



# The Coronavirus Outbreak

## US Analysis by State and County

Data as of March 28, 2020

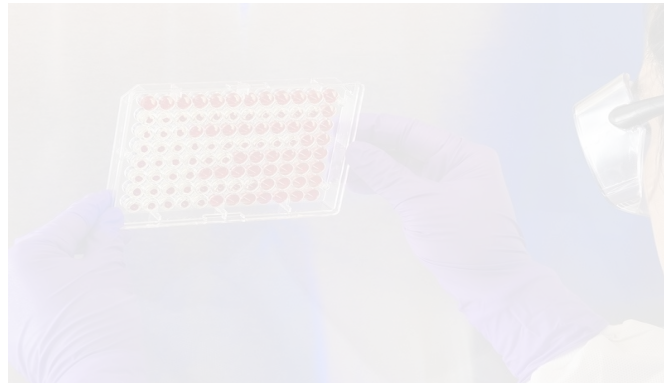
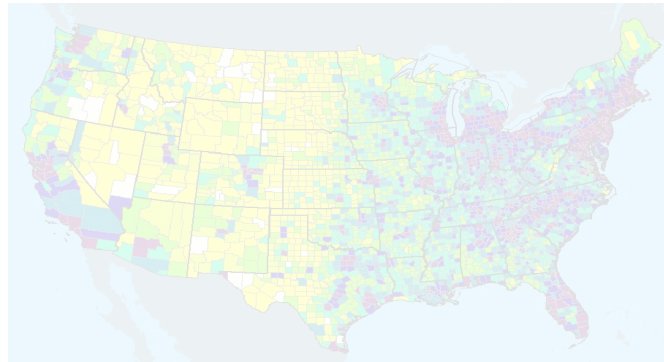
**Presented by:**

Young Hash  
Steph Hong  
Serene Yang

April 7, 2020

# Introduction: Analysis Objective

1. Determine whether there is a relationship between number of cases confirmed and population density in the US.
2. Determine whether there is a correlation between number of confirmed cases and the amount of testing occurring over time.



# Data Sources and Assumptions

- ❑ US country level history data: <https://covidtracking.com/data/>
- ❑ States level history data: <https://covidtracking.com/data/>
- ❑ Counties level history data: <https://github.com/nytimes/covid-19-data>
- ❑ US states land area: [https://tigerweb.geo.census.gov/tigerwebmain/Files/bas20/tigerweb\\_bas20\\_state\\_2010\\_us.html](https://tigerweb.geo.census.gov/tigerwebmain/Files/bas20/tigerweb_bas20_state_2010_us.html)
- ❑ US Counties land area: <https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk>
- ❑ US states and counties latitude and longitude: Google map API
- ❑ US states and counties 2018 estimated population: APIs from Census

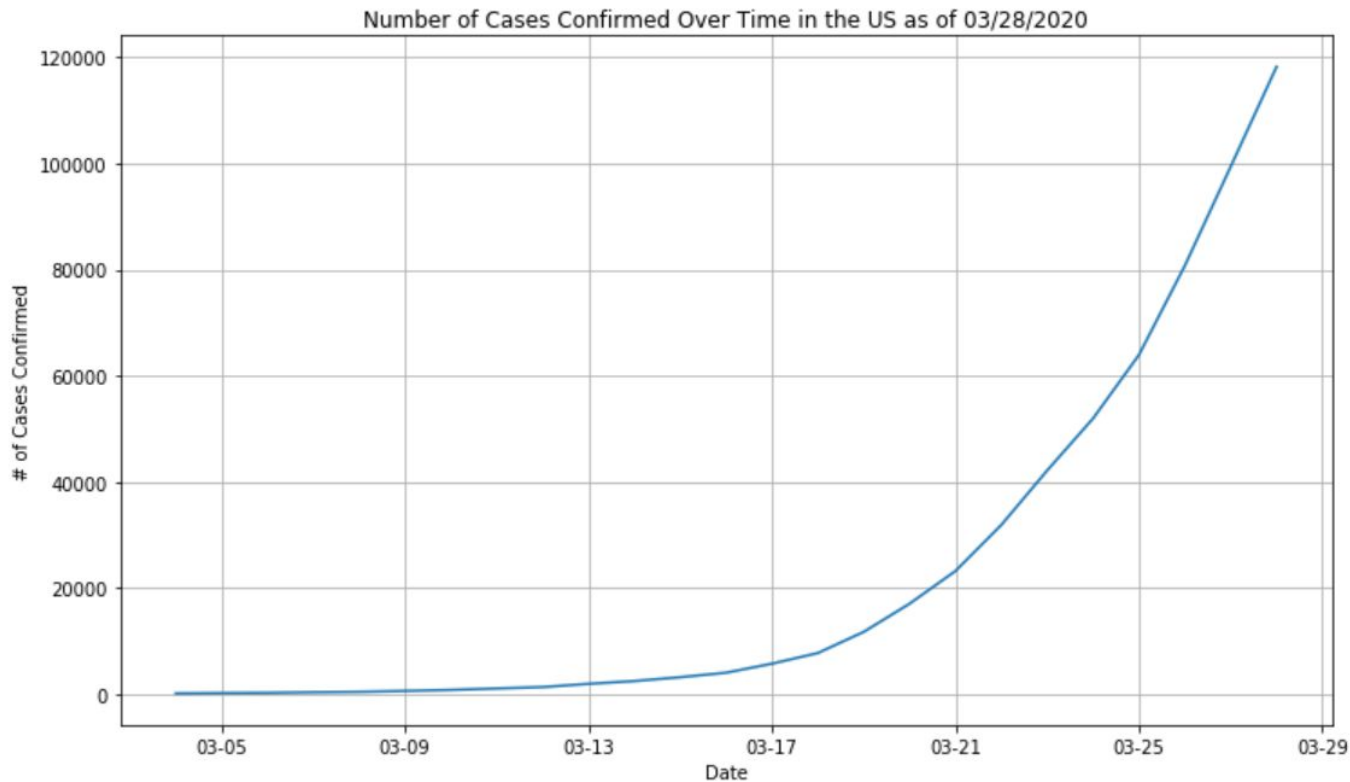
\* There are four Insular areas that have confirmed COVID-19 cases but there were no population data from Census: American Samoa(AS), Guam(GU),Northern Mariana Islands(MP),Puerto Rico(PR),U.S. Virgin Islands(VI),U.S.The total confirmed cases of these areas as of 3/28 was 179 total, less than 1% of total confirmed cases number. As such, deemed reasonable to not including in our analysis and will not impact out results.

