

```
In [1]: import pandas as pd
```

1. Read the dataset using pandas

```
In [2]: tags=pd.read_csv("tags.csv")
ratings=pd.read_csv("ratings.csv")
movies=pd.read_csv("movies.csv")
```

2. Extract the first row from tags and print its type

```
In [33]: tags.iloc[0,:]
```

```
Out[33]: userId           18
movieId          4141
tag      Mark Waters
timestamp    1240597180
Name: 0, dtype: object
```

```
In [4]: type(tags.head(0))
```

```
Out[4]: pandas.core.frame.DataFrame
```

3. Extract row 0, 11, 2000 from tags DataFrame

```
In [5]: tags.iloc[[0,11,2000],:]
```

```
Out[5]:
```

	userId	movieId	tag	timestamp
0	18	4141	Mark Waters	1240597180
11	65	1783	noir thriller	1368149983
2000	910	68554	conspiracy theory	1368043943

4. Print index, columns of the DataFrame

```
In [6]: tags.columns
```

```
Out[6]: Index(['userId', 'movieId', 'tag', 'timestamp'], dtype='object')
```

5. Calculate descriptive statistics for the 'ratings' column of the ratings DataFrame. Verify using describe()

```
In [8]: ratings.rating.std()
```

```
Out[8]: 1.051988919275684
```

```
In [9]: ratings.rating.max()
```

```
Out[9]: 5.0
```

```
In [10]: ratings.rating.count()
```

```
Out[10]: 20000263
```

```
In [11]: ratings.rating.quantile(q=0.75)
```

```
Out[11]: 4.0
```

```
In [12]: ratings.rating.mean()
```

```
Out[12]: 3.5255285642993797
```

```
In [100]: ratings.rating.describe()
```

```
Out[100]: count    2.000026e+07
mean      3.525529e+00
std       1.051989e+00
min       5.000000e-01
25%       3.000000e+00
50%       3.500000e+00
75%       4.000000e+00
max       5.000000e+00
Name: rating, dtype: float64
```

```
In [15]: movies.columns
```

```
Out[15]: Index(['movieId', 'title', 'genres'], dtype='object')
```

9. Find the average rating of movies

```
In [16]: ratings.rating.mean()
```

```
Out[16]: 3.5255285642993797
```

8. Filter out movies from the movies DataFrame that are of type 'Animation'

```
In [ ]: movies.columns
```

In [17]: `movies[movies['genres']=='Animation']`

Out[17]:

	moviedb_id	title	genres
2503	2588	Cloudland (1998)	Animation
4906	5002	Fritz the Cat (1972)	Animation
4907	5003	Nine Lives of Fritz the Cat, The (1974)	Animation
9455	27738	Cathedral, The (Katedra) (2002)	Animation
9989	32840	Vincent (1982)	Animation
13444	66335	Afro Samurai: Resurrection (2009)	Animation
13858	69469	Garfield's Pet Force (2009)	Animation
14184	71158	Immigrants (L.A. Dolce Vita) (2008)	Animation
14492	72603	Merry Madagascar (2009)	Animation
14578	72927	Donkey Xote (2007)	Animation
14580	72931	Tango (1981)	Animation
14876	74486	\$9.99 (2008)	Animation
14938	74791	Town Called Panic, A (Panique au village) (2009)	Animation
15085	76709	Spider-Man: The Ultimate Villain Showdown (2002)	Animation
15860	80469	Superman/Batman: Apocalypse (2010)	Animation
15990	81018	Illusionist, The (L'illusionniste) (2010)	Animation
16551	83603	Fern flowers (Fleur de fougère) (1949)	Animation
16649	84192	Corto Maltese: Ballad of the Salt Sea (Corto M...)	Animation
18144	90843	Lavatory Lovestory (Ubornaya istoriya - lyubov...)	Animation
18145	90845	Fall of the House of Usher, The (Zánik domu Us...)	Animation
18224	91187	Millhaven (2010)	Animation
18695	93083	Live Music (2009)	Animation
18980	94423	Disney Princess Collection: Jasmine's Enchante...	Animation
19562	96897	Bleach: Fade to Black (Burīchi Fade to Bl...)	Animation
20310	99820	Pokémon the Movie: Black - Victini and Reshira...	Animation
20311	99822	Pokémon the Movie: White - Victini and Zekrom ...	Animation
20314	99832	Hand, The (Ruka) (1966)	Animation
20527	100509	Tale of Tales (Skazka skazok) (1979)	Animation
20871	102007	Invincible Iron Man, The (2007)	Animation
20872	102009	Thor: Tales of Asgard (2011)	Animation
...
25207	118948	Blackbird (1959)	Animation
25208	118950	A Phantasy (1952)	Animation
25211	118956	Lines: Horizontal (1962)	Animation

movied		title	genres
25212	118958	Mosaic (1966)	Animation
25213	118960	Begone Dull Care (1949)	Animation
25214	118962	Synchromy (1971)	Animation
25538	120853	Fresh Guacamole (2012)	Animation
25539	120855	Adam and Dog (2012)	Animation
25663	121302	Someone's Gaze (2013)	Animation
25718	121600	Bosko's Parlor Pranks (1934)	Animation
26052	124889	The Adventures of Tom Thumb & Thumbelina (2002)	Animation
26163	125924	The Periwig-Maker (1999)	Animation
26169	125936	Crac (1981)	Animation
26171	125940	Syrinx (1966)	Animation
26181	125960	The Trip to Squash Land (1967)	Animation
26185	125968	It's Christmastime Again, Charlie Brown (1992)	Animation
26189	125976	Nocturna Artificialia (1979)	Animation
26192	125982	Trick or Treat (1952)	Animation
26207	126012	The Fat Albert Halloween Special (1977)	Animation
26226	126050	Stille Nacht I: Dramolet (1988)	Animation
26230	126058	Rehearsals for Extinct Anatomies (1987)	Animation
26236	126070	Chainsaw Maid (2007)	Animation
26237	126072	The Little Matchgirl (2006)	Animation
26246	126090	Hedgehog in the Fog (1975)	Animation
26247	126092	The Cat's Out (1931)	Animation
26248	126094	Claymation Comedy of Horrors (1991)	Animation
26319	126405	The Adventures of André and Wally B. (1984)	Animation
26809	128864	Four Sahibjade (2014)	Animation
27103	130394	The Mascot (1934)	Animation
27155	130644	The Garden of Sinners - Chapter 5: Paradox Par...	Animation

83 rows × 3 columns

10. Perform an inner join of movies and tags based on movied

In [18]: `movies_tags=pd.merge(movies,tags,how='inner',on='movieId')`

In [19]: movies_tags

Out[19]:

	movieId	title	genres	userId	tag	tim
0	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	1644	Watched	1417
1	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	1741	computer animation	1183
2	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	1741	Disney animated feature	1183
3	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	1741	Pixar animation	1183
4	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	1741	TÃ©a Leoni does not star in this movie	1245
5	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	2299	Pixar	1403
6	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	3596	animation	1290
7	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	3596	family	1290
8	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	3596	Tom Hanks	1290
9	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	6233	Pixar	1387
10	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	6233	witty	1387
11	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	8219	Pixar	1159
12	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	adventure	1277
13	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	animated	1277
14	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	animation	1277
15	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	clever	1277
16	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	comedy	1277
17	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	computer animation	1277
18	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	family	1277
19	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	fantasy	1277
20	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	10616	Tom Hanks	1277

