

# Datamining & Statistics

## Assignment 4: Recommendation

Use the **DMS – S7 - Assignment Template** notebook that is available on Moodle to complete this assignment.

### To get a 6: complete chapter 1 in the template

Create a basic item-based recommender, similar as used during the lecture, but with a different dataset. You could, for example, look into the Book dataset (<http://www2.informatik.uni-freiburg.de/~ciegler/BX/>). Note that this Book dataset might be too large to create a pivot table from, so you might want to use a subset from it.

As usual assume a student next year might get your notebook and will have to understand the steps you take, so add explanations along the way. This also goes for any of the additional steps described below.

### + 1 Negative correlations: complete chapter 2 in the template

The recommenders so far used, are looking at what if I give a high rating for this, what do you then recommend me. But how about if I give a low rating for something, what would you then recommend me?

### + 1 Content based: complete chapter 3 in the template

Add a chapter showing how a basic content based recommender might work. A starting point for this might be <https://www.datacamp.com/community/tutorials/recommender-systems-python> (at about 1/3 of the page it starts explaining this in the section called 'Content based').

### + 1 User based: complete chapter 4 in the template

Add a chapter showing how a basic user based recommender might work. A starting point for this might be <https://realpython.com/build-recommendation-engine-collaborative-filtering/>

### + 1 Combination: complete chapter 5 in the template

Combine two or more recommendation techniques to get an even better recommendation.

## Deliverables

*To submit:*

- Jupyter notebook file
- Dataset file(s) (or link to this in case it's too large)

*Method to submit:*

- One zip or rar file containing the files. Include your student number in the name of the zip/rar file.
- Upload on Moodle.

## FAQ's

*Q1: Can we work together?*

A1: No and yes. No, it is an individual assignment, yes you can of course always help each other. But no direct copy-pasting. Copy-pasting, or copy-pasting and changing just one or two words, either from fellow students, or the internet, books, etc. is considered fraud. This will be reported to the examination board.

*But...: But some lines of code will be standard, and only a few parameters or variable names will be different.....*

A-but: The copy-pasting is more on the text surrounding the code, and indeed some parts of the code will be copy-pasted. This is acceptable and probably normal in IT that you copy some code from the internet and make a few adjustments to make it work for your data.

*Q2: I'm not clear about the '+1', can you explain?*

A2: Sure. There are some criteria to get a 6. To get to a 7 add one of the three things described at a '+1', your choice which one. To get to an 8 add two, to get to a 9 add three, to get a 10 add four.