Final Project:

American Crime and Incarceration

PS I Love Yue: Oliver, Karam, Steven, Emely

Introduction

The United States currently has the highest prison population in the world at around 2,094,000. This situation can be traced back to the 1970s, when politicians like Richard Nixon, Ronald Reagan, and more began the War on Drugs and began promoting "tough on crime" agendas. All this is despite scholars of the time believing prisons might have been abolished entirely, as crime and imprisonment rates were at an all time low. Mass incarceration has resulted in many racial disparities— disproportionately affecting Black men in particular. These racial disparities are functionally seen by many as modern-day Jim Crow laws which were outlawed in the 1960s. Rates of imprisonment have continued to increase under the guise of "public safety." However, the reality is that the US's crime rates are comparable to countries with similar economies, yet the rate of incarceration is significantly higher. The US has only 5% of the world's population, and yet it houses 25% of the world's incarcerated persons. The racial, social, and economic implications of this statistic are what we intend to explore in this project.

With this context and data set, we are hoping to answer a variety of questions such as the following :

1) Is there a relationship between proportion of population incarcerated and party affiliation in American states?
2) How does policy impact incarceration? This will be explored through the lens of changing definition of rape
3) How do trends in different crimes compare to each other? Do they move together or do they change individually?
4) How does crime differ for different regions of the US?

Though it is difficult to point to a single question that we may be able to answer using this dataset, we are confident that through thorough analysis of the Crime and Incarceration Data Set, we will be able to critically analyze the rhetoric that is often used to defend the US' prison system and investigate common misconceptions regarding the root causes of crime and incarceration. We will likely have to do research outside of this data set to create a strong narrative depending on which direction our project takes us. We have found an additional dataset which includes census data that could be used to find correlations between state and prison demographics. The exact trend or correlation that we will be investigating is not perfectly clear yet, research and analysis of this dataset will reveal interesting information about the US prison system.

Data Description

"The Bureau of Justice Statistics administers the National Prisoners Statistics Program (NPS), an annual data collection effort that began in response to a 1926 congressional mandate. The Uniform Crime Report (UCR) has served as the FBI's primary national data collection tool since a 1930 congressional mandate directed the Attorney General to 'acquire, collect, classify, and preserve identification, criminal identification, crime, and other records.' The FBI collects this information voluntarily submitted by local, state, and federal law enforcement agencies." The user who updated this dataset used the raw data from the NPS and wrangled it in python to create the variables and observations. Within the dataset there are 816 rows and 17 variables.

The variables included in the set are US state, whether or not the data for that entry includes jails, the year, prisoner count on December 31st, whether or not the state in the data entry changed their system for reporting crime in comparison to previous years, whether or not crime totals are estimated, and numbers of: total state population, violent crimes, murders or manslaughters, rapes (using the old definition), rapes (using the new definition), robberies, aggravated assaults, property crime, burglaries, larceny, and vehicle theft.

Methodology

The following variables, from the original dataset, were used to address the research questions above:
* **year**:
* **prisoner_count**:
* **state_population**:
* **violent_crime_total**:
* **murder_manslaughter**:
* **rape_legacy**:
* **rape_revised**:
* **robbery**:
* **agg_assault**:
* **property_crime_total**:
* **burglary**:
* **larceny**:
* **vehicle_theft**:
* **total_prisoner**:
* **total_violent_crime**:
* **total_crime**:

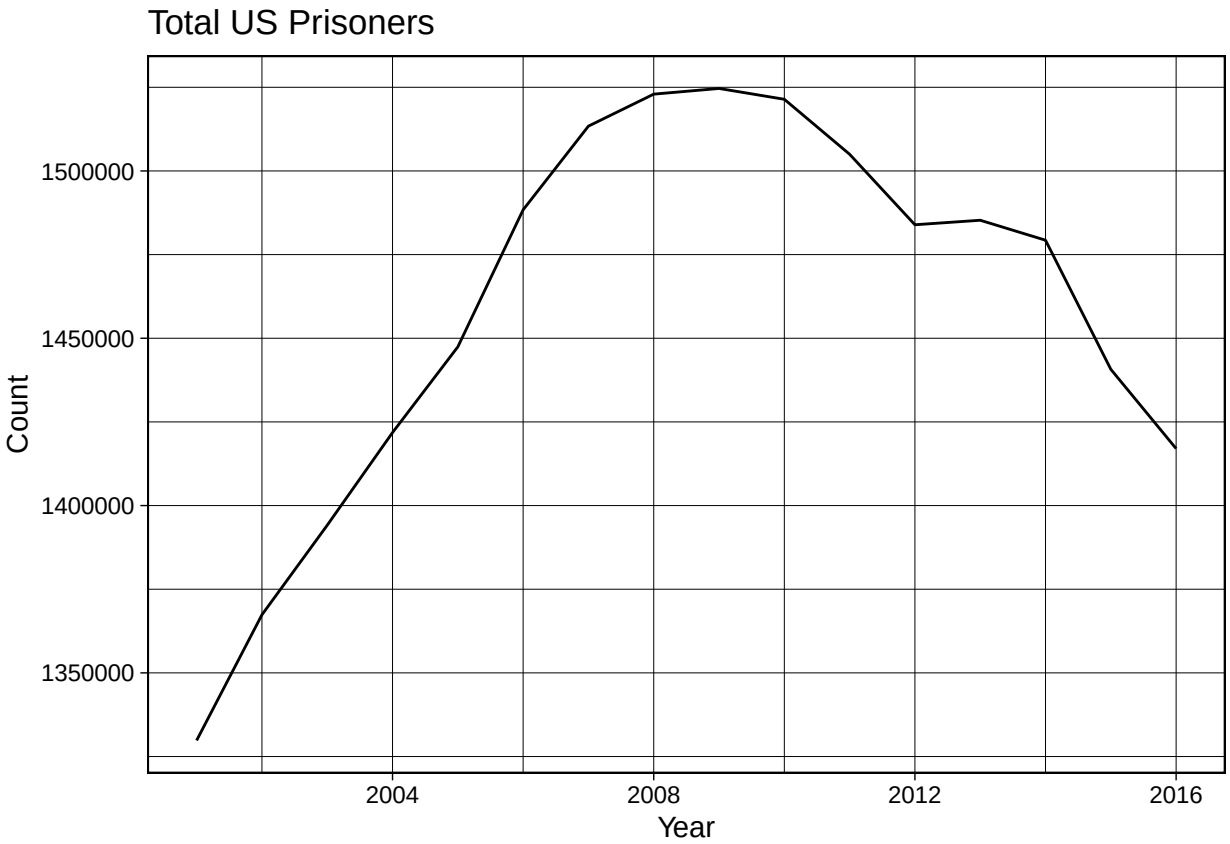
```
# install.packages("devtools")
# devtools::install_github("UrbanInstitute/urbnmapr")
```

```
# library(tidyverse)
# library(urbnmapr)
#
# states_sf <- get_urbn_map("states", sf = TRUE)
#
# states_sf %>%
#   ggplot(aes()) +
#   geom_sf(fill = "grey", color = "#ffffff")
```

