Leonard Ma and Rickey Dong Period 9
Reload, Man! - Ice Rod Kyng

NYC Restaurant Matcher

For our final project, we are going to create a website using a database of all registered and active restaurants and college dining halls in New York City along with their health violations if any, inspection ratings, zip codes, cuisine types, etc. This website will allow the user to filter out eateries based on their preferences in specific categories; for example, a list of pizza places in 10282 with an A inspection rating without evidence of mice can be requested and provided.

We're interested in this since health inspection data isn't easily found, especially specific violations. Especially now more than ever this website will be important because of the coronavirus pandemic. Once the pandemic blows over, and people are craving for some meals at a restaurant, anyone can use this site to find out the restaurants for them. This website would be an easy and accessible way to find convenient restaurants, serving the types of foods you're looking for while meeting your health expectations.

With text boxes and checkboxes, we allow the user to have a wide variety of inputs to cater towards their preferences. On the website, there will be three main sections: Home, Explore, and Match. The Home section will just be our documentation on our creation process of this project. The Explore section will be dynamic and load in 10 random choices of restaurants from the dataset everytime it is accessed. In the Match section, the user will use the text boxes and checkboxes to record their criteria, and when the "Find My Matches!" button is pressed, the Python program will process all of the data inputs, and it will bring up the top 10 best restaurant matches according to the given criteria. The way the top 10 ranking is determined is based on the restaurants' Score. The lower the Score number, the better the hygiene is. The Score value is more specific than the Grade value because the Grade value is just a letter between A, B, or C. The Score is based on the latest inspection from the City, so if the same restaurant has multiple inspections, only the most recent will be taken into account. The website will display all relevant data/information including inspection comments on one page for each individual restaurant's page, which you can find after going through the search/matching process. If no restaurants matched the user's criteria, then a message will say so. And if a user did not fill out a required field on the Match page, which will be denoted by an asterisk, a message will also say so, and they will have to refill the Match page.

NYC Restaurant Matcher

Home | Explore | Match

Leonard Ma + Rickey Dong
Period 9

Project Description:

Instructions:

How It Works:

Files:

Resources:

Bugs/Errors/Things To Be Aware Of:

NYC Restaurant Matcher

Home | Explore | Match

- Joyce's Tavern
  Staten Island
  3823 Richmond Avenue 10312
  718-961-4444
  Irish
- The Hunting Club
  Bronx
  2503 Williams Bridge Road 10469
  646-302-7210
  Café/Coffee/Tea

NYC Restaurant Matcher

Home | Explore | Match

Restaurant DBA: Carvel

Borough: ☑ Bronx ☑ Brooklyn ☐ Manhattan ☑ Queens ☐ Staten Island

Cuisine: ☐ Any ☐ American ☑ BBQ ☑ Chinese ☐ Korean ☐ Irish

General Location: 345 Chambers St.

Inspection Rating: ☑ A ☐ B ☐ C

Last Year Inspected by the NYC DOHMH: YYYY

Hygiene Inspection Comments: Select the ones you don't want ☐ Every ☑ mice ☑ cockroach ☐ improper food storage ☑ flies ☐ improper hand-washing

Find My Matches!

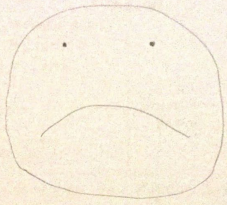
NYC RESTAURANT MATCHER

Here are the top 10 best matches given your criteria.

- Joyce's Tavern
  Staten Island
  3823 Richmond Avenue 10312
  718-961-4444
  Irish
- Niall's on 52nd
  Manhattan
  218 East 52 Street 10022
  212-888-4648
  Irish

NYC RESTAURANT MATCHER

SORRY! We could not find any matches given your criteria...



Our main source of information will be coming from the NYC Department of Health and Mental Hygiene inspection data; this can be found at https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Res ults/43nn-pn8j. This data will require us to clean a few listings for some restaurants since it includes multiple inspections from the last ~6 years, which we don't want since we'll just take the most recent data. We also have to be careful with splicing the dataset based on commas since some restaurant names such as "DUNKIN', BASKIN ROBBINS" have commas within their names. We need to be wary of restaurants that have the most recent entry marked as shut down by the DOHMH, so we need to remove those data entries. We also need to remove restaurants with empty data values, especially if the data is missing in the Score and/or Grade cells, unless there is a violation comment saying, "No violations were recorded at the time of this inspection" meaning that the restaurant was perfect, which would suit the users who want no inspection comments at all. By far the most difficult part of cleaning the data is when there are two rows of data for the same restaurant that occurred at the same date. This is possible to happen since the DOHMH issues strikes per new problem. For example, a Burger King could have a problem on the 21st of May, 2019 for plumbing issues, which would be one row in the data, but then that same exact Burger King could have another problem for cockroaches on the same exact day, but would occupy another row in the data.

We are hoping to include a map pinpointing the top suggestions our matcher finds. Since each restaurant listing in the dataset includes coordinates (in longitude, latitude), we plan to pinpoint the location of the top ten recommended restaurants in our matches using a Python module such as folium.map. The documentation source for this would be https://python-visualization.github.io/folium/.

If we have time, we hope to add in Yelp Fusion API, in order to provide things such as some reviews, the stars the business has, etc. This is free to use for up to 5,000 calls per day. This is the link to obtain an API key: https://www.yelp.com/developers/documentation/v3/authentication. If this is accomplished, we would then need to add to our Matcher, such as sorting by Yelp stars.

Development Stages/Timeline

Date	Goal:	Who
6/1	Finish creating functions to sort data in order to use	Leonard and Rickey
6/3	HTML/CSS Design for homepage, match selection page, and error page Brainstorm methods to effectively match all criteria user selects	Leonard and Rickey
6/5	Python code for displaying Explore page --- Matcher sorting: All textboxes *Collaborate on how to combine sections effectively* Matcher sorting: All checkboxes	Leonard and Rickey --- Rickey Leonard
6/8	Finish code for displaying Match HTML Begin coding for map pinpoint feature	Leonard and Rickey
6/12	Explore any relevant APIs to add, such as Yelp	Leonard and Rickey