

Predict movie ratings with MovieLens dataset

Bart Boerman

20-12-2018

Overview of capstone project

EDX MovieLens dataset

- The dataset contains 9000055 records and 6 columns
- Users: 69878
- Movies: 10677
- Titles: 8316, let's investigate how this can differ from number of movies
- Genres: 707
- Ratings: 10, in steps of 0.05 from 0 to 5
- Periode ?

Descriptive statistics

Exploratory data analysis