More information

Emma Ning and Kristy Shi

12/19/2018

Table of Contents

# Problem Statement

What is the scale of the 2016-17 increase in crime violence and what are the factors that influence NYS crime rate in 2017?

# Description of the data

## Index, Violent, Property, and Firearm Rates By County: Beginning 1990

This data is provided by the New York State Division of Criminal Justice Services and is based on crime reports from more than 500 New York State and sheriffs’ departments [[1]](#footnote-24). This annually updated dataset collects information on seven crimes classified as Index crimes. The two category of crimes are **violent crimes** (murder/non-negligent manslaughter, forcible rape, robbery, and aggravated assault) and **property crimes** (burglary, larceny, and motor vehicle theft).

## Comparison of 2000 and 2010 Populations by Race and Ethnicity

This dataset is obtained from the New York State Department of Labor, which sources its information from the 2000 and 2010 Census. The data compares the New York State, Regional, and County Populations by Race and Ethnicity (in this case, percentage of White, Black, Asian, Hispanic).

## Quarterly Census of Employment and Wages Annual Data: Beginning 2000

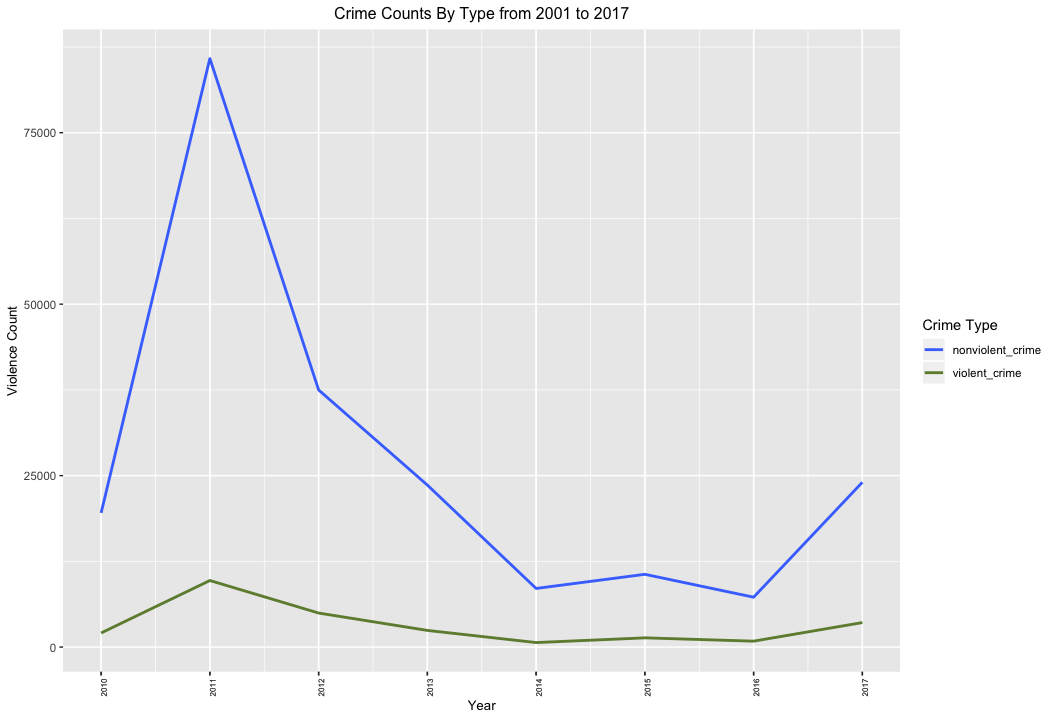
This data is collected from the Quarterly Census of Employment and Wages program, a cooperative program with the U.S. Bureau of Labor Statistics, which collects employment and wage data. This data provides a virtual census of employees and their wages as well as the most complete universe of employment and wage data, by industry, at the State, regional and county levels [[2]](#footnote-27).

# Scale of Increase: 2001-17 In-Depth

To further understand the increase of crime rate and to better visualize the seemingly linear decrease in crime rate from 1990 to 2017 (seen in the coverpage), we graphed crime count from 2001 to 2017.

