Analyzing the COVID-19 impact for New York City restorans and unemployment rate

Applied Data Science Capstone Project, September 2020, Anna Zhukova

Introduction

New York City is the biggest city of the U.S. with an estimated population of 8,336,817 in 2019 distributed over about 302.6 square miles (784 km^2), New York City is also the most densely populated major city in the United States. NYC is attracting thousands of people for tourism and living every year. In 2019 it brought in about $45 billion in annual spending and supported about 300,000 jobs.

NYC is a bright vibrant city, where you can try cuisines from all over the world. Number of restaurants in 2018 in this city was 27043, but the total number of eating and drinking places was more than 50k. In 2019 Food Services provided 865,800 jobs and covered 9% of the employment in NYC. One hundred and twenty six of NYC's restaurants have been awarded Michelin stars. There are also nine NYC restaurants that made it onto the World's 100 Best Restaurants list.

Officials had been expecting more than 67 million visitors in 2020, about one-fifth of them from outside the country. But the COVID-19 pandemic happened in March 2020. All World had already suffered from the pandemic. NYC was an epicenter in the U.S. in spring 2020 and had a strong damage from COVID-19. Let's try to find details and to do analysis of its impact on NYC.

Data Collection

The data required for this project is as follows has been collected from multiple sources. The following data is required for the project:
1) Neighborhood data of NYC (borough, neighborhood, geographical coordinates)
2) COVID-19 data in NYC
3) Unemployment rate Data
4) Venue data for neighborhoods in NYC Neighborhoods

Neighborhoods Data

The data of the neighborhoods in NYC was taken from an open json file
https://cocl.us/new_york_dataset . The data is added into a pandas dataframe. The geocoder library in python has been used to obtain latitude and longitude data for various neighborhoods in NYC. This data is relevant for plotting the map of NYC using the Folium library in python.

COVID-19 Data

It is open data on site with John Hopkins reported cases for all US:
https://www.kaggle.com/headsortails/covid19-us-county-jhu-data-demographics I will take only data for NYC.

Unemployment rate Data

Statistics per month is available on site:
https://www.bls.gov/regions/new-york-new-jersey/data/xg-tables/ro2xglausnyc.htm

Venue Data

The venue data has been extracted using the Foursquare API. This data contains venue lists, recommendations for all neighborhoods in NYC and is used to check the popular venues of different neighborhoods.