

Project Proposal for Software Design

Dean Gladish, Hashir Safdar and Tony Qiang

Working Title

An Inquiry into Restaurant Language: Analyzing Yelp Reviews State by State

Tagline: *Restaurant feedback across the U.S & Canada.*

One Sentence Summary/Site Purpose:

This site recommends and connects Epicureans around the U.S and Canada with specific restaurants based on their interests and a massive scope of previously-generated Yelp reviews, connecting them with reviewers that share the same tastes.

Audience:

User Constituencies

- Students (teenagers)
 - Students with spare time are willing to hang out after studying a whole day. They will use the reviewers' information to look for recommendations on good restaurants and interesting places for hanging out.
 - Teenagers are individuals who are always curious and want to try something new. Therefore, they might look for some brand new and fancy places to experience
- Businessman (workers)
 - Businessmen always have a lot of social engagements. Checking the reviews on the site, they can find different places in terms of restaurants, coffee houses, and hotels to hold business meetings easily.
 - After a whole day's working, workers in a big company might want to have a dinner party with their coworkers and colleagues.
- Tourism
 - When tourists travel around to some new places, they are willing to accept recommendations for local foods. Finding the local people's reviews on our site, they can pick the restaurant to taste the most native foods.

- Old people (parents)
 - Old people are individuals with more spare time and tend to be more interested in investing in restaurants with a good reputation.
 - Parents would like to find places to relax during the weekend with their family members.
- Researchers
 - Researchers might want to compare the data from the site and use to compare one specific restaurant to the other.
 - Meanwhile, they would like to evaluate how people from different countries appreciate different kinds of food specifically. By comparing the reviews from divergent users under the same restaurant, they can find a pattern with people's different tastes.
- Restaurant Owners
 - Restaurant owners typically hold a lot of stake in what their customers think of them and so want to attract customers; they want to adapt their opening or closing time (for example, to be very late at night if necessary) in order to make things more convenient for their customers.
 - Additionally, they want to be able to see how popular related restaurants (by food category) are and get in better touch with the website's clients in order to provide reservations for events and other services.
- Restaurant Customers
 - Restaurant customers are people who come from a wide variety of backgrounds; however they all have the same interest in reviewing restaurants and seeing what other people have said before trying to go to a specific place; they also want to be able to have better access to reviews without having to sift through a bunch of text in order to pick the best table if they're having a party.
 - They also might want to know if online ordering is an option or not, and all of these things can be easily provided via search/filtering.
- Food Critic
 - Food critics are especially important because they are the ones most likely to rate and review restaurants; they'll want to filter restaurants by the type of cuisine or whether it's a cafe, pub or buffet, etc., or even distinguish between vegetarian/non-vegetarian options in order to specify their preferences.

- They don't just want to rate their experience on a numeric scale - they want to see others' opinions in an easily interpretable (adjustable graphics which we will provide) format; they want to see what the most popular foods are and what they look like.
- If it's a restaurant like Dominoes, they'll want to gain some perspective on how well the delivery guys have been on following their routes and being on time.
- For food critics, it'll be crucial to provide a list of curated restaurants and maybe, for the sake of convenience/time, provide a list of the most recent or top reviews in order that they'll have the latest and most up-to-date information.

Dataset Metadata:

1. Name and Authorship

Yelp Dataset, Yelp Inc.

2. About

This dataset (the "Data") contains a subset (10 metropolitan areas across 2 countries) of Yelp's review, business, and user data as JSON files ([documentation](#)). It contains exactly 3.58 gigabytes of data, which should be an asset.

3. External Source

Yelp Website - <https://www.yelp.com/dataset/download>

4. Recommended Citation:

Yelp Inc. (2019). *Yelp Dataset*. Retrieved from:
<https://www.yelp.com/dataset>

5. Terms of Use:

- a. Dataset allows for academic use. Full terms of use [here](#).
- b. "Yelp grants you a royalty-free, non-exclusive, revocable, non-sublicensable, non-transferable, fully paid-up right and license during the Term to use, access, and create derivative works of the Data in electronic form for academic purposes only."

- c. Our use of the data is fully binding to our institution and we agree to be bound by the Data Agreement (the “Agreement”) by our access and/or use of the data.
- d. We are allowed to access the data locally for this academic project, which fulfills the requirements of being part of an ongoing course of study; Yelp reserves the right to review our usage of the data and furthermore reserves the right to make any changes to the Agreement. If we wish not to be bound by such terms we must immediately cease use of the data and delete all local copies.
- e. Yelp grants us a royalty-free, revocable right and license to use, access and create derivative works in electronic format for academic purposes only.
- f. Restrictions: using our data for promoting our business listing/any commercial purpose/disclose our analysis to third parties not covered under the Agreement/use the data in such a manner that can be *reasonably* interpreted to suggest that Yelp is the author/infringe any Yelp intellectual property or reputation/use the Data for any purpose prohibited by law.
- g. We are not entitled to support by Yelp and their liability is limited to \$50.00. We will provide written certification to Yelp that we have destroyed the data if we terminate the Agreement. Upon termination, we are still subject to the Restrictions and may contact Yelp with questions.

