

```
In [75]: import pandas as pd
```

```
In [76]: data=pd.read_csv("/home/placement/Downloads/movies.csv")
```

```
In [77]: data.describe()
```

```
Out[77]:
```

	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 [78]: data.head()
```

```
Out[78]:
```

	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 [79]: data.isna().sum()
```

```
Out[79]: srno          0  
movie          0  
year          0  
rating    38776  
time       3754  
dtype: int64
```

```
In [80]: data.head(5)
```

```
Out[80]:
```

	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 [81]: data2=data.loc[(data.time>5000)&(data.year==2000)]  
data2
```

```
Out[81]:
```

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

137 rows × 5 columns

```
In [82]: data.head(100)
```

```
Out[82]:
```

	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 [83]: `data.head(50)`

Out[83]:

	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
10	11	Broken Blossoms	1919	3.3	5367.0
11	12	Big Night	1996	3.6	6561.0
12	13	The Birth of a Nation	1915	2.9	12118.0
13	14	The Boys from Brazil	1978	3.6	7417.0
14	15	Big Doll House	1971	2.9	5696.0
15	16	The Breakfast Club	1985	4.0	5823.0
16	17	The Bride of Frankenstein	1935	3.7	4485.0
17	18	Beautiful Girls	1996	3.5	6755.0
18	19	Bustin' Loose	1981	3.7	5598.0
19	20	The Beguiled	1971	3.4	6307.0
20	21	Born on the Fourth of July	1989	3.4	8646.0
21	22	Broadcast News	1987	3.4	7940.0
22	23	Swimming with Sharks	1994	3.3	5586.0

	srno	movie	year	rating	time
23	24	Beavis and Butt-head Do America	1996	3.4	4852.0
24	25	Brighton Beach Memoirs	1986	3.4	6564.0
25	26	The Best of Times	1986	3.4	6247.0
26	27	Brassed Off	1996	3.5	6040.0
27	28	Last Tango in Paris	1972	3.1	7732.0
28	29	Leprechaun 2	1994	3.2	5125.0
29	30	Incident at Oglala: The Leonard Peltier Story	1992	3.7	5487.0
30	31	Kalifornia	1993	3.4	7095.0
31	32	The Lady Vanishes	1938	3.7	5762.0
32	33	Jingle All the Way	1996	3.6	5371.0
33	34	Killing Zoe	1993	3.4	5773.0
34	35	King of Beggars	1992	3.6	6025.0
35	36	Into the Woods	1990	4.0	9077.0
36	37	Joe Kidd	1972	3.7	5252.0
37	38	In Too Deep	1999	3.9	5823.0
38	39	King Kong	1976	3.2	8044.0
39	40	Internal Affairs	1990	3.5	6885.0
40	41	Jesus Christ Superstar	1973	3.6	6388.0
41	42	In the Name of the Father	1993	3.9	7972.0
42	43	Easy Money	1987	2.8	5794.0
43	44	Do the Right Thing	1989	3.6	7186.0
44	45	Days of Heaven	1978	3.4	5628.0
45	46	Drop Zone	1994	3.4	6087.0
46	47	Escape from L.A.	1996	3.3	6039.0
47	48	Emma	1996	3.5	7260.0



