Top Restaurant Types in Sacramento

Grant Bringham

July 10, 2020

1. Introduction

The State Capitol of California is located in Sacramento. Sacramento is a large county with population of 508,519 in 2018 according to the US Census Bureau. The Capitol is located in downtown Sacramento which has a daytime population of more than 100,000 people. Downtown Sacramento has around 150 restaurants. (cityofsacramento.org)

Our client wants to open a new restaurant near the State Capitol of California.

They want to know "Where in Downtown Sacramento would it be best to open a restaurant?" Our client will need to decide whether they want to compete with a lot of restaurants with a high market demand, or fewer restaurants with a lower market demand. --- Our client decides to go with the higher market demand and therefore would like to see what areas would be good to add a restaurant to. Our client also decides they want to add this popular restaurant type to an area with more restaurants nearby. Our client also decides they want to make a new Mexican restaurant.

Questions to answer:

What are the most common types of restaurants?

What areas have more restaurants in them?

What are proven good areas to have Mexican restaurants?

2. Data acquisition and cleaning

2.1 Zip Code Data

I started with getting a dataset in the form of a csv of all the zip codes and their geographical coordinates in the United States. I began with reading in a csv of zip code and coordinate data from a public open data site (https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/). I then organized the dataset to show only zip code, that they were all in Sacramento, and their latitude and longitude. I then trimmed this list down to only the 5 zip codes around the downtown area of Sacramento.

2.2 Foursquare location data

After cleaning and organizing data of the zip codes, I then proceeded to get geolocator data to get the coordinates of Sacramento to later be used in the Folium map of the data. I then created a dataframe from Foursquare of the restaurants in each zip code.

3. Exploratory Data Analysis/ Methodology

For my exploratory data analysis, I started by using the data from the Foursquare API to mark on a map using folium, the restaurants locations along with the zip codes. Foursquare is a powerful API and is used by both Apple Maps and Uber. The map shown in Figure 1 shows the 5 zip codes as green dots and the restaurant data for each zip code is a different color.
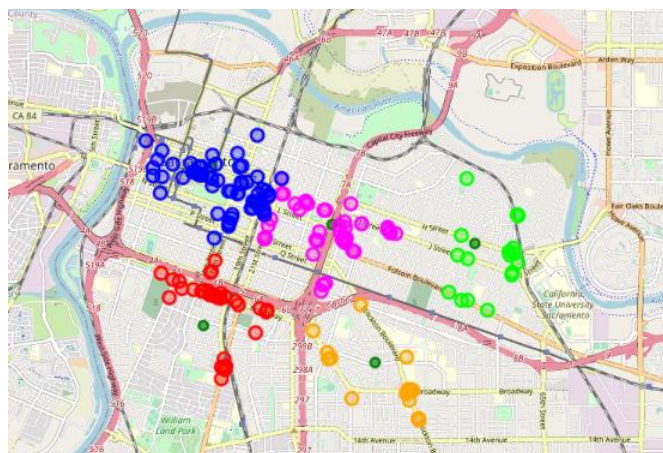

Figure 1

I then proceeded by using the data from Foursquare to make a vertical bar chart of the number of restaurants per zip code as seen in Figure 2.
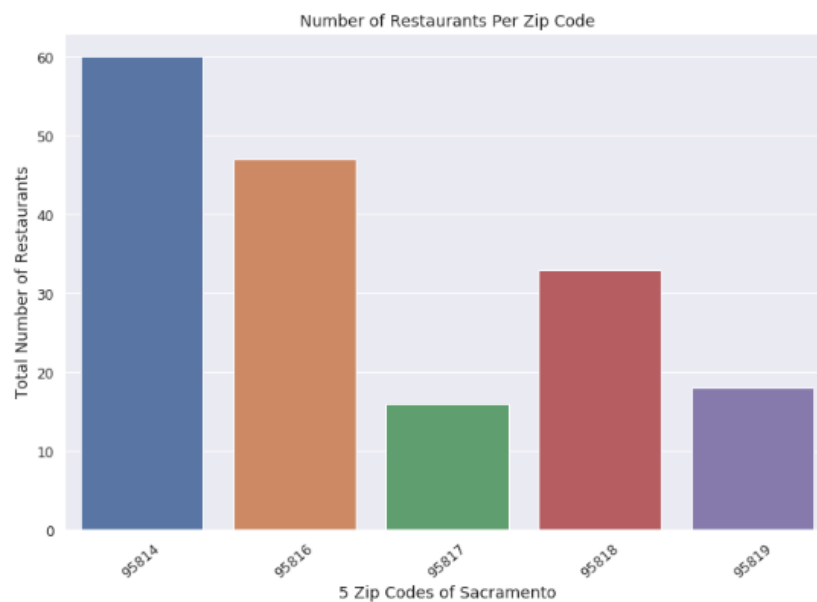

Figure 2

I then proceeded to get a count of the type of each of the venues for overall reference. This organized count data can be seen in Figure 3 below.

	Venue_Category	Frequency
0	American Restaurant	29
1	Mexican Restaurant	25
2	Chinese Restaurant	18
3	Café	16
4	Pizza Place	15
5	Restaurant	14
6	Thai Restaurant	13
7	BBQ Joint	9
8	Breakfast Spot	9
9	Japanese Restaurant	8

Figure 3

Following getting the count data, I proceeded to make a horizontal bar seen in Figure 4 chart to better visualize this count.
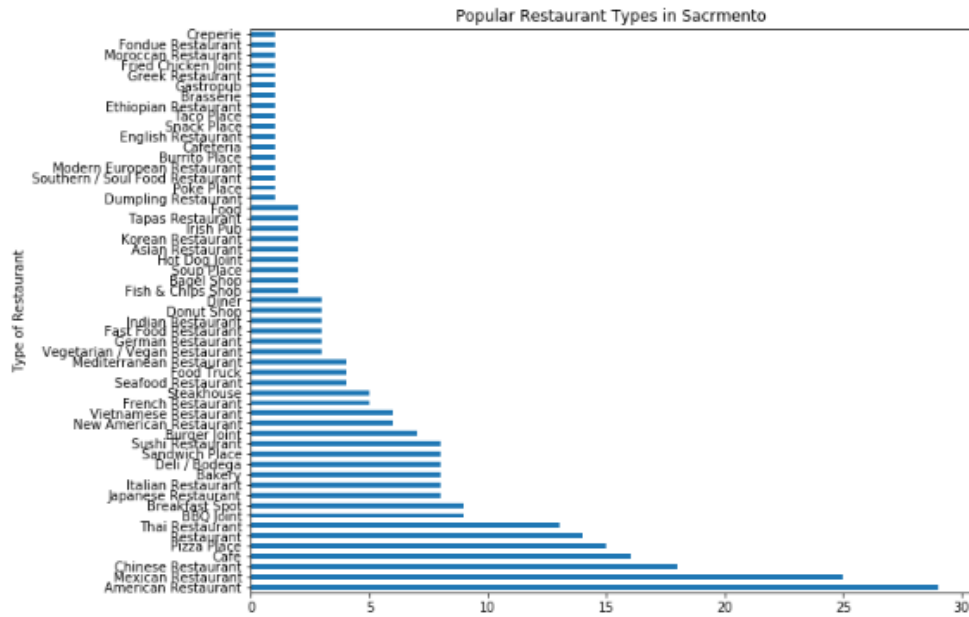
Figure 4

I then proceeded to narrow this count of all the restaurants in the downtown Sacramento area down to 10 by making a vertical bar chart of the top 10 most frequently occurring venues in the 5 zip codes of downtown, Sacramento as seen in Figure 5.
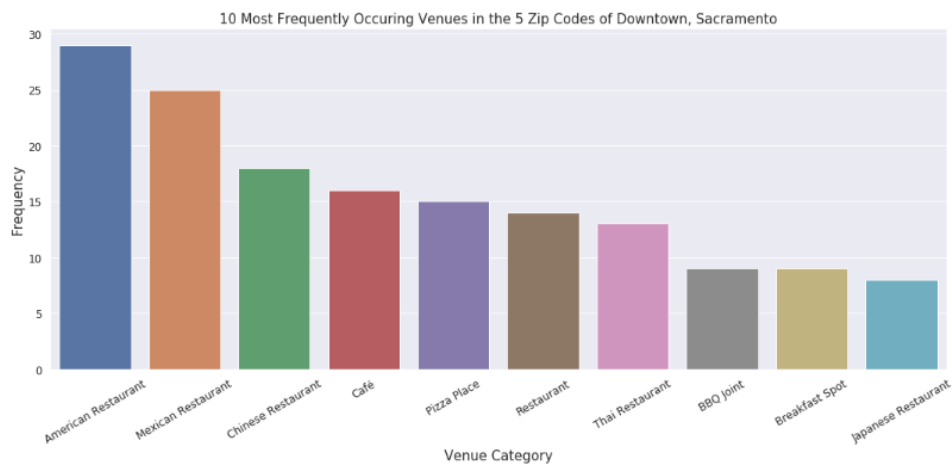

Figure 5

I then used hot encoding to get more information about the venue categories for each zip code and I used the hot encoding to produce the top 10 most common venues for each zip code as seen in Figure 6 and the Violin plots in Figure 7.

|	Zip_codes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	95814	American Restaurant	Mexican Restaurant	Café	Deli / Bodega	Italian Restaurant	Thai Restaurant	Pizza Place	BBQ Joint	Sushi Restaurant	Steakhouse
2	95816	Mexican Restaurant	American Restaurant	Breakfast Spot	Café	Sandwich Place	Pizza Place	Thai Restaurant	BBQ Joint	Burger Joint	Chinese Restaurant
3	95817	Chinese Restaurant	Restaurant	Steakhouse	Vegetarian / Vegan Restaurant	Food Truck	American Restaurant	Bakery	BBQ Joint	Cafeteria	Café
4	95818	Chinese Restaurant	Restaurant	Thai Restaurant	Pizza Place	Vietnamese Restaurant	Breakfast Spot	Asian Restaurant	Bakery	Japanese Restaurant	Mexican Restaurant
5	95819	American Restaurant	Pizza Place	Chinese Restaurant	Café	Restaurant	Mexican Restaurant	Taco Place	Fish & Chips Shop	Food Truck	French Restaurant

