

COVID

An introduction to coronavirus in New York City

Intro: US on April 9, 2020

Intro: New York State and New York City on April 9, 2020

Descriptive Stats: NYC borough case counts per day

Descriptive Stats: different rates of infection in NYC

Descriptive Stats: 4 factors

Descriptive Stats: household...

A Study of Covid in New York City



Intro Descriptive Statistics Inferential Statistics

Since the start of the year, Corvid-19 has spread across the country, resulting in many cases of infections, hospitalizations, and deaths. New York (especially New York City) emerged as an early hot spot, and there are trends among those who are infected.

A Capstone Project for Springboard by Emily Rice

Sources:

github.com/nychealth/coronavirus-data
github.com/BuzzFeedNews/2020-05-covid-city-zip-codes
github.com/nytimes/covid-19-data
2015-2019 American Community Survey 5-Year Estimates

COVID

An introduction to coronavirus in New York City

Intro: US on April 9, 2020

Intro: New York State and New York City on April 9, 2020

Descriptive Stats: NYC borough case counts per day

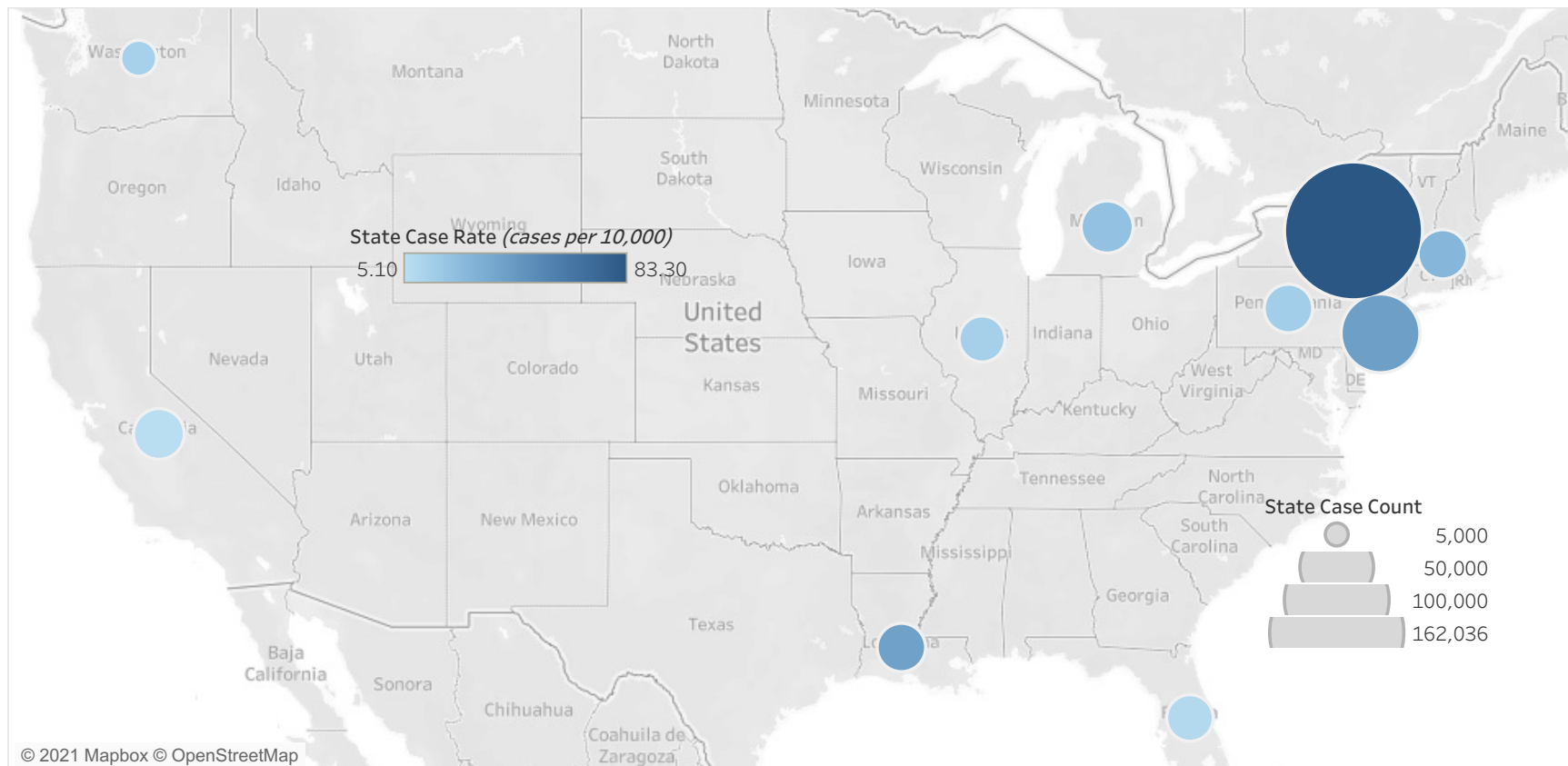
Descriptive Stats: different rates of infection in NYC