Research Question #1

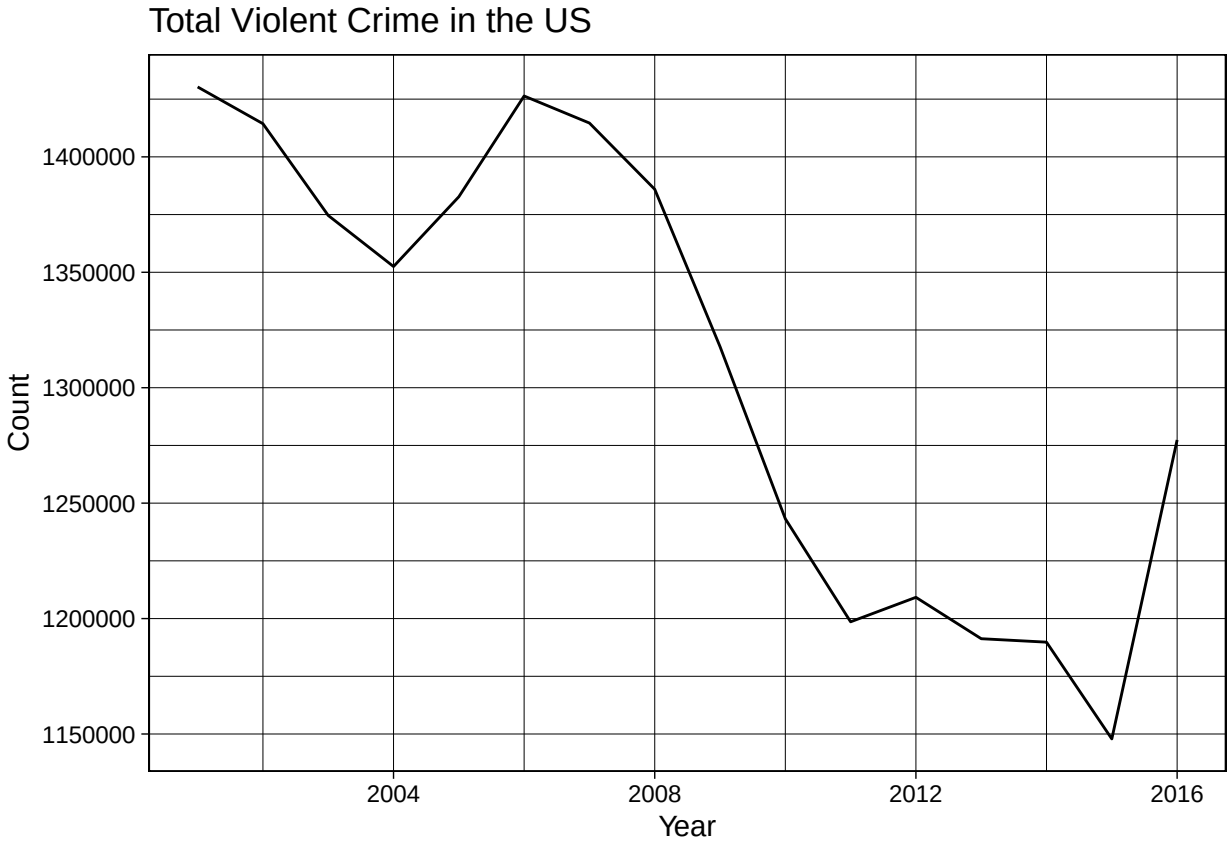
Is there a relationship between proportion of population incarcerated and party affiliation in American states?

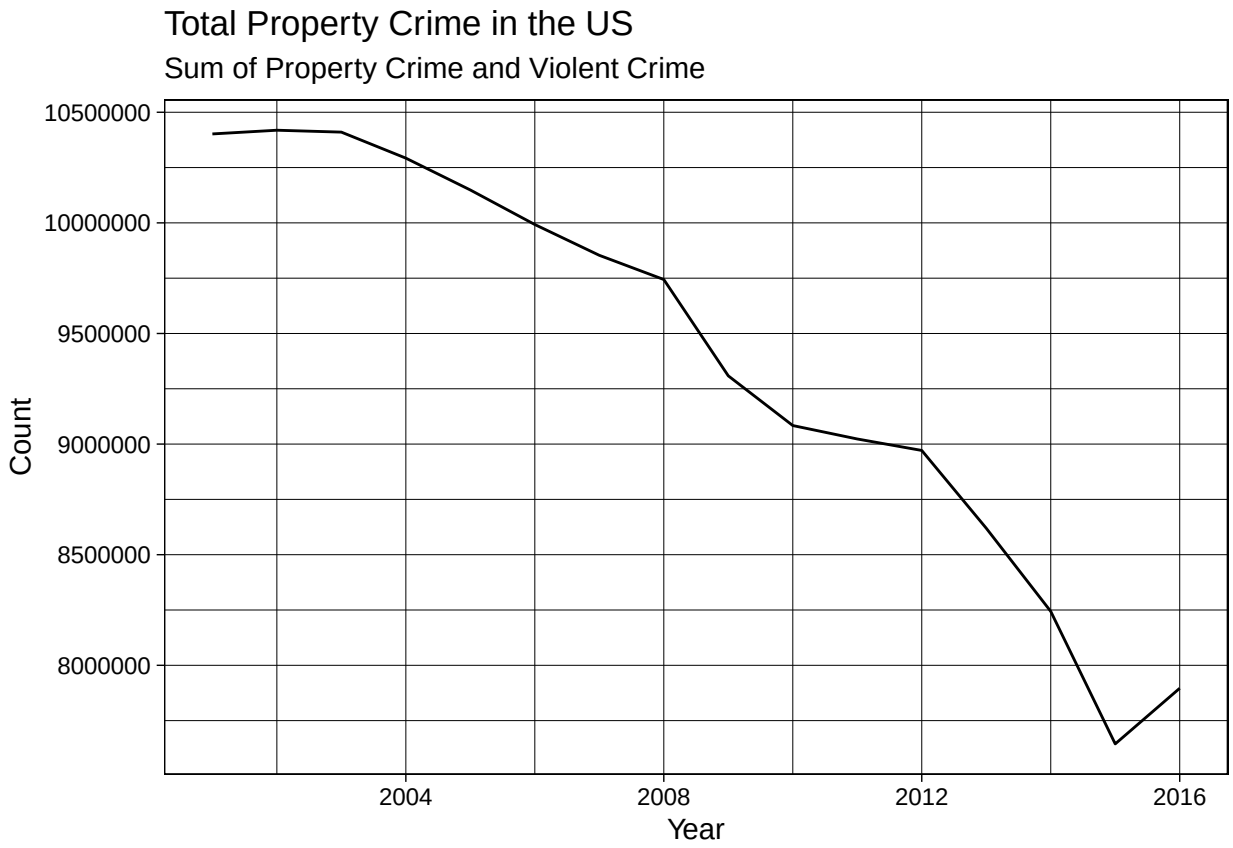
Political rhetoric that has come to the forefront in recent decades is the idea of law and order, being tough on crime, and ensuring punishment to the fullest extent of the law. We went into this project knowing that the number of incarcerated individuals in the US is the highest of any other developed country with a similar economy and governmental system. Knowing this, we would assume that if the US has a strong belief that we need to be tough on crime, the high rates of incarceration are due to high crime rates, violent or otherwise. With this in mind, we wanted to find if the data would support this hypothesis.

