SI 507 Final Project Document

May Chang

Project code

Github repo for final project code

https://github.com/ruqinch99/SI-507-Final-Project

README

The program uses Ann Arbor's businesses data from the Yelp Fusion and Open Table, provides an interactive command-line prompt for users to choose data/visualization options, and provides restaurant recommendations that meet their needs. To run the program, please install all required python packages, change into the directory where you store the python and data file, and then run the python file. The interactive command-line prompt will inform and ask users' options for visualization and recommendation preferences.

The Yelp Fusion API provides access to Yelp content and data, which is ideal for the restaurant recommendation application project. I created an app on Yelp Fusion and received a private API key authentication to access all endpoints' information. To create your private API Key, please see detailed instructions here. I used my private API Key to access Ann Arbor's restaurant data for this project and saved it as a non-dynamic JSON file (yelp.json). I will also scrape data from OpenTable for restaurants in Ann Arbor. The data include information about names, ratings, reviews, price, cuisine, and location and will be stored as a CSV file (opentable.csv).

Required Python packages for my project to work

JSON, random, requests, webbrowser, pandas, plotly.express, selenium webdriver, and bs4 BeautifulSoup.

Data sources

Yelp Fusion API

Origin

Data: https://api.yelp.com/v3/businesses/search?location=Ann Arbor
Business search endpoint documentation:
https://www.yelp.com/developers/documentation/v3/business_search

Format

JSON file (yelp.json)

Data collection

I created an app on Yelp Fusion and received a private API key authentication to access all endpoints' information. I will be using an endpoint of Business Search to retrieve the information of businesses in Ann Arbor. It allows search for businesses by keyword, category, location, price level, etc. And the results include information about businesses, categories, location, phone, price, rating, review count, transaction types, etc. Ann Arbor's restaurants' data is saved as a non-dynamic JSON file drawn from Yelp Fusion API. The function get_yelp() sends GET requests to the Yelp API, returns a dictionary with information about the business in Ann Arbor, and saves data as a JSON file.

Summary of data

240 records available
240 records retrieved
Important fields/attributes of each record for my project

- name: restaurant's name
- url: URL for restaurant's page on Yelp
- image_url: URL of the photo for this restaurant
- review_count: number of reviews for this restaurant
- categories: list of category title and alias pairs associated with this business
- rating: rating for this restaurant (value ranges from 1, 1.5, ... 4.5, 5)
- coordinates: coordinates of this business
- Transactions: list of Yelp transactions that the business is registered for, currently supported values are pickup, delivery, and restaurant_reservation
- price: price level of the business, value is one of $, $$, $$$, and $$$$.
- location: location of this restaurant, including address, city, state, zip code, and country
- phone: phone number of the business

OpenTable

Origin

https://www.opentable.com/ann-arbor-restaurant-listings

Format

CVS file (opentable.csv)

Data collection

I scraped an OpenTable page of restaurants in Ann Arbor. Functions parse_html and get_opentable are being used to collect and cache data from one OpenTable site. The data is stored in a data frame and saved as a CSV file (opentable.csv).

Summary of data

365 records available
100 records retrieved
Important fields/attributes of each record for my project

- Name: restaurant's name
- Rating: rating for the restaurant, e.g., 4.7 stars out of 5
- Reviews: reviews of the restaurant, value in the data frame is one of Exceptional, Awesome, Excellent, Very Good, Good, and NA.
- Price: price level of the restaurant, value in the data frame is one of 2, 3, 4.
- Cuisine: cuisine of the restaurant
- Location: city of the restaurant

Data Structure

README describing two tree data structure

Tree by price level
The class Business is an object that stores all the information about a specific restaurant. Function tree_price(json_str) constructs a tree data structure, organized in the format of the dictionary, storing businesses by price levels. The Attribute (internal) nodes of tree_price includes $, $$, $$$, $$$$, NA. Restaurant (leaf) nodes are Business objects.

Tree by price level and rating
Function tree_price_rating constructs another tree data structure, storing business by price levels and rating. Attribute (internal) nodes are $, $$, $$$, $$$$, NA, and different ratings (e.g. 4.0, 4.5, 5.0). Restaurant (leaf) nodes are Business objects.

A python file that constructs trees from stored data

tree.py
This python file contains class Business, methods to constructs trees: tree_price(json_str) and tree_price_rating(json_str), and methods to save trees as JSON file: save_tree_price(json_str) and save_tree_price_rating(json_str).

