

importing numpy and pandas

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

reading & describing data

```
In [2]: kc=pd.read_csv("/home/placement/Downloads/movies.csv")  
kc.describe()
```

Out[2]:

	srno	year	rating	time
count	49590.000000	49590.000000	10814.000000	45836.000000
mean	24795.500000	2002.303428	3.451248	2628.445436
std	14315.544261	12.534555	0.495601	1604.646265
min	1.000000	1913.000000	1.400000	52.000000
25%	12398.250000	1999.000000	3.100000	1356.000000
50%	24795.500000	2007.000000	3.500000	2563.000000
75%	37192.750000	2010.000000	3.800000	2877.000000
max	49590.000000	2014.000000	4.500000	28813.000000

```
In [3]: kc.head(10)
```

```
Out[3]:
```

	srno	movie	year	rating	time
0	1	The Nightmare Before	1993	3.9	4568.0
1	2	The Mummy	1932	3.5	4388.0
2	3	Orphans of the Storm	1921	3.2	9062.0
3	4	The Object of Beauty	1991	2.8	6150.0
4	5	Night Tide	1963	2.8	5126.0
5	6	One Magic Christmas	1985	3.8	5333.0
6	7	Muriel's Wedding	1994	3.5	6323.0
7	8	Mother's Boys	1994	3.4	5733.0
8	9	Nosferatu: Original Version	1929	3.5	5651.0
9	10	Nick of Time	1995	3.4	5333.0

checking nulls

```
In [4]: kc.isna().sum()
```

```
Out[4]: srno      0
movie      0
year       0
rating    38776
time      3754
dtype: int64
```

removing movie column

```
In [5]: a=kc.drop("movie",axis=1)  
a
```

Out[5]:

	srno	year	rating	time
0	1	1993	3.9	4568.0
1	2	1932	3.5	4388.0
2	3	1921	3.2	9062.0
3	4	1991	2.8	6150.0
4	5	1963	2.8	5126.0
...
49585	49586	2013	2.8	1812.0
49586	49587	2013	NaN	6822.0
49587	49588	2010	NaN	3610.0
49588	49589	2010	2.7	NaN
49589	49590	2010	2.7	NaN

49590 rows × 4 columns

filling zeros at empty spaces

```
In [6]: b=kc.fillna(kc.median)
b
```

Out[6]:

	srno	movie	year	rating	time
0	1	The Nightmare Before	1993	3.9	4568.0
1	2	The Mummy	1932	3.5	4388.0
2	3	Orphans of the Storm	1921	3.2	9062.0
3	4	The Object of Beauty	1991	2.8	6150.0
4	5	Night Tide	1963	2.8	5126.0
...
49585	49586	Winter Wonderland	2013	2.8	1812.0
49586	49587	Top Gear: Series 19: Africa Special	2013	<bound method NDFrame._add_numeric_operations....	6822.0
49587	49588	Fireplace For Your Home: Crackling Fireplace w...	2010	<bound method NDFrame._add_numeric_operations....	3610.0
49588	49589	Kate Plus Ei8ht	2010	2.7	<bound method NDFrame._add_numeric_operations....
49589	49590	Kate Plus Ei8ht: Season 1	2010	2.7	<bound method NDFrame._add_numeric_operations....

49590 rows × 5 columns

installing & importing seaborn

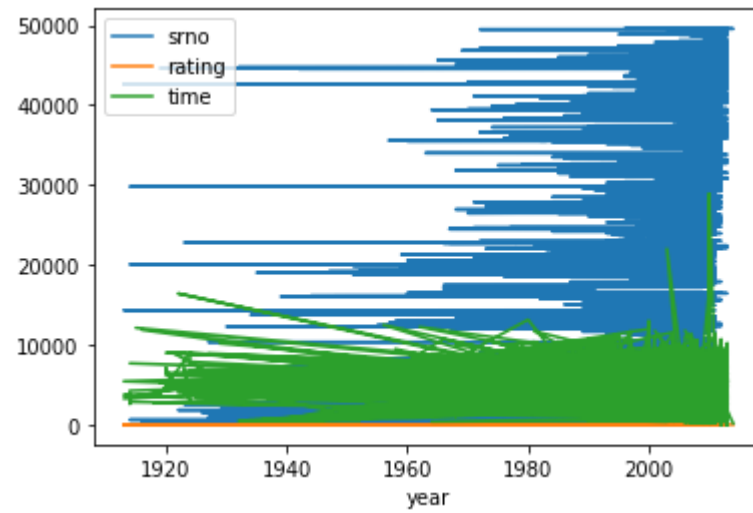
```
In [7]: !pip3 install seaborn
import seaborn as sb
```

```
Requirement already satisfied: seaborn in ./local/lib/python3.8/site-packages (0.12.2)
Requirement already satisfied: matplotlib!=3.6.1,>=3.1 in ./local/lib/python3.8/site-packages (from seaborn) (3.7.1)
Requirement already satisfied: pandas>=0.25 in ./local/lib/python3.8/site-packages (from seaborn) (2.0.2)
Requirement already satisfied: numpy!=1.24.0,>=1.17 in ./local/lib/python3.8/site-packages (from seaborn) (1.24.3)
Requirement already satisfied: cycler>=0.10 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (0.11.0)
Requirement already satisfied: importlib-resources>=3.2.0; python_version < "3.10" in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (5.12.0)
Requirement already satisfied: kiwisolver>=1.0.1 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (1.4.4)
Requirement already satisfied: contourpy>=1.0.1 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (1.1.0)
Requirement already satisfied: pillow>=6.2.0 in /usr/lib/python3/dist-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (7.0.0)
Requirement already satisfied: packaging>=20.0 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (23.1)
Requirement already satisfied: python-dateutil>=2.7 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (2.8.2)
Requirement already satisfied: pyparsing>=2.3.1 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (3.0.9)
Requirement already satisfied: fonttools>=4.22.0 in ./local/lib/python3.8/site-packages (from matplotlib!=3.6.1,>=3.1->seaborn) (4.40.0)
Requirement already satisfied: pytz>=2020.1 in ./local/lib/python3.8/site-packages (from pandas>=0.25->seaborn) (2023.3)
Requirement already satisfied: tzdata>=2022.1 in ./local/lib/python3.8/site-packages (from pandas>=0.25->seaborn) (2023.3)
Requirement already satisfied: zipp>=3.1.0; python_version < "3.10" in ./local/lib/python3.8/site-packages (from importlib-resources>=3.2.0; python_version < "3.10"->matplotlib!=3.6.1,>=3.1->seaborn) (3.15.0)
Requirement already satisfied: six>=1.5 in /usr/lib/python3/dist-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.1->seaborn) (1.14.0)
```