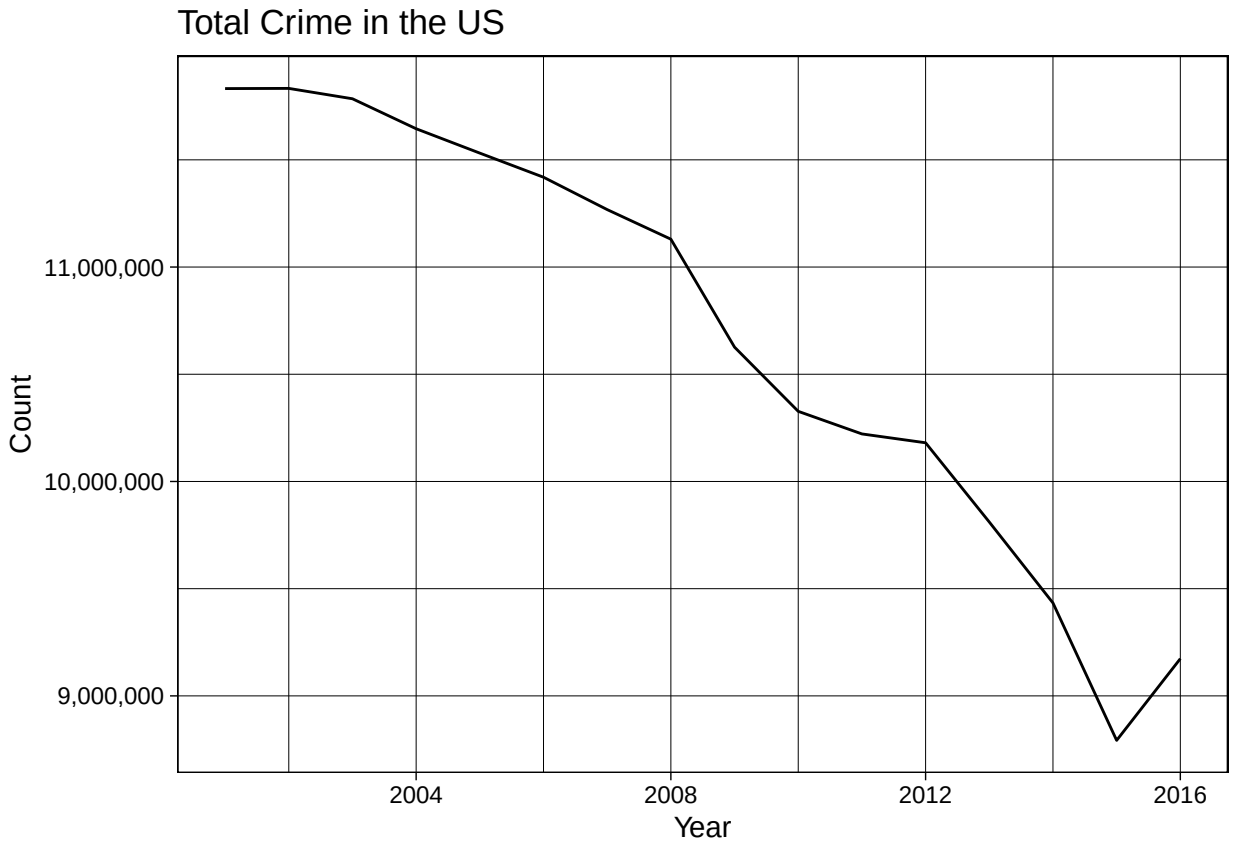


Above is the total prisoners in all 50 US states from 2001 to 2016. The graph shows a steady incline from 2001 to 2008 and then a steady decline followed by a sharp decline thereafter. Drawing attention to the

range of the y-axis, we can see that despite these sharp declines, the number of incarcerated individuals at the lowest point was still above 1.19 million, which is still much higher than the rate of incarceration in other countries.



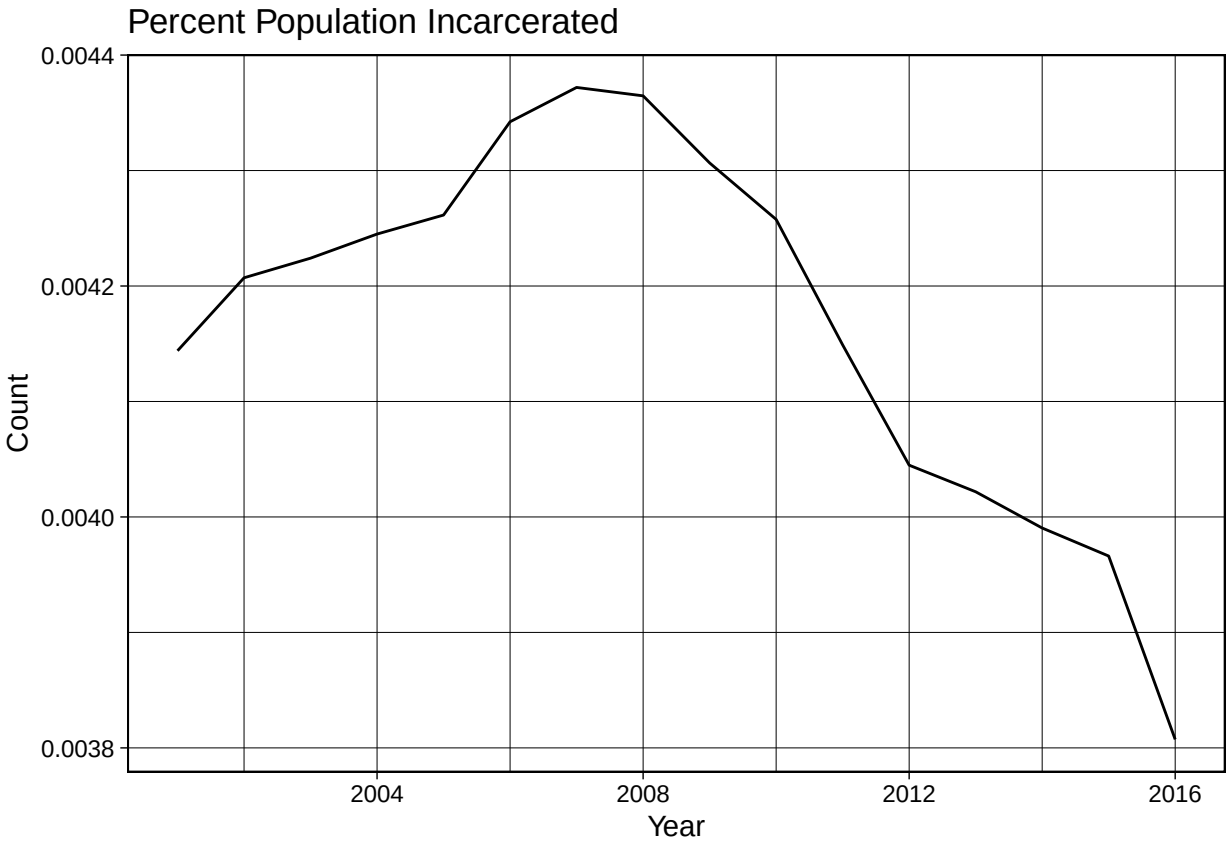




The three graphs above show the change in total violent crime (defined as murder/manslaughter, robbery,

rape, and aggravated assault), total property crime (defined as burglary, larceny, and vehicle theft) and the total crime (sum of property and violent crime) committed between 2001 and 2016. What is immediately evident is that even when crime of all types was sharply decreasing, the number of incarcerated individuals was still steadily decreasing. Based on this alone, though we cannot definitively draw this conclusion, it would appear that rates of incarceration are actually not due to how much or how little crime is being committed and is instead due to other factors which we will explore later in our analyses. Possible explanations might be policy changes, new sentencing guidelines, amount of policing, and a plethora of other reasons.

###Wouldn't this show a correlation between decreasing incarceration and decreasing crime rates?



After drawing the previous conclusion, we thought there may be another explanation, even if crime was going down it could be that US population was increasing, when we graphed it above, it would appear that the percent of the US population which is imprisoned shows very strong similarities to the Total US Prison Population graph. However, that being said, even if the population is increasing, if the crime rate is staying the same, or decreasing as we observed previously, there is still no valid reason to incarcerate more people.

###I feel like this doesn't answer the question above, so change the question or add political stuff?

Research Question #2

How does policy impact incarceration? This will be explored through the lens of changing definition of rape

Impact of Changing the Definition of Rape

In 2012 the FBI changed their definition of rape from: "The old definition was 'The carnal knowledge of a female forcibly and against her will.' Many agencies interpreted this definition as excluding a long list of sex offenses that are criminal in most jurisdictions, such as offenses involving oral or anal penetration, penetration with objects, and rapes of males.

to: The new Summary definition of Rape is: 'Penetration, no matter how slight, of the vagina or anus with any body part or object, or oral penetration by a sex organ of another person, without the consent of the victim.'"

