

Python for Data Analysis and Scientific Computing
COMPSCI X433.3 - 005
David Lim Cheng (X017427)
davidlimcheng@gmail.com

Project Proposal

For my final project, I would like to build an application that I can demo on my personal portfolio. So, I searched for good projects for learning beginner data analytics, and found one that I think fits perfectly.

The idea is that I will build a recommendation system based on a dataset of 100,000 ratings from 943 users on 1682 different movies. Along with these reviews, there will be demographic data for the users, as well as genre information of movies.

The dataset is provided here:

<https://grouplens.org/datasets/movielens/1m/>

A tutorial for this project is provided here:

<https://www.analyticsvidhya.com/blog/2016/06/quick-guide-build-recommendation-engine-python/>

I will be using the tutorial as a guideline but will be attempting to figure out the core problems on my own. In addition, I would like to create plots of each recommendation – maybe a bar chart of recommended movies, as well as the factors that go in to each recommendation and the weight associated with them.

Please let me know if this will suffice – thanks!