```
In [84]: data.tail(50)
```

```
Out[84]:
```

	srno	movie	year	rating	time
49540	49541	Life With Boys: Season 1: This Time the Proble...	2011	NaN	1368.0
49541	49542	Life With Boys: Season 1: In the Principal's O...	2011	NaN	1368.0
49542	49543	El Fuente: 5994 MP	2013	3.5	471.0
49543	49544	Life With Boys: Season 1: Monkey Talk with Boys	2011	NaN	1368.0
49544	49545	El Fuente: 50 MP	2013	3.3	465.0
49545	49546	Bo Burnham: what.	2013	4.1	3614.0
49546	49547	Life With Boys: Season 1: A Perfect Life with ...	2011	NaN	1368.0
49547	49548	Life With Boys: Season 1: Wrestling with Boys	2011	NaN	1369.0
49548	49549	Life With Boys: Season 1	2011	4.1	NaN
49549	49550	El Fuente: 30 MP	2013	3.3	471.0
49550	49551	El Fuente: 2997 MP	2013	3.2	471.0
49551	49552	El Fuente: 25 MP	2013	3.3	465.0
49552	49553	El Fuente: 24 MP	2013	3.5	484.0
49553	49554	Max Steel	2013	4.1	NaN
49554	49555	Saving Santa	2013	3.8	5038.0
49555	49556	Lilyhammer: Season 1 (Recap)	2013	4.2	194.0
49556	49557	Shinobi Girl	2012	2.0	NaN
49557	49558	GLOW: The Story of the Gorgeous Ladies of Wres...	2012	3.7	4593.0
49558	49559	Mitt (Trailer)	2013	3.0	138.0
49559	49560	My Hope America with Billy Graham: The Cross	2013	NaN	1706.0
49560	49561	The Square (Trailer)	2014	3.6	154.0
49561	49562	My Hope America with Billy Graham	2013	3.9	NaN
49562	49563	My Hope America with Billy Graham: Defining Mo...	2013	NaN	1791.0



	srno	movie	year	rating	time
49563	49564	My Hope America with Billy Graham: Lose to Gain	2013	NaN	1400.0
49564	49565	American Addict	2013	3.5	5377.0
49565	49566	My Hope America with Billy Graham	2013	3.9	NaN
49566	49567	El Fuente: 2997 MP10	2013	2.7	470.0
49567	49568	El Fuente: 25 MP10	2013	2.9	464.0
49568	49569	El Fuente: 24 MP10	2013	2.8	484.0
49569	49570	El Fuente: 23976 MP10	2013	2.9	484.0
49570	49571	The Short Game (Trailer)	2013	4.1	156.0
49571	49572	El Fuente: 5994 MP10	2013	2.8	471.0
49572	49573	El Fuente: 50 MP10	2013	2.9	464.0
49573	49574	El Fuente: 30 MP10	2013	2.8	470.0
49574	49575	Greg Fitzsimmons: Life on Stage	2013	3.3	3671.0
49575	49576	Dave Foley: Relatively Well	2013	3.2	3446.0
49576	49577	Barbie: Life in the Dreamhouse: Barbie Life in...	2013	NaN	1390.0
49577	49578	Barbie: Life in the Dreamhouse: Barbie Life in...	2013	NaN	1458.0
49578	49579	Transformers Prime Beast Hunters: Predacons Ri...	2013	4.2	3950.0
49579	49580	Underground: The Julian Assange Story	2012	3.7	5665.0
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 &#38; 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

```
In [85]: data1=data.loc[(data.rating==4)&(data.year<=2010)&(data.year>2000)]
```

```
In [86]: data1
```

```
Out[86]:
```

	srno	movie	year	rating	time
816	817	Dil Chahta Hai	2001	4.0	11110.0
865	866	Black Hawk Down	2001	4.0	8660.0
920	921	We Were Soldiers	2002	4.0	8309.0
1329	1330	Farscape: Season 3	2001	4.0	NaN
1362	1363	The Forsyte Saga: Series 1	2002	4.0	NaN
...	...	...	...	...	...
47718	47719	Killer Whales	2010	4.0	2620.0
47793	47794	King Tut Unwrapped: Season 1	2010	4.0	NaN
48066	48067	Police Women of Dallas: Season 1	2010	4.0	NaN
48784	48785	King Tut Unwrapped	2010	4.0	NaN
49028	49029	Police Women of Dallas	2010	4.0	NaN

278 rows × 5 columns

```
In [87]: data1=data.sort_values('time')
```

In [88]:

```
data1
```

Out[88]:

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

49590 rows × 5 columns

In [89]:

```
data1=data.sort_values('rating')
```

In [90]:

```
data1
```

Out[90]:

	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
...	...	...	...	...	...
49563	49564	My Hope America with Billy Graham: Lose to Gain	2013	NaN	1400.0
49576	49577	Barbie: Life in the Dreamhouse: Barbie Life in...	2013	NaN	1390.0
49577	49578	Barbie: Life in the Dreamhouse: Barbie Life in...	2013	NaN	1458.0
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

49590 rows × 5 columns

In [91]:

```
data.isna().sum()
```

Out[91]:

srno	0
movie	0
year	0
rating	38776
time	3754
dtype:	int64

In [92]:

```
data1=data.fillna(data.median())
data1
```

/tmp/ipykernel\_6220/1028561032.py:1: FutureWarning: The default value of numeric\_only in DataFrame.median is deprecated. In a future version, it will default to False. In addition, specifying 'numeric\_only=None' is deprecated. Select only valid columns or specify the value of numeric\_only to silence this warning.

