

# MTA RIDERSHIP IMPACT OF COVID-19

HOW COVID-19 EFFECTED MTA USAGE IN NYC

# PRESENTER

- ❑ Thomas Joy II - Data Scientist
- ❑ Bachelor's of Arts in Mathematics with minor in Computer Science from Hunter College - CUNY
- ❑ Pursuing Master's of Arts in Applied Mathematics
- ❑ Former Active Duty Marine Corps, Presently Air Force Reservist
- ❑ I have a dog: Cinny





# BACKGROUND INFORMATION

- The first laboratory-confirmed case of COVID-19 in the United States was confirmed on January 20, 2020, and reported to CDC on January 22, 2020.
- As of September 30, 2020, there has been 8.93M COVID-19 cases and 228K COVID-19 deaths across the United States.
- The NYC MTA subway service has been and still is suspended from 1 a.m. to 5 a.m. for disinfection of stations and trains.
- It is and has been a requirement by law to wear a mask while on MTA subways.

# ECONOMIC IMPACT OF CORONAVIRUS FOR NEW YORKERS



## NYC

Coronavirus likely to cost NYC over \$10 billion in lost revenue.

## MTA


The New York MTA projects total coronavirus losses between \$7 - \$8.5 billion. There are talks about a possible 40% cut in MTA services.

## INDIVIDUAL

NYC unemployment rate is 16%, twice as high as the rest of the country. Personal income tax revenue is expected to drop by \$2 billion this fiscal year.