plotting graph to year

```
In [8]: kc.plot("year")
```

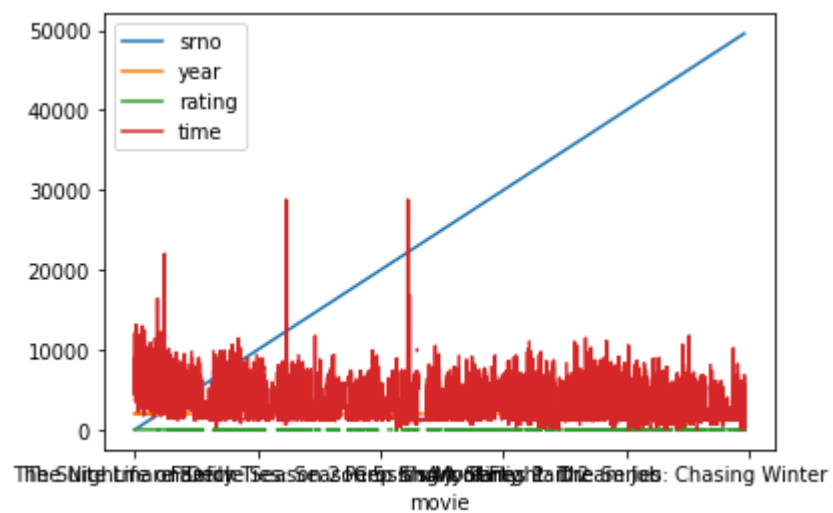
```
Out[8]: <Axes: xlabel='year'>
```



plotting graph to movie

```
In [9]: kc.plot("movie")
```

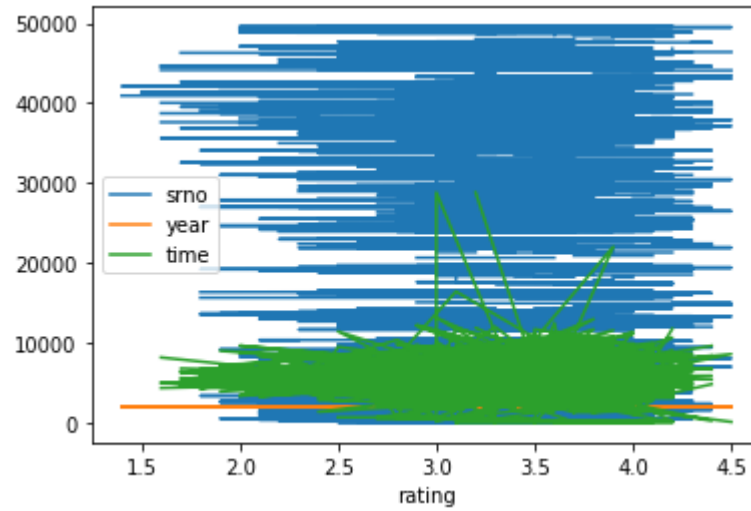
```
Out[9]: <Axes: xlabel='movie'>
```



plotting graph to rating

```
In [10]: kc.plot("rating")
```

```
Out[10]: <Axes: xlabel='rating'>
```



installing and importing matplotlib

```
In [11]: !pip3 install matplotlib
import matplotlib as mp
```

```
Requirement already satisfied: matplotlib in ./local/lib/python3.8/site-packages (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in ./local/lib/python3.8/site-packages (from matplotlib) (1.1.0)
Requirement already satisfied: packaging>=20.0 in ./local/lib/python3.8/site-packages (from matplotlib) (23.1)
Requirement already satisfied: pyparsing>=2.3.1 in ./local/lib/python3.8/site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in ./local/lib/python3.8/site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: importlib-resources>=3.2.0; python_version < "3.10" in ./local/lib/python3.8/site-packages (from matplotlib) (5.12.0)
Requirement already satisfied: cyclor>=0.10 in ./local/lib/python3.8/site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in ./local/lib/python3.8/site-packages (from matplotlib) (4.40.0)
Requirement already satisfied: kiwisolver>=1.0.1 in ./local/lib/python3.8/site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: pillow>=6.2.0 in /usr/lib/python3/dist-packages (from matplotlib) (7.0.0)
Requirement already satisfied: numpy>=1.20 in ./local/lib/python3.8/site-packages (from matplotlib) (1.24.3)
Requirement already satisfied: six>=1.5 in /usr/lib/python3/dist-packages (from python-dateutil>=2.7->matplotlib) (1.14.0)
Requirement already satisfied: zipp>=3.1.0; python_version < "3.10" in ./local/lib/python3.8/site-packages (from importlib-resources>=3.2.0; python_version < "3.10"->matplotlib) (3.15.0)
```

removing movie & year from table

```
In [12]: a=kc.drop(["movie","year"],axis=1)  
a
```

Out[12]:

	srno	rating	time
0	1	3.9	4568.0
1	2	3.5	4388.0
2	3	3.2	9062.0
3	4	2.8	6150.0
4	5	2.8	5126.0
...
49585	49586	2.8	1812.0
49586	49587	NaN	6822.0
49587	49588	NaN	3610.0
49588	49589	2.7	NaN
49589	49590	2.7	NaN

