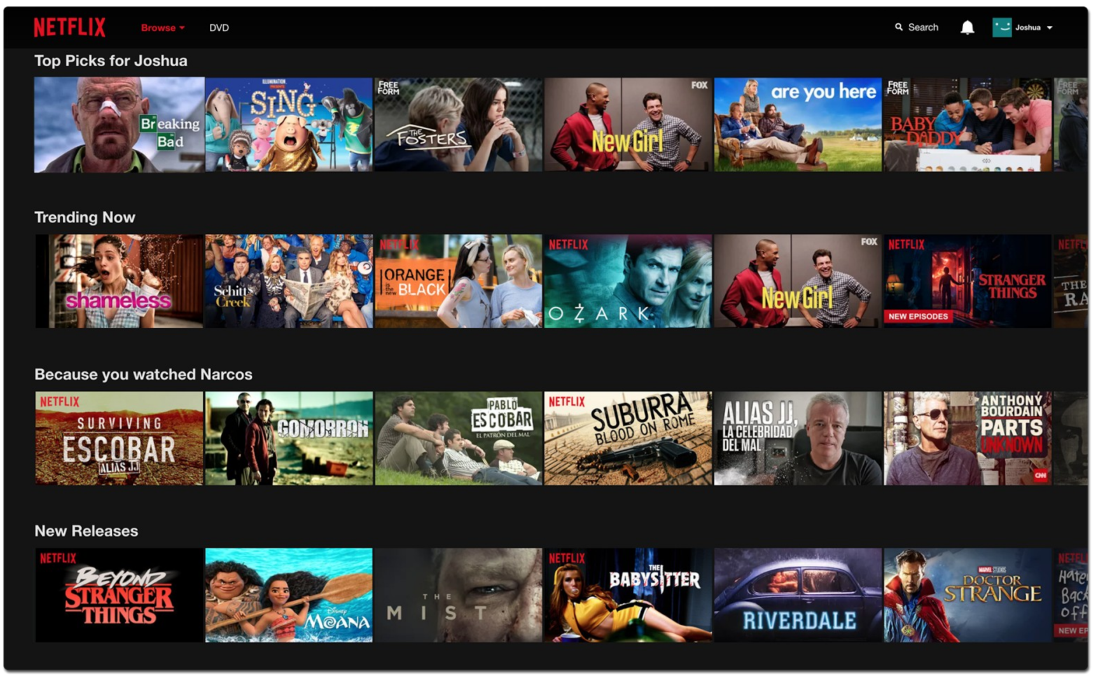
Movie Recommendation System

Recommender systems are utilized in a variety of areas, and are most commonly recognized as playlist generators for video and music services like Netflix, YouTube and Spotify, product recommenders for services such as Amazon, or content recommenders for social media platforms such as Facebook and Twitter. Recommender systems usually make use of either or both Collaborative Filtering or Content-Based filtering approach.



Here we will be recommending movies to the users who have watched movie 'Jurassic Park (1993)' using Collaborative Filtering

Import necessary libraries

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
```

Import the data set

Importing the rating data set which contains ratings given by the users to the movies they watched.

```python
rating = pd.read_csv('ratings.csv')
```

```
rating.head()
```

```
   userId  movieId  rating  timestamp
0       1        1     4.0  964982703
1       1        3     4.0  964981247
2       1        6     4.0  964982224
3       1       47     5.0  964983815
4       1       50     5.0  964982931
```

Importing the movie dataset which contains the description about all the movies

```
movies = pd.read_csv("movies.csv")
```

```
movies.head()
```

```
   movieId                               title  \
0        1                    Toy Story (1995)
1        2                      Jumanji (1995)
2        3             Grumpier Old Men (1995)
3        4            Waiting to Exhale (1995)
4        5  Father of the Bride Part II (1995)

                                        genres
0  Adventure|Animation|Children|Comedy|Fantasy
1                   Adventure|Children|Fantasy
2                               Comedy|Romance
3                         Comedy|Drama|Romance
4                                       Comedy
```

Let's merge both the dataset so that in ratings dataset we have complete information about the movies apart from the movie id.

```
# merging both the datasets on 'movieId' column
movie_rating = pd.merge(left=rating,right=movies,on='movieId')
```

```
movie_rating.head()
```

```
   userId  movieId  rating   timestamp             title  \
0       1        1     4.0   964982703  Toy Story (1995)
1       5        1     4.0   847434962  Toy Story (1995)
2       7        1     4.5  1106635946  Toy Story (1995)
3      15        1     2.5  1510577970  Toy Story (1995)
4      17        1     4.5  1305696483  Toy Story (1995)

                                        genres
0  Adventure|Animation|Children|Comedy|Fantasy
1  Adventure|Animation|Children|Comedy|Fantasy
2  Adventure|Animation|Children|Comedy|Fantasy
3  Adventure|Animation|Children|Comedy|Fantasy
4  Adventure|Animation|Children|Comedy|Fantasy
```

```
movie_rating.columns
```

```
Index(['userId', 'movieId', 'rating', 'timestamp', 'title', 'genres'],
dtype='object')
```

Getting the columns of the movie_rating dataframe in proper order

```
movie_rating = movie_rating[['userId', 'movieId', 'title', 'genres',
'rating', 'timestamp']]
```

```
movie_rating.head()
```

```
   userId  movieId             title  \
0       1        1  Toy Story (1995)
1       5        1  Toy Story (1995)
2       7        1  Toy Story (1995)
3      15        1  Toy Story (1995)
4      17        1  Toy Story (1995)

                                        genres  rating   timestamp
0  Adventure|Animation|Children|Comedy|Fantasy     4.0   964982703
1  Adventure|Animation|Children|Comedy|Fantasy     4.0   847434962
2  Adventure|Animation|Children|Comedy|Fantasy     4.5  1106635946
3  Adventure|Animation|Children|Comedy|Fantasy     2.5  1510577970
4  Adventure|Animation|Children|Comedy|Fantasy     4.5  1305696483
```

Exploratory Data Analysis

```
movie_rating.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 100836 entries, 0 to 100835
Data columns (total 6 columns):
 #   Column     Non-Null Count   Dtype
---  ------     --------------   -----
 0   userId     100836 non-null  int64
 1   movieId    100836 non-null  int64
 2   title      100836 non-null  object
 3   genres     100836 non-null  object
 4   rating     100836 non-null  float64
 5   timestamp  100836 non-null  int64
dtypes: float64(1), int64(3), object(2)
memory usage: 5.4+ MB
```

```
movie_rating.isnull().sum()
```

```
userId       0
movieId      0
title        0
genres       0
rating       0
timestamp    0
dtype: int64
```

Let's create a dataframe with number of ratings and average rating for each movie

```
movie_rating.head(2)
```

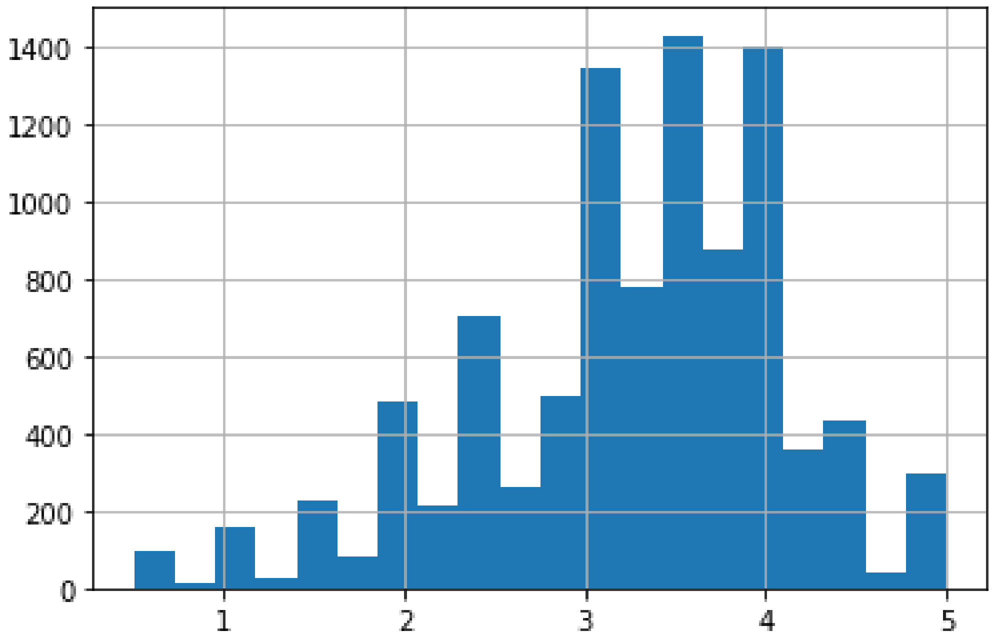
```
   userId  movieId             title  \
0       1        1  Toy Story (1995)
1       5        1  Toy Story (1995)

                                        genres  rating  timestamp
0  Adventure|Animation|Children|Comedy|Fantasy     4.0  964982703
1  Adventure|Animation|Children|Comedy|Fantasy     4.0  847434962
```

```
# grouping the movies based on average rating
average_rating_movies = movie_rating.groupby('title')
['rating'].mean().sort_values(ascending=False)
```

```
average_rating_movies.head(10)
```

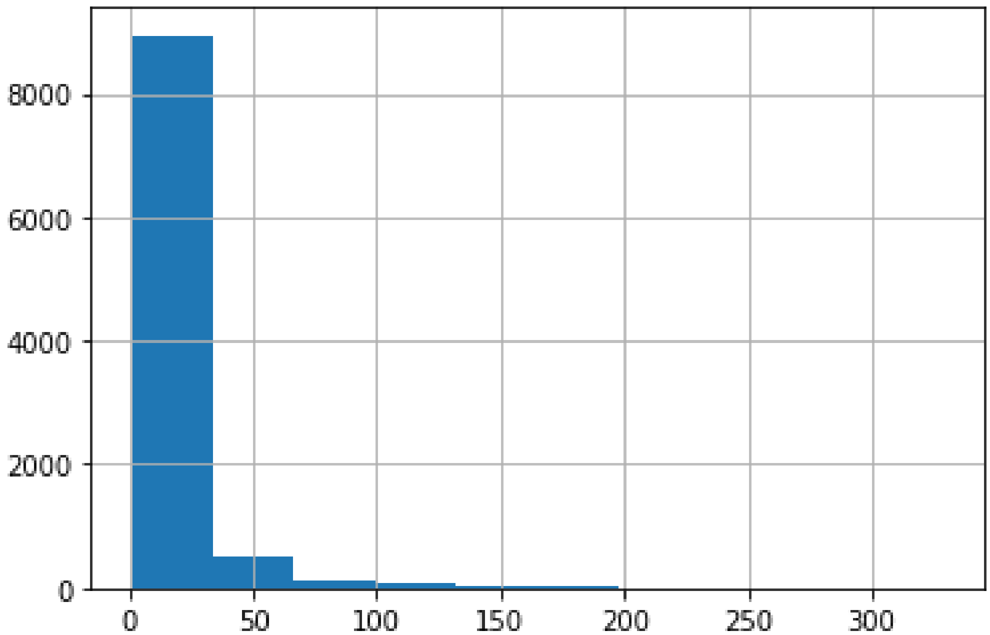
```
title
Gena the Crocodile (1969)                       5.0
True Stories (1986)                             5.0
Cosmic Scrat-tastrophe (2015)                   5.0
Love and Pigeons (1985)                         5.0
Red Sorghum (Hong gao liang) (1987)             5.0
Thin Line Between Love and Hate, A (1996)       5.0
Lesson Faust (1994)                             5.0
Eva (2011)                                      5.0
Who Killed Chea Vichea? (2010)                  5.0
Siam Sunset (1999)                              5.0
Name: rating, dtype: float64
```

```
average_rating_movies.hist(bins=20)
plt.show()
```



Maximum movies have average rating in the range 3 to 4. The movies which have average = 5.0 may be the ones which may have been watched once or twice.

```
# grouping the movies based on count of users who rated the movies
count_userid = movie_rating.groupby('title')
['userId'].count().sort_values(ascending=False)

count_userid

title
Forrest Gump (1994)                             329
Shawshank Redemption, The (1994)                317
Pulp Fiction (1994)                             307
Silence of the Lambs, The (1991)                279
Matrix, The (1999)                              278
                                               ...
King Solomon's Mines (1950)                       1
King Solomon's Mines (1937)                       1
King Ralph (1991)                                 1
King Kong Lives (1986)                            1
À nous la liberté (Freedom for Us) (1931)         1
Name: userId, Length: 9719, dtype: int64

count_userid.hist()
plt.show()
```



Maximum movies have been viewed in the range 0 - 40 views

The movies which have average = 5.0 may be the ones which may have been watched once or twice. Let's see number of ratings given to movies which have average rating = 5.0

```
for movie in average_rating_movies[average_rating_movies==5.0].index:
    print(movie,count_userid[movie])
```

```
Gena the Crocodile (1969) 1
True Stories (1986) 1
Cosmic Scrat-tastrophe (2015) 1
Love and Pigeons (1985) 1
Red Sorghum (Hong gao liang) (1987) 1
Thin Line Between Love and Hate, A (1996) 1
Lesson Faust (1994) 2
Eva (2011) 1
Who Killed Chea Vichea? (2010) 1
Siam Sunset (1999) 1
Ex Drummer (2007) 1
Reform School Girls (1986) 1
Buzzard (2015) 1
Hellbenders (2012) 1
Heidi Fleiss: Hollywood Madam (1995) 2
My Left Eye Sees Ghosts (Ngo joh aan gin diy gwai) (2002) 1
Animals are Beautiful People (1974) 1
My Life as McDull (Mak dau goo si) (2001) 1
My Love (2006) 1
My Man Godfrey (1957) 1
Continental Divide (1981) 1
My Sassy Girl (Yeopgijeogin geunyeo) (2001) 1
Trinity and Sartana Are Coming (1972) 1
```

Calcium Kid, The (2004) 1
Connections (1978) 1
Superman/Batman: Public Enemies (2009) 1
Mystery of the Third Planet, The (Tayna tretey planety) (1981) 1
Mr. Skeffington (1944) 1
Shogun Assassin (1980) 1
Tickling Giants (2017) 1
Entertaining Angels: The Dorothy Day Story (1996) 1
Crippled Avengers (Can que) (Return of the 5 Deadly Venoms) (1981) 1
Tales of Manhattan (1942) 1
Monster Squad, The (1987) 1
SORI: Voice from the Heart (2016) 1
Awfully Big Adventure, An (1995) 1
Brother (Brat) (1997) 1
Light Years (Gandahar) (1988) 1
Moonlight 1
There Once Was a Dog (1982) 1
More (1998) 1
Holy Motors (2012) 1
Moscow Does Not Believe in Tears (Moskva slezam ne verit) (1979) 1
Mother (Madeo) (2009) 1
Hollywood Shuffle (1987) 1
Sherlock Holmes and Dr. Watson: Acquaintance (1979) 1
When Worlds Collide (1951) 1
Hollywood Chainsaw Hookers (1988) 1
Empties (2007) 1
Jane Eyre (1944) 1
Craig Ferguson: I'm Here To Help (2013) 1
Asterix and the Vikings (Astérix et les Vikings) (2006) 1
Assignment, The (1997) 1
English Vinglish (2012) 1
Runaway Brain (1995) 1
Enter the Void (2009) 2
Supercop 2 (Project S) (Chao ji ji hua) (1993) 1
Willy/Milly (1986) 1
Travels of an Ant (1983) 1
Last Year's Snow Was Falling (1983) 1
All Yours (2016) 1
Slumber Party Massacre III (1990) 1
Junior and Karlson (1968) 1
Strictly Sexual (2008) 1
Won't You Be My Neighbor? (2018) 1
Particle Fever (2013) 1
Last Hurrah for Chivalry (Hao xia) (1979) 1
Alien Contamination (1980) 1
Wonder Woman (2009) 1
Justice League: Doom (2012) 1
Larry David: Curb Your Enthusiasm (1999) 1
Alesha Popovich and Tugarin the Dragon (2004) 1
Karlson Returns (1970) 1

Snowflake, the White Gorilla (2011) 1
Five Senses, The (1999) 1
Nine Lives of Tomas Katz, The (2000) 1
Three from Prostokvashino (1978) 1
World of Glory (1991) 1
Story of Women (Affaire de femmes, Une) (1988) 1
Return to Treasure Island (1988) 1
Wow! A Talking Fish! (1983) 1
Lamerica (1994) 2
Advise and Consent (1962) 1
Laggies (2014) 1
Paper Birds (Pájaros de papel) (2010) 1
Jump In! (2007) 1
Stuart Little 3: Call of the Wild (2005) 1
Slumber Party Massacre II (1987) 1
Sun Alley (Sonnenallee) (1999) 1
Nasu: Summer in Andalusia (2003) 1
National Lampoon's Bag Boy (2007) 1
Peaceful Warrior (2006) 1
Harlan County U.S.A. (1976) 1
Jonah Who Will Be 25 in the Year 2000 (Jonas qui aura 25 ans en l'an 2000) (1976) 2
Faster (2010) 1
Sisters (Syostry) (2001) 1
Wings, Legs and Tails (1986) 1
Winnie Pooh (1969) 1
Winnie the Pooh Goes Visiting (1971) 1
Winnie the Pooh and the Day of Concern (1972) 1
American Friend, The (Amerikanische Freund, Der) (1977) 1
Woman Under the Influence, A (1974) 1
Come and See (Idi i smotri) (1985) 2
Winter in Prostokvashino (1984) 1
Summer's Tale, A (Conte d'été) (1996) 1
Happy Feet Two (2011) 1
Trailer Park Boys (1999) 1
Colourful (Karafuru) (2010) 1
Thousand Clowns, A (1965) 1
Passenger, The (Professione: reporter) (1975) 1
Guy X (2005) 1
Woman Is a Woman, A (femme est une femme, Une) (1961) 1
All the Vermeers in New York (1990) 1
Eichmann (2007) 1
Babes in Toyland (1934) 1
What Men Talk About (2010) 1
Marriage of Maria Braun, The (Ehe der Maria Braun, Die) (1979) 1
Bitter Lake (2015) 1
The Girl with All the Gifts (2016) 1
Very Potter Sequel, A (2010) 1
The Fox and the Hound 2 (2006) 1
Denise Calls Up (1995) 1

Radio Day (2008) 1
Bill Hicks: Revelations (1993) 1
Man and a Woman, A (Un homme et une femme) (1966) 1
Big Top Scooby-Doo! (2012) 1
Villain (1971) 1
The Eye: Infinity (2005) 1
Bobik Visiting Barbos (1977) 1
The Editor (2015) 1
Unicorn City (2012) 1
Man with the Golden Arm, The (1955) 1
Idiots and Angels (2008) 1
Ice Age: The Great Egg-Scapade (2016) 1
Unfaithfully Yours (1948) 1
The Love Bug (1997) 1
Maniac Cop 2 (1990) 1
Rain (2001) 1
I, the Jury (1982) 1
Vovka in the Kingdom of Far Far Away (1965) 1
I'm the One That I Want (2000) 1
Saving Face (2004) 1
Delirium (2014) 1
The Girls (1961) 1
Blue Planet II (2017) 1
Love Exposure (Ai No Mukidashi) (2008) 1
Loving Vincent (2017) 1
Vacations in Prostokvashino (1980) 1
The Jinx: The Life and Deaths of Robert Durst (2015) 1
Death Note: Desu nôto (2006–2007) 1
Vagabond (Sans toit ni loi) (1985) 1
Bloodsucking Bastards (2015) 1
Valet, The (La doublure) (2006) 1
Lumberjack Man (2015) 1
In the blue sea, in the white foam. (1984) 1
Vampire in Venice (Nosferatu a Venezia) (Nosferatu in Venice) (1986) 1
In the Realm of the Senses (Ai no corrida) (1976) 1
De platte jungle (1978) 1
Scooby-Doo! and the Samurai Sword (2009) 1
Deathgasm (2015) 1
Black Tar Heroin: The Dark End of the Street (2000) 1
Madame Sousatzka (1988) 1
Scooby-Doo Goes Hollywood (1979) 1
Decalogue, The (Dekalog) (1989) 1
Louis Theroux: Law & Disorder (2008) 1
Scooby-Doo! Abracadabra-Doo (2010) 1
Scooby-Doo! Curse of the Lake Monster (2010) 1
Saving Santa (2013) 1
Scooby-Doo! and the Loch Ness Monster (2004) 1
Black Mirror 1
Raise Your Voice (2004) 1
The Bremen Town Musicians (1969) 1

What Love Is (2007) 1
Martin Lawrence Live: Runteldat (2002) 1
Live Nude Girls (1995) 1
Watching the Detectives (2007) 1
Tyler Perry's I Can Do Bad All by Myself (2009) 1
Branded to Kill (Koroshi no rakuin) (1967) 1
Investigation Held by Kolobki (1986) 1
Little Murders (1971) 1
Sandpiper, The (1965) 1
Watermark (2014) 1
Two Family House (2000) 1
Human Condition III, The (Ningen no joken III) (1961) 1
Mickey's Once Upon a Christmas (1999) 1
Human (2015) 1
Seve (2014) 1
Dragons: Gift of the Night Fury (2011) 1
Breed, The (2006) 1
Little Dieter Needs to Fly (1997) 1
Dream of Light (a.k.a. Quince Tree Sun, The) (Sol del membrillo, El) (1992) 1
Tenchi Muyô! In Love (1996) 1
Duel in the Sun (1946) 1
Ballad of Narayama, The (Narayama bushiko) (1983) 1
Miss Nobody (2010) 1
Dylan Moran: Monster (2004) 1
Cruel Romance, A (Zhestokij Romans) (1984) 1
Crossing Delancey (1988) 1
What Happened Was... (1994) 1
Hunting Elephants (2013) 1
Battle For Sevastopol (2015) 1
Battle Royale 2: Requiem (Batoru rowaiaru II: Chinkonka) (2003) 1
I Am Not Your Negro (2017) 1
Belle époque (1992) 2
Bossa Nova (2000) 1
Seems Like Old Times (1980) 1
Match Factory Girl, The (Tulitikkutehtaan tyttö) (1990) 1
The Big Bus (1976) 1
Max Manus (2008) 1
Satin Rouge (2002) 1
Priklyucheniya Kapitana Vrungelya (1979) 1
Into the Abyss (2011) 1
The Adventures of Sherlock Holmes and Dr. Watson: Bloody Signature (1979) 1
The Adventures of Sherlock Holmes and Doctor Watson: The Treasures of Agra (1983) 1
The Adventures of Sherlock Holmes and Doctor Watson: King of Blackmailers (1980) 1
Dr. Goldfoot and the Bikini Machine (1965) 1
The Adventures of Sherlock Holmes and Doctor Watson 1
Meantime (1984) 1

Boy Eats Girl (2005) 1
Umberto D. (1952) 1
Presto (2008) 1
Hype! (1996) 1
Ugly Duckling and Me!, The (2006) 1
Down Argentine Way (1940) 1
Mephisto (1981) 1
Into the Forest of Fireflies' Light (2011) 1
Into the Woods (1991) 1
Adventures Of Sherlock Holmes And Dr. Watson: The Twentieth Century Approaches (1986) 1
Indignation (2016) 1
Tom Segura: Mostly Stories (2016) 1
Ghost Graduation (2012) 1
Knock Off (1998) 1
Four Seasons, The (1981) 1
Kung Fu Panda: Secrets of the Masters (2011) 1
Spellbound (2011) 1
Get Low (2009) 1
Four Days in September (O Que É Isso, Companheiro?) (1997) 1
One I Love, The (2014) 1
A Flintstones Christmas Carol (1994) 1
Lady Jane (1986) 1
61* (2001) 1
George Carlin: You Are All Diseased (1999) 1
George Carlin: Life Is Worth Losing (2005) 1
Galaxy of Terror (Quest) (1981) 1
Act of Killing, The (2012) 1
Chinese Puzzle (Casse-tête chinois) (2013) 1
Formula of Love (1984) 1
12 Chairs (1976) 1
Sorority House Massacre II (1990) 1
A Detective Story (2003) 1
7 Faces of Dr. Lao (1964) 1
12 Angry Men (1997) 1
Rivers and Tides (2001) 1
Zeitgeist: Moving Forward (2011) 1
Odd Life of Timothy Green, The (2012) 1
Tom and Jerry: Shiver Me Whiskers (2006) 1
Chump at Oxford, A (1940) 1
Sonatine (Sonachine) (1993) 1
Che: Part One (2008) 1
Front of the Class (2008) 1
Che: Part Two (2008) 1
PK (2014) 1
Son of the Bride (Hijo de la novia, El) (2001) 1
On the Other Side of the Tracks (De l'autre côté du périph) (2012) 1
Cheburashka (1971) 1
On the Ropes (1999) 1
Fugitives (1986) 1

```
Going Places (Valseuses, Les) (1974) 1
On the Trail of the Bremen Town Musicians (1973) 1
Stand, The (1994) 1
Cherish (2002) 1
A Plasticine Crow (1981) 1
King of Hearts (1966) 1
20 Million Miles to Earth (1957) 1
Tokyo Tribe (2014) 1
Tom and Jerry: A Nutcracker Tale (2007) 1
A Perfect Day (2015) 1
George Carlin: Jammin' in New York (1992) 1
Girls About Town (1931) 1
Sorority House Massacre (1986) 1
Ooops! Noah is Gone... (2015) 1
L.A. Slasher (2015) 1
Palindromes (2004) 1
Goodbye Charlie (1964) 1
'Salem's Lot (2004) 1
Garfield's Pet Force (2009) 1
George Carlin: Back in Town (1996) 1
9/11 (2002) 1
Only Lovers Left Alive (2013) 1
Obsession (1965) 1
Oscar (1967) 1
Tom Segura: Completely Normal (2014) 1
George Carlin: It's Bad for Ya! (2008) 1
Garden of Words, The (Koto no ha no niwa) (2013) 1
Go for Zucker! (Alles auf Zucker!) (2004) 1
Open Hearts (Elsker dig for evigt) (2002) 1
```

```
# grouping the movie_rating based on count on userId and mean on
rating
userid_rating = movie_rating.groupby('title')
[['userId','rating']].agg({'userId':'count','rating':'mean'}).round(2)
.sort_values(by='userId',ascending=False)

userid_rating.head()
```

```
                                  userId  rating
title
Forrest Gump (1994)                  329    4.16
Shawshank Redemption, The (1994)     317    4.43
Pulp Fiction (1994)                  307    4.20
Silence of the Lambs, The (1991)     279    4.16
Matrix, The (1999)                   278    4.19
```

Building Recommendation System

```
# creating pivot table to create item by item collaborative filtering
movie_rating_pivot =
```

```
pd.pivot_table(index='userId',columns='title',values='rating',data=mov
ie_rating)
```

There will be many Nan values because users have watched only few of the movies and given ratings only to those movies

```
movie_rating_pivot.head()
```

```
title   '71 (2014)  'Hellboy': The Seeds of Creation (2004)  \
userId
1              NaN                                      NaN
2              NaN                                      NaN
3              NaN                                      NaN
4              NaN                                      NaN
5              NaN                                      NaN

title   'Round Midnight (1986)  'Salem's Lot (2004)  \
userId
1                          NaN                  NaN
2                          NaN                  NaN
3                          NaN                  NaN
4                          NaN                  NaN
5                          NaN                  NaN

title   'Til There Was You (1997)  'Tis the Season for Love (2015)  \
userId
1                             NaN                              NaN
2                             NaN                              NaN
3                             NaN                              NaN
4                             NaN                              NaN
5                             NaN                              NaN

title   'burbs, The (1989)  'night Mother (1986)  (500) Days of Summer
(2009)  \
userId

1                      NaN                   NaN
NaN
2                      NaN                   NaN
NaN
3                      NaN                   NaN
NaN
4                      NaN                   NaN
NaN
5                      NaN                   NaN
NaN

title   *batteries not included (1987)  ...  Zulu (2013)  [REC] (2007)
\
userId                                  ...
```

```
1                                  NaN  ...          NaN           NaN

2                                  NaN  ...          NaN           NaN

3                                  NaN  ...          NaN           NaN

4                                  NaN  ...          NaN           NaN

5                                  NaN  ...          NaN           NaN


title   [REC]² (2009)  [REC]³ 3 Génesis (2012)  \
userId
1                 NaN                        NaN
2                 NaN                        NaN
3                 NaN                        NaN
4                 NaN                        NaN
5                 NaN                        NaN

title   anohana: The Flower We Saw That Day - The Movie (2013)  \
userId
1                                                     NaN
2                                                     NaN
3                                                     NaN
4                                                     NaN
5                                                     NaN

title   eXistenZ (1999)  xXx (2002)  xXx: State of the Union (2005)  \
userId
1                   NaN         NaN                             NaN
2                   NaN         NaN                             NaN
3                   NaN         NaN                             NaN
4                   NaN         NaN                             NaN
5                   NaN         NaN                             NaN

title   ¡Three Amigos! (1986)  À nous la liberté (Freedom for Us)
(1931)
userId

1                         4.0
NaN
2                         NaN
NaN
3                         NaN
NaN
4                         NaN
NaN
5                         NaN
```

```
NaN

[5 rows x 9719 columns]
```

Most Rated movies:

```
userid_rating.head(10)
```

```
                                           userId  rating
title
Forrest Gump (1994)                           329    4.16
Shawshank Redemption, The (1994)              317    4.43
Pulp Fiction (1994)                           307    4.20
Silence of the Lambs, The (1991)              279    4.16
Matrix, The (1999)                            278    4.19
Star Wars: Episode IV - A New Hope (1977)     251    4.23
Jurassic Park (1993)                          238    3.75
Braveheart (1995)                             237    4.03
Terminator 2: Judgment Day (1991)             224    3.97
Schindler's List (1993)                       220    4.22
```

Let's find which movies to recommend to the users who have watched 'Jurassic Park (1993)'. To do this we have to find correlation of 'Jurassic Park (1993)' with other movies which have been rated in a similar way by the users.

```
# assigning ratings of movie 'Jurassic Park (1993)' to a new variable
from movie_rating_pivot
jurassic_park = movie_rating_pivot['Jurassic Park (1993)'].head(10)
```

```
jurassic_park.head(10)
```

```
userId
1     4.0
2     NaN
3     NaN
4     NaN
5     NaN
6     5.0
7     5.0
8     4.0
9     NaN
10    NaN
Name: Jurassic Park (1993), dtype: float64
```

Find the correlation with other movies from movie_rating_pivot table

```
correlation_jurassicpark =
pd.DataFrame(movie_rating_pivot.corrwith(jurassic_park))
```

```
/usr/local/lib/python3.8/dist-packages/numpy/lib/
function_base.py:2821: RuntimeWarning: Degrees of freedom <= 0 for
slice
```

```
  c = cov(x, y, rowvar, dtype=dtype)
/usr/local/lib/python3.8/dist-packages/numpy/lib/function_base.py:2680
: RuntimeWarning: divide by zero encountered in true_divide
  c *= np.true_divide(1, fact)
```

```
correlation_jurassicpark.head()
```

```
                                          0
title
'71 (2014)                              NaN
'Hellboy': The Seeds of Creation (2004) NaN
'Round Midnight (1986)                  NaN
'Salem's Lot (2004)                     NaN
'Til There Was You (1997)               NaN
```

Removing Nan values and naming the column as 'Correlation'

```
correlation_jurassicpark.columns = ['Correlation']
correlation_jurassicpark.dropna(inplace=True,axis=0)
```

```
correlation_jurassicpark.sort_values(by='Correlation',ascending=True).
head()
```

```
                                                    Correlation
title
X-Men (2000)                                               -1.0
Austin Powers: International Man of Mystery (1997)         -1.0
Enemy of the State (1998)                                  -1.0
Gladiator (2000)                                           -1.0
Interview with the Vampire: The Vampire Chronic...         -1.0
```

There may be movies which might have been watched only once or twice by the users who have watched 'Jurassic Park (1993)' and those movies will show high correlation. We will consider only those movies which have been viewed more than 100 times. Let's add views column in the correlation_jurassicpark data frame

```
correlation_jurassicpark['Views'] = userid_rating['userId']
```

Now filtering out top 20 movies which have views greater than 100

```
correlation_jurassicpark[correlation_jurassicpark['Views'] >
100].sort_values(by='Correlation',ascending=False).head(20)
```

```
                                         Correlation  Views
title
Jurassic Park (1993)                        1.000000    238
Mission: Impossible (1996)                  1.000000    162
Twister (1996)                              1.000000    123
Speed (1994)                                1.000000    171
Pretty Woman (1990)                         1.000000    135
Outbreak (1995)                             1.000000    101
Toy Story (1995)                            1.000000    215
```

Twelve Monkeys (a.k.a. 12 Monkeys) (1995)	1.000000	177
GoldenEye (1995)	1.000000	132
Stargate (1994)	1.000000	140
Clear and Present Danger (1994)	1.000000	110
Batman Forever (1995)	1.000000	137
Independence Day (a.k.a. ID4) (1996)	0.970725	202
Mrs. Doubtfire (1993)	0.944911	144
Forrest Gump (1994)	0.904534	329
Braveheart (1995)	0.866025	237
Silence of the Lambs, The (1991)	0.577350	279
Apollo 13 (1995)	0.500000	201
Mask, The (1994)	0.500000	157
Fugitive, The (1993)	0.500000	190