# THE DATA

Note: All time-series data was filtered between 03/01 - 09/30



## TURNSTILE DATA FROM 2019

- 11 Columns
- 10.5M Rows



## TURNSTILE DATA FROM 2020

- 11 Columns
- 11.2M Rows



## COVID-19 DATA BY ZIP CODE

- 10 Columns
- 178 Rows



## COVID-19 DATA BY BOROUGH

- 7 Columns
- 7 Rows



## COVID-19 TESTING DATA BY DAY

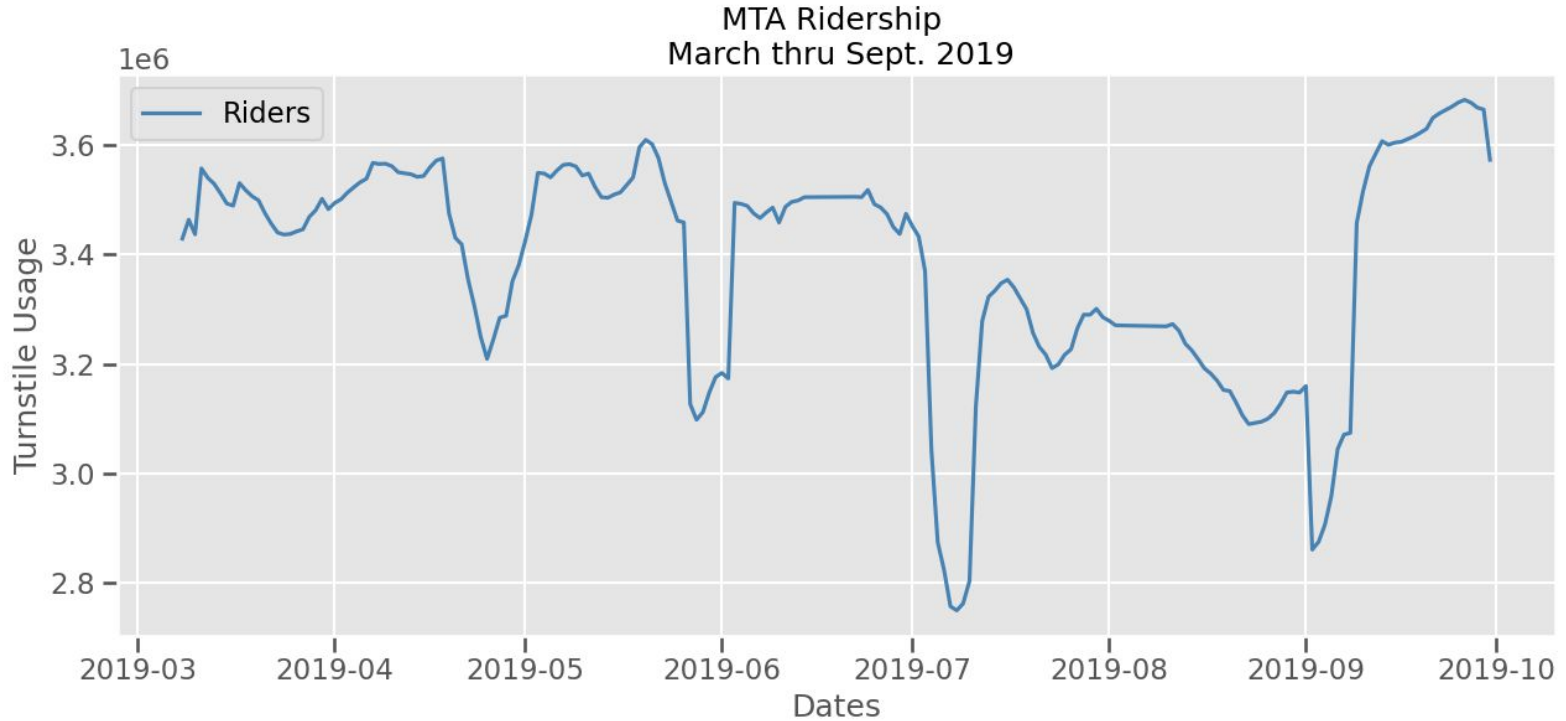
- 8 Columns
- 235 Rows



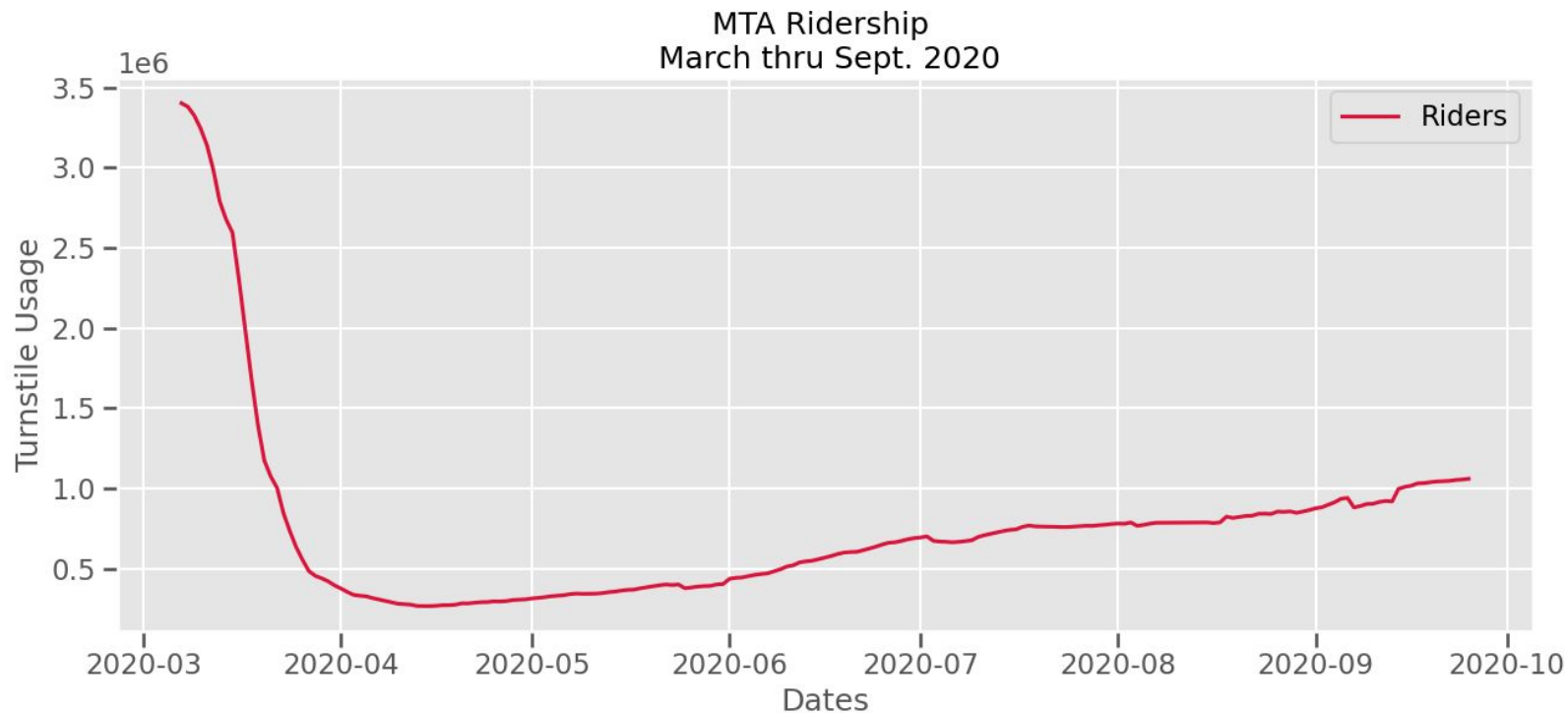
## MEDIAN INCOME BY ZIP CODE

- 6 Columns
- 178 Rows

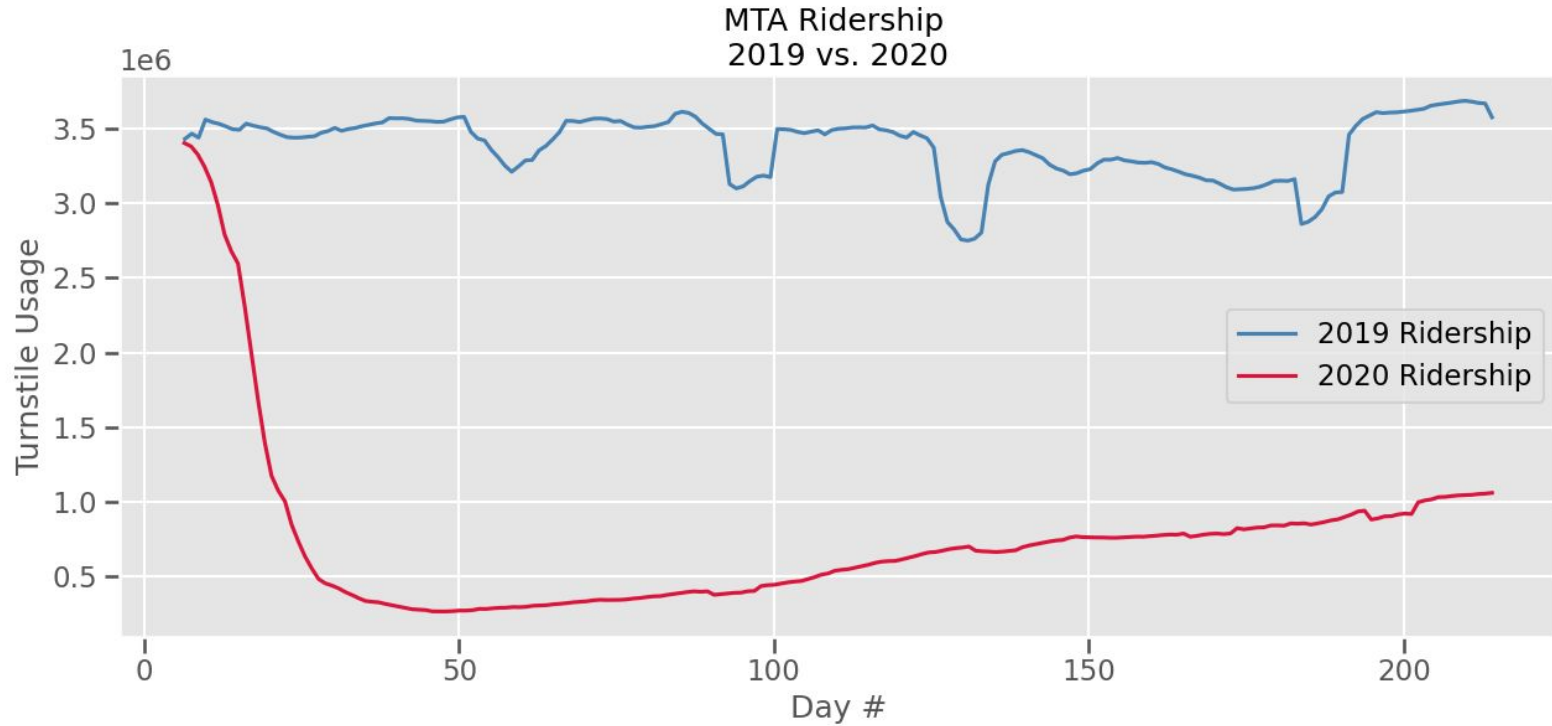
# MTA TURNSTILE USAGE 2019



# MTA TURNSTILE USAGE 2020

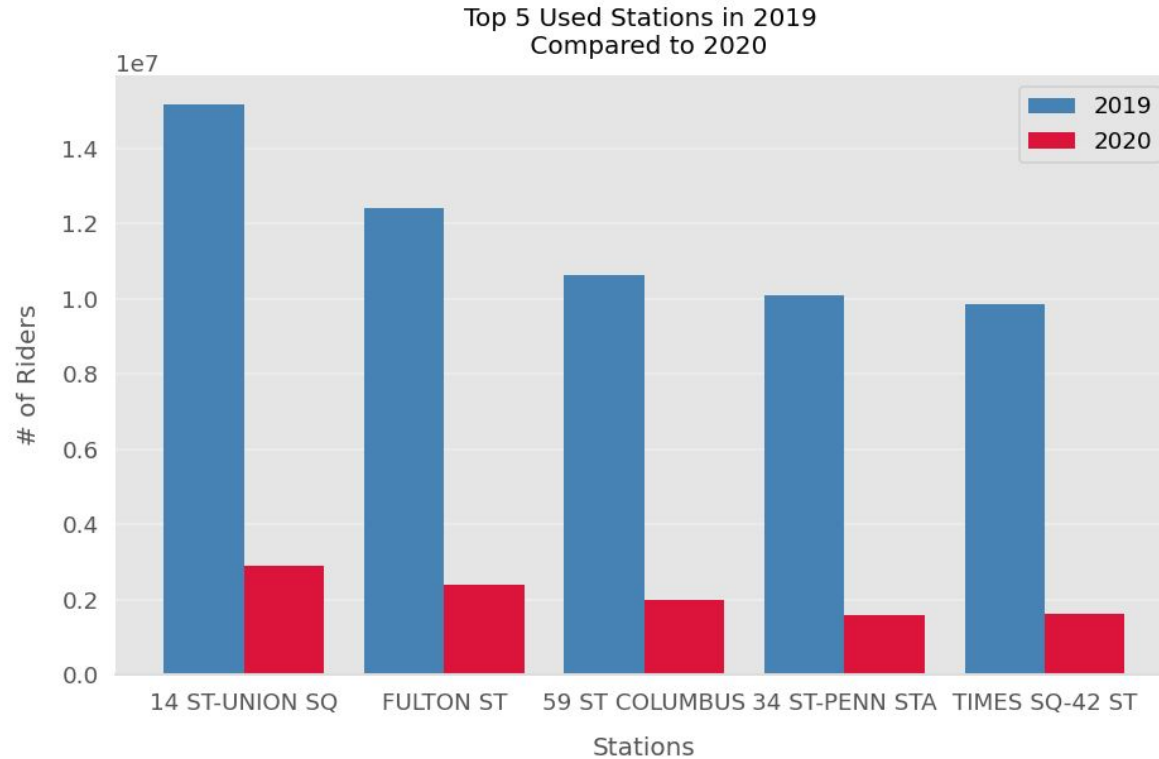


# MTA TURNSTILE USAGE COMPARISON





# STATION TRAFFIC FOR TOP 5 STATIONS IN 2019



## HYPOTHESIS TESTING

Collect &  
Clean Data

Null Hypothesis: There was NO difference in MTA turnstile usage between 03/01/2019 - 09/30/2019 and 03/01/2020 - 09/30/2020.

Alternate Hypothesis: There was a difference in MTA turnstile usage between 03/01/2019 - 09/30/2019 and 03/01/2020 - 09/30/2020.

Significance Level:  $\alpha = 0.05$

## CONCLUSION

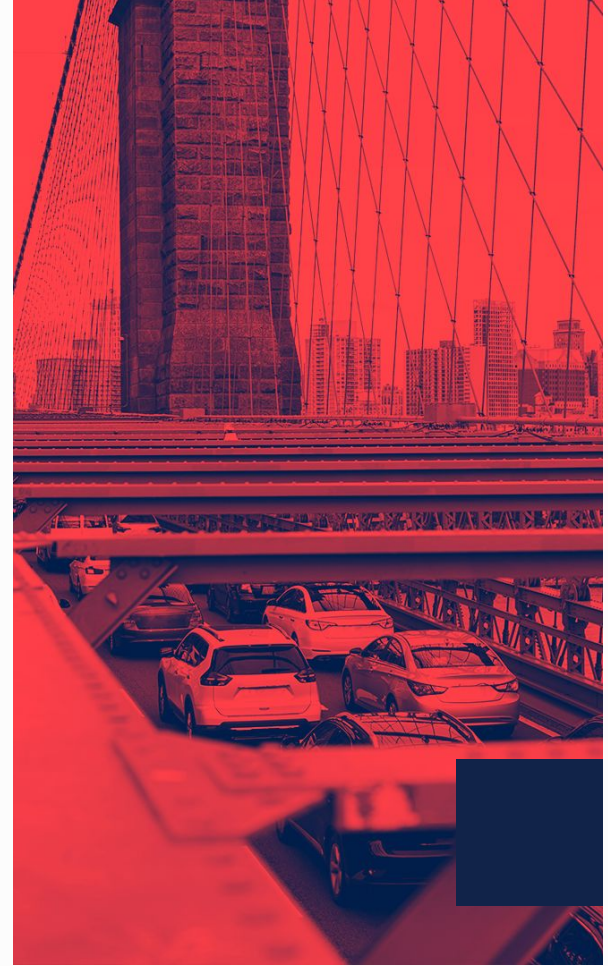
Since the p-value is lower than our alpha, we reject the Null Hypothesis and can say there is sufficient statistical evidence that there was a difference in MTA Turnstile usage between March 01, 2019 thru September 30, 2019 and March 01, 2020 thru September 30, 2020.

**58.645**  
Test Statistic

**1.1265...833e-145**  
P-Value

# EDA OF NYC COVID-19 DATA

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. Most people infected with the COVID-19 virus will experience mild to moderate respiratory illness and recover without requiring special treatment.