# US Overview as of 3-28-2020

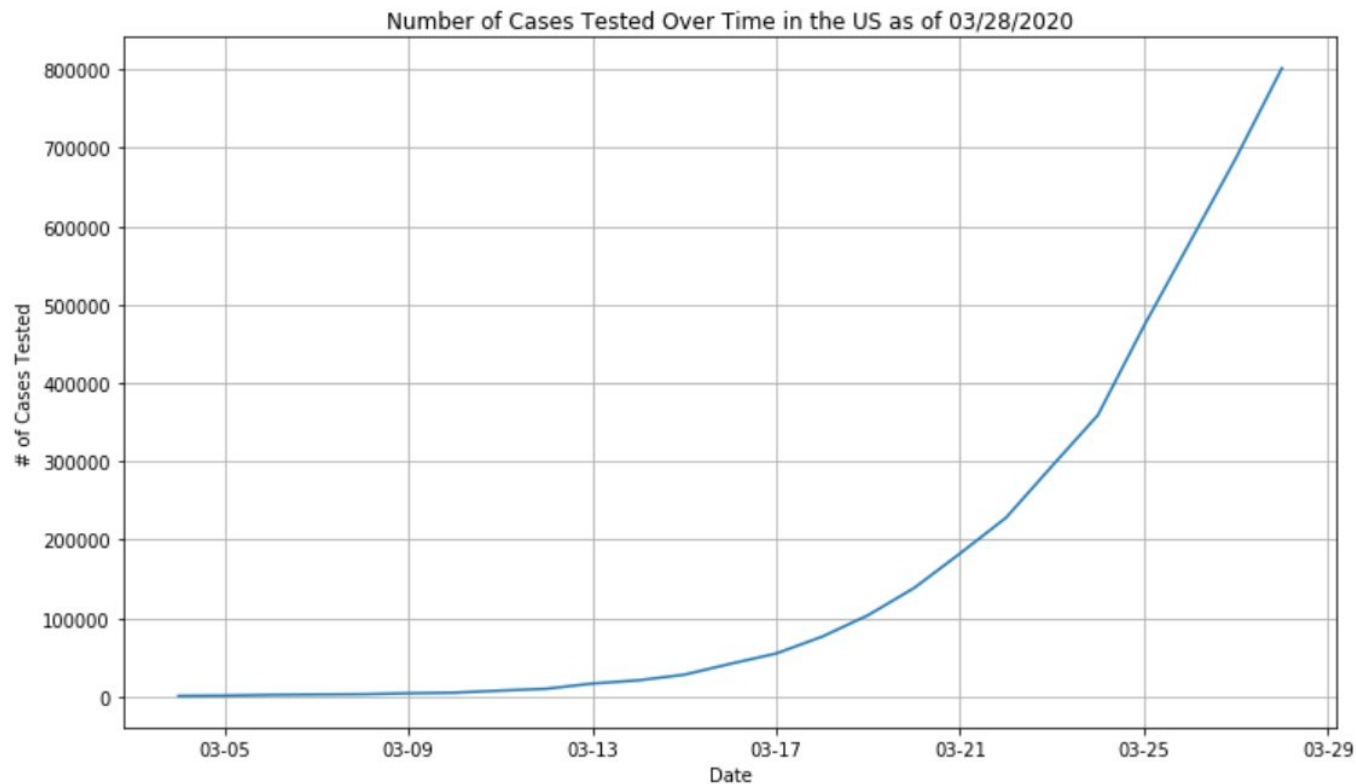
date	states	positive	negative	pending	hospitalized	death	total	totalTestResults
2020-03-28	56	118234	617470	65712	16729.0	1965.0	801416	735704

	positive	negative	hospitalized	death	totalTestResults
mean	2,111	11,026	299	35	13,138
std	7,073	16,042	1,347	102	22,432
min	0	0	0	0	0
max	52,318	103,616	10,054	728	155,934

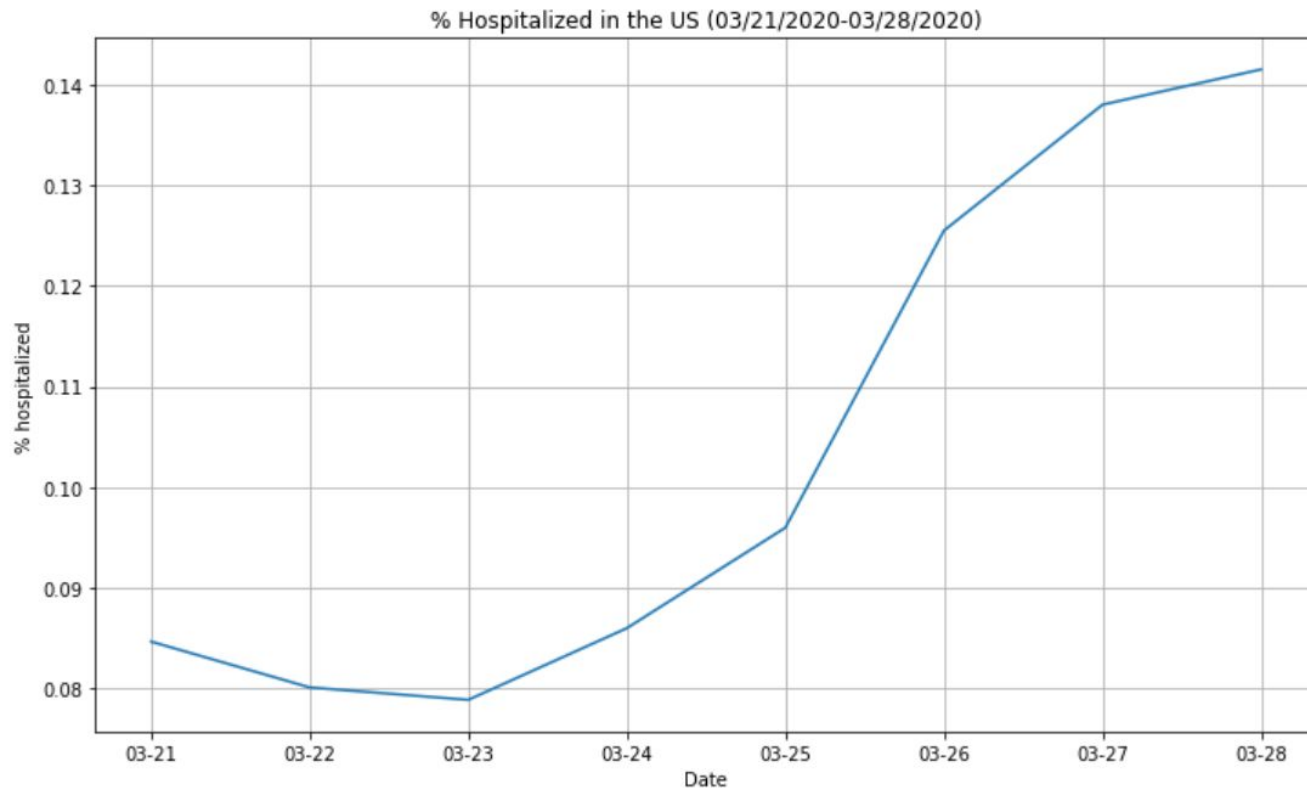
# Time Analysis - US



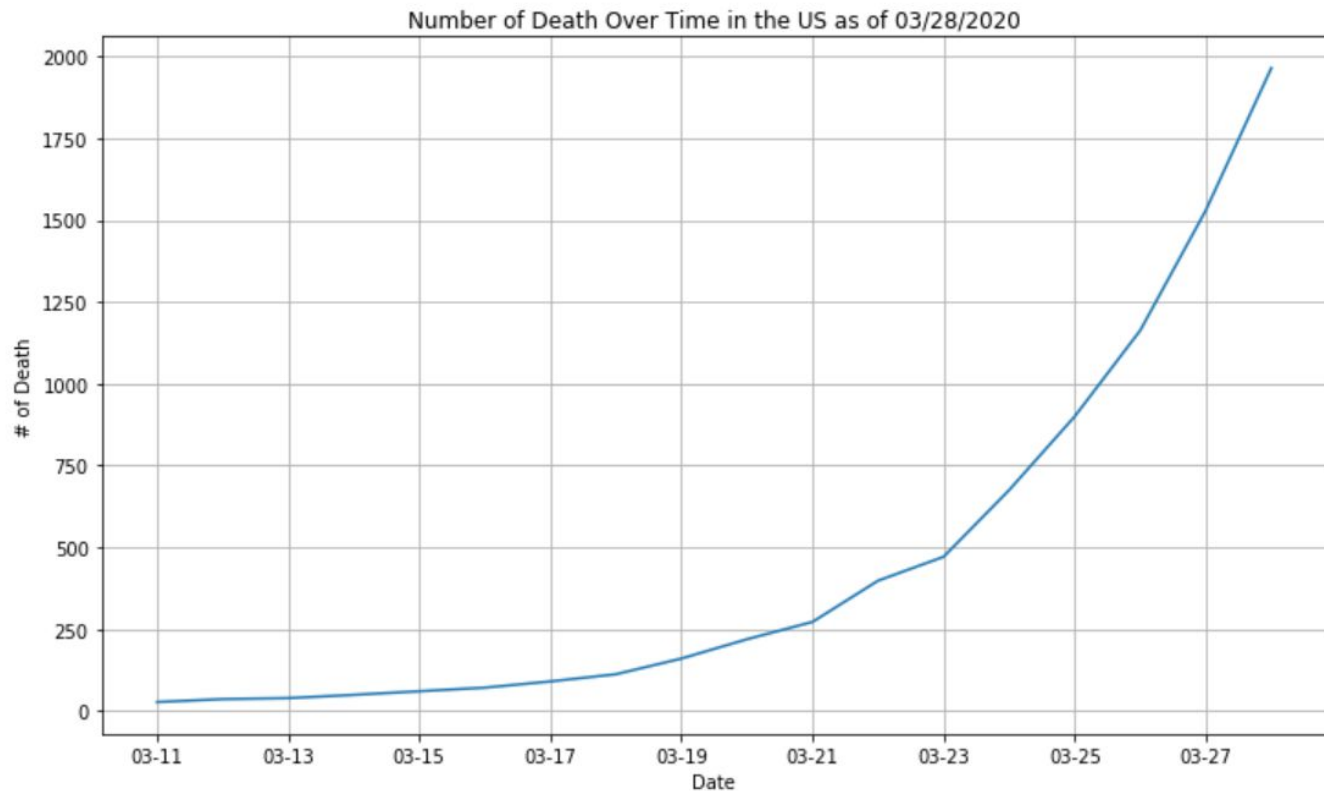
# Time Analysis - US



# Time Analysis - US

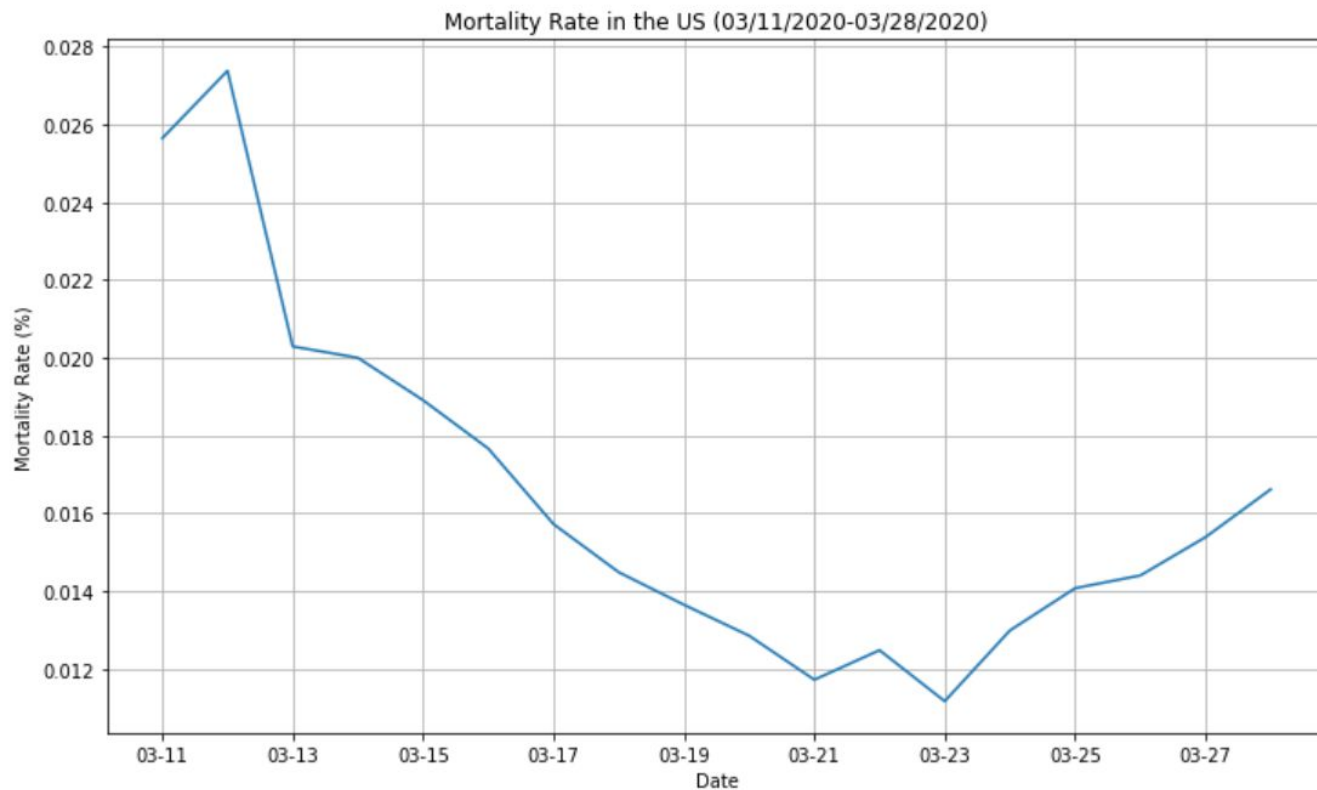


# Time Analysis - US

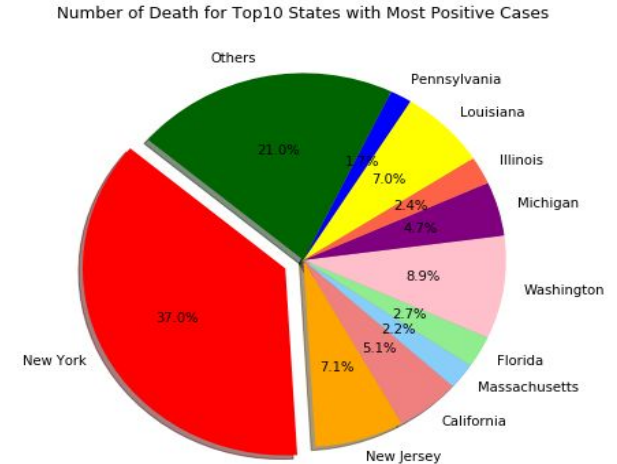
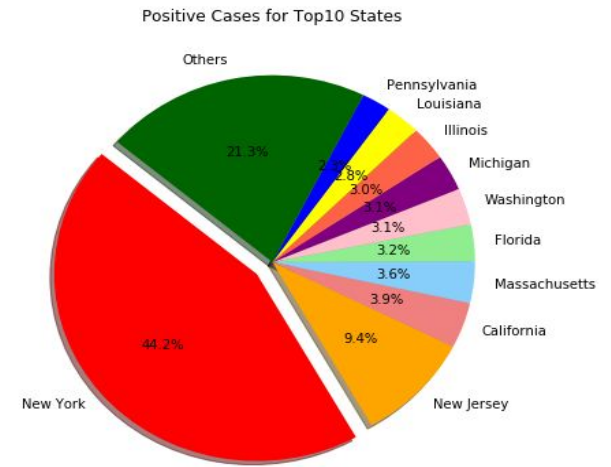
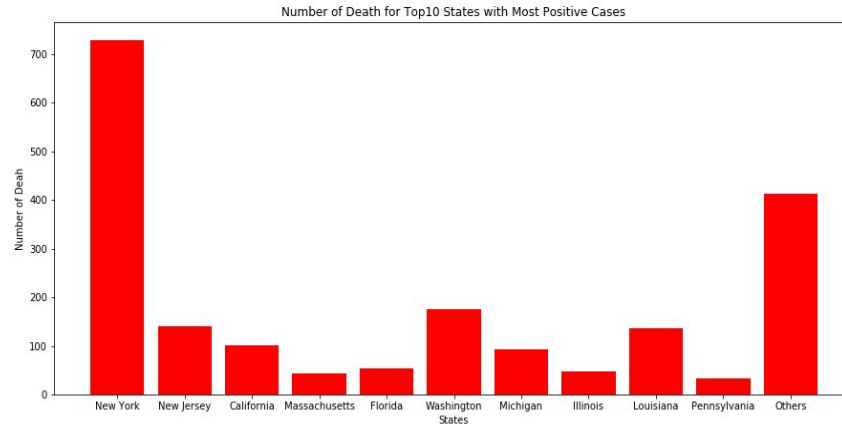
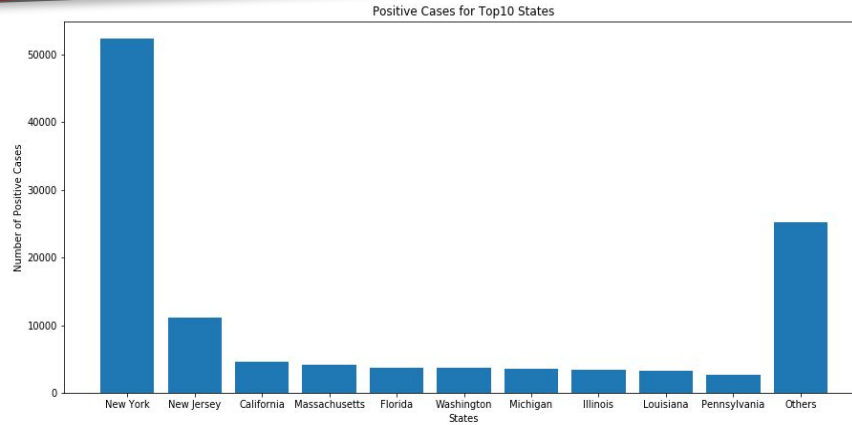




# Time Analysis - US

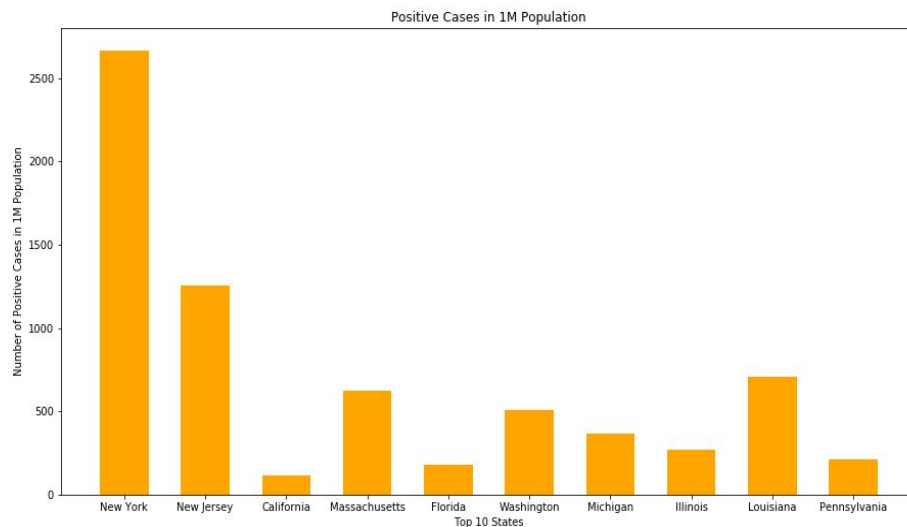
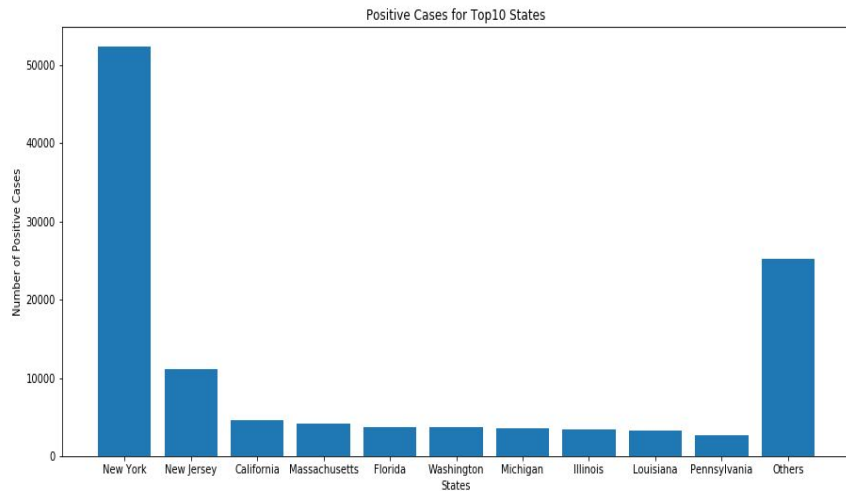


# US States Overview



# US States Overview

## - Positive Cases per 1M



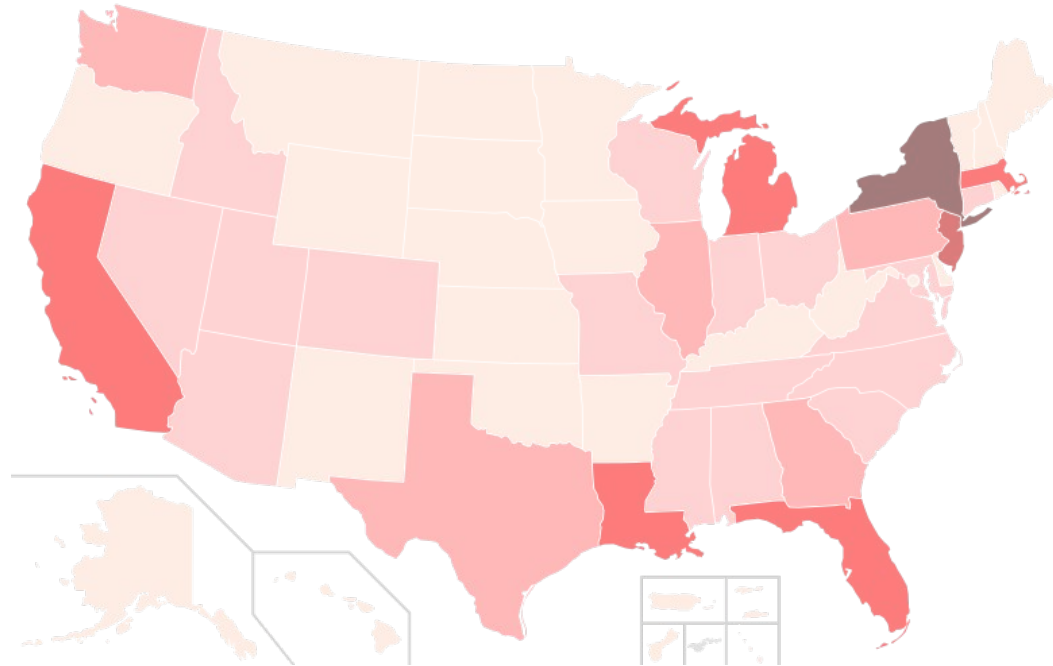
# US States Overview



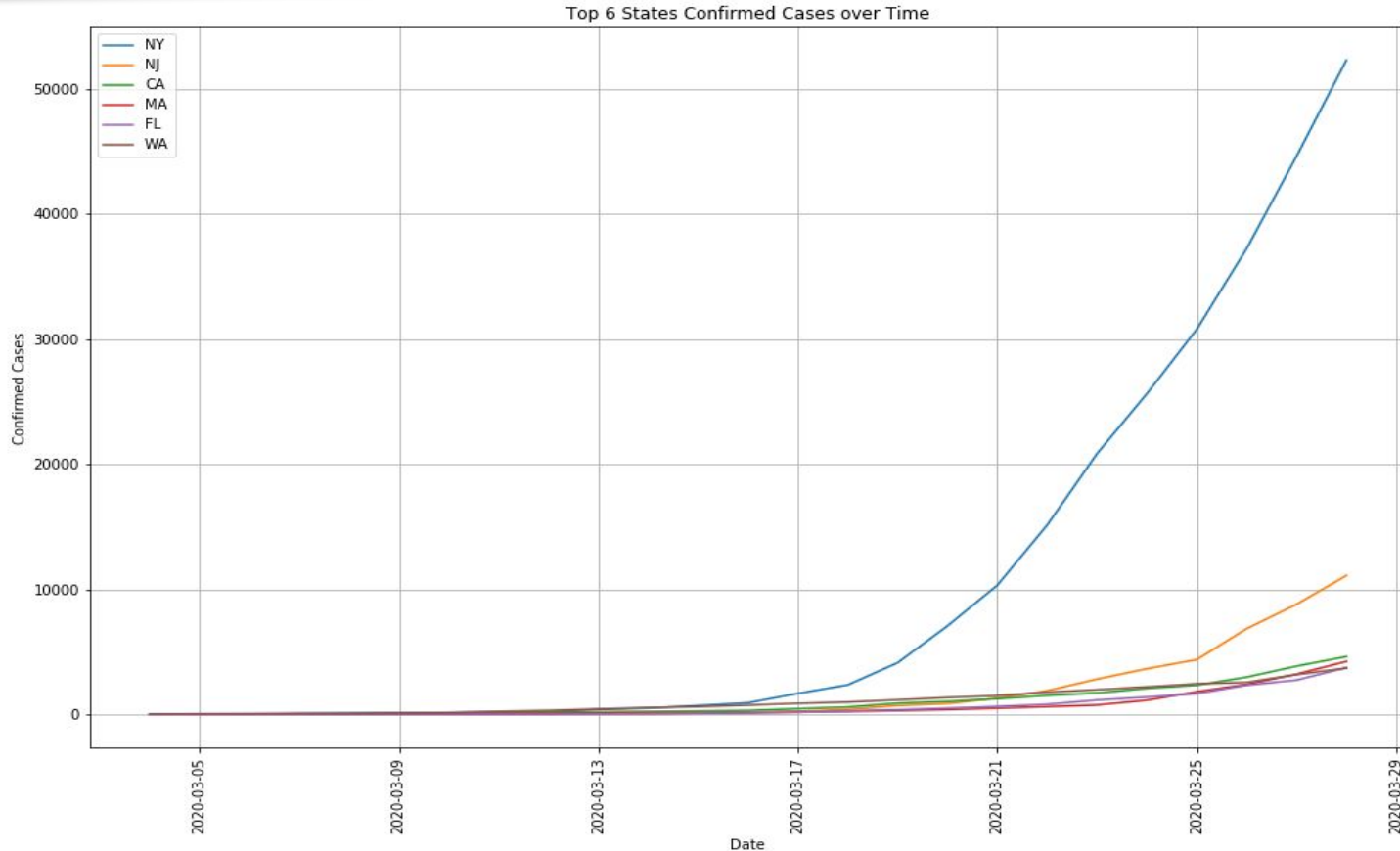
# Hot Spots Analysis

Top 6 States: Ranked by Confirmed Cases (data retrieved up to March 28, 2020)

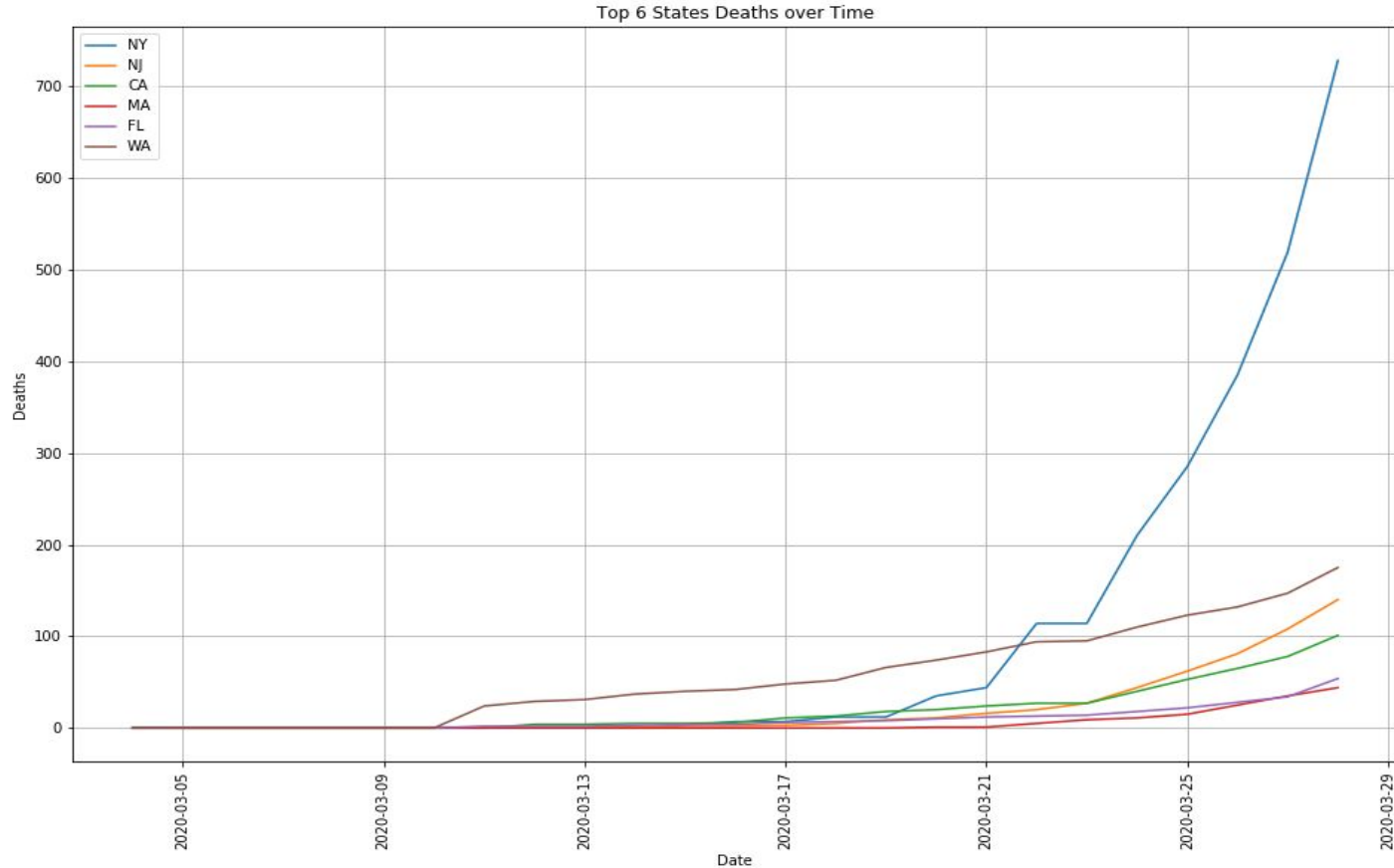
1. New York
2. New Jersey
3. California
4. Massachusetts
5. Florida
6. Washington