# Crime Time-series

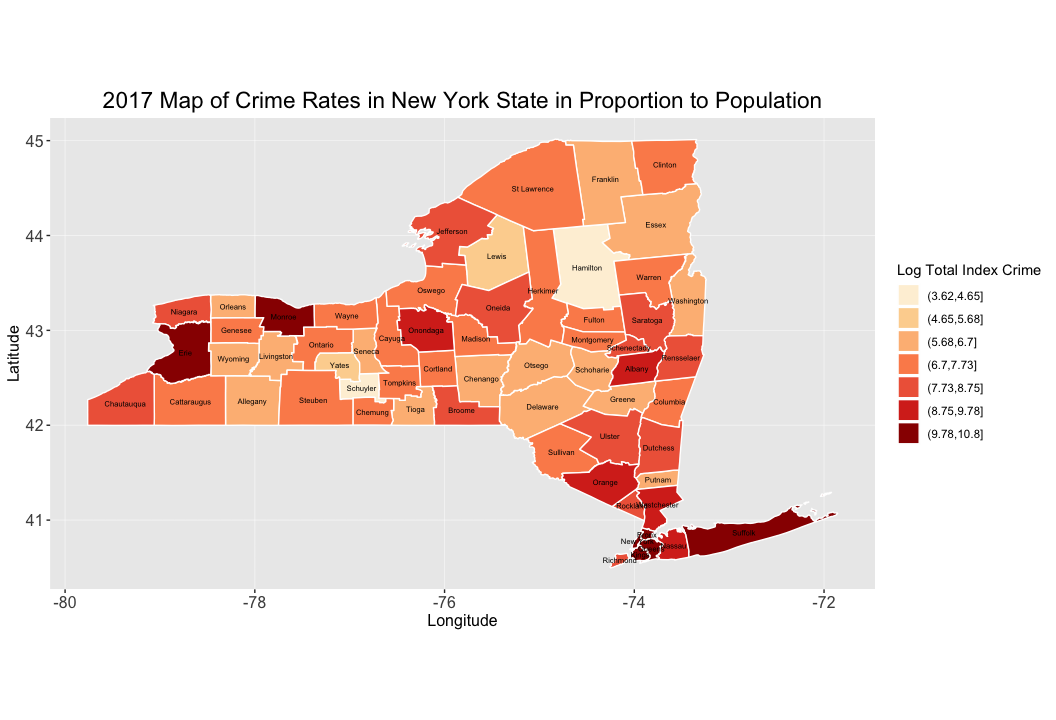
# violent and non-violent crimes  
crime\_types <- violence %>%   
 filter(Year==c(2010:2017)) %>%   
 group\_by(Year) %>%   
 drop\_na() %>%   
 mutate(violent\_crime=sum(Violent.Count),  
 nonviolent\_crime=sum(Property.Count)) %>%   
 distinct(violent\_crime, nonviolent\_crime) %>%   
 gather(key="crime\_type", value="count", violent\_crime:nonviolent\_crime)  
  
ggplot(crime\_types, aes(x=Year, y=count)) +   
 geom\_line(aes(color=crime\_type), size=1)+  
 labs(x="Year", y="Violence Count",   
 title="Crime Counts By Type from 2001 to 2017", color="Crime Type")+  
 scale\_x\_continuous(breaks=(crime\_types$Year))+  
 theme(plot.title = element\_text(size = 12, hjust = 0.5),  
 axis.title.x = element\_text(size = 10),  
 axis.text.x = element\_text(angle = 90, size=6),  
 axis.title.y = element\_text(size = 10))+  
 scale\_color\_manual(values = c("royalblue1", "darkolivegreen4"))



This time series line graph shows that the 2016-17 increase in crime rates seem to be less significant than initially expected. In the overall 1990 picture, peaks in crime rate were counterbalanced with declines; as a result, the upward slope in 2016-17 was especially visible. This means that surveillance and increased defense spending did not reduce crime violence to the extent that there was no increase in crime during 1990 to 2016 in NYS.

# Map of Crime Rates per Capita

# some data wrangling  
# for year 2017  
violence\_county <- violence %>%   
 select(County, Index.Rate, Year) %>%   
 filter(Year==2017) %>%   
 right\_join(newyork\_pop, by="County") %>%   
 # "num\_unit" stands for the number of units (100,000)  
 mutate(num\_unit=estimate/100000,  
 # "total\_index" stands for the total index crime for that county  
 total\_index=Index.Rate\*num\_unit,  
 log\_index=round(log(total\_index), 2)) %>%   
 # make it discrete  
 mutate(index\_discrete=cut(log\_index, breaks=7))  
  
# plot using sf  
ggplot() +  
 geom\_sf(data = violence\_county, aes(fill = index\_discrete), size = 0.5, color = "white") +  
 geom\_text(data = violence\_county, aes(x=lon, y=lat, label = County), color = "black", size = 2) +  
 scale\_fill\_brewer(type = "seq", palette = "OrRd", direction = 1) +   
 labs(title = "2017 Map of Crime Rates in New York State in Proportion to Population",  
 fill = "Log Total Index Crime",  
 x = "Longitude", y = "Latitude")+  
 theme(plot.title = element\_text(size = 17, hjust = 0.5),  
 axis.text.x = element\_text(size = 12), axis.title.x = element\_text(size = 12),  
 axis.text.y = element\_text(size = 12), axis.title.y = element\_text(size = 12))

 We graphed the Index Rate for the year 2017 to see counties with the most crime per capita.