Descriptive Stats: 4 factors

Descriptive Stats: household...

COVID-19 in the United States on April 9, 2020

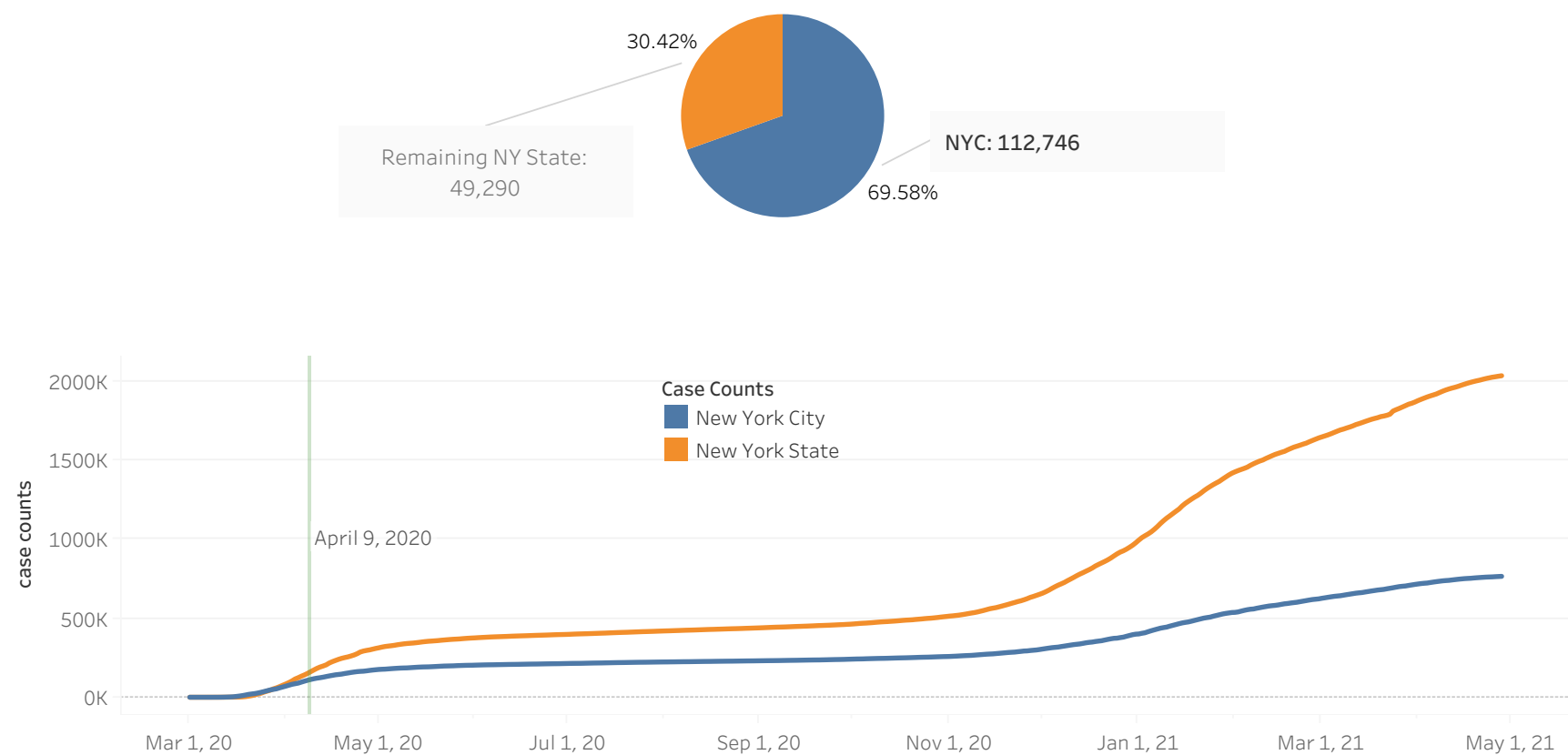
At the start of the pandemic, most of the cases were focused in a handful of hotspots, New York being the main one



COVID

An introduction to coronavirus in New York City	Intro: US on April 9, 2020	Intro: New York State and New York City on April 9, 2020	Descriptive Stats: NYC borough case counts per day	Descriptive Stats: different rates of infection in NYC	Descriptive Stats: 4 factors	Descriptive Stats: household...
---	----------------------------	--	--	--	------------------------------	---------------------------------

At the start of the pandemic in New York, most of the cases were located in NYC



COVID

An introduction
to coronavirus
in New York City

Intro: US on April 9,
2020

Intro: New York State
and New York City on
April 9, 2020

Descriptive Stats: NYC
borough case counts
per day

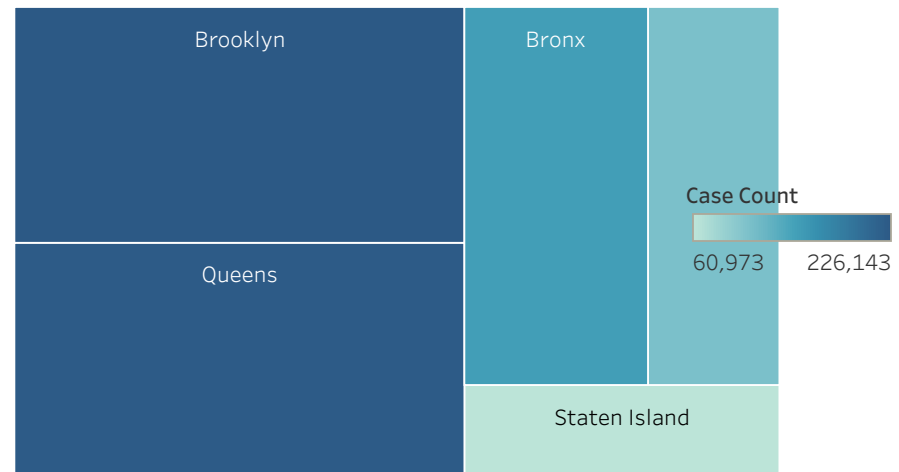
Descriptive Stats:
different rates of
infection in NYC

