SI 507 Final Project Proposal

Uniqname:yukuanz

Brief Introduction:

In this project, I will collect the restaurant data in Michigan and then make a search engine that allows users to search by distance, rating or store name.

1. Data Sources

I will collect my data from the Yelp Fusion API. I've already created my APP and gained my API key. From the API, I will be able to gain data about the name, the rating and the position of the restaurants. Then I plan to use BeautifulSoup to scrap some pages on the Yelp website to collect data on restaurant reviews.

2. Data processing

I plan to create a 2-d tree to store data on restaurant ratings and geolocation. The review data will be stored in a dictionary whose key will be the name of the restaurants. Then I'll create a barchart whose x-axis will be the restaurant's rating, from 0 to 5.0, in 0.5 steps. In addition, on this barchart I'm going to create a user interaction that when the user clicks on one of the bars of this barchart, it will reveal a list of all the restaurant names and urls corresponding to this score range. Finally, I plan to apply the Polygon library to create a map showing the location of the restaurants.

3. Data presentation and interaction

There will be an interactive interface for the user to choose how to interrogate the data. They will have a total of four choices:

1. Map
   By choosing 'map', the user will be able to see a map showing the location of the restaurants.
2. Rank by distance
   By choosing 'rank by distance', the user will input their positions. Then they will be able to see the top 20 restaurants which are closest to them.
3. Rank by ratings
   By choosing 'rank by ratings', the user will input two parameters, which are their positions and the maximum distance that they can accept. Then the users will be able to see the top 20 restaurants with the highest ratings within their acceptable distance.
4. Bar chart
   By choosing 'bar chart', the user will see a bar chart. As is introduced in the data processing part, the user will be able to see the list of restaurants within the rating range by clicking on the bar.