**White(1)= lowest ; Dark-Red(7)= highest**

# Possible Crime Influences- 2017

## Income and Crime Rates

# Mean wage by county, year 2017  
income\_short <- income %>%   
 filter(Area.Type=="County", Year=="2017") %>%   
 select(-NAICS, -NAICS.Title) %>%   
 group\_by(Area) %>%   
 mutate(county\_mean\_wage=mean(Annual.Average.Salary)) %>%   
 distinct(Area, county\_mean\_wage) %>%   
 arrange(county\_mean\_wage)  
# top 10 lowest income counties  
income\_short1 <- income\_short[1:10, ]  
# rename vars for display  
# 10 lowest wage county  
income\_short1 <- income\_short1 %>%   
 rename(County=Area, Mean\_wage=county\_mean\_wage)  
# 10 highest wage county  
income\_short2 <- income\_short[53:62, ]  
income\_short2 <- income\_short2 %>%   
 rename(County=Area, Mean\_wage=county\_mean\_wage)  
# side-by-side  
income\_comparison <- cbind(income\_short1, income\_short2)  
income\_comparison <- income\_comparison %>%   
 rename(County\_low=County, County\_high=County1,  
 Low\_mean\_wage=Mean\_wage, Top\_mean\_wage=Mean\_wage1)  
income\_comparison

## # A tibble: 10 x 4  
## # Groups: County\_low [10]  
## County\_low Low\_mean\_wage County\_high Top\_mean\_wage  
## <fct> <dbl> <fct> <dbl>  
## 1 Hamilton County 28950. Kings County 50589.  
## 2 Schuyler County 31505. Bronx County 51139.  
## 3 Yates County 32289. Monroe County 51895.  
## 4 Allegany County 33870. Albany County 53947.  
## 5 Tioga County 34053. Rockland County 54103.  
## 6 Lewis County 34188. Queens County 54484.  
## 7 Oswego County 34511. Suffolk County 63133.  
## 8 Delaware County 34710 Nassau County 63333.  
## 9 Livingston County 35020. Westchester County 72193.  
## 10 Orleans County 35077. New York County 95290.

The table depicts counties with the 10 highest and lowest mean wage in 2017. Hamilton and Schuyler, the lightest colored counties on the map, have the least mean wages. The lowest waged 10 counties range from white(1) to orange(4), and the highest from orange(4) to dark-red(7).

The data is surprising as people usually associate poorer regions with more crime, but the data seems to demonstrates that the higher the wage, the higher the crime rates per capita. Does this correlation prove causation, or is this association a mere coincidence?

## Population and Crime Rates

# year 2017  
pop <- violence %>%   
 filter(Year=="2017") %>%   
 select(County, Population) %>%   
 arrange(desc(Population))  
  
pop1 <- pop[53:62, ]  
pop1 <- pop1 %>%   
 rename(Low\_population=Population)  
pop2 <- pop[1:10, ]  
pop2 <- pop2 %>%   
 rename(Top\_population=Population)  
pop\_comparison <- cbind(pop1, pop2)  
row.names(pop\_comparison) <- NULL  
pop\_comparison