	movieId	title	genres	userId	tag	tim
21	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	bright	1173
22	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	DARING RESCUES	1173
23	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	fanciful	1173
24	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	HEROIC MISSION	1173
25	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	humorous	1173
26	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	light	1173
27	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	rousing	1173
28	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	TOYS COME TO LIFE	1173
29	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	11081	UNLIKELY FRIENDSHIPS	1173
...
465534	131015	Hellgate (2011)	Horror Thriller	25978	Cary Elwes	1427
465535	131015	Hellgate (2011)	Horror Thriller	25978	Demons	1427
465536	131015	Hellgate (2011)	Horror Thriller	25978	Exorcism	1427
465537	131015	Hellgate (2011)	Horror Thriller	25978	Ghosts	1427
465538	131015	Hellgate (2011)	Horror Thriller	25978	Monks	1427
465539	131015	Hellgate (2011)	Horror Thriller	25978	nurse	1427
465540	131015	Hellgate (2011)	Horror Thriller	25978	psychic	1427
465541	131015	Hellgate (2011)	Horror Thriller	25978	slow	1427
465542	131015	Hellgate (2011)	Horror Thriller	25978	special effects	1427
465543	131015	Hellgate (2011)	Horror Thriller	25978	surfing	1427
465544	131015	Hellgate (2011)	Horror Thriller	25978	Thailand	1427
465545	131015	Hellgate (2011)	Horror Thriller	25978	William Hurt	1427
465546	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	Barcelona	1427

	movieid	title	genres	userId	tag	tim
465547	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	exploitation	1427
465548	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	female football	1427
465549	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	football	1427
465550	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	Rome	1427
465551	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	shorts	1427
465552	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	soccer	1427
465553	131031	La liga no es cosa de hombres (1972)	(no genres listed)	33323	transvestite	1427
465554	131054	Dinotopia: Quest for the Ruby Sunstone (2005)	Children Fantasy Sci-Fi	88144	dinosaurs	1427
465555	131082	Playground (2009)	(no genres listed)	92885	documentary	1427
465556	131082	Playground (2009)	(no genres listed)	92885	Yoshitomo Nara	1427
465557	131164	Vietnam in HD (2011)	War	54560	Vietnam War	1427
465558	131170	Parallels (2015)	Sci-Fi	88738	alternate reality	1427
465559	131258	The Pirates (2014)	Adventure	28906	bandits	1427
465560	131258	The Pirates (2014)	Adventure	28906	Korea	1427
465561	131258	The Pirates (2014)	Adventure	28906	mutiny	1427
465562	131258	The Pirates (2014)	Adventure	28906	pirates	1427

movield	title	genres	userId	tag	tim
465563	131258 The Pirates (2014)	Adventure	28906	whale	1427

465564 rows × 6 columns

Perform an inner join of movies and tags based on movield

```
In [20]: movies_rating=pd.merge(ratings,movies,how='inner',on='movieId')
```

In [21]: movies_rating

Out[21]:

	userId	movieId	rating	timestamp	title	genres
0	1	2	3.5	1112486027	Jumanji (1995)	Adventure Children Fantasy
1	5	2	3.0	851527569	Jumanji (1995)	Adventure Children Fantasy
2	13	2	3.0	849082742	Jumanji (1995)	Adventure Children Fantasy
3	29	2	3.0	835562174	Jumanji (1995)	Adventure Children Fantasy
4	34	2	3.0	846509384	Jumanji (1995)	Adventure Children Fantasy
5	54	2	3.0	974918176	Jumanji (1995)	Adventure Children Fantasy
6	88	2	1.0	1098277938	Jumanji (1995)	Adventure Children Fantasy
7	91	2	3.5	1112061358	Jumanji (1995)	Adventure Children Fantasy
8	116	2	2.0	1132728068	Jumanji (1995)	Adventure Children Fantasy
9	119	2	4.0	845110667	Jumanji (1995)	Adventure Children Fantasy
10	120	2	1.0	838153425	Jumanji (1995)	Adventure Children Fantasy
11	124	2	2.0	1134476330	Jumanji (1995)	Adventure Children Fantasy
12	127	2	3.0	847127455	Jumanji (1995)	Adventure Children Fantasy
13	128	2	3.0	900720958	Jumanji (1995)	Adventure Children Fantasy
14	129	2	3.0	1283448701	Jumanji (1995)	Adventure Children Fantasy
15	131	2	1.0	1238326861	Jumanji (1995)	Adventure Children Fantasy
16	132	2	3.0	1114172997	Jumanji (1995)	Adventure Children Fantasy
17	137	2	3.0	992797771	Jumanji (1995)	Adventure Children Fantasy
18	142	2	4.0	833458658	Jumanji (1995)	Adventure Children Fantasy
19	152	2	3.0	1161215251	Jumanji (1995)	Adventure Children Fantasy
20	156	2	5.0	1040937649	Jumanji (1995)	Adventure Children Fantasy
21	158	2	3.0	842734429	Jumanji (1995)	Adventure Children Fantasy
22	164	2	3.0	845248397	Jumanji (1995)	Adventure Children Fantasy
23	165	2	3.0	835483502	Jumanji (1995)	Adventure Children Fantasy
24	170	2	3.0	940253294	Jumanji (1995)	Adventure Children Fantasy
25	184	2	3.0	840016743	Jumanji (1995)	Adventure Children Fantasy
26	220	2	3.0	1056172947	Jumanji (1995)	Adventure Children Fantasy
27	232	2	3.5	1162148886	Jumanji (1995)	Adventure Children Fantasy
28	235	2	3.0	939225131	Jumanji (1995)	Adventure Children Fantasy
29	238	2	3.0	1111509612	Jumanji (1995)	Adventure Children Fantasy
...
20000233	136015	95578	4.0	1341625654	Stoplight Society, The (La Sociedad del Semáforo...)	Drama

	userId	movieId	rating	timestamp	title	genres
20000234	136015	107147	3.5	1386962880	Last Winter (L'hiver dernier) (2011)	Drama
20000235	136054	101820	3.0	1365472258	Fighting Prince of Donegal, The (1966)	Action Adventure
20000236	136196	129313	4.0	1424987707	Reality (2014)	Comedy
20000237	136268	65621	3.0	1232604498	Liberty Kid (2007)	Drama
20000238	136268	114174	2.5	1411048756	American Dreams in China (2013)	Comedy Drama
20000239	136268	125656	3.0	1421693709	Another Day, Another Time: Celebrating the Mus...	Documentary
20000240	136571	126815	4.0	1422587314	Anotherworld (2008)	Drama Romance
20000241	136684	130522	3.5	1426970944	The Brave Little Toaster Goes to Mars (1998)	Animation Children
20000242	137031	117442	3.5	1421807244	The Monkey King (2014)	Action Adventure Children IMAX
20000243	137031	121699	4.0	1421807493	Katt Williams: Priceless: Afterlife (2014)	Comedy
20000244	137031	123188	2.5	1421883946	Ironclad 2: Battle for Blood (2014)	Action Adventure
20000245	137277	110052	3.5	1395643640	Stay as You Are (1978)	Drama Romance
20000246	137717	120831	2.5	1420430755	Dealin' with Idiots (2013)	Comedy
20000247	137717	120833	3.0	1420602817	Super Capers (2009)	Action Adventure Comedy Fantasy Sci-Fi
20000248	137717	120839	3.5	1420444839	Comes a Bright Day (2012)	Romance
20000249	137717	127445	4.0	1423802804	Saturday Night (2010)	Documentary
20000250	137893	65909	4.0	1233561753	Maria (2003)	Drama
20000251	137893	71300	3.0	1305104853	Conspiracy of Torture, The (Beatrice Cenci) (1...	Drama
20000252	138301	103010	3.5	1416786218	Just Like Brothers (Comme des frères) (2012)	Comedy Drama Romance

	userId	movieId	rating	timestamp	title	genres
20000253	138301	104307	4.0	1415227515	Souper, Le (Supper, The) (1992)	Drama
20000254	138301	106170	4.0	1383873162	Patriots, The (Patriotes, Les) (1994)	Thriller
20000255	138301	106401	3.0	1394152234	Le convoyeur (2004)	Crime Drama Thriller
20000256	138301	113539	4.0	1408674064	First Man, The (Le Premier Homme) (2011)	Drama
20000257	138301	118856	4.0	1418447976	Paris by Night (2012)	Crime Drama
20000258	138301	121017	3.5	1420558479	The Gentleman from Epsom (1962)	Comedy Crime
20000259	138301	121019	4.5	1420558606	The Great Spy Chase (1964)	Action Comedy Thriller
20000260	138301	121021	4.5	1420558687	Taxi for Tobruk (1961)	Drama War
20000261	138406	110167	4.5	1396184127	Judge and the Assassin, The (Juge et l'assassin...) (1969)	Crime Drama
20000262	138406	110510	4.5	1397597539	Série noire (1979)	Film-Noir

20000263 rows × 6 columns

11. Print out the 5 movies that belong to the Comedy genre and have rating greater than 4

```
In [62]: rating5=movies_rating[movies_rating['genres']=='Comedy'][movies_rating['rating']>4
rating5.groupby(by='title')['title'].head()
```

C:\Users\Acer\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: UserWarning:
 g: Boolean Series key will be reindexed to match DataFrame index.
 """Entry point for launching an IPython kernel.

```
Out[62]: 190505          Clerks (1994)
190508          Clerks (1994)
190509          Clerks (1994)
190510          Clerks (1994)
190512          Clerks (1994)
766651  Monty Python's Life of Brian (1979)
766656  Monty Python's Life of Brian (1979)
766661  Monty Python's Life of Brian (1979)
766663  Monty Python's Life of Brian (1979)
766664  Monty Python's Life of Brian (1979)
1837152  Ferris Bueller's Day Off (1986)
1837162  Ferris Bueller's Day Off (1986)
1837165  Ferris Bueller's Day Off (1986)
1837170  Ferris Bueller's Day Off (1986)
1837171  Ferris Bueller's Day Off (1986)
2380426  Monty Python's The Meaning of Life (1983)
2380428  Monty Python's The Meaning of Life (1983)
2380429  Monty Python's The Meaning of Life (1983)
2380435  Monty Python's The Meaning of Life (1983)
2380436  Monty Python's The Meaning of Life (1983)
2929204          Airplane! (1980)
2929208          Airplane! (1980)
2929222          Airplane! (1980)
2929223          Airplane! (1980)
2929224          Airplane! (1980)
3015291  Grumpy Old Men (1993)
3015296  Grumpy Old Men (1993)
3015297  Grumpy Old Men (1993)
3015308  Grumpy Old Men (1993)
3015311  Grumpy Old Men (1993)
...
19993415          Gorko! (2013)
19993461  Parasites, Les (1999)
19993569  World's Greatest Lover, The (1977)
19993572  World's Greatest Lover, The (1977)
19993981  How to Be a Man (2013)
19994180  Craig Ferguson: A Wee Bit o' Revolution (2009)
19994182  Craig Ferguson: A Wee Bit o' Revolution (2009)
19994809  Public Eye, The (Follow Me!) (1972)
19994877  A Short History of Decay (2014)
19994881  Down and Out with the Dolls (2001)
19994997  George Carlin: Life Is Worth Losing (2005)
19994998  George Carlin: Life Is Worth Losing (2005)
19995044  Hollywood Sex Wars (2011)
19995593  Eddie Izzard: Circle (2002)
19996624  Schtonk! (1992)
19996880  On a marché sur Bangkok (2014)
19997116  Kevin Nealon: Whelmed, But Not Overly (2012)
19997132  What About Dick? (2012)
```

```
19997140           The great match (2007)
19997542           Dans la peau d'une grande (2011)
19998011   Flight of the Conchords: A Texan Odyssey (2006)
19998286           People, Places, Things (2015)
19998955           Late George Apley, The (1947)
19999104           The Admirable Crichton (1957)
19999249           Bill Hicks: Sane Man (1989)
19999504           This Is the Night (1932)
19999904           Living with the Fosters (2002)
19999949   Tale from the Past, A (Përralle Nga e Kaluara)...
20000197           Inquire Within (2012)
20000206           Armless (2010)
Name: title, Length: 5405, dtype: object
```

13. Extract year from title e.g. (1995)

In [99]:

```
tags=tags_ts
movies_tags=pd.merge(movies,tags,how='inner',on='movieId')
mt_1=movies_tags.iloc[:,[1,6]]
mt_1
```

Out[99]:

	title	year
0	Toy Story (1995)	2006
1	Toy Story (1995)	2006
2	Toy Story (1995)	2006
3	Toy Story (1995)	2006
4	Toy Story (1995)	2006
5	Toy Story (1995)	2006
6	Toy Story (1995)	2006
7	Toy Story (1995)	2006
8	Toy Story (1995)	2006
9	Toy Story (1995)	2006
10	Toy Story (1995)	2006
11	Toy Story (1995)	2006
12	Toy Story (1995)	2006
13	Toy Story (1995)	2006
14	Toy Story (1995)	2006
15	Toy Story (1995)	2006
16	Toy Story (1995)	2006
17	Toy Story (1995)	2006
18	Toy Story (1995)	2006
19	Toy Story (1995)	2006
20	Toy Story (1995)	2006
21	Toy Story (1995)	2006
22	Toy Story (1995)	2006
23	Toy Story (1995)	2006
24	Toy Story (1995)	2006
25	Toy Story (1995)	2006
26	Toy Story (1995)	2006
27	Toy Story (1995)	2006
28	Toy Story (1995)	2006
29	Toy Story (1995)	2006

	title	year
...
465534	Hellgate (2011)	2015
465535	Hellgate (2011)	2015
465536	Hellgate (2011)	2015
465537	Hellgate (2011)	2015
465538	Hellgate (2011)	2015
465539	Hellgate (2011)	2015
465540	Hellgate (2011)	2015
465541	Hellgate (2011)	2015
465542	Hellgate (2011)	2015
465543	Hellgate (2011)	2015
465544	Hellgate (2011)	2015
465545	Hellgate (2011)	2015
465546	La liga no es cosa de hombres (1972)	2015
465547	La liga no es cosa de hombres (1972)	2015
465548	La liga no es cosa de hombres (1972)	2015
465549	La liga no es cosa de hombres (1972)	2015
465550	La liga no es cosa de hombres (1972)	2015
465551	La liga no es cosa de hombres (1972)	2015
465552	La liga no es cosa de hombres (1972)	2015
465553	La liga no es cosa de hombres (1972)	2015
465554	Dinotopia: Quest for the Ruby Sunstone (2005)	2015
465555	Playground (2009)	2015
465556	Playground (2009)	2015
465557	Vietnam in HD (2011)	2015
465558	Parallels (2015)	2015
465559	The Pirates (2014)	2015
465560	The Pirates (2014)	2015
465561	The Pirates (2014)	2015
465562	The Pirates (2014)	2015
465563	The Pirates (2014)	2015

465564 rows × 2 columns

15. Sort the tags DataFrame based on timestamp

```
In [48]: tags['timestamp']=tags['timestamp'].apply(pd.to_datetime,unit='s')
tags_ts=tags.sort_values(by='timestamp')
tags_ts
```

Out[48]:

	userId	movieId	tag	timestamp
333932	100371	2788	monty python	2005-12-24 13:00:10
333927	100371	1732	coen brothers	2005-12-24 13:00:36
333924	100371	1206	stanley kubrick	2005-12-24 13:00:48
333923	100371	1193	jack nicholson	2005-12-24 13:02:51
333939	100371	5004	peter sellers	2005-12-24 13:03:19
333922	100371	47	morgan freeman	2005-12-24 13:03:32
333921	100371	47	brad pitt	2005-12-24 13:03:32
333936	100371	4011	brad pitt	2005-12-24 13:03:51
333937	100371	4011	guy ritchie	2005-12-24 13:03:51
333920	100371	32	bruce willis	2005-12-24 13:04:02
333933	100371	2858	kevin spacey	2005-12-24 13:04:26
333931	100371	2329	fuckoff nazi	2005-12-24 13:05:08
333930	100371	2329	edward norton	2005-12-24 13:05:08
333944	100371	6016	brazil	2005-12-24 13:05:20
15903	3797	110	overrated	2005-12-26 09:24:35
256663	75141	4642	Terrible	2005-12-27 12:54:55
15926	3797	8573	Nonlinear Surrealism	2005-12-30 09:50:23
68928	16050	38886	twee	2005-12-31 13:43:32
68924	16050	38886	eighties	2005-12-31 13:43:32
68925	16050	38886	family	2005-12-31 13:43:32
68926	16050	38886	new york	2005-12-31 13:43:32
68927	16050	38886	overrated	2005-12-31 13:43:32
68919	16050	7060	christian	2005-12-31 14:28:21
68921	16050	7060	musical	2005-12-31 14:28:21
68920	16050	7060	kitsch	2005-12-31 14:28:21
68922	16050	7060	religious	2005-12-31 14:28:21
68918	16050	6898	harsh	2005-12-31 14:29:28
68916	16050	6898	british	2005-12-31 14:29:28
68917	16050	6898	drugs	2005-12-31 14:29:28
81444	20513	36529	truth	2006-01-02 22:57:03
...
119997	28906	115149	violent	2015-03-30 20:07:11
119991	28906	115149	lack of story	2015-03-30 20:07:33

userId	movieId		tag	timestamp
119993	28906	115149	over the top	2015-03-30 20:07:40
119989	28906	115149	cheesy	2015-03-30 20:07:54
405155	123297	128642	CLV	2015-03-30 20:12:17
399877	123297	4211	CLV	2015-03-30 20:18:29
400993	123297	6419	CLV	2015-03-30 20:21:04
401017	123297	6481	CLV	2015-03-30 20:27:00
405123	123297	118722	CLV	2015-03-30 20:28:41
400106	123297	4573	CLV	2015-03-30 20:31:02
150873	42640	115122	mockumentary	2015-03-30 20:52:58
150871	42640	115122	comedy	2015-03-30 20:53:17
150872	42640	115122	dark comedy	2015-03-30 20:53:28
392905	122523	5004	USA	2015-03-30 21:51:29
392904	122523	5004	party	2015-03-30 21:51:32
290536	88044	106782	unsatisfying ending	2015-03-30 22:03:20
290534	88044	106782	Orgies	2015-03-30 22:05:11
290529	88044	106782	ENFP	2015-03-30 22:05:47
290532	88044	106782	Left-Wing Propaganda	2015-03-30 22:07:46
290533	88044	106782	Left-Wing Writers	2015-03-30 22:13:41
290528	88044	106782	Economically Illiterate Writers	2015-03-30 22:15:39
290530	88044	106782	inaccurate	2015-03-30 22:16:46
290531	88044	106782	Jonah Hill	2015-03-30 22:17:29
290526	88044	106782	Amoral	2015-03-30 22:18:33
290527	88044	106782	crime	2015-03-30 22:18:41
290535	88044	106782	profanity	2015-03-30 22:21:36
288375	87797	215	Vienna	2015-03-30 22:50:01
158763	46072	3409	premonition	2015-03-31 00:12:06
158780	46072	6058	premonition	2015-03-31 00:12:44
339178	102853	115149	russian mafia	2015-03-31 03:09:12

465564 rows × 4 columns

14. Select rows based on timestamps later than 2015-02-01

```
In [60]: start_date = '2015-02-01'
mask=(tags_ts['timestamp']>start_date)
tags_ts1=tags_ts.loc[mask]
tags_ts1
```

Out[60]:

	userId	movieId	tag	timestamp
431961	130446	113565	christianity	2015-02-01 00:04:03
431964	130446	113565	cult	2015-02-01 00:04:11
431965	130446	113565	drama	2015-02-01 00:04:18
431966	130446	113565	fake documentary	2015-02-01 00:04:26
431967	130446	113565	found footage	2015-02-01 00:04:32
431958	130446	113565	born-again	2015-02-01 00:04:41
431970	130446	113565	Jonestown	2015-02-01 00:04:50
431969	130446	113565	horror	2015-02-01 00:04:55
431963	130446	113565	community	2015-02-01 00:05:10
431962	130446	113565	closed community	2015-02-01 00:05:29
431974	130446	113565	suicide	2015-02-01 00:05:41
431960	130446	113565	charismatic leader	2015-02-01 00:05:55
431971	130446	113565	journalism	2015-02-01 00:06:54
431972	130446	113565	journalist	2015-02-01 00:06:58
431968	130446	113565	helicopter	2015-02-01 00:07:09
431973	130446	113565	mass murder	2015-02-01 00:07:29
431959	130446	113565	brainwashing	2015-02-01 00:08:18
124475	31782	318	tim robbins	2015-02-01 00:09:53
124474	31782	318	steven king	2015-02-01 00:09:53
124473	31782	318	prison break	2015-02-01 00:09:53
383326	119989	114935	mindfuck	2015-02-01 00:35:27
370204	114601	95890	improvisation	2015-02-01 00:47:05
370206	114601	95890	shane meadows	2015-02-01 00:47:05
370205	114601	95890	paddy considine	2015-02-01 00:47:05
17886	4274	109374	ridiculous	2015-02-01 01:56:44
17884	4274	109374	colourful sets	2015-02-01 01:56:44
17885	4274	109374	life-long relationship	2015-02-01 01:56:44
361771	111473	6539	action	2015-02-01 02:48:38
361773	111473	6539	pirates	2015-02-01 02:48:38
361772	111473	6539	fantasy	2015-02-01 02:48:38
...
119997	28906	115149	violent	2015-03-30 20:07:11

userId	movieId		tag	timestamp
119991	28906	115149	lack of story	2015-03-30 20:07:33
119993	28906	115149	over the top	2015-03-30 20:07:40
119989	28906	115149	cheesy	2015-03-30 20:07:54
405155	123297	128642	CLV	2015-03-30 20:12:17
399877	123297	4211	CLV	2015-03-30 20:18:29
400993	123297	6419	CLV	2015-03-30 20:21:04
401017	123297	6481	CLV	2015-03-30 20:27:00
405123	123297	118722	CLV	2015-03-30 20:28:41
400106	123297	4573	CLV	2015-03-30 20:31:02
150873	42640	115122	mockumentary	2015-03-30 20:52:58
150871	42640	115122	comedy	2015-03-30 20:53:17
150872	42640	115122	dark comedy	2015-03-30 20:53:28
392905	122523	5004	USA	2015-03-30 21:51:29
392904	122523	5004	party	2015-03-30 21:51:32
290536	88044	106782	unsatisfying ending	2015-03-30 22:03:20
290534	88044	106782	Orgies	2015-03-30 22:05:11
290529	88044	106782	ENFP	2015-03-30 22:05:47
290532	88044	106782	Left-Wing Propaganda	2015-03-30 22:07:46
290533	88044	106782	Left-Wing Writers	2015-03-30 22:13:41
290528	88044	106782	Economically Illiterate Writers	2015-03-30 22:15:39
290530	88044	106782	inaccurate	2015-03-30 22:16:46
290531	88044	106782	Jonah Hill	2015-03-30 22:17:29
290526	88044	106782	Amoral	2015-03-30 22:18:33
290527	88044	106782	crime	2015-03-30 22:18:41
290535	88044	106782	profanity	2015-03-30 22:21:36
288375	87797	215	Vienna	2015-03-30 22:50:01
158763	46072	3409	premonition	2015-03-31 00:12:06
158780	46072	6058	premonition	2015-03-31 00:12:44
339178	102853	115149	russian mafia	2015-03-31 03:09:12

12130 rows × 4 columns

6. Filter out ratings with rating > 5

```
In [94]: ratings_5=ratings[(ratings['rating'])>5]
ratings_5.count()
```

```
Out[94]: userId      0
movieId     0
rating      0
timestamp   0
dtype: int64
```

7. Find how many null values, missing values are present. Deal with them. Print out how many rows have been modified.

```
In [108]: tags.isnull().sum()
```

```
Out[108]: userId      0
movieId     0
tag         16
timestamp   0
year        0
dtype: int64
```

```
In [110]: tags.isna().sum()
```

```
Out[110]: userId      0
movieId     0
tag         16
timestamp   0
year        0
dtype: int64
```

```
In [113]: tags[tags['tag'].isnull()].index.tolist()
```

```
Out[113]: [373289,
 373334,
 373339,
 373325,
 373276,
 373332,
 373340,
 373291,
 373301,
 373299,
 373277,
 373288,
 373281,
 373319,
 373303,
 454615]
```

```
In [ ]:
```

