

FALL 2021 - 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:

- *Restaurants.db* - a database with local restaurant data collected from Yelp.
- 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](#)

[Part 1: Process the data](#)

[Part 2: Visualize the data](#)

[Extra credit: Visualize more data](#)

[Grading](#)

Part 0: Look at the database

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

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

Database Structure Browse Data Edit Pragas Execute SQL								
Table: Restaurants					Filter in any column			
	restaurant_id	name	address	zip	category_id	rating	price	
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	
1	Xcuu9bTjW62Q36yUkv-JVQ	Dom Bakeries	1305 Washtenaw Rd,Ypsilanti	48197	0	4.5	1	
2	RRtWTpa15xaDDBtKly8ksw	White Lotus Farms	7217 W Liberty Rd,Ann Arbor	48103	1	5.0	2	
3	JXcDdgg4RZhSyr4ojCyY5A	NeoPapalis	500 E William St,Ann Arbor	48104	2	4.5	2	
4	5yrNbYde_PmHfmNv1fCjDw	Aamani's Smokehouse & Pizza	2529 Dexter Ave,Ann Arbor	48103	3	4.5	1	
5	qWMflzMymWBxgGR_NAaTsA	Zingerman's Bakehouse	3711 Plaza Dr,Ann Arbor	48108	0	4.5	2	
6	4REtzXpQYy8dVev8RjWbSQ	Mani Osteria & Bar	341 E Liberty St,Ann Arbor	48104	4	4.0	3	
7	8Ww_4J_4pqXh5EFb0bhsTQ	Wolverine State Brewing Co	2019 W Stadium,Ann Arbor	48103	5	4.5	2	
8	fQ8c9S6jitKS5RT6S-ziGA	Zingerman's Delicatessen	422 Detroit St,Ann Arbor	48104	6	4.0	2	
9	-h_zeuiMCWWkSIAYhw25EA	Anthony's Gourmet Pizza	1508 N Maple Rd,Ann Arbor	48103	7	4.5	2	
10	yNIYH9041m1JEyRS-N_LNw	Aventura	216 E Washington St,Ann Arbor	48104	8	4.0	3	
11	Lb3kPdkKFjpcbq_JwvXFlg	Pizza Perfect	332 S Ford Blvd,Ypsilanti	48198	7	4.5	1	
12	ZJVhCAjBeRlzlHgRVVJD5Q	Jolly Pumpkin Cafe & Brewery	311 S Main St,Ann Arbor	48104	5	4.0	2	
13	XFlmIH0ZKR3wTxKpDevhcQ	Detroit Street Filling Station	300 Detroit St,Ann Arbor	48104	9	4.0	1	
14	WQGUTUmR5uge1Z6PZYd4Rw	Mr Spots	808 S State St,Ann Arbor	48104	10	4.0	2	
15	p4Nad3u6PD03cJEn8VtiEw	Tippins Market	4845 Ann Arbor Saline Rd,Ann Arbor	48113	11	4.5	2	
16	qkw4xWWgTufvBs1NcxsFnw	Vinology Restaurant & Event Space	110 S Main St,Ann Arbor	48104	12	4.0	3	
17	s8x9YIRRASt8h_l41VzVCw	Joe's Pizza	1107 S University Ave,Ann Arbor	48104	7	4.5	1	
18	a-mgl_xrcFdmUyjeF0mS2Q	Jolly Pumpkin Artisan Ales + Kitchen	2319 Bishop Cr E,Dexter	48130	13	4.5	3	
19	EtjYm9PGiPZkdBA5mBCieA	Stadium Market	1423 E Stadium Blvd,Ann Arbor	48113	11	4.5	2	
20	ILsG4YKZfM1c2QPKeZFvgQ	Ashley's Restaurant	338 S State St,Ann Arbor	48104	5	3.5	3	
21	M3wssPEKzZb3hm4P0UmCdQ	Grizzly Peak Brewing Company	120 W Washington St,Ann Arbor	48104	13	3.5	2	