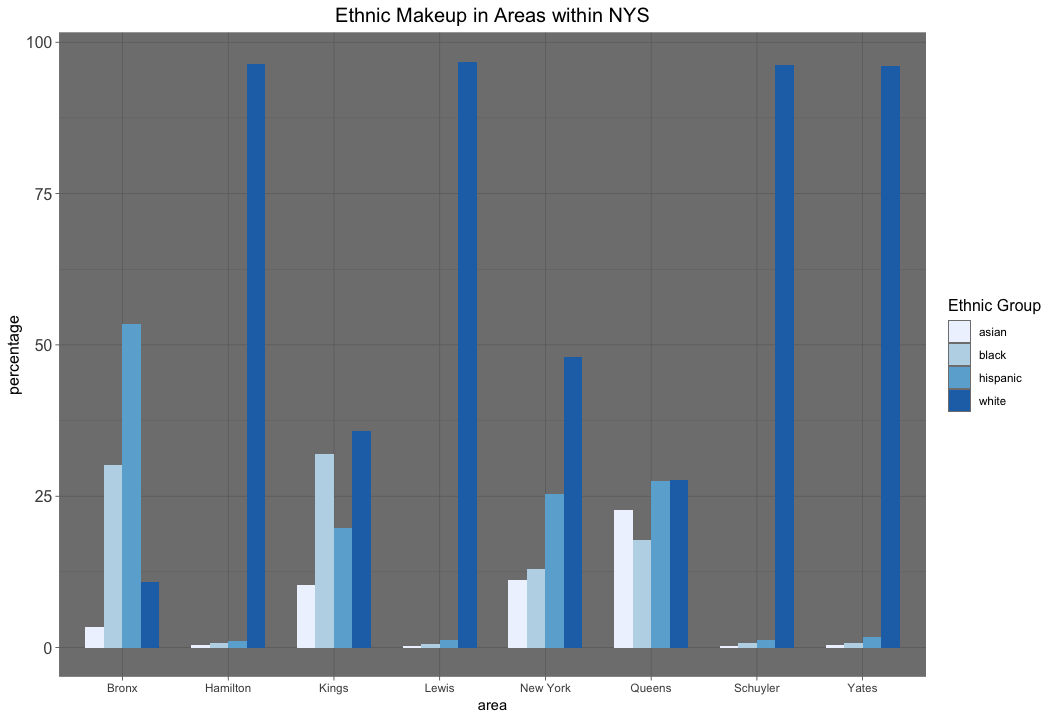
## County Low\_population County Top\_population  
## 1 Delaware 44468 Kings 2646816  
## 2 Orleans 41194 Queens 2356841  
## 3 Wyoming 40666 New York 1663498  
## 4 Essex 36491 Suffolk 1495619  
## 5 Seneca 34779 Bronx 1470074  
## 6 Schoharie 31166 Nassau 1367901  
## 7 Lewis 26893 Westchester 980819  
## 8 Yates 24910 Erie 923447  
## 9 Schuyler 18106 Monroe 749892  
## 10 Hamilton 4505 Richmond 479104

In 2017, 8 out of 10 counties with the highest population had the highest mean wage per worker, and 6 out of 10 lowest population had the lowest mean wage. The two counties, Erie and Richmond are dark-red(7) and dark-orange(5) respectively. Hamilton and Schuyler have the lowest population. Note that the crime range for highest population is dark-orange(5) to dark-red(7).

Population seems to have a closer correlation to danger than Income since the color gradient for crime rate per capita is more compact (5-7 as opposed to 4-7).

## Ethnicity and Crime Rates

# we picked out several dangerous regions and look at the ethnic make-up  
# from 2010 census  
ethnicity1 <- ethnicity %>%   
 mutate(Area=as.factor(X)) %>%   
 filter(Area %in% c("Hamilton", "Schuyler", "Lewis", "Yates",  
 "Kings", "Queens", "New York", "Bronx")) %>%   
 select(Area, Population, X2010:X.9) %>%   
 rename(white=X2010, black=X.7, asian=X.8, hispanic=X.9) %>%   
 gather(key="group", value="percentage", white:hispanic) %>%   
 mutate(percentage=parse\_number(percentage))  
  
  
ggplot(data=ethnicity1, aes(x=Area, y=percentage, fill=group))+  
 geom\_bar(position="dodge", stat="identity", width=0.7)+  
 scale\_fill\_brewer()+theme\_dark()+  
 labs(title="Ethnic Makeup in Areas within NYS",  
 fill="Ethnic Group",  
 x="area", y="percentage")+  
 theme(plot.title=element\_text(size=15, hjust=0.5),  
 axis.text.y=element\_text(size=12), axis.title.y=element\_text(size=12),  
 legend.title=element\_text(size=12))



# Conclusion

The scale of the 2016-17 increase in NYS crime violence is negligible in closer examination. Wage, population and crime rate per capita seem to have a correlation: the higher the earnings and the higher the county’s population size, the higher the crime per capita. The correlation between wage and population is self-evident; more populated counties provide higher quantities and varieties of job opportunities. A possible explanation for the correlation between population and crime rates could be that job opportunities vary (thus, inequality), or that denser populations lead to increased competition and a higher cost of living (thus, cost of living negates wages).

1. “Index, Violent, Property, and Firearm Rates By County: Beginning 1990.” State University of New York (SUNY) Campus Locations with Websites, Enrollment and Select Program Offerings | Open Data NY, DATA.NY.GOV, 25 Nov. 2018, data.ny.gov/Public-Safety/Index-Violent-Property-and-Firearm-Rates-By-County/34dd-6g2j. [↑](#footnote-ref-24)
2. “Quarterly Census of Employment and Wages Annual Data: Beginning 2000.” State University of New York (SUNY) Campus Locations with Websites, Enrollment and Select Program Offerings | Open Data NY, DATA.NY.GOV, 14 Nov. 2018, data.ny.gov/Economic-Development/Quarterly-Census-of-Employment-and-Wages-Annual-Da/shc7-xcbw. [↑](#footnote-ref-27)