

For my final project, I am planning to develop a program that provides a user with average ratings of restaurants from two different sources: Google and Yelp. To do so, I will connect to both Google Places and Yelp Fusion API's to extract ratings of restaurants based on a specific city, state, and restaurant type indicated by the user. After doing so, my program will organize the data in an SQL database, creating two different tables for data acquired from Google and Yelp. The two tables will be bridged together based on the city, state and name of the restaurant, which will allow me to analyze data points from both of those sources together. Additionally, I plan on using food inspection data from data.gov in order to calculate the number of food inspection violations for a specific state or city, if that data is available. Here are the data sources I plan to use:

- <https://www.yelp.com/fusion>
- <https://developers.google.com/places/>
- Data.gov

Here are the anticipated point allocations of my project

- Web API you haven't used before that requires API key or HTTP Basic authorization ⇕ (Yelp Fusion)
 - **4 points**
- Web API you've used before (Google Places)
 - **2 points**
- CSV or JSON file you haven't used before with > 1000 records
 - **2 points**
- Total score: **8**

Here's an example of what my program would do. A user would enter a city and state name in addition to two types of restaurants in those city (Indian and Mexican for example). My program would then calculate the average ratings of overall Indian restaurants and average ratings of overall Mexican restaurants within a 40000 meter radius. Then, my program would plot those visuals comparing restaurant ratings from both of those from the two sources – Yelp and Google and will display that through graphs using plotly. The data presentation tools are the interactive command prompt that allows the user to choose a type of graph they want to have displayed. Additionally, the user will have the option of seeing the number of food inspection complaints in that city extracted from the data.gov website.

For my final project—I am planning on developing a tool that allows a user to input the name of a company they hope to explore and a location in which they hope to do so. My program will scrape LinkedIn to find the employees that work for that company in the specified location using the advance filter options. After doing so, program will create a dictionary of universities that the employees of these companies graduated from. The value corresponding to each university (key) will be the total number of employees that come from that university who are currently working for that company. The program will finally output the top 10 represented universities at that company. Additionally, my program will scrape a college ranking website (U.S. News Best College Rankings) to extract and output the ratings of those top 10 universities in the united states. My main stat source will be LinkedIn: <https://www.linkedin.com> and U.S. News Best College Rankings <https://www.usnews.com/best-colleges/rankings/national-universities>. Here are my data challenge score components:

- Web API you haven't used before that requires OAuth (6 points): LinkedIn **+6**
- Scrape a new single page: U.S. News Best College Rankings **+4**
- Total score: **10**

This project will allow me to scrape at least 100 records that have 5 corresponding fields. My presentation will include an interactive prompt that allows a user to look up a company and then choose a presentation option for the top 10 schools represented at the company whether that is a pie chart or a bar or a map graphic. Additionally, a user will be allowed to select one of the schools to display their rankings in addition to having a graphic that compares the ratings of the top 10 represented schools.

The data presentation tools are the interactive command prompt that allows the user to choose a type of graph they want to have displayed—whether that's a bar graph, or a pie chart or a map graphic. I will be using plotly graphics in addition to the Flask app to allow the user to choose data visualization options and display those in HTML tables.