```
In [2]:
import pandas as pd
In [3]:
movie=pd.read_csv('movies1.csv')
In [4]:
ratings=pd.read csv('ratings.csv')
In [5]:
df=pd.merge(movie, ratings).drop(['genres', 'timestamp'], axis=1)
In [6]:
df.head()
Out[6]:
                          userld rating
    movield
                      title
 0
         1 Toy Story (1995)
         1 Toy Story (1995)
                               5
                                     4.0
 1
 2
         1 Toy Story (1995)
                                     4.5
 3
         1 Toy Story (1995)
                               15
                                     2.5
         1 Toy Story (1995)
                                     4.5
In [7]:
userrating=df.pivot table(index=['userId'],columns=['title'],values='rating')
In [8]:
userrating.head()
Out[8]:
               'Hellboy':
                                              'Til
                                                  'Tis the
                                                                            (500)
                                                                                   *batteries
                                                                                                 Zulu [REC] [REC]<sup>2</sup> [REC]<sup>3</sup> 3
                          'Round 'Salem's
                                           There
                                                          'burbs,
                                                                   'night
                    The
                                                  Season
                                                                          Days of
                                                                                       not
                                                             The Mother
   title
               Seeds of
                         Midnight
                                      Lot
                                            Was
                                                     for
        (2014)
                                                                                    included
                                                                                                (2013) (2007) (2009)
                                                                         Summer
                                    (2004)
                                                           (1989)
                                                                                                                       (2012)
                                             You
                Creation
                           (1986)
                                                    Love
                                                                  (1986)
                                                                            (2009)
                                                                                     (1987)
                  (2004)
                                           (1997)
                                                   (2015)
 userld
          NaN
                   NaN
                             NaN
                                     NaN
                                             NaN
                                                     NaN
                                                            NaN
                                                                    NaN
                                                                             NaN
                                                                                       NaN ...
                                                                                                        NaN
                                                                                                               NaN
                                                                                                                        NaN
                                                                                                 NaN
     2
          NaN
                   NaN
                             NaN
                                      NaN
                                             NaN
                                                     NaN
                                                            NaN
                                                                    NaN
                                                                             NaN
                                                                                       NaN
                                                                                                 NaN
                                                                                                        NaN
                                                                                                               NaN
                                                                                                                        NaN
     3
          NaN
                   NaN
                             NaN
                                     NaN
                                             NaN
                                                     NaN
                                                            NaN
                                                                    NaN
                                                                             NaN
                                                                                       NaN
                                                                                                 NaN
                                                                                                        NaN
                                                                                                               NaN
                                                                                                                        NaN
     4
          NaN
                   NaN
                             NaN
                                     NaN
                                             NaN
                                                     NaN
                                                            NaN
                                                                    NaN
                                                                             NaN
                                                                                       NaN
                                                                                                 NaN
                                                                                                        NaN
                                                                                                               NaN
                                                                                                                        NaN
     5
          NaN
                   NaN
                             NaN
                                                                                                                        NaN
                                      NaN
                                             NaN
                                                     NaN
                                                            NaN
                                                                    NaN
                                                                             NaN
                                                                                       NaN
                                                                                                 NaN
                                                                                                        NaN
                                                                                                                NaN
5 rows × 9719 columns
4
                                                                                                                          F
In [9]:
userrating=userrating.dropna\,(thresh=10\,\text{,axis}=1)\,\text{.fillna}\,(0)
```

In [10]:

userrating.head()

Out[10]:

	title	'burbs, The (1989)	(500) Days of Summer (2009)	10 Cloverfield Lane (2016)	Things I Hate About You (1999)	10,000 BC (2008)	101 Dalmatians (1996)	Dalmatians (One Hundred and One Dalmatians) (1961)	12 Angry Men (1957)	12 Years a Slave (2013)	127 Hours (2010)	 Zack and Miri Make a Porno (2008)	Zero Dark Thirty (2012)	Zero Effect (1998)	Zodia: (2007
us	serld														
	1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0
	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0
	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0
	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	 0.0	0.0	0.0	0.0
	5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0

5 rows × 2269 columns

In [11]:

itemsimilaritydf=userrating.corr(method='pearson')

In [12]:

itemsimilaritydf.head(15)