Descriptive Stats: 4
factors

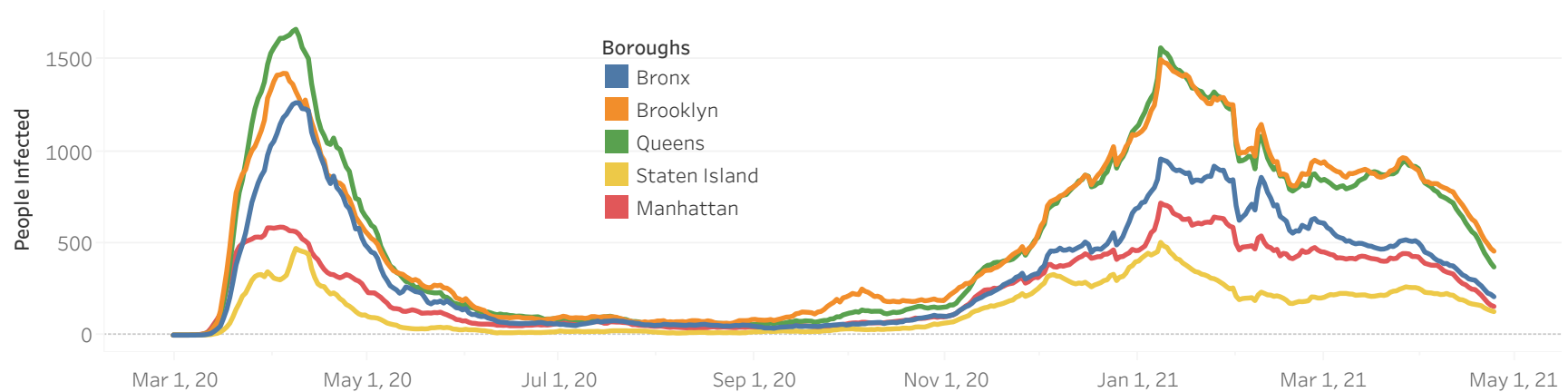
Descriptive Stats:
household mean
income (negative)

Roughly very large
case counts corresponded with
large populations

total cases per borough (by May 2020)



cases per day (based on 7 day avg)



COVID

Intro: US on
April 9, 2020

Intro: New York State
and New York City on
April 9, 2020

Descriptive Stats: NYC
borough case counts
per day

Descriptive Stats:
different rates of
infection in NYC

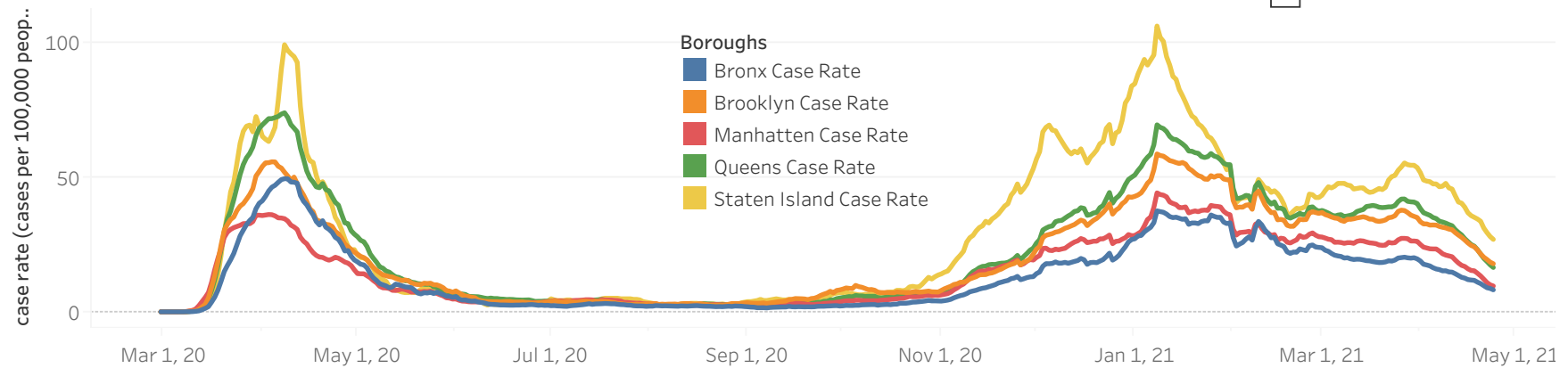
Descriptive Stats: 4
factors

Descriptive Stats:
household mean
income (negative)

Descriptive Stats:
service jobs
(positive)

However, although Staten Island had the fewest case counts per day,
it consistently had the highest case rate per day

case rate in borough per day *(based on 7 day avg)*



COVID

Intro: New York State and New York City on April 9, 2020	Descriptive Stats: NYC borough case counts per day	Descriptive Stats: different rates of infection in NYC	Descriptive Stats: 4 factors	Descriptive Stats: household mean income (negative)	Descriptive Stats: service jobs (positive)	Descriptive Stats: percent of color (indeterminate)
--	--	--	------------------------------	---	--	---

4 Factors of Coronavirus Spread and their Correlation and their Predicted Correlation

(taken from Social Vulnerability Index - an index of factors related to a population's ability to handle public health emergencies)

1. Household Mean Income -
2. Percent Service Jobs +
3. Percent of Color +
4. Median Age +

COVID

Descriptive Stats: NYC borough case counts per day

Descriptive Stats: different rates of infection in NYC

Descriptive Stats: 4 factors

Descriptive Stats: household mean income (negative)

Descriptive Stats: service jobs (positive)

Descriptive Stats: percent of color (indeterminate)

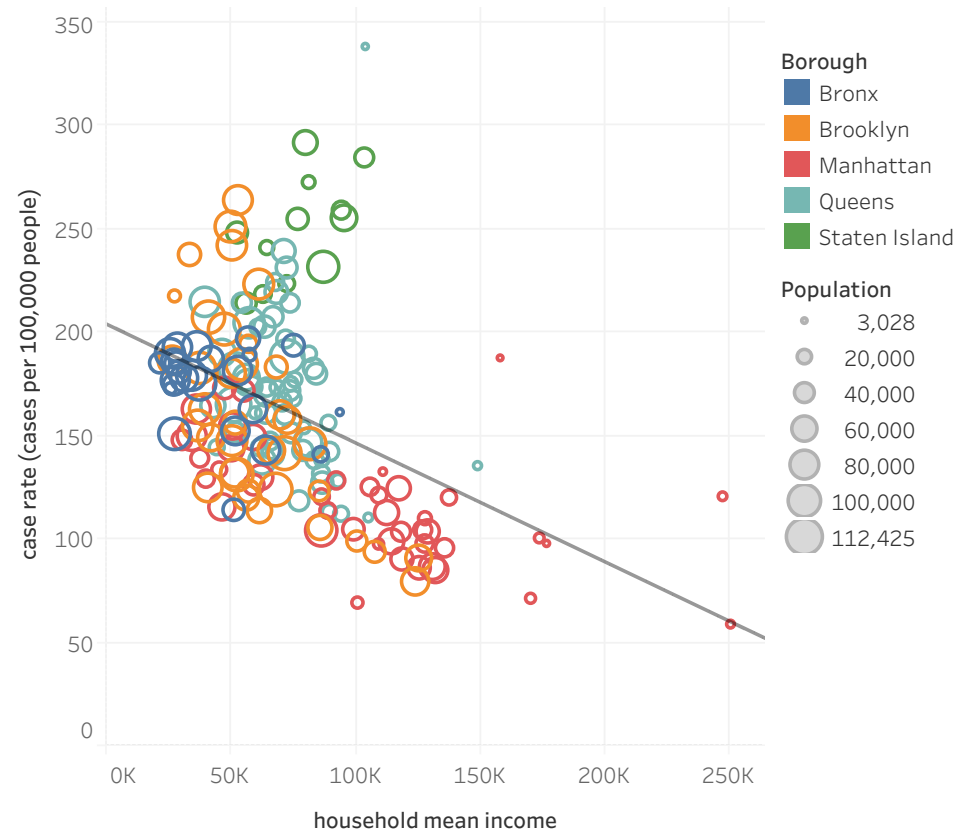
Descriptive Stats: thoughts on median age (indeterminate)

Insights in scatter plots:

- 1) Staten Island's case rate usually higher than normal
- 2) Manhattan condenses around high income, low percent service jobs, and low percent color
- 3) Bronx, Brooklyn, and Queens (largest boroughs) generally intermixed with each other

Household income is negatively correlated

household mean income (by zip code)
(using avg NYC case rates from 8/2020 to 5/2021)



COVID

Descriptive
Stats: different
rates of
infection in NYC

Descriptive Stats: 4
factors

Descriptive Stats:
household mean
income (negative)

Descriptive Stats:
service jobs (positive)

Descriptive Stats:
percent of color
(indeterminate)

