

# **Covid in Metro Nashville**

# How do we get a handle on the spread of COVID-19 in Metro Nashville?

- More consistent mask use?
- Responding to hubNashville reported violations?
- Something else?

# Available Data

hubNashville 311 Service Requests - Violations of Covid related rules and regulations are contained in this dataset where Request Type is **COVID-19** and Subrequest Type is **COVID-19 Violations**

**clusters.csv** - scraped from the [wsmv website](#) and geocoded using the Google API. The dataset provides detail for COVID-19 clusters (groupings of 10 or more cases identified by the Metro Public Health Department through its case investigation process).

**Cluster Name** - the name of the place where a cluster of cases has been traced to

**Facility Type** - the type of place the cluster location is

**Cluster Start Date** - date of the first known case traced to the cluster

**# Cases** - the count of cases traced to the cluster

**Latitude** - latitude of the cluster

**Longitude** - longitude off the cluster

**businesses.csv** is another geocoded dataset created for this project using the Google API. This dataset includes:

**address** - the address from hubNashville of a reported violation

**vicinity** - the address for a business found near the address (the closest five businesses are included)

**name** - the name of the business located at the vicinity.

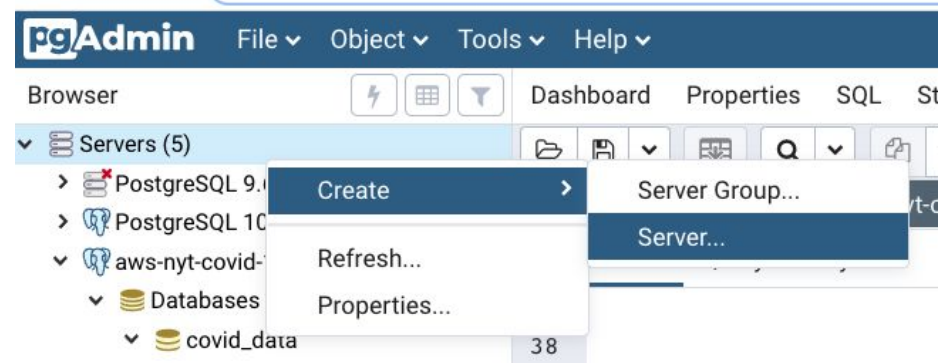
**type** - Google's classification of that business

data from the **nytimes/covid-19-data** repository is loaded to a PostgreSQL database on AWS. Tables in this database include:

**counties** table - the running total of cases and deaths for US counties reported daily since 3/8/2020

**mask\_use** table - results from online interviews conducted by Dynata for the NYT between July 2 and July 14, 2020. Participants were asked *How often do you wear a mask in public when you expect to be within six feet of another person?*

**To Access the NYT data on AWS:**  
Create a connection to the AWS data in PgAdmin by creating a new server.

A screenshot of the 'Create - Server' dialog box in PgAdmin, showing the 'General' tab. The 'Name' field contains 'aws-nyt-covid-1'. The 'Server group' dropdown is set to 'Servers'. There are checkboxes for 'Background' and 'Foreground' (both unchecked). The 'Connect now?' checkbox is checked. There is a 'Comments' text area at the bottom.A screenshot of the 'Create - Server' dialog box in PgAdmin, showing the 'Connection' tab. The 'Host name/address' field contains 'nyt-covid-1.csdmctdmvs7t.us-east-1.rds.amazonaws.c'. The 'Port' field contains '5432'. The 'Maintenance database' field contains 'postgres'. The 'Username' field contains 'nss\_user'. The 'Password' field is masked with dots. The 'Save password?' checkbox is checked. There are empty fields for 'Role' and 'Service'.



**Server name:** aws-nyt-covid-1 (or similar)

**Host name:** nyt-covid-1.csdmctdmvs7t.us-east-1.rds.amazonaws.com

**port:** 5432

**user:** nss\_user

**password:** nssAWS2020

1. Write and test your queries in PostgreSQL 
2. Use sqlalchemy to create an engine (connection to your database) 
3. Use pandas read\_sql() method to load the results of your queries to pandas dataframes





## Some Questions to Explore

How have the COVID-19 violations reported through hubNashville changed over time?

Do violations get reported more in certain areas? (**hubNashville data**)

Is there a relationship between the number of cases reported in Davidson County over time and the number of violations reported to hubNashville? (**NYT data**, **hubNashville data**)

How does mask use in Davidson County compare to other counties in TN? How does it compare to other counties across the US? (**NYT data**)

Do violations reported to hubNashville presage clusters of COVID-19? (**hubNashville data**, **clusters.csv**, **businesses.csv**)