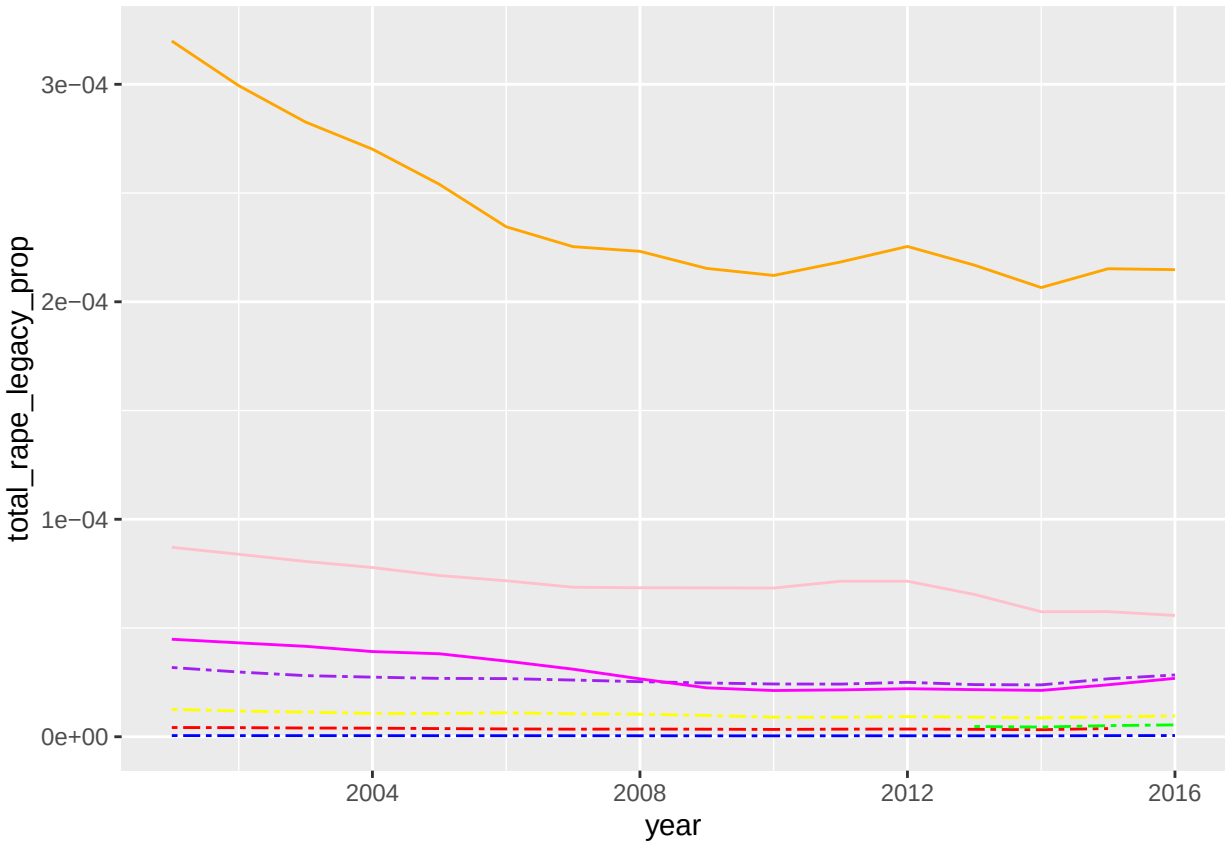
We wanted to see how the change in definition affected the number of prisoners convicted for rape. Based on the graph, it seems like after the new definition was implemented, the number of prisoners who are incarcerated for rape increased. We will perform tests to find the p-value and try to determine if this is a correlation relationship or a causal relationship.

https://ucr.fbi.gov/recent-program-updates/new-rape-definition- frequently-asked-questions

Proportions of Crime in the United States

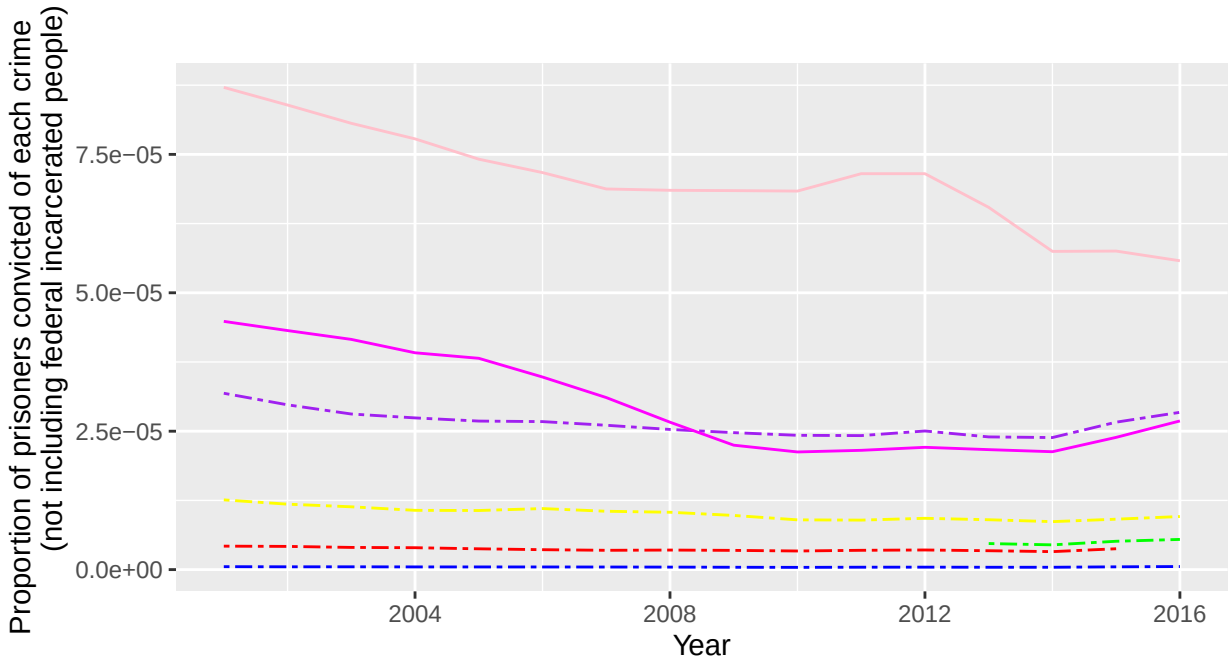
```
## Warning: Removed 50 row(s) containing missing values (geom_path).
```

```
## Warning: Removed 600 row(s) containing missing values (geom_path).
```

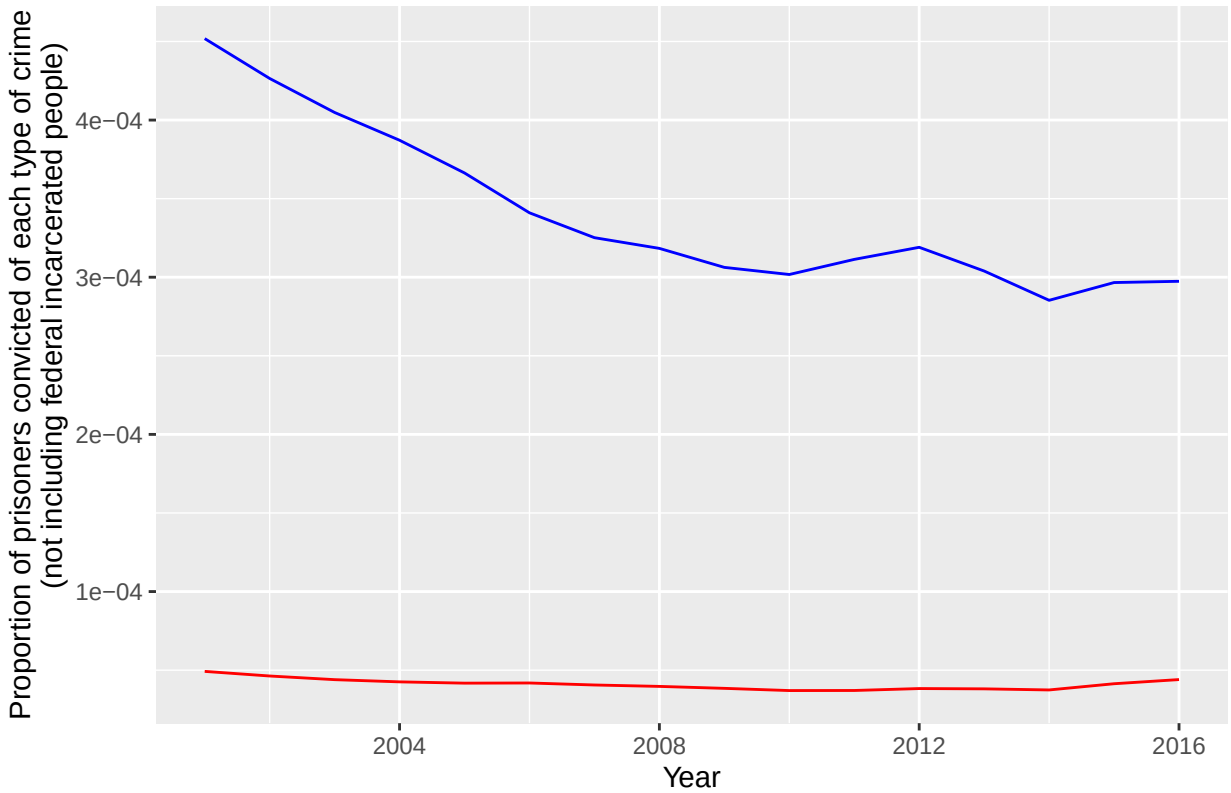



```
## $x
## [1] "Year"
##
## $y
## [1] "Proportion of prisoners convicted of each crime\n           (not including federal incarcerate
##
## $title
## [1] ""
##
## $caption
## [1] "Red = Proportion of Rape by legacy definiton\n               Blue = Proportion of murder/mansl
##
## attr(,"class")
## [1] "labels"
```

###STEVEN CAN YOU FIX THIS GRAPH TO NOT HAVE COMMENT LOOK WEIRD?



Red = Proportion of Rape by legacy definiton
Blue = Proportion of murder/manslaughter
Green = Proportion of Rape by revised definition
Yellow = Proportion of robbery
Purple = Proportion of aggravated assault
Pink = Proportion of burglary
Magenta = Proportion of vehicle theft



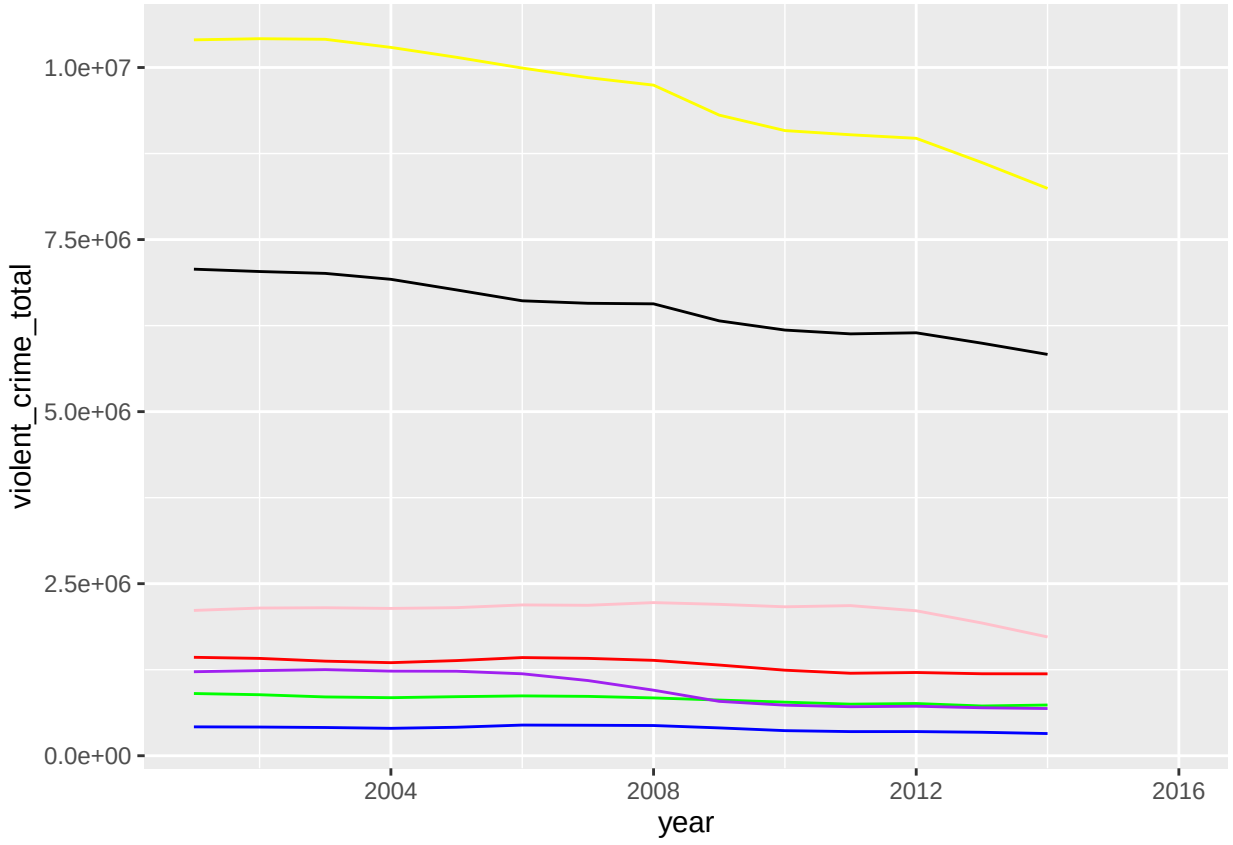
We wanted to see how the proportions of crimes were divided among the prior population. Also we wanted to see whether violent or non-violent crimes were represented more among the prison population. It appears

that non violent crimes make up a larger proportion of the prison population than violent crimes. Generally, violent and non-violent crimes are decreasing in regards to proportion of the prison population. Nonviolent crimes seem to be decreasing at a faster rate than violent crimes. Larceny makes up the largest proportion of crimes in the prison population. So we took it out for the second visualization to get a better look at the other crimes. Murder and manslaughter make up the smallest proportion.

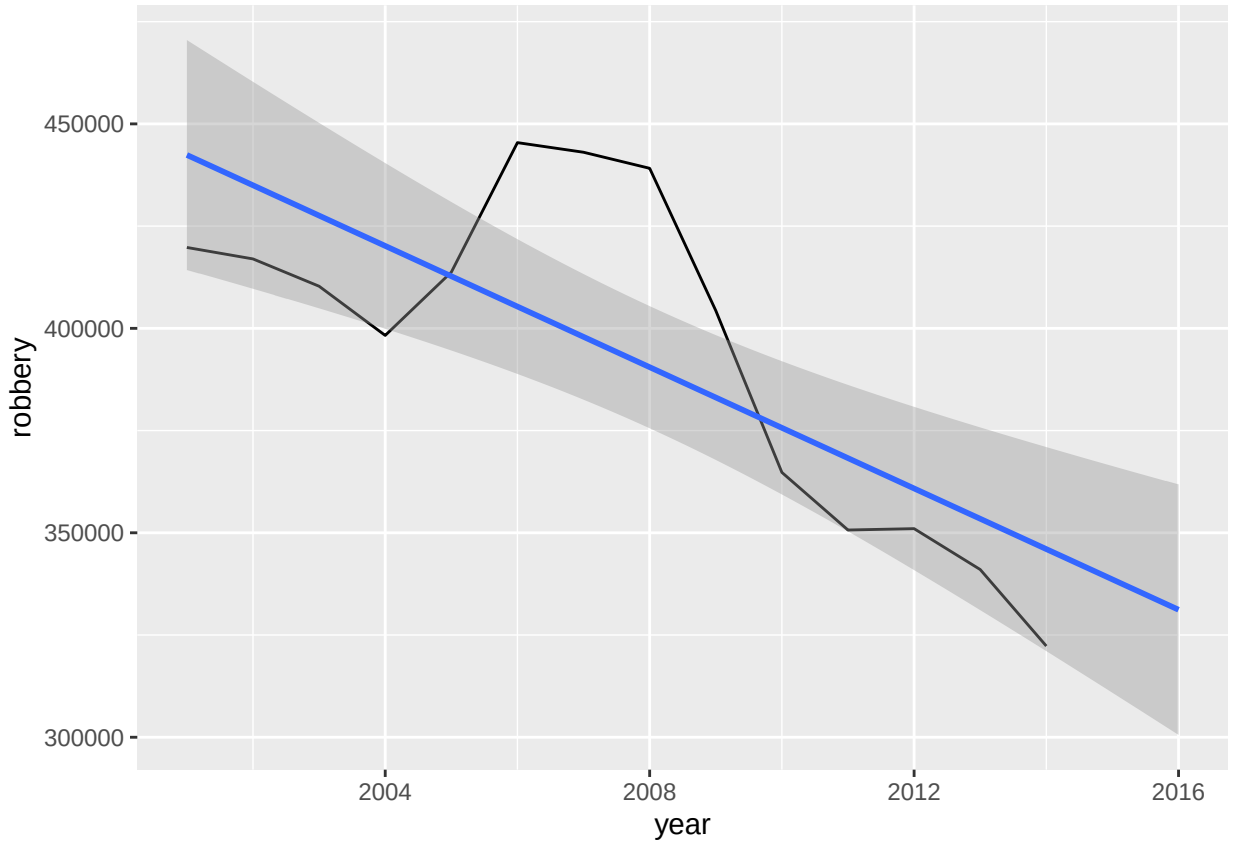
###How does what you did tie into the original rape question you were finding answers to?

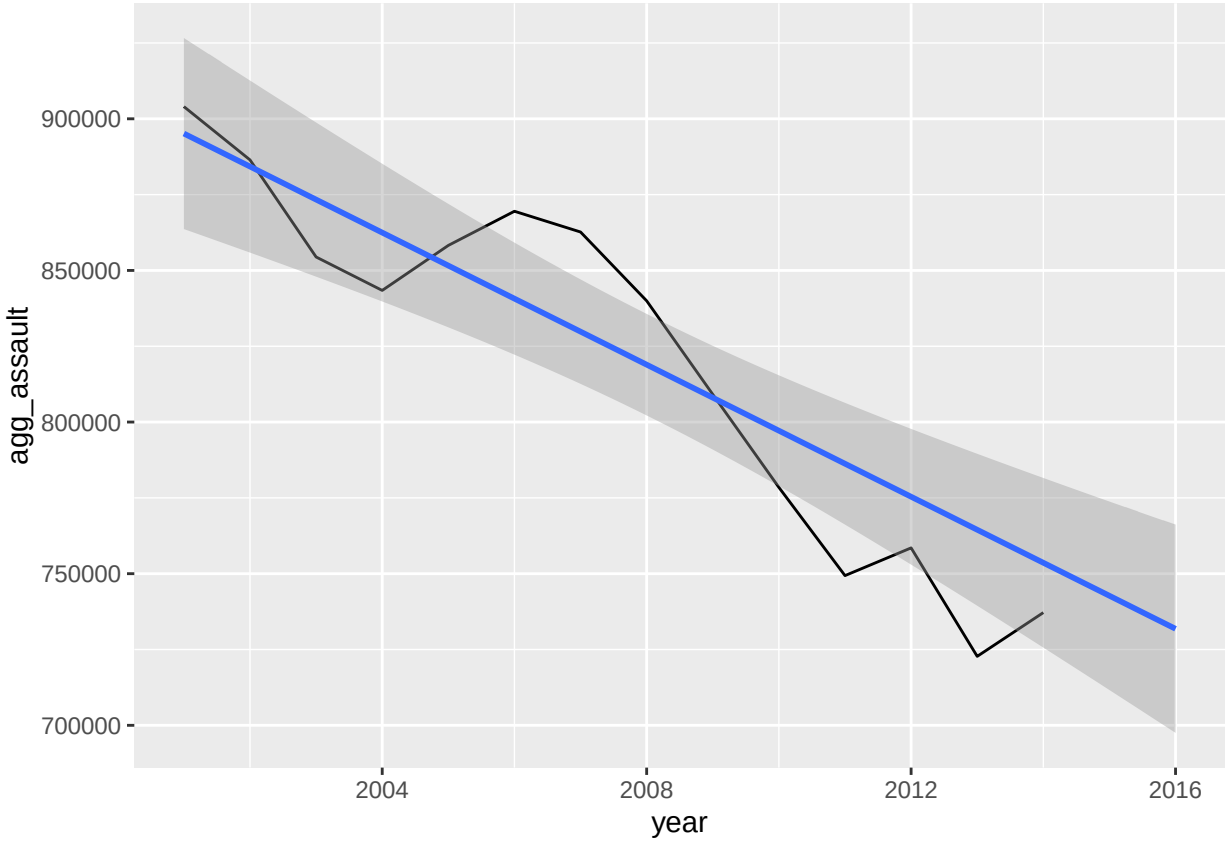
Research Question #3

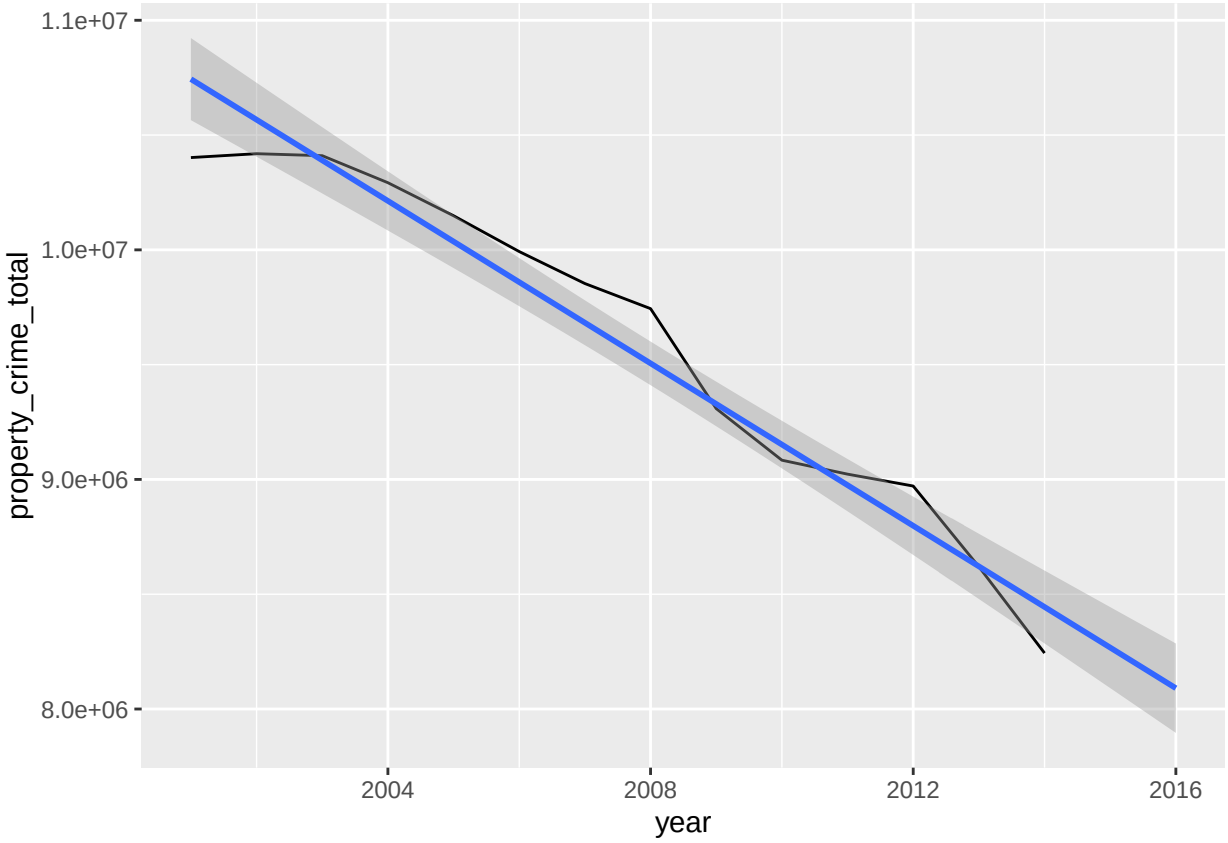
How do trends in different crimes compare to each other? Do they move together or do they change individually?

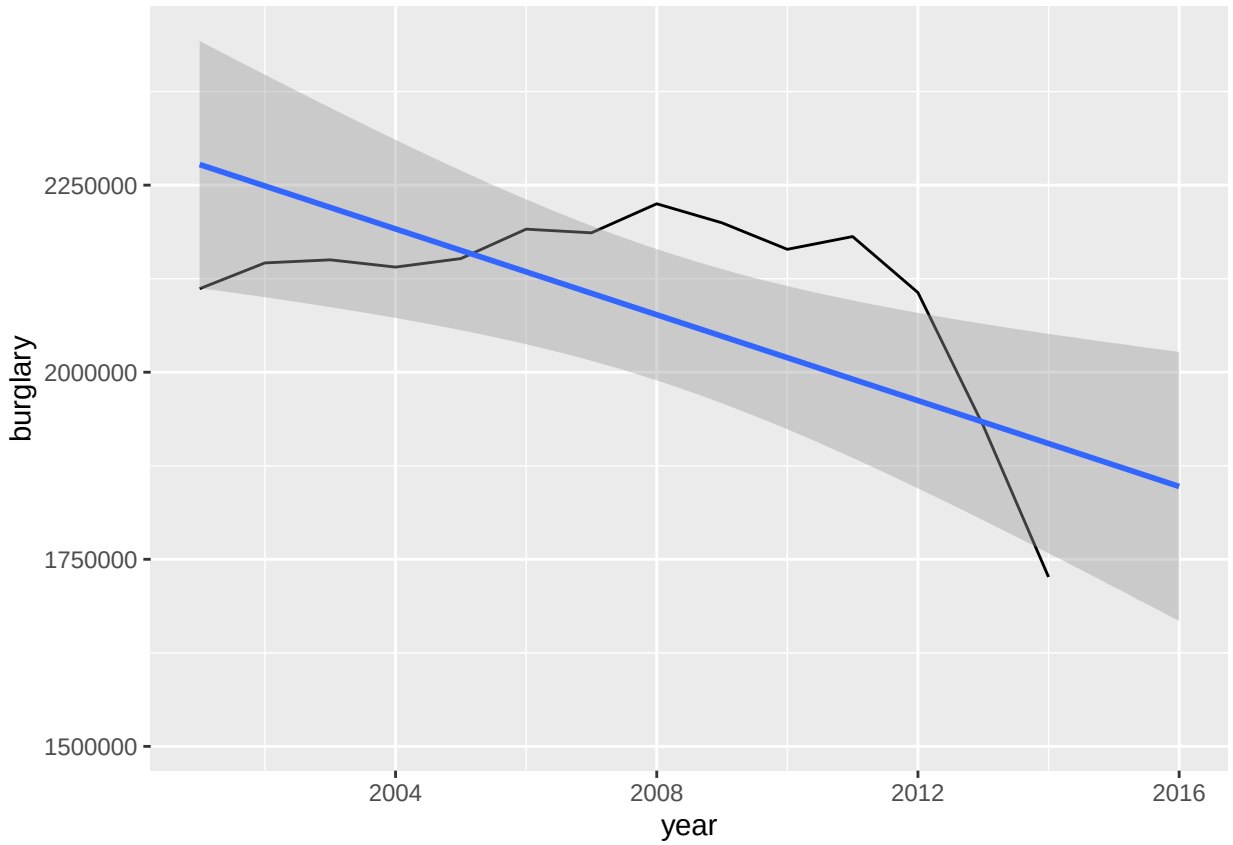


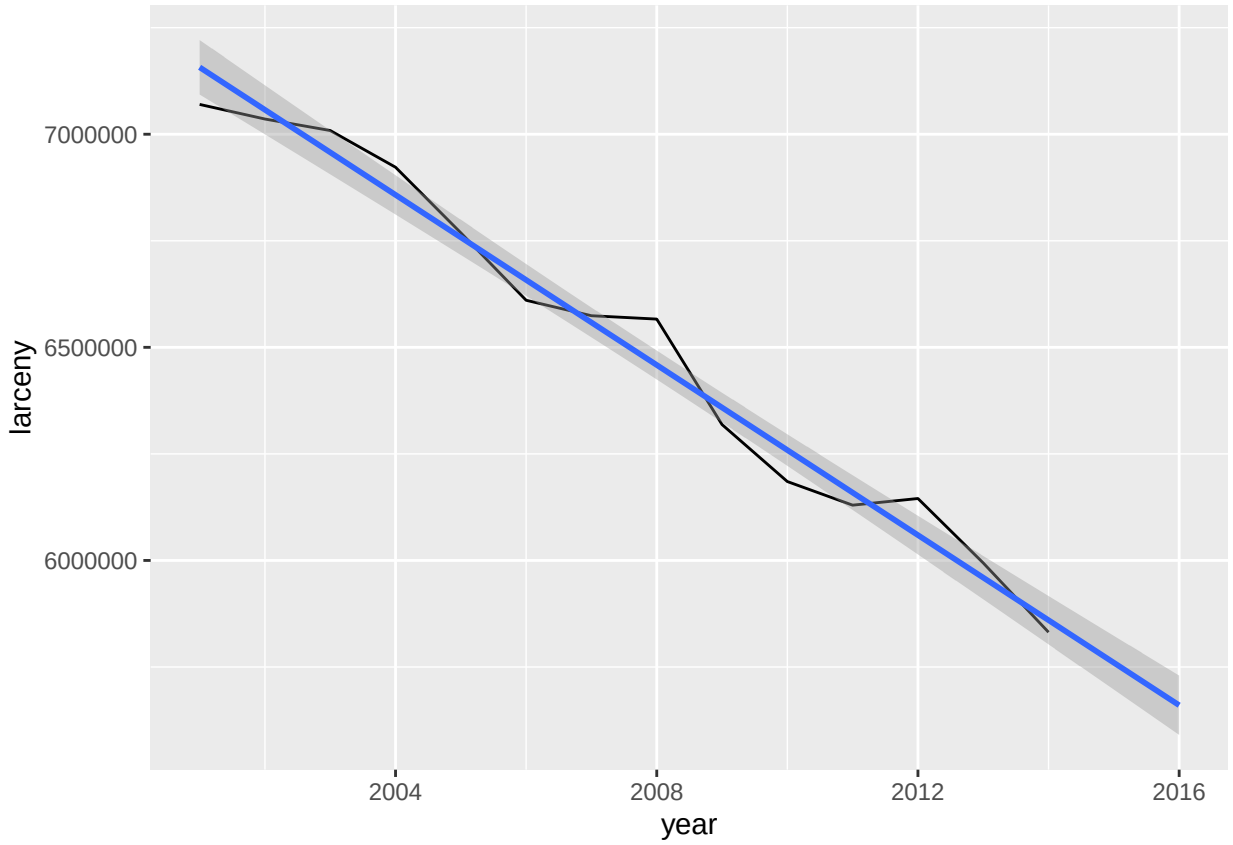

1500000
1400000
1300000
1200000
violent_crime_total
2004
2008
2012
2016
year

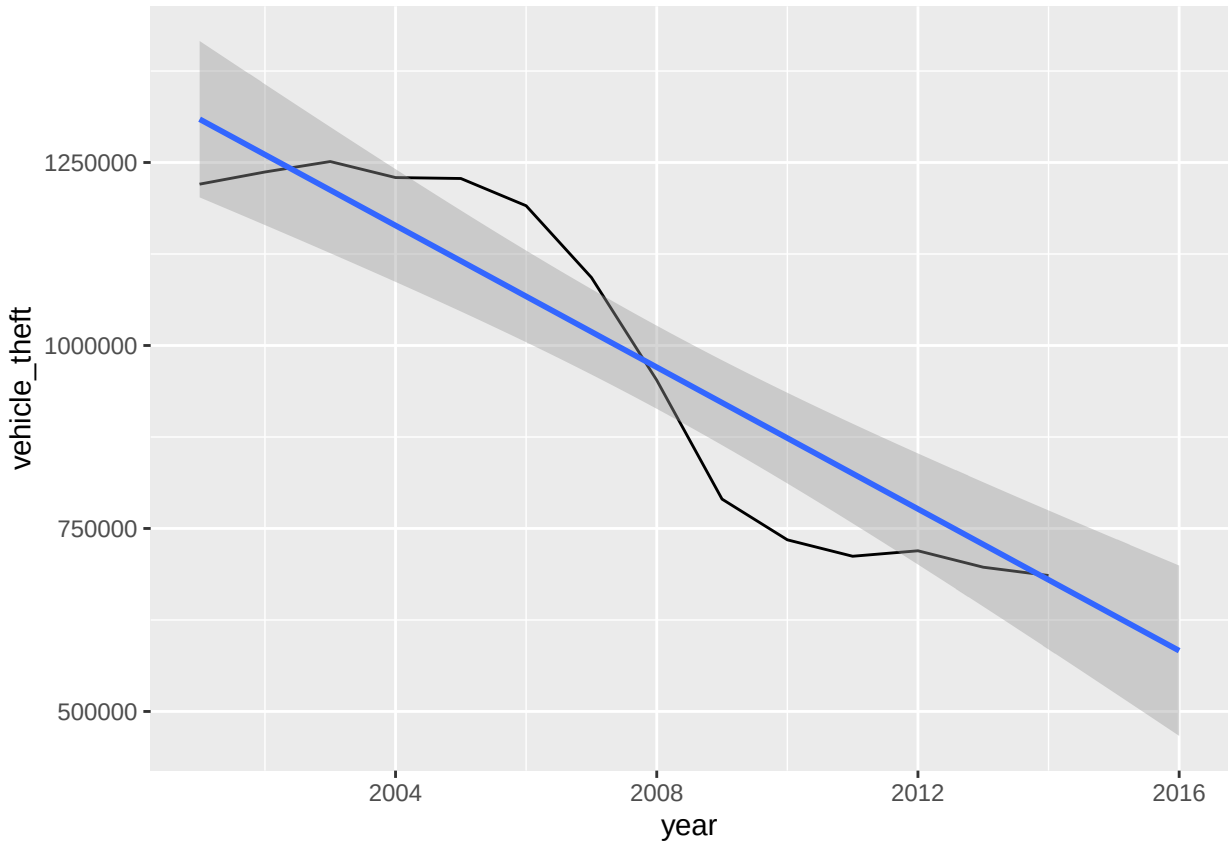

450000
400000
350000
300000
robbery
2004
2008
2012
2016
year


900000
850000
800000
750000
700000
agg_assault
2004
2008
2012
2016
year


1.1e+07
1.0e+07
9.0e+06
8.0e+06
property_crime_total
2004
2008
2012
2016
year


burglary
2250000
2000000
1750000
1500000
2004
2008
2012
2016
year


larceny
7000000
6500000
6000000
2004
2008
2012
2016
year




```
## # A tibble: 2 x 5
##   term         estimate std.error statistic   p.value
##   <chr>           <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept) 35134355.  6345033.      5.54 0.0000959
## 2 year          -16839.     3160.     -5.33 0.000137
```

```
## # A tibble: 2 x 5
##   term         estimate std.error statistic  p.value
##   <chr>           <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept) 15278231.  3131011.      4.88 0.000301
## 2 year           -7414.     1559.     -4.76 0.000376
```

```
## # A tibble: 2 x 5
##   term         estimate std.error statistic   p.value
##   <chr>           <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept) 22681002.  3509767.      6.46 0.0000213
## 2 year          -10887.     1748.     -6.23 0.0000307
```

```
## # A tibble: 2 x 5
##   term          estimate std.error statistic  p.value
##   <chr>            <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept) 364663825. 19917209.      18.3 1.15e-10
## 2 year          -176872.     9919.     -17.8 1.61e-10
```

```
## # A tibble: 2 x 5
```

```
##   term         estimate std.error statistic p.value
##   <chr>           <dbl>     <dbl>     <dbl>   <dbl>
## 1 (Intercept) 59667554. 18388187.      3.24 0.00639
## 2 year          -28681.     9157.     -3.13 0.00794
```

```
## # A tibble: 2 x 5
##   term          estimate std.error statistic  p.value
##   <chr>            <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept) 206818475.  7120397.      29.0 3.29e-13
## 2 year           -99781.     3546.     -28.1 4.93e-13
```

```
## # A tibble: 2 x 5
##   term         estimate std.error statistic    p.value
##   <chr>           <dbl>     <dbl>     <dbl>      <dbl>
## 1 (Intercept) 98177796. 11901185.      8.25 0.00000160
## 2 year          -48410.     5927.     -8.17 0.00000178
```

Here we attempted to answer whether different types of crime change together over the years or if different crimes fluctuate in different years. We used data from the variables `violent_crime_total`, `robbery`, `agg_assault`, `property_crime_total`, `burglary`, `larceny`, and `vehicle_theft`.

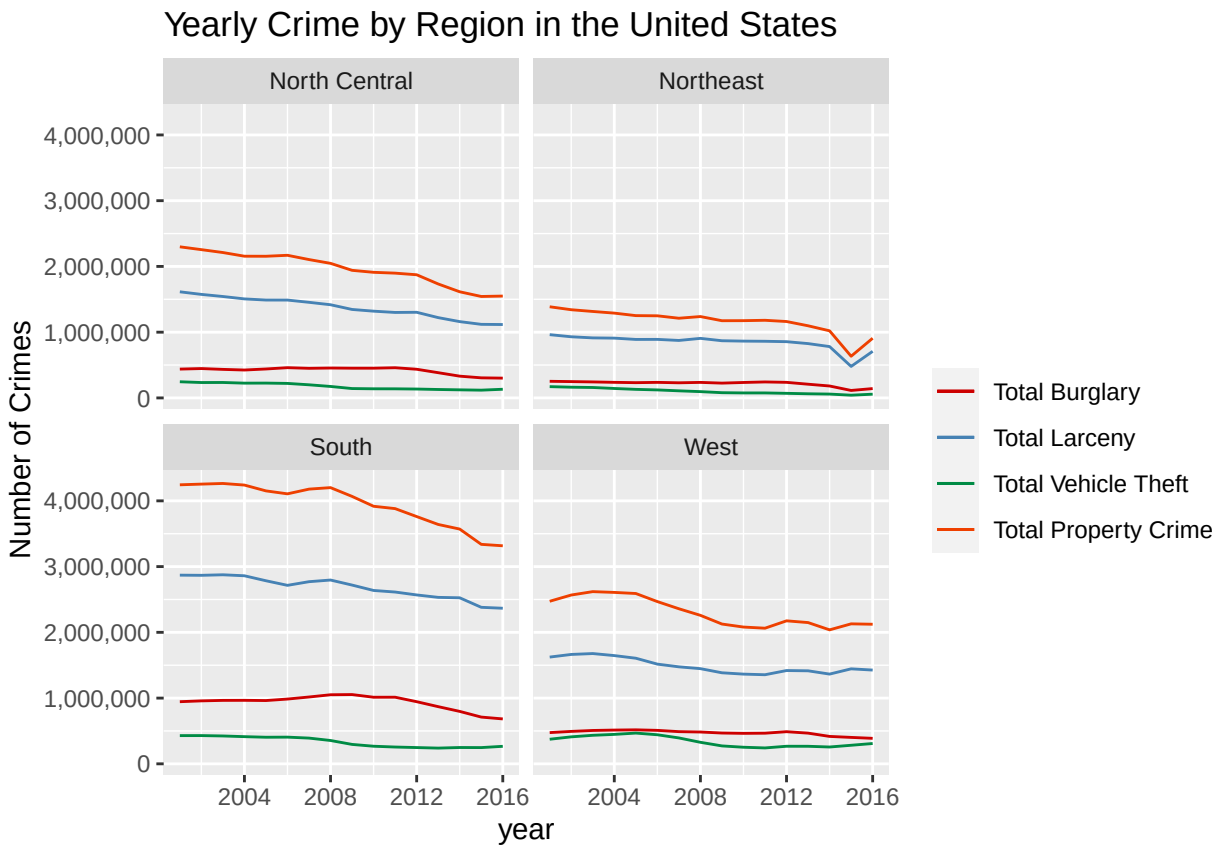