Part 1: Process the data

Complete the `get_restaurant_dict(db_filename)` function that accepts the filename of the database as a parameter, and returns a dictionary with the number of restaurants with a price level of 2 (“\$\$”) for each category. The keys should be the category names and the values should be the number of restaurants with a price level of 2. The dictionary should look like:

Expected output:

```
{'Farmers Market': 1, 'Salad': 2, 'Bakeries': 1, 'Pubs': 2, 'Delis': 2, 'Pizza': 8, 'Sandwiches': 1, 'Beer, Wine & Spirits': 2, 'Breweries': 1, 'Sports Bars': 1, 'American (Traditional)': 1, 'Italian': 1, 'Beer Bar': 1}
```

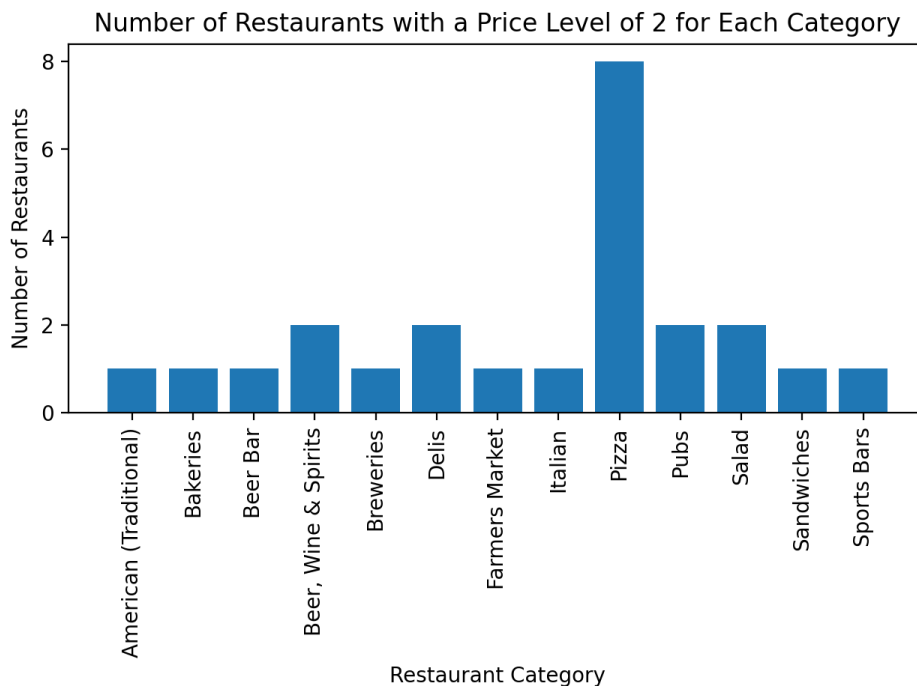
Hint: Use WHERE clause, refer to this week’s discussion.

Your function must pass all the unit tests to get full credit.

Part 2: Visualize the data

Complete the function `barchart_restaurants_and_price(cat_dict)`, which takes in a dictionary created by the function in Part 2, and uses matplotlib functions to draw a bar chart with restaurant categories on the x axis, and the number of restaurants with a price level of 2 in that category on the y axis. The chart must have appropriate axis labels and a title.

Sort the X axis alphabetically from left-to-right. Your chart should look like this:



*Some categories do not have any restaurant of price 2. Based on the result, not all categories are plotted.

Finally, this function should sort the dictionary items alphabetically and return the resulting list of tuples. Your list should look like:

```
answer_data = [  
    ('American (Traditional)', 1), ('Bakeries', 1), ('Beer Bar', 1),  
    ('Beer, Wine & Spirits', 2), ('Breweries', 1), ('Delis', 2), ('Farmers Market', 1),  
    ('Italian', 1), ('Pizza', 8), ('Pubs', 2), ('Salad', 2), ('Sandwiches', 1), ('Sports Bars', 1)]
```

