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

reading a file

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

describe data

In [3]:
```
pavan.describe()
```

Out[3]:

	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 [4]:
```
pavan.head()
```

Out[4]:

	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 [5]:
```
pavan.tail()
```

Out[5]:

	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 [6]:
```
pavan.tail(10)
```

Out[6]:

	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 [7]:
```
pavan.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
```

unique values

In [8]:
```
pavan['movie'].unique()
```

Out[8]:
```
array(['The Nightmare Before', 'The Mummy', 'Orphans of the Storm', ...,
       'Fireplace For Your Home: Crackling Fireplace with Music',
       'Kate Plus Ei8ht', 'Kate Plus Ei8ht: Season 1'], dtype=object)
```

loc keyword using

In [9]:
```
k=pavan.loc[(pavan.time>5000)&(pavan.year==2000)]
k
```

Out[9]:

	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 [10]:
```
k.count()
```

Out[10]:
```
srno      137
movie     137
year      137
rating    133
time      137
dtype: int64
```

In [20]:
```
k.head()
```

Out[20]:

	srno	year	time
srno	1.000000	0.401153	-0.286048
year	0.401153	1.000000	-0.058444
time	-0.286048	-0.058444	1.000000

In [12]:
```
g=pavan.loc[pavan.movie=='Quills']
g
```

Out[12]:

	srno	movie	year	rating	time
558	559	Quills	2000	3.5	7444.0

In [13]:
```
u=pavan.loc[(pavan.rating>=4)&(pavan.year>=2000)&(pavan.year<=2010)]
u
```

Out[13]:

	srno	movie	year	rating	time
484	485	Coming to Light: Edward S. Curtis and the Nort...	2000	4.0	5027.0
560	561	The Emperor's New Groove	2000	4.0	4703.0
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
...	...	...	...	...	...
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

725 rows × 5 columns

In [14]:
```
u.count()
```

Out[14]:
```
srno      725
movie     725
year      725
rating    725
time      147
dtype: int64
```

In [15]:
```
k=pavan.sort_values('rating')
k.head(15)
```

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
41083	41084	Meri Toh Lag Gayi Naukri	2011	1.5	NaN
35597	35598	1313: UFO Invasion	2011	1.6	5053.0
40035	40036	Champ	2008	1.6	NaN
38729	38730	The Tulse Luper Suitcases - The TV Series	2004	1.6	NaN
41350	41351	Sid vs. Varun	2009	1.6	NaN
44595	44596	Where Is Parsifal?	1984	1.6	4937.0
37613	37614	1313: Hercules Unbound	2012	1.6	4339.0
40183	40184	Champ	2008	1.6	NaN
44154	44155	The Tulse Luper Suitcases - The TV Series	2004	1.6	NaN
44656	44657	Makers Our Story	2011	1.6	8162.0

eliminating columns

In [16]:
```
s=pavan.drop(['movie','rating'],axis=1)
s
```

Out[16]:

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

49590 rows × 3 columns

correlation

In [17]:
```
k=s.corr()
k
```

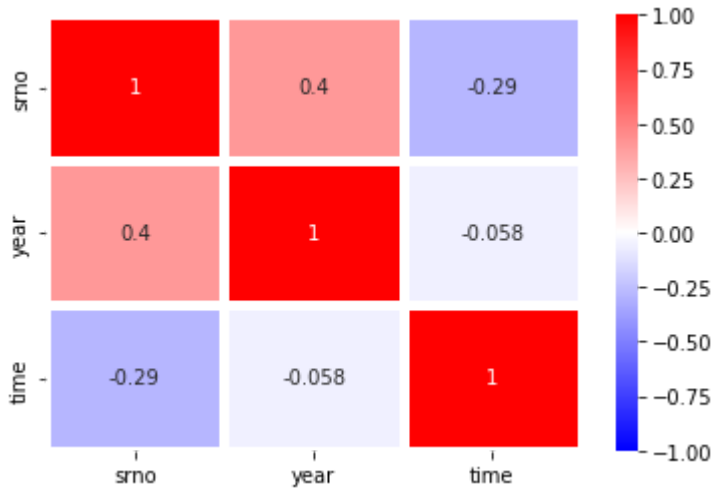
Out[17]:

	srno	year	time
srno	1.000000	0.401153	-0.286048
year	0.401153	1.000000	-0.058444
time	-0.286048	-0.058444	1.000000

correlation plot

In [19]:
```
sns.heatmap(k,vmax=1,vmin=-1,annot=True,linewidth=5,cmap='bwr')
```

Out[19]: <Axes: >



In [ ]: