Impact of Neighborhood Composition on COVID-19 Rates in Los Angeles

Matthew Morris

Impact of COVID-19 On Businesses

- COVID-19 has forced changes in the way most businesses operate
- Some areas were hit much harder than others
- How can I learn from lessons around COVID-19 as:
  - A small business owner with locations throughout LA, and
  - As a municipality or local government
- What factors influence COVID case rates?

Data Acquisition, Processing, and Cleaning

- Data on COVID case rates by neighborhood was obtained from data publicly available on the LA county website. For the purposes of this report, we will only focus on LA counties. http://publichealth.lacounty.gov/media/coronavirus/locations.htm
- Geographical data about each neighborhood were obtained from USC's neighborhood data for social change, located here, which can be exported as a CSV: https://usc.data.socrata.com/dataset/Los-Angeles-Neighborhood-Map/r8qd-yxsr
- Neighborhood composition data (ie, the kinds of venues in each neighborhood and how frequently they occur), was obtained using the FourSquare API to pull venues per neighborhood (using latitude and longitude data from USC).
- Data from LA county and USC was merged, and data was filtered down to neighborhoods starting with “Los Angeles -”, and the “Los Angeles” prefix was dropped.
- Incorrect values for latitude and longitude columns were moved to the proper column so that the data appeared properly on the map.

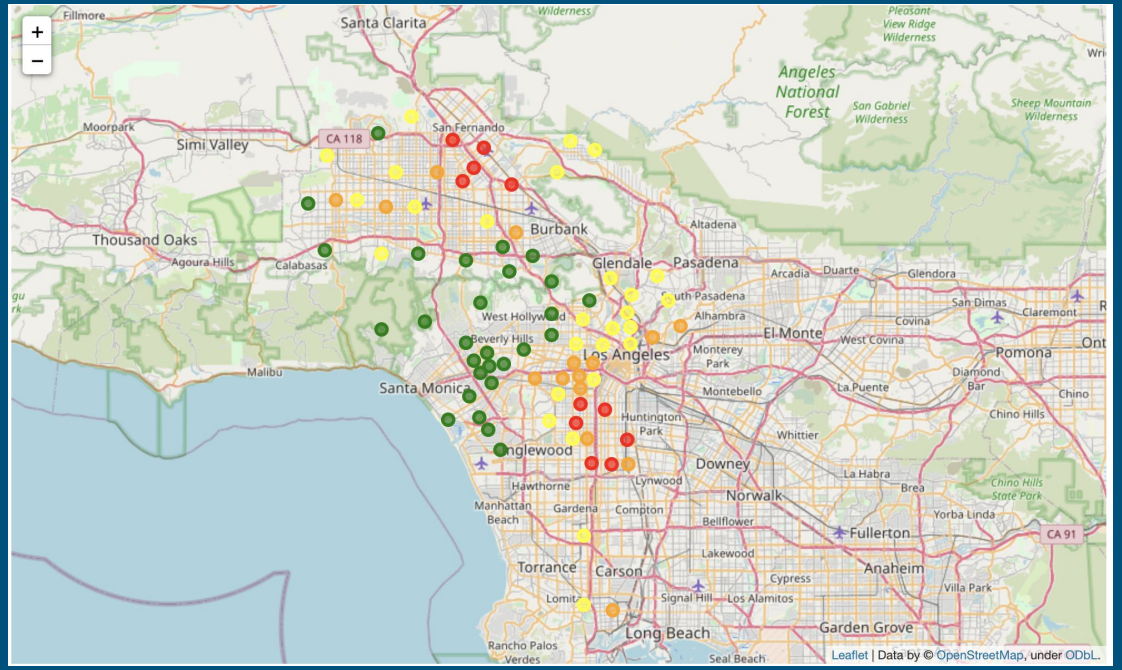
Mapping LA: A Visualization of COVID-19 Impact

Key:

- **Green**: 0-8062 Cases/100k residents
- **Yellow**: 8063-12878 Cases/100k residents
- **Orange**: 12878-17263 Cases/100k residents
- **Red**: 17263-22119 Cases/100k residents

