Winter 2022 - HW8

In this homework, you will select data from a database, process it, and create a visualization using Matplotlib. This is similar to the final steps of your pipeline for the final project.

We have provided:

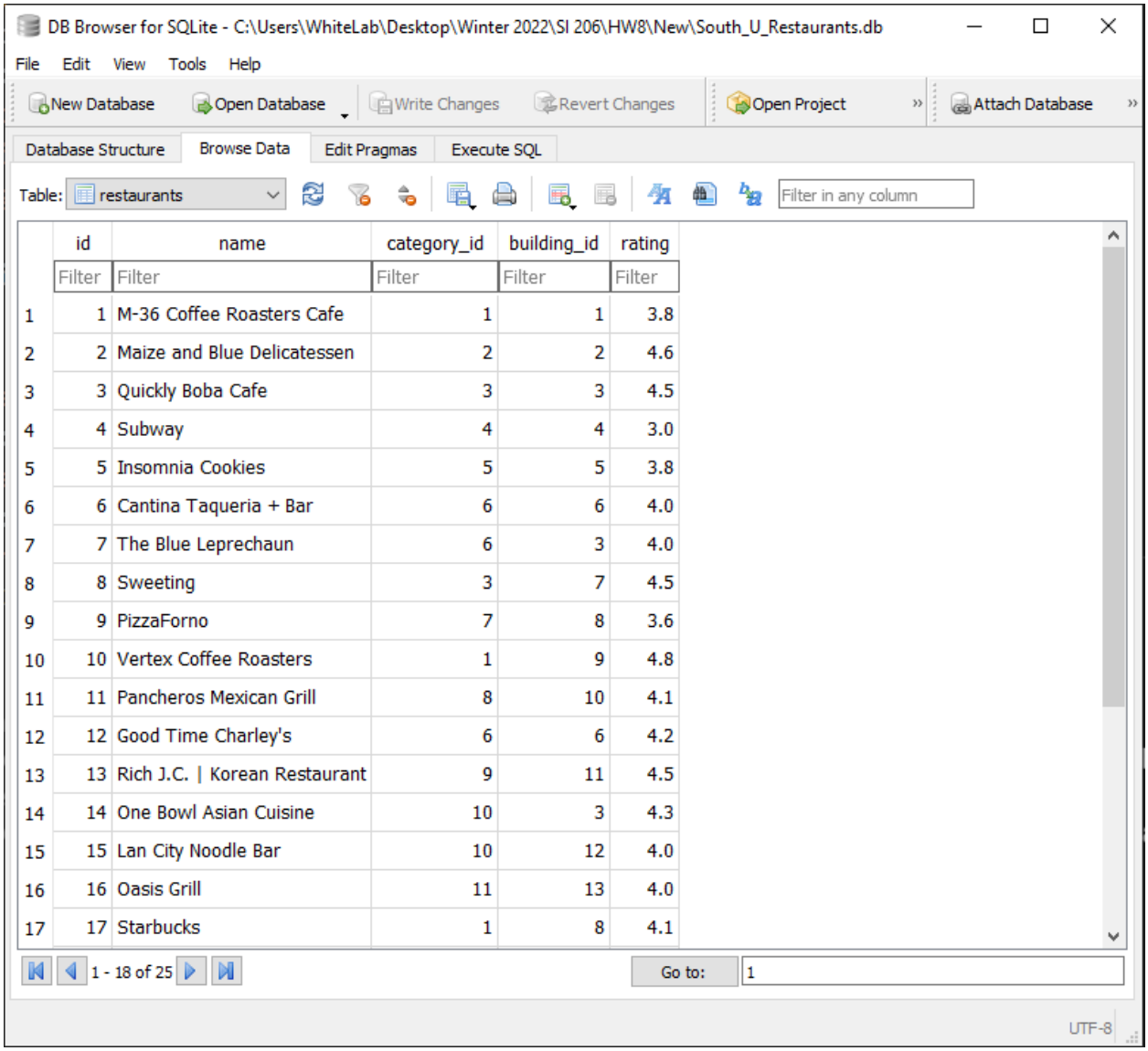
- *South_U_Restaurants.db* - a database with local restaurant data collected from Google.
- HW8.py - starter code for the functions below.

Make sure you are using Anaconda python for this assignment (preferred), or have installed Matplotlib on your own (using `pip install matplotlib` or another installation method).

Part 0: Look at the database

Check out *south_u_restaurants.db* in your DB Browser for SQLite program.

