

Project 4

2022-12-12

System 1 - Recommendations by Genre

Group Members: Sarah Yang (sarahxy2), Ryan Berry (rpberry2)

We created three different algorithms for recommending movies by genre:

1. Highest Rated
2. Highest Rated and Recently Created
3. Most Popular

We will detail each of the three algorithms further below. All algorithms went through the same data loading and preprocessing steps.

Data Loading

Load data to prepare for preprocessing.

In [1]:

```
import pandas as pd

genre_list = ["Action", "Adventure", "Animation",
              "Children's", "Comedy", "Crime",
              "Documentary", "Drama", "Fantasy",
              "Film-Noir", "Horror", "Musical",
              "Mystery", "Romance", "Sci-Fi",
              "Thriller", "War", "Western"]

def load_data():
    movies = pd.read_csv('movies.dat', sep='::', names=['MovieID',
                                                         'Title', 'Genre', 'Year'],
                        encoding='latin-1')
    ratings = pd.read_csv('ratings.dat', sep='::',
                          names=['UserID', 'MovieID', 'Rating',
                                 'Timestamp'])
    return movies, ratings
```

Data Preprocessing

For all algorithms we undergo the same preprocessing steps, however, algorithm 2 has an additional step to remove movies produced after a certain year.

Our steps are as follows:

- Remove year from title and assign to own column
- Remove movies that were reviewed less than 30 times, because that is not a significant sample size
- (Algo 2 only) - remove movies created after 1995

- Assign averages for ratings for each movie

In [2]:

```
def preprocess(movies, ratings, year_filter=None):
    # assign year
    movies['Year'] = movies['Title'].str[-5:-1]
    movies['Year'] = movies['Year'].str.extract('(\d+)', expand=False)
    movies = movies[movies['Year'].notna()]
    movies['Year'] = movies['Year'].astype(int)

    # remove year from title
    movies['Title'] = movies['Title'].str[:-6]

    # remove movies where there are less than 30 total reviews
    ratings_grouped = ratings.groupby('MovieID',
as_index=False).size()
    movies = movies.join(ratings_grouped.set_index('MovieID'),
on='MovieID')
    movies = movies.rename(columns={'size': 'RatingCount'})
    size1 = len(movies)
    movies = movies[movies['RatingCount'] >= 30]
    size2 = len(movies)
    print('\n\nRemoved', size1-size2, 'movies because there aren\'t
enough reviews\n\n')

    # remove movies from before a certain year
    if year_filter:
        movies = movies[movies['Year'] >= year_filter]

    # assign average ratings
    ratings_avg = ratings.groupby('MovieID', as_index=False)
```

```
['Rating'].mean()
    movies = movies.join(ratings_avg.set_index('MovieID'),
on='MovieID')
    movies = movies.rename(columns={'Rating': 'AvgRating'})

    return movies
```

Data Analysis

In this step, we identify the top n movies based on a certain category.

In [3]:

```
def get_top_by_genre(movies, category, top_n):
    top_dict = {}

    for genre in genre_list:
        genre_movies = movies[movies['Genre'].str.contains(genre)]
        genre_top = genre_movies.sort_values(by=[category], ascending=
[False]).head(top_n)
        top_dict[genre] = genre_top.to_dict('records')

    return top_dict
```

Data Display

This step helps to display recommendation results more clearly.

In [4]:

```
def display_top_results(top_dict, title=None):
    print(title)
    for key, value in top_dict.items():
        print('Genre:', key)
        for movie in value:
            print('Movie:', movie)
        print('\n')
    print('\n')
```

Testing Algorithm 1 (Highest Rated)

Highest Rated refers to the overall highest average rated movies per genre. To get this metric, we take the mean rating across all reviews for each movie that survived preprocessing.

Notice how the AvgRating dictates the order of results.

In [5]:

```
# Get the top 5 movies per genre by average rating
movies_df, ratings_df = load_data()

cleaned_movies_df = preprocess(movies_df, ratings_df)
genre_dict = get_top_by_genre(cleaned_movies_df, 'AvgRating', top_n=5)
display_top_results(genre_dict, '\n\nEvaluating movies by average
rating\n')

/tmp/ipykernel_41716/3701943215.py:12: ParserWarning: Falling back to
the 'python' engine because the 'c' engine does not support regex
```

separators (separators > 1 char and different from '\s+' are interpreted as regex); you can avoid this warning by specifying engine='python'.

```
movies = pd.read_csv('movies.dat', sep='::', names=['MovieID',
'Title', 'Genre', 'Year'],
/tmp/ipykernel_41716/3701943215.py:14: ParserWarning: Falling back to
the 'python' engine because the 'c' engine does not support regex
separators (separators > 1 char and different from '\s+' are
interpreted as regex); you can avoid this warning by specifying
engine='python'.
```

```
ratings = pd.read_csv('ratings.dat', sep='::',
```

Removed 1047 movies because there aren't enough reviews

Evaluating movies by average rating

Genre: Action

```
Movie: {'MovieID': 2905, 'Title': 'Sanjuro ', 'Genre':
'Action|Adventure', 'Year': 1962, 'RatingCount': 69.0, 'AvgRating':
4.608695652173913}
```

```
Movie: {'MovieID': 2019, 'Title': 'Seven Samurai (The Magnificent
Seven) (Shichinin no samurai) ', 'Genre': 'Action|Drama', 'Year':
1954, 'RatingCount': 628.0, 'AvgRating': 4.560509554140127}
```

```
Movie: {'MovieID': 858, 'Title': 'Godfather, The ', 'Genre':
'Action|Crime|Drama', 'Year': 1972, 'RatingCount': 2223.0,
```

```
'AvgRating': 4.524966261808367}
Movie: {'MovieID': 1198, 'Title': 'Raiders of the Lost Ark ', 'Genre':
'Action|Adventure', 'Year': 1981, 'RatingCount': 2514.0, 'AvgRating':
4.477724741447892}
Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope
', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977,
'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}
```

Genre: Adventure

```
Movie: {'MovieID': 2905, 'Title': 'Sanjuro ', 'Genre':
'Action|Adventure', 'Year': 1962, 'RatingCount': 69.0, 'AvgRating':
4.608695652173913}
Movie: {'MovieID': 1198, 'Title': 'Raiders of the Lost Ark ', 'Genre':
'Action|Adventure', 'Year': 1981, 'RatingCount': 2514.0, 'AvgRating':
4.477724741447892}
Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope
', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977,
'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}
Movie: {'MovieID': 1204, 'Title': 'Lawrence of Arabia ', 'Genre':
'Adventure|War', 'Year': 1962, 'RatingCount': 831.0, 'AvgRating':
4.401925391095066}
Movie: {'MovieID': 1262, 'Title': 'Great Escape, The ', 'Genre':
'Adventure|War', 'Year': 1963, 'RatingCount': 696.0, 'AvgRating':
4.376436781609195}
```

Genre: Animation

```
Movie: {'MovieID': 745, 'Title': 'Close Shave, A ', 'Genre':
'Animation|Comedy|Thriller', 'Year': 1995, 'RatingCount': 657.0,
```

```
'AvgRating': 4.52054794520548}
Movie: {'MovieID': 1148, 'Title': 'Wrong Trousers, The ', 'Genre':
'Animation|Comedy', 'Year': 1993, 'RatingCount': 882.0, 'AvgRating':
4.507936507936508}
Movie: {'MovieID': 720, 'Title': 'Wallace & Gromit: The Best of
Aardman Animation ', 'Genre': 'Animation', 'Year': 1996,
'RatingCount': 438.0, 'AvgRating': 4.426940639269406}
Movie: {'MovieID': 1223, 'Title': 'Grand Day Out, A ', 'Genre':
'Animation|Comedy', 'Year': 1992, 'RatingCount': 473.0, 'AvgRating':
4.361522198731501}
Movie: {'MovieID': 3429, 'Title': 'Creature Comforts ', 'Genre':
'Animation|Comedy', 'Year': 1990, 'RatingCount': 274.0, 'AvgRating':
4.335766423357664}
```

Genre: Children's

```
Movie: {'MovieID': 919, 'Title': 'Wizard of Oz, The ', 'Genre':
"Adventure|Children's|Drama|Musical", 'Year': 1939, 'RatingCount':
1718.0, 'AvgRating': 4.247962747380675}
Movie: {'MovieID': 3114, 'Title': 'Toy Story 2 ', 'Genre':
"Animation|Children's|Comedy", 'Year': 1999, 'RatingCount': 1585.0,
'AvgRating': 4.218927444794953}
Movie: {'MovieID': 1, 'Title': 'Toy Story ', 'Genre':
"Animation|Children's|Comedy", 'Year': 1995, 'RatingCount': 2077.0,
'AvgRating': 4.146846413095811}
Movie: {'MovieID': 2761, 'Title': 'Iron Giant, The ', 'Genre':
"Animation|Children's", 'Year': 1999, 'RatingCount': 674.0,
'AvgRating': 4.0474777448071215}
Movie: {'MovieID': 1023, 'Title': 'Winnie the Pooh and the Blustery
Day ', 'Genre': "Animation|Children's", 'Year': 1968, 'RatingCount':
```


221.0, 'AvgRating': 3.986425339366516}

Genre: Comedy

Movie: {'MovieID': 745, 'Title': 'Close Shave, A ', 'Genre':
'Animation|Comedy|Thriller', 'Year': 1995, 'RatingCount': 657.0,
'AvgRating': 4.52054794520548}

Movie: {'MovieID': 1148, 'Title': 'Wrong Trousers, The ', 'Genre':
'Animation|Comedy', 'Year': 1993, 'RatingCount': 882.0, 'AvgRating':
4.507936507936508}

Movie: {'MovieID': 3030, 'Title': 'Yojimbo ', 'Genre':
'Comedy|Drama|Western', 'Year': 1961, 'RatingCount': 215.0,
'AvgRating': 4.404651162790698}

Movie: {'MovieID': 3307, 'Title': 'City Lights ', 'Genre':
'Comedy|Drama|Romance', 'Year': 1931, 'RatingCount': 271.0,
'AvgRating': 4.387453874538745}

Movie: {'MovieID': 3022, 'Title': 'General, The ', 'Genre': 'Comedy',
'Year': 1927, 'RatingCount': 206.0, 'AvgRating': 4.368932038834951}

Genre: Crime

Movie: {'MovieID': 858, 'Title': 'Godfather, The ', 'Genre':
'Action|Crime|Drama', 'Year': 1972, 'RatingCount': 2223.0,
'AvgRating': 4.524966261808367}

Movie: {'MovieID': 50, 'Title': 'Usual Suspects, The ', 'Genre':
'Crime|Thriller', 'Year': 1995, 'RatingCount': 1783.0, 'AvgRating':
4.517106001121705}

Movie: {'MovieID': 3435, 'Title': 'Double Indemnity ', 'Genre':
'Crime|Film-Noir', 'Year': 1944, 'RatingCount': 551.0, 'AvgRating':
4.415607985480944}

```
Movie: {'MovieID': 1221, 'Title': 'Godfather: Part II, The ', 'Genre':  
'Action|Crime|Drama', 'Year': 1974, 'RatingCount': 1692.0,  
'AvgRating': 4.357565011820331}  
Movie: {'MovieID': 1234, 'Title': 'Sting, The ', 'Genre':  
'Comedy|Crime', 'Year': 1973, 'RatingCount': 1049.0, 'AvgRating':  
4.320305052430887}
```

Genre: Documentary

```
Movie: {'MovieID': 1147, 'Title': 'When we were Kings ', 'Genre':  
'Documentary', 'Year': 1996, 'RatingCount': 277.0, 'AvgRating':  
4.32129963898917}  
Movie: {'MovieID': 3679, 'Title': 'Decline of Western Civilization,  
The ', 'Genre': 'Documentary', 'Year': 1981, 'RatingCount': 62.0,  
'AvgRating': 4.274193548387097}  
Movie: {'MovieID': 363, 'Title': 'Wonderful, Horrible Life of Leni  
Riefenstahl, The (Die Macht der Bilder) ', 'Genre': 'Documentary',  
'Year': 1993, 'RatingCount': 62.0, 'AvgRating': 4.258064516129032}  
Movie: {'MovieID': 1189, 'Title': 'Thin Blue Line, The ', 'Genre':  
'Documentary', 'Year': 1988, 'RatingCount': 269.0, 'AvgRating':  
4.25278810408922}  
Movie: {'MovieID': 2859, 'Title': 'Stop Making Sense ', 'Genre':  
'Documentary', 'Year': 1984, 'RatingCount': 239.0, 'AvgRating':  
4.2343096234309625}
```

Genre: Drama

```
Movie: {'MovieID': 2019, 'Title': 'Seven Samurai (The Magnificent  
Seven) (Shichinin no samurai) ', 'Genre': 'Action|Drama', 'Year':  
1954, 'RatingCount': 628.0, 'AvgRating': 4.560509554140127}
```

Movie: {'MovieID': 318, 'Title': 'Shawshank Redemption, The ',
 'Genre': 'Drama', 'Year': 1994, 'RatingCount': 2227.0, 'AvgRating':
 4.554557700942973}

Movie: {'MovieID': 858, 'Title': 'Godfather, The ', 'Genre':
 'Action|Crime|Drama', 'Year': 1972, 'RatingCount': 2223.0,
 'AvgRating': 4.524966261808367}

Movie: {'MovieID': 527, 'Title': 'Schindler's List ', 'Genre':
 'Drama|War', 'Year': 1993, 'RatingCount': 2304.0, 'AvgRating':
 4.510416666666667}

Movie: {'MovieID': 1178, 'Title': 'Paths of Glory ', 'Genre':
 'Drama|War', 'Year': 1957, 'RatingCount': 230.0, 'AvgRating':
 4.473913043478261}

Genre: Fantasy

Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope
 ', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977,
 'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}

Movie: {'MovieID': 1097, 'Title': 'E.T. the Extra-Terrestrial ',
 'Genre': 'Children's|Drama|Fantasy|Sci-Fi', 'Year': 1982,
 'RatingCount': 2269.0, 'AvgRating': 3.9651828999559275}

Movie: {'MovieID': 247, 'Title': 'Heavenly Creatures ', 'Genre':
 'Drama|Fantasy|Romance|Thriller', 'Year': 1994, 'RatingCount': 477.0,
 'AvgRating': 3.8658280922431865}

Movie: {'MovieID': 1073, 'Title': 'Willy Wonka and the Chocolate
 Factory ', 'Genre': 'Adventure|Children's|Comedy|Fantasy', 'Year':
 1971, 'RatingCount': 1313.0, 'AvgRating': 3.8613861386138613}

Movie: {'MovieID': 2797, 'Title': 'Big ', 'Genre': 'Comedy|Fantasy',
 'Year': 1988, 'RatingCount': 1491.0, 'AvgRating': 3.8558014755197854}

Genre: Film-Noir

Movie: {'MovieID': 922, 'Title': 'Sunset Blvd. (a.k.a. Sunset Boulevard) ', 'Genre': 'Film-Noir', 'Year': 1950, 'RatingCount': 470.0, 'AvgRating': 4.491489361702127}

Movie: {'MovieID': 3435, 'Title': 'Double Indemnity ', 'Genre': 'Crime|Film-Noir', 'Year': 1944, 'RatingCount': 551.0, 'AvgRating': 4.415607985480944}

Movie: {'MovieID': 913, 'Title': 'Maltese Falcon, The ', 'Genre': 'Film-Noir|Mystery', 'Year': 1941, 'RatingCount': 1043.0, 'AvgRating': 4.395973154362416}

Movie: {'MovieID': 1252, 'Title': 'Chinatown ', 'Genre': 'Film-Noir|Mystery|Thriller', 'Year': 1974, 'RatingCount': 1185.0, 'AvgRating': 4.339240506329114}

Movie: {'MovieID': 1267, 'Title': 'Manchurian Candidate, The ', 'Genre': 'Film-Noir|Thriller', 'Year': 1962, 'RatingCount': 765.0, 'AvgRating': 4.333333333333333}

Genre: Horror

Movie: {'MovieID': 1278, 'Title': 'Young Frankenstein ', 'Genre': 'Comedy|Horror', 'Year': 1974, 'RatingCount': 1193.0, 'AvgRating': 4.250628667225482}

Movie: {'MovieID': 1219, 'Title': 'Psycho ', 'Genre': 'Horror|Thriller', 'Year': 1960, 'RatingCount': 1263.0, 'AvgRating': 4.218527315914489}

Movie: {'MovieID': 1214, 'Title': 'Alien ', 'Genre': 'Action|Horror|Sci-Fi|Thriller', 'Year': 1979, 'RatingCount': 2024.0, 'AvgRating': 4.159584980237154}

Movie: {'MovieID': 1258, 'Title': 'Shining, The ', 'Genre': 'Horror',

```
'Year': 1980, 'RatingCount': 1087.0, 'AvgRating': 4.104875804967802}
Movie: {'MovieID': 1387, 'Title': 'Jaws ', 'Genre': 'Action|Horror',
'Year': 1975, 'RatingCount': 1697.0, 'AvgRating': 4.089569829110195}
```

Genre: Musical

```
Movie: {'MovieID': 899, 'Title': 'Singin' in the Rain ', 'Genre':
'Musical|Romance', 'Year': 1952, 'RatingCount': 751.0, 'AvgRating':
4.2836218375499335}
```

```
Movie: {'MovieID': 919, 'Title': 'Wizard of Oz, The ', 'Genre':
"Adventure|Children's|Drama|Musical", 'Year': 1939, 'RatingCount':
1718.0, 'AvgRating': 4.247962747380675}
```

```
Movie: {'MovieID': 1288, 'Title': 'This Is Spinal Tap ', 'Genre':
'Comedy|Drama|Musical', 'Year': 1984, 'RatingCount': 1118.0,
'AvgRating': 4.179785330948121}
```

```
Movie: {'MovieID': 1066, 'Title': 'Shall we Dance? ', 'Genre':
'Comedy|Musical|Romance', 'Year': 1937, 'RatingCount': 175.0,
'AvgRating': 4.1657142857142855}
```

```
Movie: {'MovieID': 914, 'Title': 'My Fair Lady ', 'Genre':
'Musical|Romance', 'Year': 1964, 'RatingCount': 636.0, 'AvgRating':
4.154088050314465}
```

Genre: Mystery

```
Movie: {'MovieID': 904, 'Title': 'Rear Window ', 'Genre':
'Mystery|Thriller', 'Year': 1954, 'RatingCount': 1050.0, 'AvgRating':
4.476190476190476}
```

```
Movie: {'MovieID': 1212, 'Title': 'Third Man, The ', 'Genre':
'Mystery|Thriller', 'Year': 1949, 'RatingCount': 480.0, 'AvgRating':
4.452083333333333}
```

```
Movie: {'MovieID': 913, 'Title': 'Maltese Falcon, The ', 'Genre':  
'Film-Noir|Mystery', 'Year': 1941, 'RatingCount': 1043.0, 'AvgRating':  
4.395973154362416}  
Movie: {'MovieID': 1252, 'Title': 'Chinatown ', 'Genre': 'Film-  
Noir|Mystery|Thriller', 'Year': 1974, 'RatingCount': 1185.0,  
'AvgRating': 4.339240506329114}  
Movie: {'MovieID': 1284, 'Title': 'Big Sleep, The ', 'Genre': 'Film-  
Noir|Mystery', 'Year': 1946, 'RatingCount': 541.0, 'AvgRating':  
4.312384473197782}
```

Genre: Romance

```
Movie: {'MovieID': 912, 'Title': 'Casablanca ', 'Genre':  
'Drama|Romance|War', 'Year': 1942, 'RatingCount': 1669.0, 'AvgRating':  
4.412822049131217}  
Movie: {'MovieID': 3307, 'Title': 'City Lights ', 'Genre':  
'Comedy|Drama|Romance', 'Year': 1931, 'RatingCount': 271.0,  
'AvgRating': 4.387453874538745}  
Movie: {'MovieID': 1197, 'Title': 'Princess Bride, The ', 'Genre':  
'Action|Adventure|Comedy|Romance', 'Year': 1987, 'RatingCount':  
2318.0, 'AvgRating': 4.3037100949094045}  
Movie: {'MovieID': 898, 'Title': 'Philadelphia Story, The ', 'Genre':  
'Comedy|Romance', 'Year': 1940, 'RatingCount': 582.0, 'AvgRating':  
4.3006872852233675}  
Movie: {'MovieID': 930, 'Title': 'Notorious ', 'Genre': 'Film-  
Noir|Romance|Thriller', 'Year': 1946, 'RatingCount': 445.0,  
'AvgRating': 4.29438202247191}
```

Genre: Sci-Fi

```

Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope
', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977,
'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}
Movie: {'MovieID': 750, 'Title': 'Dr. Strangelove or: How I Learned to
Stop Worrying and Love the Bomb ', 'Genre': 'Sci-Fi|War', 'Year':
1963, 'RatingCount': 1367.0, 'AvgRating': 4.4498902706656915}
Movie: {'MovieID': 2571, 'Title': 'Matrix, The ', 'Genre':
'Action|Sci-Fi|Thriller', 'Year': 1999, 'RatingCount': 2590.0,
'AvgRating': 4.315830115830116}
Movie: {'MovieID': 1196, 'Title': 'Star Wars: Episode V - The Empire
Strikes Back ', 'Genre': 'Action|Adventure|Drama|Sci-Fi|War', 'Year':
1980, 'RatingCount': 2990.0, 'AvgRating': 4.292976588628763}
Movie: {'MovieID': 541, 'Title': 'Blade Runner ', 'Genre': 'Film-
Noir|Sci-Fi', 'Year': 1982, 'RatingCount': 1800.0, 'AvgRating':
4.273333333333333}

```

Genre: Thriller

```

Movie: {'MovieID': 745, 'Title': 'Close Shave, A ', 'Genre':
'Animation|Comedy|Thriller', 'Year': 1995, 'RatingCount': 657.0,
'AvgRating': 4.52054794520548}
Movie: {'MovieID': 50, 'Title': 'Usual Suspects, The ', 'Genre':
'Crime|Thriller', 'Year': 1995, 'RatingCount': 1783.0, 'AvgRating':
4.517106001121705}
Movie: {'MovieID': 904, 'Title': 'Rear window ', 'Genre':
'Mystery|Thriller', 'Year': 1954, 'RatingCount': 1050.0, 'AvgRating':
4.476190476190476}
Movie: {'MovieID': 1212, 'Title': 'Third Man, The ', 'Genre':
'Mystery|Thriller', 'Year': 1949, 'RatingCount': 480.0, 'AvgRating':
4.452083333333333}

```

Movie: {'MovieID': 2762, 'Title': 'Sixth Sense, The ', 'Genre': 'Thriller', 'Year': 1999, 'RatingCount': 2459.0, 'AvgRating': 4.406262708418057}

Genre: War

Movie: {'MovieID': 527, 'Title': 'Schindler's List ', 'Genre': 'Drama|War', 'Year': 1993, 'RatingCount': 2304.0, 'AvgRating': 4.510416666666667}

Movie: {'MovieID': 1178, 'Title': 'Paths of Glory ', 'Genre': 'Drama|War', 'Year': 1957, 'RatingCount': 230.0, 'AvgRating': 4.473913043478261}

Movie: {'MovieID': 750, 'Title': 'Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb ', 'Genre': 'Sci-Fi|War', 'Year': 1963, 'RatingCount': 1367.0, 'AvgRating': 4.4498902706656915}

Movie: {'MovieID': 912, 'Title': 'Casablanca ', 'Genre': 'Drama|Romance|War', 'Year': 1942, 'RatingCount': 1669.0, 'AvgRating': 4.412822049131217}

Movie: {'MovieID': 1204, 'Title': 'Lawrence of Arabia ', 'Genre': 'Adventure|War', 'Year': 1962, 'RatingCount': 831.0, 'AvgRating': 4.401925391095066}

Genre: Western

Movie: {'MovieID': 3030, 'Title': 'Yojimbo ', 'Genre': 'Comedy|Drama|Western', 'Year': 1961, 'RatingCount': 215.0, 'AvgRating': 4.404651162790698}

Movie: {'MovieID': 1304, 'Title': 'Butch Cassidy and the Sundance Kid ', 'Genre': 'Action|Comedy|Western', 'Year': 1969, 'RatingCount': 1419.0, 'AvgRating': 4.215644820295983}


```
Movie: {'MovieID': 1283, 'Title': 'High Noon ', 'Genre': 'Western',  
       'Year': 1952, 'RatingCount': 403.0, 'AvgRating': 4.1786600496277915}  
Movie: {'MovieID': 1201, 'Title': 'Good, The Bad and The Ugly, The ',  
       'Genre': 'Action|Western', 'Year': 1966, 'RatingCount': 822.0,  
       'AvgRating': 4.133819951338199}  
Movie: {'MovieID': 1266, 'Title': 'Unforgiven ', 'Genre': 'Western',  
       'Year': 1992, 'RatingCount': 997.0, 'AvgRating': 4.0902708124373115}
```

Testing Algorithm 2 (Highest Rated and Recently Created)

This algorithm is identical to algorithm 1, but includes a condition to remove older movies. All movies created before 1995 are excluded. Since this algorithm is so similar to the first, we added an extra algorithm (algo 3) for this assignment.

Notice how the year is always above 1995.

In [6]:

```
# Get the top 5 recently created movies per genre by average rating  
(only movies after 1995)  
movies_df, ratings_df = load_data()
```

```
cleaned_movies_df_year = preprocess(movies_df, ratings_df,  
year_filter=1995)
```

```
genre_year_dict = get_top_by_genre(cleaned_movies_df_year,
    'AvgRating', top_n=5)
display_top_results(genre_year_dict, '\n\nEvaluating movies after 1995
by average rating\n')
```

```
/tmp/ipykernel_41716/3701943215.py:12: ParserWarning: Falling back to
the 'python' engine because the 'c' engine does not support regex
separators (separators > 1 char and different from '\s+' are
interpreted as regex); you can avoid this warning by specifying
engine='python'.
```

```
movies = pd.read_csv('movies.dat', sep='::', names=['MovieID',
'Title', 'Genre', 'Year'],
/tmp/ipykernel_41716/3701943215.py:14: ParserWarning: Falling back to
the 'python' engine because the 'c' engine does not support regex
separators (separators > 1 char and different from '\s+' are
interpreted as regex); you can avoid this warning by specifying
engine='python'.
```

```
ratings = pd.read_csv('ratings.dat', sep='::',
```

Removed 1047 movies because there aren't enough reviews

Evaluating movies after 1995 by average rating

Genre: Action

Movie: {'MovieID': 2028, 'Title': 'Saving Private Ryan ', 'Genre':

```
'Action|Drama|War', 'Year': 1998, 'RatingCount': 2653.0, 'AvgRating':  
4.337353938937053}  
Movie: {'MovieID': 2571, 'Title': 'Matrix, The ', 'Genre':  
'Action|Sci-Fi|Thriller', 'Year': 1999, 'RatingCount': 2590.0,  
'AvgRating': 4.315830115830116}  
Movie: {'MovieID': 110, 'Title': 'Braveheart ', 'Genre':  
'Action|Drama|War', 'Year': 1995, 'RatingCount': 2443.0, 'AvgRating':  
4.234957020057307}  
Movie: {'MovieID': 2692, 'Title': 'Run Lola Run (Lola rennt) ',  
'Genre': 'Action|Crime|Romance', 'Year': 1998, 'RatingCount': 1072.0,  
'AvgRating': 4.224813432835821}  
Movie: {'MovieID': 3000, 'Title': 'Princess Mononoke, The (Mononoke  
Hime) ', 'Genre': 'Action|Adventure|Animation', 'Year': 1997,  
'RatingCount': 345.0, 'AvgRating': 4.147826086956521}
```

Genre: Adventure

```
Movie: {'MovieID': 3000, 'Title': 'Princess Mononoke, The (Mononoke  
Hime) ', 'Genre': 'Action|Adventure|Animation', 'Year': 1997,  
'RatingCount': 345.0, 'AvgRating': 4.147826086956521}  
Movie: {'MovieID': 29, 'Title': 'City of Lost Children, The ',  
'Genre': 'Adventure|Sci-Fi', 'Year': 1995, 'RatingCount': 403.0,  
'AvgRating': 4.062034739454094}  
Movie: {'MovieID': 3175, 'Title': 'Galaxy Quest ', 'Genre':  
'Adventure|Comedy|Sci-Fi', 'Year': 1999, 'RatingCount': 1728.0,  
'AvgRating': 3.771412037037037}  
Movie: {'MovieID': 1580, 'Title': 'Men in Black ', 'Genre':  
'Action|Adventure|Comedy|Sci-Fi', 'Year': 1997, 'RatingCount': 2538.0,  
'AvgRating': 3.739952718676123}  
Movie: {'MovieID': 733, 'Title': 'Rock, The ', 'Genre':
```

```
'Action|Adventure|Thriller', 'Year': 1996, 'RatingCount': 1340.0,  
'AvgRating': 3.723134328358209}
```

Genre: Animation

```
Movie: {'MovieID': 745, 'Title': 'Close Shave, A ', 'Genre':  
'Animation|Comedy|Thriller', 'Year': 1995, 'RatingCount': 657.0,  
'AvgRating': 4.52054794520548}
```

```
Movie: {'MovieID': 720, 'Title': 'Wallace & Gromit: The Best of  
Aardman Animation ', 'Genre': 'Animation', 'Year': 1996,  
'RatingCount': 438.0, 'AvgRating': 4.426940639269406}
```

```
Movie: {'MovieID': 3114, 'Title': 'Toy Story 2 ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1999, 'RatingCount': 1585.0,  
'AvgRating': 4.218927444794953}
```

```
Movie: {'MovieID': 3000, 'Title': 'Princess Mononoke, The (Mononoke  
Hime) ', 'Genre': 'Action|Adventure|Animation', 'Year': 1997,  
'RatingCount': 345.0, 'AvgRating': 4.147826086956521}
```

```
Movie: {'MovieID': 1, 'Title': 'Toy Story ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1995, 'RatingCount': 2077.0,  
'AvgRating': 4.146846413095811}
```

Genre: Children's

```
Movie: {'MovieID': 3114, 'Title': 'Toy Story 2 ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1999, 'RatingCount': 1585.0,  
'AvgRating': 4.218927444794953}
```

```
Movie: {'MovieID': 1, 'Title': 'Toy Story ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1995, 'RatingCount': 2077.0,  
'AvgRating': 4.146846413095811}
```

```
Movie: {'MovieID': 2761, 'Title': 'Iron Giant, The ', 'Genre':
```

```
"Animation|Children's", 'Year': 1999, 'RatingCount': 674.0,  
'AvgRating': 4.0474777448071215}  
Movie: {'MovieID': 34, 'Title': 'Babe ', 'Genre':  
"Children's|Comedy|Drama", 'Year': 1995, 'RatingCount': 1751.0,  
'AvgRating': 3.8914905768132497}  
Movie: {'MovieID': 3751, 'Title': 'Chicken Run ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 2000, 'RatingCount': 1329.0,  
'AvgRating': 3.8796087283671934}
```

Genre: Comedy

```
Movie: {'MovieID': 745, 'Title': 'Close Shave, A ', 'Genre':  
'Animation|Comedy|Thriller', 'Year': 1995, 'RatingCount': 657.0,  
'AvgRating': 4.52054794520548}  
Movie: {'MovieID': 2324, 'Title': 'Life Is Beautiful (La vita è bella)  
, 'Genre': 'Comedy|Drama', 'Year': 1997, 'RatingCount': 1152.0,  
'AvgRating': 4.329861111111111}  
Movie: {'MovieID': 2858, 'Title': 'American Beauty ', 'Genre':  
'Comedy|Drama', 'Year': 1999, 'RatingCount': 3428.0, 'AvgRating':  
4.3173862310385065}  
Movie: {'MovieID': 3897, 'Title': 'Almost Famous ', 'Genre':  
'Comedy|Drama', 'Year': 2000, 'RatingCount': 994.0, 'AvgRating':  
4.22635814889336}  
Movie: {'MovieID': 3114, 'Title': 'Toy Story 2 ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1999, 'RatingCount': 1585.0,  
'AvgRating': 4.218927444794953}
```

Genre: Crime

```
Movie: {'MovieID': 50, 'Title': 'Usual Suspects, The ', 'Genre':
```

```
'Crime|Thriller', 'Year': 1995, 'RatingCount': 1783.0, 'AvgRating':
4.517106001121705}
Movie: {'MovieID': 608, 'Title': 'Fargo ', 'Genre':
'Crime|Drama|Thriller', 'Year': 1996, 'RatingCount': 2513.0,
'AvgRating': 4.254675686430561}
Movie: {'MovieID': 2692, 'Title': 'Run Lola Run (Lola rennt) ',
'Genre': 'Action|Crime|Romance', 'Year': 1998, 'RatingCount': 1072.0,
'AvgRating': 4.224813432835821}
Movie: {'MovieID': 1617, 'Title': 'L.A. Confidential ', 'Genre':
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1997, 'RatingCount':
2288.0, 'AvgRating': 4.219405594405594}
Movie: {'MovieID': 47, 'Title': 'Seven (Se7en) ', 'Genre':
'Crime|Thriller', 'Year': 1995, 'RatingCount': 1137.0, 'AvgRating':
4.106420404573439}
```

Genre: Documentary

```
Movie: {'MovieID': 1147, 'Title': 'When we were Kings ', 'Genre':
'Documentary', 'Year': 1996, 'RatingCount': 277.0, 'AvgRating':
4.32129963898917}
Movie: {'MovieID': 3077, 'Title': '42 Up ', 'Genre': 'Documentary',
'Year': 1998, 'RatingCount': 88.0, 'AvgRating': 4.2272727272727275}
Movie: {'MovieID': 1361, 'Title': 'Paradise Lost: The Child Murders at
Robin Hood Hills ', 'Genre': 'Documentary', 'Year': 1996,
'RatingCount': 159.0, 'AvgRating': 4.213836477987422}
Movie: {'MovieID': 2330, 'Title': 'Hands on a Hard Body ', 'Genre':
'Documentary', 'Year': 1996, 'RatingCount': 92.0, 'AvgRating':
4.163043478260869}
Movie: {'MovieID': 116, 'Title': 'Anne Frank Remembered ', 'Genre':
'Documentary', 'Year': 1995, 'RatingCount': 41.0, 'AvgRating':
```

4.097560975609756}

Genre: Drama

Movie: {'MovieID': 2028, 'Title': 'Saving Private Ryan ', 'Genre': 'Action|Drama|War', 'Year': 1998, 'RatingCount': 2653.0, 'AvgRating': 4.337353938937053}

Movie: {'MovieID': 2324, 'Title': 'Life Is Beautiful (La Vita è bella) ', 'Genre': 'Comedy|Drama', 'Year': 1997, 'RatingCount': 1152.0, 'AvgRating': 4.329861111111111}

Movie: {'MovieID': 2858, 'Title': 'American Beauty ', 'Genre': 'Comedy|Drama', 'Year': 1999, 'RatingCount': 3428.0, 'AvgRating': 4.3173862310385065}

Movie: {'MovieID': 2360, 'Title': 'Celebration, The (Festen) ', 'Genre': 'Drama', 'Year': 1998, 'RatingCount': 117.0, 'AvgRating': 4.3076923076923075}

Movie: {'MovieID': 2357, 'Title': 'Central Station (Central do Brasil) ', 'Genre': 'Drama', 'Year': 1998, 'RatingCount': 215.0, 'AvgRating': 4.283720930232558}

Genre: Fantasy

Movie: {'MovieID': 2628, 'Title': 'Star Wars: Episode I - The Phantom Menace ', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1999, 'RatingCount': 2250.0, 'AvgRating': 3.409777777777778}

Movie: {'MovieID': 2622, 'Title': 'Midsummer Night's Dream, A ', 'Genre': 'Comedy|Fantasy', 'Year': 1999, 'RatingCount': 415.0, 'AvgRating': 3.3493975903614457}

Movie: {'MovieID': 1654, 'Title': 'FairyTale: A True Story ', 'Genre': 'Children's|Drama|Fantasy', 'Year': 1997, 'RatingCount': 87.0,

```
'AvgRating': 3.310344827586207}
Movie: {'MovieID': 653, 'Title': 'Dragonheart ', 'Genre':
'Action|Adventure|Fantasy', 'Year': 1996, 'RatingCount': 612.0,
'AvgRating': 3.2254901960784315}
Movie: {'MovieID': 60, 'Title': 'Indian in the Cupboard, The ',
'Genre': "Adventure|Children's|Fantasy", 'Year': 1995, 'RatingCount':
357.0, 'AvgRating': 3.212885154061625}
```

Genre: Film-Noir

```
Movie: {'MovieID': 1617, 'Title': 'L.A. Confidential ', 'Genre':
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1997, 'RatingCount':
2288.0, 'AvgRating': 4.219405594405594}
Movie: {'MovieID': 1748, 'Title': 'Dark City ', 'Genre': 'Film-
Noir|Sci-Fi|Thriller', 'Year': 1998, 'RatingCount': 874.0,
'AvgRating': 3.822654462242563}
Movie: {'MovieID': 164, 'Title': 'Devil in a Blue Dress ', 'Genre':
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1995, 'RatingCount':
379.0, 'AvgRating': 3.5672823218997363}
Movie: {'MovieID': 707, 'Title': 'Mulholland Falls ', 'Genre':
'Crime|Film-Noir|Thriller', 'Year': 1996, 'RatingCount': 275.0,
'AvgRating': 3.0872727272727274}
Movie: {'MovieID': 1783, 'Title': 'Palmetto ', 'Genre': 'Film-
Noir|Mystery|Thriller', 'Year': 1998, 'RatingCount': 220.0,
'AvgRating': 2.9272727272727272}
```