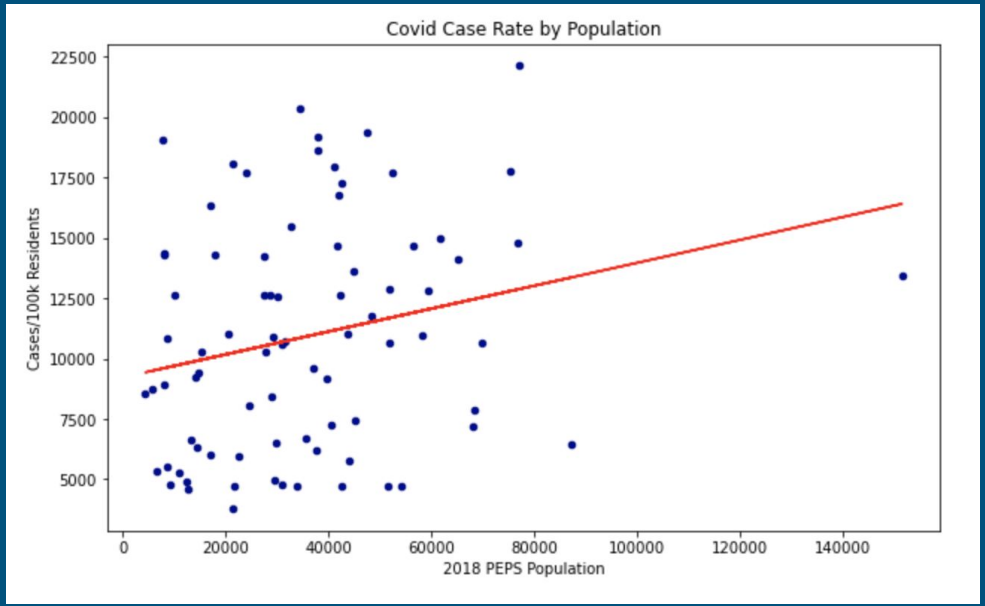
Observations:

- There are visible clusterings of neighborhoods near each other with high and low COVID-19 rates
- Do these neighborhoods share something in common?
- What made neighborhoods in the red or orange colors have higher case rates?



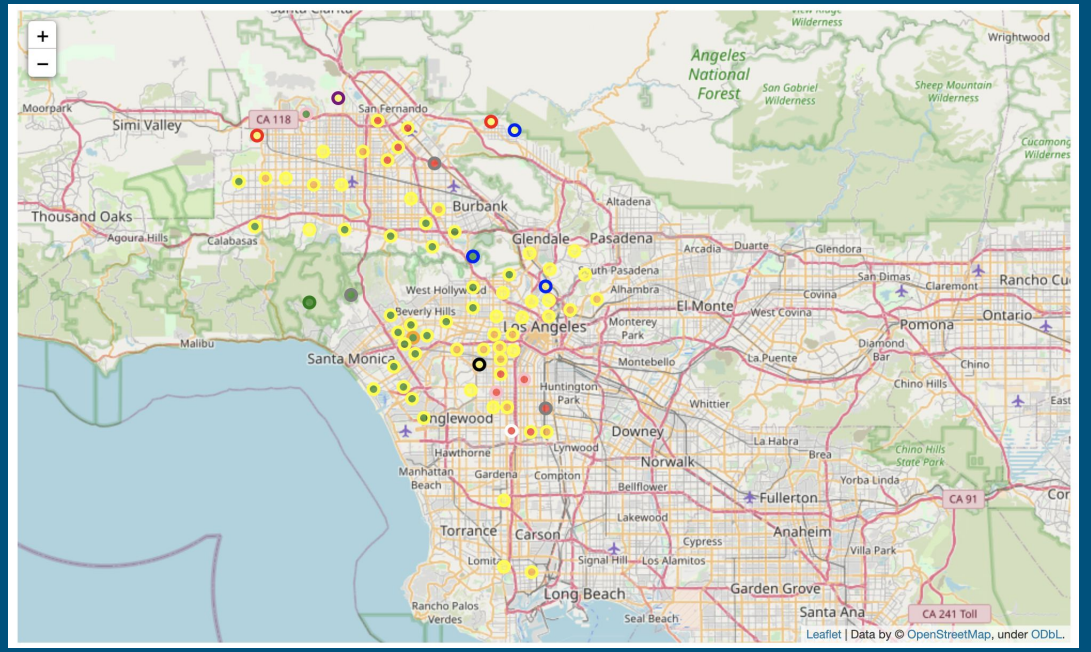
COVID-19 By Population Density

- An initial hypothesis was that COVID-19 rates were affected mostly by population density
- Do dense urban areas have higher COVID rates?
- This chart shows that there is only a weak correlation between population in a given neighborhood and cases. The R^2 is a very low 0.056.
- We will discount population as the primary factor affecting COVID rates for the time being, since the R^2 is so low.



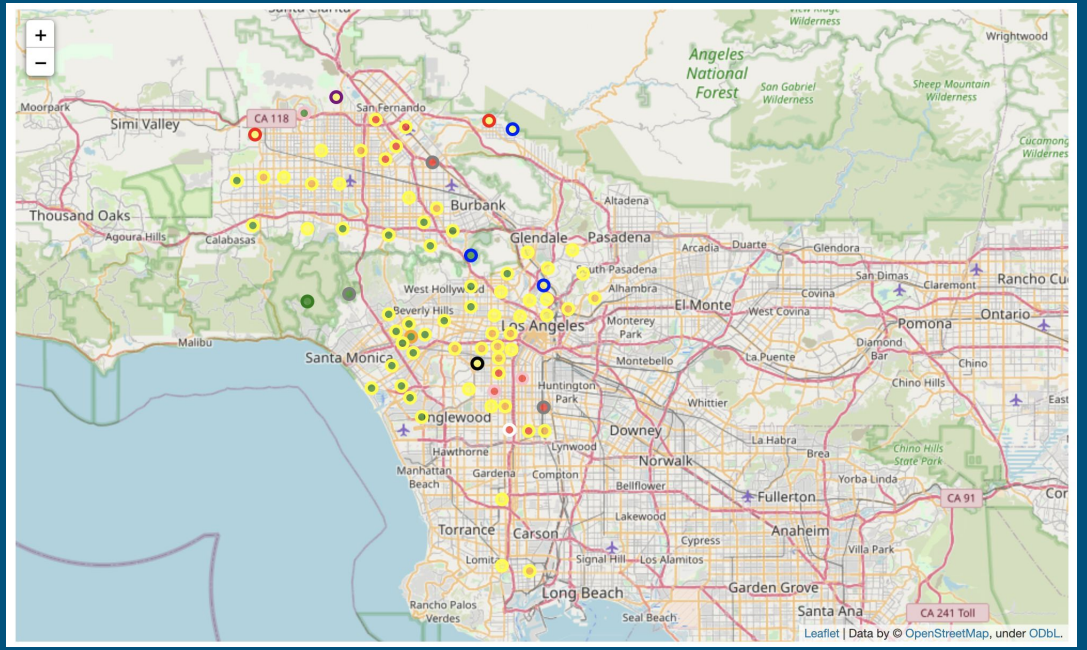
K-Means Clusters Round 1: Outlier Detection

- This is a map of neighborhoods clustered by the prevalence of certain venues.
- Outer circle colors on the map below represent different clusters
- Fill color still represents COVID case rates
- Only one primary (yellow border) cluster consisting of most urban areas
- Most clusters not in the set of points outlined in yellow consist of one or 2 neighborhoods only, mostly on the outskirts of LA proper
- These appear to be suburban or exurban communities of significantly different character than the greater LA area
- **Hypothesis**: we've performed outlier detection and found mostly suburbs.



K-Means Clusters Round 1: Outlier Detection (Cont'd)

- The points in South Park and Harvard Park (pink), Florence-Firestone and Sun Valley (grey), Leimert Park (black), and Vermont Vista (white) bear further discussion - they are urban and have high COVID case rates
- Otherwise, the hypothesis is correct - we've essentially detected 2 things with this clustering:
    - 1) The boundaries of LA where the character of neighborhoods change dramatically; and
    - 2) The affluent, diffuse neighborhood of Cheviot Hills
- These are not the primary interest of this report, since they will have few restaurants and other venues.
- I chose to perform another K-means analysis on more urban areas outlined in yellow



K-Means Clusters Round 1: Discussion of Results

- The urban areas with high COVID case rates are groups 8, 4, and 9 (white, pink, and gray).

	k-means-labels	Cases/100k Residents	label_color
0	8	17958.000000	white
1	4	14841.000000	pink
2	9	13964.333333	gray
3	5	10930.000000	purple
4	2	10909.682540	yellow
5	6	10305.000000	black
6	3	10283.500000	red
7	0	9292.666667	blue
8	7	4798.000000	orange
9	1	3790.000000	green

K-Means Clusters Round 1: Discussion of Results

- These neighborhoods have a high concentration of outdoor and physical activities, as well as generic food venues
- Parks, Scenic lookouts, yoga studios, and food courts seem to dominate
- These could have high COVID rates due to demographics or due to people flocking to locations where activities are allowed
- Other groups are essentially one big cluster we will dig into separately

	k-means-labels	Neighborhoods	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
4	9	Brentwood	Food Truck	Scenic Lookout	Ethiopian Restaurant	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop	Fire Station	Financial or Legal Service
20	9	Florence-Firestone	Food	Other Repair Shop	Music Venue	Grocery Store	Yoga Studio	Food Court	Food & Drink Shop	Flower Shop	Fire Station	Financial or Legal Service
57	9	Sun Valley	Food	Food Truck	Electronics Store	Taco Place	Fire Station	Furniture / Home Store	Seafood Restaurant	Donut Shop	Convenience Store	Film Studio

```
la_venue_types_by_neighborhood_sorted[la_venue_types_by_neighborhood_sorted["k-means-labels"]==8]
```

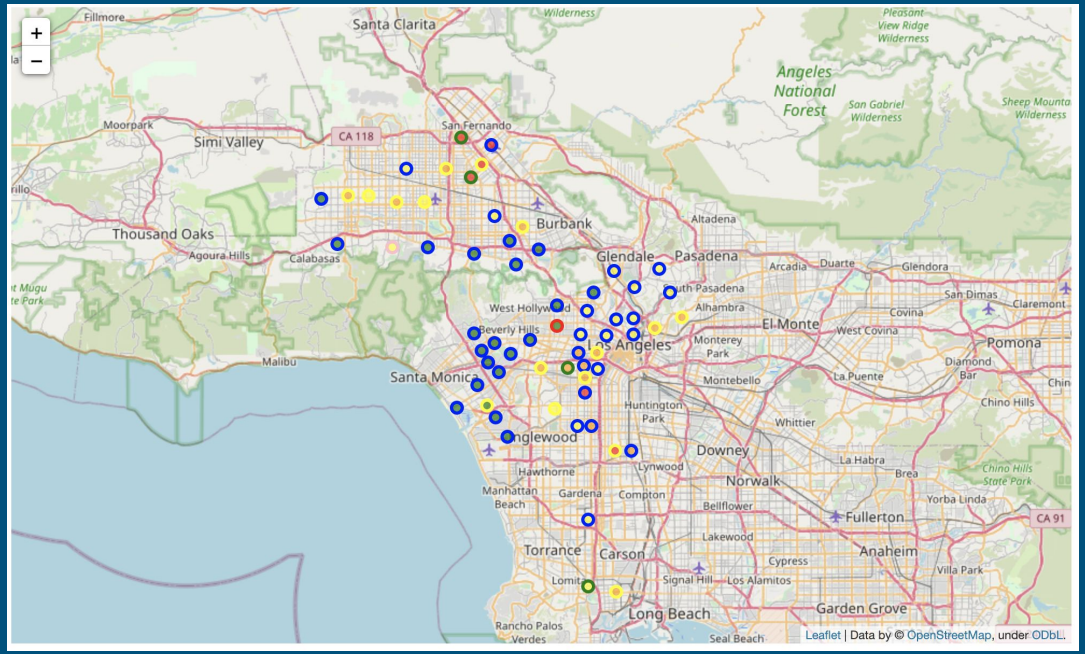
	k-means-labels	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
68	8	Vermont Vista	Burger Joint	Yoga Studio	Farmers Market	Food Stand	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop	Fire Station

```
la_venue_types_by_neighborhood_sorted[la_venue_types_by_neighborhood_sorted["k-means-labels"]==4]
```