6. Local Copy

Downloaded on 1/26, 6 .json files:

business.json

An assortment of information. Rows are businesses described by address, city, state, postal code, latitude, longitude, stars, review count, open status, good for kids, noise level, delivery service, accepts credit cards, by appointment only, etc.

checkin.json

A list of check-in dates, organized by business id.

Dataset_Challenge_Dataset_Agreement.pdf

All terms of service, which are delineated in the same 5. “Terms of Use” above.

photo.json

A basic row-by-row list of photo id, caption (if it exists), business id, and label (indoors or outdoors, restaurant or not).

review.json

Full customer reviews.

tip.json

Contains information, based on user id and business id, on short blurbs of advice on specific dates, and includes a count of compliments.

user.json

Full user information.

Yelp_Dataset_Challenge_Round_13.pdf

This describes the 13th round of the Yelp Dataset Challenge, which took place from January 15, 2019 to December 31, 2019. In it, there's more information about requirements and individual/team requisites (18 or older, enrolled in an accredited institution, not restricted by the Office of Foreign Assets Control).

Summary of Goals:

Users of this site have a variety of wishes including but not limited to:

- Pre-select restaurants based on the type of food/beverages they offer, and have their decision sent to them by email (.pdf format).
- Compare the same restaurants between a particular city and another city
- Filter restaurant results by open/close time (both day and time of day) in order to make informed choices
- If possible, restaurant customers might be given information on how to make a reservation (or whether such an option is currently available or is likely to be made available in the future) at a specific venue.
- Find and follow certain Yelp users whose tastes line up with the users, to help find appropriate recommendations.
 - We'll do this by building tables, making interactive graphs and providing lists of reviews based on specific search/filtering features.

Requirements

Functional

- A search function to find restaurants that are related to certain tags.
 - tags obtained from common words used in reviews for the business
- Provide tailored recommendations based on user's tastes.
- Allow the user to save restaurants for a personal list.

Non-functional

- Create a taste profile for a user based on restaurants they like.
- Store user data securely and remember users

Team Collaboration Plan:

We intend to make this project fully collaborative and crystal-clear with regard to our intention (which is to hold people accountable, set specific deadlines for each stage of the project while being fully adaptable and ask questions as needed) in order to accomplish what we want. Dean, Hashir and Tony will focus mainly for now on the translation of the dataset (which is in .json format currently) into .csv or similar format, which will then be posted on GitHub. We will establish bi-weekly meeting patters; we will meet once in-class and once outside of class in order to establish progress. While working on our own, we will ensure contact amongst the group through our team Slack channel. We don't know exactly which role we will participate in, since we don't know what we need to do for the project yet. However, we will all engage collaboratively in the project and assign roles if needed as the project progresses.

Personas (also in the */Personas* directory of the repository)

- **Fatima, 42-year old entrepreneur:**

- Fatima is a 42-year old entrepreneur living in New York City. She works with a range of clothing and goods businesses that serve primarily under resourced communities. As part of her job, she needs to meet with her partners in different areas of the city, often for lunch. When she's not working, she likes to spend time with her family and indulge in their shared love for food, whether at home or outside.
- As Fatima is self-employed, she spends most of her time with her work, and does not have much energy to search for restaurants, for work or for home. She really enjoys Mexican food, but not ones that are too Americanized. Other cuisines she's a fan of includes Thai and Indian. Her love for food brings her to many cuisines, and often she will ask who she is going with for their cuisine preference.

- **Bill, 24-year old Epicurean:**

- Bill is a 24 year old Epicurean food connoisseur who wants to filter restaurants by cuisine, restaurant type, vegetarian/non-vegetarian preference, see the most popular items, critique online payment options, plot restaurants by distance and location, and have the easiest method of delivering his critiques (e.g. sending them back to Yelp).

- As an avid food traveller, he likes to plan ahead for his trips. He wants to be able to save his preferences for later with the click of a button and see up-to-date information such as the most recent/best reviews. Just before travelling, he likes having a paper copy of the list of restaurants he wants to eat at.

Persona:

Bobby is a stay-at-home dad who wants to find restaurants which are close enough that he can get food delivered straight to his address without hassle. He wants to know how long the food delivery typically takes in order to make the best choice. He also wants to know if he can pay cash on delivery and whether he can contact the delivery person directly.

Persona:

Esa is a food critic who wants to add or remove various food items to find the restaurant which best fits those choices.

Persona:

Marisa is a legal assistant who doesn't have the time to actually visit the restaurant but still wants to call them periodically so that she can check the progress of her order.

User Stories (*also in the /UserStories directory of the repository*)

- As a student, I want to find restaurant choices which I can hang out with my friends after studying for a whole day.
- As a businessman, I want to look for places where I can set up social engagements with some partners.
- As Fatima, I want to find restaurants that others with similar tastes have liked so that I can take my family out to eat tonight.
- As Bill, I want to use filters to find specific types of restaurants so that I can read the best reviews by trusted users and save a list of places I want to go.
- As a tourist, I want to experience the local customs by searching for local restaurants.