1. Open DB Browser for SQLite
2. Click on “Open Database” and choose South_U_*Restaurants.db*.
3. Click on Browse Data
4. Take some time to familiarize yourself with the table and column names



	id	name	category_id	building_id	rating
	Filter	Filter	Filter	Filter	Filter
1	1	M-36 Coffee Roasters Cafe	1	1	3.8
2	2	Maize and Blue Delicatessen	2	2	4.6
3	3	Quickly Boba Cafe	3	3	4.5
4	4	Subway	4	4	3.0
5	5	Insomnia Cookies	5	5	3.8
6	6	Cantina Taqueria + Bar	6	6	4.0
7	7	The Blue Leprechaun	6	3	4.0
8	8	Sweeting	3	7	4.5
9	9	PizzaForno	7	8	3.6
10	10	Vertex Coffee Roasters	1	9	4.8
11	11	Pancheros Mexican Grill	8	10	4.1
12	12	Good Time Charley's	6	6	4.2
13	13	Rich J.C. | Korean Restaurant	9	11	4.5
14	14	One Bowl Asian Cuisine	10	3	4.3
15	15	Lan City Noodle Bar	10	12	4.0
16	16	Oasis Grill	11	13	4.0
17	17	Starbucks	1	8	4.1

Part 1: Process the data

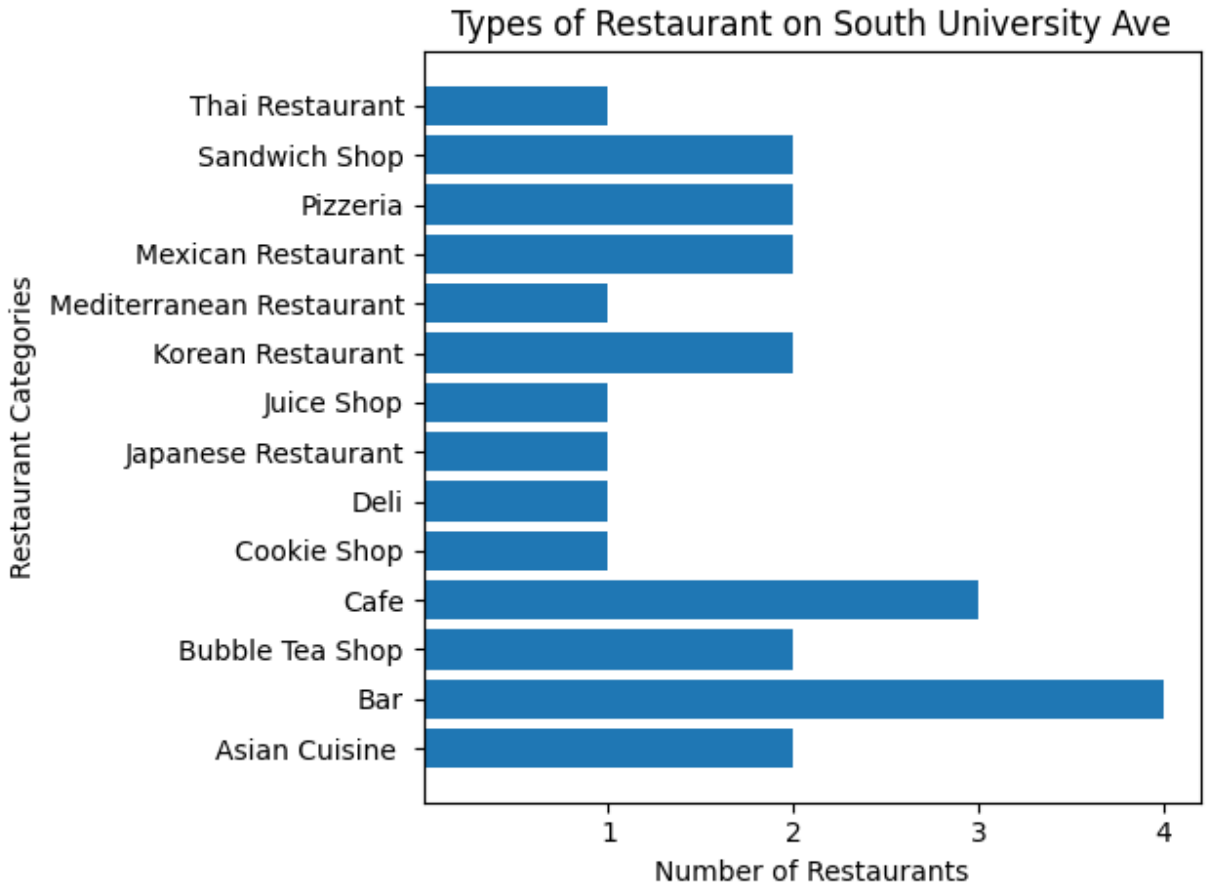
Complete the *get_restaurant_data(db_filename)* function that accepts the filename of the database as a parameter, and returns a list of dictionaries. The key:value pairs should be the name, category, building, and rating for each restaurant. The list should look like:

Expected output:
[{'name': 'M-36 Coffee Roasters Cafe', 'category': 'Cafe', 'building': 1101, 'rating': 3.8}, . . .]

Your function must pass all the unit tests to get full credit.
Note: Because all of the restaurants are on the same street (in this case, South University Ave), the addresses only contain the building numbers.

Part 2: Visualize the data

Complete the function *barchart_retaurant_categories(db_filename)*, which accepts the filename of the database as a parameter and returns a dictionary. The keys should be the restaurant categories and the values should be the number of restaurants in each category (**hint:** use the SQL COUNT keyword). The function should also create a horizontal bar chart with restaurant categories along the y-axis and the counts along the x-axis.

