Project Checkpoint

Fall 2021

SI 507 Final Project

Wenjie Wu

Data Checkpoint:

I have two data sources: Yelp and Wikipedia. I use Yelp's API to get the restaurant data, and use Wikipedia to scrap data of the cities by using the BeautifulSoup. As for cashing, I use JSON file to cache data from API.

- Yelp Fusion API: https://www.yelp.com/developers/documentation/v3/business_search.
  The data is stored in JSON file by using the location name that the user entered. For example, if the user wants to get the restaurant information in Ann Arbor, the name of the JSON file will be "yelp_ann arbor.json". Sample data is below:


{"businesses": [{"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=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 1900, "categories": [{"alias": "cuban", "title": "Cuban"}, {"alias": "burgers", "title": "Burgers"}], "rating": 4.5, "coordinates": {"latitude": 42.2803651, "longitude": -83.7491532}, "transactions": ["pickup", "delivery"], "price": "$$", "location": {"address1": "117 W Washington St", "address2": "", "address3": "", "city": "Ann Arbor", "zip_code": "48104", "country": "US", "state": "MI", "display_address": ["117 W Washington St", "Ann Arbor, MI 48104"]}, "phone": "+17347612882", "display_phone": "(734) 761-2882", "distance": 783.8571139726386}, {"id": "fQRc956jitK55RT6S-riGA", "alias": "zingermans-delicatessen-ann-arbor-2", "name": "Zingerman's Delicatessen", "image_url": "https://s3-media1.fl.yelpcdn.com/bphoto/IzbO8i-Teq6khNRrZSGU4A/o.jpg", "is_closed": false, "url": "https://www.yelp.com/biz/zingermans-delicatessen-ann-arbor-2?adjust_creative=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 2243, "categories": [{"alias": "delis", "title": "Delis"}, {"alias": "salad", "title": "Salad"}, {"alias": "sandwiches", "title": "Sandwiches"}], "rating": 4.0, "coordinates": {"latitude": 42.2844740041972, "longitude": -83.74527753066885}, "transactions": ["pickup", "delivery"], "price": "$$", "location": {"address1": "422 Detroit St", "address2": "", "address3": "", "city": "Ann Arbor", "zip_code": "48104", "country": "US", "state": "MI", "display_address": ["422 Detroit St", "Ann Arbor, MI 48104"]}, "phone": "+17346633354", "display_phone": "(734) 663-3354", "distance": 920.7909068298882}, {"id": "4REtzXpQYy8dVev8RjWbSQ", "alias": "mani-osteria-and-bar-ann-arbor", "name": "Mani Osteria & Bar", "image_url": "https://s3-media4.fl.yelpcdn.com/bphoto/-xLufUqnuyQ6M5kDVpalMQ/o.jpg", "is_closed": false, "url": "https://www.yelp.com/biz/mani-osteria-and-bar-ann-arbor?adjust_creative=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 853, "categories": [{"alias": "italian", "title": "Italian"}, {"alias": "pizza", "title": "Pizza"}, {"alias": "cocktailbars", "title": "Cocktail Bars"}], "rating": 4.0, "coordinates": {"latitude": 42.279579374976 3, "longitude": -83.74424470261l}, "transactions": ["delivery"], "price": "$$", "location": {"address1": "341 E Liberty St", "address2": "", "address3": "", "city": "Ann Arbor", "zip_code": "48104", "country": "US", "state": "MI", "display_address": ["341 E Liberty St", "Ann Arbor, MI 48104"]}, "phone": "+17347696700", "display_phone": "(734) 769-6700", "distance": 400.4088835060728}, {"id": "fv2VLaVj9ATLc7bFehTDjg", "alias": "savas-ann-arbor", "name": "Sava's", "image_url": "https://s3-media4.fl.yelpcdn.com/bphoto/DwLnoBuT5eDuAZLIKdxB-Q/o.jpg", "is_closed": false, "url": "https://www.yelp.com/biz/savas-ann-arbor?adjust_creative=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 1200, "categories": [{"alias": "bars", "title": "Bars"}, {"alias": "breakfast_brunch", "title": "Breakfast & Brunch"}, {"alias": "tradamerican", "title": "American (Traditional)"}], "rating": 4.0, "coordinates": {"latitude": 42.279605, "longitude": -83.7409649}, "transactions": ["delivery", "price": "$$", "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": "W4WNypFs--aQfpG0V5gsxQ", "alias": "blank-slate-creamery-ann-arbor", "name": "Blank Slate Creamery", "image_url": "https://s3-media3.fl.yelpcdn.com/bphoto/YVKmdXyCUwVEYT3YES7YNw/o.jpg", "is_closed": false, "url": "https://www.yelp.com/biz/blank-slate-creamery-ann-arbor?adjust_creative=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 470, "categories": [{"alias": "icecream", "title": "Ice Cream & Frozen Yogurt"}], "rating": 4.5, "coordinates": {"latitude": 42.279645, "longitude": -83.7513053}, "transactions": [], "price": "$$", "location": {"address1": "300 W Liberty St", "address2": "", "address3": "", "city": "Ann Arbor", "zip_code": "48103", "country": "US", "state": "MI", "display_address": ["300 W Liberty St", "Ann Arbor, MI 48103"]}, "phone": "+17349225383", "display_phone": "(734) 922-5383", "distance": 889.3056548088025}, {"id": "ZJVhCAjBeKlzLhgRVVJD5Q", "alias": "jolly-pumpkin-cafe-and-brewery-ann-arbor-8", "name": "Jolly Pumpkin Cafe & Brewery", "image_url": "https://s3-media1.fl.yelpcdn.com/bphoto/CKYWheQyrANzs9qIfnectA/o.jpg", "is_closed": false, "url": "https://www.yelp.com/biz/jolly-pumpkin-cafe-and-brewery-ann-arbor-8?adjust_creative=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 923, "categories": [{"alias": "pubs", "title": "Pubs"}, {"alias": "breweries", "title": "Breweries"}, {"alias": "tradamerican", "title": "American (Traditional)"}], "rating": 4.0, "coordinates": {"latitude": 42.279168, "longitude": -83.748462}, "transactions": ["delivery"], "price": "$$", "location": {"address1": "311 S Main St", "address2": "", "address3": "", "city": "Ann Arbor", "zip_code": "48104", "country": "US", "state": "MI", "display_address": ["311 S Main St", "Ann Arbor, MI 48104"]}, "phone": "+17349132730", "display_phone": "(734) 913-2730", "distance": 652.9924714967 9}, {"id": "yWIYH9O4lmlJEyR8-M_LNw", "alias": "aventura-ann-arbor", "name": "Aventura", "image_url": "https://s3-media2.fl.yelpcdn.com/bphoto/ofTEQ919bA0DOIoKxAEpew/o.jpg", "is_closed": false, "url": "https://www.yelp.com/biz/aventura-ann-arbor?adjust_creative=uQ24qiN_PuXAYAjOTwz9DA&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=uQ24qiN_PuXAYAjOTwz9DA", "review_count": 761, "categories": [{"alias": "tapas", "title": "Tapas Bars"}, {"alias": "spanish", "title": "Spanish"}, {"alias": "bars", "title": "Bars"}], "rating": 4.0, "coordinates": {"latitude": 42.280283, "longitude": -83.746494}, "transactions":

