SI 507 – Final Project

SICHEN LIU sichenl@umich.edu

Project code:

1. Link to github: https://github.com/elainesichen/Final_project_si507

2. Demo video link: https://www.youtube.com/watch?v=NNVNUFTkpvQ

3. Special instructions for running code: ①Get daily weather information in Boston for the past six years through National Centers for Environmental Information(NOAA) API key; ②Two options for users to interact: search for restaurants information or weather and bike-sharing information in Boston.

4. Required Python packages:

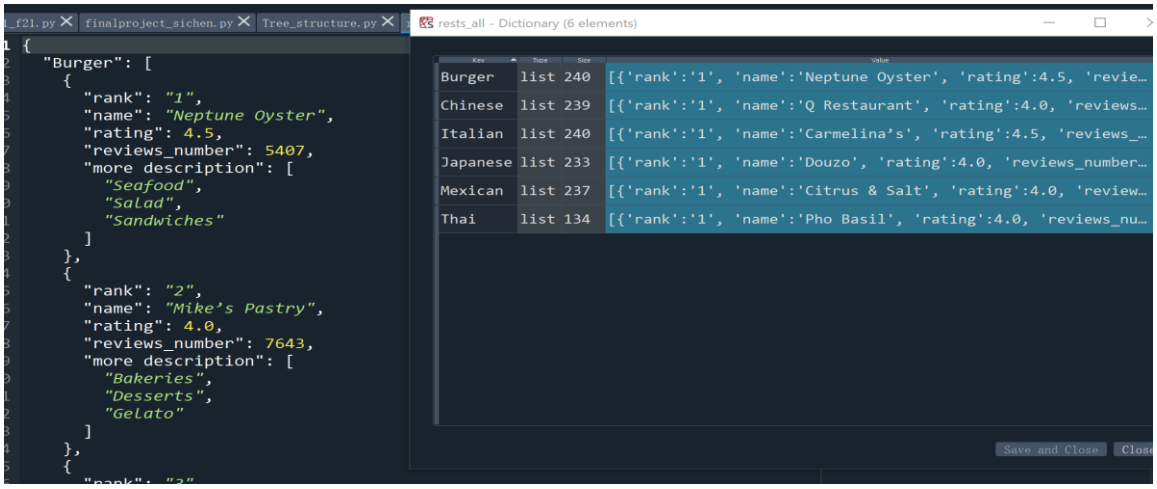
```
import noaa_api_v2
from time import sleep
import requests
from bs4 import BeautifulSoup
import secret_final
import json
import csv
import plotly.graph_objects as go
import plotly.express as px
import plotly.io as pio
pio.renderers.default='browser'
import Tree_structure
import print_tree
import pandas as pd
```

Data sources:

1. Yelp(restaurant_data): Web scraping Yelp restaurant listings using Requests, BeautifulSoup. Get details of restaurants that are listed for the city of Boston, Massachusetts.

Summary of data(json file: use web scraping and caching):

There are five attributes: ①restaurant's rank, ②name, ③rating(full score is 5), ④review numbers, ⑤more description(In addition to the main cuisine categories, more labels about the restaurant)



2. National Centers for Environmental Information(weather_data): Get daily weather information in Boston for the past six years through API key.

Summary of data (json file: use API key)

There are seven attributes: ①Date: 2015-01-01 to 2020-12-31; ②Min_temp: Minimum temperature (degrees C); ③ Max_temp: Maximum temperature (degrees C); ④ Avg_temp: Average temperature (degrees C) ; ⑤PRCP = Precipitation (mm); ⑥SNOW = Snowfall (mm); ⑦ AWND = Average daily wind speed (tenths of meters per second);


```
[
  {
    "date": "2015-01-01",
    "min_temp": -5.5,
    "max_temp": 0.6,
    "avg_temp": -3.3,
    "prcp": 0.0,
    "snow": 0,
    "awnd": 64
  },
  {
    "date": "2015-01-02",
    "min_temp": -0.5,
    "max_temp": 5.0,
    "avg_temp": 1.3,
    "prcp": 0.0,
    "snow": 0,
    "awnd": 57
  },
  {
    "date": "2015-01-03",
    "min_temp": -5.5,
    "max_temp": 2.8,
    "avg_temp": -2.1,
    "prcp": 15.7,
    "snow": 36,
    "awnd": 46
  },
  {
    "date": "2015-01-04",
    "min_temp": 1.7,
    "max_temp": 11.1,
    "avg_temp": 3.5,
    "prcp": 14.5,
    "snow": 0,
    "awnd": 37
```


3. Blue Bikes(bike_data): Obtain bike usage in Boston every day for the past six years. (Company 'Blue Bikes' shares the data through https://s3.amazonaws.com/hubway-data/index.html, and the data is publicly available from Umich's phpMyAdmin).