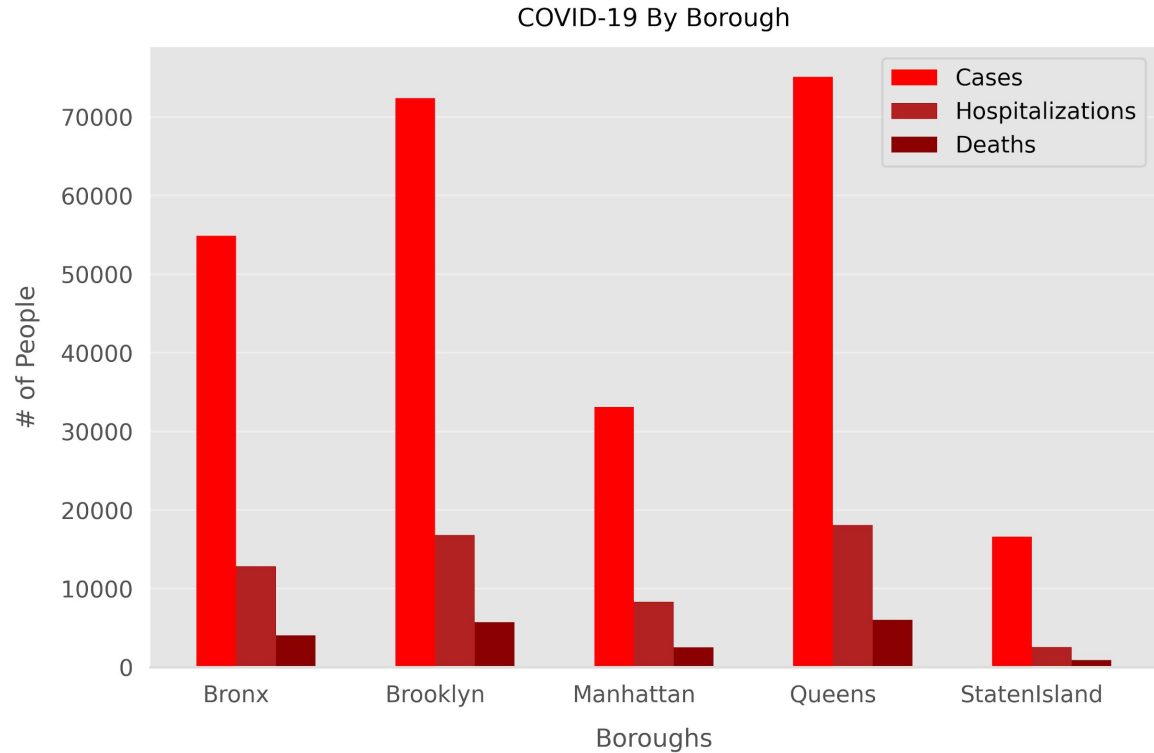
## BY BOROUGH

### Citywide Numbers:

Cases: 252,281

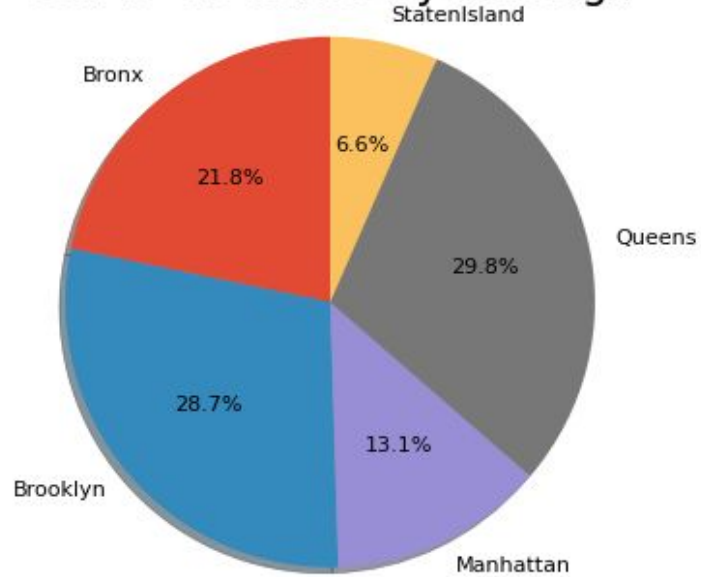
Hospitalizations: 58,695

Deaths: 19,314



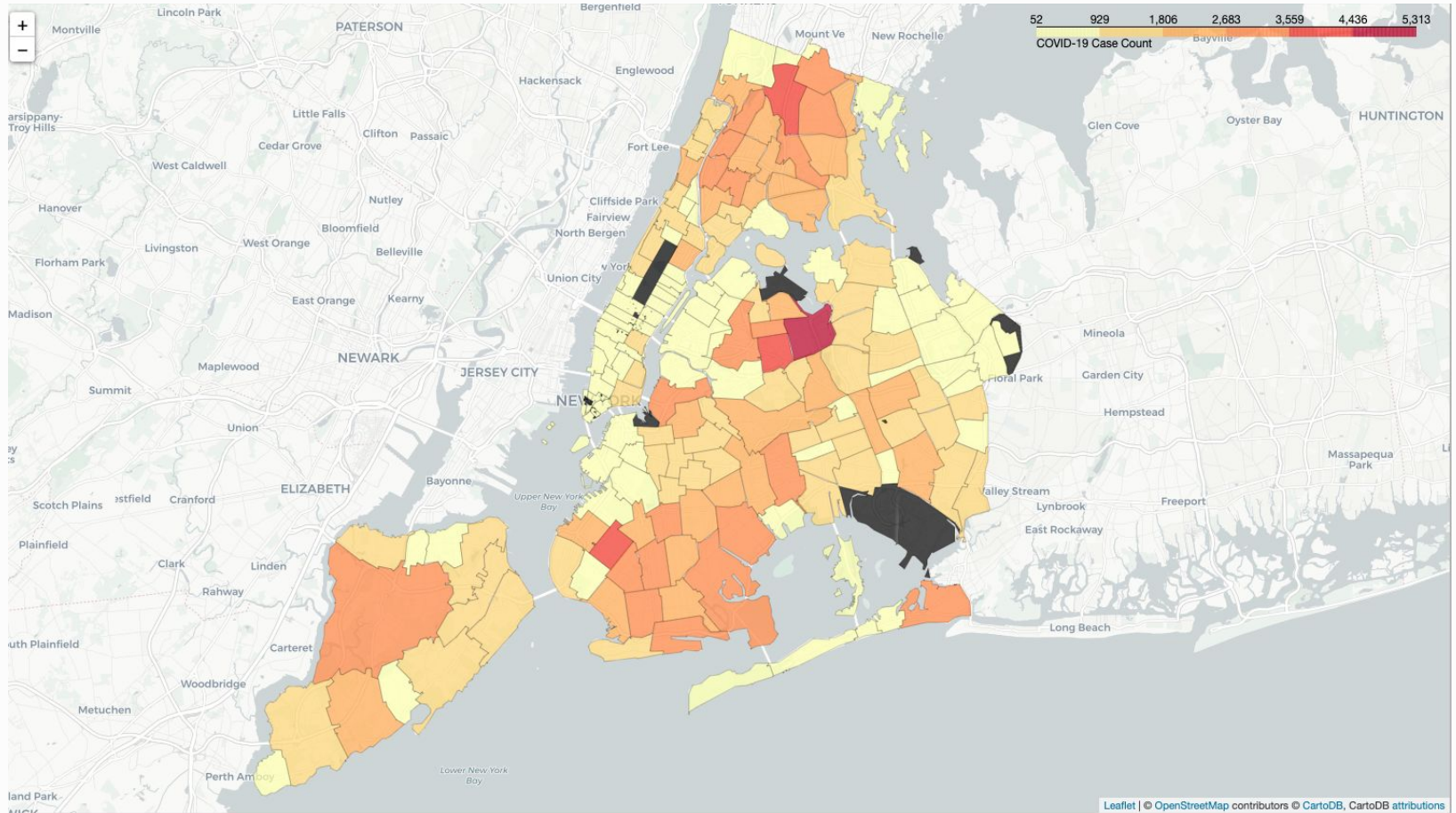
## BY BOROUGH

COVID-19 Cases by Borough



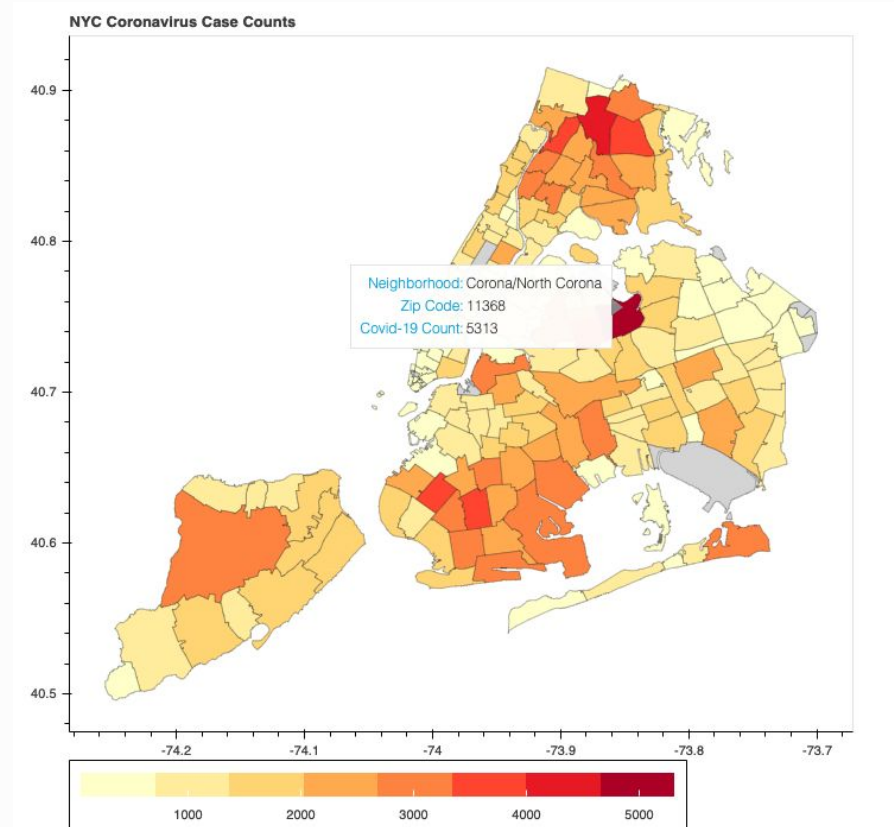
# BY ZIP CODE

## COVID-19 CASE COUNTS



# BREAKDOWN OF COVID-19 CASE COUNTS

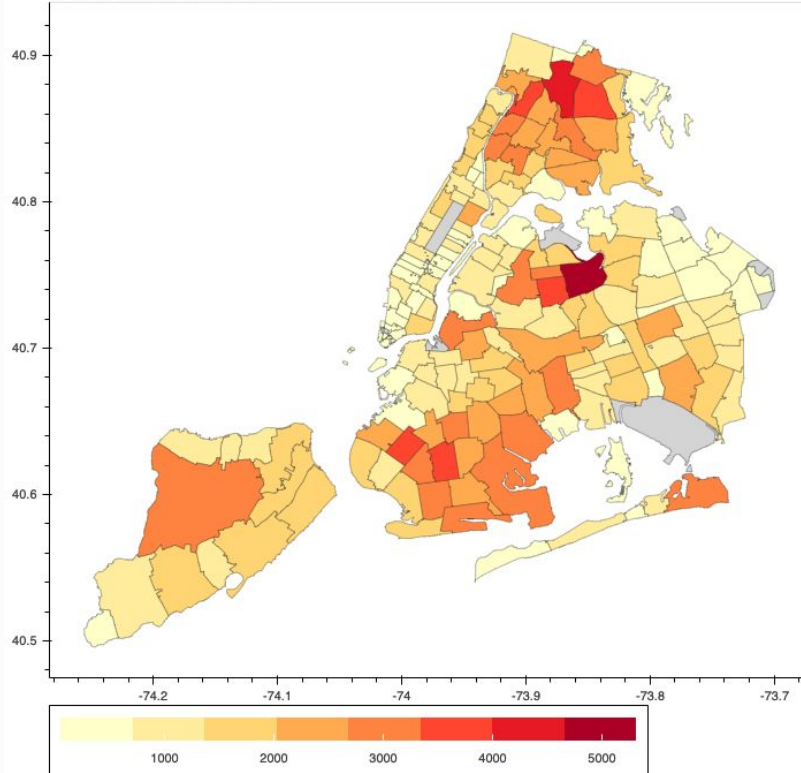
- ❑ Zip code with the most COVID-19 cases:
  - ❑ Zip Code: 11368
  - ❑ Neighborhood: Corona, Queens
  - ❑ COVID-19 Cases: 5,313
  
- ❑ Zip code with the least COVID-19 cases:
  - ❑ Zip Code: 10006
  - ❑ Neighborhood: Financial District, NY
  - ❑ COVID-19 Cases: 52