Genre: Horror

```
Movie: {'MovieID': 1407, 'Title': 'Scream ', 'Genre':
'Horror|Thriller', 'Year': 1996, 'RatingCount': 886.0, 'AvgRating':
```


3.5756207674943568}

Movie: {'MovieID': 3081, 'Title': 'Sleepy Hollow ', 'Genre':
'Horror|Romance', 'Year': 1999, 'RatingCount': 1066.0, 'AvgRating':
3.516885553470919}

Movie: {'MovieID': 2167, 'Title': 'Blade ', 'Genre':
'Action|Adventure|Horror', 'Year': 1998, 'RatingCount': 809.0,
'AvgRating': 3.477132262051916}

Movie: {'MovieID': 1645, 'Title': 'Devil's Advocate, The ', 'Genre':
'Crime|Horror|Mystery|Thriller', 'Year': 1997, 'RatingCount': 826.0,
'AvgRating': 3.4358353510895885}

Movie: {'MovieID': 799, 'Title': 'Frighteners, The ', 'Genre':
'Comedy|Horror', 'Year': 1996, 'RatingCount': 338.0, 'AvgRating':
3.3491124260355027}

Genre: Musical

Movie: {'MovieID': 3910, 'Title': 'Dancer in the Dark ', 'Genre':
'Drama|Musical', 'Year': 2000, 'RatingCount': 200.0, 'AvgRating':
3.82}

Movie: {'MovieID': 73, 'Title': 'Misérables, Les ', 'Genre':
'Drama|Musical', 'Year': 1995, 'RatingCount': 224.0, 'AvgRating':
3.8169642857142856}

Movie: {'MovieID': 3159, 'Title': 'Fantasia 2000 ', 'Genre':
'Animation|Children's|Musical', 'Year': 1999, 'RatingCount': 453.0,
'AvgRating': 3.8057395143487858}

Movie: {'MovieID': 2394, 'Title': 'Prince of Egypt, The ', 'Genre':
'Animation|Musical', 'Year': 1998, 'RatingCount': 579.0, 'AvgRating':
3.5716753022452505}

Movie: {'MovieID': 1057, 'Title': 'Everyone Says I Love You ',
'Genre': 'Comedy|Musical|Romance', 'Year': 1996, 'RatingCount': 274.0,

'AvgRating': 3.5583941605839415}

Genre: Mystery

Movie: {'MovieID': 1617, 'Title': 'L.A. Confidential ', 'Genre':
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1997, 'RatingCount':
2288.0, 'AvgRating': 4.219405594405594}

Movie: {'MovieID': 800, 'Title': 'Lone Star ', 'Genre':
'Drama|Mystery', 'Year': 1996, 'RatingCount': 630.0, 'AvgRating':
4.082539682539682}

Movie: {'MovieID': 2686, 'Title': 'Red Violin, The (Le Violon rouge)
, 'Genre': 'Drama|Mystery', 'Year': 1998, 'RatingCount': 583.0,
'AvgRating': 3.9622641509433962}

Movie: {'MovieID': 2810, 'Title': 'Perfect Blue ', 'Genre':
'Animation|Mystery', 'Year': 1997, 'RatingCount': 41.0, 'AvgRating':
3.926829268292683}

Movie: {'MovieID': 1625, 'Title': 'Game, The ', 'Genre':
'Mystery|Thriller', 'Year': 1997, 'RatingCount': 882.0, 'AvgRating':
3.9138321995464853}

Genre: Romance

Movie: {'MovieID': 2692, 'Title': 'Run Lola Run (Lola rennt) ',
'Genre': 'Action|Crime|Romance', 'Year': 1998, 'RatingCount': 1072.0,
'AvgRating': 4.224813432835821}

Movie: {'MovieID': 2396, 'Title': 'Shakespeare in Love ', 'Genre':
'Comedy|Romance', 'Year': 1998, 'RatingCount': 2369.0, 'AvgRating':
4.127479949345715}

Movie: {'MovieID': 1046, 'Title': 'Beautiful Thing ', 'Genre':
'Drama|Romance', 'Year': 1996, 'RatingCount': 90.0, 'AvgRating':

4.066666666666666}

Movie: {'MovieID': 28, 'Title': 'Persuasion ', 'Genre': 'Romance',
'Year': 1995, 'RatingCount': 179.0, 'AvgRating': 4.055865921787709}

Movie: {'MovieID': 17, 'Title': 'Sense and Sensibility ', 'Genre':
'Drama|Romance', 'Year': 1995, 'RatingCount': 835.0, 'AvgRating':
4.027544910179641}

Genre: Sci-Fi

Movie: {'MovieID': 2571, 'Title': 'Matrix, The ', 'Genre':
'Action|Sci-Fi|Thriller', 'Year': 1999, 'RatingCount': 2590.0,
'AvgRating': 4.315830115830116}

Movie: {'MovieID': 741, 'Title': 'Ghost in the Shell (Kokaku kidotai)
, 'Genre': 'Animation|Sci-Fi', 'Year': 1995, 'RatingCount': 314.0,
'AvgRating': 4.089171974522293}

Movie: {'MovieID': 29, 'Title': 'City of Lost Children, The ',
'Genre': 'Adventure|Sci-Fi', 'Year': 1995, 'RatingCount': 403.0,
'AvgRating': 4.062034739454094}

Movie: {'MovieID': 2594, 'Title': 'Open Your Eyes (Abre los ojos) ',
'Genre': 'Drama|Romance|Sci-Fi', 'Year': 1997, 'RatingCount': 62.0,
'AvgRating': 3.9516129032258065}

Movie: {'MovieID': 32, 'Title': 'Twelve Monkeys ', 'Genre':
'Drama|Sci-Fi', 'Year': 1995, 'RatingCount': 1511.0, 'AvgRating':
3.945731303772336}

Genre: Thriller

Movie: {'MovieID': 745, 'Title': 'Close Shave, A ', 'Genre':
'Animation|Comedy|Thriller', 'Year': 1995, 'RatingCount': 657.0,
'AvgRating': 4.52054794520548}

Movie: {'MovieID': 50, 'Title': 'Usual Suspects, The ', 'Genre':
'Crime|Thriller', 'Year': 1995, 'RatingCount': 1783.0, 'AvgRating':
4.517106001121705}

Movie: {'MovieID': 2762, 'Title': 'Sixth Sense, The ', 'Genre':
'Thriller', 'Year': 1999, 'RatingCount': 2459.0, 'AvgRating':
4.406262708418057}

Movie: {'MovieID': 2571, 'Title': 'Matrix, The ', 'Genre':
'Action|Sci-Fi|Thriller', 'Year': 1999, 'RatingCount': 2590.0,
'AvgRating': 4.315830115830116}

Movie: {'MovieID': 608, 'Title': ' Fargo ', 'Genre':
'Crime|Drama|Thriller', 'Year': 1996, 'RatingCount': 2513.0,
'AvgRating': 4.254675686430561}

Genre: War

Movie: {'MovieID': 2028, 'Title': 'Saving Private Ryan ', 'Genre':
'Action|Drama|War', 'Year': 1998, 'RatingCount': 2653.0, 'AvgRating':
4.337353938937053}

Movie: {'MovieID': 110, 'Title': 'Braveheart ', 'Genre':
'Action|Drama|War', 'Year': 1995, 'RatingCount': 2443.0, 'AvgRating':
4.234957020057307}

Movie: {'MovieID': 41, 'Title': 'Richard III ', 'Genre': 'Drama|War',
'Year': 1995, 'RatingCount': 242.0, 'AvgRating': 3.958677685950413}

Movie: {'MovieID': 3746, 'Title': 'Butterfly (La Lengua de las
Mariposas) ', 'Genre': 'Drama|War', 'Year': 2000, 'RatingCount': 69.0,
'AvgRating': 3.8550724637681157}

Movie: {'MovieID': 2890, 'Title': 'Three Kings ', 'Genre':
'Drama|War', 'Year': 1999, 'RatingCount': 1021.0, 'AvgRating':
3.8070519098922624}

Genre: Western

Movie: {'MovieID': 714, 'Title': 'Dead Man ', 'Genre': 'Western',
'Year': 1995, 'RatingCount': 153.0, 'AvgRating': 3.888888888888889}

Movie: {'MovieID': 210, 'Title': 'Wild Bill ', 'Genre': 'Western',
'Year': 1995, 'RatingCount': 41.0, 'AvgRating': 3.1463414634146343}

Movie: {'MovieID': 303, 'Title': 'Quick and the Dead, The ', 'Genre':
'Action|Adventure|Western', 'Year': 1995, 'RatingCount': 418.0,
'AvgRating': 2.9760765550239237}

Movie: {'MovieID': 996, 'Title': 'Last Man Standing ', 'Genre':
'Action|Drama|Western', 'Year': 1996, 'RatingCount': 256.0,
'AvgRating': 2.90625}

Movie: {'MovieID': 2701, 'Title': 'Wild wild west ', 'Genre':
'Action|Sci-Fi|Western', 'Year': 1999, 'RatingCount': 902.0,
'AvgRating': 2.158536585365854}

Testing Algorithm 3 (Most Popular)

Most Popular represents the most rated movies per genre. To get this metric, we measure the quantity of reviews for each movie. So, the movies that have the most reviews are considered the most popular.

Notice how the RatingCount takes priority.

In [7]:

```
# Get the top 5 movies per genre by popularity (quantity of reviews)
movies_df, ratings_df = load_data()
```

```
cleaned_movies_df_pop = preprocess(movies_df, ratings_df)
genre_pop_dict = get_top_by_genre(cleaned_movies_df_pop,
    'RatingCount', top_n=5)
display_top_results(genre_pop_dict, '\n\nEvaluating movies by
popularity\n')
```

```
/tmp/ipykernel_41716/3701943215.py:12: ParserWarning: Falling back to
the 'python' engine because the 'c' engine does not support regex
separators (separators > 1 char and different from '\s+' are
interpreted as regex); you can avoid this warning by specifying
engine='python'.
```

```
movies = pd.read_csv('movies.dat', sep='::', names=['MovieID',
'Title', 'Genre', 'Year'],
/tmp/ipykernel_41716/3701943215.py:14: ParserWarning: Falling back to
the 'python' engine because the 'c' engine does not support regex
separators (separators > 1 char and different from '\s+' are
interpreted as regex); you can avoid this warning by specifying
engine='python'.
```

```
ratings = pd.read_csv('ratings.dat', sep='::',
```

Removed 1047 movies because there aren't enough reviews

Evaluating movies by popularity

Genre: Action

Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977, 'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}

Movie: {'MovieID': 1196, 'Title': 'Star Wars: Episode V - The Empire Strikes Back', 'Genre': 'Action|Adventure|Drama|Sci-Fi|War', 'Year': 1980, 'RatingCount': 2990.0, 'AvgRating': 4.292976588628763}

Movie: {'MovieID': 1210, 'Title': 'Star Wars: Episode VI - Return of the Jedi', 'Genre': 'Action|Adventure|Romance|Sci-Fi|War', 'Year': 1983, 'RatingCount': 2883.0, 'AvgRating': 4.022892819979188}

Movie: {'MovieID': 480, 'Title': 'Jurassic Park', 'Genre': 'Action|Adventure|Sci-Fi', 'Year': 1993, 'RatingCount': 2672.0, 'AvgRating': 3.7638473053892216}

Movie: {'MovieID': 2028, 'Title': 'Saving Private Ryan', 'Genre': 'Action|Drama|War', 'Year': 1998, 'RatingCount': 2653.0, 'AvgRating': 4.337353938937053}

Genre: Adventure

Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977, 'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}

Movie: {'MovieID': 1196, 'Title': 'Star Wars: Episode V - The Empire Strikes Back', 'Genre': 'Action|Adventure|Drama|Sci-Fi|War', 'Year': 1980, 'RatingCount': 2990.0, 'AvgRating': 4.292976588628763}

Movie: {'MovieID': 1210, 'Title': 'Star Wars: Episode VI - Return of the Jedi', 'Genre': 'Action|Adventure|Romance|Sci-Fi|War', 'Year': 1983, 'RatingCount': 2883.0, 'AvgRating': 4.022892819979188}

Movie: {'MovieID': 480, 'Title': 'Jurassic Park', 'Genre': 'Action|Adventure|Sci-Fi', 'Year': 1993, 'RatingCount': 2672.0,

```
'AvgRating': 3.7638473053892216}  
Movie: {'MovieID': 1580, 'Title': 'Men in Black ', 'Genre':  
'Action|Adventure|Comedy|Sci-Fi', 'Year': 1997, 'RatingCount': 2538.0,  
'AvgRating': 3.739952718676123}
```

Genre: Animation

```
Movie: {'MovieID': 1, 'Title': 'Toy Story ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1995, 'RatingCount': 2077.0,  
'AvgRating': 4.146846413095811}  
Movie: {'MovieID': 2987, 'Title': 'Who Framed Roger Rabbit? ',  
'Genre': 'Adventure|Animation|Film-Noir', 'Year': 1988, 'RatingCount':  
1799.0, 'AvgRating': 3.6798221234018897}  
Movie: {'MovieID': 2355, 'Title': "Bug's Life, A ", 'Genre':  
"Animation|Children's|Comedy", 'Year': 1998, 'RatingCount': 1703.0,  
'AvgRating': 3.854374633000587}  
Movie: {'MovieID': 3114, 'Title': 'Toy Story 2 ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1999, 'RatingCount': 1585.0,  
'AvgRating': 4.218927444794953}  
Movie: {'MovieID': 588, 'Title': 'Aladdin ', 'Genre':  
"Animation|Children's|Comedy|Musical", 'Year': 1992, 'RatingCount':  
1351.0, 'AvgRating': 3.788304959289415}
```

Genre: Children's

```
Movie: {'MovieID': 1097, 'Title': 'E.T. the Extra-Terrestrial ',  
'Genre': "Children's|Drama|Fantasy|Sci-Fi", 'Year': 1982,  
'RatingCount': 2269.0, 'AvgRating': 3.9651828999559275}  
Movie: {'MovieID': 1, 'Title': 'Toy Story ', 'Genre':  
"Animation|Children's|Comedy", 'Year': 1995, 'RatingCount': 2077.0,
```



```
'AvgRating': 4.146846413095811}
Movie: {'MovieID': 34, 'Title': 'Babe ', 'Genre':
"Children's|Comedy|Drama", 'Year': 1995, 'RatingCount': 1751.0,
'AvgRating': 3.8914905768132497}
Movie: {'MovieID': 919, 'Title': 'Wizard of Oz, The ', 'Genre':
"Adventure|Children's|Drama|Musical", 'Year': 1939, 'RatingCount':
1718.0, 'AvgRating': 4.247962747380675}
Movie: {'MovieID': 2355, 'Title': "Bug's Life, A ", 'Genre':
"Animation|Children's|Comedy", 'Year': 1998, 'RatingCount': 1703.0,
'AvgRating': 3.854374633000587}
```

Genre: Comedy

```
Movie: {'MovieID': 2858, 'Title': 'American Beauty ', 'Genre':
'Comedy|Drama', 'Year': 1999, 'RatingCount': 3428.0, 'AvgRating':
4.3173862310385065}
Movie: {'MovieID': 1270, 'Title': 'Back to the Future ', 'Genre':
'Comedy|Sci-Fi', 'Year': 1985, 'RatingCount': 2583.0, 'AvgRating':
3.9903213317847466}
Movie: {'MovieID': 1580, 'Title': 'Men in Black ', 'Genre':
'Action|Adventure|Comedy|Sci-Fi', 'Year': 1997, 'RatingCount': 2538.0,
'AvgRating': 3.739952718676123}
Movie: {'MovieID': 2396, 'Title': 'Shakespeare in Love ', 'Genre':
'Comedy|Romance', 'Year': 1998, 'RatingCount': 2369.0, 'AvgRating':
4.127479949345715}
Movie: {'MovieID': 1197, 'Title': 'Princess Bride, The ', 'Genre':
'Action|Adventure|Comedy|Romance', 'Year': 1987, 'RatingCount':
2318.0, 'AvgRating': 4.3037100949094045}
```

Genre: Crime

Movie: {'MovieID': 608, 'Title': 'Fargo ', 'Genre':
'Crime|Drama|Thriller', 'Year': 1996, 'RatingCount': 2513.0,
'AvgRating': 4.254675686430561}

Movie: {'MovieID': 1617, 'Title': 'L.A. Confidential ', 'Genre':
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1997, 'RatingCount':
2288.0, 'AvgRating': 4.219405594405594}

Movie: {'MovieID': 858, 'Title': 'Godfather, The ', 'Genre':
'Action|Crime|Drama', 'Year': 1972, 'RatingCount': 2223.0,
'AvgRating': 4.524966261808367}

Movie: {'MovieID': 296, 'Title': 'Pulp Fiction ', 'Genre':
'Crime|Drama', 'Year': 1994, 'RatingCount': 2171.0, 'AvgRating':
4.278212805158913}

Movie: {'MovieID': 50, 'Title': 'Usual Suspects, The ', 'Genre':
'Crime|Thriller', 'Year': 1995, 'RatingCount': 1783.0, 'AvgRating':
4.517106001121705}

Genre: Documentary

Movie: {'MovieID': 2064, 'Title': 'Roger & Me ', 'Genre':
'Comedy|Documentary', 'Year': 1989, 'RatingCount': 798.0, 'AvgRating':
4.0739348370927315}

Movie: {'MovieID': 246, 'Title': 'Hoop Dreams ', 'Genre':
'Documentary', 'Year': 1994, 'RatingCount': 716.0, 'AvgRating':
4.18854748603352}

Movie: {'MovieID': 162, 'Title': 'Crumb ', 'Genre': 'Documentary',
'Year': 1994, 'RatingCount': 491.0, 'AvgRating': 4.063136456211812}

Movie: {'MovieID': 3007, 'Title': 'American Movie ', 'Genre':
'Documentary', 'Year': 1999, 'RatingCount': 295.0, 'AvgRating':
4.013559322033898}

Movie: {'MovieID': 1147, 'Title': 'When we were Kings ', 'Genre': 'Documentary', 'Year': 1996, 'RatingCount': 277.0, 'AvgRating': 4.32129963898917}

Genre: Drama

Movie: {'MovieID': 2858, 'Title': 'American Beauty ', 'Genre': 'Comedy|Drama', 'Year': 1999, 'RatingCount': 3428.0, 'AvgRating': 4.3173862310385065}

Movie: {'MovieID': 1196, 'Title': 'Star Wars: Episode V - The Empire Strikes Back ', 'Genre': 'Action|Adventure|Drama|Sci-Fi|War', 'Year': 1980, 'RatingCount': 2990.0, 'AvgRating': 4.292976588628763}

Movie: {'MovieID': 2028, 'Title': 'Saving Private Ryan ', 'Genre': 'Action|Drama|War', 'Year': 1998, 'RatingCount': 2653.0, 'AvgRating': 4.337353938937053}

Movie: {'MovieID': 593, 'Title': 'Silence of the Lambs, The ', 'Genre': 'Drama|Thriller', 'Year': 1991, 'RatingCount': 2578.0, 'AvgRating': 4.3518231186966645}

Movie: {'MovieID': 608, 'Title': ' Fargo ', 'Genre': 'Crime|Drama|Thriller', 'Year': 1996, 'RatingCount': 2513.0, 'AvgRating': 4.254675686430561}

Genre: Fantasy

Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope ', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977, 'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}

Movie: {'MovieID': 1097, 'Title': 'E.T. the Extra-Terrestrial ', 'Genre': 'Children's|Drama|Fantasy|Sci-Fi', 'Year': 1982, 'RatingCount': 2269.0, 'AvgRating': 3.9651828999559275}

```
Movie: {'MovieID': 2628, 'Title': 'Star Wars: Episode I - The Phantom  
Menace ', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1999,  
'RatingCount': 2250.0, 'AvgRating': 3.409777777777778}  
Movie: {'MovieID': 2174, 'Title': 'Beetlejuice ', 'Genre':  
'Comedy|Fantasy', 'Year': 1988, 'RatingCount': 1495.0, 'AvgRating':  
3.5678929765886287}  
Movie: {'MovieID': 2797, 'Title': 'Big ', 'Genre': 'Comedy|Fantasy',  
'Year': 1988, 'RatingCount': 1491.0, 'AvgRating': 3.8558014755197854}
```

Genre: Film-Noir

```
Movie: {'MovieID': 1617, 'Title': 'L.A. Confidential ', 'Genre':  
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1997, 'RatingCount':  
2288.0, 'AvgRating': 4.219405594405594}  
Movie: {'MovieID': 541, 'Title': 'Blade Runner ', 'Genre': 'Film-  
Noir|Sci-Fi', 'Year': 1982, 'RatingCount': 1800.0, 'AvgRating':  
4.273333333333333}  
Movie: {'MovieID': 2987, 'Title': 'Who Framed Roger Rabbit? ',  
'Genre': 'Adventure|Animation|Film-Noir', 'Year': 1988, 'RatingCount':  
1799.0, 'AvgRating': 3.6798221234018897}  
Movie: {'MovieID': 1252, 'Title': 'Chinatown ', 'Genre': 'Film-  
Noir|Mystery|Thriller', 'Year': 1974, 'RatingCount': 1185.0,  
'AvgRating': 4.339240506329114}  
Movie: {'MovieID': 913, 'Title': 'Maltese Falcon, The ', 'Genre':  
'Film-Noir|Mystery', 'Year': 1941, 'RatingCount': 1043.0, 'AvgRating':  
4.395973154362416}
```

Genre: Horror

```
Movie: {'MovieID': 2716, 'Title': 'Ghostbusters ', 'Genre':
```

```
'Comedy|Horror', 'Year': 1984, 'RatingCount': 2181.0, 'AvgRating':  
3.9055479138010085}  
Movie: {'MovieID': 1214, 'Title': 'Alien ', 'Genre':  
'Action|Horror|Sci-Fi|Thriller', 'Year': 1979, 'RatingCount': 2024.0,  
'AvgRating': 4.159584980237154}  
Movie: {'MovieID': 1387, 'Title': 'Jaws ', 'Genre': 'Action|Horror',  
'Year': 1975, 'RatingCount': 1697.0, 'AvgRating': 4.089569829110195}  
Movie: {'MovieID': 1219, 'Title': 'Psycho ', 'Genre':  
'Horror|Thriller', 'Year': 1960, 'RatingCount': 1263.0, 'AvgRating':  
4.218527315914489}  
Movie: {'MovieID': 2710, 'Title': 'Blair Witch Project, The ',  
'Genre': 'Horror', 'Year': 1999, 'RatingCount': 1237.0, 'AvgRating':  
3.0315278900565885}
```

Genre: Musical

```
Movie: {'MovieID': 919, 'Title': 'Wizard of Oz, The ', 'Genre':  
"Adventure|Children's|Drama|Musical", 'Year': 1939, 'RatingCount':  
1718.0, 'AvgRating': 4.247962747380675}  
Movie: {'MovieID': 588, 'Title': 'Aladdin ', 'Genre':  
"Animation|Children's|Comedy|Musical", 'Year': 1992, 'RatingCount':  
1351.0, 'AvgRating': 3.788304959289415}  
Movie: {'MovieID': 1220, 'Title': 'Blues Brothers, The ', 'Genre':  
'Action|Comedy|Musical', 'Year': 1980, 'RatingCount': 1341.0,  
'AvgRating': 3.9395973154362416}  
Movie: {'MovieID': 2657, 'Title': 'Rocky Horror Picture Show, The ',  
'Genre': 'Comedy|Horror|Musical|Sci-Fi', 'Year': 1975, 'RatingCount':  
1233.0, 'AvgRating': 3.291159772911598}  
Movie: {'MovieID': 364, 'Title': 'Lion King, The ', 'Genre':  
"Animation|Children's|Musical", 'Year': 1994, 'RatingCount': 1121.0,
```

'AvgRating': 3.8608385370205176}

Genre: Mystery

Movie: {'MovieID': 1617, 'Title': 'L.A. Confidential ', 'Genre':
'Crime|Film-Noir|Mystery|Thriller', 'Year': 1997, 'RatingCount':
2288.0, 'AvgRating': 4.219405594405594}

Movie: {'MovieID': 924, 'Title': '2001: A Space Odyssey ', 'Genre':
'Drama|Mystery|Sci-Fi|Thriller', 'Year': 1968, 'RatingCount': 1716.0,
'AvgRating': 4.068764568764569}

Movie: {'MovieID': 648, 'Title': 'Mission: Impossible ', 'Genre':
'Action|Adventure|Mystery', 'Year': 1996, 'RatingCount': 1527.0,
'AvgRating': 3.432874918140144}

Movie: {'MovieID': 3176, 'Title': 'Talented Mr. Ripley, The ',
'Genre': 'Drama|Mystery|Thriller', 'Year': 1999, 'RatingCount':
1331.0, 'AvgRating': 3.503380916604057}

Movie: {'MovieID': 1252, 'Title': 'Chinatown ', 'Genre': 'Film-
Noir|Mystery|Thriller', 'Year': 1974, 'RatingCount': 1185.0,
'AvgRating': 4.339240506329114}

Genre: Romance

Movie: {'MovieID': 1210, 'Title': 'Star Wars: Episode VI - Return of
the Jedi ', 'Genre': 'Action|Adventure|Romance|Sci-Fi|War', 'Year':
1983, 'RatingCount': 2883.0, 'AvgRating': 4.022892819979188}

Movie: {'MovieID': 2396, 'Title': 'Shakespeare in Love ', 'Genre':
'Comedy|Romance', 'Year': 1998, 'RatingCount': 2369.0, 'AvgRating':
4.127479949345715}

Movie: {'MovieID': 1197, 'Title': 'Princess Bride, The ', 'Genre':
'Action|Adventure|Comedy|Romance', 'Year': 1987, 'RatingCount':

```
2318.0, 'AvgRating': 4.3037100949094045}
Movie: {'MovieID': 1265, 'Title': 'Groundhog Day ', 'Genre':
'Comedy|Romance', 'Year': 1993, 'RatingCount': 2278.0, 'AvgRating':
3.953028972783143}
Movie: {'MovieID': 356, 'Title': 'Forrest Gump ', 'Genre':
'Comedy|Romance|War', 'Year': 1994, 'RatingCount': 2194.0,
'AvgRating': 4.087967183226983}
```

Genre: Sci-Fi

```
Movie: {'MovieID': 260, 'Title': 'Star Wars: Episode IV - A New Hope
', 'Genre': 'Action|Adventure|Fantasy|Sci-Fi', 'Year': 1977,
'RatingCount': 2991.0, 'AvgRating': 4.453694416583082}
Movie: {'MovieID': 1196, 'Title': 'Star Wars: Episode V - The Empire
Strikes Back ', 'Genre': 'Action|Adventure|Drama|Sci-Fi|War', 'Year':
1980, 'RatingCount': 2990.0, 'AvgRating': 4.292976588628763}
Movie: {'MovieID': 1210, 'Title': 'Star Wars: Episode VI - Return of
the Jedi ', 'Genre': 'Action|Adventure|Romance|Sci-Fi|War', 'Year':
1983, 'RatingCount': 2883.0, 'AvgRating': 4.022892819979188}
Movie: {'MovieID': 480, 'Title': 'Jurassic Park ', 'Genre':
'Action|Adventure|Sci-Fi', 'Year': 1993, 'RatingCount': 2672.0,
'AvgRating': 3.7638473053892216}
Movie: {'MovieID': 589, 'Title': 'Terminator 2: Judgment Day ',
'Genre': 'Action|Sci-Fi|Thriller', 'Year': 1991, 'RatingCount':
2649.0, 'AvgRating': 4.058512646281616}
```

