

```
In [1]: import pandas as pd
columns = ['user_id', 'item_id', 'rating', 'timestamp']
df = pd.read_csv("C:/Users/Admin/Desktop/data11.txt", sep='\t', names=columns)
```

```
In [2]: df.head(10)
```

Out[2]:

	user_id	item_id	rating	timestamp
0	0	50	5	881250949
1	0	172	5	881250949
2	0	133	1	881250949
3	196	242	3	881250949
4	186	302	3	891717742
5	22	377	1	878887116
6	244	51	2	880606923
7	166	346	1	886397596
8	298	474	4	884182806
9	115	265	2	881171488

```
In [3]: movie_titles = pd.read_csv("C:/Users/Admin/Desktop/Movie_Id_Titles.csv")
movie_titles.head()
data = pd.merge(df, movie_titles, on='item_id')
```

```
In [4]: data.head(10)
```

Out[4]:

	user_id	item_id	rating	timestamp	title
0	0	50	5	881250949	Star Wars (1977)
1	290	50	5	880473582	Star Wars (1977)
2	79	50	4	891271545	Star Wars (1977)
3	2	50	5	888552084	Star Wars (1977)
4	8	50	5	879362124	Star Wars (1977)
5	274	50	5	878944679	Star Wars (1977)
6	227	50	4	879035347	Star Wars (1977)
7	99	50	5	885679998	Star Wars (1977)
8	305	50	5	886321799	Star Wars (1977)
9	108	50	4	879879739	Star Wars (1977)

```
In [6]: data.groupby('title')['rating'].mean().sort_values(ascending=False).head(10)
```

Out[6]:

title	
Marlene Dietrich: Shadow and Light (1996)	5.0
Prefontaine (1997)	5.0
Santa with Muscles (1996)	5.0
Star Kid (1997)	5.0
Someone Else's America (1995)	5.0
Entertaining Angels: The Dorothy Day Story (1996)	5.0
Saint of Fort Washington, The (1993)	5.0
Great Day in Harlem, A (1994)	5.0
They Made Me a Criminal (1939)	5.0
Aiqing wansui (1994)	5.0
Name: rating, dtype: float64	

```
In [7]: data.groupby('title')['rating'].mean().sort_values(ascending=False).tail(8)
```

Out[7]:

