Capstone Project

Introduction \ Problem

It's hard to track down some food for takeout during the times of COVID restrictions. Some restaurants have closed down completely, some have outdoor dining, some are exclusively takeout. Which neighborhoods are surviving the lockdown the best? Which neighborhoods are going to need new restaurants to open once restrictions are gone, and will probably have better deals on leases?

Target Audience

I've been interested in this information myself. I try to support the local places in my neighborhood so that the restaurants can survive this period of restrictions, and more employees can keep their jobs. There are many consumers in Seattle that feel and act in the same way in their own neighborhoods. Restaurateurs may look to capitalize on "restaurant deserts" once consumers can go back to full indoor dining.

Datasets

I will be using a number of datasets for this project. I will use venue data from foursquare using their API. I will use a downloaded csv dataset from the city of Seattle that contains data from restaurants self reporting their opening hours, and the types of service that they offer. I will use data scraped from wikipedia to get a list of Seattle neighborhoods so that I can do some comparison neighborhoods.

https://developer.foursquare.com/docs/api-reference/venues/search/#request

https://data-seattlecitygis.opendata.arcgis.com/datasets/restaurants-operating-during-covid19/data?showData=true

https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Seattle

Methodology

My methodology was to get the datasets, do some data cleaning, do some exploratory analysis including visualizations, and then analyze the neighborhoods for additional findings on their services provided during COVID.

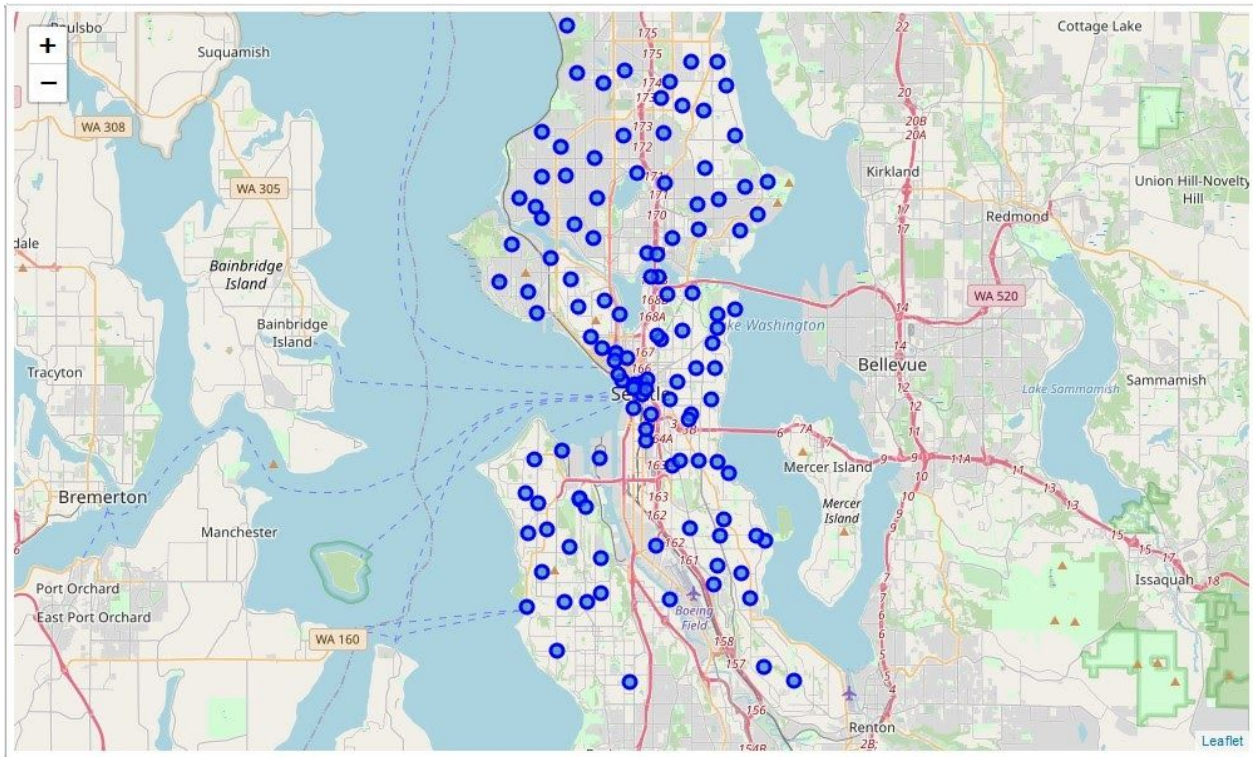
Results

Initial neighborhood data from wikipedia scrape.

```
req = requests.get("https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Seattle")

soup = BeautifulSoup(req.content,'lxml')

table = soup.find_all('table')[0]

df1 = pd.read_html(str(table))

hoods = pd.DataFrame(df1[0])
hoods.head()
```

	Unnamed: 0	Neighborhood name	Within larger district	Annexed[41]	Locator map	Street map	Image	Notes
0	1	North Seattle	Seattle	Various	NaN	NaN	NaN	North of the Lake Washington Ship Canal[42]
1	2	Broadview	North Seattle[42]	1954[43]	NaN	NaN	NaN	[44]
2	3	Bitter Lake	North Seattle[42]	1954[43]	NaN	NaN	NaN	[45]
3	4	North Beach / Blue Ridge	North Seattle[42]	1940,[43] 1954[43]	NaN	NaN	NaN	[46]
4	5	Crown Hill	North Seattle[42]	1907,[47] 1952,[43] 1954[43]	NaN	NaN	NaN	[48]

I cleaned the data, added coordinates, and produced a folium map of Seattle neighborhoods.



Using the foursquare API, I produced a venue dataframe.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bitter Lake	47.71868	-122.3503	Rain Cafe	47.724808	-122.343854	Café
3	Crown Hill	47.69520	-122.3741	Dick's Drive-In	47.696500	-122.371748	Burger Joint
4	Crown Hill	47.69520	-122.3741	Wild Mountain Cafe	47.690779	-122.374559	American Restaurant
5	Crown Hill	47.69520	-122.3741	Crown Hill Broiler	47.687647	-122.377145	Restaurant
7	Northgate	47.71310	-122.3193	Boud's Pinehurst Pub	47.715641	-122.312633	Bar

Next was the city of Seattle restaurant COVID data. I downloaded the entire dataset. The set includes all of King county so there was some more cleaning to do.

```
cos_df = pd.read_csv('Restaurants_Operating_during_COVID19.csv')
```

```
cos_df.head()
```

	X	Y	objectid	globalid	business_name	address_help_your_customer_find	phone_number_to_place_an_order	
0	-122.318799	47.613362	3	12aba10d-7825-45cb-ac38-46b22e670698	Bateau	1040 E Union St, Seattle, WA 98122	206-900-8699	https://www.r
1	-122.314425	47.614776	4	4b271d79-3283-4e16-bd59-cfadca726051	Nue	1519 14th Ave	206-257-0312	https:/
2	-122.328084	47.614436	5	f4928649-888c-41be-9726-4eb86ae05a08	Terra Plata	1501 Melrose Ave Seattle, WA 98122	(206) 325-1501	http
3	-122.326795	47.598599	6	0745b48d-87a3-4c4e-9910-24340784a738	A+ Hong Kong Kitchen	419 6th Ave S, Seattle, WA 98104	206-682-1267	http
4	-122.321425	47.581182	7	0afb2fdb-9ef1-430a-8757-beafc610fd24	Gourmondo	2415 Airport Way S. Seattle WA 98134	206-587-0190	https://cateri

```
cos_df[cos_df['business_name'].str.contains("Random")]
```

	X	Y	objectid	globalid	business_name	address_help_your_customer_find	phone_number_to_place_an_order	business_webs
721	-122.349791	47.615518	818	74138a75-dff3-428a-8c71-b44f2a6fd6f3	Some Random Bar	2604 1st Ave	2067286621	http://somerandombar.c

Discussion

The city of Seattle set is compiled from restaurants self reporting, or answering the city's surveys. It's not going to be complete for obvious reasons, but the exercise, for me, was about merging the multiple datasets. Using foursquare data was a requirement for the project.