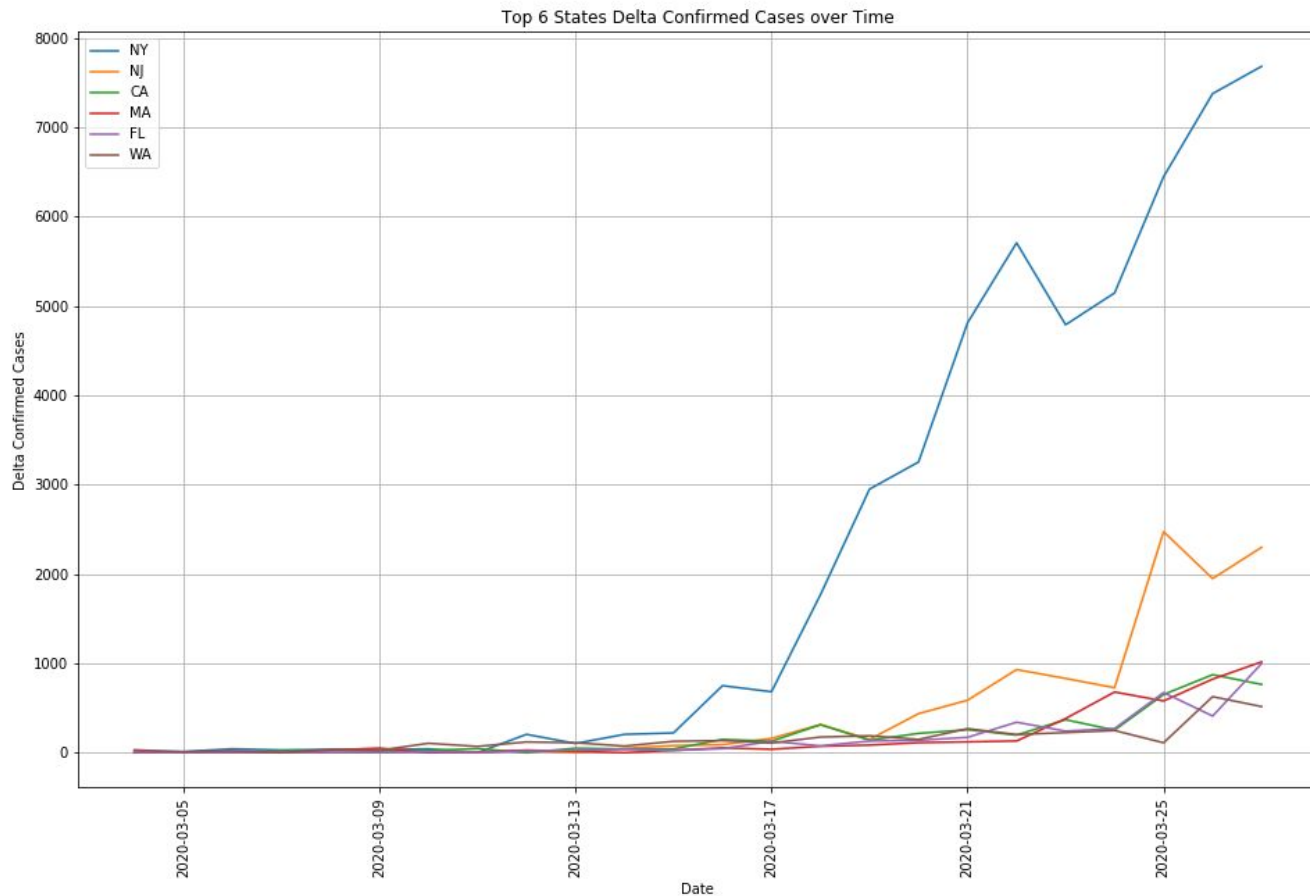
# Confirmed Cases over Time - Top 6 States