Out[12]:

title	'burbs, The (1989)	(500) Days of Summer (2009)	10 Cloverfield Lane (2016)	10 Things I Hate About You (1999)	10,000 BC (2008)	101 Dalmatians (1996)	101 Dalmatians (One Hundred and One Dalmatians) (1961)	12 Angry Men (1957)	12 Years a Slave (2013)	127 Hours (2010)	 Zack and Miri Make a Porno (2008)	
title												_
'burbs, The (1989)	1.000000	0.063117	-0.023768	0.143482	0.011998	0.087931	0.224052	0.034223	0.009277	0.008331	 0.017477	(
(500) Days of Summer (2009)	0.063117	1.000000	0.142471	0.273989	0.193960	0.148903	0.142141	0.159756	0.135486	0.200135	 0.374515	(
10 Cloverfield Lane (2016)	0.023768	0.142471	1.000000	0.005799	0.112396	0.006139	-0.016835	0.031704	0.024275	0.272943	 0.242663	(
10 Things I Hate About You (1999)	0.143482	0.273989	-0.005799	1.000000	0.244670	0.223481	0.211473	0.011784	0.091964	0.043383	 0.243118	(
10,000 BC (2008)	0.011998	0.193960	0.112396	0.244670	1.000000	0.234459	0.119132	0.059187	0.025882	0.089328	 0.260261	(
101 Dalmatians (1996)	0.087931	0.148903	0.006139	0.223481	0.234459	1.000000	0.285112	0.119843	0.072399	0.029967	 0.114968	(
101 Dalmatians (One Hundred and One Dalmatians) (1961)	0.224052	0.142141	-0.016835	0.211473	0.119132	0.285112	1.000000	0.134037	0.017264	- 0.046277	 0.120302	(
12 Angry Men (1957)	0.034223	0.159756	0.031704	0.011784	0.059187	0.119843	0.134037	1.000000	0.132979	0.058862	 0.104518	(
12 Years a Slave	0.009277	0.135486	-0.024275	0.091964	- 0 025882	0.072399	0.017264	0.132979	1.000000	0.249931	 0.024045	(

```
101
                                           10
                                                                  Dalmatians
  127 Hours
                   0.20(500)
Days of
                                                                            0.058862 9249934 1.000929 ...
Angry a Slave Hours
                            0.272943
Cloverfield
           0.0000000
                                                         0.029967
                                                                    -0.0462772
                                                                                                         0.2231165
               The
                                                   ВС
                                                       Dalmatians
                                                                    Hundred
                                                                                      a Slave
                                                                                                Hours
                                                                                                          Make a
                    Summer
                                Lane
                                        About
                                                                               Men
             (1989)
                                                 (2008)
                                                                    and One
                                                                                       (2013)
                                                                                                (2010)
   13 Going
                                                           (1996)
           0.049700 0.297152
                                                                               (1957)
                                              0.167098
                                                                  Dalmatians)
                                                                                                        0.1639992) (
                                                         0.188467
                                                                                     0.092133 0.043314 ...
     on 30
                                                                             0.027672
                                                                      (1961)
     (2004)
      tista
    Warrior,
           0.176071 0.072835
                             -0.029983 0.215828 0.163148
                                                         0.004213
                                                                    0.135894 0.084760 0.040699 0.005800 ... 0.104176
  The (1999)
      1408
           0.131993 0.226574
                             0.181245 0.069470 0.133706
                                                         0.159777
                                                                    0.227193  0.189497  0.227912  0.226178  ...  0.255607  (
     (2007)
  15 Minutes
           0.104500 0.087240
                             0.017315 0.223197 0.190482
                                                         0.023682
                                                                    (2001)
   17 Again
           0.047271 0.342948
                            -0.019197 0.231261 0.174499
                                                         0.064464
                                                                    0.121410  0.049762  0.177934  0.054016  ...  0.263390  (
     (2009)
15 rows × 2269 columns
4
In [13]:
def getsimilarmovies(moviename, userrating):
     similarscore=itemsimilaritydf[moviename] * (userrating-2.5)
     similarscore=similarscore.sort values(ascending=False)
     return similarscore
In [14]:
actionlover=[('10 Things I Hate About You (1999)',5),('10,000 BC (2008)',3),('Zombieland (2009)',5
) ]
In [15]:
similarmovies=pd.DataFrame()
In [16]:
for movie, rating in actionlover:
     similarmovies=similarmovies.append(getsimilarmovies(movie,rating),ignore index=True)
In [17]:
similarmovies.head()
similarmovies.sum().sort values(ascending=False)
Out[17]:
10 Things I Hate About You (1999)
                                           3.018927
Zombieland (2009)
                                           3.017741
Scott Pilgrim vs. the World (2010)
                                           2.098950
Pineapple Express (2008)
                                           2.059345
Bring It On (2000)
                                           2.018599
                                          -0.286697
Drop Zone (1994)
Specialist, The (1994)
                                          -0.290607
Clear and Present Danger (1994)
                                          -0.302525
Madness of King George, The (1994)
                                          -0.309482
Disclosure (1994)
                                          -0.317898
Length: 2269, dtype: float64
In [ ]:
```

0.020002

(2013)