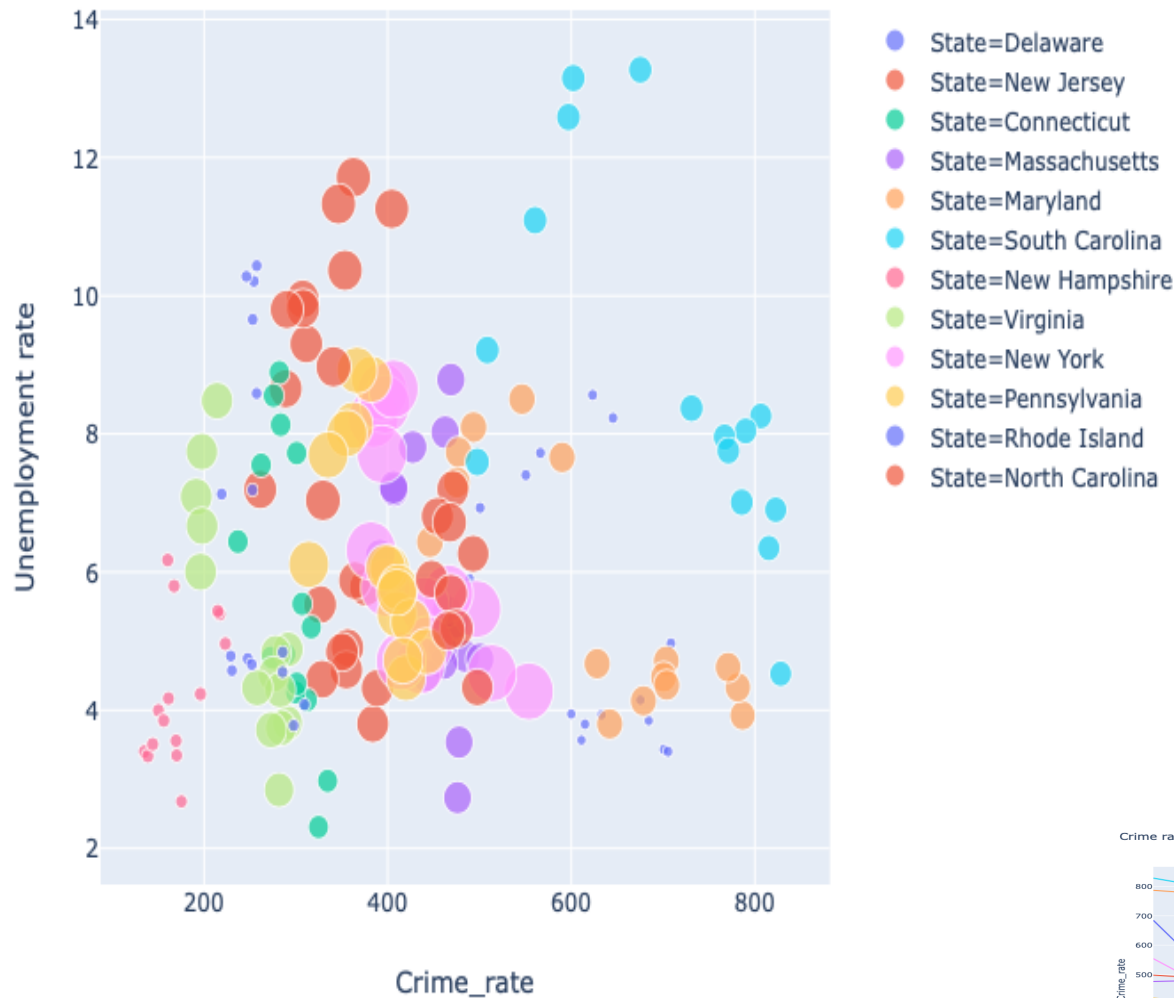
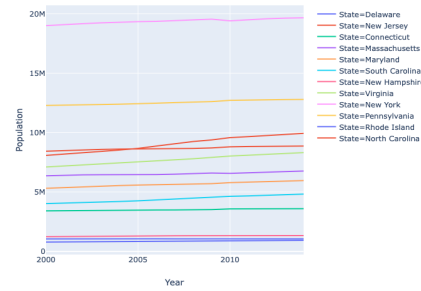


# Is unemployment source of crime?

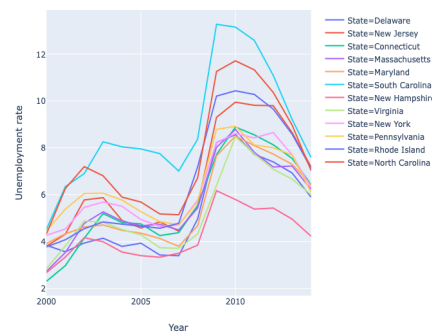
Unemployment rate vs Crime rate of First 13 colonies of USA