# Deaths over Time - Top 6 States

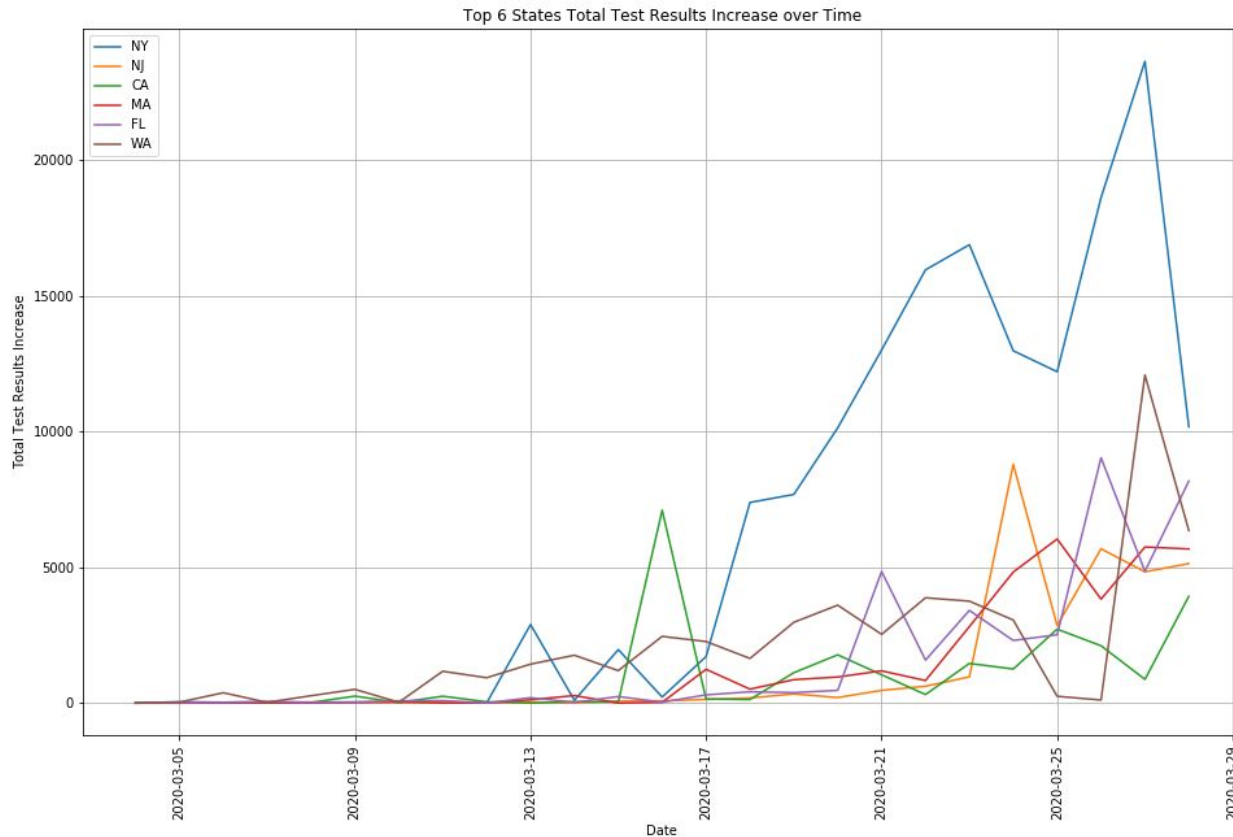


# Change ( $\Delta$ ) of Confirmed Cases over Time - Top 6 States





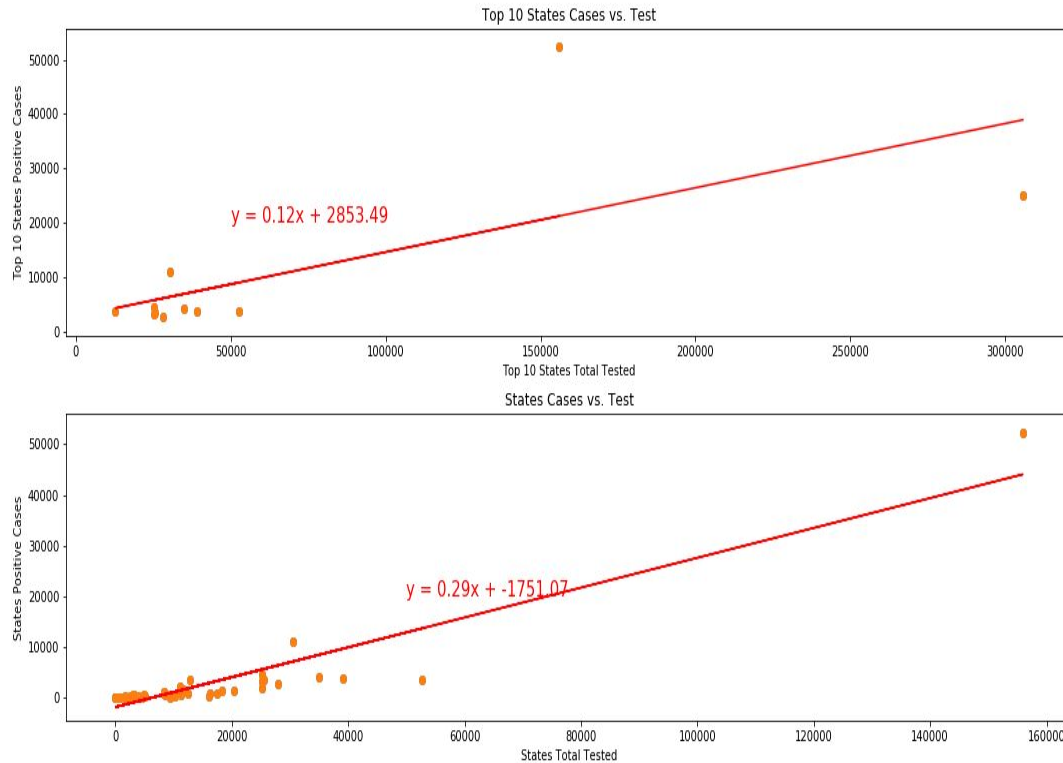
# Total Test Results Increase over Time - Top 6 States



# Positive Cases vs. Total Tests - States



Is there a relationship between number of confirmed cases vs. amount of testing occurring over time?



The correlation between both factors for top 10 states is 0.681  
The correlation between both factors for all states is 0.932

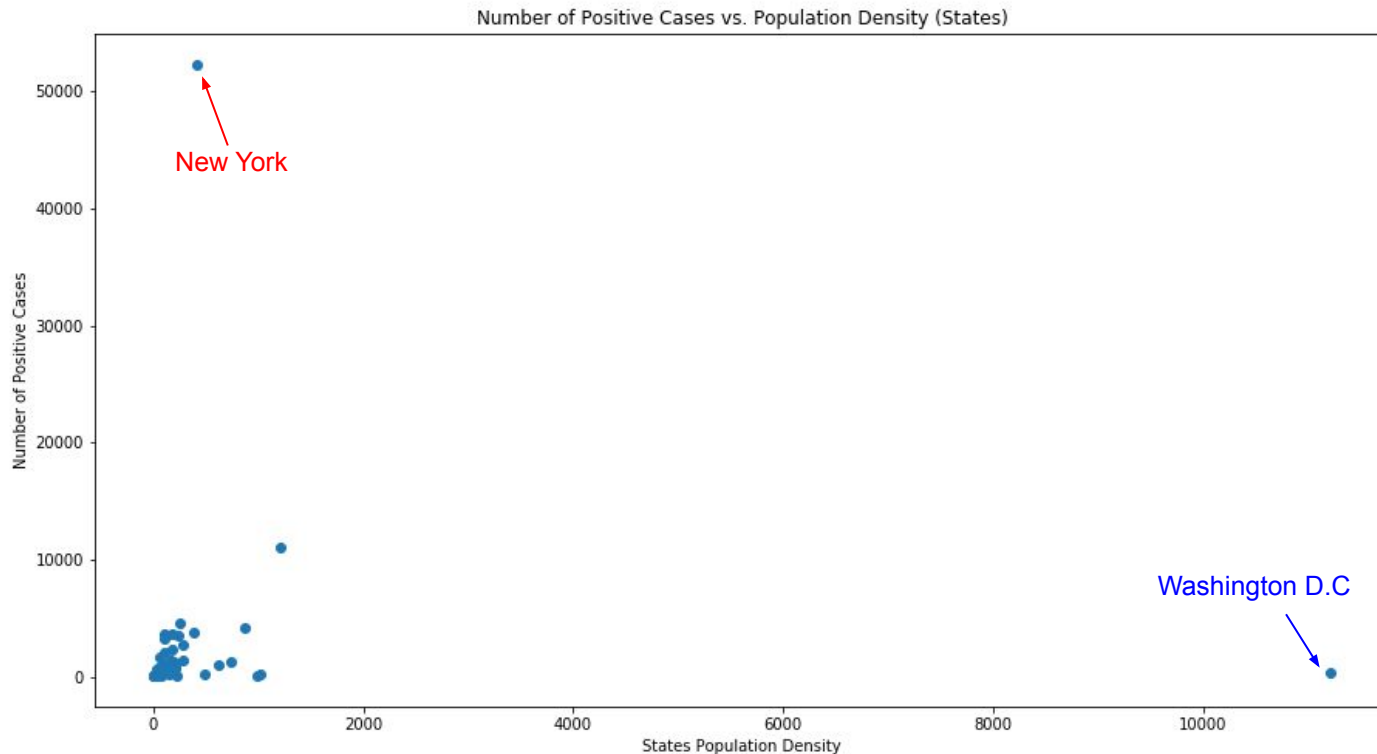


# Positive Cases vs. Pop Density - States



Is there a relationship between number of confirmed cases vs. population density for all states?

