In [2]:
```
import pandas as pd
```

In [3]:
```
data= pd.read_csv('/home/placement/Desktop/movies.csv')   ## Reading the data of csv file
```

In [4]:
```
print(data)
```

```
        srno                                              movie  year  rating  \
0          1                          The Nightmare Before  1993     3.9
1          2                                    The Mummy  1932     3.5
2          3                          Orphans of the Storm  1921     3.2
3          4                          The Object of Beauty  1991     2.8
4          5                                    Night Tide  1963     2.8
...      ...                                           ...   ...     ...
49585  49586                               Winter Wonderland  2013     2.8
49586  49587           Top Gear: Series 19: Africa Special  2013     NaN
49587  49588  Fireplace For Your Home: Crackling Fireplace w...  2010     NaN
49588  49589                                 Kate Plus Ei8ht  2010     2.7
49589  49590                      Kate Plus Ei8ht: Season 1  2010     2.7

         time
0      4568.0
1      4388.0
2      9062.0
3      6150.0
4      5126.0
...       ...
49585  1812.0
49586  6822.0
49587  3610.0
49588     NaN
49589     NaN

[49590 rows x 5 columns]
```

In [5]:
```
data.describe()     ## DEscribing the data
```

Out[5]:

	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 [6]:
```
data.columns     ## Columns of the data
```

Out[6]:
```
Index(['srno', 'movie', 'year', 'rating', 'time'], dtype='object')
```

In [7]:
```
data.head(5)   ## Head Top Data will we get
```

Out[7]:

	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

In [8]:
```
data.head(10)
```

Out[8]:

	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

In [9]:
```
data.tail(10)     ## Tail Down Data will we get
```

Out[9]:

	srno	movie	year	rating	time
49580	49581	Curious George: A Very Monkey Christmas	2009	3.8	3438.0
49581	49582	Mumfie's White Christmas	1996	2.4	1350.0
49582	49583	Lady Gaga & The Muppets' Holiday Spectacular	2013	3.1	3496.0
49583	49584	Sunset Strip	2012	3.0	5770.0
49584	49585	Silver Bells	2013	3.5	5287.0
49585	49586	Winter Wonderland	2013	2.8	1812.0
49586	49587	Top Gear: Series 19: Africa Special	2013	NaN	6822.0
49587	49588	Fireplace For Your Home: Crackling Fireplace w...	2010	NaN	3610.0
49588	49589	Kate Plus Ei8ht	2010	2.7	NaN
49589	49590	Kate Plus Ei8ht: Season 1	2010	2.7	NaN

In [10]:
```
data1=data[data['time']>5000]       ## seperating the data that are greater than 5000 in time
data1
```

Out[10]:

	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 [11]:
```
data2=data.loc[(data.time>5000)]
data2
```

Out[11]:

	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 [12]:
```
data3=data.loc[(data.year==2000)&(data.time>5000)]      ## seperating the data that are greater than 5000 in time and yera =2000
data3
```

Out[12]:

	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

In [13]:
```
data.head(100)
```

Out[13]:

	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
...	...	...	...	...	...
95	96	The Hunted	1995	3.4	6605.0
96	97	The Great Waldo Pepper	1975	3.5	6467.0
97	98	Godzilla: King of the Monsters	1956	3.5	4828.0
98	99	Highlander 2: Renegade Version	1991	3.1	6585.0
99	100	High Noon	1952	3.9	5087.0

100 rows × 5 columns

In [34]:
```
data.tail(5)
```

Out[34]:

	srno	movie	year	rating	time
49585	49586	Winter Wonderland	2013	2.8	1812.0
49586	49587	Top Gear: Series 19: Africa Special	2013	NaN	6822.0
49587	49588	Fireplace For Your Home: Crackling Fireplace w...	2010	NaN	3610.0
49588	49589	Kate Plus Ei8ht	2010	2.7	NaN
49589	49590	Kate Plus Ei8ht: Season 1	2010	2.7	NaN

