YELP ME!

Rajas Deshpande, Omkar Buchade, Preet Dalsania

December 13, 2018

Abstract

This document contains information pertaining to the YELP ME application submitted as the final project for *CIS 600 - Principles of Social Media and Data Mining*. This application is a restaurant recommender system based on a user preferred cuisine. Data mining along with Natural Language Processing is used to build this project. The different aspects of design and development of this project are portrayed in this documentation. The content contained within this document includes the following

- Application Overview
- Runtime Instructions
- User Interface
- Technical Components
- Analysis

In addition, in the appendix at the end of the document there are the following references -

- List of Helper Functions

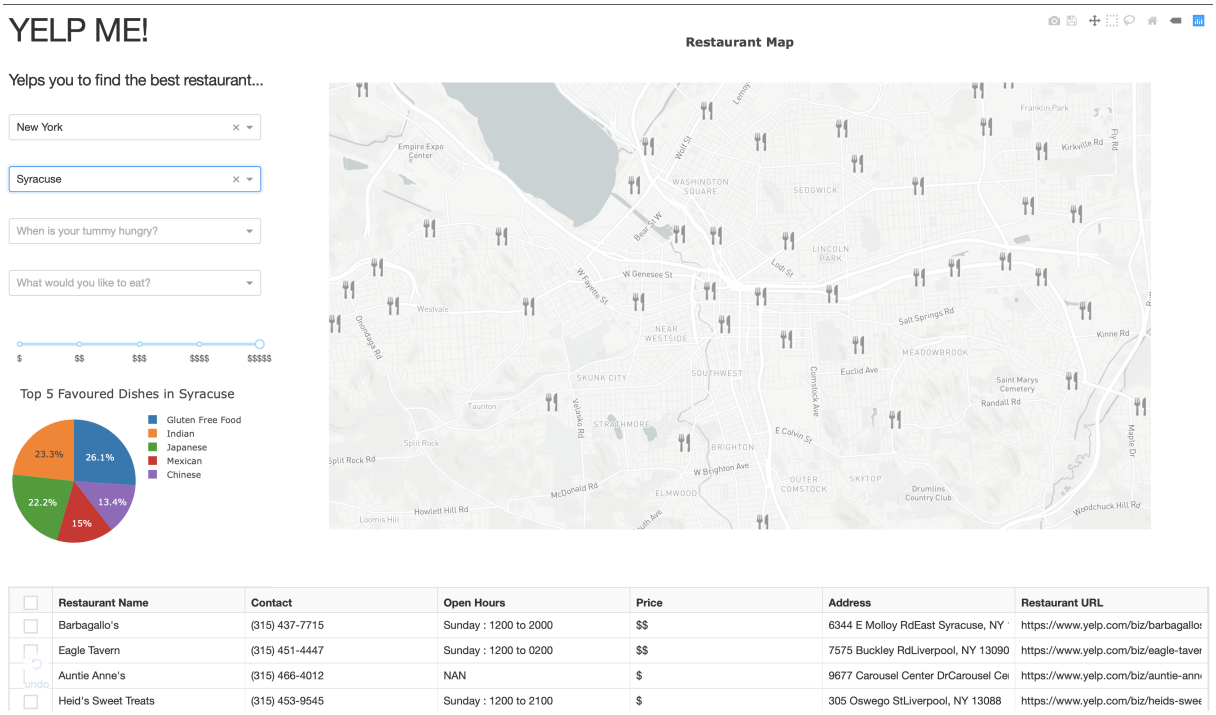
Contents

Chapter 1

Application Overview

1.1 Overview

The goal of the Yelp Me! application is to the recommend the best restaurant based on the user entered cuisine. Our proposed application will analyze user reviews of all the restaurants that correspond to that user entered cuisine and provide a sentiment score i.e. our system will recommend the best restaurant for that cuisine which the user desires to eat.



Users have the flexibility to provide a geographic location ranging from different states to cities. In addition, filter criteria can be added to refine the search. For example, users could search from the cheapest restaurants to the expensive ones.

Once this information is entered, our application finds all the restaurants with their information and displays this information on the location map and also in the form of a table. User can hover on the map to check the names of these restaurants and also select a restaurant from the table to check it on the location map. The data returned is in the form of a location map that displays all the restaurants in the requested state and city.

Further non-technical and technical details are covered in the next two chapters.

