Jett Gemmer, Andrew Song, Amanda Chen, and Lauren Cali
CS3300: Data-Driven Web Applications
October 6, 2023

Project 1 Report

Investigating the Relationship Between Unemployment and Total Crime in the United States

United States Rates of Unemployment vs. Total Crime

US Unemployment and Crime 1981-2018

How have unemployment rates in the US changed over the past 40 years? In which states is unemployment highest, and where has it improved the most over time?
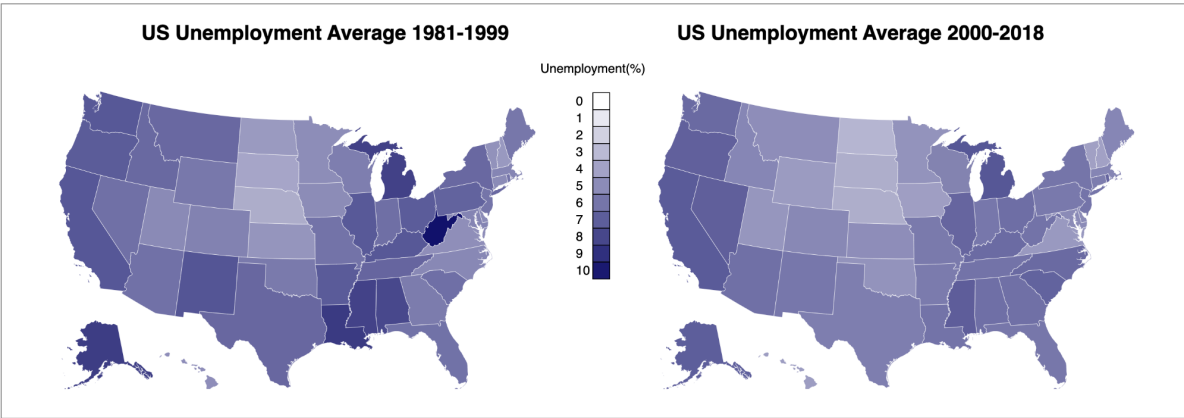

Figure 1. US Unemployment Averages from 1981-1999 and 2000-2018.
These maps were generated with unemployment data by collected by state, which was then transformed to calculate ~20 year averages, and state geometric data. A sequential color scale shows unemployment ranging from 0-10%. This visualization shows an overall decrease in average unemployment rate between the to ranges in most states.
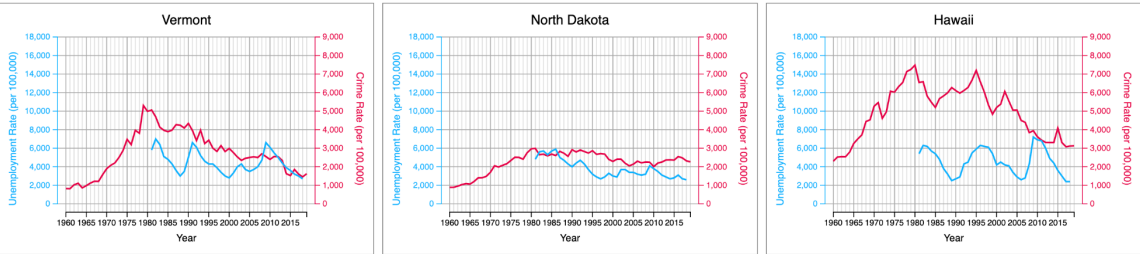
US Unemployment and Crime 1960-2019

What is the relationship between unemployment rates and crime rates? What is the overall trend for the United States, and how does it differ by state? Are there any consistent patterns between them?

US Total
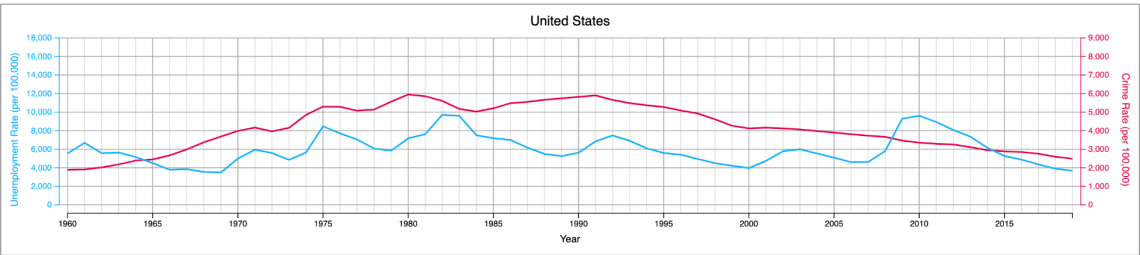
Top 3 States with Lowest Unemployment Rates in 2019