In [15]:
```
data4=data.loc[(data.year>=2000)&(data.year<=2010)&(data.rating>3.0)]
data4
```

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
431	432	The Tigger Movie	2000	3.5	4626.0
437	438	Requiem for a Dream	2000	3.9	6087.0
...	...	...	...	...	...
49300	49301	Signing Time	2007	3.8	NaN
49320	49321	Dave Gorman Stand-Up. Live.	2010	3.4	5090.0
49334	49335	Signing Time: Collection 1	2007	3.8	NaN
49394	49395	A Liga	2010	3.8	NaN
49580	49581	Curious George: A Very Monkey Christmas	2009	3.8	3438.0

4263 rows × 5 columns

In [16]:
```
dataq=data.sort_values('time')   ## sorting the data with time
```

In [17]:
```
dataq
```

Out[17]:

	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 [18]:
```
dataa=data.sort_values('rating')   ## sorting the data with rating
```

In [33]:
```
dataa.head(5)
```

Out[33]:

	srno	movie	year	rating	time
40934	40935	Lagegi	2007	1.4	NaN
42115	42116	Sun Yaar Chill Maar	2007	1.4	NaN
40826	40827	Lagegi	2007	1.4	NaN
42160	42161	Sun Yaar Chill Maar	2007	1.4	NaN
41396	41397	Meri Toh Lag Gayi Naukri	2011	1.5	NaN

In [20]:
```
data['rating'].max()   ## max value
```

Out[20]:
```
4.5
```

In [21]:
```
data['rating'].min()   # min value
```

Out[21]: 1.4

In [22]:
```
data7=data.drop(columns=['movie'])   ## removing the data using drop function
```

In [23]:
```
data7
```

Out[23]:

	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

In [24]:
```
data7.corr()       ## correlation to the data
```

Out[24]:

	srno	year	rating	time
srno	1.000000	0.401153	0.054288	-0.286048
year	0.401153	1.000000	0.156210	-0.058444
rating	0.054288	0.156210	1.000000	0.015285
time	-0.286048	-0.058444	0.015285	1.000000

In [32]:
```
data.sample(10)  ## sampling the data
```

Out[32]:

	srno	movie	year	rating	time
43542	43543	Dexter's Laboratory: Season 1: Double Trouble ...	1996	NaN	1345.0
36785	36786	I Predator: Season 1: Crocodile vs. Wildebeest	2011	NaN	2544.0
28061	28062	Wild Recon: Season 1	2010	3.8	NaN
6241	6242	The Incredible Hulk: Season 2: Vendetta Road	1978	NaN	2894.0
39308	39309	Locked Up Abroad: Season 2: Lima	2008	NaN	2759.0
27794	27795	Frasier: Season 10: Some Assembly Required	2003	NaN	1297.0
22588	22589	Star Trek: The Next Generation: Season 6: Lessons	1992	NaN	2707.0
16021	16022	Twin Peaks	1990	3.9	NaN
33545	33546	The Best Adventures of Sonic the Hedgehog: Sub...	1993	NaN	1326.0
47625	47626	Pit Bulls & Parolees: Season 3: Prison Break	2011	NaN	2608.0

In [27]:
```
data8=data.groupby(['year']).count()   ## grouping the data having year
data8
```

Out[27]:

year	srno	movie	rating	time
1913	3	3	3	3
1914	20	20	5	18
1915	1	1	1	1
1916	1	1	1	1
1918	1	1	1	1
...	...	...	...	...
2010	5107	5107	1102	4671
2011	5511	5511	1346	4992
2012	4339	4339	1130	3978
2013	981	981	345	901
2014	1	1	1	1

101 rows × 4 columns

In [30]:
```
data8.to_csv('file.csv')      ## plotting the graph to the data
data8.plot()
```

Out[30]: <Axes: xlabel='year'>