1.2 Limitations

1.2.1 MapBox API

The location map provided by the MapBox API does not take the correct latitude and longitude values when the map is zoomed out to the maximum level during initialization.

1.2.2 Yelp API

The Yelp Fusion API has a daily limit of 5000 requests. Additionally, Yelp Fusion API limits the number of reviews to only 3 reviews per restaurant. Also, additional details of the restaurants such as the ambience, amenities, happy hours and other related aspects of the restaurant are only available via Yelp Fusion VIP API and not the regular Fusion API.

Chapter 2

Runtime Instructions

2.1 Configuration

2.1.1 API Accounts

1. **Mapbox API:**
   To use the mapbox API we first need to create a mapbox account and then use the access token. The steps to create a mapbox account and generate the access tokens is given below-

   Click here to Get Started with the MapBox API.

2. **Yelp API:**
   To use the Yelp fusion API, user must have a Yelp Fusion API key to make requests to the API. The steps to create a Yelp developer account and generate the API key is given below-

   Click here to Get Started with the Yelp API.

   Now copy these access token to the *OAuth_Keys.json* file.

2.1.2 Authentication File

Once accounts are setup with the Yelp API and MapBox API services, the access tokens will need to be stored in JSON format in a file called *OAuth_Keys.json*. This file should be placed in the same directory as the YelpMe.py. A sample file containing dummy keys currently exists in the YelpMe zip file.

The file must contain the following key value pairs all at the initial level in the JSON file.

Key	**Value**
Token	Mapbox API access token
Key	Yelp Fusion API key

2.1.3 Supporting Data

The data sub-directory must be located in the directory where the python code resides. This folder, contains files about businesses (restaurants) and business review (restaurants reviews) for different businesses across the United States in JSON format.