title	
Somebody to Love (1994)	1.0
Invitation, The (Zaproszenie) (1986)	1.0
Lotto Land (1995)	1.0
Touki Bouki (Journey of the Hyena) (1973)	1.0
JLG/JLG - autoportrait de décembre (1994)	1.0
Daens (1992)	1.0
Butterfly Kiss (1995)	1.0
Eye of Vichy, The (Oeil de Vichy, L') (1993)	1.0
Name: rating, dtype: float64	

```
In [8]: data.groupby('title')['rating'].mean().sort_values(ascending=False).head(50)
```

```
Out[8]: title
Marlene Dietrich: Shadow and Light (1996)      5.000000
Prefontaine (1997)                             5.000000
Santa with Muscles (1996)                     5.000000
Star Kid (1997)                               5.000000
Someone Else's America (1995)                 5.000000
Entertaining Angels: The Dorothy Day Story (1996) 5.000000
Saint of Fort Washington, The (1993)           5.000000
Great Day in Harlem, A (1994)                 5.000000
They Made Me a Criminal (1939)                5.000000
Aiqing wansui (1994)                         5.000000
Pather Panchali (1955)                       4.625000
Anna (1996)                                   4.500000
Everest (1998)                               4.500000
Maya Lin: A Strong Clear Vision (1994)         4.500000
Some Mother's Son (1996)                     4.500000
Close Shave, A (1995)                        4.491071
Schindler's List (1993)                      4.466443
Wrong Trousers, The (1993)                   4.466102
Casablanca (1942)                           4.456790
Wallace & Gromit: The Best of Aardman Animation (1996) 4.447761
Shawshank Redemption, The (1994)             4.445230
Rear Window (1954)                          4.387560
Usual Suspects, The (1995)                   4.385768
Star Wars (1977)                             4.359589
12 Angry Men (1957)                         4.344000
Third Man, The (1949)                        4.333333
Letter From Death Row, A (1998)              4.333333
Bitter Sugar (Azucar Amargo) (1996)          4.333333
Citizen Kane (1941)                         4.292929
Some Folks Call It a Sling Blade (1993)      4.292683
To Kill a Mockingbird (1962)                 4.292237
One Flew Over the Cuckoo's Nest (1975)        4.291667
Silence of the Lambs, The (1991)             4.289744
North by Northwest (1959)                   4.284916
Godfather, The (1972)                       4.283293
Secrets & Lies (1996)                       4.265432
Good Will Hunting (1997)                    4.262626
Shall We Dance? (1996)                      4.260870
Manchurian Candidate, The (1962)             4.259542
Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1963) 4.252577
Raiders of the Lost Ark (1981)               4.252381
Vertigo (1958)                              4.251397
Innocents, The (1961)                       4.250000
Kaspar Hauser (1993)                        4.250000
Crossfire (1947)                            4.250000
Titanic (1997)                              4.245714
Lawrence of Arabia (1962)                   4.231214
Paths of Glory (1957)                       4.212121
Maltese Falcon, The (1941)                  4.210145
Empire Strikes Back, The (1980)              4.206522
Name: rating, dtype: float64
```

```
In [9]: data.groupby('title')['rating'].mean().sort_values(ascending=True).head(50)
```

```
Out[9]: title
Eye of Vichy, The (Oeil de Vichy, L') (1993)    1.0
Butterfly Kiss (1995)                          1.0
Daens (1992)                                  1.0
JLG/JLG - autoportrait de décembre (1994)      1.0
Touki Bouki (Journey of the Hyena) (1973)      1.0
Lotto Land (1995)                             1.0
Invitation, The (Zaproszenie) (1986)           1.0
Somebody to Love (1994)                       1.0
Low Life, The (1994)                          1.0
Police Story 4: Project S (Chao ji ji hua) (1993) 1.0
Power 98 (1995)                               1.0
I, Worst of All (Yo, la peor de todas) (1990)  1.0
Promise, The (Versprechen, Das) (1994)         1.0
Death in the Garden (Mort en ce jardin, La) (1956) 1.0
Hungarian Fairy Tale, A (1987)                1.0
New Age, The (1994)                           1.0
Pharaoh's Army (1995)                        1.0
Hostile Intentions (1994)                    1.0
Stefano Quantestorie (1993)                  1.0
Liebelei (1933)                              1.0
King of New York (1990)                      1.0
The Courtyard (1995)                         1.0
Nobody Loves Me (Keiner liebt mich) (1994)    1.0
Terror in a Texas Town (1958)                1.0
T-Men (1947)                                1.0
Office Killer (1997)                         1.0
Tigrero: A Film That Was Never Made (1994)    1.0
Boys in Venice (1996)                       1.0
Baton Rouge (1988)                          1.0
Symphonie pastorale, La (1946)               1.0
Amityville: Dollhouse (1996)                 1.0
Amityville: A New Generation (1993)           1.0
Leopard Son, The (1996)                      1.0
Crude Oasis, The (1995)                      1.0
Amityville 1992: It's About Time (1992)       1.0
To Cross the Rubicon (1991)                  1.0
Man from Down Under, The (1943)              1.0
Lashou shentan (1992)                       1.0
Quartier Mozart (1992)                       1.0
Mille bolle blu (1993)                       1.0
Further Gesture, A (1996)                    1.0
Wend Kuuni (God's Gift) (1982)               1.0
Modern Affair, A (1995)                      1.0
Careful (1992)                              1.0
Babyfever (1994)                            1.0
Homage (1995)                               1.0
3 Ninjas: High Noon At Mega Mountain (1998)  1.0
Every Other Weekend (1990)                   1.0
Woman in Question, The (1950)                1.0
Beyond Bedlam (1993)                         1.0
Name: rating, dtype: float64
```

```
In [10]: data.groupby('title')['rating'].count().sort_values(ascending=False).head(15)
```

```
Out[10]: title
Star Wars (1977)          584
Contact (1997)            509
Fargo (1996)              508
Return of the Jedi (1983) 507
Liar Liar (1997)          485
English Patient, The (1996) 481
Scream (1996)             478
Toy Story (1995)          452
Air Force One (1997)      431
Independence Day (ID4) (1996) 429
Raiders of the Lost Ark (1981) 420
Godfather, The (1972)     413
Pulp Fiction (1994)       394
Twelve Monkeys (1995)     392
Silence of the Lambs, The (1991) 390
Name: rating, dtype: int64
```

```
In [11]: ratings = pd.DataFrame(data.groupby('title')['rating'].mean())
```

```
In [12]: ratings.head(10)
```