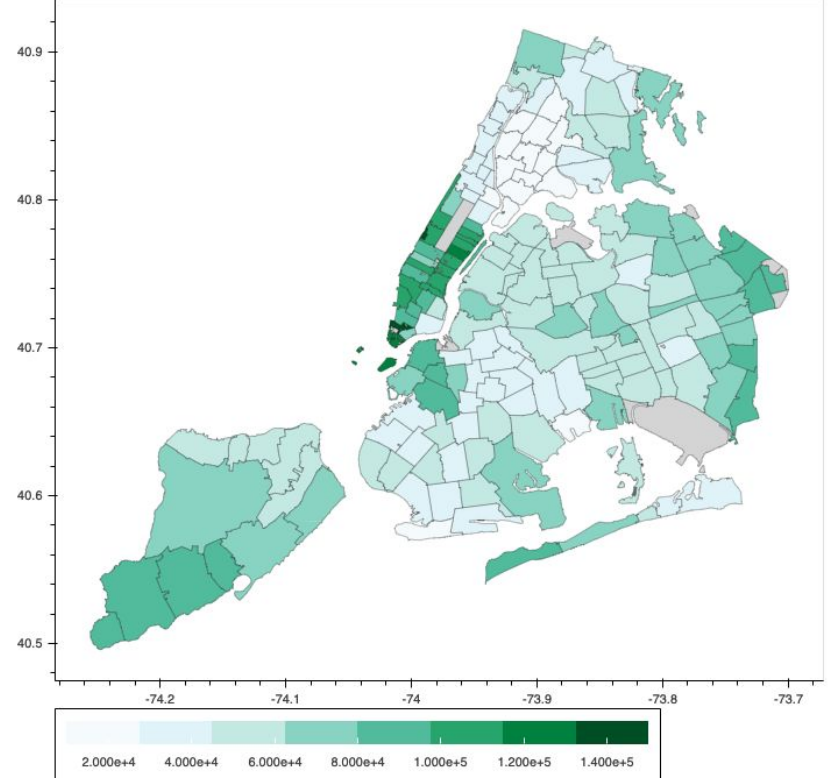


# COVID-19 CASE COUNT VS MEDIAN INCOME LEVELS

NYC Coronavirus Case Counts

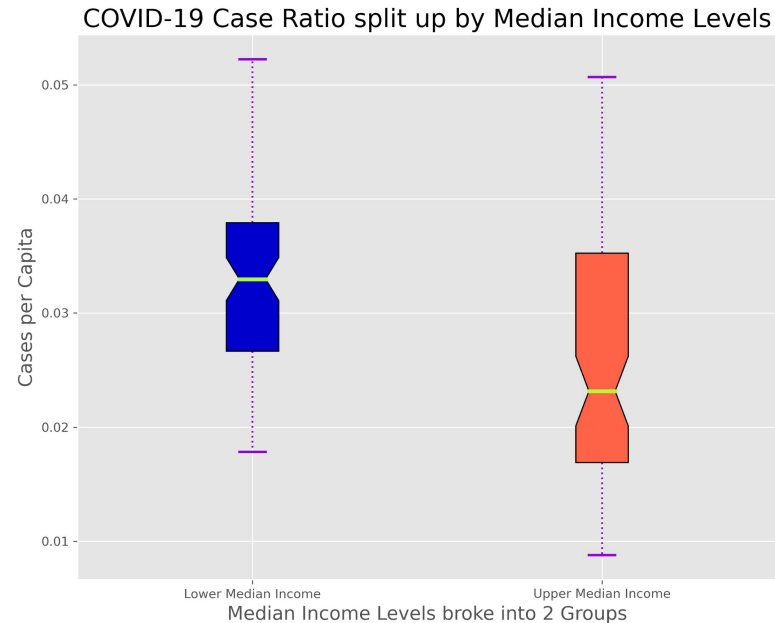


NYC Median Income Levels





# COVID-19 CASE COUNT PER CAPITA COMPARED TO MEDIAN HOUSEHOLD INCOMES



# HYPOTHESIS TEST

## NULL HYPOTHESIS



There is No difference in the covid-19 case count ratio and median household incomes.

## ALTERNATE HYPOTHESIS



There is a difference in the covid-19 case count ratio and median household incomes.

