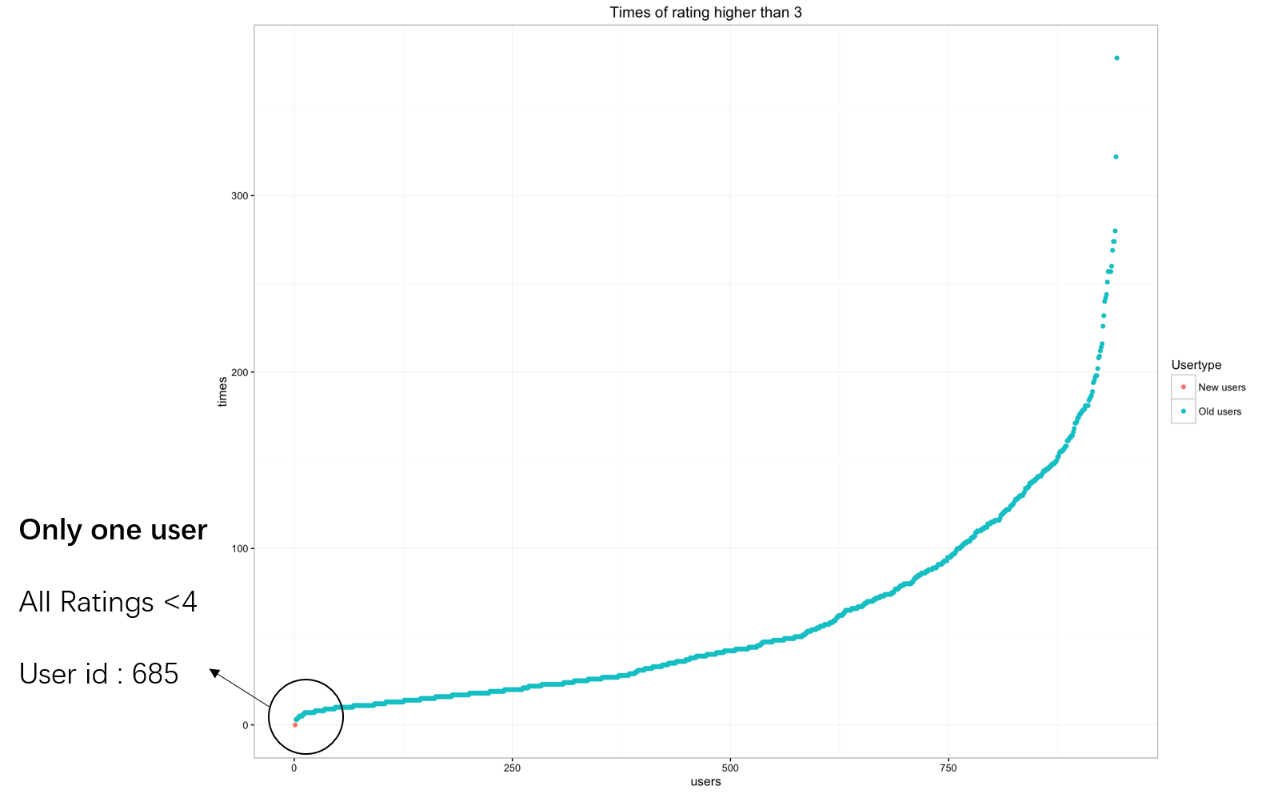
The main difference between old users and new users is that we can get the information of movie preference directly by using the ratings, which means we need to make sure that every old user provides enough information. As for the Movie Lens data, we have a filter for classification.

1. The user has rated over 10 movies and these ratings contain 4 or 5.
2. The user provides any information about age, gender and occupation.

The first condition above ensures that the user has enough rating information including preferred movies, which is needed to apply the method for old users. When (i) is not satisfied, we may consider the target user as a new user even if he has rating information (ratings are all below 4, contains no preference information). As for the new users, if (ii) is satisfied, then we’ll use new user method with personal information, if not, we’ll just recommend the most popular movies.



In the 100K data, among 943 users, only one user has no rating over 3 and we’ll treat him as “new user” even he has enough rating information. By the way, we may also user the information of user’s “dislike”, namely, ratings below 3 to avoid recommend “bad” movies. We’ll discuss that in the challenging part.