Filter values -> Movies database

November 17, 2016

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

In [2]: movies = pd.read_csv('http://bit.ly/imdbratings')

In [3]: movies.head()
```

```
Out[3]:    star_rating                      title content_rating    genre  duration
        0          9.3   The Shawshank Redemption              R    Crime       142
        1          9.2              The Godfather              R    Crime       175
        2          9.1      The Godfather: Part II             R    Crime       200
        3          9.0            The Dark Knight          PG-13   Action       152
        4          8.9               Pulp Fiction              R    Crime       154

                                        actors_list
        0  [u'Tim Robbins', u'Morgan Freeman', u'Bob Gunt...
        1    [u'Marlon Brando', u'Al Pacino', u'James Caan']
        2  [u'Al Pacino', u'Robert De Niro', u'Robert Duv...
        3  [u'Christian Bale', u'Heath Ledger', u'Aaron E...
        4  [u'John Travolta', u'Uma Thurman', u'Samuel L...
```

```
In [4]: movies.shape

Out[4]: (979, 6)

In [6]: type(True)

Out[6]: bool

In [7]: # create list of Bools that tells us which rows match a condition
        booleans = []
        for length in movies.duration:
            if length >= 200:
                booleans.append(True)
            else:
                booleans.append(False)

In [10]: booleans[0:5] # check first five values

Out[10]: [False, False, True, False, False]
```

```
In [11]: len(booleans)

Out[11]: 979

In [12]: # convert booleans list into pandas series
         is_long = pd.Series(booleans)

In [13]: is_long.head()

Out[13]: 0    False
         1    False
         2     True
         3    False
         4    False
         dtype: bool

In [15]: # movies[is_long] # longer version
         # movies[movies.duration >= 200]
         is_long = movies.duration >= 200 # more efficient version
         is_long.head()

Out[15]: 0    False
         1    False
         2     True
         3    False
         4    False
         Name: duration, dtype: bool

In [16]: movies[movies.duration >= 200]

Out[16]:      star_rating                                          title  \
         2            9.1                          The Godfather: Part II
         7            8.9  The Lord of the Rings: The Return of the King
         17           8.7                                   Seven Samurai
         78           8.4                      Once Upon a Time in America
         85           8.4                              Lawrence of Arabia
         142          8.3                  Lagaan: Once Upon a Time in India
         157          8.2                                 Gone with the Wind
         204          8.1                                          Ben-Hur
         445          7.9                             The Ten Commandments
         476          7.8                                           Hamlet
         630          7.7                                        Malcolm X
         767          7.6                  It's a Mad, Mad, Mad, Mad World

              content_rating       genre   duration  \
         2                 R       Crime        200
         7             PG-13   Adventure        201
         17          UNRATED       Drama        207
         78                R       Crime        229
```

```
85            PG  Adventure      216
142           PG  Adventure      224
157            G      Drama      238
204            G  Adventure      212
445      APPROVED  Adventure      220
476         PG-13      Drama      242
630         PG-13  Biography      202
767      APPROVED     Action      205

                                actors_list
2    [u'Al Pacino', u'Robert De Niro', u'Robert Duv...
7    [u'Elijah Wood', u'Viggo Mortensen', u'Ian McK...
17   [u'Toshir\xf4 Mifune', u'Takashi Shimura', u'K...
78   [u'Robert De Niro', u'James Woods', u'Elizabet...
85   [u"Peter O'Toole", u'Alec Guinness', u'Anthony...
142  [u'Aamir Khan', u'Gracy Singh', u'Rachel Shell...
157  [u'Clark Gable', u'Vivien Leigh', u'Thomas Mit...
204  [u'Charlton Heston', u'Jack Hawkins', u'Stephe...
445  [u'Charlton Heston', u'Yul Brynner', u'Anne Ba...
476  [u'Kenneth Branagh', u'Julie Christie', u'Dere...
630  [u'Denzel Washington', u'Angela Bassett', u'De...
767  [u'Spencer Tracy', u'Milton Berle', u'Ethel Me...
```

In [17]: `movies[movies.duration >= 200].genre` *# or ['genre']*

```
Out[17]: 2        Crime
         7    Adventure
         17       Drama
         78       Crime
         85   Adventure
         142  Adventure
         157      Drama
         204  Adventure
         445  Adventure
         476      Drama
         630  Biography
         767     Action
         Name: genre, dtype: object
```

In [18]: `movies.loc[movies.duration >= 200, 'genre']` *# loc*

```
Out[18]: 2        Crime
         7    Adventure
         17       Drama
         78       Crime
         85   Adventure
         142  Adventure
         157      Drama
         204  Adventure
```

```
445     Adventure
476         Drama
630     Biography
767        Action
Name: genre, dtype: object
```

In [ ]: