

**Microsoft:** DAT210x Programming with Python for Data Science

Help

**Outline** > Final Exam and Course Wrap-up > Final Project > Option C

## **Option C**

☐ Bookmark this page

Final Project: Option C

## **Snack Search**

One of the main uses of computers is to help us solve problems quickly and effectively. And a problem we often run into is figuring out what to eat, or what to make. This problem is solvable using data and recommendation engines.

Recommendation engines work on two levels. The first level is on the personal level. Let's say you create a dataset of foods and rank how much you enjoy or dislike them, 1-10. Given an unseen food and its set of features (such as the inclusion of ingredients, or perhaps the percentage of that meal the ingredient takes up), you should be able to train a machine learning algorithm to figure out if and how much you'd like it. The other way recommendation engines work is on the group level. A machine learning algorithm should be able to recommend new foods to you, given a set of people who share your similar food preferences.

The goal for this final project option is to build a system that allows you to identify and then recommend, recipes you're likely to enjoy.

## Hints

- Nail down your features and how you want to represent them. Will your features be the binary existence of an ingredient? How much of that ingredient is included a particular the recipe? Or something else? Will you use sparse dataframes, or is that not necessary?
- There are many freely available food websites that you can use as a basis to build your dataset. Alternatively, there are also many pre-built food datasets online, some organized by culture, by recipe name, or even by chemical properties. If you aren't able to find a dataset that represents recipes in the way you need, then you may have to do some pre-processing and wrangling to get it into your desired format.
- Share discovered resources, datasets, and approaches on the discussion forum with your fellow course-mates.
- Think about what machine learning algorithms are needed to accomplish your goal. Is it just one? Or could you leverage multiple techniques to get a more comprehensive answer?

© All Rights Reserved



© 2012-2017 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc. | 粵ICP备17044299号-2

