Out[12]:

	rating
title	
'Til There Was You (1997)	2.333333
1-900 (1994)	2.600000
101 Dalmatians (1996)	2.908257
12 Angry Men (1957)	4.344000
187 (1997)	3.024390
2 Days in the Valley (1996)	3.225806
20,000 Leagues Under the Sea (1954)	3.500000
2001: A Space Odyssey (1968)	3.969112
3 Ninjas: High Noon At Mega Mountain (1998)	1.000000
39 Steps, The (1935)	4.050847

```
In [13]: ratings = pd.DataFrame(data.groupby('title')['rating'].mean())
ratings['num of ratings'] = pd.DataFrame(data.groupby('title')['rating'].count())
```

```
In [14]: ratings.head(15)
```

Out[14]:

	rating	num of ratings
title		
'Til There Was You (1997)	2.333333	9
1-900 (1994)	2.600000	5
101 Dalmatians (1996)	2.908257	109
12 Angry Men (1957)	4.344000	125
187 (1997)	3.024390	41
2 Days in the Valley (1996)	3.225806	93
20,000 Leagues Under the Sea (1954)	3.500000	72
2001: A Space Odyssey (1968)	3.969112	259
3 Ninjas: High Noon At Mega Mountain (1998)	1.000000	5
39 Steps, The (1935)	4.050847	59
8 1/2 (1963)	3.815789	38
8 Heads in a Duffel Bag (1997)	3.250000	4
8 Seconds (1994)	3.750000	4
A Chef in Love (1996)	4.125000	8
Above the Rim (1994)	3.000000	5

```
In [15]: moviemat = data.pivot_table(index='user_id', columns='title', values='rating')
moviemat.head()
ratings.sort_values('num of ratings', ascending=False).head(10)
```

Out[15]:

	rating	num of ratings
title		
Star Wars (1977)	4.359589	584
Contact (1997)	3.803536	509
Fargo (1996)	4.155512	508
Return of the Jedi (1983)	4.007890	507
Liar Liar (1997)	3.156701	485
English Patient, The (1996)	3.656965	481
Scream (1996)	3.441423	478
Toy Story (1995)	3.878319	452
Air Force One (1997)	3.631090	431
Independence Day (ID4) (1996)	3.438228	429

```
In [16]: data.groupby('title')['rating'].mean().sort_values(ascending=False).tail(10)
```

```
Out[16]: title
Police Story 4: Project S (Chao ji ji hua) (1993)    1.0
Low Life, The (1994)                                1.0
Somebody to Love (1994)                             1.0
Invitation, The (Zaproszenie) (1986)                 1.0
Lotto Land (1995)                                    1.0
Touki Bouki (Journey of the Hyena) (1973)            1.0
JLG/JLG - autoportrait de décembre (1994)            1.0
Daens (1992)                                          1.0
Butterfly Kiss (1995)                                1.0
Eye of Vichy, The (Oeil de Vichy, L') (1993)         1.0
Name: rating, dtype: float64
```

```
In [17]: ratings.sort_values('num of ratings', ascending = False).tail(10)
```

Out[17]:

	rating	num of ratings
title		
<hr/>		
Girl in the Cadillac (1995)	1.0	1
He Walked by Night (1948)	4.0	1
Hana-bi (1997)	3.0	1
Object of My Affection, The (1998)	4.0	1
Office Killer (1997)	1.0	1
Great Day in Harlem, A (1994)	5.0	1
Other Voices, Other Rooms (1997)	3.0	1
Good Morning (1971)	1.0	1
Girls Town (1996)	3.0	1
Á köldum klaka (Cold Fever) (1994)	3.0	1

```
In [ ]:
```