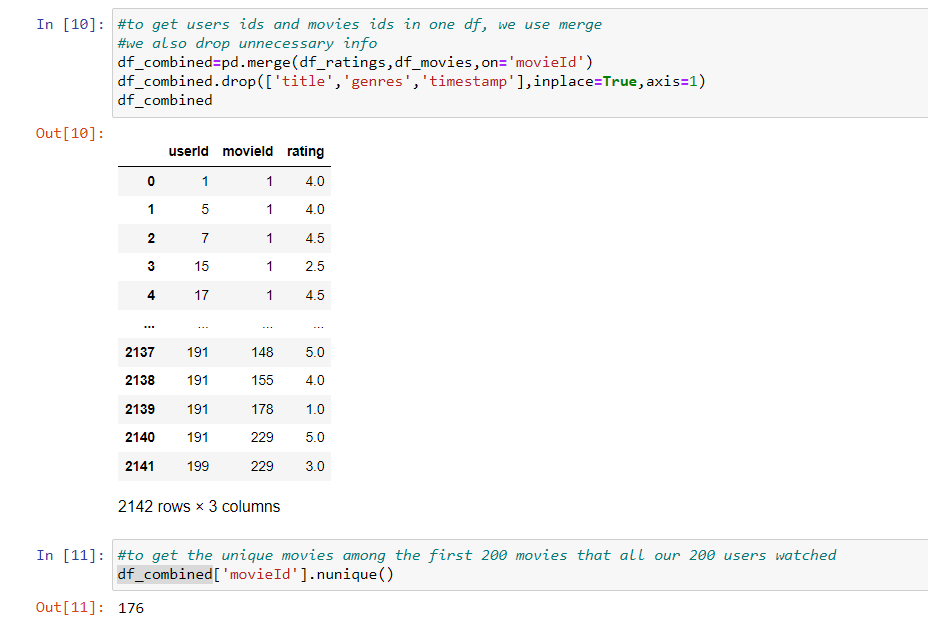
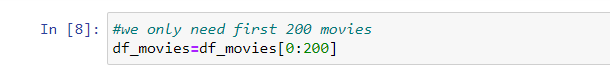
**Name: Yara Hossam El-Din**