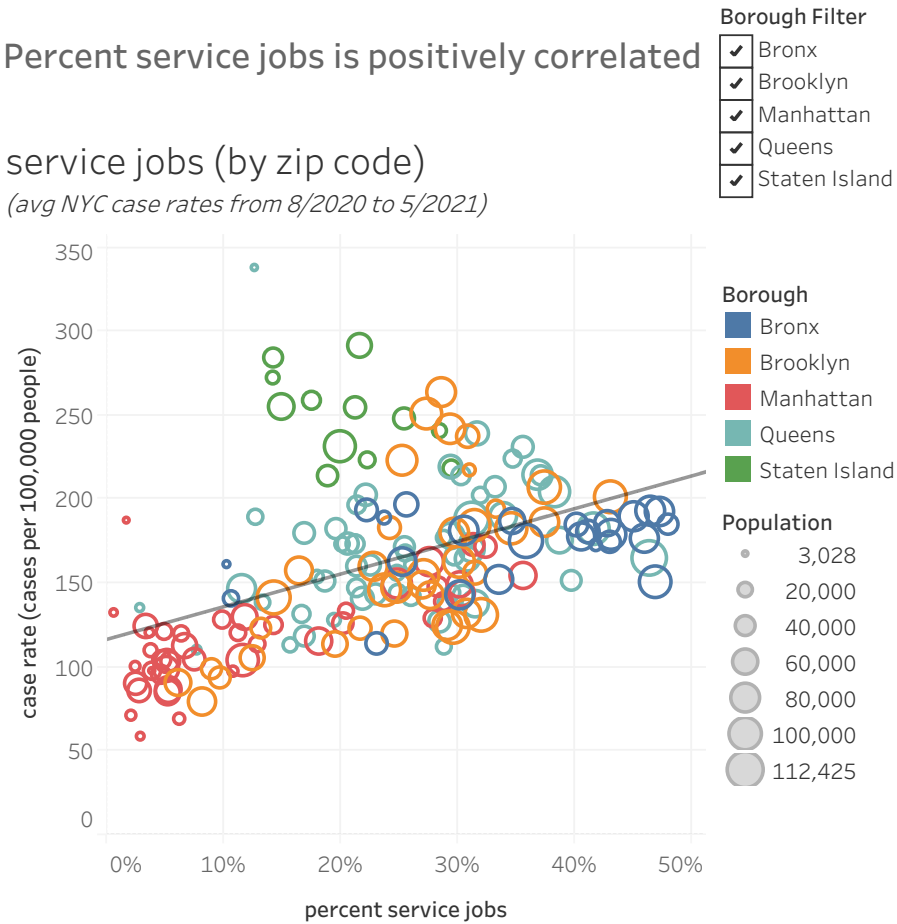
Descriptive Stats:
thoughts on median
age (indeterminate)

Inferential Stats:
linear regression
using
demographics

Percent service jobs is positively correlated

service jobs (by zip code)

(avg NYC case rates from 8/2020 to 5/2021)



COVID

Descriptive Stats: different..

Descriptive Stats: 4 factors

Descriptive Stats: household mean income (negative)

Descriptive Stats: service jobs (positive)

Descriptive Stats: percent of color (indeterminate)