Summary of data (csv file :obtain from public resource)

To enrich the data information by join weather_data through date.

	A	B
1	date	bike_usage
2	1/1/2015	127
3	1/2/2015	213
4	1/3/2015	142
5	1/4/2015	112
6	1/5/2015	337
7	1/6/2015	267
8	1/7/2015	296
9	1/8/2015	247
10	1/9/2015	275
11	1/10/2015	162
12	1/11/2015	134
13	1/12/2015	279
14	1/13/2015	387
15	1/14/2015	379
16	1/15/2015	287

Data structure:

Organize the restaurant_data into a tree, the data of restaurants are firstly classified by cuisines, and then each cuisine category is classified by binary tree with a star rating of four or above, or below four stars.

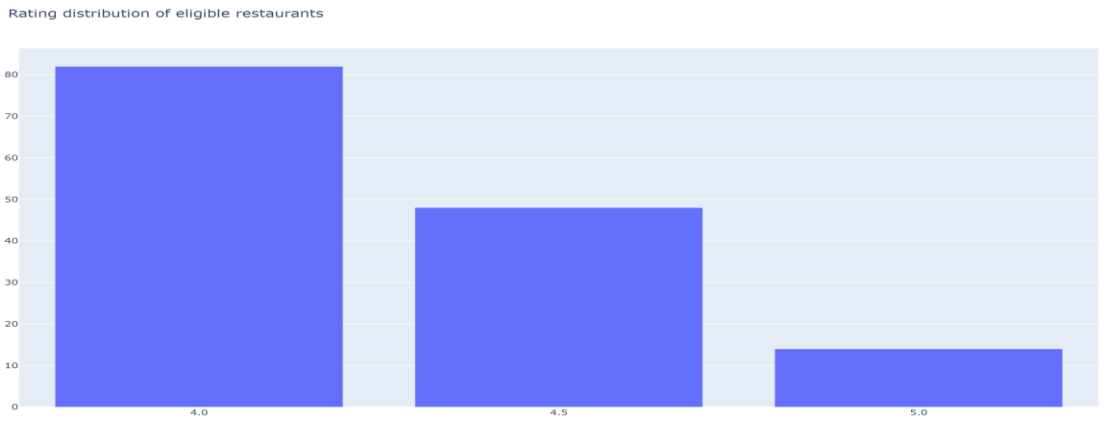
- python file that constructs trees: *Tree_structure.py*
- python file that print the tree: *print_tree.py*
- JSON file with the tree: *tree_restaurants.json*