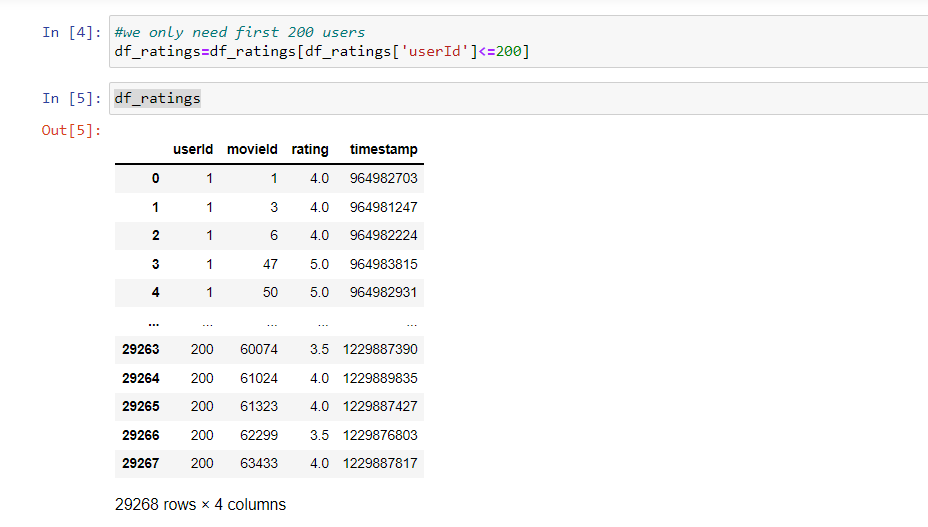
**Sec :2**

**BN: 48**

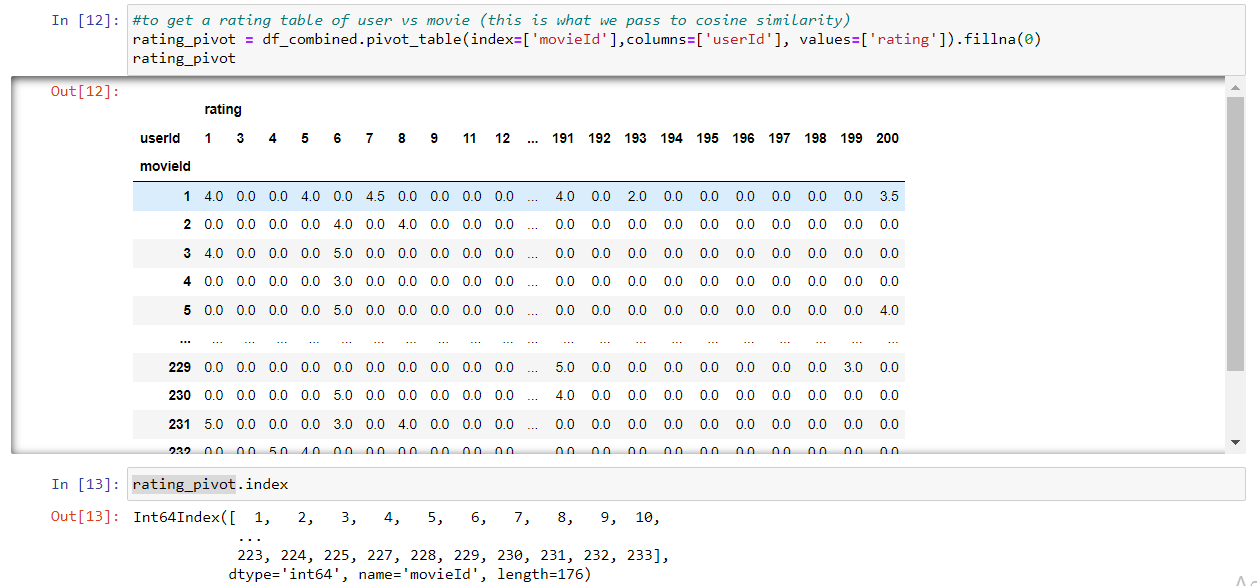
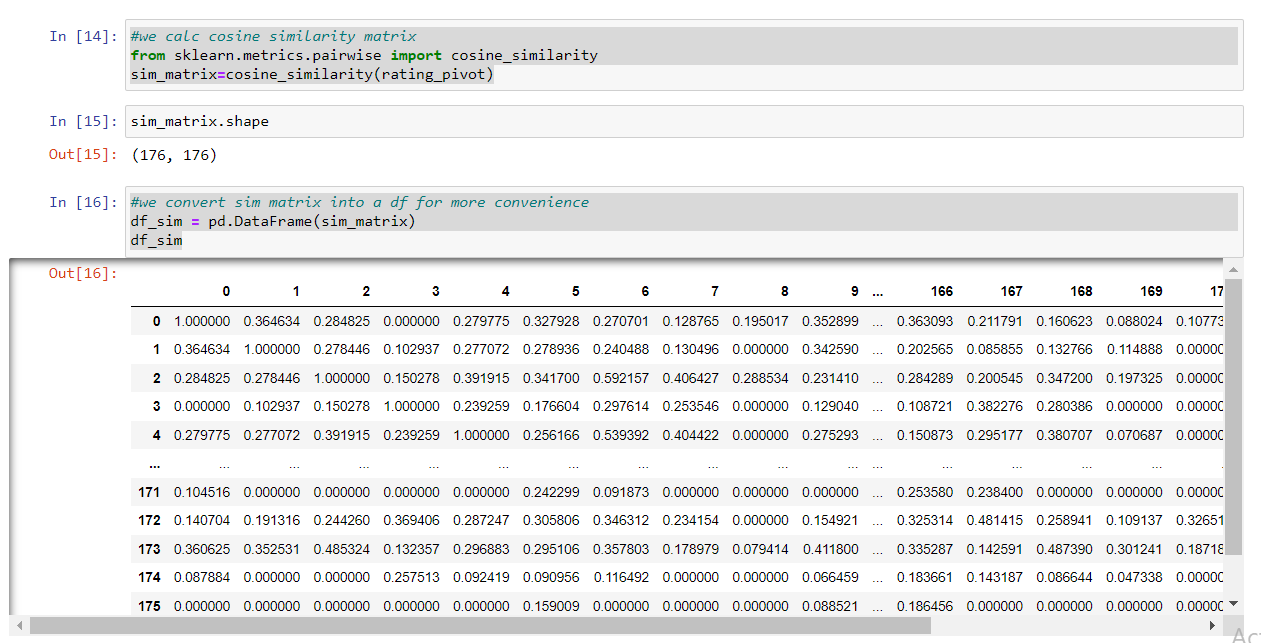
Data Preprocessing

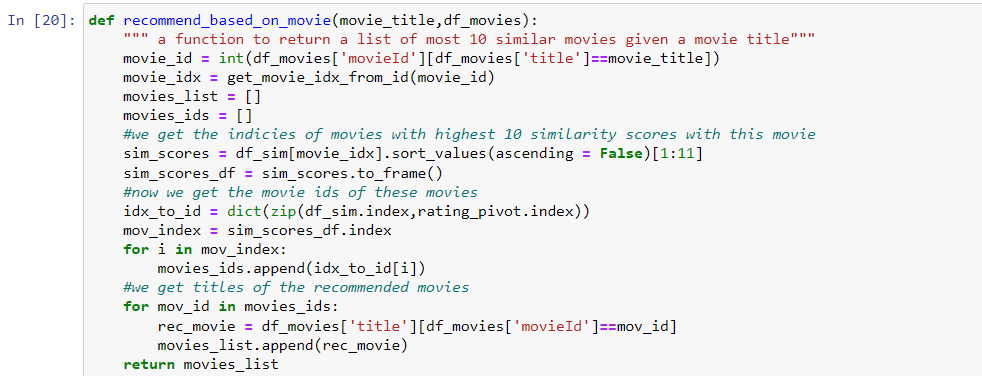
we have 2 csv files, but since we’re only interested on 200 users and 200 movies, we sample our files into smaller dataframes



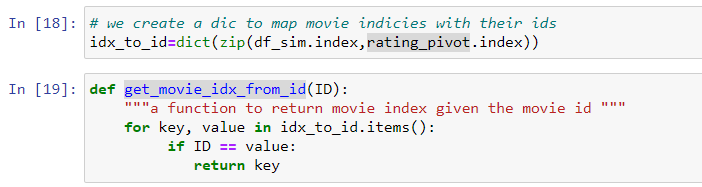
Then we merge the two dataframes 

Implementation

* we get the pivot table that contains user ids vs movie ids with each cell representing the rating
* This is what we need to pass to our cosine similarity function.
* given a movie title, we get the movie id and the movie index, then we retrieve the 10 highest similarity scores to this movie
* Now we get indices of these 10 movies, then their ids.
* finally we return titles of these 10 movies in a list



This is the function we use to map indices to ids



Testing

Testing our implementation on two movie titles, we get the following results:

