



# Netflix Movie Recommendation System



## What is recommendation system? Why important?

- Recommended systems are one of the most common used and easily understandable applications of data science. Recommender systems are algorithms aimed at suggesting relevant items to users.
- These systems help users find items they want to buy from a business. Recommender systems benefit users by enabling them to find items they like. Additionally, they help businesses by generating more sales.

# Netflix Prize Data Set

	movie_id	year	name
0	1	2003	Dinosaur Planet
1	2	2004	Isle of Man TT 2004 Review
2	3	1997	Character
3	4	1994	Paula Abdul's Get Up & Dance
4	5	2004	The Rise and Fall of ECW

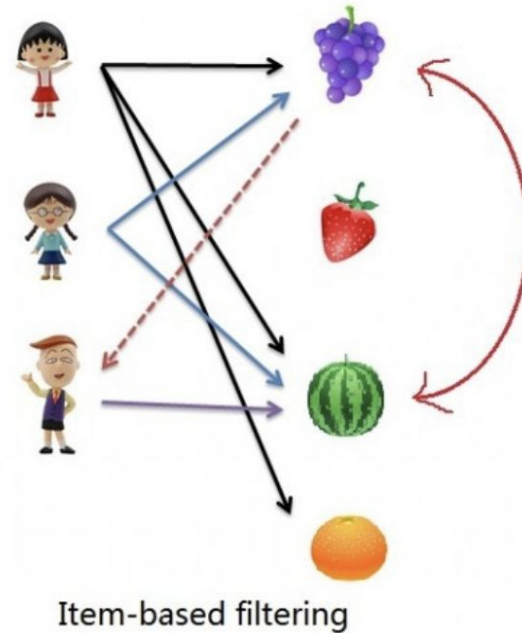
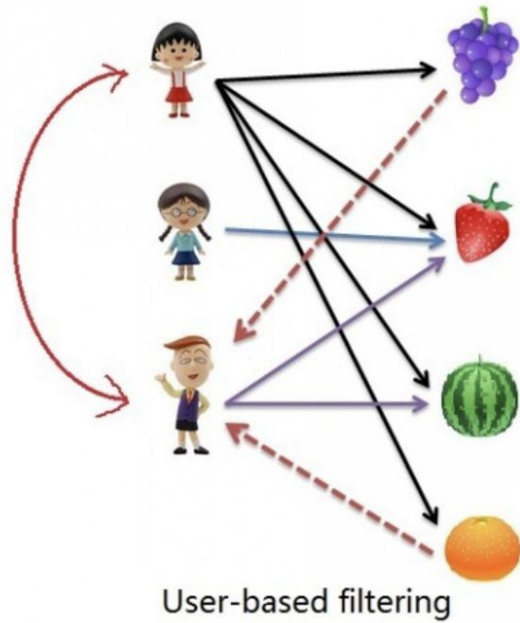
	movie_id	user_id	rating	date
100480502	17770	1790158	4	2005-11-01
100480503	17770	1608708	3	2005-07-19
100480504	17770	234275	1	2004-08-07
100480505	17770	255278	4	2004-05-28
100480506	17770	453585	2	2005-03-10

There are 480,189 users and 17,700 movies.  
There are a total of 100,480,507 ratings.

# **There are 2 popular recommendations systems.**

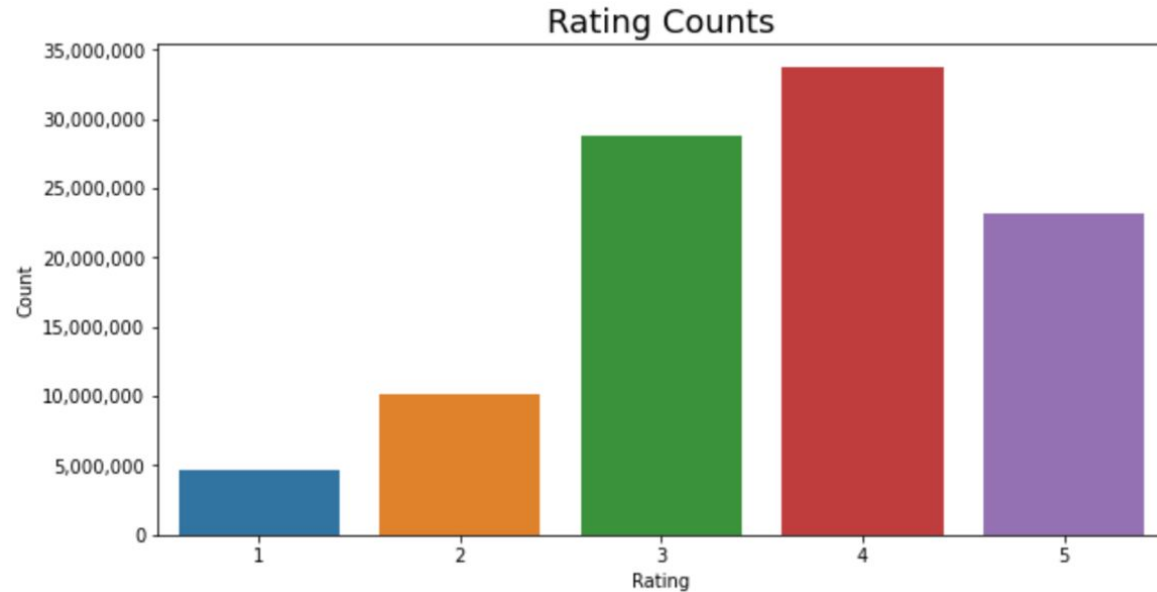
1. Content-based Recommendation Systems
2. Collaborative-based Recommendations Systems
  - a. User-User Filtering
  - b. Item-Item Filtering

# Comparison of User-User and Item-Item Filtering

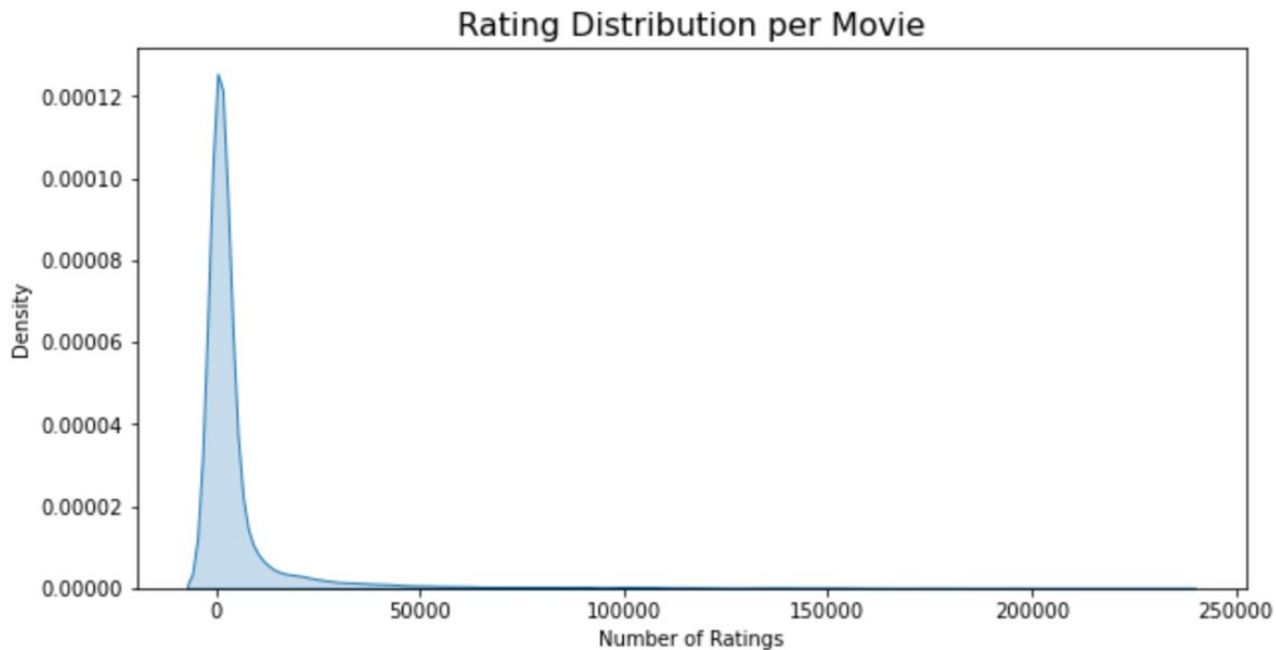


# Exploratory Data Analysis

**Most of the people rate 3 or 4.**

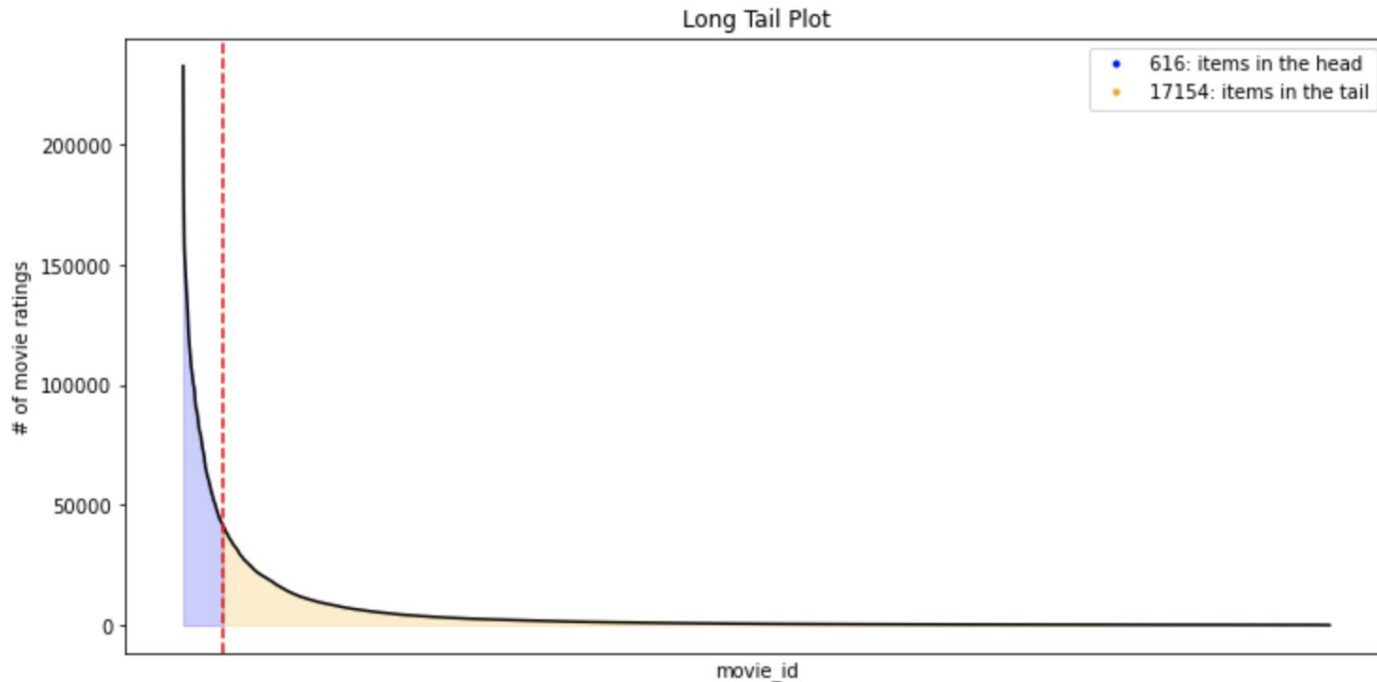


**Most of the movies don't get ratings or have less number of ratings.**

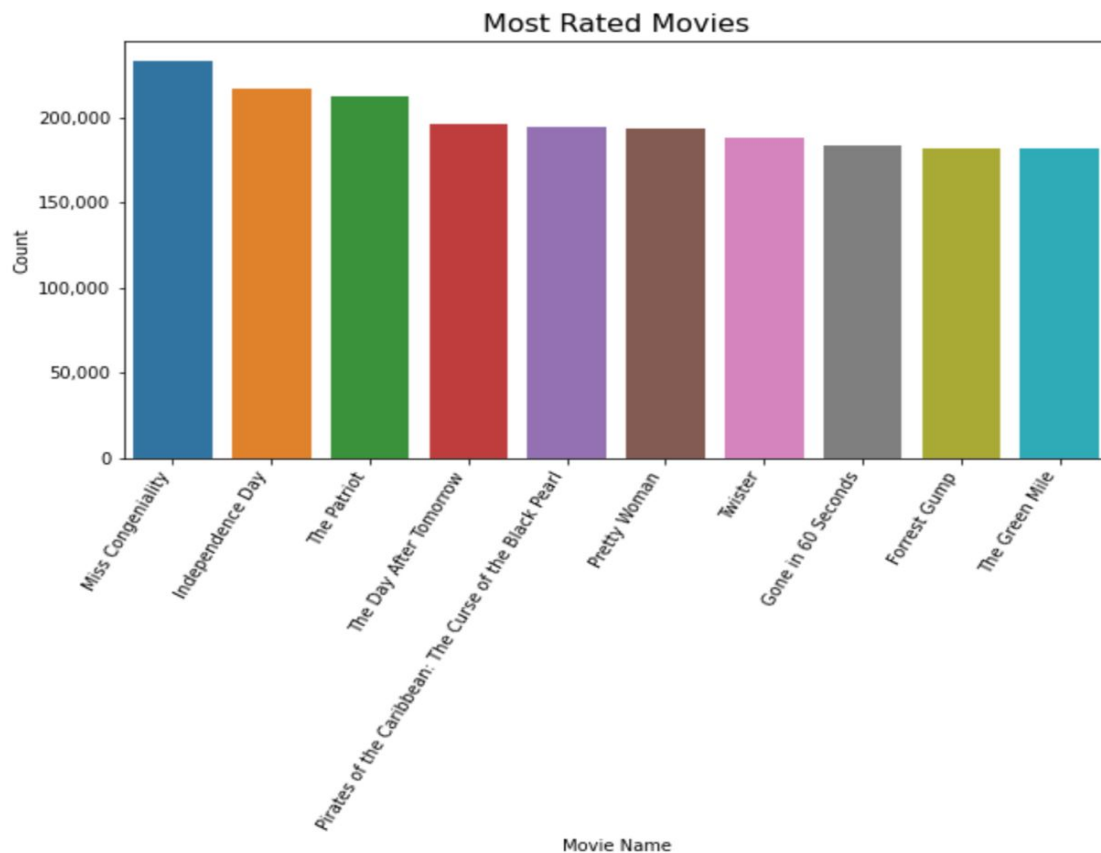




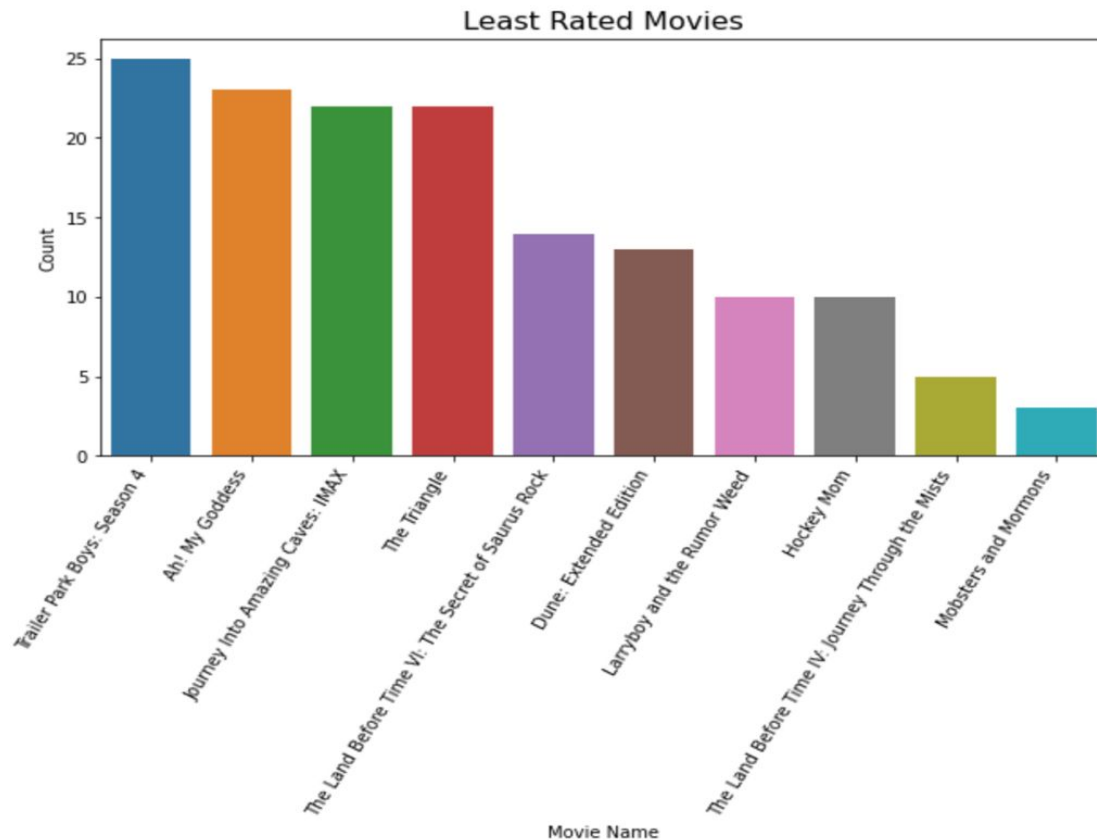
**4% of total movies received half of all ratings.**



# Top 10 most common rated movies.

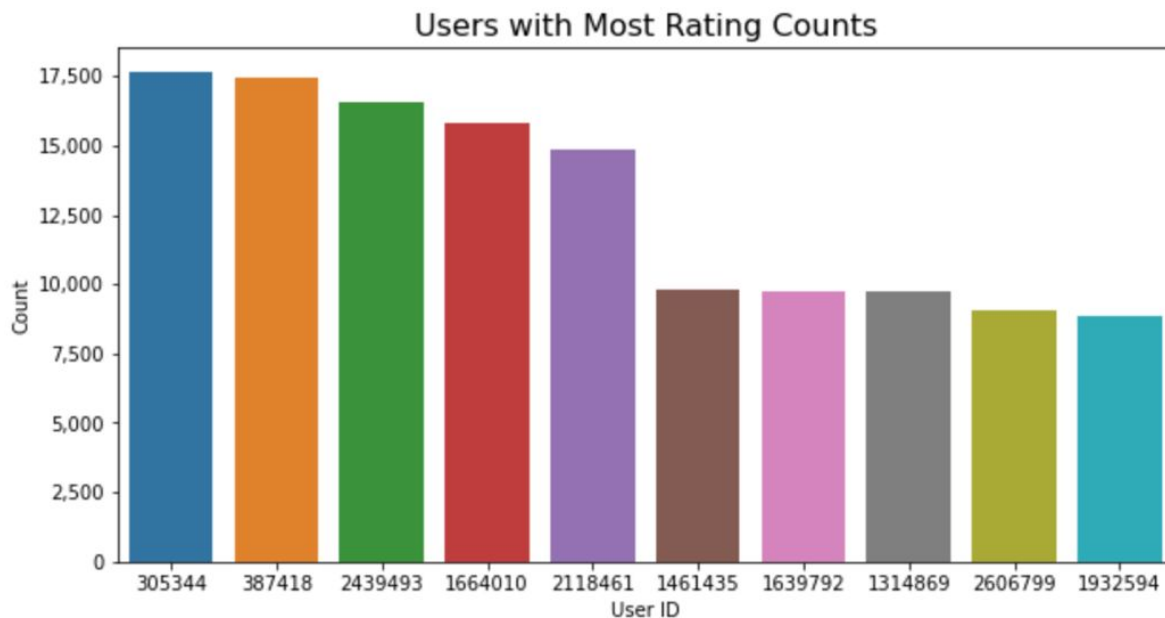


## Top 10 least common rated movies.

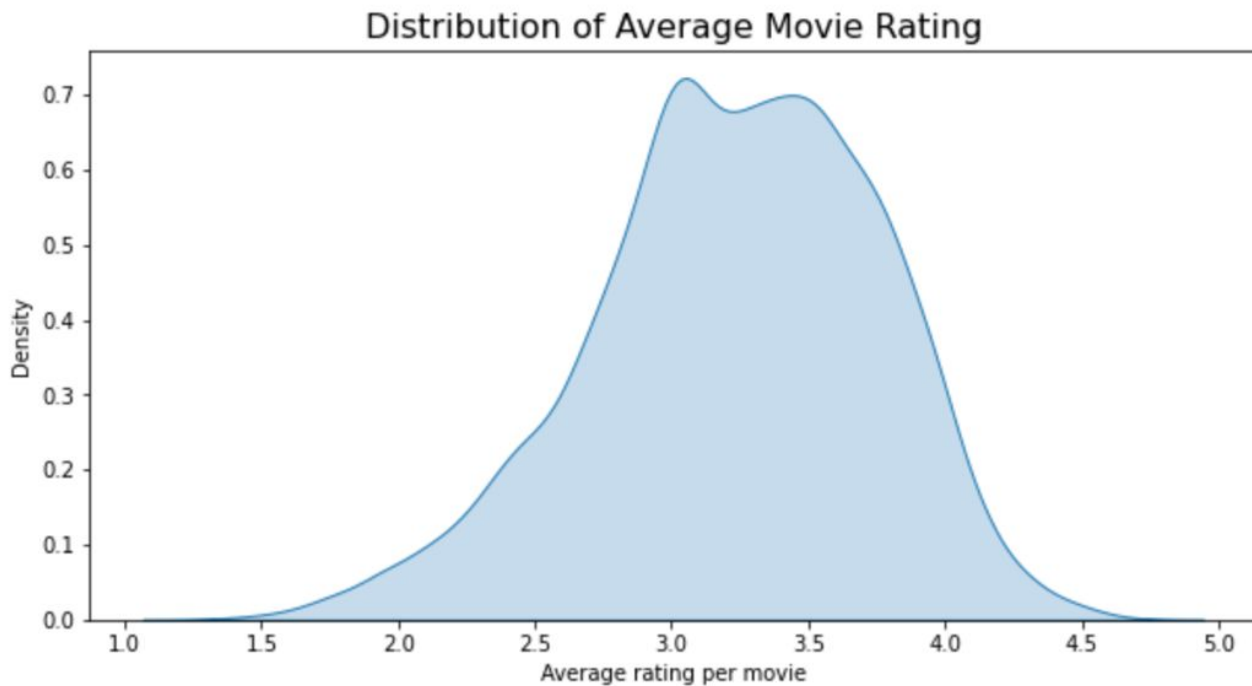




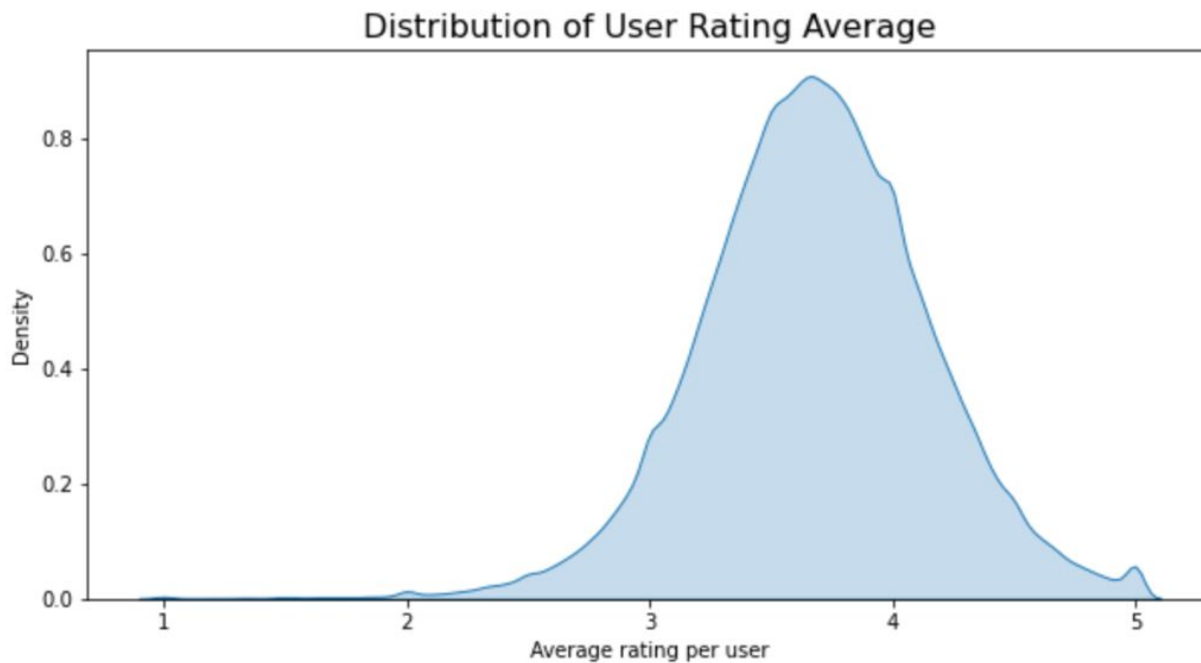
**The user with highest number of ratings: 17,500 ratings.**



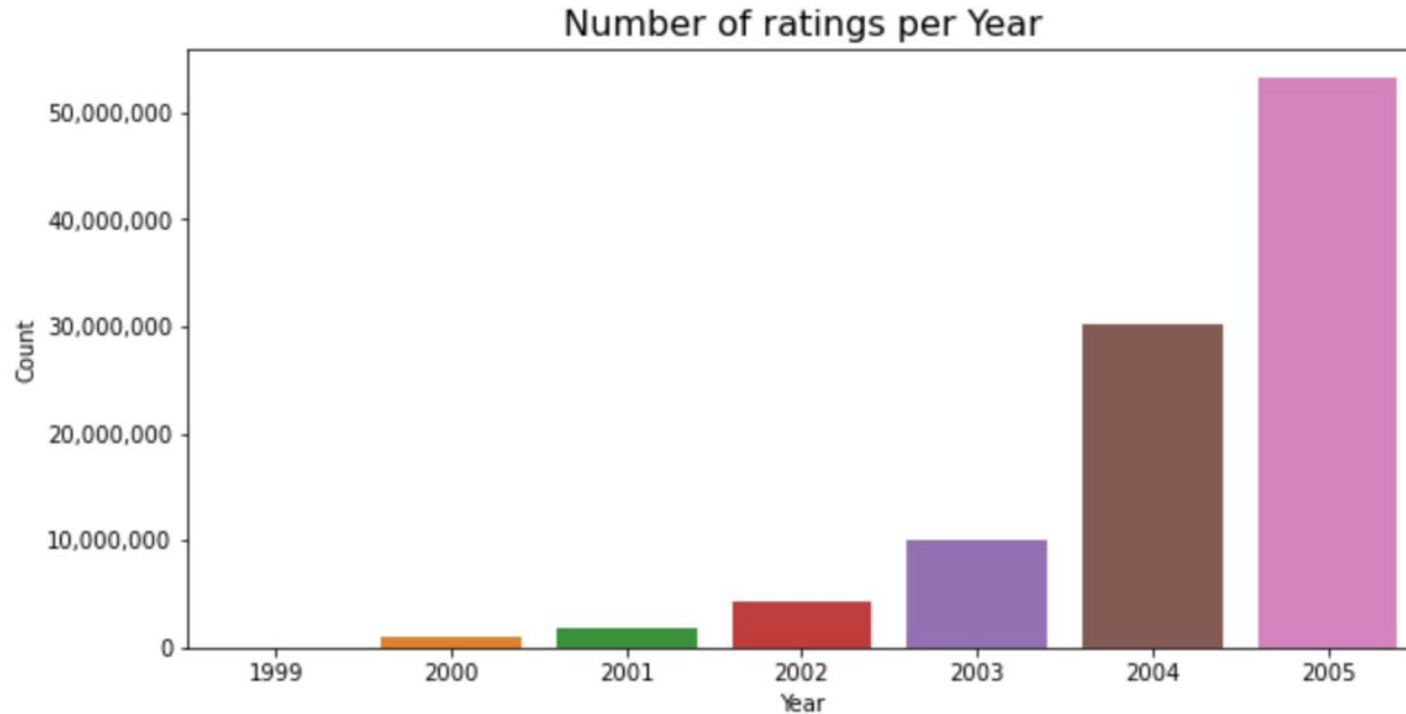
**Most of the movies are rated 3.2 on average.**