## SIGNIFICANCE LEVEL



$\alpha = 0.05$

## TEST RESULTS

TEST STATISTIC



4.7844

P-VALUE



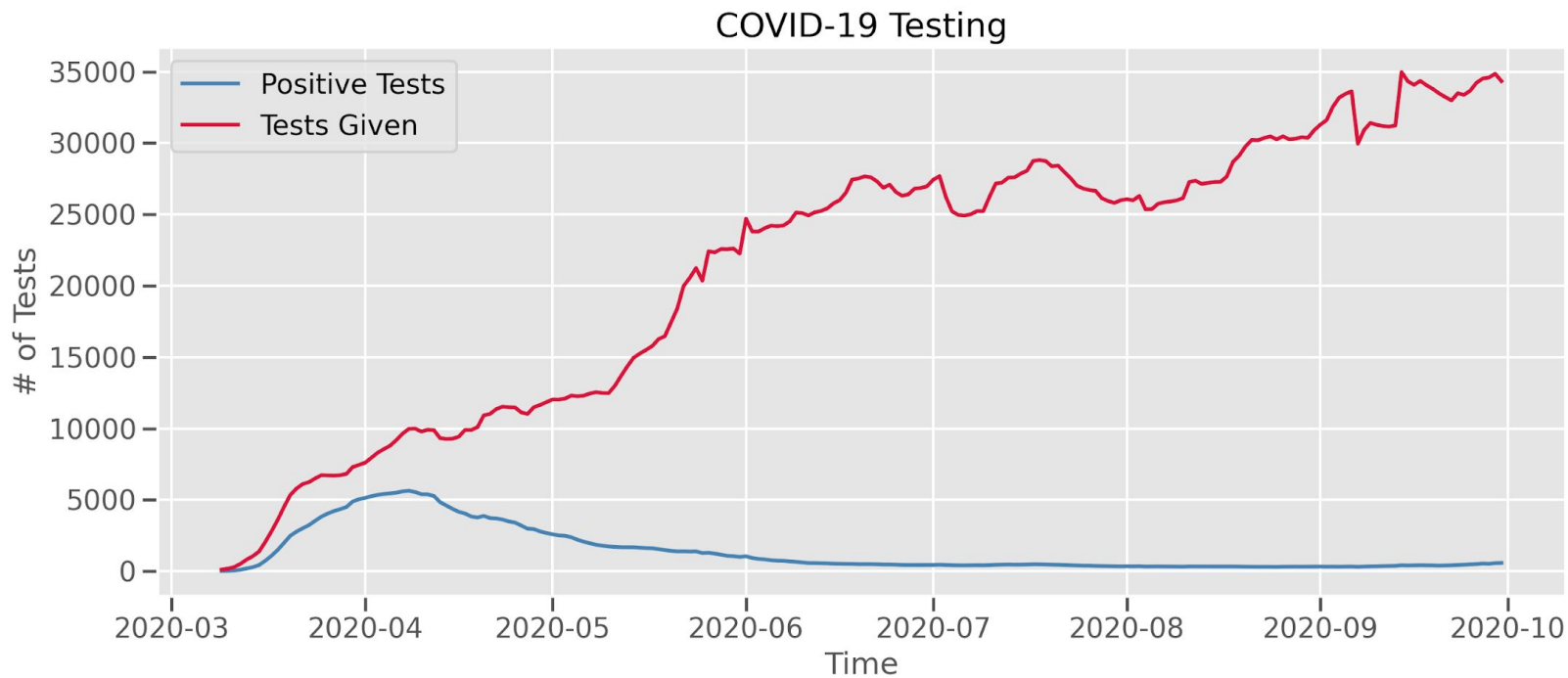
3.871...707e-06

CONCLUSION

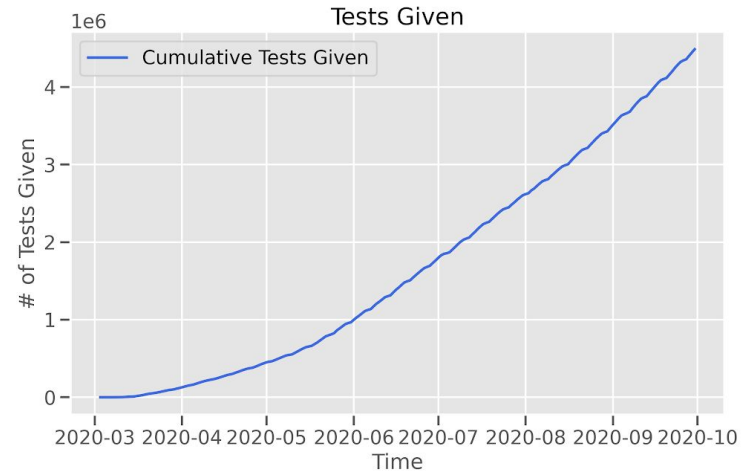
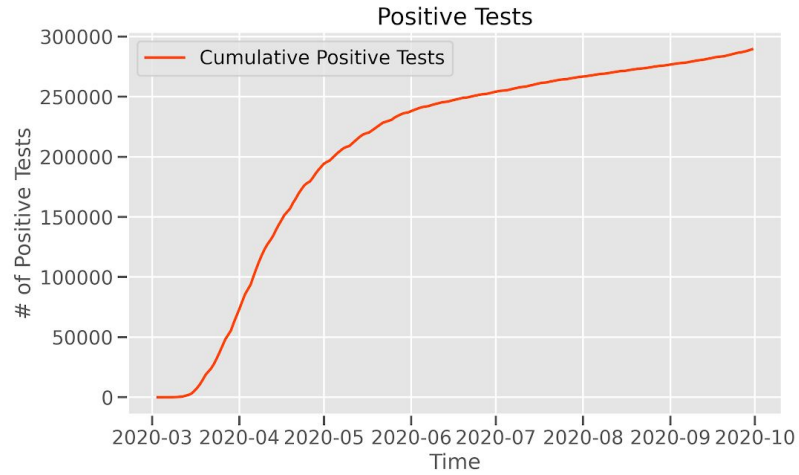


Since the p-value is lower than our alpha, we can reject the Null Hypothesis and say that there is sufficient statistical data to say there is a difference in the COVID-19 case count ratio and median household incomes.

