6M	100+	Free	0	Everyone	Education	July 6, 2018	1	4.0 and up
	Parkinson Exercises FR	MEDICAL	NaN	3	9.5M	1,000+	Free	0	Everyone	Medical	January 20, 2017	1	2.2 and up
	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
	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 × 13 columns

```
In [3]: data.head()#Display top 5 of the rows
```

Out[3]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
	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
	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
	U Launcher Lite æ™ FREE Live Cool Themes, Hid...	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
	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
	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

```
In [4]: data.tail()# display last 5 rows
```

Out[4]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
	Sya9a Maroc - FR	FAMILY	4.5	38	53M	5,000+	Free	0	Everyone	Education	July 25, 2017	1.48	4.1 and up
	Fr. Mike Schmitz Audio Teachings	FAMILY	5.0	4	3.6M	100+	Free	0	Everyone	Education	July 6, 2018	1	4.1 and up
	Parkinson Exercises FR	MEDICAL	NaN	3	9.5M	1,000+	Free	0	Everyone	Medical	January 20, 2017	1	2.2 and up
	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
	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

```
In [5]: data.shape# Display the shape of our database columns and rows
```

Out[5]: (10841, 13)

```
In [6]: data.describe()# get overall sattistics about the data frame
```

Out[6]:

	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

```
In [7]: data.columns
```

Out[7]: Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type', 'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver', 'Android Ver'], dtype='object')

```
In [8]: len(data[data['App'].str.contains('Astrology',na=False)])#Total number of App Titles contains Astrology
```

Out[8]: 3

```
In [9]: data['Rating'].mean()#Find Avarage App Rating
```

Out[9]: 4.193338315362448

```
In [10]: data['Category'].nunique()#Find total Unique Category
```

Out[10]: 34

```
In [11]: data.groupby('Category')['Rating'].mean().sort_values(ascending=False)# Whicce Category getting the heighest avarage Rating
```

Out[11]:

Category	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	

```
In [12]: len(data[data['Rating']==5.0])#Find Total number of apps having 5 star rating
```

Out[12]: 274

```
In [13]: data.columns
```

Out[13]: Index(['App', 'Category', 'Rating', 'Reviews', 'Size', 'Installs', 'Type', 'Price', 'Content Rating', 'Genres', 'Last Updated', 'Current Ver', 'Android Ver'], dtype='object')

```
In [14]: data['Reviews'].dtype#Find Avarage values of Reviews
```

Out[14]: dtype('O')

```
In [16]: data['Reviews'].astype('float')#Find Avarage values of Reviews
```

```
-----
ValueError                                Traceback (most recent call last)
Input In [16], in <cell line: 1>()
----> 1 data['Reviews'].astype('float')

File ~\anaconda3\lib\site-packages\pandas\core\generic.py:5912, in NDFrame.astype(self, dtype, copy, errors)
   5905         results = [
   5906             self.iloc[:, i].astype(dtype, copy=copy)
   5907             for i in range(len(self.columns))
   5908         ]
   5910     else:
   5911         # else, only a single dtype is given
-> 5912         new_data = self._mgr.astype(dtype=dtype, copy=copy, errors=errors)
   5913         return self._constructor(new_data).__finalize__(self, method="astype")
   5915 # GH 33113: handle empty frame or series

File ~\anaconda3\lib\site-packages\pandas\core\internals\managers.py:419, in BaseBlockManager.astype(self, dtype, copy, errors)
   418     def astype(self, T, dtype, copy: bool = False, errors: str = "raise") -> T:
-> 419         return self.apply("astype", dtype=dtype, copy=copy, errors=errors)

File ~\anaconda3\lib\site-packages\pandas\core\internals\managers.py:304, in BaseBlockManager.apply(self, f, align_keys, ignore_failures, **kwargs)
   302         applied = b.apply(f, **kwargs)
-> 304         applied = getattr(b, f)(**kwargs)
   305     except (TypeError, NotImplementedError):
   306         if not ignore_failures:

File ~\anaconda3\lib\site-packages\pandas\core\internals\blocks.py:580, in Block.astype(self, dtype, copy, errors)
   562     """
   563     Coerce to the new dtype.
   564     (...)
   576     Block
   577     """
-> 580     new_values = astype_array_safe(values, dtype, copy=copy, errors=errors)
   582     new_values = maybe_coerce_values(new_values)
   583     newb = self.make_block(new_values)

File ~\anaconda3\lib\site-packages\pandas\core\dtypes\cast.py:1292, in astype_array_safe(values, dtype, copy, errors)
   1289         dtype = dtype.numpy_dtype
   1291     try:
-> 1292         new_values = astype_array(values, dtype, copy=copy)
   1293     except (ValueError, TypeError):
   1294         # e.g. astype_nansafe can fail on object-dtype of strings
   1295         # trying to convert to float
   1296         if errors == "ignore":

File ~\anaconda3\lib\site-packages\pandas\core\dtypes\cast.py:1237, in astype_array(values, dtype, copy)
   1234     values = values.astype(dtype, copy=copy)
   1236     else:
-> 1237         values = astype_nansafe(values, dtype, copy=copy)
   1239     # in pandas we don't store numpy str dtypes, so convert to object
   1240     if isinstance(dtype, np.dtype) and issubclass(values.dtype.type, str):

File ~\anaconda3\lib\site-packages\pandas\core\dtypes\cast.py:1181, in astype_nansafe(arr, dtype, copy, skipna)
   1177         raise ValueError(msg)
   1179     if copy or is_object_dtype(arr.dtype) or is_object_dtype(dtype):
   1180         # Explicit copy, or required since NumPy can't view from / to object.
-> 1181         return arr.astype(dtype, copy=True)
   1183     return arr

ValueError: could not convert string to float: '3.0M'
```

```
In [17]: data[data['Reviews']=='3.0M']
```

Out[17]:

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

```
In [18]: data['Reviews']=data['Reviews'].replace('3.0M',3.0)#Find Avarage values of Reviews
```

```
In [19]: data['Reviews']=data['Reviews'].astype('float')#Find Avarage values of Reviews
```

```
In [20]: data['Reviews'].dtype#Find Avarage values of Reviews
```

Out[20]: dtype('float64')

```
In [21]: data['Reviews'].mean()#Find Avarage values of Reviews
```

Out[21]: 444111.9265750392

```
In [22]: data['Type'].value_counts()#Find total Number of Paid and Free Apps
```

Out[22]: Free 18039
Paid 8000
0 1
Name: Type, dtype: int64

```
In [23]: data[data['Reviews'].max()==data['Reviews']]#Which App Has the Maximum Reviews?
```

Out[23]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
	2544 Facebook	SOCIAL	4.1	78158306.0	Varies with device	1,000,000,000+	Free	0	Teen	Social	August 3, 2018	Varies with device	Varies with device

```
In [27]: index=data['Reviews'].sort_values(ascending=False).head().index#Display top 5 Apps Having Highest Reviews?
```

```
In [25]: data.iloc[index]['App']#Display top 5 Apps Having Highest Reviews?
```

Out[25]: 2544 Facebook
3943 Facebook
381 WhatsApp Messenger
336 WhatsApp Messenger
3904 WhatsApp Messenger
Name: App, dtype: object

```
In [26]: data.groupby('Type')['Rating'].mean()#Find Avarage Rating of Free and Paid Apps?
```

Out[26]: Type
0 19.000000
Free 4.186203
Paid 4.266615
Name: Rating, dtype: float64