49590 rows × 3 columns

corrilating the values in the table

```
In [13]: b=a.corr()  
b  
b.describe()
```

Out[13]:

	srno	rating	time
count	3.000000	3.000000	3.000000
mean	0.256080	0.356524	0.243079
std	0.666348	0.557607	0.672605
min	-0.286048	0.015285	-0.286048
25%	-0.115880	0.034786	-0.135381
50%	0.054288	0.054288	0.015285
75%	0.527144	0.527144	0.507642
max	1.000000	1.000000	1.000000

checking the table

In [14]: `kc.loc[kc.time>5000]`

Out[14]:

	srno	movie	year	rating	time	
	2	3	Orphans of the Storm	1921	3.2	9062.0
	3	4	The Object of Beauty	1991	2.8	6150.0
	4	5	Night Tide	1963	2.8	5126.0
	5	6	One Magic Christmas	1985	3.8	5333.0
	6	7	Muriel's Wedding	1994	3.5	6323.0

49564	49565	American Addict	2013	3.5	5377.0	
49579	49580	Underground: The Julian Assange Story	2012	3.7	5665.0	
49583	49584	Sunset Strip	2012	3.0	5770.0	
49584	49585	Silver Bells	2013	3.5	5287.0	
49586	49587	Top Gear: Series 19: Africa Special	2013	NaN	6822.0	

5897 rows × 5 columns

In [15]:

```
kc.loc[(kc.year==2000) & (kc.time>5000)]
```

Out[15]:

	srno	movie	year	rating	time
409	410	Believe	2000	3.3	5767.0
416	417	The Prophecy 3: The Ascent	2000	3.4	5048.0
430	431	Scream 3	2000	3.2	7013.0
432	433	Holy Smoke	2000	3.0	6855.0
437	438	Requiem for a Dream	2000	3.9	6087.0
...
32557	32558	Shaded Places	2000	2.9	5350.0
36229	36230	The Three Stooges	2000	3.7	5256.0
37333	37334	Les Miserables: Pt. 2	2000	NaN	5170.0
37336	37337	Les Miserables: Pt. 1	2000	NaN	5194.0
39493	39494	The Prophet's Game	2000	3.2	6486.0

137 rows × 5 columns

making the table in order on the basis of time

```
In [16]: kc.sort_values("time")
```

Out[16]:

	srno	movie	year	rating	time
40150	40151	Trailer: Pain	2012	3.6	52.0
41081	41082	Trailer: Get to Work	2012	3.3	55.0
41082	41083	Trailer: Give and Take	2012	3.3	66.0
43166	43167	Trailer: Emperor	2013	3.1	67.0
43330	43331	Trailer: Blood Angel	2013	4.2	69.0
...
49556	49557	Shinobi Girl	2012	2.0	NaN
49561	49562	My Hope America with Billy Graham	2013	3.9	NaN
49565	49566	My Hope America with Billy Graham	2013	3.9	NaN
49588	49589	Kate Plus Ei8ht	2010	2.7	NaN
49589	49590	Kate Plus Ei8ht: Season 1	2010	2.7	NaN

49590 rows × 5 columns

```
In [17]: kc.loc[(kc.rating>=4) & (kc.year>=1993) & (kc.year <=2010)]
```

Out[17]:

	srno	movie	year	rating	time
117	118	Fresh	1994	4.0	6829.0
138	139	Pulp Fiction	1994	4.1	9265.0
248	249	Hercules	1997	4.0	5572.0
287	288	Life Is Beautiful	1997	4.2	6973.0
302	303	Mulan: Special Edition	1998	4.2	5270.0
...
48784	48785	King Tut Unwrapped	2010	4.0	NaN
48854	48855	Pit Bulls & Parolees	2009	4.3	NaN
48874	48875	Brew Masters	2010	4.1	NaN
49025	49026	Cake Boss: Next Great Baker	2010	4.1	NaN
49028	49029	Police Women of Dallas	2010	4.0	NaN

802 rows × 5 columns

In []: