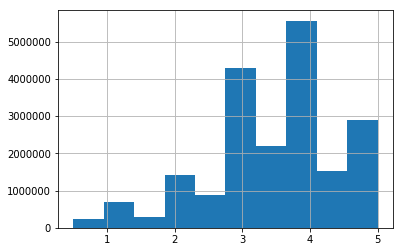
**Introduction**

Recommendation systems use ratings that users have given items to make specific recommendations. Companies nowadays are building smart and intelligent Recommendation Systems by studying the past behaviour of their users. Hence providing them recommendations and choices of their interest in terms of “Relevant Job postings”, “Movies of Interest”, “Suggested Videos”, “Facebook friends that you may know” and “People who bought this also bought this” etc. The aim of this project is to create a movie recommendation system using the MovieLens dataset that includes the previous ratings of movies by the users.

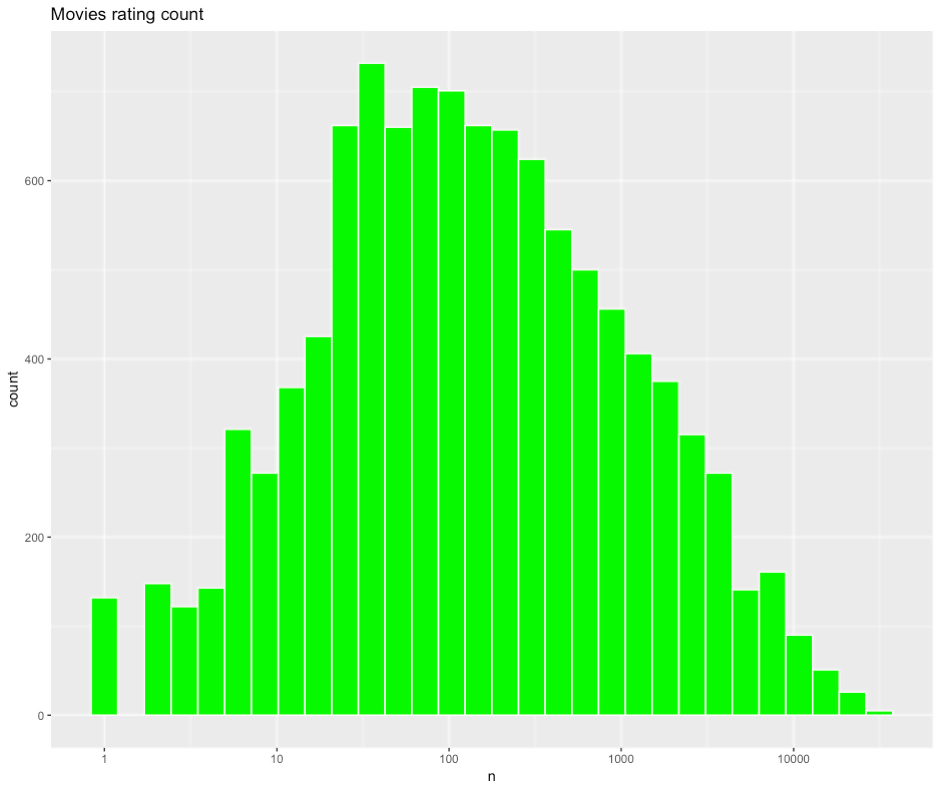
**Methods:**

The data used contains 20000263 ratings of 27278 movies. So, firstly an exploratory analysis of the data is performed to look into insights with different features affecting the rating as this will help in the modelling process. It’s done by data visualisation in the form different bar charts to gain useful insights.

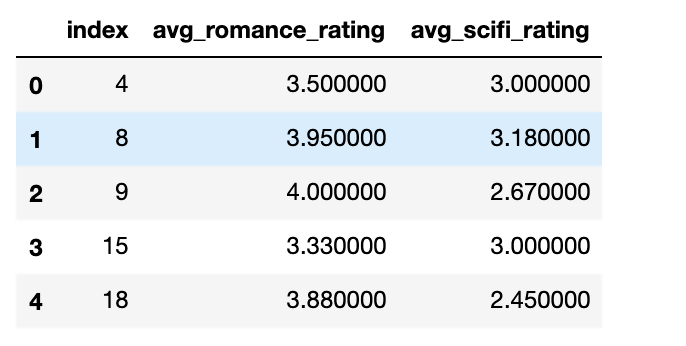


The above bar graph shows that the users mostly rate the movies between 3 and 4. But this is just the initial analysis, for making a good model different feature effects need to be studied too.

The dataset contains: 20000263 ratings of 27278 movies.



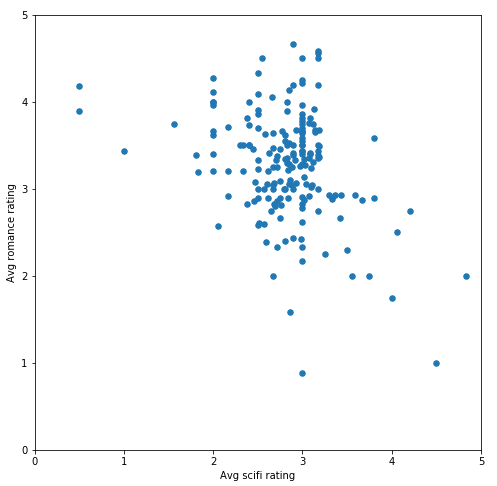
The above plot shows the movie rating count. It shows that some movies are rated very less and so are of least importance in the modelling process.

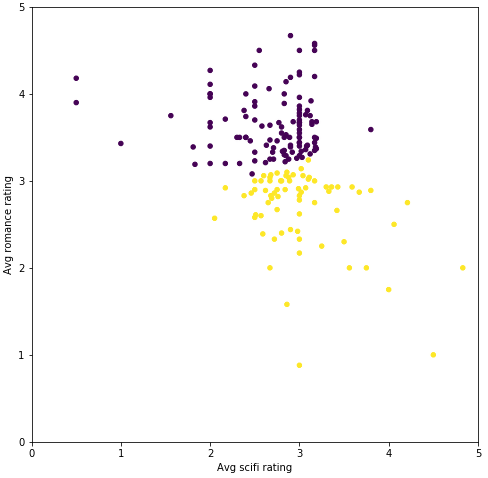


The above tables shows subset of rating for a romance and science fiction movie

Firstly, a subset of users and discovering what are their favourite genre is explored. It is done by calculating each user’s average rating for all science fiction and romance movies. Then only those results are filtered of users who only like romance or science fiction.

Now the results are visualised to look at the characteristics and can be seen the bias is clear now. Now k-means clustering can be applied for grouping.





From the above it is clear that clusters have been formed on how each person rated romance movies. People that averaged a rating on romance movies of 3 or higher will belong to one group, and people who averaged a rating of less than 3 will belong to the other.