Crime Prevalence on the State Level

Data Science Capstone Project
Ankita Hanmiah

Why Is This Important Information for You?

- Understanding where and what types of crime are most prevalent
  - Is useful to state governments
    - For developing solutions and pinpointing problem areas
  - Is useful for the general public
    - Helpful in determining which states are "safer" for people concerned and wanting to settle down in areas of less crime
    - People wanting to be more informed and aware of potential crimes where they live
- Observing the rate of poverty and the relationship of poverty to crime
  - Understanding the relationship between poverty and crime would be useful for working at prevention
    - Developing helpful solutions to prevent impoverished people from turning to crime

Where is This Information Sourced From?

- Crime Data
  - From the Federal Bureau of Investigation's Uniform Crime Reporting program
    - Collects the data of committed crimes reported to the FBI every year and releases around 24 tables
      - Breakdowns of regions, states, counties, cities, metropolitan versus non-metropolitan areas, suburban versus non-suburban areas, generalized location of the offenses, weapons used in violent crime, time of day
      - Table 4 of 2017 used for this project (includes data for all 50 states plus District of Columbia
        - Counts and rates of different criminal offenses
        - Counts and rates for violent and property crime
- Poverty Data
  - From the United States Census' compilation of poverty data
  - Measure of who falls under the poverty line:
    - Sets of income thresholds that differs based on the size of the family and their composition
      - Total income before taxes and minus benefits should fall beneath their specific threshold

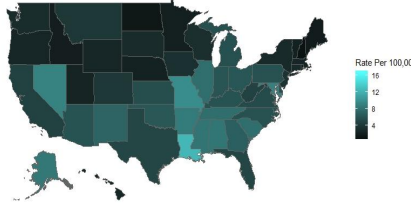# Observing Individual Criminal Offenses

Violent Crimes:

Rate of Aggravated Assault by State, 2017

Rate of Murder by State, 2017

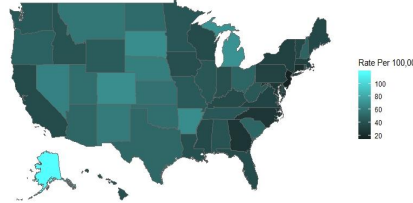Rate of Rape by State, 2017

Rate of Robbery by State, 2017
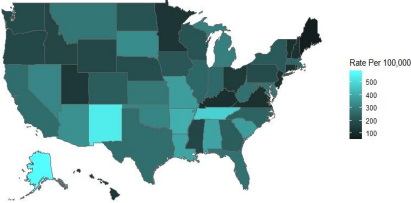
Aggravated Assault
- Highest rates (over 500 cases per 100,000)
  - New Mexico, Alaska, and Tennessee
- Rates of 400 to 500
  - Arkansas, Louisiana, Missouri, Alabama, South Carolina, and South Dakota
- 100 to 200 per 100,000 rates
  - New Hampshire, Vermont, Connecticut, New Jersey, Virginia, Kentucky, Ohio, Minnesota, and Utah
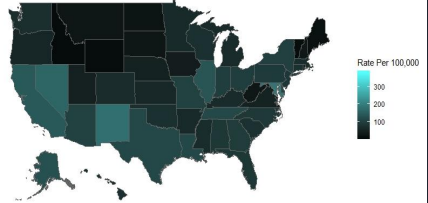- Lowest prevalence (below 100) seems
  - Maine

Murder
- Least prevalent crime reported to the FBI
  - Lowest of crime rates 0 to over 16 cases per 100,000 people
- Mostly mid to low rates (from 0 to 12) throughout the United States
- Higher murder rates
  - Louisiana, Nevada, Missouri, and Maryland
- Lowest rates below 4
  - New England/northeast region, Hawaii, and states in the northern midwest and northwest (Minnesota, North Dakota, Nebraska, Idaho, and Utah)

Rape
- Highest prevalence rating over 100 per 100,000 in Alaska
- Most other states show middle range rates (between 40 to 90 per 100,000)
- Georgia and North Carolina in the 20 to 40 range.
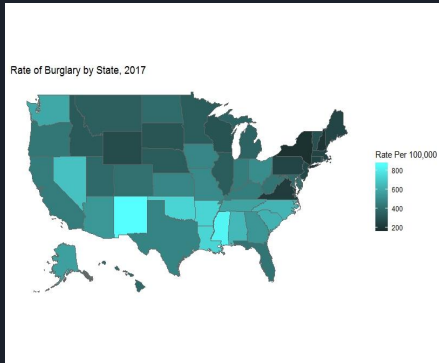- Connecticut and New Jersey are the lowest rated states (rates below 20)

Robbery
- Fairly low with rates below 150
- Except for states like Nevada, New Mexico, Maryland, and Illinois (rates over 200)