# TESTING DATA



# TESTING DATA

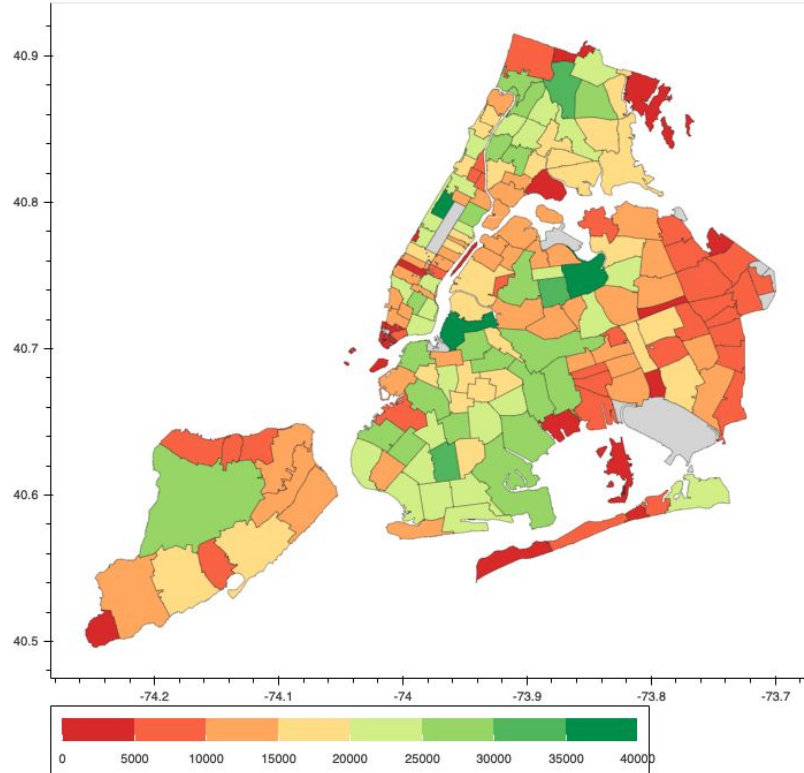


Total Tests Given: 4,487,431

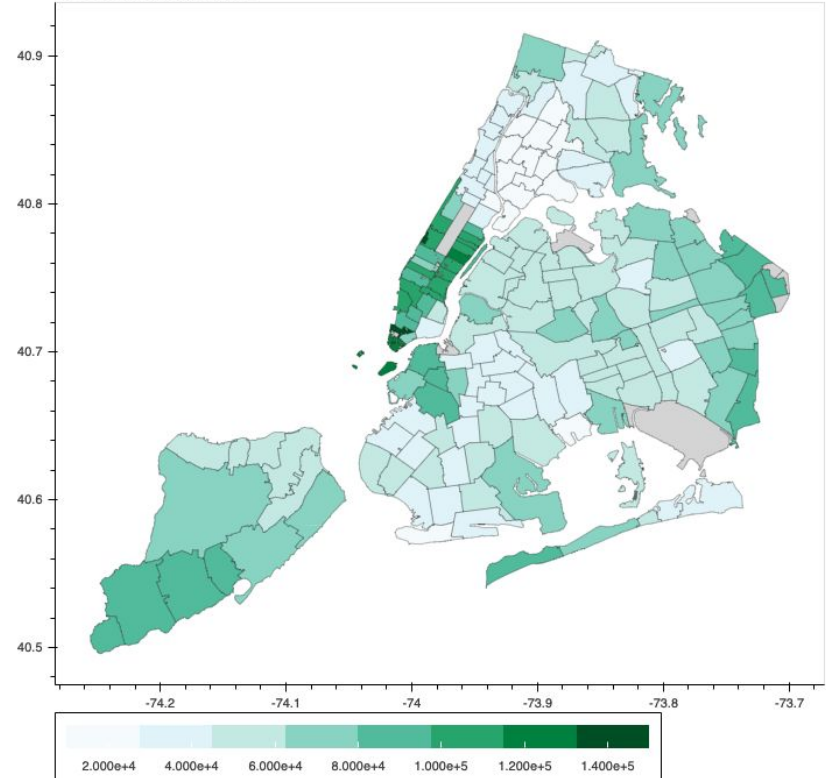
Total Positive Test Results: 289,475

# COVID-19 TESTS GIVEN VS MEDIAN INCOME LEVELS

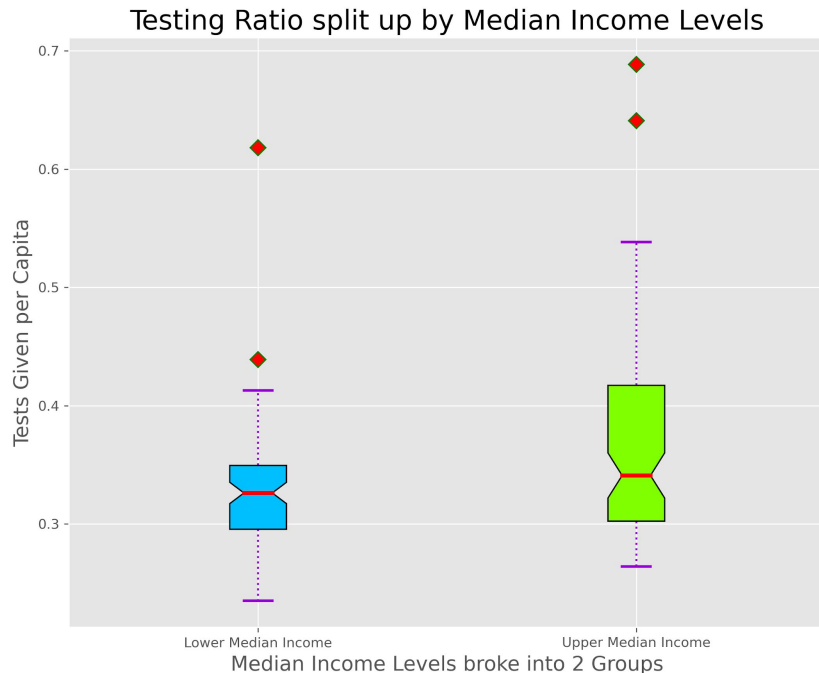
NYC Coronavirus Tests Given



NYC Median Income Levels



# COVID-19 TESTING PER CAPITA COMPARED TO MEDIAN HOUSEHOLD INCOMES



## Hypothesis Test

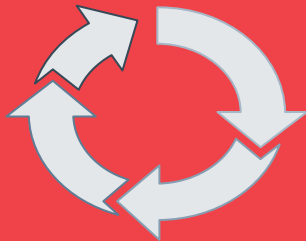
Null Hypothesis: There is NO difference in the testing rate given and median household incomes.

Alternate Hypothesis: There is a difference in the testing rate given and median household incomes.

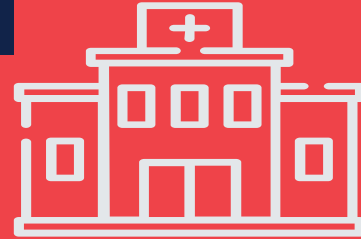
Significance Level:  $\alpha = 0.05$

TEST STATISTIC: -3.8996

P-VALUE: 0.0001



## TEST RESULTS



### CONCLUSION

Since the p-value is lower than our alpha, we can reject the Null Hypothesis and say that there is sufficient statistical data to say there is a difference in the testing rate given and median household incomes.



The background of the slide is a photograph of a city street, likely New York City, with a red overlay. The image shows a street with a brick-paved sidewalk in the foreground, a person walking on the right, and a building with a sign that says "HOTEL" on the left. The red overlay is semi-transparent, allowing the city scene to be visible behind the text.

# CONCLUSION

- ❑ There was indeed a difference in MTA usage between 2019 and 2020. In 2020, turnstile usage dropped significantly.
- ❑ Covid-19 was a main contributor to this.
- ❑ Lower Median Household Income levels do coincide with less tests available and more COVID-19 cases.

## NEXT STEPS

- ❑ Collect more data after this winter to see if trends stay the same.
- ❑ Look into other causes why people may have stopped taking the subways (i.e. MTA policies, Safety due to new laws, etc.)

# RESOURCES

## LINKS

- ❑ <https://data.ny.gov/Transportation/Turnstile-Usage-Data-2019/xfn5-qji9>
- ❑ <https://data.ny.gov/Transportation/Turnstile-Usage-Data-2020/py8k-a8wg>
- ❑ <https://github.com/nychealth/coronavirus-data>
- ❑ <https://www1.nyc.gov/site/doh/covid/covid-19-data.page>
- ❑ <https://data.cityofnewyork.us/browse?category=Health&q=covid>
- ❑ <https://pypi.org/project/uszipcode/>

# THANKS!

Do you have any questions?

Thomas Joy II - Data Scientist  
tomjjoy123@gmail.com  
+1 718 869 9409

<https://github.com/tomjjoy>  
<https://www.linkedin.com/in/thomas-joy-ii-37a246153/>