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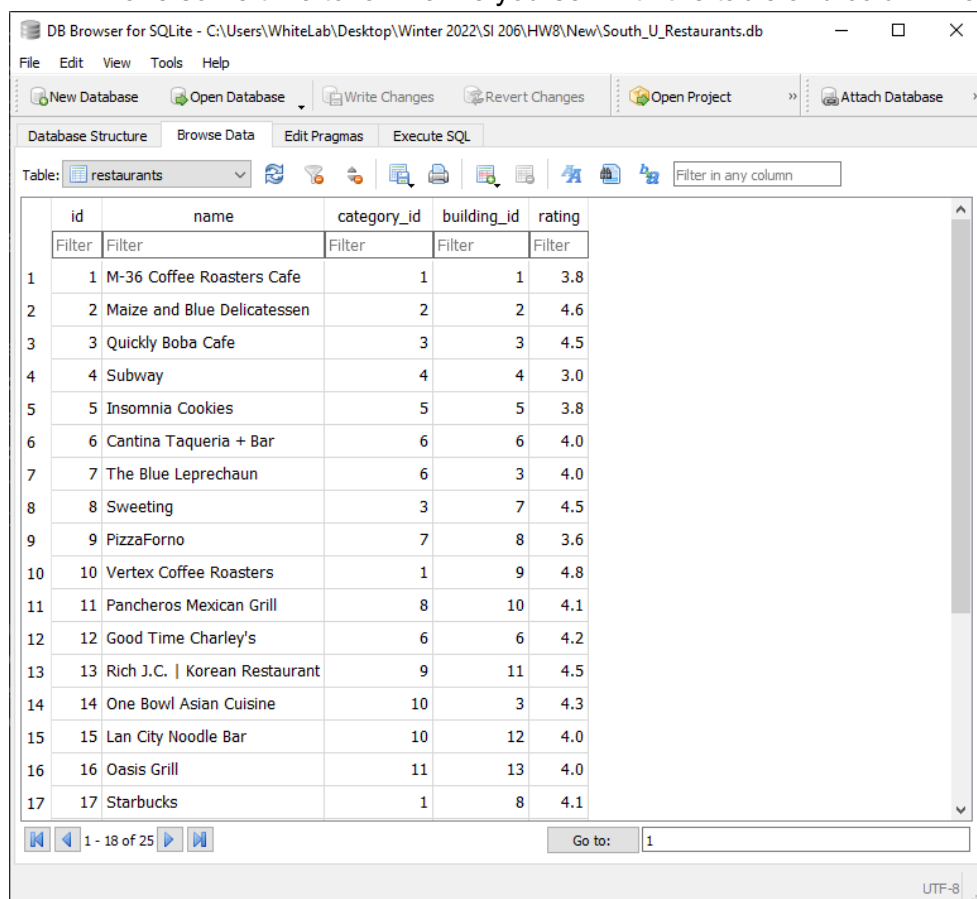
We have provided:
- *South_U_Restaurants.db* - a database with local restaurant data collected from Google.
- HW8.py - starter code for the functions below.

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).

Before you start: 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

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 return value:
[{'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_restaurant_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).

Expected return value:
{'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}

The function should also create a bar chart (horizontal or vertical – figure out which one gives a better visualization) with restaurant categories along one axis and the counts along the other axis. In the chart, the counts should be descending from one side to the other.

Example chart:

Types of Restaurant on South University Ave
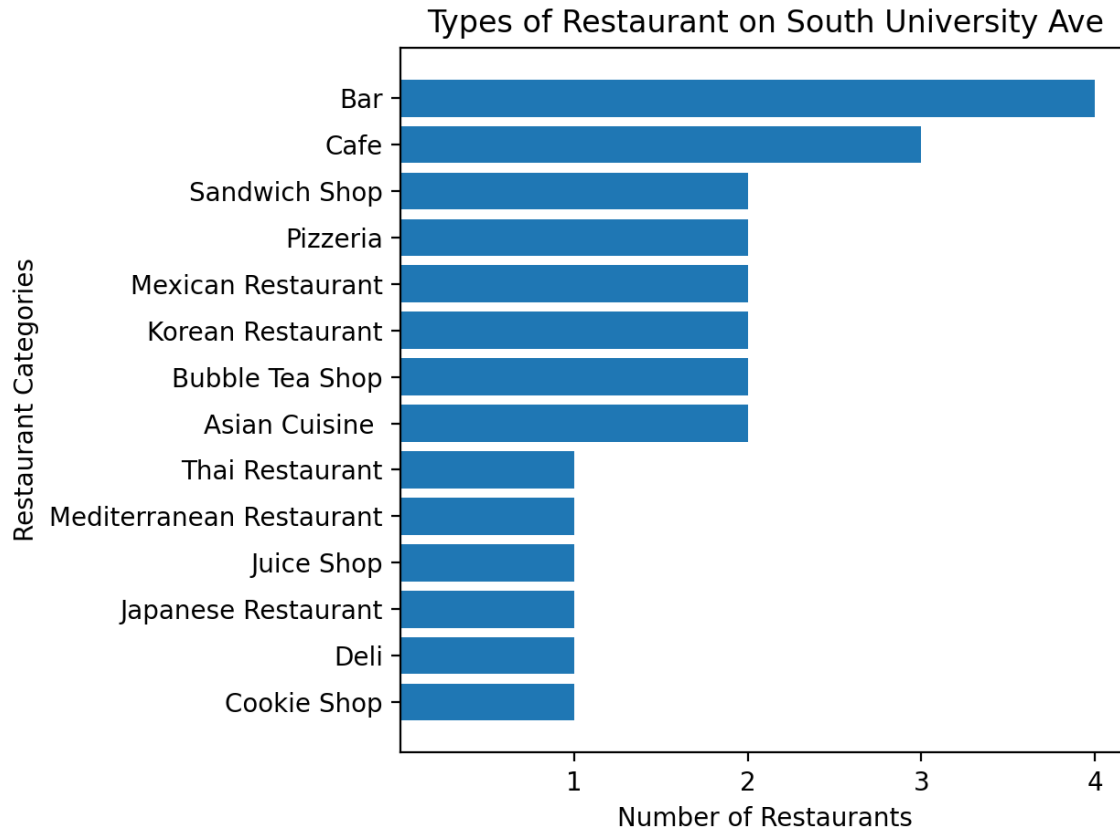
Submit an image file of your bar chart to Canvas, along with your repository link.

Extra credit: Visualize more data

Let's write a function to determine which category of restaurant has (on average) the highest rated restaurants.

Complete function *highest_rated_category(db_filename)* to plot a barchart. The y-axis will be the numbers of the different types. The x-axis will be the average rating for the restaurants in each type (**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:

Average Restaurant Ratings by Category
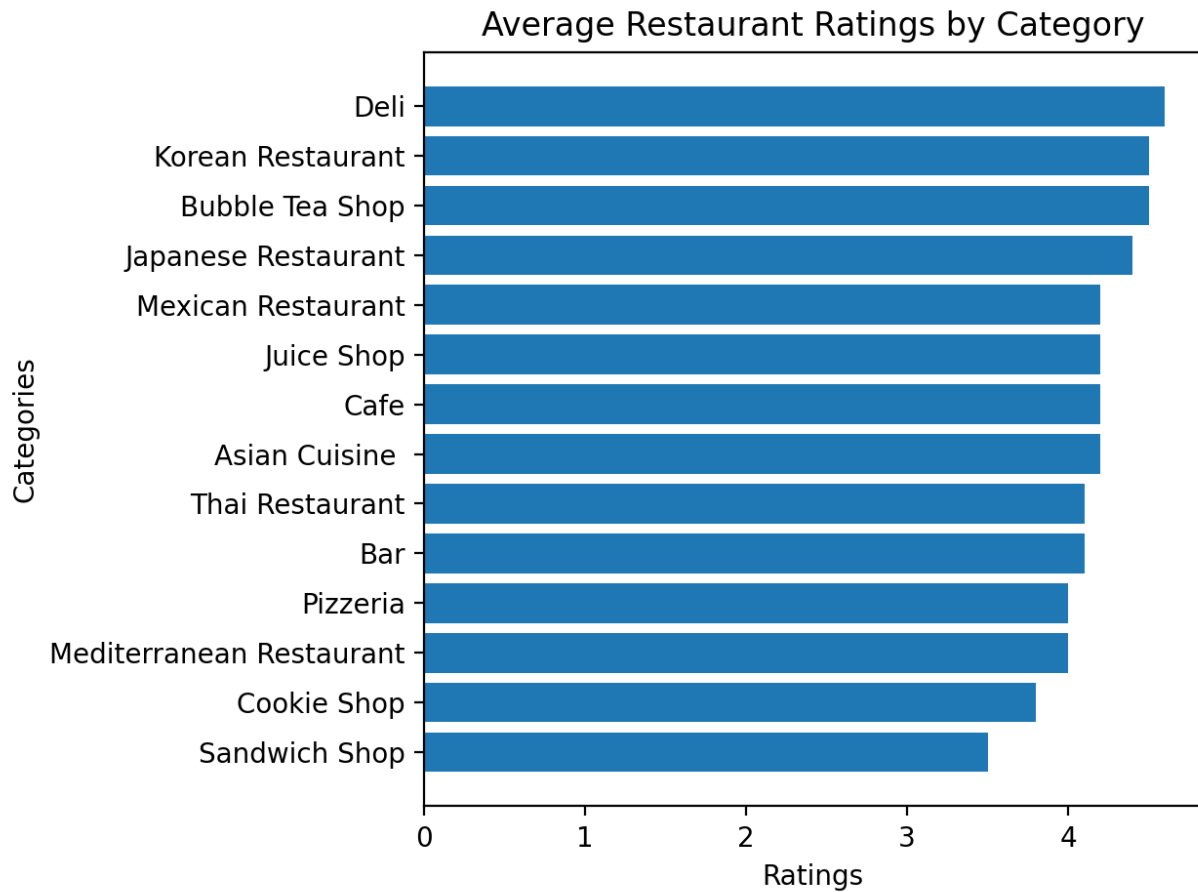
Finally, this function should return a tuple containing the highest rated restaurant category name and the average rating of the restaurants in that category.

Expected Output
('Deli', 4.6)

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**