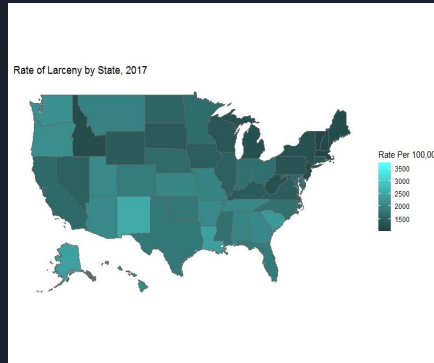# Observing Individual Criminal Offenses

Property Crimes:

Rate of Burglary by State, 2017

Rate Per 100,000

Rate of Larceny by State, 2017

Rate Per 100,000

Rate of Motor Vehicle Theft by State, 2017

Rate Per 100,000

Burglary
- High rates throughout the south (500 to over 800 cases per 100,000)
  - Most highlighted states being New Mexico and Mississippi, followed by states like Oklahoma, Arkansas, and Louisiana.
- The lowest rates of burglary (200 to 250 per 100,000)
  - New York, New Hampshire, and Virginia

Larceny
- Highest rates (2500 to 3500 per 100,000)
  - New Mexico, Louisiana, South Dakota, and Alaska
  - Fairly widespread throughout the United States
- Most prevalent of the individual criminal offenses reported to the FBI.
- Lowest rates (below 1500 per 100,000)
  - Parts of New England/Northeast regions
  - Parts of the midwest (such as Wisconsin and Michigan), and Idaho
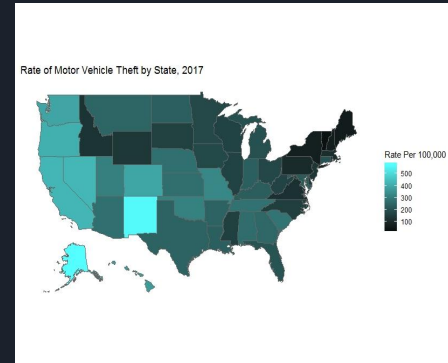
Motor Vehicle Theft
- Most prevalent in New Mexico and Alaska (over 500 per 100,000)
- Followed by west coast states (Washington, Oregon, California) in the 400 to 500 range
  - Also Nevada and Colorado
- Lowest rates (below 100) look to be in the New England/Northeast states (Maine, New Hampshire, Vermont, New York, and Pennsylvania)

Maximum Crime States

For each offense, we observed which states had the maximum and minimum crime rate. The maximum crime rate for:

- Aggravated Assault
    - Alaska with a rate of 575 per 100,000 people
- Burglary
    - New Mexico with a rate of 858 per 100,000
- Larceny
    - District of Columbia with a rate of 3650 per 100,000
- Motor Vehicle Theft
    - Alaska with a rate of 576 per 100,000
- Murder
    - District of Columbia with a rate of 16.7 per 100,000
- Rape
    - Alaska with a rate of 117 per 100,000
- Robbery
   - District of Columbia with a rate of 378 per 100,000

Minimum Crime States

The minimum crime rate states for each offense are:

- Aggravated Assault
  - Maine with a rate of 65.3 per 100,000 people
- Burglary
  - New York with a rate of 176 per 100,000
- Larceny
  - Massachusetts with a rate of 1078 per 100,000
- Motor Vehicle Theft
  - Vermont with a rate of 31.1 per 100,000
- Murder
  - New Hampshire with a rate of 1 per 100,000
- Rape
  - New Jersey with a rate of 16.7 per 100,000
- Robbery
 - Idaho with a rate of 11.4 per 100,000

