Project Proposal for CS171
Jia Zhang

Background and Motivation

I work in a research group where our goal is to make collective portraits of aspects of life in the city. As a group, we are producing maps designed in ways to help people understand small things about their city. A large part of this broader research is communicating by visualizing publically available data. My interests are in datasets that are rich places for data analysis but whose primary objective is to provide a service. Some examples of these are google streetview images, google business, 311 call centers, ebay, craigslist, and menupages.

Project Objectives.

For this project, I would like to build a visual navigation to menupages where the food item, not the restaurant or cuisine is used as the starting point of a user's inquiry. I will build an interactive wordtree of all the menu items in all the restaurants for a city. The user coming to this visualization will type in a starting word, such as "salad", or "chicken" and the visualization will construct for them a word tree of the items before this word(usually how it is cooked, ethnicity, descriptors), and after this word(usually what it comes with, etc). The user then click to travel along a certain path until the choices he/she has made eliminates all restaurants except for a few.

Data

I will be taking data from the menupages site, this data includes the locations of the restaurants as well as menu in html format. The project will
List of restaurants for Cambridge
http://boston.menupages.com/restaurants/cambridge/all-neighborhoods/all-cuisines/
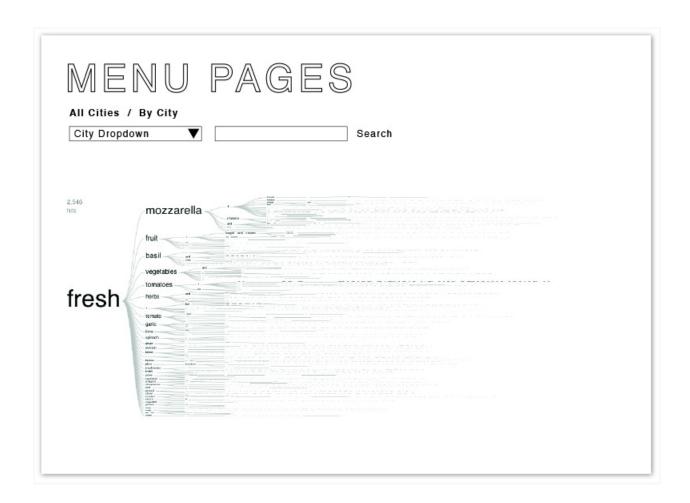Sample individual menu
http://boston.menupages.com/restaurants/1369-coffee-house-2/menu
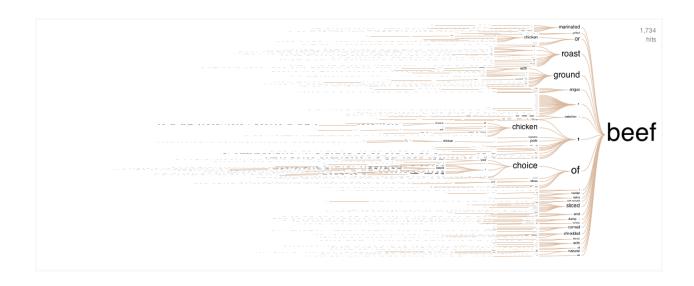
Data Processing

Although the menu data will have the same format, and will be simple to clean, the data processing portion will be the most labor intensive. I plan to construct traversable word trees from the menu items for each city. A example of a word tree implemented by Jason Davis in D3 and can be seen here:
http://www.jasondavies.com/wordtree/?source=obama.inauguration.2013.txt&prefix=We

Visualization.

I would like to display the data in 2 ways. The main interface, the wordtree will be a simple search box for the starting point which will generate a word tree that is navigatable. A second interface will be summary infographics across cities, anecdotal graphics of how restaurants prepare foods. Below are some examples of wordtrees generated from data I scraped from Menupages.

MENU PAGES

All Cities / By City

City Dropdown ▼   [   ]   Search

2,546
hits

fresh — mozzarella, fruit, basil, vegetables, tomatoes, herbs, tomato, garlic, lime, spinach, ginger, avocado, lemon

---

2,546
hits

fresh — mozzarella, fruit, basil, vegetables, tomatoes, herbs, tomato, garlic, lime, spinach, ginger, avocado, lemon

Must-Have Features and Optional Features

Must-Have: Search and display interactive tree.

Optional: A geographical map feature to this for which the user can filter using location. Voronoi maps generated by the user's selections.

Project Schedule

Week 7, 8, 9	Data downloading, cleaning, parsing - getting right http call and downloading files for all major cities from menupages site - cleaning html data - doing preliminary data analysis such as word frequency, error checking, make summary statistics for overview visualization. - building tree structure for data - learning related d3 visualization examples to prepare - make mockup of design layout, styles, etc
Week 10,11	Visualization - building search element - building tree structure - add traceback feature from hw 3 - allow people to send each other particular views.
Week 12	- building website to tie together all visuals - writing about page, additional non-visualization things - implement style and design of site
week 13	Final revisions