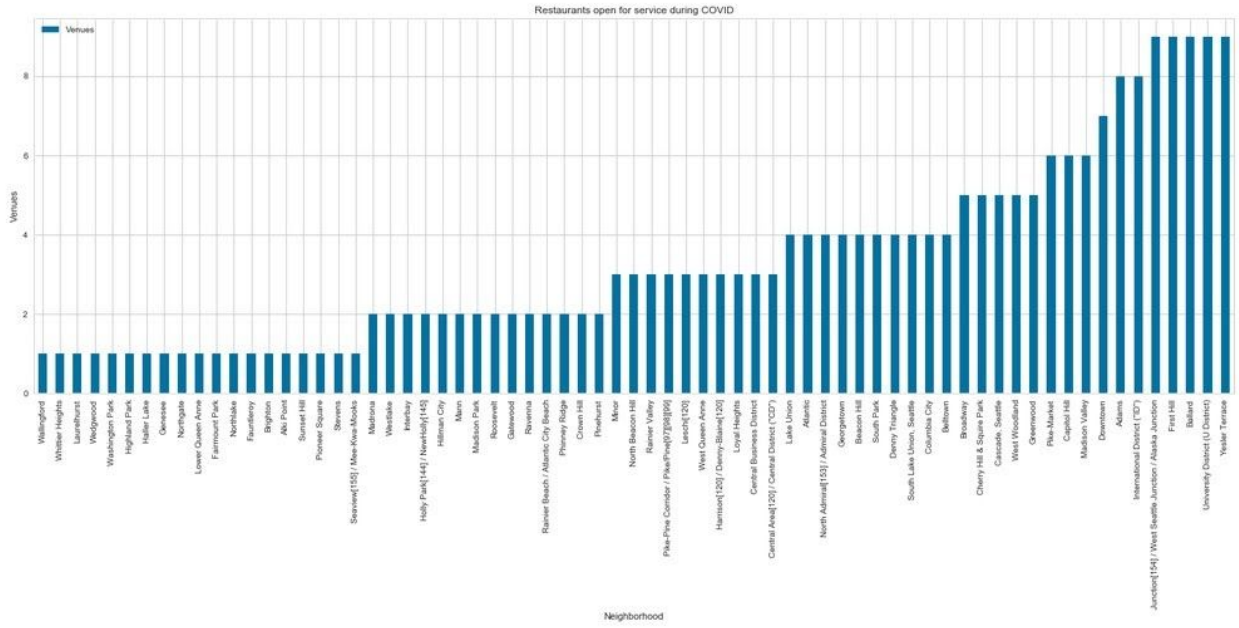
Here is the merged data for a larger dataset that includes more than burgers. It will be reduced to the info present in the city of Seattle dataset with an inner join.

```
print(merged_df.shape)
merged_df.head()
```

```
(187, 13)
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	business_name	phone	business_website
0	Crown Hill	47.69520	-122.37410	Wild Mountain Cafe	47.690779	-122.374559	American Restaurant	Wild Mountain Cafe	(206) 297-9453	https://www.thewildmountaincafe.com/
1	Crown Hill	47.69520	-122.37410	Papa John's Pizza	47.694423	-122.375077	Pizza Place	Papa John's Pizza	(206) 985-0000	https://locations.papajohns.com/united-states/...
2	Seaview[155] / Mee-Kwa-Mooks	47.55217	-122.39175	Papa John's Pizza	47.554182	-122.386628	Pizza Place	Papa John's Pizza	(206) 985-0000	https://locations.papajohns.com/united-states/...
3	Fairmount Park	47.55341	-122.38160	Papa John's Pizza	47.554182	-122.386628	Pizza Place	Papa John's Pizza	(206) 985-0000	https://locations.papajohns.com/united-states/...
4	Greenwood	47.69082	-122.35529	Flying Bike Cooperative Brewery	47.692123	-122.355175	Brewery	Flying Bike Cooperative Brewery	(206) 428-7709	https://flyingbike.coop/

With the merged dataset, I summarized the reported open venues by neighborhood.



Conclusions

Using the self-reported data from the city, you can see that a number of neighborhoods have very few, or no restaurant service at all. The chart from the merged data set doesn't show all of the neighborhoods that were not found in the city data. An outer join on that merge would retain the original rows for the neighborhood dataframe.