

Get Data from the following link: <http://files.grouplens.org/dataset/ml-movie-lens/>

We will be using the following files for this exercise:

ratings.csv : userId, movieId, rating, timestamp

tags.csv : userId, movieId, tag, timestamp

movies.csv : movieId, title, genres

1. Read the dataset using pandas.

2. Extract the first row from tags and print its type.

3. Extract row 0, 11, 2000 from tags DataFrame.

4. Print index, columns of the DataFrame.

5. Calculate descriptive statistics for the 'ratings' column of the ratings DataFrame. Verify using describe().

6. Filter out ratings with rating > 5

7. Find how many null values, missing values are present. Deal with them. Print out how many rows have been modified.

8. Filter out movies from the movies DataFrame that are of type 'Animation'.

9. Find the average rating of movies.

10. Perform an inner join of movies and tags based on movieId.

11. Print out the 5 movies that belong to the Comedy genre and have rating greater than 4.

12. Split 'genres' into multiple columns.

13. Extract year from title e.g. (1995).

14. Select rows based on timestamps later than 2015-02-01.

15. Sort the tags DataFrame based on timestamp.