Genre: Thriller

```
Movie: {'MovieID': 589, 'Title': 'Terminator 2: Judgment Day ',
'Genre': 'Action|Sci-Fi|Thriller', 'Year': 1991, 'RatingCount':
```

```
2649.0, 'AvgRating': 4.058512646281616}
Movie: {'MovieID': 2571, 'Title': 'Matrix, The ', 'Genre':
'Action|Sci-Fi|Thriller', 'Year': 1999, 'RatingCount': 2590.0,
'AvgRating': 4.315830115830116}
Movie: {'MovieID': 593, 'Title': 'Silence of the Lambs, The ',
'Genre': 'Drama|Thriller', 'Year': 1991, 'RatingCount': 2578.0,
'AvgRating': 4.3518231186966645}
Movie: {'MovieID': 608, 'Title': 'Fargo ', 'Genre':
'Crime|Drama|Thriller', 'Year': 1996, 'RatingCount': 2513.0,
'AvgRating': 4.254675686430561}
Movie: {'MovieID': 2762, 'Title': 'Sixth Sense, The ', 'Genre':
'Thriller', 'Year': 1999, 'RatingCount': 2459.0, 'AvgRating':
4.406262708418057}
```

Genre: War

```
Movie: {'MovieID': 1196, 'Title': 'Star Wars: Episode V - The Empire
Strikes Back ', 'Genre': 'Action|Adventure|Drama|Sci-Fi|War', 'Year':
1980, 'RatingCount': 2990.0, 'AvgRating': 4.292976588628763}
Movie: {'MovieID': 1210, 'Title': 'Star Wars: Episode VI - Return of
the Jedi ', 'Genre': 'Action|Adventure|Romance|Sci-Fi|War', 'Year':
1983, 'RatingCount': 2883.0, 'AvgRating': 4.022892819979188}
Movie: {'MovieID': 2028, 'Title': 'Saving Private Ryan ', 'Genre':
'Action|Drama|War', 'Year': 1998, 'RatingCount': 2653.0, 'AvgRating':
4.337353938937053}
Movie: {'MovieID': 110, 'Title': 'Braveheart ', 'Genre':
'Action|Drama|War', 'Year': 1995, 'RatingCount': 2443.0, 'AvgRating':
4.234957020057307}
Movie: {'MovieID': 527, 'Title': 'Schindler's List ', 'Genre':
'Drama|War', 'Year': 1993, 'RatingCount': 2304.0, 'AvgRating':
```


4.510416666666667}

Genre: Western

Movie: {'MovieID': 590, 'Title': 'Dances with wolves ', 'Genre':
'Adventure|Drama|Western', 'Year': 1990, 'RatingCount': 1451.0,
'AvgRating': 3.9152308752584424}

Movie: {'MovieID': 1304, 'Title': 'Butch Cassidy and the Sundance Kid
, 'Genre': 'Action|Comedy|Western', 'Year': 1969, 'RatingCount':
1419.0, 'AvgRating': 4.215644820295983}

Movie: {'MovieID': 2012, 'Title': 'Back to the Future Part III ',
'Genre': 'Comedy|Sci-Fi|Western', 'Year': 1990, 'RatingCount': 1148.0,
'AvgRating': 3.2421602787456445}

Movie: {'MovieID': 3671, 'Title': 'Blazing Saddles ', 'Genre':
'Comedy|Western', 'Year': 1974, 'RatingCount': 1119.0, 'AvgRating':
4.047363717605005}

Movie: {'MovieID': 1266, 'Title': 'Unforgiven ', 'Genre': 'Western',
'Year': 1992, 'RatingCount': 997.0, 'AvgRating': 4.0902708124373115}

Part 2: User-based and Item-based Collaborative Filtering

2022-12-12

UBCF

#Initial Code

`library(recommenderlab)``## warning: package 'recommenderlab' was built under R version 4.2.2``## Loading required package: Matrix``## Loading required package: arules``## warning: package 'arules' was built under R version 4.2.2``##``## Attaching package: 'arules'``## The following objects are masked from 'package:base':``##
abbreviate, write``## Loading required package: proxy``## warning: package 'proxy' was built under R version 4.2.2``##``## Attaching package: 'proxy'``## The following object is masked from 'package:Matrix':``##`

```
##      as.matrix

## The following objects are masked from 'package:stats':
##
##      as.dist, dist

## The following object is masked from 'package:base':
##
##      as.matrix

## Loading required package: registry

## Registered S3 methods overwritten by 'registry':
##      method          from
##      print.registry_field proxy
##      print.registry_entry proxy

myurl = "https://liangfgithub.github.io/MovieData/"
ratings = read.csv(paste0(myurl, 'ratings.dat?raw=true'),
                  sep = ': ',
                  colClasses = c('integer', 'NULL'),
                  header = FALSE)
colnames(ratings) = c('UserID', 'MovieID', 'Rating', 'Timestamp')
i = paste0('u', ratings$UserID)
j = paste0('m', ratings$MovieID)
x = ratings$Rating
tmp = data.frame(i, j, x, stringsAsFactors = T)
Rmat = sparseMatrix(as.integer(tmp$i), as.integer(tmp$j), x = tmp$x)
rownames(Rmat) = levels(tmp$i)
colnames(Rmat) = levels(tmp$j)
```

```
Rmat = new('realRatingMatrix', data = Rmat)

train = Rmat[1:500, ]
test = Rmat[501, ]

#normalize
newdata = as(test, "matrix")
data = as(train, "matrix")

user.means = rowMeans(data, na.rm = TRUE)
data = data - user.means

newuser.means = rowMeans(newdata, na.rm = TRUE)
newdata = newdata - newuser.means

#Cosine Similarity
sim = rep(0, dim(data)[1])
for(i in 1:length(sim)) {
  tmp.y = as.vector(newdata)
  ind.y = which(!is.na(tmp.y))
  tmp.x = data[i, ]
  ind.x = which(!is.na(tmp.x))
  ind = intersect(ind.x, ind.y)
  if (length(ind) > 0){
    tmp.x = tmp.x[ind]
    tmp.y = tmp.y[ind]
    sim[i] = sum(tmp.x * tmp.y) / sqrt(sum(tmp.x^2) * sum(tmp.y^2))
  }
}
sim = (1 + sim)/2
```

```
#####
sort = tail(sort(sim), 20)
val = sort[1]
sim[sim < val] = 0

mypred = colSums(data * sim, na.rm = TRUE) / colSums(!is.na(data)) *
        sim)
mypred = mypred + newuser.means

mypred[is.na(newdata)] = NA
mypred[is.infinite(mypred)] = NA
mypred[is.nan(mypred)] = NA

# UBCF
recommender.UBCF <- Recommender(train, method = "UBCF",
                                parameter = list(normalize = 'center',
                                                  method = 'Cosine',
                                                  nn = 20))

p.UBCF <- predict(recommender.UBCF, test, type="ratings")
p.UBCF <- as.numeric(as(p.UBCF, "matrix"))

sum(is.na(p.UBCF) != is.na(mypred)) ### should be zero

## [1] 0

max(abs(p.UBCF - mypred), na.rm = TRUE) ### should be less than 1e-06

## [1] 4.953572e-07
```

IBCF

#Initial Code

```
library(recommenderlab)
myurl = "https://liangfgithub.github.io/MovieData/"
ratings = read.csv(paste0(myurl, 'ratings.dat?raw=true'),
                   sep = ': ',
                   colClasses = c('integer', 'NULL'),
                   header = FALSE)
colnames(ratings) = c('UserID', 'MovieID', 'Rating', 'Timestamp')
i = paste0('u', ratings$UserID)
j = paste0('m', ratings$MovieID)
x = ratings$Rating
tmp = data.frame(i, j, x, stringsAsFactors = T)
Rmat = sparseMatrix(as.integer(tmp$i), as.integer(tmp$j), x = tmp$x)
rownames(Rmat) = levels(tmp$i)
colnames(Rmat) = levels(tmp$j)
Rmat = new('realRatingMatrix', data = Rmat)
train = Rmat[1:500, ]
test = Rmat[501, ]
```

#normalize

```
newdata = as(test, "matrix")
data = as(train, "matrix")
user.means = rowMeans(data, na.rm = TRUE)
data = data - user.means
```

#Cosine Similarity

```

sim = proxy::simil(data, method = "cosine", by_rows = FALSE)
sim = (1 + sim)/2
sim_mat = as.matrix(sim)

for(i in 1:ncol(data)){
  ord1 = order(sim_mat[i, ], decreasing = FALSE, na.last = FALSE)
  ord = tail(ord1, 30)
  sim_mat[i, -ord] = NA
}

#####
nonNA = which(!is.na(newdata))
ent = as.vector(newdata[nonNA])
sim = t(sim_mat[, nonNA])
mypred = colSums(sim * ent, na.rm = T)/rowSums(sim_mat[, nonNA], na.rm
      = T)

mypred = as(mypred, "matrix")
mypred[!is.na(newdata)] = NA
mypred[is.infinite(mypred)] = NA
mypred[is.nan(mypred)]=NA

# IBCF
recommender.IBCF <- Recommender(train, method = "IBCF",
                                parameter = list(normalize = 'center',
                                                    method = 'Cosine',
                                                    k = 30))

p.IBCF <- predict(recommender.IBCF, test, type="ratings")
p.IBCF <- as.numeric(as(p.IBCF, "matrix"))

```

```
## first output: should be less than 10
sum(is.na(p.IBCF) != is.na(mypred))

## [1] 1

## second output: should be less than 10%
mydiff = abs(p.IBCF - mypred)
sum(mydiff[!is.na(mydiff)] > 1e-6) / sum(!is.na(mydiff))

## [1] 0.08243052
```

