

Egypt University of Informatics

Computer and Information Systems

Data Analysis Course

The Analysis of **The Relationship Between Unemployment Rates and Crime in the U.S.**

Submitted by: Omar Mohamed Abdelsalam 23-101207 || Kirollos Rafaat 23-101223 ||

Youssef Walid Salah 23-101117 || Ali Tamer 23-101210

24/05/2025

# Introduction

Understanding the connection between economic conditions and crime is crucial for shaping effective public policies. In particular, unemployment has long been viewed as a factor that may influence crime rates due to its impact on financial stability, social stress, and opportunity structures. This study investigates the statistical relationship between unemployment rates and crime levels in the United States, utilizing data from all 50 states over several decades. Historical trends, including the 2008 financial crisis and the COVID-19 pandemic, have underscored the relevance of this issue, making it timely and significant for policymakers, criminologists, and economists alike.

# Research Question

Is there a statistically significant relationship between unemployment rates and crime levels across U.S. states?

# Hypothesis:

*Null Hypotheses (H₀):*

1. There is no difference in violent crime rates between states with high and low unemployment.
2. There is no difference in property crime rates between states with high and low unemployment.

*Alternative Hypotheses (H₁):*

1. States with higher unemployment have significantly different (higher) violent crime rates.
2. States with higher unemployment have significantly different (higher) property crime rates.

# Population of Interest:

The population of interest includes all 50 U.S. states, with data ranging from the 1970s through October 2023. The dataset comprises annual crime and unemployment statistics, yielding a total of 1,989 observations.

# Sampling Method:

This study utilizes a convenience sample drawn from publicly available state-level datasets. While not random, the sample is comprehensive across all states and years, enhancing its generalizability.

# Bias Identification:

Potential biases include:

* **Temporal bias**: Economic and crime trends vary across decades.
* **Reporting bias**: Differences in crime reporting standards between states.
* **Seasonal/Environmental confounders**: Variables such as temperature or seasonality may influence crime and employment independently.

Mitigation strategies include examining trends over time, using linear regression with control variables, and identifying outliers.

# Survey Questions/Collected Data/Dataset:

The dataset used is a cleaned version of crimebystatecombinedwithunemployment.csv, containing the following variables:

* **Unemployment** – Percentage of the labor force that is unemployed and actively seeking work.
* **Violent Crime** – Incidents such as assault, robbery, and homicide.
* **Property Crime** – Incidents like burglary, theft, and motor vehicle theft.
* **Total Crime** – Sum of violent and property crime incidents.
* **Population** – Used for visual scaling in bubble plots.

**Number of Samples:** n = xx (replace with actual number when known)

# Analysis:

The relationship between unemployment and crime was explored using:

* **Scatterplots**: Showed positive trends between unemployment and both violent and property crimes.
* **Linear Regression**: Demonstrated a statistically significant positive association.
* **Boxplots and Time Series Plots**: Helped assess variation across time and states.
* **Bubble Plots**: Illustrated regional differences.
* **T-tests**: Confirmed statistically significant differences in crime levels between high- and low-unemployment states.

# Hypothesis Testing Steps

* **Defined null and alternative hypotheses.**
* **Divided the dataset into high and low unemployment groups.**
* **Calculated mean crime rates for each group.**
* **Applied T-tests to compare group means.**
* **Interpreted p-values to determine statistical significance (p < 0.05).**

**Results: The null hypotheses were rejected, indicating a significant association between higher unemployment and higher crime rates.**

# Conclusion

The analysis reveals a clear positive relationship between unemployment and crime in the U.S. While causality cannot be confirmed, the findings support the integration of employment initiatives within crime prevention strategies. Policymakers should consider these insights when designing economic and criminal justice reforms.

# Any potential issues

* **Data Quality**: The dataset would have been more accurate with a larger number of rows.
* **Confounding Variables**: Other factors not accounted for in the dataset, such as education, healthcare access, or poverty levels, may influence crime rates.
* **Generalizability**: The findings may not be applicable to all populations if the dataset is not representative.
* **Incomplete Data**: There was some missing data; the mode was used to fill missing values for specific metrics.
* **Data Volatility**: Due to the dataset's limited spread in numeric values, it was challenging to derive more decisive or nuanced statistical conclusions regarding the strength and nature of the correlation between unemployment and crime.