Figure 6
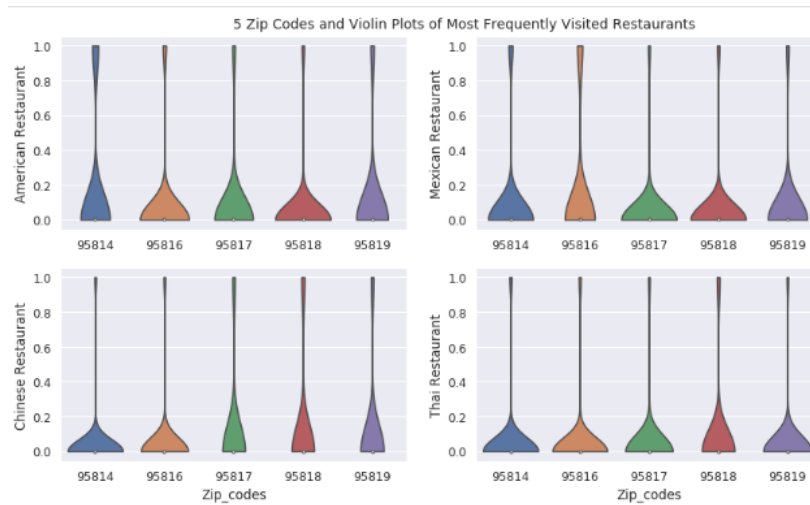
Figure 7

Finally, I clustered these 5 zip codes based on the restaurant types and used K-Means clustering. This would show what areas have similarities in restaurant category types. The K-Means clustering results are seen in Figure 8 below.
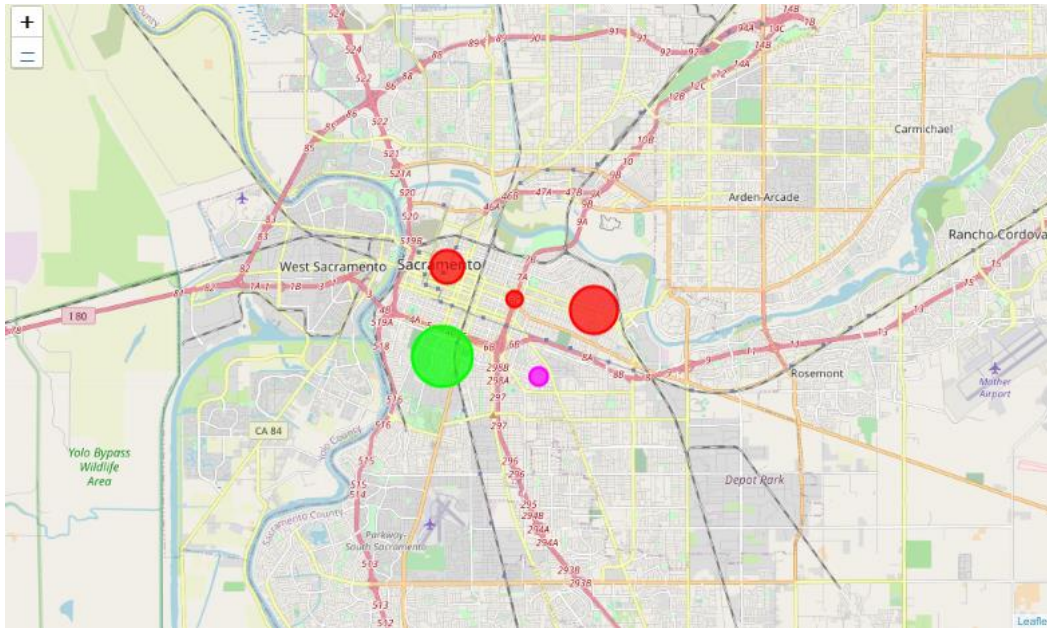
Figure 8

4. Results and Discussion

The three questions we wanted to answer:

- What areas have the most restaurants in them?
- What are the most common types of restaurants?
- What are proven good areas to have Mexican restaurants?

Tackling the first question, "What areas have the most restaurants in them?" we can refer to Figure 2 where we see a vertical bar chart of the number of restaurants per zip code. We see that 95814 has the most restaurants with 60, followed by, 95816 with 47, 95818 with 33, 95819 with 18 and 95817 with 16.

Tackling the second question, "What are the most common types of restaurants?" we can refer to Figure 4 and Figure 5 where we can see the horizontal plot of the count of all the restaurant types in Sacramento and the vertical bar plot of the count of the top 10 restaurant types in Sacramento respectively. The top 10 were American, Mexican, Chinese, Pizza, Restaurant*, Thai, BBQ Joint, Breakfast Spot, and Japanese, respectively. Restaurant* is a category that foursquare has for restaurants with multiple types of food at them.

Tackling the third question, "What are proven good areas to have Mexican restaurants?" we can refer to Figure 6 and Figure 7. In Figure 7, we can see violin plots for American, Mexican, Chinese, and Thai restaurants. The violin plot further shows which restaurants would be better suited for certain zip codes. From the violin plot, we see Mexican restaurants are more frequent in 95817 and 95818 than 95814 (largest number of restaurants) and more Mexican restaurants in 95814 than in 95816 and 95819. We can see in Figure 8 the three northern zip codes: 95819, 95816, and 95614 are more similar to each other than the southern 2 zip codes: 95818 and 95817. Therefore, it would be better to have a Mexican in one of the three northern zip codes, 95814, 95816 or 95819. In Figure 6, we can see the most popular categories/ restaurant types for each zip code. Figure 6 shows Mexican Restaurants being most popular in 95816 and 95814. Our client should have their new Mexican restaurant in either 95816 or 95814 based on the data presented above.

5. Conclusion

In conclusion, I have gotten a glimpse of a real-life data-science project. I learned how to use Foursquare API to get location data around certain areas. I can then format this data to learn more about the number of restaurants in each area and further get the frequency of restaurants in each zip code. This kind of analysis provides great initial guidance to take on more real-life challenges using data science.