JSON file with trees

tree_price.json and tree_price_rating.json

A python file that reads the json of trees

read_tree.py

Screenshots showing your data and data structures

Data from Yelp Fusion API


```
[{'id': 'uTMqhmpgfpDMLN3W3YvMeQ',
  'alias': 'frita-batidos-ann-arbor',
  'name': 'Frita Batidos',
  'image_url': 'https://s3-media3.fl.yelpcdn.com/bphoto/Dxn825cYU1eFEoUXClZ_Dg/o.jpg',
  'is_closed': False,
  'url': 'https://www.yelp.com/biz/frita-batidos-ann-arbor?adjust_creative=SR0bivTjF3Dxrqp6F5KI_A&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=SR0bivTjF3Dxrqp6F5KI_A',
  'review_count': 1901,
  'categories': [{'alias': 'cuban', 'title': 'Cuban'},
   {'alias': 'burgers', 'title': 'Burgers'}],
  'rating': 4.5,
  'coordinates': {'latitude': 42.2803651, 'longitude': -83.7491532},
  'transactions': ['pickup', 'delivery'],
  'price': '$$',
  'location': {'address1': '117 W Washington St',
   'address2': '',
   'address3': '',
   'city': 'Ann Arbor',
   'zip_code': '48104',
   'country': 'US',
   'state': 'MI',
   'display_address': ['117 W Washington St', 'Ann Arbor, MI 48104']},
  'phone': '+17347612882',
  'display_phone': '(734) 761-2882',
  'distance': 783.857113972638 6},
 {'id': 'Fv2VLzVj9ATLcTbFehTDjg',
  'alias': 'savas-ann-arbor',
  'name': "Sava's",
  'image_url': 'https://s3-media4.fl.yelpcdn.com/bphoto/DwLnoBuT5eDuAZLI8dxH-Q/o.jpg',
  'is_closed': False,
  'url': 'https://www.yelp.com/biz/savas-ann-arbor?adjust_creative=SR0bivTjF3Dxrqp6F5KI_A&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=SR0bivTjF3Dxrqp6F5KI_A',
  'review_count': 1201,
  'categories': [{'alias': 'bars', 'title': 'Bars'},
   {'alias': 'breakfast_brunch', 'title': 'Breakfast & Brunch'},
   {'alias': 'tradamerican', 'title': 'American (Traditional)'}],
  'rating': 4.0,
  'coordinates': {'latitude': 42.279605, 'longitude': -83.7409649},
  'transactions': ['delivery'],
  'price': '$$',
  'location': {'address1': '216 S State St',
   'address2': '',
   'address3': '',
   'city': 'Ann Arbor',
   'zip_code': '48104',
   'country': 'US',
   'state': 'MI',
   'display_address': ['216 S State St', 'Ann Arbor, MI 48104']},
  'phone': '+17346232233',
  'display_phone': '(734) 623-2233',
  'distance': 325.65864392541806},
 {'id': 'j7Ukf82vXmCYImj2smZkWQ',
```


Data from the OpenTable page

	Name	Rating	Reviews	Price	Cuisine	Location
0	The Earle	4.7 stars out of 5	Exceptional	3	Italian	Ann Arbor
1	Jim Brady's - Ann Arbor	4.4 stars out of 5	Awesome	2	American	Ann Arbor
2	Ashley's Restaurant	4.1 stars out of 5	Excellent	2	American	Ann Arbor
3	Shalimar Restaurant	4.3 stars out of 5	Awesome	2	Indian	Ann Arbor
4	Taste Kitchen	4.7 stars out of 5	Exceptional	3	Contemporary American	Ann Arbor
...	...	...	...	...	...	...
95	eve, Ann Arbor	4.5 stars out of 5	Awesome	3	Contemporary American	Ann Arbor
96	Mtotoro	NaN	NaN	2	Japanese	Ann Arbor
97	Nagomi Sushi - Downtown	NaN	NaN	3	Sushi	Ann Arbor
98	Black Pearl Ann Arbor	4.4 stars out of 5	Awesome	2	Seafood	Ann Arbor
99	Grange Kitchen & Bar	4.4 stars out of 5	Awesome	2	Farm-to-table	Ann Arbor

100 rows × 6 columns

Tree by price level

```
 <__main__.Business at 0x11ab40fa0>,
 <__main__.Business at 0x11ab40490>,
 <__main__.Business at 0x11ab40400>],
'$$$': [<__main__.Business at 0x11ab1e7c0>,
 <__main__.Business at 0x11ab1e940>,
 <__main__.Business at 0x11ab1e9d0>,
 <__main__.Business at 0x11ab1ef40>,
 <__main__.Business at 0x11ab1e580>,
 <__main__.Business at 0x11ab1e4c0>,
 <__main__.Business at 0x11ab1e070>,
 <__main__.Business at 0x11ab1e100>,
 <__main__.Business at 0x11ab2f250>,
 <__main__.Business at 0x11ab40520>,
 <__main__.Business at 0x11ab40940>],
'$$$$': [<__main__.Business at 0x11ab1ea30>,
 <__main__.Business at 0x11ab2f040>,
 <__main__.Business at 0x11ab40e80>],
'NA': [<__main__.Business at 0x11ab1e280>,
 <__main__.Business at 0x11ab1e850>,
 <__main__.Business at 0x11ab1e670>,
 <__main__.Business at 0x11ab1e5e0>,
 <__main__.Business at 0x11ab1e3d0>,
```