  Moreover, if the user wants to get the restaurant reviews in Ann Arbor, the name of the file will be "review_ann arbor.json". Sample data is below:


[{"name": "Frita Batidos", "reviews": ["One of the best places to get burgers in Ann Arbor! Opens late so you can stop by after a long day. My favorite has to be the hibiscus Batidos, a bit sweet,...", "on first look, the man at the counter didn't give me a birthday churro, after we asked for said churro we received it. if you come here on you birthday do...", "What can I say about Frita's. If my Yelp check-in count is correct, I have eaten here 43 times (over the past 8 years). Frita's has (in my very scientific...")}, {"name": "Zingerman's Delicatessen", "reviews": ["First off, this place is delicious! Sometimes the line can be ungodly long, however, it is well worth the wait. The staff is delightful and efficient. The...", "I mean, obviously this place is 5 stars. Best deli in the Midwest, amazing delivery options (got my daughter's 19th birthday cake delivered to her at...", "We were here a while back, since my brother likes their sandwiches, pastrami mostly. I don't eat beef so Turkey it is for me. \n\nMy husband and I tried their...")}, {"name": "Mani Osteria & Bar", "reviews": ["I've been wanting to check out Mani for a while because of the pedigree (it's sister restaurant Isalita is good too). Finally got the chance and it didn't...", "I'm born and raised in New Jersey and I lived in Rome, so I have pretty high standards when it comes to Italian food. That being said, I thought Mani was a...", "I read great reviews on Yelp and the pictures looked great so my friends and I decided to come here to eat on a Saturday night. I was hoping the experience...")}, {"name": "Sava's", "reviews": ["Sava's is a must if you're in downtown Ann Arbor! I've been here a few times and it never disappoints me. If you're looking for a restaurant to go on a...", "I am so happy I found Sava's on Yelp and Trip Advisor. This is definitely going to be the go-to restaurant for me whenever I'm in town visiting my son at...", "Sava's is clearly popular, but I have to say that I think it is overrated. Overall it was probably 3.5 stars for the food but I rounded down since we got..."]}, {"name": "Blank Slate Creamery", "reviews": ["First things first, if you haven't been to Blank Slate in Ann Arbor then you're missing out!\n\nSo I've been here a few times and it's always a hit. If you've...", "What a delightful treat! Popped by here on a warm Saturday afternoon in Sept and not surprisingly the line was out the door. It moved fairly quickly and...", "Their lavender lemon honey is sooo good. I love lavender ice cream flavors and this was super good. For those of you who say lavender anything tastes like...")}, {"name": "Jolly Pumpkin Cafe & Brewery", "reviews": ["Beer? Wonderful. Come here for drinks and nice ambiance.\nFood? Wack. That tofu bibimbap sandwich....the tofu wasn't really well flavored. The truffle fries...", "This place is pretty reliable. Emphasis on reliable. They've got an array of beer options and a smaller cocktail option. \n\nA few times I've been they were...", "What a bummer of a dinner plan. We swung through Ann Arbor on our way to northern Michigan for a night and were excited to grab a bite to eat here. Upon...")}, {"name": "Aventura", "reviews": ["We had reservations for four for a going away dinner. Our waitress was very friendly and knowledgeable about the menu, which is divided into four sections:...", "Pre-covid this used to be my go to spot for happy hour. They had amazing sangria fro $5 dollars and discount on some of their tapas. Before covid I would...", "Came here on a Saturday early evening. The noise level is a little higher than some places but you could still have a good conversation. Tables were..."]}, {"name": "Zingerman's Roadhouse", "reviews": ["Wait time was about a half hour but we had a party of 7. \n\nEggnog latte was really good. Glad I got it.\n\nMy daughter got the 4oz steak and mashed potatoes...", "Came here when I wanted some comfort food. I wanted fried chicken and mac and cheese. I was ecstatic to see there was a fried chicken mac and cheese option,...", "This review is based on the fact that the menu clearly identifies what is gluten free or adaptable to be gluten free. I drove over an hour for brunch and..."]