After attempting to compare the shapes of the crime vs. year graphs all together, we found that the incredibly different scales make this task impossible to do on a single graph. We then graphed these different crimes on different graphs with linear models, making it easier to compare shape but still difficult to make objective claims about this data. One notable feature we noticed was that there was a peak for many crimes between 2004 and 2008. We plan to research further to find possible causations for these trends.

###Do the researchˆˆ?

Last, we created linear models for each crime to see how they each changed between 2001 and 2016. Notice, all of these 7 crimes have a negative slope meaning they overall decreased during this time period. The slopes vary in magnitude from about -7000 crimes/year to about -180,000 crimes/year.

Research Question #4

How does crime differ for different regions of the US?



To see how crime differs by region, we made a variable called `region` which splits up the United States into 4 distinct regions (North Central, Northeast, South, and West). We decided to look at only property crime (burglary, larceny, and vehicle theft) and so we plotted the total buglary, total larceny, total vehicle theft, and overall total property crime from 2000-2016 on each regional graph to see what the trends looked like.

One positive note with these graphs is that each region has a overall general decline in the amount of property crime throughout 2000-2016. This is only a speculation, but this would suggest that trends of property crime does not differ throughout each region but it does for each crime.

The south overwhelmingly has the total highest property crime total burglary, and total larceny.

An outlier seems to be in the northeast graph in 2015 when there was sharp decrease with larceny and burglary, and thus total property crime. There must have been policy that affected the Northeast only during that time period that lessened crime

Because of the high number of crimes in the south and the downward spime for the northwest graph that the north central and west regions are most similar in numbers of crimes committed over this 16 year span.

It is also interesting to see that the types of crimes themselves are seemingly related. But this we mean that for each graph, from the highest number to the lowest number, the crimes are larceny, burglary, then vehicle theft.

If you were looking at a region to live in the united states and buy a house or car, it seems that the best place to do that would be in the North East, followed closely by the north central and west region. Thus it would be advised to not choose the south, although they are on a decline in property crimes committed.

###Should we ask Yue if it's ok to have multiple research questions that are seperate or do we need to tie them together with one central one?