

Choose six recent popular movies. Ask at least five people that you know (friends, family, classmates, imaginary friends) to rate each of these movies that they have seen on a scale of 1 to 5. There should be at least one movie that not everyone has seen!

Take the results (observations) and store them somewhere (like a SQL database, or a .CSV file). Load the information into a pandas dataframe. Your solution should include Python and pandas code that accomplishes the following:

1. Load the ratings by user information that you collected into a pandas dataframe.
2. Show the average ratings for each user and each movie.
3. Create a new pandas dataframe, with *normalized* ratings for each user. Again, show the average ratings for each user and each movie.
4. Provide a text-based conclusion: explain what might be advantages and disadvantages of using normalized ratings instead of the actual ratings.
5. [Extra credit] Create another new pandas dataframe, with *standardized* ratings for each user. Once again, show the average ratings for each user and each movie.

You may find this short article on normalization and standardization to be useful:

<http://bi-analytics.org/topic/9-standardization-vs-normalization/>

Your deliverables should include your source data and a Jupyter Notebook, posted to GitHub.

This is by design a very open ended assignment. A variety of reasonable approaches are acceptable.

You may work in a small group on this assignment. If you work in a group, each group member should indicate who they worked with, and all group members should individually submit their assignment.

Please start early, and do work that you would want to include in a “presentations portfolio” that you might share in a job interview with a potential employer! You are encouraged to share thoughts, ask, and answer clarifying questions in the “Week 7: Data Aggregation” forum.

Here are some examples of recommender systems for movies:

<http://www.cnet.com/news/top-10-movie-recommendation-engines/>

	American Sniper	Edge of Tomorrow	Groundhog Day	Jurassic World	Lost in Translation	Lucy
John	5	4	3			4
Logan	4		3	3		
Modesto			4		4	4
Malcolm			2		4	
Maurice	5	4	4	2	3	3
	4.7	4.0	3.2	2.5	3.7	3.7