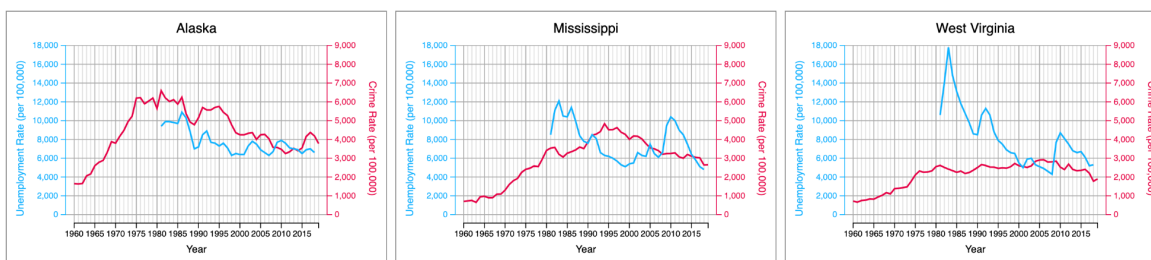**Top 3 States with Highest Unemployment Rates in 2019**

Figure 2. Unemployment Rates and Crime Rates from 1960 to 2020. State datasets contain unemployment data from 1980 onwards.
These graphs show a weak, positive relationship between unemployment and crime rates. Analysis of the overall United States shows that 5/6 major unemployment spikes occur near-synchronously with increases and decreases in crime rate. However, analysis at the state level reveals greater variability. States with stable rates like Vermont, North Dakota and Alaska show closely linked unemployment and crime, while states with greater fluctuation like Mississippi and Hawaii show weaker relationships.

ac2262 (Amanda), ajs557 (Andrew), jmg575 (Jett), lmc339 (Lauren)
Cornell University
CS3300: Data-Driven Web Applications
Fall 2023

Data Sets Utilized

Our visualizations utilized four data sets, two for US Total and two for individual States.

US Total:
- United States Crime Rates 1960 - 2019: https://www.disastercenter.com/crime/uscrime.htm
  - Used an HTML to CSV converter to get data: https://www.convertcsv.com/html-table-to-csv.htm
  - *us-crime-rate.csv*
- Labor Force Statistics from the Current Population Survey: https://data.bls.gov/timeseries/LNS14000000
  - Modified file in Excel and exported to CSV: Added a "Total (per 100K)" column by averaging 12 months and multiplying by 1000
  - *emp-unemployment.xls >> us-unemployment-rate.csv*

By State:
- State Crime: https://corgis-edu.github.io/corgis/csv/state_crime/
  - Used Data.Rates.Property.All and Data.Rates.Violent.All for total crime
  - *state_crime.csv*
- Annual Unemployment Rates by State: https://www.icip.iastate.edu/tables/employment/unemployment-states
  - Wrote python script to create CSV with Year, State, Rate columns
  - *us-unemployment.csv >> us-unemployment.py >> us-unemployment-2.csv*

The original source of the State Crime data set is the United Crime Reporting Program website, which exists to provide "...reliable statistics for use in law enforcement. It also provides information for students of criminal justice, researchers, the media, and the public." This gives us confidence in the data's accuracy. The State Crime data set contains crime reports ranging from 1960 to 2019. While it measures several statistics, the primary ones we chose to focus on were the Violent Crime Rate

(measured per 100,000 people as a float) and the Property Crime Rate (measured per 100,000 people as a float) as well as the year they occurred. To get a combined crime rate, we summed the two variables. These statistics are specified in the data source table as *Data.Rates.Violent.All* and *Data.Rates.Property.All.* This data set was used for the visualizations above labeled Figure 2. The crime rate is indicated on the right vertical axis of each graph and is measured in the same units as the original data set.

The next data source we used was the U.S. State Crime Rates. This data set measured the total number of several different types of crimes across all of the United States from 1960 to 2019. We elected to analyze the total rate of all crimes in a given year rather than one specific category. In order to make the data align with our other data sets (and subsequently easier to compare), we decided to use the version that reported the crime rate per 100,000 people. To get a usable version of this data, we used an HTML to CSV converter. This data set was used for our line graph that analyzed unemployment rates and crime rates in the United States as a whole from 1960 to 2019. The crime rate is indicated on the right vertical axis and measured in the same units as the original data set.

The third data set we used came from the U.S. Bureau of Labor Statistics (BLS). It measured the unemployment rate of the U.S. population as a percentage every year for each individual month. For the sake of this data set, the population was considered anyone 16 or older. The available data ranged from 1948 to 2023, but we elected to only analyze from 1960 to 2019 to fit with our other datasets. This data set was used for both our visualizations of U.S. unemployment.

The final data set measured the Annual Unemployment Rates by State. This data set was obtained from Iowa State University's website, but was originally sourced from the U.S. Bureau of Labor Statistics (BLS). The unemployment rate was measured as a percentage of the state population and, like our previous data set, considered employable citizens to be those 16 or older. This data set was utilized for our visualizations in Figure 2 and is represented by the left vertical axis.

In designing our choropleth map, we used a TopoJSON file containing a template for the United States map, which was sourced from GitHub.

Design Rationale

Our project contains two sets of visualizations: our US unemployment choropleth maps, and our 3-axis line charts. The choropleth maps used the "Annual Unemployment By States" dataset to determine average unemployment rates, while using a state geographic data json to generate the map. Mapping was then conducted to set each state's value to its corresponding unemployment average. The choropleth maps were positioned side-by-side to give an easy comparison of 20 year unemployment averages. A blue sequential color scale was used to indicate each states' unemployment percentage. This visualization channel was chosen because unemployment ranged from 0-10% during this period with each percentage being equally relevant, thus a sequential scale was more fitting then a divergen scale. Blue was the chosen hue because of its clear visibility. A map of the United States was chosen as the chart area, because it already contained clear state border lines as marks. This was a more effective visual mark than, say, indicating each state as a circle across some scale.

For comparing unemployment rates and crime rates, a 3-axis line chart with time as the x-axis was the visual channel selected. This allows us to plot both unemployment vs time and crime vs time on the same graph, allowing for the direct comparison of the two statistics. Red and blue were chosen as color strokes for the different lines, since they contrast and are easily distinguishable. Linear scales were used for our data transformations, since both unemployment and crime rates did not change drastically enough to necessitate a nonlinear scale, nor did changing the scale reveal relational patterns. The line chart of the overall United States was placed at the top of the other charts, since we wanted viewers to observe this chart first. State charts were then placed underneath in a 2x3 format. States were chosen based on 2019 unemployment rates, selecting for the 3 highest and 3 lowest rates.

Story

The purpose of this project was to uncover any relationship between unemployment and crime rates in the United States. Our initial hypothesis was that there would be a strong, positive and proportional correlation between the two. Our reasoning behind this was that unemployment means loss of income, and thus would likely lead to an increase in property theft such as robbery, burglary, and shoplifting.

Our data visualizations tell us a somewhat different story. In our graphs depicting "Unemployment Rate (per 100,000)" and "Crime Rate (per 100,000)" over "Year", for the United States there appears to be a weak, positive relationship between the two variables. From 1960 to 1990, crime rose steadily from 2,000 to 6,000 and then decreased to 2,500 by 2020. Unemployment on the other hand, increased and decreased in waves during that time period. From 1960 to 2020 unemployment rates saw 6 major spikes, usually returning to approximately 5,000 or so every decade. Tracking the peak of each unemployment spike, from 1970 to 2005 the unemployment rate does appear to increase and decrease correspondingly with crime rates. In 2010, there was a major spike in unemployment, but surprisingly the crime rate continued to decrease. This spike was likely caused by the 2008 housing crisis and subsequent recession, so it is interesting that although millions of Americans lost their jobs, crime rates did not rise. It appears that there is a weak and positive correlation between unemployment and crime rates, but there is not enough evidence from our data visualizations to conclusively state any causality between the two.

This determination is supported through further investigation of individual states. In states like Alaska, North Dakota, and Vermont, unemployment rates appear closely tied with crime rates. However, the rates in these states appear stable over time. This seems more indicative of socioeconomic stability overall in these states however, rather than necessarily a causal relationship between unemployment and crime. In states like Mississippi and West Virginia, with large fluctuations in unemployment, there is no clear relationship with crime rates. Hawaii is an interesting case where unemployment and crime have large fluctuations, and somewhat match in trends. Data analysis at the state level reveals very different relationships between unemployment and crime rates over this 40 year period. While unemployment and crime rates appear to have a weak correlation, this indicates that there are very likely other socioeconomic factors that determine these two statistics.

Team Contributions

Team Member	Work Completed	Hours Worked
Amanda	- Discussed + chose topic over zoom meeting - Sketched plan of idea - Created line graphs of crime rate vs. unemployment rate for US Total and States - Fixed alignment of maps and charts - Added styling to/reworked layout of site	11
Andrew	- Discussed + chose topic over zoom meeting -Wrote up project report design rationale and story - Presented project in class - Added figure captions	4.5
Jett	- Discussed + chose topic over zoom meeting - Created first map of unemployment % by states - Finished maps of unemployment rates	7
Lauren	- Discussed + chose topic over zoom meeting - Wrote up project report descriptions	4.5