Deal or No Deal: UChicago Course Edition

By Christian Bone, Khephren Chambers, Maddie Kingan, and Maddie Stamos

Concept

When it comes time to pick classes for the upcoming quarter (or for the first time), knowing the structure of your coursework and the professor's reputation are two integral components in selecting a class. Though UChicago already has the course evaluations system, it is inefficient and often time-consuming to manually sort through all evaluations to gather the information you need. Furthermore, you may know a professor's reputation, but do you also want to take a class that may have a final worth an overwhelming percentage your grade, or frequently has assignments due on an off-day or during a break? With this in mind, we built a website that, given a professor and/or course by a user on the frontend, returns info on various aspects of a course (professor reputation, instructor availability, time commitment, due dates, grading structure, etc.) in order to give a UChicago student a more integrated and simplistic way of creating their schedule. Additionally, students may upload their own syllabi to the project's database, which we can scrape data from and add to the backend in order to build the database.

This project was completed for UChicago's CMSC 12200: Computer Science with Applications-II course, over a six-week timeframe.

Website Output

From UChicago course evaluations:

- sentiment analysis of student evaluations (of the specified professor, or of the specified course)
- number of hours based on average within courses
- enrollment for the course*

From course syllabi:

- total office hours*
- if an assignment is due during an off-day/break day
- grade breakdown

*if both office hour and enrollment data were available for any course specified by the user, the website would instead provide the office hour per student ratio. This is more interesting and useful data for the user.

Description of Git Repository Project Folders

There are several folders pertaining to this project in the home directory of the "dealornodeal" Git repository. The following is a general description of each folder, and some of their contents:

- **database_creation:** folder for creating the SQL database on the backend, plus a couple of CSVs containing all courses taught at UChicago
- **eval_scraping:** contains python code for scraping and cleaning UChicago course evaluations
- **syllabi_scraping:** contains python and SQL code for scraping data from UChicago course syllabi (in PDF format)
- **website:** contains python, HTML, and CSS code used to create the website for the project, titled "takethisnotthat". We used Django to build and design a development server for the project. All of our work is located in the outer "takethisnotthat" folder. A description of the sub-folders:
    - media: where uploaded syllabi are stored
    - res: contains a CSV file listing all departments at UChicago
    - searcher: the primary app within our Django project, which returns results to the user based on their input on the frontend
    - static: contains a CSS file for designing the website
    - takethisnotthat: the projects main folder with the settings, urls, and WSGI configuration for the project
    - uploader: app for uploading syllabi, which go to the "media" folder

Data Sources

UChicago Course Evaluations
UChicago Academic Calendar (2015 to 2019)
Course Syllabi (in PDF format)**

- UChicago Canvas
- physical & digital copies collected from students

**Course syllabi are extremely variable, especially in the case of grade breakdowns. We acknowledged this difficulty when working on the project, writing code to gather data from as many syllabi as possible while acknowledging time constraints or deadlines for the project that limited us from scraping every single syllabus that was collected.