Leah Krasnick

SI 206

Final Project Proposal

03/25/2018

For my final project I intend to use the website allrecipes.com (link <https://www.allrecipes.com>”) by crawling through pages and scraping data from them. Allrecipes.com is a website that has thousands of pages of recipes that can be accessed by searching or setting preferences on the home page. The challenge score for this task would be 8 points because it is “Crawling [and scraping] multiple pages in a site you haven’t used before”. My project will be interactive and users will be prompted to enter a food that they want a recipe for. After entering their desired food/dish I will search the allrecipes website for that keyword and crawl to the results page from this search. Then by scraping the data from the results page I will crawl to each of the top 5 recipes in the results for the food that the user entered. I will then scrape each page with the recipe on it and gather data from each recipe. From each recipe I will be gathering data about the amount of each ingredient needed, nutrition facts (calories, carbohydrates, sodium, fat, protein, cholesterol), the recipe title, and the amount of time it takes to make it. Once all the data is gathered, users will be prompted with how they would like to see the their results for their keyword search. They will be asked if they want to choose the recipes by health, time, number of ingredients, or all. If the user chooses that they want the recipes by health, the data will be presented in bar graphs by plotly comparing the calories, carbs, protein, sodium, etc. of each recipe. If the user chooses that they want the recipes by time, a pie chart of the time for each recipe will be presented showing the prep time and cook time out of the overall time it takes to make it. If the user chooses the recipes by number of ingredients the data will be presented in a bar graph that compares the amount of ingredients and how much of each ingredient is needed for each recipe. Lastly if the user wants to see all of these components of the recipes, they can choose “all” and the program will present all of the graphs for the recipe including showing the nutrition data comparisons, the time data comparisons, and the ingredient data comparisons. From this information presented, there will be links available to each recipe that the user can click to then direct them back to the full recipe on the allrecipes website.