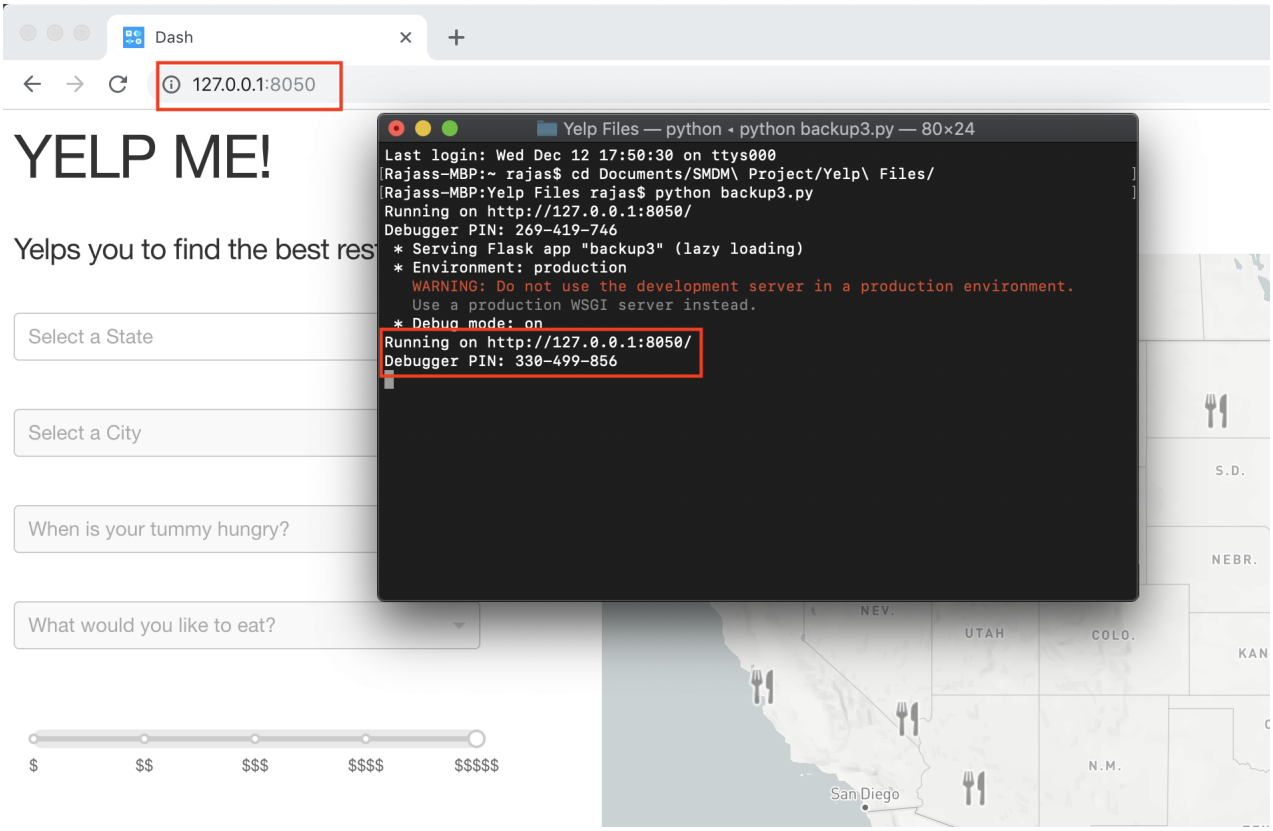
2.2 Running the Code

After all of the configuration steps have been completed, the Yelp Me application is ready to be run. In order to start the application, open your terminal and run the following command -
Command: python YelpMe.py

The output of this command will be as follows -

Running on http://127.0.0.1:8050/
Debugger PIN: 269-419-746
* Serving Flask app "backup3" (lazy loading)
* Environment: production
WARNING: Do not use the development server in a production environment.
Use a production WSGI server instead.
* Debug mode: on
Running on http://127.0.0.1:8050/
Debugger PIN: 330-499-856

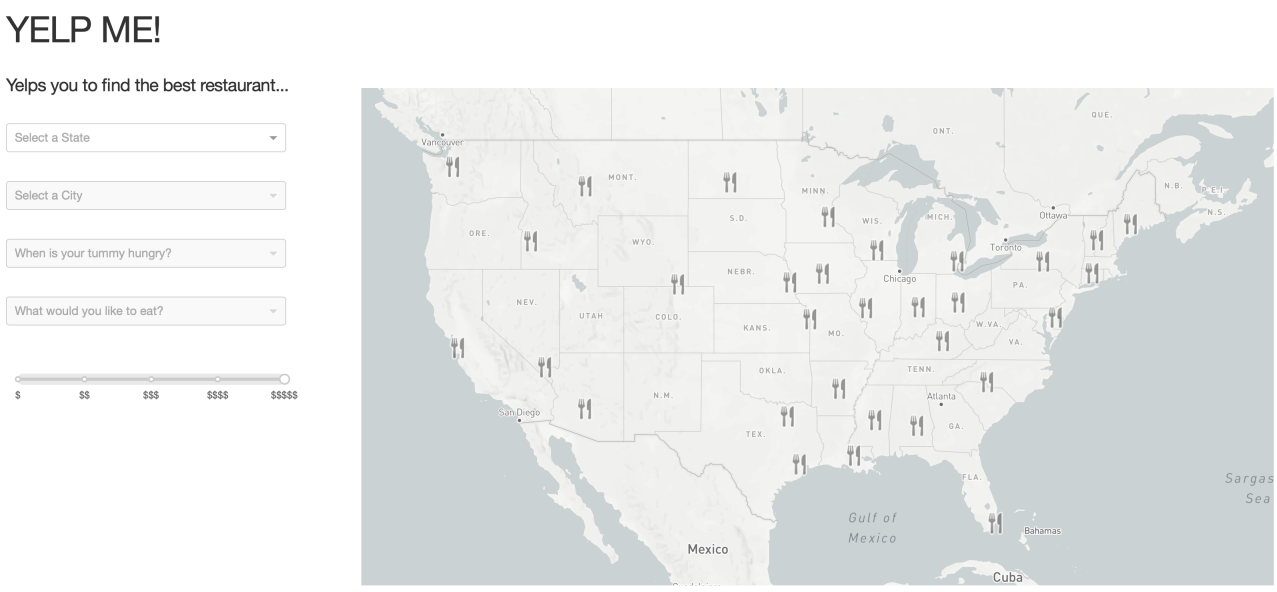
Enter the url mentioned after "Running On" on your browser and the application is started. The image below highlights the two locations where the port number must match.



Chapter 3

User Interface

The user interface is comprised of two sections, the user input menu on the left and the results that the application displays visually on the right. Initially, since their are no results to display, the right side of the application will just display the location map of all the restaurants across the United States. The initial display is shown below.