}, {"name": "Juicy Kitchen", "reviews": ["Great smoothies and breakfast! My daughter and I shared BLT omelette and Spanish chocolate smoothy. Take out service only. Also nice menu of coffee choices.", "Gluten-free, vegan, and vegetarians look no further!!! This healthy breakfast cafe offers so many yummy options for all diets. For being a smaller...", "I've written it in other reviews but one of the things I miss most about dining out is breakfast and brunch. A hearty plate of eggs, pancakes, and...")}, {"name": "TeaHaus", "reviews": ["TeaHaus is a fantastic spot. Their macarons are very tasty, their tea selection is expansive, and they have great tea items and mugs. During the pandemic...", "FYI--the last weekend of every month they offer carryout high tea for $35 per service. It's AMAZING and it's so nice to get all my faves from the tea...", "Such a cute place to catch up with friends or study. The food is quite pricey and sometimes doesn't warrant the cost, but the daintiness of it all certainly..."]}, {"name": "Tomukun Noodle Bar", "reviews": ["This place is great. I gave them 4 stars because their communication isn't too good with each other. The other night the kitchen closed earlier than usual...", "I am trying to post reviews not just for places I'm in love with. Therefore, I must be completely honest and say that I wished this place focused more on...", "I had leveled my expectations going in, because let's be honest, one look at the menu and anyone can tell that it's not going to be authentic.\n\nThere are..."]}, {"name": "Pacific Rim", "reviews": ["Savory Pan Asian!\n\nWhen the waiter tells you the tuna arrived that morning, it's a no brainer what to order. \pacificrim_a2 definitely delivered during our...", "Went on a date on a Tuesday night. The gentleman who sat us was very friendly and nice. The decor is elegant, and they even have a room they can block out...", "Place:\nThis place is on main street. This place is definitely bring your parents for college students or on a special occasion. The place has some asian..."]}, {"name": "Monahan's Seafood Market", "reviews": ["Their fish and chips is incredible. The fish is mild and moist, and the breading stays crisp even through the fifteen minute drive home. The tartar sauce is...", "I stopped in late on Friday (1:45 pm) after getting off the train. Hungry, I stopped by the area knowing I could find options. Even though it's past the...", "Are you a seafood fanatic? Then get yourself to Monahan's. Are you a HARDCORE seafood fanatic? Then head to their website and sign up for the \"Friday Fish..."]}, {"name": "RoosRoast Coffee", "reviews": ["Good service nice sitting area outside social distancing. Street parking in front and parking in the rear. The area is very quiet. They have umbrellas to...", "I really do love this charming neighborhood coffee spot. They are doing the best they can in spite of this dreadfully prolonged pandemic. And their flavored...", "Best coffee shop in Ann Arbor! Not only the coffee is fantastic but friendly baristas! Cosy indoor decoration, social distance and nice outdoor seating...."]}, {"name": "Isalita", "reviews": ["Great food and the drinks are amazing ! The staff also is really helpful guiding you to the tastiest experience.", "First things first, the vibe here is incredible! We got seated immediately, I came with my mom :). \n\nNE music + decor is really nice, I enjoyed it a...", "We enjoyed sitting outside and eating tacos and guac while on a vist to the University of Michigan. \n\nIt was easy to walk to from campus, and we didn't

- Wikipedia: https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population.
  The name of the page is "List of United States cities by population". I use the BeautifulSoup to scrap the information of cities from the page. I will not store the data I

scrapped as a file. I stored the data inside the python file as a dictionary after data cleaning.

Interactive Presentation Design:

I use interactive command shell to present the data. I will give users many choices to help users go through the tree. The user can choose whether or not to get the data and get which part of data. The user should be able to choose to get the information of restaurants and cities of location they are interested in. Also, I have several graphs to visualize the data I got. I will try to use different types of graphs to present the data by using different libraries like matplotlib or plotly. In order to improve the readability of my data structure, I will try to visualize my tree in a better way.