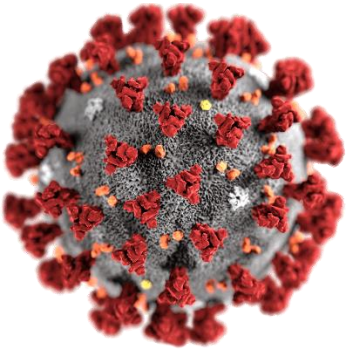
Data Science Capstone Project: August 2020

Tracking the Covid-19 pandemic collaboratively and setup Impacted, Recovery and Testing sites possibilities

1) Introduction/Business Problem

As all we are navigate the ongoing COVID-19 crisis, there are a number of key challenges observed by public health authorities and government to count real time total impacted peoples as well as how many peoples are recovered, in addition to identify new covid-19 testing location based on data science technology.

"The idea of this study is to help public health authorities planning to launch more Covid-19 testing sites to choose the right location and plan to collect recovery counts by collecting historic data about the Covid-19 impacted peoples of each locality and major more preventive actions for spread the covid-19 cases"



COVID-19 case counts are alarming in both their volume and widening geographic scope. There are also concerns about the COVID-19 case counts, recovery rates. Specifically, concerns regarding COVID-19 underreporting are influenced by factors such as lack of access to testing sites. The trajectories show the daily number of confirmed cases. But the widely available data on confirmed cases only becomes meaningful when it can be interpreted in light of how much a country is testing.

2) Downloading and Preparing data

The Downloading and Preparing data involve gather data, discover and assess data, cleanse and validate data, transform and enrich data and store data.

The COVID Tracking testing data available for US states and territories. It has been cited in and used by major public health authorities and community users . The daily number of cases, including states, U.S. territories can be found in below links:

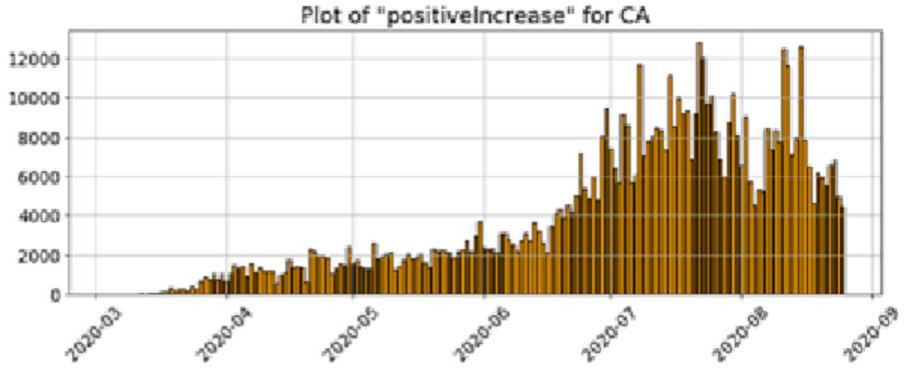
http://covidtracking.com/api/states/daily.csv

This source data need to be cleaning and transforming prior to processing and analysis, additionally using some data science techniques reformatting data (For example - Replacing the NaN), making corrections to data and the combining of data sets to get outcome results like : planning to launch more Covid-19 testing sites to choose the right location and plan to collect recovery counts by collecting historic data about the Covid-19.

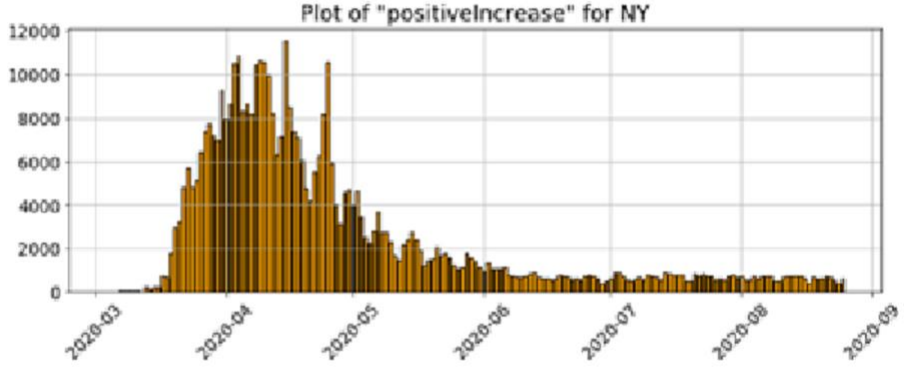
3) Methodology

For this problem and solution I have used iterative system of methods that guides me on the ideal approach to solving problems with data science, through a specified sequence of steps.

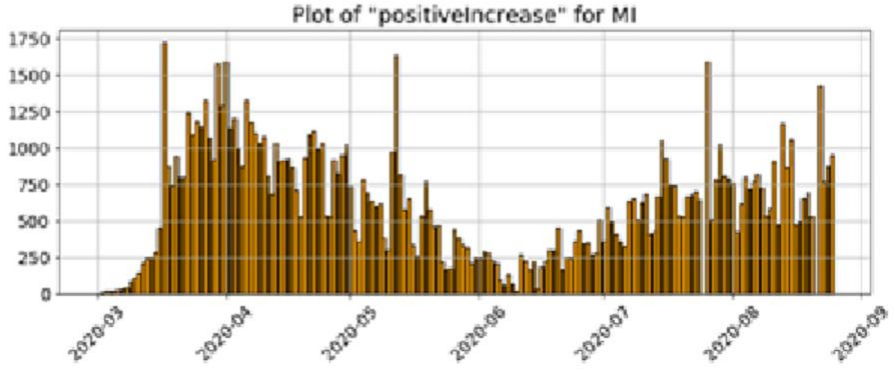
For this report I used a few different maps that could help public health authorities to decide and choose the right location for more testing and plan to collect recovery counts in different states(CA, NY, MI) of USA . I have created visual graph and data presentation based on Covid-19 pandemic location wise data for USA states CA, NY and MI, in order to do that I've used the current Covid-19 pandemic information and data science tools to combined maps for Positive Cases Increase in CA, NY and MI states, Plot for Positive VS Recovered , Total Recovered counts in Last 6 months etc.



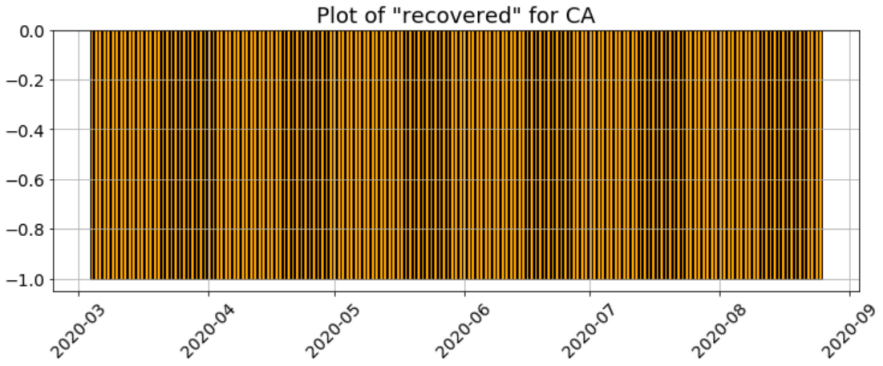
In [163]:
```
plot_var('positiveIncrease','NY')
```



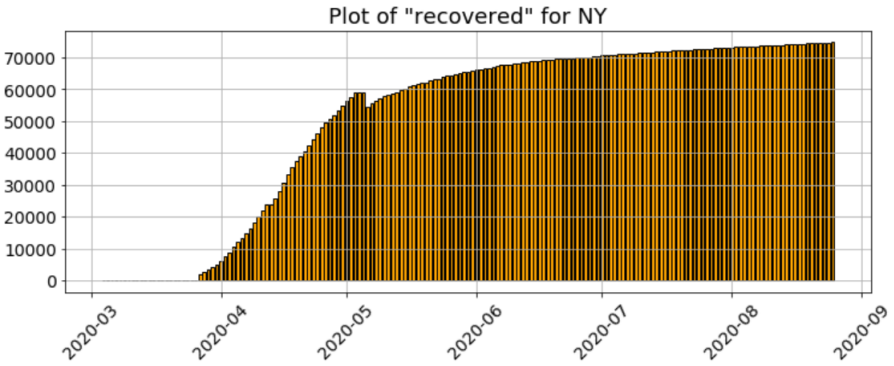
In [165]:
```
plot_var('positiveIncrease','MI')
```



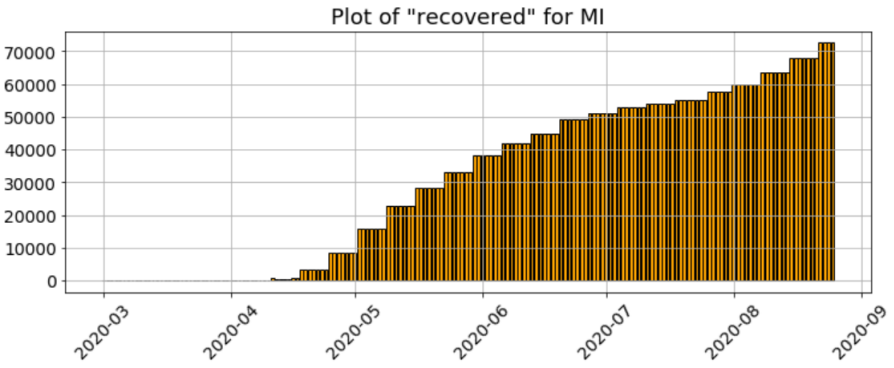
In [170]:
```
plot_var('recovered','CA')
```



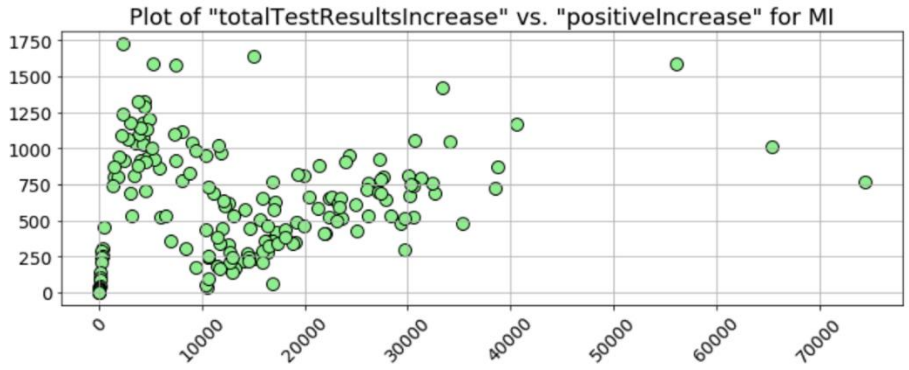
In [171]:
```
plot_var('recovered','NY')
```



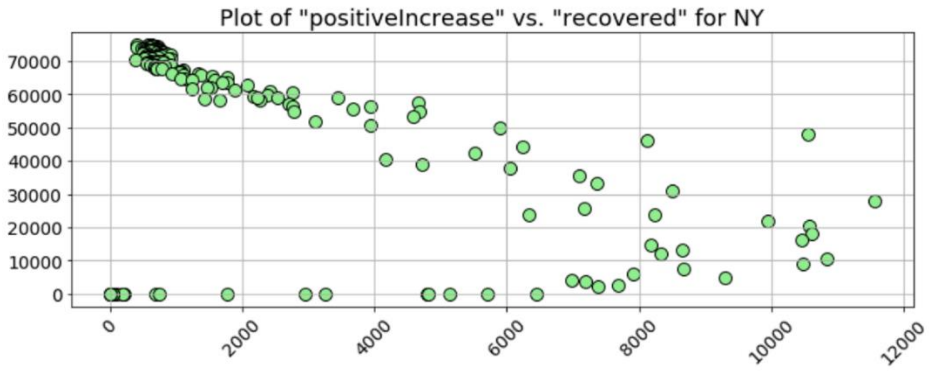
In [172]:
```
plot_var('recovered','MI')
```



