

MovieLens pandas exercise

1. How many unique movies and users are in the dataset?
2. Who is the user who saw the highest number of movies? How many movies did he see? What is the average number of ratings (movies) per user?
3. How many movies have an average rating higher than 4.8? What do they have in common?
4. How many movies are classified as 'Comedy'?
5. How many movies are rated with each rating? Draw a barplot of the result?
6. What is the most common movie genre?
7. What is the average rating per user? Draw a histogram of this data.
8. Draw a histogram of the number of movies per user. Repeat this, but filter out users who saw more than 500 movies.
9. Part I: Which movies (names) have the highest average rating? Have you heard about them? Why?
Part II: Repeat the previous question, but filter out movies with less than 30 watchers.