

Where to Open a Cafe: New York or San Francisco

Data Science Capstone Project Report by Burce Bogmak Sit

May 2021

I. Introduction

Coffee is very popular beverage especially during day time. Especially, in the past year, the popularity of coffee-to-go has risen. Now, people prefers coffee-to-go not only on the way to work, but also on the daily walks or socializing. Due to the Covid-19 outbreak, even if cafes are open, people are more likely to drink their coffee or eat something in a open places, like parks.

The study on this report is about which city would be best to open a small cafe offering coffee and some snacks. The analysis is focused on two crowded cities on different coasts of US: New York, NY and San Francisco, CA. The aim is help the potential entrepreneurs to decide which city to open the cafe.

II. Data Acquisition

In this study, the two cities are compared in terms of parks and coffee shops. It is desired to high number and density of parks but low number of coffee shops, which are potential rivals. So, we need two information: the parks and the coffee shops at the center of the related cities. Foursquare can provide all these information with names and locations.

In order to, get the desired information “query=park” and “section=coffee” parameters are added to url’s built for parks and coffee shops. And the limit is set to 100. After cleaning the unnecessary columns, we have four data frames. The first 10 columns are shown below as a representation:

	Park	Park Latitude	Park Longitude	City
0	Monsignor McGolrick Park	40.724546	-73.943654	Brooklyn
1	McCarren Park	40.721640	-73.952579	Brooklyn
2	WNYC Transmitter Park	40.729958	-73.960733	Brooklyn
3	Hunter's Point South Park	40.742632	-73.960701	Queens
4	Cooper Park	40.715940	-73.937280	Brooklyn
5	Bushwick Inlet Park	40.722559	-73.961549	Brooklyn
6	Thomas P. Noonan, Jr. Playground	40.741053	-73.922213	Sunnyside
7	Msgr. McGolrick Park Dog Run	40.723291	-73.943459	Brooklyn
8	Four Freedoms Park	40.750744	-73.960465	New York
9	John F Murray Playground	40.747272	-73.948608	Queens

Table 1: Parks in New York

	Park	Park Latitude	Park Longitude	City
0	Alamo Square	37.775881	-122.434412	San Francisco
1	Painted Ladies	37.776120	-122.433389	San Francisco
2	Duboce Park	37.769578	-122.432960	San Francisco
3	Patricia's Green	37.776369	-122.424479	San Francisco
4	Alamo Square Dog Park	37.775878	-122.435740	San Francisco
5	Waller Park	37.771648	-122.426626	San Francisco
6	Buena Vista Park	37.768338	-122.440501	San Francisco
7	Dog Park	37.769503	-122.432709	San Francisco
8	Corona Heights Park	37.765023	-122.438831	San Francisco
9	Daniel E. Koshland Community Park	37.773218	-122.427146	San Francisco

Table 2: Parks in San Francisco

	Venue	Venue Latitude	Venue Longitude	City
0	Alkemy	40.725925	-73.939136	Brooklyn
1	Variety Coffee Roasters	40.723169	-73.944223	Brooklyn
2	Crema BK	40.723001	-73.945727	Brooklyn
3	Café Grumpy	40.728563	-73.948639	Brooklyn
4	Doughnut Plant	40.742989	-73.935292	Queens
5	Peter Pan Donut & Pastry Shop	40.726102	-73.952252	Brooklyn
6	Moe's Doughs	40.724532	-73.948517	Brooklyn
7	Charlotte Patisserie	40.723051	-73.950386	Brooklyn
8	Tar Pit	40.717615	-73.941401	Brooklyn
9	The Blue Stove	40.717512	-73.944908	Brooklyn

Table 3: Coffee Places in New York

	Venue	Venue Latitude	Venue Longitude	City
0	The Center SF	37.774545	-122.430730	San Francisco
1	Réveille Coffee Co.	37.770978	-122.432029	San Francisco
2	Lady Falcon Coffee Club	37.777039	-122.431946	San Francisco
3	Cafe International	37.772156	-122.430691	San Francisco
4	Sightglass Coffee	37.772287	-122.437419	San Francisco
5	The Mill	37.776425	-122.437970	San Francisco
6	Ritual Coffee Roasters	37.776476	-122.424281	San Francisco
7	Wise Sons Bagel & Coffee	37.777284	-122.424958	San Francisco
8	Blue Bottle Coffee	37.776430	-122.423224	San Francisco
9	Duboce Park Cafe	37.769334	-122.431561	San Francisco

Table 4: Coffee Places in San Francisco

III. Data Exploration and Analysis

Since the limit is set 100, we know that the obtained dataframes cannot have more than 100 rows. When the row count is calculated it is seen that park numbers are hit to 100, whereas the coffee shop numbers are 50 in NY and 32 in SF. At first glance, it is noticed that the number of coffee shops is higher in New York. But, in order to see the density of the locations of these parks and coffee shops, they are marked on the map. The following maps are in the same zoom level, 13.

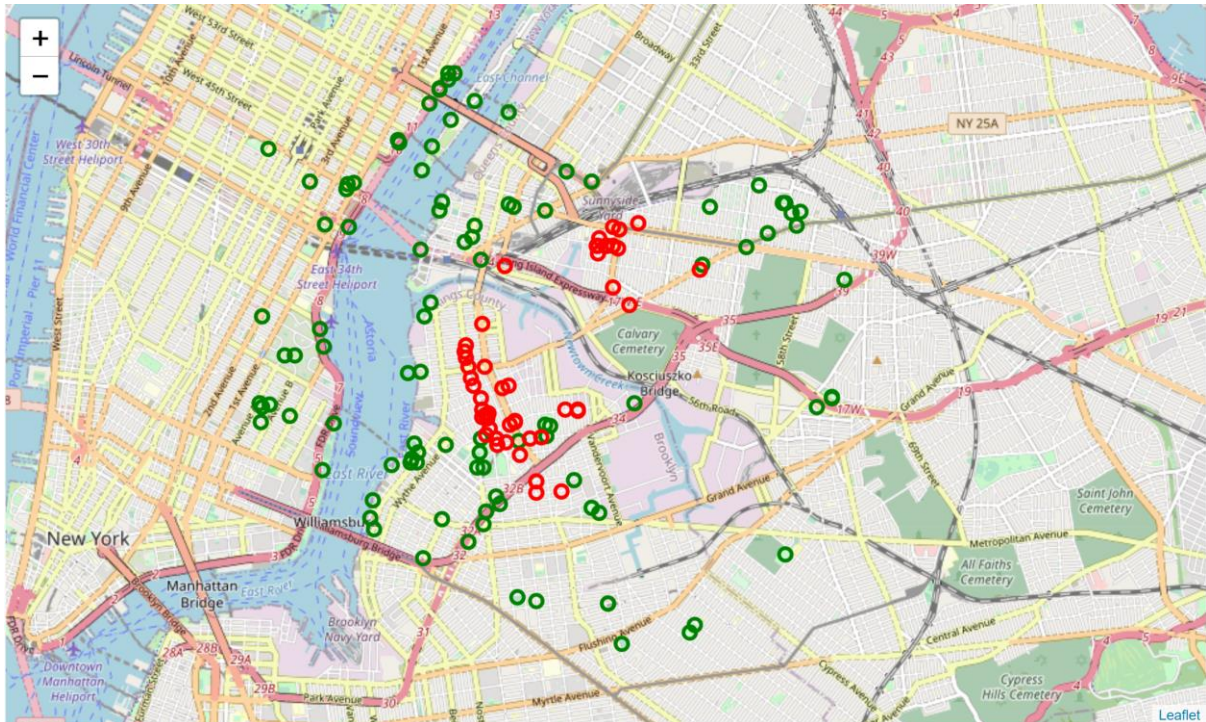


Figure 1: The map of New York with Coffee shops and parks marked red and green respectively

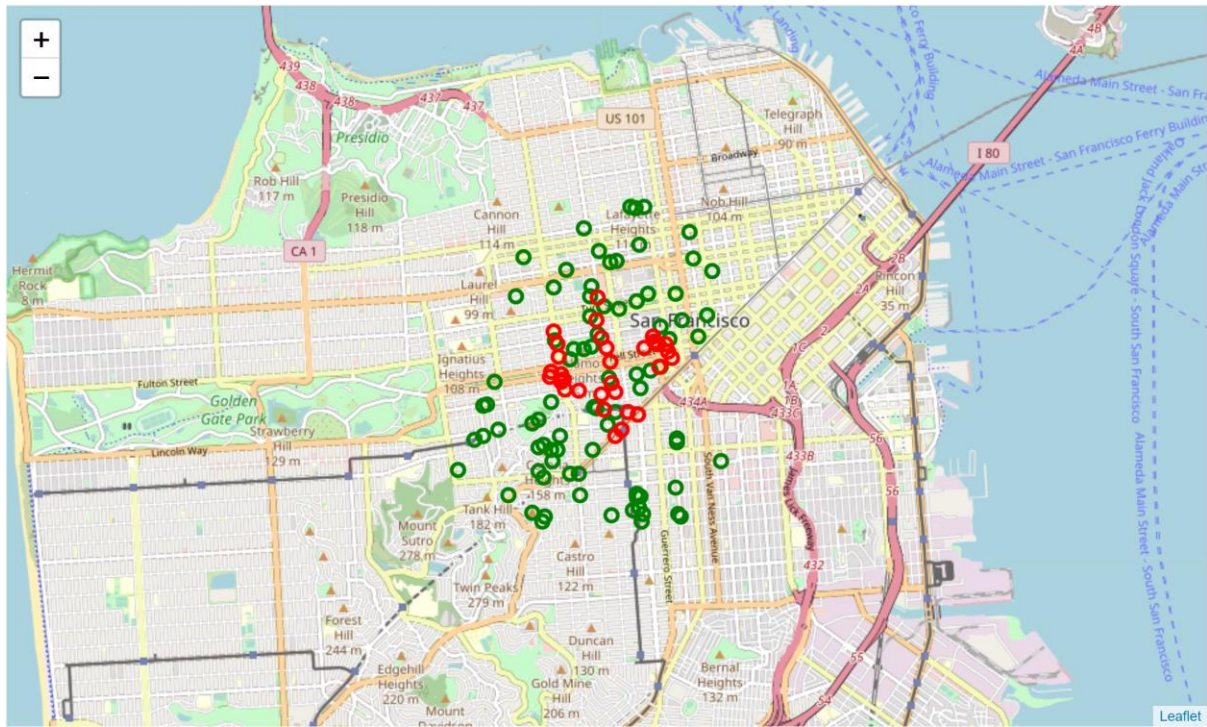


Figure 2: The map of San Francisco with Coffee shops and parks marked red and green respectively

IV. Outcomes and Conclusion

According to our criteria of the location and density of parks and coffee shops, it is seen that the venues both park and cafes are more densely located in San Francisco. In addition, considering the less number of coffee shops are present in the center of San Francisco, it can be concluded that opening the cafe in San Francisco would be recommended. Note that, this recommendation is made based on only the number and density of parks and coffee places.

As a future study, the other parameters could be taken into account, for more complex analysis. For example, the demographic data can be taken into consideration so that the target customer group is also studied.