In [65]:
```
plot_xy(state='MI')
```

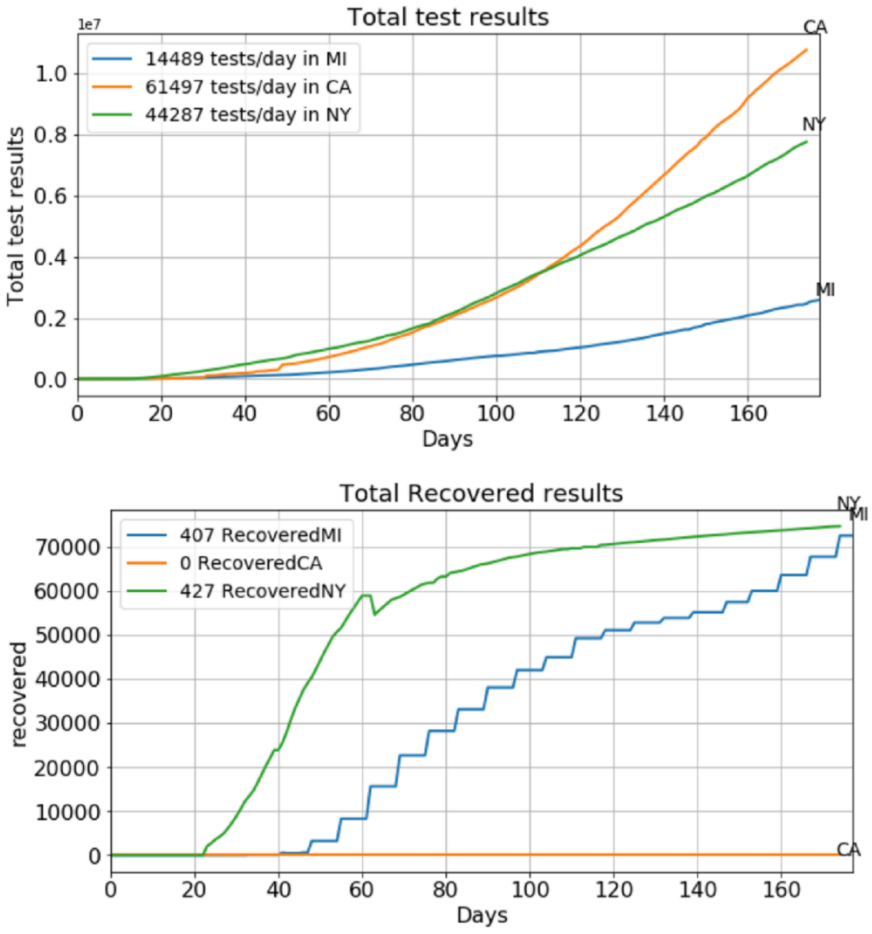


In [173]:
```
plot_xy('positiveIncrease','recovered','NY')
```



4) Results

Comparing the maps we can notice the MI State Test Results are less compare to other state of testing like CA and NY, Recovered rate is higher in NY compare to other state although CA state has no more details for recovered results. Considering Recovery and testing MI and CA states needed more Testing and Recovery plan.



5) Discussion

When I collected Covid-19 datasets, I was observed that more optimized process needs to evaluate for collect recovered data and some of states needed more sites for launch the testing.

6) Conclusion

This report may be helpful for take the decision by public health authorities for launch more Covid-19 testing sites and prioritize collecting recovered plan for particular state. They can easy to compare positive test results counts, and recovered results, however it may not cover all data as reports are based on certain sample data sets and prototype for specific states.