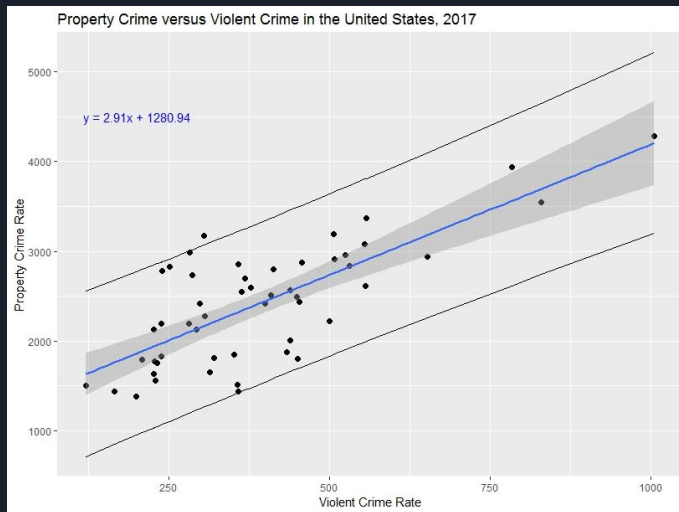# Do Poverty and Crime Go Hand in Hand?
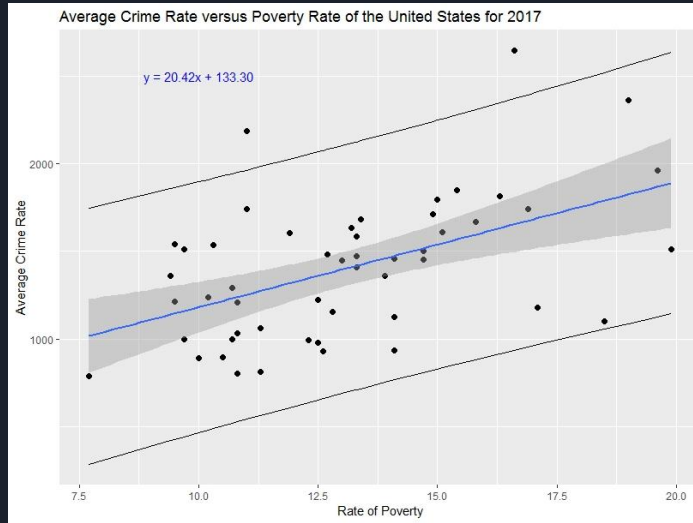
- Gray shading indicates confidence interval
  - Shows range within which there is a 95% probability that the actual observed dependent variable (average crime rate) will be found
- Black lines indicate prediction interval
  - Estimated range within which future predicted values of the dependent variable may possibly be found, based on the actual observed data

Do Poverty and Crime Go Hand in Hand? (Continued)

- Correlation coefficient = + 0.510
  - Indicates the strength of correlation between two variables as a value between -1 (strong negative correlation) and +1 (strong positive correlation)
  - Moderately strong positive correlation shown for poverty and crime
- R squared = 0.2596
  - Relatively measures how good the fit of the linear model is
  - Fairly low
- Root mean squared errors = 341.518
  - The root mean squared errors value is the square root of the residuals' variance and is in same units as the dependent variable
    - Measures the absolute fit of the model
    - Residuals are the difference between the observed and predicted values of the dependent variable
  - The lower the value is, the better the fit

Average Crime Rate versus Poverty Rate of the United States for 2017

y = 20.42x + 133.30
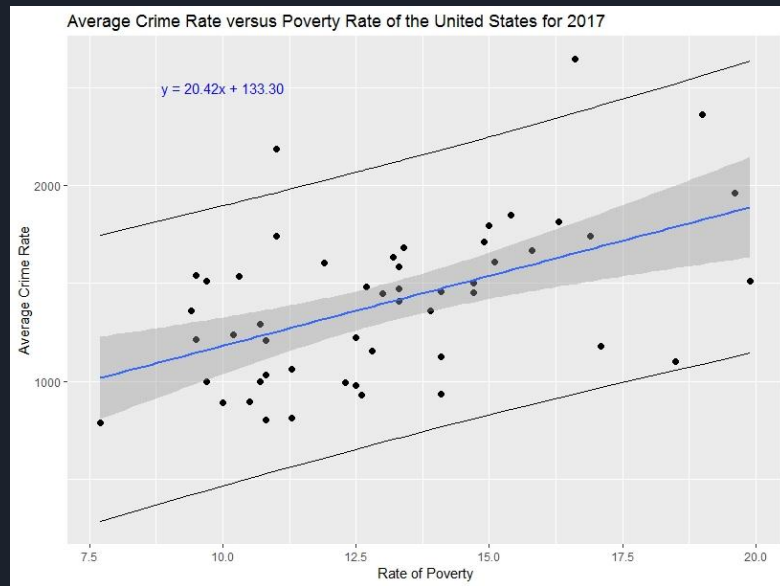
Is Property Crime Related to Violent Crime?

Property Crime versus Violent Crime in the United States, 2017

- Correlation Coefficient = + 0.7515
  - Fairly strong positive correlation
- R Squared = 0.5647
  - Fairly good fit
- RMSE = 434.227

Going Forward...What Can Be Done?

- General Public
  - Can choose the "safer" states to settle their families in
  - Can invest in safety measures to prevent themselves from being victims of property crimes
- State Governments/Law Enforcement
  - States with higher poverty and crime
    - Could build mentorship programs for children belonging to families of poverty
      - Attempt to safeguard children from dangerous environments and promote education and future career opportunities
  - States with higher crime rates
    - Invest in more surveillance or have more patrolling officers be present

Further Investigation

- Would delve into the other tables provided by the FBI UCR
  - Look more specifically into the states on a city level
    - Determine which cities within the higher crime rated states are more responsible for higher crime
    - Whether or not there is a difference between metropolitan and non-metropolitan areas
    - Urban areas versus rural areas
  - Doing a study on crime rates over the years (in the last decade)
    - Time series plot to understand trend of crime over time
    - Could help in making predictions for whether crime will increase or decrease