The correlation between both factors is 0.00358

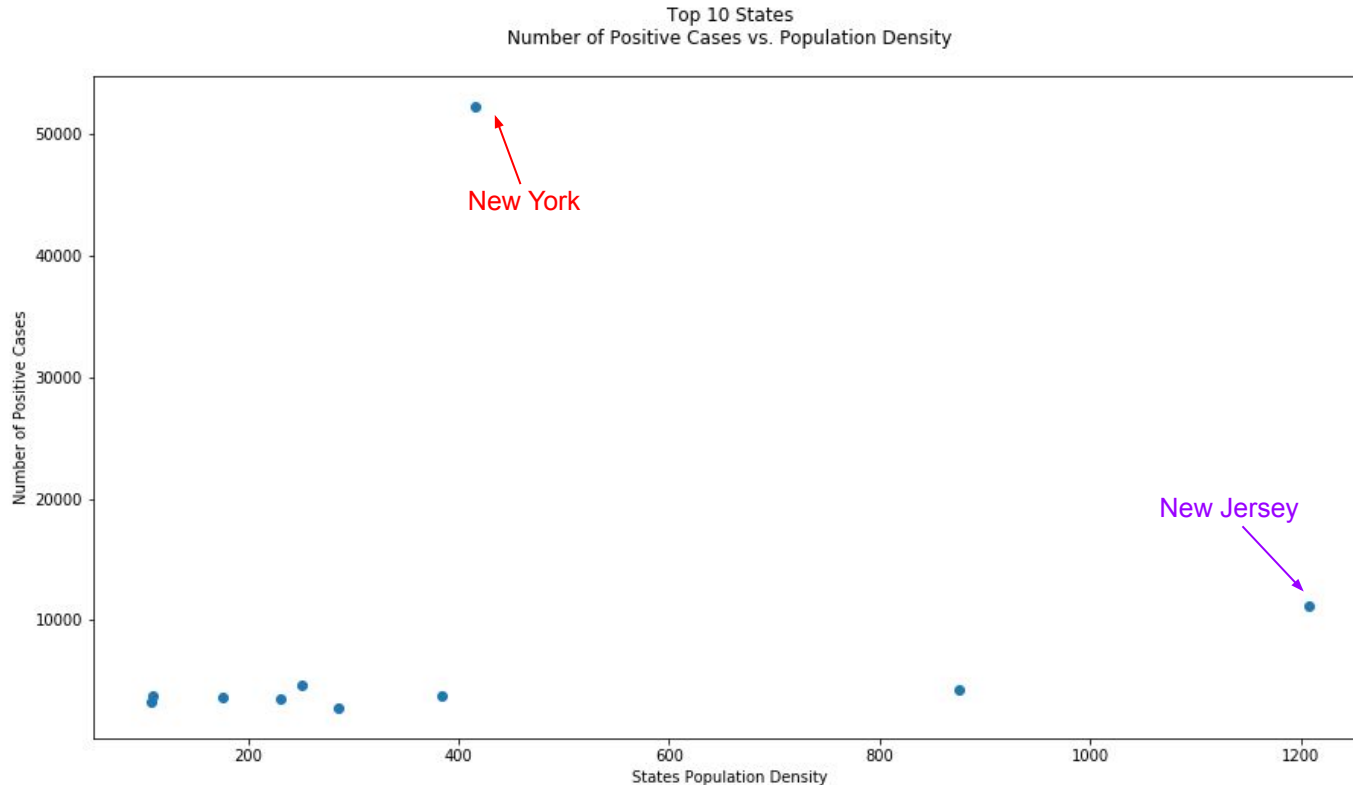


# Positive Cases vs. Pop Density - Top 10 States

The correlation between both factors is 0.13973



Is there a relationship between number of confirmed cases vs. population density for the top 10 states?

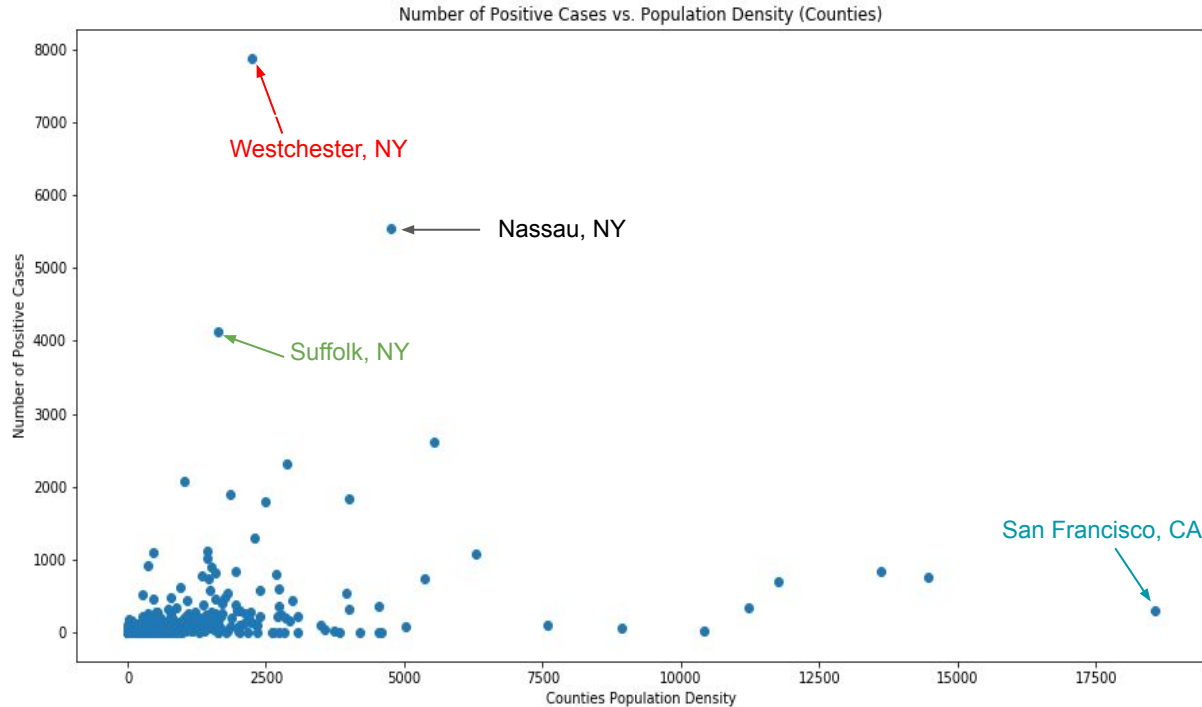


# Positive Cases vs. Pop Density - Counties



Is there a relationship between number of confirmed cases vs. population density on county level?

The correlation between both factors is 0.32498



\*New York City - All cases for the five boroughs of New York City (New York, Kings, Queens, Bronx and Richmond counties) are assigned to a single area called New York City. The city has positive cases of 37066 and population density of 27899 per square miles.

# Summary of Analysis

- As of March 28, 2020, **total number of confirmed cases** is 118,234 in the US
- As of March 28, 2020, **total number of death** is 1,965 in the US
- **Highest mortality rate** recorded on March 12, 2020 at 2.7%
- **Highest hospitalized rate** recorded on March 28, 2020 at 14.1%
- **Top 10 states** number of confirmed cases represents 78.7% of the entire country
- **Top 10 states** number of death represents 79% of the entire country
- There is strong positive correlation between confirmed positive cases vs. total tests.
- There is no strong correlation between confirmed positive cases vs. population density.

**KEEP  
CALM  
AND  
STAY  
AT HOME**