

SI 206 Final Project Plan

Jean Gamboa Fuentes (jcgamboa)

Henry Collins-Thompson (henryct)

December 16, 2024

1. Group Name

Team Long Last Names

2. People in Group

Henry Collins-Thompson: henryct@umich.edu

Jean Gamboa Fuentes: jcgamboa@umich.edu

3. What APIs/websites will we be gathering data from?

- a. Yelp Fusion API: <https://api.yelp.com/v3>
- b. Census Bureau API (ACS 5- Year Data): <https://api.census.gov/data/>

4. What data will we collect from each API and store?

Yelp Fusion API:

- Business name
- Zip Code
- Average Rating (1-5 scale)
- Number of Reviews
- Business Category

Census Bureau API:

- Zip Code
- Median Household Income
- Population

5. What data will we be calculating from the data in the database?

We will first start by calculating the Average Yelp Rating per zip code. From there, we will try to find correlation between the median household income and other demographic data and the average yelp rating per zip code. Finally, we will find the distribution of businesses by number of reviews or income range.

6. What visualization package will we use?

We will use Matplotlib for our visualization creation.

7. What graphs/charts will we be creating?

- a. Scatter plot: Visualizing the correlation between median household income and average Yelp ratings by zip code
- b. Bar chart/histogram: Showing the number of businesses in income ranges

8. Who is responsible for what?

Jean is responsible for obtaining the data from the Yelp API and storing it in the table. Henry will be responsible for obtaining the census data from the API and storing that information into the table. Both of us will pair-program together on the calculations, and each one of us will be in charge of a visualization. Both will then work on the final report together