```
data1=data.fillna(data.median())
```

Out[92]:

	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	3.5	6822.0
49587	49588	Fireplace For Your Home: Crackling Fireplace w...	2010	3.5	3610.0
49588	49589	Kate Plus Ei8ht	2010	2.7	2563.0
49589	49590	Kate Plus Ei8ht: Season 1	2010	2.7	2563.0

49590 rows × 5 columns

In [93]: data.shape

Out[93]: (49590, 5)

In [94]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 49590 entries, 0 to 49589
Data columns (total 5 columns):
#   Column  Non-Null Count  Dtype
---  -
0   srno    49590 non-null    int64
1   movie   49590 non-null    object
2   year    49590 non-null    int64
3   rating  10814 non-null    float64
4   time    45836 non-null    float64
dtypes: float64(2), int64(2), object(1)
memory usage: 1.9+ MB
```

In [95]:

```
show Lines in check type of columns
```

Cell In[95], line 1

```
show Lines in check type of columns
```

^

SyntaxError: invalid syntax

In [96]:

```
data2=data.groupby(['year']).count()
```

In [97]:

```
data2
```

Out[97]:

	srno	movie	rating	time
year				
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 [98]:

```
data2.to_csv("Downloads.csv")
```

In [99]:

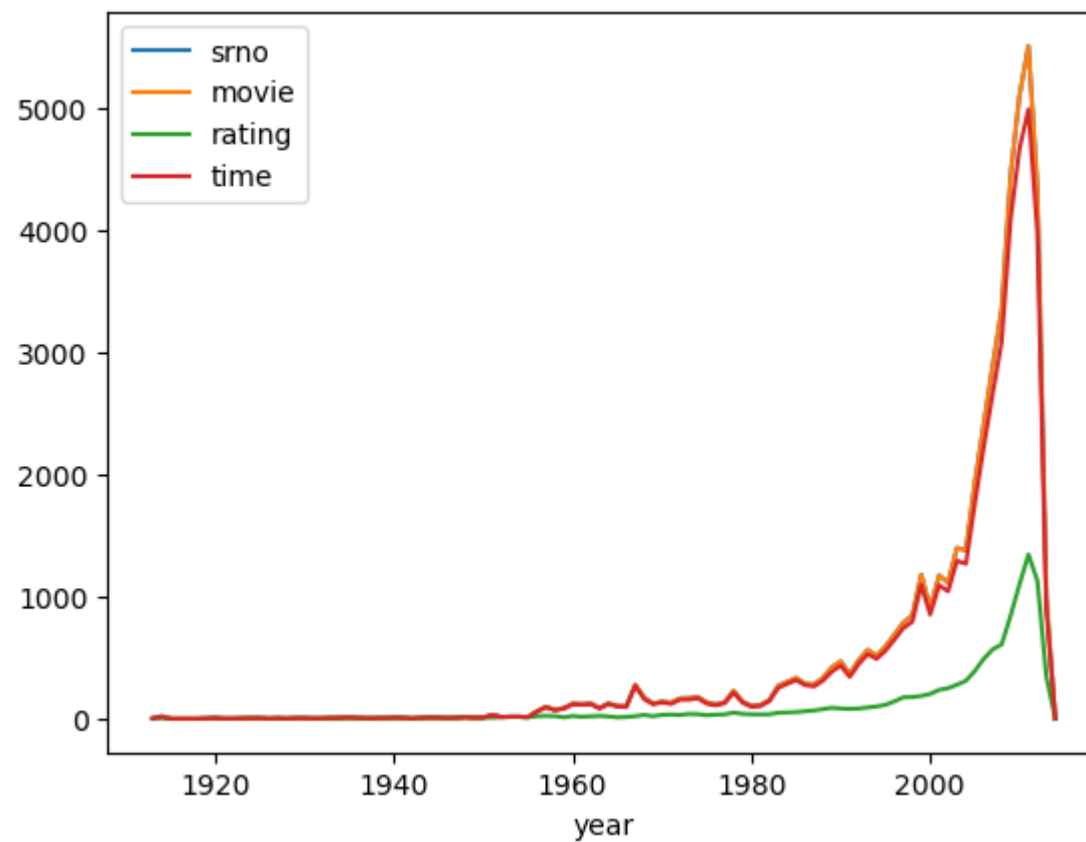
```
data.isna().sum()
```

Out[99]:

srno	0
movie	0
year	0
rating	38776
time	3754
dtype:	int64

```
In [100]: data2.plot()
```

```
Out[100]: <Axes: xlabel='year'>
```



```
In [ ]:
```