**Most of the users rated 3.7 on average.**

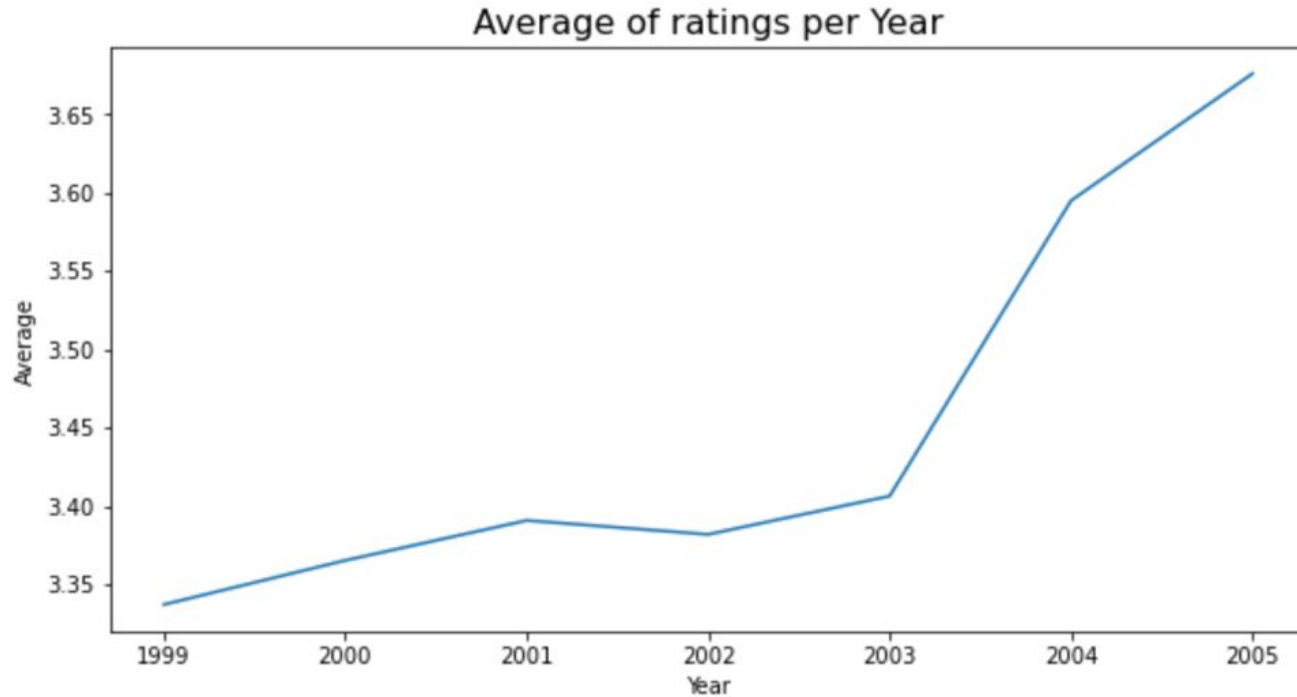


**The number of ratings has increased exponentially over the years.**

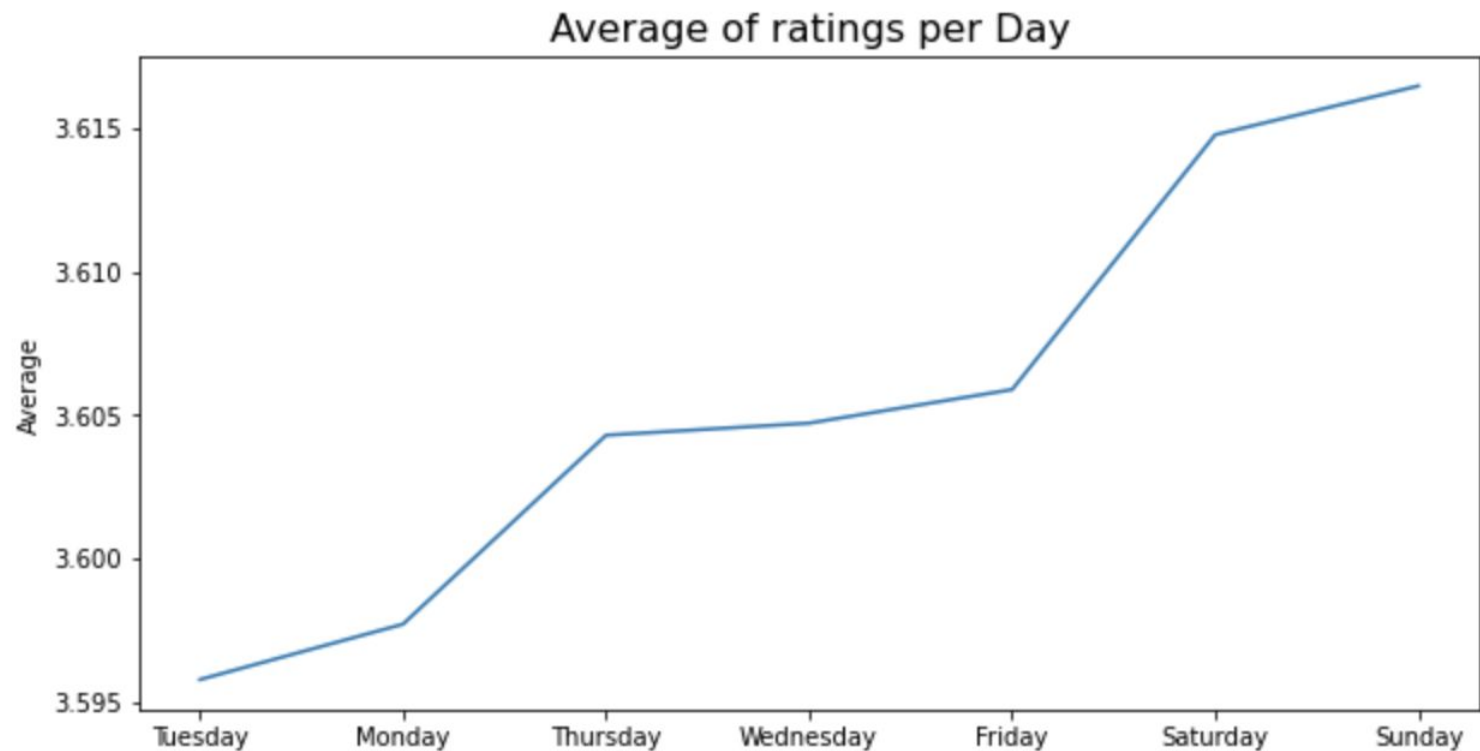




**Average rating per year also slightly increasing.**



# People are more generous on Sundays!



# Modeling

# Models for Recommendation System...

1. Models from Surprise Package
  - a. Basic Models
  - b. K-NN Models
  - c. Matrix Factorization-based Models
2. Linear Regression

**RMSE is the model evaluation metric.**

	test_rmse
<b>LinearRegression</b>	0.8913
<b>SVDpp</b>	0.9576
<b>BaselineOnly</b>	0.9677
<b>SVD</b>	0.9732
<b>KNNBaseline</b>	0.9782
<b>KNNBasic</b>	1.0245
<b>SlopeOne</b>	1.0333
<b>KNNWithZScore</b>	1.0500
<b>KNNWithMeans</b>	1.0530
<b>NormalPredictor</b>	1.4560

**Lowest RMSE**

Root Mean Squared Error Formula

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

# Movie Recommendation Web Application.

Rate Your Favorite Movies  
Then Click on the Recommend Button

Lord of the Rings: The Fellowship of the Ring  
Release year 2001

★ ★ ★ ★ ★ [Clear All](#)

Double Jeopardy  
Release year 1999

★ ★ ★ ★ ★ [Clear All](#)

Indiana Jones and the Last Crusade  
Release year 1989

★ ★ ★ ★ ★ [Clear All](#)

A Cry in the Dark  
Release year 1988

★ ★ ★ ★ ★ [Clear All](#)

The Waterdance  
Release year 1992

★ ★ ★ ★ ★ [Clear All](#)

Elizabeth R  
Release year 1971

★ ★ ★ ★ ★ [Clear All](#)

Crosby Stills & Nash: The Acoustic Concert  
Release year 2004

★ ★ ★ ★ ★ [Clear All](#)

Vlad  
Release year 2003

★ ★ ★ ★ ★ [Clear All](#)

West Point: The First 200 Years  