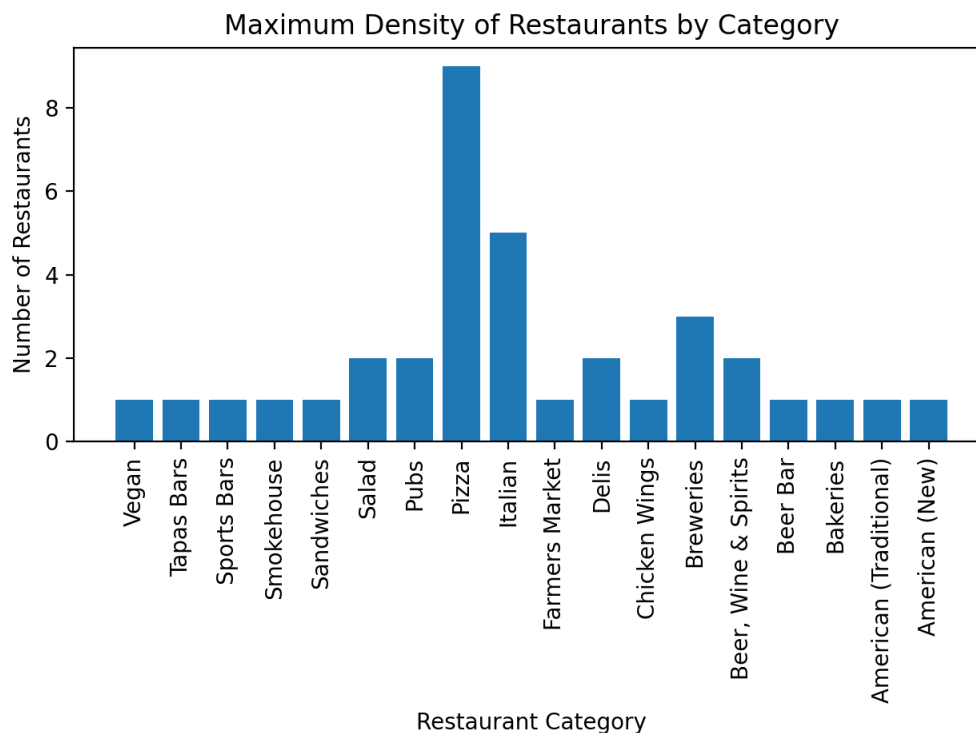
Submit an image file of your bar chart to Canvas, along with your repository link.

Extra credit: Visualize more data

One way to analyze restaurant data is to understand how dense it is for a specific category in one zip area. For example, the pizza category has 2 restaurants at 48104, and 5 at 48106. In this case, pizza is more dense at 48106 than 48104. The more restaurants with the same category in one zip area, the more competition there would be.

Complete function *maximum_density_by_category(.)* to plot a barchart. The x axis will be the restaurant category. The y axis will be the number of restaurants at their most densely located zip code (i.e., maximum density).

Sort the x axis in **alphabetically descending order** from left-to-right. The chart must have appropriate axis labels and a title. Your chart should look like this:



Finally, this function should return a dictionary where the keys are restaurant categories. The values should be a nested dictionary, where the keys are zip code and values are the number of restaurants for that zip code and corresponding category. Your dictionary should look like this:

```
{ 'Bakeries': { '48197': 1, '48108': 1 }, 'Farmers Market': { '48103': 1 }, 'Salad': { '48104': 2 }, 'Smokehouse': { '48103': 1 }, 'Italian': { '48104': 5 }, 'Pubs': { '48103': 1, '48104': 2 }, 'Delis': { '48104': 2 }
```

```
2}, 'Pizza': {'48103': 2, '48198': 3, '48104': 9, '48108': 3,
'48111': 1, '48197': 2}, 'Tapas Bars': {'48104': 1}, 'Vegan':
{'48104': 1}, 'Sandwiches': {'48104': 1}, 'Beer, Wine & Spirits':
{'48113': 2}, 'American (New)': {'48104': 1}, 'Breweries': {'48130':
1, '48104': 3}, 'Sports Bars': {'48176': 1}, 'American
(Traditional)': {'48104': 1}, 'Chicken Wings': {'48130': 1}, 'Beer
Bar': {'48103': 1}}
```

Submit an image file of your chart to Canvas.

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
<i>Correct code and image file for extra credit</i>	<i>6 pts extra credit</i>
Total	60 pts + 6 pts extra credit