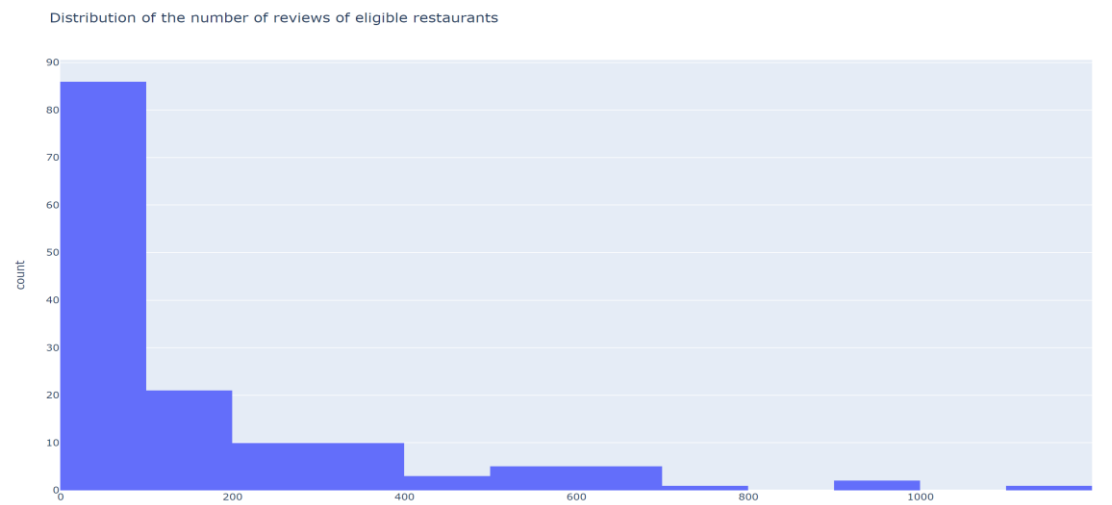
Screenshots:

```
This is a summary of information about the restaurant in Boston:
Is it Asian food?
+-Yes: Is it Chinese food?
| +-Yes: It is Chinese food
|   rating>=4.0 restaurants: Q Restaurant, Dumpling Xuan...
|   rating<4.0 restaurants: Gourmet Dumpling House, Taiwan Café...
| `-No:  Is it Japanese food?
|   +-Yes: It is Japanese food
|     rating>=4.0 restaurants: Douzo, Nagomi Izakaya...
|     rating<4.0 restaurants: Yamato II, Tsurutontan Udon Noodle Brasserie - Boston...
|   `-No:  It is Thai food
|     rating>=4.0 restaurants: Pho Basil, Kala Thai Cookery...
|     rating<4.0 restaurants: Thai Basil Restaurant, Montien Thai Restaurant...
`-No:  Is it Burger food?
  +-Yes: It is Burger food
    rating>=4.0 restaurants: Neptune Oyster, Mike’s Pastry...
    rating<4.0 restaurants: The Beehive, Union Oyster House...
  `-No:  Is it Mexican food?
    +-Yes: It is Mexican food
      rating>=4.0 restaurants: Citrus & Salt, Tenoch Mexican...
      rating<4.0 restaurants: Casa Romero, Temazcal Tequila Cantina...
    `-No:  It is Italian food
      rating>=4.0 restaurants: Carmelina’s, Giacomo’s Ristorante - Boston...
      rating<4.0 restaurants: Strega Italiano - Back Bay, Maggiano’s Little Italy...
```

<u>Interaction and Presentation Options</u>

1. If the user chooses to query restaurant information(for example: mexcian restaurants with star rating of 4 or above)

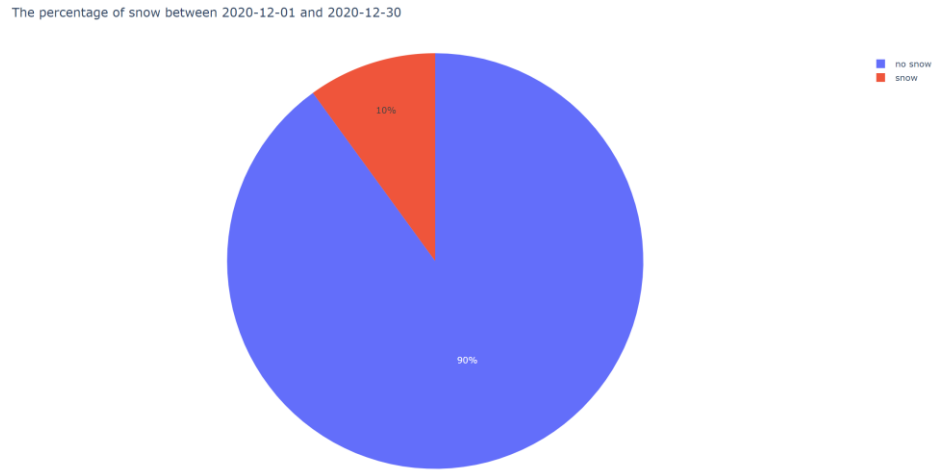




2. If the user chooses to query weather information and bicycles usage. ***For example:***

- The user wants to search for the average temperature on 2020.08.20, then show the change of the month that the date belongs to

- The user wants to search for the percentage of snow over a certain period of time



- The user wants to search for the bike usage over a certain period of time