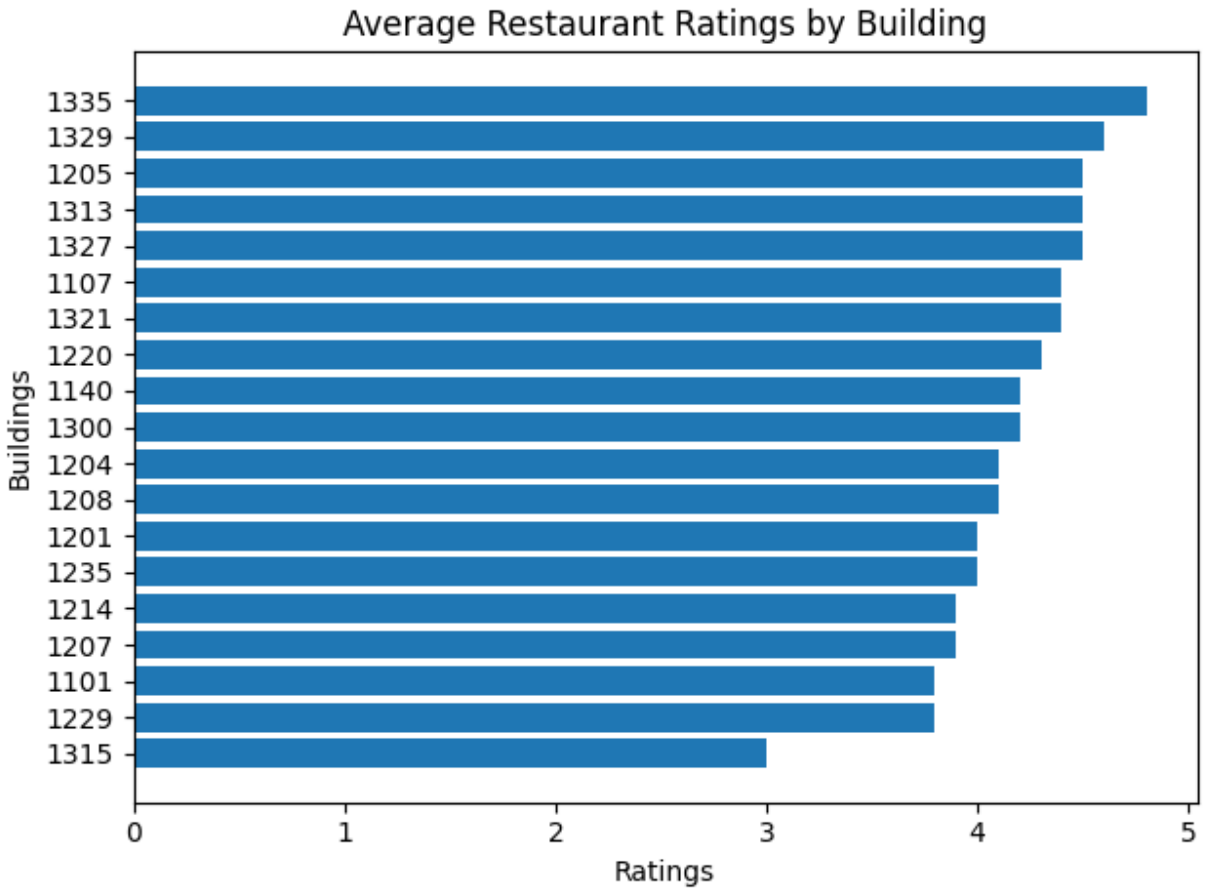


Expected Output:
{'Asian Cuisine ': 2, 'Bar': 4, 'Bubble Tea Shop': 2, 'Cafe': 3, 'Cookie Shop': 1, 'Deli': 1, 'Japanese Restaurant': 1, 'Juice Shop': 1, 'Korean Restaurant': 2, 'Mediterranean Restaurant': 1, 'Mexican Restaurant': 2, 'Pizzeria': 2, 'Sandwich Shop': 2, 'Thai Restaurant': 1}
Submit an image file of your bar chart to Canvas, along with your repository link.

Extra credit: Visualize more data

A lot of the restaurants on South U share buildings. Let's write a function to determine which building has (on average) the highest rated restaurants.

Complete function *highest_rated_building(db_filename)* to plot a barchart. The y-axis will be the numbers of the different buildings. The x-axis will be the average rating for the restaurants in each building (**hint**: use the AVG keyword when writing your query). The average values should be rounded to one decimal place. Sort the y-axis in **descending order** from top-to-bottom by rating. The chart must have appropriate axis labels and a title. Your chart should look like this:



Finally, this function should return a tuple containing the number of the building with the highest rated restaurants and the average rating of the restaurants in that building.

Expected Output
('1335', 4.8)

Grading

Code passes all unit tests	14 pts (2 pts per unit test with 7 tests)
Submission of bar chart image file	5 pts
Created a bar chart from the data	26 pts
Title on bar chart	5 pts
Informative X-axis label on bar chart	5 pts
Informative Y-axis label on bar chart	5 pts
Correct code and image file for extra credit	*6 pts extra credit*
Total	**60 pts + 6 pts extra credit**