	k-means-labels	Neighborhoods	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
28	4	Harvard Park	Shipping Store	Park	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fire Station	Financial or Legal Service
51	4	Porter Ranch	Park	Yoga Studio	Ethiopian Restaurant	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop	Fire Station	Financial or Legal Service
55	4	South Park	Park	Yoga Studio	Ethiopian Restaurant	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop	Fire Station	Financial or Legal Service

K-Means Clusters Round 2: Urban Areas of LA

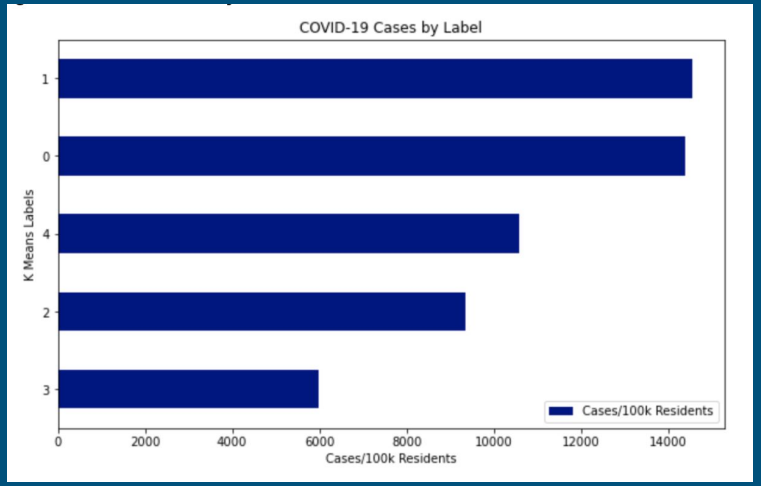
- Performed a second k-means, only on the urban areas, removing outliers
- Once again clustered by venue type; each border color represents a cluster
- Attempted to determine some larger clusters by neighborhood character
- This gave us a few neighborhood clusters that seemed to follow the boundaries of COVID impact a little more closely



Primary Clusters with High Incidence of COVID-19

- In order to get a more accurate, non-visual grouping of the data by cluster, I tabulated and sorted as well as graphed the data by cluster
- It is immediately obvious that clusters 0 and 1 (yellow and green respectively) have far higher COVID rates per 100k residents.
- Let's take a look at neighborhood composition to see if we can determine any trends.

	yellow-k-means-labels	Cases/100k Residents	label_color
0	1	14542.250000	green
1	0	14367.000000	yellow
2	4	10568.000000	pink
3	2	9334.512195	blue
4	3	5986.000000	red



K-Means Round 2: Primary Clusters with High Incidence of COVID-19

- Cluster 1 (green) has the highest COVID rate.
- We notice 2 things fairly quickly:
  - The first is outdoor locations (park, plaza, skating rink, curling ice).
  - The second is ethnic restaurants, predominantly Mexican.
- This will prove to be an interesting hybrid between the neighborhoods from the first round of clustering and Cluster 0 from this round, discussed in the next slide.

	yellow-k-means-labels	Neighborhoods	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
19	1	Harbor City	Mexican Restaurant	Bakery	Park	Spanish Restaurant	Yoga Studio	Farmers Market	Food Truck	Food Stand	Food Service	Food Court
25	1	Jefferson Park	Park	Neighborhood	Mexican Restaurant	Home Service	Taco Place	Fried Chicken Joint	Convenience Store	Yoga Studio	Fast Food Restaurant	Filipino Restaurant
32	1	Mission Hills	Plaza	Church	Park	Ethiopian Restaurant	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop	Financial or Legal Service
38	1	Panorama City	Mexican Restaurant	Skating Rink	Curling Ice	Automotive Shop	Park	Yoga Studio	Falafel Restaurant	Food Service	Food Court	Food & Drink Shop

K-Means Round 2: Primary Clusters with High Incidence of COVID-19

- Cluster 0 has the second highest COVID rate from this round of k-means.
- The dominating feature of this cluster is a high prevalence of ethnic restaurants (and in particular Mexican and Latin restaurants).

	yellow-k-means-labels	Neighborhoods	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	0	Arleta	Video Store	Historic Site	Bakery	Convenience Store	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Yoga Studio	Ethiopian Restaurant
4	0	Canoga Park	Mexican Restaurant	Ice Cream Shop	Sports Bar	Restaurant	Furniture / Home Store	Sushi Restaurant	Liquor Store	Farmers Market	Food Service	Food Court
8	0	Del Rey	Mexican Restaurant	Hobby Shop	Bakery	Pizza Place	Donut Shop	Convenience Store	Coffee Shop	Sandwich Place	Café	Smoke Shop
12	0	El Sereno	Mexican Restaurant	ATM	Trail	Convenience Store	Restaurant	Neighborhood	Seafood Restaurant	South American Restaurant	Liquor Store	Thrift / Vintage Store
15	0	Exposition Park	Coffee Shop	Design Studio	Intersection	Mexican Restaurant	Food	Yoga Studio	Food Stand	Food Service	Food Court	Food & Drink Shop
17	0	Green Meadows	Food Stand	Pizza Place	Food	Donut Shop	Sandwich Place	Ethiopian Restaurant	Food Service	Food Court	Food & Drink Shop	Flower Shop
24	0	Hyde Park	Caribbean Restaurant	Convenience Store	Bookstore	Grocery Store	Yoga Studio	Falafel Restaurant	Food Stand	Food Service	Food Court	Food & Drink Shop
27	0	Lake Balboa	Sandwich Place	Steakhouse	Fast Food Restaurant	Convenience Store	Donut Shop	Automotive Shop	Seafood Restaurant	Mexican Restaurant	Flower Shop	Ethiopian Restaurant
28	0	Lincoln Heights	Mexican Restaurant	Fast Food Restaurant	Convenience Store	Burger Joint	Fried Chicken Joint	Music Venue	Gas Station	Sandwich Place	Food Truck	Video Store
33	0	North Hills	Pizza Place	River	Fast Food Restaurant	Baseball Field	Yoga Studio	Electronics Store	Food Service	Food Court	Food & Drink Shop	Food
34	0	North Hollywood	Latin American Restaurant	Shoe Store	Sandwich Place	Electronics Store	Mobile Phone Shop	Thrift / Vintage Store	Fast Food Restaurant	Donut Shop	Pizza Place	Insurance Office
39	0	Pico-Union	Latin American Restaurant	Mexican Restaurant	South American Restaurant	Cuban Restaurant	Clothing Store	Storage Facility	Convenience Store	Park	Grocery Store	Food
42	0	Reseda	Vietnamese Restaurant	Fast Food Restaurant	Mexican Restaurant	Furniture / Home Store	Chinese Restaurant	Supermarket	Convenience Store	Pawn Shop	Thai Restaurant	Falafel Restaurant
54	0	West Adams	Mexican Restaurant	Gym / Fitness Center	Fried Chicken Joint	Fast Food Restaurant	Latin American Restaurant	Bar	Performing Arts Venue	Café	Wine Bar	Deli / Bodega
60	0	Wilmington	Pizza Place	Fast Food Restaurant	Latin American Restaurant	Convenience Store	Discount Store	Mexican Restaurant	Museum	Park	Sandwich Place	Falafel Restaurant

Conclusion

- COVID case rates were highest in neighborhoods with a high number of parks and outdoor activities, which may have stayed open during COVID.
- Case rates were next highest in areas with high numbers of ethnic restaurants. This likely would need further examination to determine the root cause.
- Municipalities may want to examine demographics in these areas to determine why they were so unequally affected by COVID.