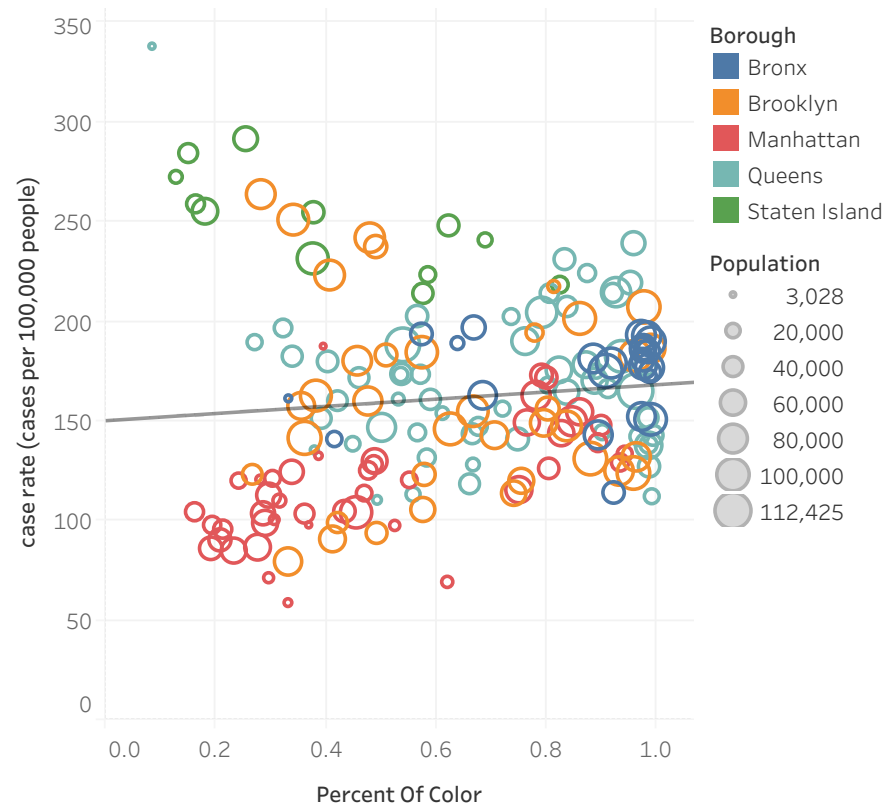
Descriptive Stats: thoughts on median age (indeterminate)

Inferential Stats: linear regression using demographics

Percent of color is indeterminate

percent color (by zip code)

(avg NYC case rates from 8/2020 to 5/2021)



COVID

Descriptive Stats: different..

Descriptive Stats: 4 factors

Descriptive Stats: household mean income (negative)

Descriptive Stats: service jobs (positive)

Descriptive Stats: percent of color (indeterminate)

Descriptive Stats: thoughts on median age (indeterminate)