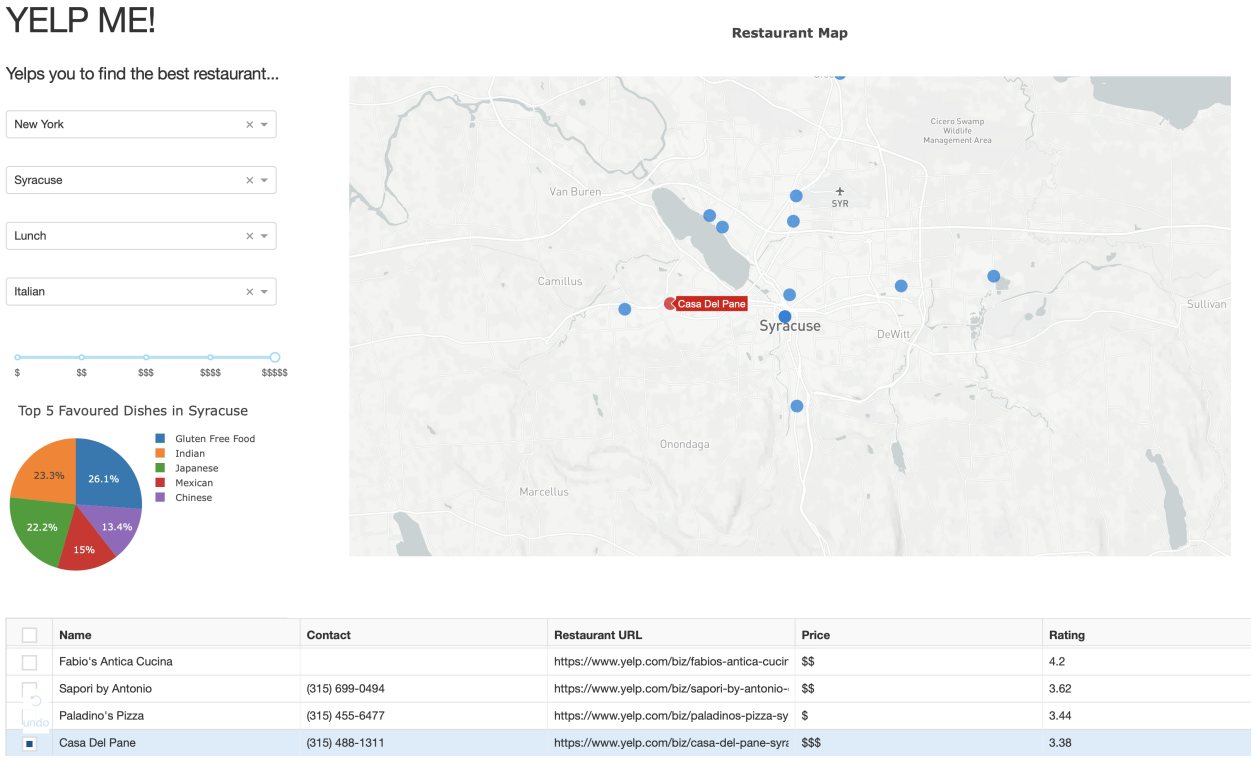


When a user enters a state and a city, the location will display all the restaurants in that city along with a table providing the information of all restaurants displayed on the location Map. Additionally, a pie chart is provided which suggests the best cuisine which the people of that city prefer. An example of the final display is shown below.

Restaurant Name	Contact	Open Hours	Price	Address	Restaurant URL
Barbagallo's	(315) 437-7715	Wednesday : 1100 to 2300	$$	6344 E Molloy RdEast Syracuse, NY	https://www.yelp.com/biz/barbagallo
Eagle Tavern	(315) 451-4447	Wednesday : 1130 to 0200	$$	7575 Buckley RdLiverpool, NY 13090	https://www.yelp.com/biz/eagle-taver
Auntie Anne's	(315) 466-4012	NAN	$	9677 Carousel Center DrCarousel Cer	https://www.yelp.com/biz/auntie-ann
Heid's Sweet Treats	(315) 453-9545	Wednesday : 1200 to 2100	$	305 Oswego StLiverpool, NY 13088	https://www.yelp.com/biz/heids-swee

3.1 User Input

The left hand menu allows users to enter the input to the Yelp Me application. This includes the geographic location around which the restaurants will be gathered. Additionally, a dropdown filter is provided in which the user chooses the preferred cuisine which he/she desires to eat. Also, a price filter is provided that allows the user to choose between the most expensive to least expensive restaurants. The output is shown below.