Tree by price level and rating

```
...  {'$': {4.0: [<__main__.Business at 0x11ab40310>],
       4.5: [<__main__.Business at 0x11ab40070>],
       3.5: [<__main__.Business at 0x11ab40550>],
       2.5: [<__main__.Business at 0x11abc10a0>],
       3.0: [<__main__.Business at 0x11abc10d0>]},
      '$$': {4.5: [<__main__.Business at 0x11aac1f70>],
       4.0: [<__main__.Business at 0x11aac1a30>],
       5.0: [<__main__.Business at 0x11ab40250>],
       3.5: [<__main__.Business at 0x11ab40730>],
       3.0: [<__main__.Business at 0x11ab402b0>],
       2.0: [<__main__.Business at 0x11abc1040>]},
      '$$$': {4.0: [<__main__.Business at 0x11ab40190>],
       4.5: [<__main__.Business at 0x11ab40220>],
       3.5: [<__main__.Business at 0x11ab40280>]},
      '$$$$': {4.0: [<__main__.Business at 0x11ab40760>],
       3.5: [<__main__.Business at 0x11ab40160>]},
      'NA': {4.5: [<__main__.Business at 0x11ab401f0>],
       4.0: [<__main__.Business at 0x11ab40130>],
       5.0: [<__main__.Business at 0x11ab402e0>],
       3.5: [<__main__.Business at 0x11ab405b0>],
       2.5: [<__main__.Business at 0x11abc1070>],
       1.0: [<__main__.Business at 0x11abc1100>],
       3.0: [<__main__.Business at 0x11abc1130>]}}
```

Interaction and Presentation Options

The program provides three different data visualizations of Ann Arbor's restaurants, created using Plotly, with an interactive command-line prompt.

1) A scatter plot showing Ann Arbor's restaurants' locations
2) A pie chart showing Ann Arbor's restaurants' price level distribution
3) A histogram showing Ann Arbor's restaurants' rating distribution

```
Welcome to Ann Arbor Restaurant Picker!
The program will provide you with visualizations about Ann Arbor's restaurants and pick restaurants according to your price preference.

Would you like to see a scatter plot showing Ann Arbor's restaurants' locations? (yes/no) yes
Please see the graph in pop-up window.

Would you like to see a pie chart showing Ann Arbor's restaurants' price level distribution? (yes/no) yes
Please see the graph in pop-up window.

Would you like to see a histogram showing Ann Arbor's restaurants' rating distribution? (yes/no) y
Please see the graph in pop-up window.
```

The program will provide an interactive command-line prompt for users to choose their price level and rating preferences. The recommendations will be described via text, and the detailed information of restaurants from Yelp and OpenTable will be provided via text. See the below screenshot for demonstration.

```
What is your preferred price level? ($, $$, $$$, $$$$, NA, Exit) $$
$$ restaurants have the following rating: [2.0, 3.0, 3.5, 4.0, 4.5, 5.0].

Would you want me to pick a restaurant with a specific rating for you? [2.0, 3.0, 3.5, 4.0, 4.5, 5.0], or any string for no preference. 3.0
I will pick a $$ restaurant with 3.0 rating for you.

Name: Zola Bistro
Location : 3030 Washtenaw Ave, Ste 101, Ann Arbor, MI 48104
Phone: (734) 477-8088
Price Level: $$
Rating: 3.0
Review Count: 344
Category(s): French, Turkish, Italian
Transactions type(s): delivery

Would you like to open Zola Bistro's Yelp site? (yes/no) no
Sure.

Would you like to see more information of Zola Bistro from OpenTable? (yes/no) yes

Name: Zola Bistro
Rating: 3.9 stars out of 5
Review: Excellent
Price: 2
Cuisine: Contemporary French / American
```

In addition to command-line prompts and different Plotly visualizations, the program will also ask if the user would like to open the yelp site of a specific restaurant. If the user says yes, the program will open the site in a web browser.

```
Would you like to open P.F. Chang's's Yelp site? (yes/no) yes
Please see a pop-up window.
```

All interactions are done with simple command-line prompts and the user's inputs.

Demo video

https://drive.google.com/file/d/1XiDULC-DAyTpM-jpv_SQpUFbF0rOE15o/view?usp=sharing