Inferential Stats: linear regression using demographics

Borough

- Bronx
- Brooklyn
- Manhattan
- Queens
- Staten Island

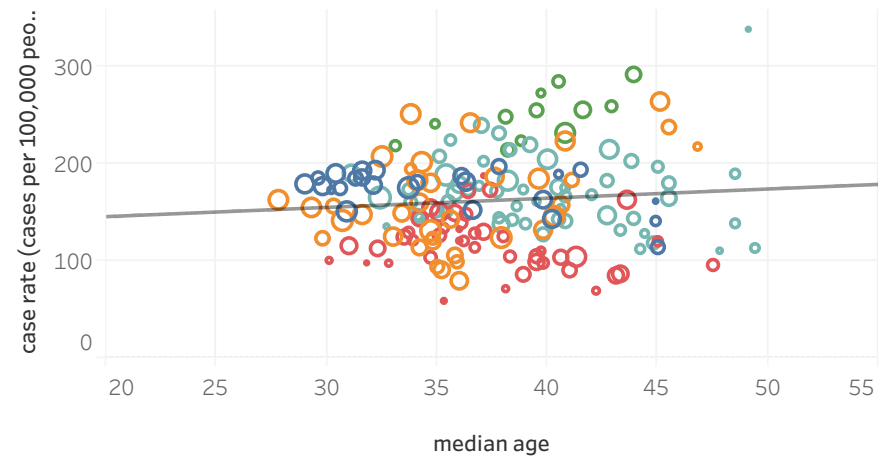
Borough Filter

- ☒ Bronx
- ☒ Brooklyn
- ☒ Manhattan
- ☒ Queens
- ☒ Staten Island

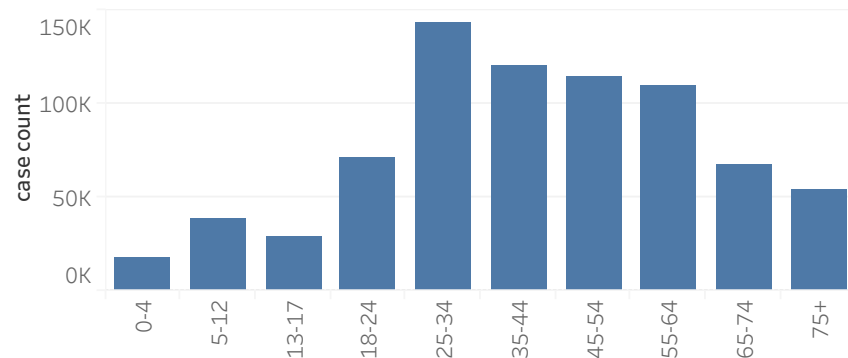
note: While median age has a positive correlation with the covid case rate, by late April 2021 the case counts age distribution reflected the population age distribution, indicating that by this date (diagnosed) Covid had affected the ages equally

median age (by zip code)

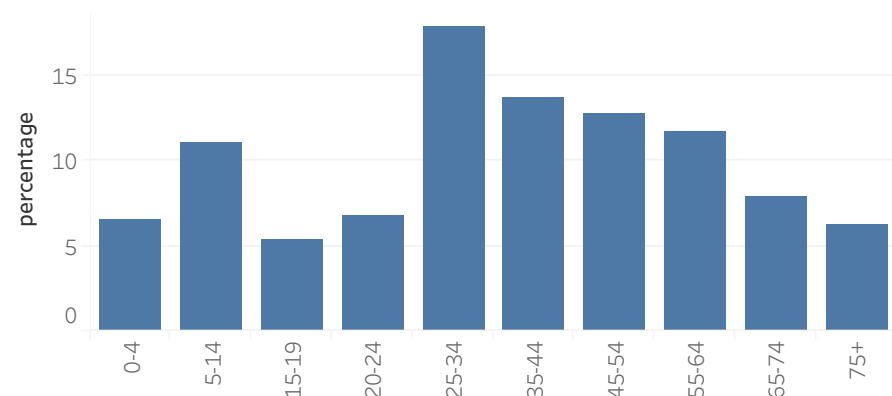
(using avg NYC case rates from 8/2020 to 5/2021)



covid case age distribution in NYC (by late April 2021)



population age distribution of NYC



COVID

Descriptive Stats: different..	Descriptive Stats: 4 factors	Descriptive Stats: household mean income (negative)	Descriptive Stats: service jobs (positive)	Descriptive Stats: percent of color (indeterminate)	Descriptive Stats: thoughts on median age (indeterminate)	Inferential Stats: linear regression using demographics
--------------------------------	------------------------------	---	--	---	---	---

+ **coefficients:** median age, percent service jobs - **coefficients:** mean household income, percent..

regression results
(using standardized variables)

Variables	Coefficients
intercept	163.64
mean household income	-20.89
median age	12.87
percent of color	-42.76
percent service jobs	42.80

context

Variables	Mean	Standard Devia
household mean income	60,440.00	27,321
median age	36.30	4
percent of color	0.69	(
percent service jobs	0.27	(

R squared = 0.69

dependent variable:

cases per 100,000 people (by zip code)

Each coefficient represents the change in the case rate per change in the number of standard deviations of the independent variable (value minus mean divided by std dev). R squared may be 0.69 because it is hard to predict mass human phenomenon or/and because additional significant variables could be inc..

The coefficient for percent of color is negative, enough though it is commonly believed that being non-white is positively correlated with contracting covid. However, if one combines Brooklyn, Queens, the Bronx, and Staten Island together, NYC has a negative correlation between case rate and percent of color. One explanation is that this study uses case rate instead of positivity rate. Case rate is positive covid results per number of people, while the positivity rate is positive covid results per tests. Hence, case rates usually underestimate the rate of infection, whereas positivity rates overestimate it (since people are more likely to be tested if they have symptoms). The calculated correlation between percent of color and test rate is -0.53, meaning there is less testing where the..

* only used 73 records (out of 177) where zip code population was above 50,000 (reducing the effect of a zip code with a pop. of 3000 having the same impact ..