1. Ramya Nagapudi: ramyasn2 (captain)

2. **Free Topic:** Vegan-Friendly Restaurants Recommender

Task Description: I will be scraping and reviewing yelp reviews to recommend a ranked list of restaurants with the best vegan food options in the desired area (that the user inputs).

Importance: As a vegan it's difficult to find restaurants on yelp that have **good** vegan options. Some issues with Yelp's vegan restaurants finder/recommender are the following:

- When you search for "vegan restaurants" Yelp tends to include "Vegetarian Friendly" restaurants in some of the top results. Because these are two different diets, even if a restaurant is vegetarian friendly it can still have 0 vegan options.
- Yelp doesn't directly indicate the quality and taste of the vegan food sold at a restaurant. To determine the taste of the vegan options I have to take the time to manually search through the various reviews in yelp with some keywords.
- When you search for "vegan restaurants" Yelp's results are ranked in order of popularity and ratings. Most of the time, the ratings of a restaurant are reflective of the food on the entire menu (which is mostly non-vegan) instead of the few vegan options.

Planned Approach:

- Task 1: Take user input, which is the user's desired location, and search for all the vegan friendly restaurants in the desired location (according to Yelp) in order to gather a list of all potential vegan friendly restaurants nearby.

Then for each restaurant in that list, I will scrape and review its Yelp reviews which discuss the vegan food sold.

- Task 2: Implement a good search query. In this case the best search query finds the most Yelp reviews that review vegan food. It should not be an over-constrained or under-constrained query. For example it could be "vegan" or "plant-based".
- Task 3: Implement a Document Selection algorithm to determine if a document is relevant or not. In this case, the collection of documents are all the yelp reviews for a restaurant and each document is a single review. Using my Document Selection system and my search query, I will retrieve a list of all the relevant documents (reviews). The relevant documents are the reviews which discuss the vegan food at the restaurant.
- Task 4: Evaluate my document selection system with precision.
- Task 5: With this narrowed down list of reviews, I want to identify if each review is positive or negative. To do this, I will perform sentiment analysis to determine if a review on the vegan food was positive/negative/neutral.
- Task 6: Aggregate the results from the sentiment analysis to compile a ranked list (the restaurants with the most positive reviews will be ranked higher on this list).
- Task 7: Compute average precision of the ranked list for evaluation.

Tools and Datasets: jupyter notebook and Yelp's datasets/API
Expected Outcome: The expected outcome is a ranked list of the restaurants in the area with the best vegan options.

Evaluation: I can compute the average precision against an average of a few vegan peers' relevance judgements to evaluate my final ranked list.

3. **Programming Language**: Python

4. **20+ hour workload**:
   - Task 1 = 2 Hours
   - Task 2 = 1 Hour
   - Task 3 = 7 Hours
   - Task 4 = 1 Hour
   - Task 5 = 9 Hours
   - Task 6 = 1 Hour
   - Task 7 = 1 Hour