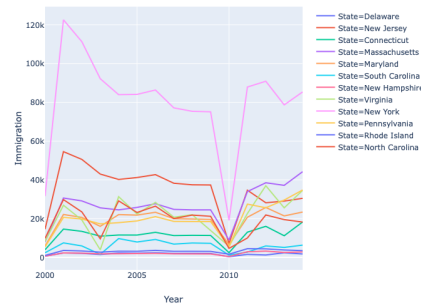
Population of First 13 colonies of USA from 2000 to 2014



Unemployment rate of First 13 Colonies of USA from 2000 to 2014



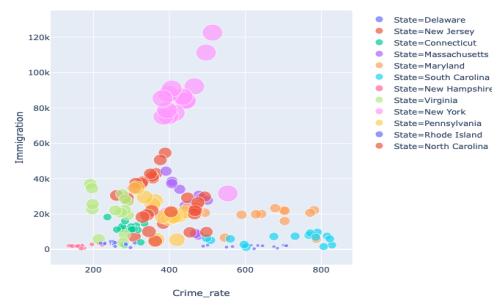
Immigration in First 13 colonies of USA from 2000 to 2014



Crime rate of First 13 Colonies of USA from 2000 to 2014



Immigration VS Crime rate of First 13 Colonies of USA



## LEGEND

The above figure is representation of Unemployment rate VS Crime rate of First 13 Colonies of United States of America with respect to Population from 2000 to 2014. The main dataset is from Kaggle dataset with 885548 number of rows that give the unemployment rate of the state, county and year. I combined four datasets for this project. I got the dataset from Kaggle, UCR data tool and US Census. I was curious about the first 13 colonies of United States since it was one of the questions that was asked by interviewer for my naturalization. I wanted to know how it changed throughout the year. Since the data was large, I sliced the data to the first 13 colonies of USA. Comparing to seaborn and plotly, I found plotly much better since it is interactive visualization and tells us a bit of information while hovering at the graph. I combined the data of unemployment and population so as to find out the relation between them. For the above figure I used plotly for the plot, 'Crime rate' for x axis, 'Unemployment rate' for y axis, 'State' as a color, 'Population' for the size of the plot. Here every component has a purpose in the figure.

## SIGNIFICANCE

From this figure I can tell that the unemployment rate and Crime Rate has little to no correlation. There is a significant jump in unemployment rate due to the recession in 2010. There is also a significant drop in immigration in 2010 possibly due to recession. For most of the states the highest crime rate was at 2000 even though the unemployment rate was all time low during that year. From my analysis, the rate of unemployment has always been in sinusoidal wave pattern whereas the crime rate has always been decreasing as the year increase except "New Hampshire". I found it interesting that even though the population of "New York" has been the highest, the unemployment rate remained average out of these states.

## FINDINGS

After looking at the diagram I found that my first data is missing entire information/values about the State "Georgia". Therefore, there is no information about the state "Georgia" even though it is one of the first 13 colonies of USA. From my analysis, the crime rate and unemployment rate have little correlation. The crime rate and the immigration have some positive correlation. The unemployment rate increases significantly in year 2010 because of the recession. Crime rate is decreasing every year maybe due to increase and improvement of technologies.

## DATA

For the unemployment rate I used the dataset from this link: <https://www.kaggle.com/jayrav13/unemployment-by-county-us>. This is a dataset that 'Jay Ravaliya' built by scraping the United States Department of Labor's Bureau of Labor Statistics. The dataset has a data of unemployment rate from all of the county of United States. For the population dataset I used this link: [https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html#par\\_textimage](https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html#par_textimage) and <https://www.census.gov/data/tables/time-series/demo/popest/intercensal-2000-2010-state.html>. For the crime rate dataset I used this link: <https://www.ucrdatatool.gov/Search/Crime/State/StateCrime.cfm>.

## METHOD

I wanted to work with unemployment rate of USA and therefore I chose the unemployment dataset as my main dataset. After I worked with the unemployment dataset, I merged the population of USA dataset from 2010 to 2018 that I got from US census website. I inner merged it with state and the year. So, with these two datasets I w

as only getting few information. Therefore, I merged more population dataset from 2000 to 2009 to the existing merged dataset. After this my results were completely different than that of only one dataset of population. To find out the relation of crime and unemployment, I merged the crime rate dataset to the merged dataset. Since my dataset was very large, I sliced it down to focus only on the first 13 colonies of United States of America. The unemployment dataset has the data from the county, so I replaced the mean value of unemployment rate to that of state. My final dataset had 696 rows and 6 columns. I used the plotly package since it was giving me the best visualization as well as some tools to work with the data. It also has the interactive visualization so as to find out the exact information of the data in the figure.

## **DISCUSSION**

When I started to work with the unemployment dataset, I was curious about the causes of the unemployment of the country. Therefore I started to merge a lot of other dataset in hope of getting some insight out of these dataset. I was confident that crime and immigration was some of the causes of the unemployment and therefore I used those dataset. But on the contrary my results showed that it might be different. The results showed that there was little to no correlation between crime rate and unemployment. Therefore I can't say that unemployment is the source of crime. I understand that the actual cause cannot be just one thing. It depends on multiple factors and most of the time it changes every year. We live in a dynamic world so the way we live in changes all the time and so the unemployment.