Name	Contact	Restaurant URL	Price	Rating
Fabio's Antica Cucina		https://www.yelp.com/biz/fabios-antica-cucir	$$	4.2
Sapori by Antonio	(315) 699-0494	https://www.yelp.com/biz/sapori-by-antonio-	$$	3.62
Paladino's Pizza	(315) 455-6477	https://www.yelp.com/biz/paladinos-pizza-sy	$	3.44
Casa Del Pane	(315) 488-1311	https://www.yelp.com/biz/casa-del-pane-syr	$$$	3.38

3.1.1 Search Filters

The user can also apply search criteria to filter the restaurants at the given location. The different filters and their usage are as follows.

Filter	Example	Description
States	New York	Drop-Down menu of states
City	Syracuse	Drop-Down menu of cities
Meal	Lunch	Drop-Down to select the meal
Cuisine	American	Drop-Down of different cuisines
Price	$$	Slider bar to select budget

3.2 Visualizations

The results of a user request are displayed graphically in the form of a geographical map. The output is divided into three sections - the location map, the tabular information and the pie chart.

3.2.1 Location Map

The first section plots all the restaurants of the city that the user has provided as input. Each data point (restaurant) is displayed using a 'Fork and Knife' marker.

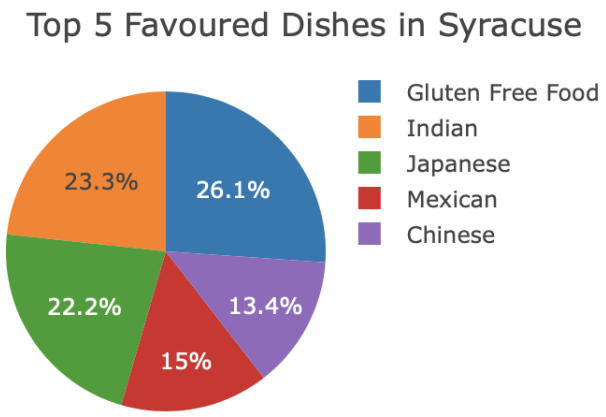
3.2.2 Tabular Information

The second section displays a table of restaurants that are displayed on the location map. The table displays different information about the restaurants such as location, contact number, restaurant URL, price, rating, open hours.

Restaurant Name	Contact	Open Hours	Price	Address	Restaurant URL
Barbagallo's	(315) 437-7715	Wednesday : 1100 to 2300	$$	6344 E Molloy RdEast Syracuse, NY	https://www.yelp.com/biz/barbagallo
Eagle Tavern	(315) 451-4447	Wednesday : 1130 to 0200	$$	7575 Buckley RdLiverpool, NY 13090	https://www.yelp.com/biz/eagle-taver
Auntie Anne's	(315) 466-4012	NAN	$	9677 Carousel Center DrCarousel Ce	https://www.yelp.com/biz/auntie-ann
Heid's Sweet Treats	(315) 453-9545	Wednesday : 1200 to 2100	$	305 Oswego StLiverpool, NY 13088	https://www.yelp.com/biz/heids-swee
The Chickadee Human Eatery	(315) 699-7044	Wednesday : 0800 to 1800	$$	8140 Brewerton RdCicero, NY 13039	https://www.yelp.com/biz/the-chicka
Bruegger's Bagels	(315) 449-2528	Wednesday : 0600 to 1800	$	3065 Erie Blvd EDeWitt, NY 13224	https://www.yelp.com/biz/brueggers-
Saltine Warrior Sports Pub	(315) 314-7740	Wednesday : 1100 to 0000	$$	214 W Water StSyracuse, NY 13202	https://www.yelp.com/biz/saltine-war
OYE at Cowboys Saloon	(315) 472-0102	NAN		306 Hiawatha Blvd WSyracuse, NY 1	https://www.yelp.com/biz/oye-at-cow
The Beginning II	(315) 463-5080	Wednesday : 1130 to 0200	$	6897 Manlius Center RdEast Syracus	https://www.yelp.com/biz/the-beginn
Doc's Pizzaria & Fishfry	(315) 487-2698	NAN	$$	3797 Milton AveCamillus, NY 13031	https://www.yelp.com/biz/docs-pizza

3.2.3 Pie Chart

The third section shows the overall sentiment score of a cuisine based on the user reviews collected using the Yelp Fusion API.