Release year 2002

★ ★ ★ ★ ★ [Clear All](#)

Personalized Movie  
Recommendation List



## Your Recommended Movies

Trailer Park Boys: Season 3  
Release year 2003  
Predicted Rating 4.72

Little Murders  
Release year 1971  
Predicted Rating 4.69

Screaming Dead  
Release year 2003  
Predicted Rating 4.66

The Blair Thumb  
Release year 2001  
Predicted Rating 4.66

Gantz  
Release year 2004  
Predicted Rating 4.66

Midsomer Murders: Painted in Blood  
Release year 2003  
Predicted Rating 4.66

Dream Theater: Images and Words: Live in Tokyo / 5 Years in a LIVEtime  
Release year 2004  
Predicted Rating 4.66

Harvie Krumpet  
Release year 2003  
Predicted Rating 4.66

The Most Terrible Time in My Life  
Release year 2000  
Predicted Rating 4.64

Hamlet  
Release year 2000  
Predicted Rating 4.64

Rated movies that are  
previously watched.

# Conclusion

The movie recommender system is useful to any business that makes money via recommendation. This includes Netflix, DisneyPlus, Hulu, etc. These companies could be potential clients of this recommendation engine. The system will give a good recommendation since the model gives minimal error compared to other methods. Giving good recommendations directly entails one or many of the following:

1. Customers buy a particular product or service leading to increased revenue or sales.
2. Customers use the platform more frequently due to the quality and relevance of content shown to them.
3. Better User Experience. Customers spend less time searching and more time watching. The pain of discovery is eliminated.