

Ques 08: How do I apply multiple filter criteria to a pandas DataFrame?

```
In [3]: # IMDB - Internet Movie Database (Dataset)
```

```
import pandas as pd
movie_ratings_ds = pd.read_csv('http://bit.ly/imdbratings')
movie_ratings_ds.head(4)
```

Out[3]:

	star_rating	title	content_rating	genre	duration	actors_list
0	9.3	The Shawshank Redemption	R	Crime	142	[u'Tim Robbins', u'Morgan Freeman', u'Bob Gunt...]
1	9.2	The Godfather	R	Crime	175	[u'Marlon Brando', u'Al Pacino', u'James Caan']
2	9.1	The Godfather: Part II	R	Crime	200	[u'Al Pacino', u'Robert De Niro', u'Robert Duv...]
3	9.0	The Dark Knight	PG-13	Action	152	[u'Christian Bale', u'Heath Ledger', u'Aaron E...]

```
In [4]: movie_ratings_ds[(movie_ratings_ds.duration >= 200) & (movie_ratings_ds.genre == 'Drama')]
```

Out[4]:

	star_rating	title	content_rating	genre	duration	actors_list
17	8.7	Seven Samurai	UNRATED	Drama	207	[u'Toshirô Mifune', u'Takashi Shimura', u'K...]
157	8.2	Gone with the Wind	G	Drama	238	[u'Clark Gable', u'Vivien Leigh', u'Thomas Mit...]
476	7.8	Hamlet	PG-13	Drama	242	[u'Kenneth Branagh', u'Julie Christie', u'Dere...]

```
In [5]: movie_ratings_ds[(movie_ratings_ds.duration >= 200) && (movie_ratings_ds.genre == 'Drama')].head()
```

Out[5]:

	star_rating	title	content_rating	genre	duration	actors_list
2	9.1	The Godfather: Part II	R	Crime	200	[u'Al Pacino', u'Robert De Niro', u'Robert Duv...]
5	8.9	12 Angry Men	NOT RATED	Drama	96	[u'Henry Fonda', u'Lee J. Cobb', u'Martin Bals...]
7	8.9	The Lord of the Rings: The Return of the King	PG-13	Adventure	201	[u'Elijah Wood', u'Viggo Mortensen', u'Ian McK...]
9	8.9	Fight Club	R	Drama	139	[u'Brad Pitt', u'Edward Norton', u'Helena Bonh...]
13	8.8	Forrest Gump	PG-13	Drama	142	[u'Tom Hanks', u'Robin Wright', u'Gary Sinise']

```
In [6]: # Specify multiple-values for a column
```

```
movie_ratings_ds[ movie_ratings_ds.genre.isin(['Drama', 'Crime', 'Adventure'])].head(10)
```

Out[6]:

	star_rating	title	content_rating	genre	duration	actors_list
0	9.3	The Shawshank Redemption	R	Crime	142	[u'Tim Robbins', u'Morgan Freeman', u'Bob Gunt...]
1	9.2	The Godfather	R	Crime	175	[u'Marlon Brando', u'Al Pacino', u'James Caan']
2	9.1	The Godfather: Part II	R	Crime	200	[u'Al Pacino', u'Robert De Niro', u'Robert Duv...]
4	8.9	Pulp Fiction	R	Crime	154	[u'John Travolta', u'Uma Thurman', u'Samuel L....]
5	8.9	12 Angry Men	NOT RATED	Drama	96	[u'Henry Fonda', u'Lee J. Cobb', u'Martin Bals...]
7	8.9	The Lord of the Rings: The Return of the King	PG-13	Adventure	201	[u'Elijah Wood', u'Viggo Mortensen', u'Ian McK...]
9	8.9	Fight Club	R	Drama	139	[u'Brad Pitt', u'Edward Norton', u'Helena Bonh...]
10	8.8	The Lord of the Rings: The Fellowship of the Ring	PG-13	Adventure	178	[u'Elijah Wood', u'Ian McKellen', u'Orlando Bl...]
13	8.8	Forrest Gump	PG-13	Drama	142	[u'Tom Hanks', u'Robin Wright', u'Gary Sinise']
14	8.8	The Lord of the Rings: The Two Towers	PG-13	Adventure	179	[u'Elijah Wood', u'Ian McKellen', u'Viggo Mort...]

How to read only some specific columns & rows from a CSV file?

```
In [7]: ufo_ds = pd.read_csv('http://bit.ly/uforeports', usecols = [0,3], nrows = 5)
# Picking 1st and 4th column [ City and State ]
# picking first 5 rows
ufo_ds
```

Out[7]:

	City	State
0	Ithaca	NY
1	Willingboro	NJ
2	Holyoke	CO
3	Abilene	KS
4	New York Worlds Fair	NY

How to iterate a Series and Dataframe ?

```
In [8]: for c in ufo_ds.State:  
        print(c)
```

```
NY  
NJ  
CO  
KS  
NY
```

```
In [9]: for index, row in ufo_ds.iterrows():  
        print(index, row.City, row.State)
```

```
0 Ithaca NY  
1 Willingboro NJ  
2 Holyoke CO  
3 Abilene KS  
4 New York Worlds Fair NY
```