Chapter 4

Technical Components

On the back end, the Yelp Me Application gathers restaurant data, performs text analysis of user reviews and plots the restaurant data on the fly when a user request is received. The following sections describe each of the major components of application.

4.1 Data Acquisition

(For development purpose: This is not necessary to execute code and is just required to mine data using the Yelp Fusion API.)

Gathering restaurant data is done through the Yelp Fusion API. The results from the API are in JSON format. These results are further cleaned, processed and stored in pandas dataframe. The columns contained in the dataframe are listed below.

Data Field	Description
id	Unique Yelp ID of the restaurant.
name	Name of the Restaurant.
coordinates	Coordinates of the restaurant.
display_phone	Contact of the restaurant in the standard phone number format.
hours	Opening hours of the restaurant.
display_address	Provides the address.
rating	Rating of the restaurant ranging from 1 to 5.
Price	Price level of the restaurant ranging from $ to $$$$$.
url	URL for restaurant page on YELP.

To use the yelp API in python we can simply instantiate yelp.client.Client with our API key, and we can start making requests. Following code can be used to create the client object to make requests.

```
from yelp.client import Client
API_KEY = "abcefghijklmnowxy"  #Replace this with your yelp fusion API key
client = Client(API_KEY)
```

Name	Path	Request	Description
Business Search	/businesses/search	GET https://api.yelp.com/v3/businesses/search	Search for businesses by keyword, category, location, price level, etc.
Business Details	/businesses/{id}	GET https://api.yelp.com/v3/businesses/{id}	Get rich business data, such as name, address, phone number, photos, Yelp rating, price levels and hours of operation.
Reviews	/businesses/{id}/reviews	GET https://api.yelp.com/v3/businesses/{id}/reviews	Get up to three review excerpts for a business.

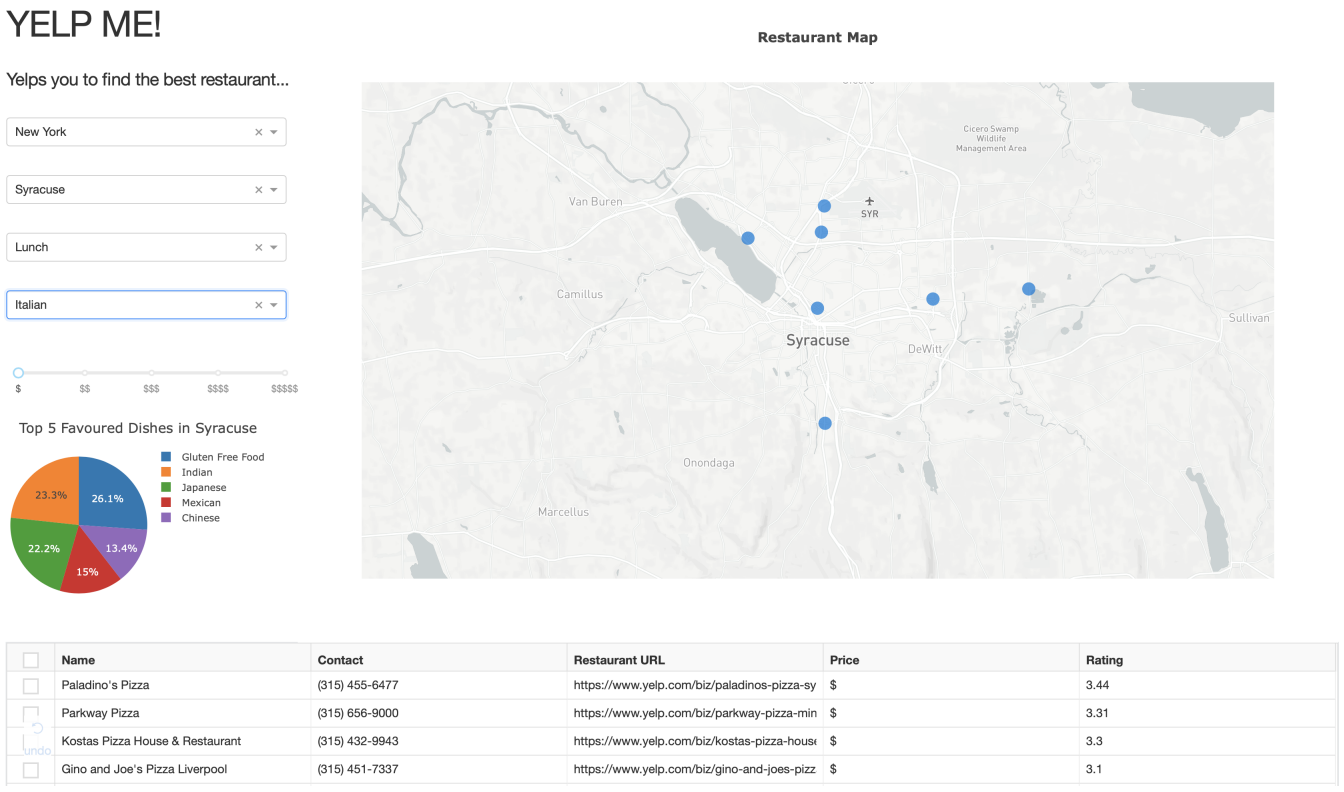
4.2 Sentiment Analysis

The sentiment score of a review is found out using TextBlob python package. This score ranges from -1 to +1. The sentiment score of all the reviews is calculated for a particular restaurant and a cumulative average of all these reviews is calculated and is further normalized on a scale of 0 to 5.

id	TB_score R1	TB_score R2	TB_score R3	TB_score_avg	TB_score_normalized
HIH8Xr7ePF1SjDTotvNqCg	0.34	0.4	0	0.25	3.12
GUGIFOAeTrj6bj-bwB025w	0.8	0.3	-0.4	0.23	3.08
lp-yx5mUB776EGUuXbouxA	0	0	0	0.00	2.50
55mbrvxSMijLVrE2qKP3Fw	0.4	0.1	-0.3	0.07	2.67
5w7Cv7JIWffBwi3ElYHH0Q	0.6	0.1	-0.3	0.13	2.83
r5Rvd27eHLb_NUq0HhcS2A	0	0.3	0.32	0.21	3.02
JJ5wlYTVENVnLVZhr7p1bQ	0.87	0.57	0.9	0.78	4.45
ZbDpNHcUtu9SvXT7hfS9g	-0.32	-0.5	0.3	-0.17	2.07
XWD1X5COj9HpY4jNb2q9-w	0.13	0.46	-0.42	0.06	2.64
l2sEA8T4kPgJOicXqI8KXA	0.5	-0.32	0.13	0.10	2.76

4.3 Visualizations

All of the visualizations are plotted using Plotly and Dash which is a web interface platfrom provided by Plotly. The visualizations provided by Dash include the dropdown menus, tables, sliders and pie charts. The geographic map is provided by Mapbox API which provides a mapbox access token and this token is given to plotly's Layout object to plot all the restaurants. Each plot has an associated update function that is called when different filters are applied. The update function changes the location map, table and the pie chart depending upon the input received from the user.



	Name	Contact	Restaurant URL	Price	Rating
☐	Paladino's Pizza	(315) 455-6477	https://www.yelp.com/biz/paladinos-pizza-sy	$	3.44
☐	Parkway Pizza	(315) 656-9000	https://www.yelp.com/biz/parkway-pizza-min	$	3.31
☐	Kostas Pizza House & Restaurant	(315) 432-9943	https://www.yelp.com/biz/kostas-pizza-hous	$	3.3
☐	Gino and Joe's Pizza Liverpool	(315) 451-7337	https://www.yelp.com/biz/gino-and-joes-pizz	$	3.1

Appendix A

List of Helper Functions

Function Name	Description
filter_function	Retrieves relevant data from restaurant database based on user applied filters .
displayOnTable	Generates data to be displayed on the table.
display_on_pie	Generates data to be displayed on the pie chart.
update_cities	Updates cities based on selected state name.
update_Disable	Sets disable value of city dropdown to False.
update_cuisine	Sets disable value of cuisine dropdown to False.
update_food	Sets disable value of food dropdown to False.
update_slider	Sets disable value of price slider to False.
update_table_col	Updates Table columns using data received from displayOnTable function
update_table_row	Updates Table rows using data received from displayOnTable function
update_selected_row_indices	Updates table when clicked on the data point on the location map.
display_content	Displays the pie chart.
update_graph	Updates the entire location graph when filters are selected.
request	HTTP request to Yelp API.
get_business	Query the Yelp Business API by a business ID.
get_reviews	Query the Yelp Review API by a business ID.
search	Query the Yelp Search API by a search term and location.