CAPSTONE KAROLINA NIKOLAJEW

**‘A Curious Case of Crime: The Impact of the Death Penalty and Unemployment on Crime Rates in Chicago, Houston, and U.S. States’**

**Executive Summary**

*This project will explore the differences in crime rates between* ***Chicago, IL*** *and* ***Houston, TX****, with a focus on understanding the impact of the death penalty. Houston, which has the death penalty, will be compared to Chicago, which does not, to see if there is a noticeable difference in crime patterns.*

*As a secondary analysis, crime data from all U.S. states will be analyzed, distinguishing between states that have the death penalty and those that do not.*

*The dataset for state-level crimes is already clean and will be used to visualize crime trends over the years, based on the presence or absence of the death penalty in each state.*

*Additionally, the project will incorporate unemployment data for all states to investigate the relationship between unemployment rates and crime rates, providing a broader understanding of the factors affecting crime. (stretch one)*

**Motivation**

*The motivation behind this project is to understand how the death penalty may affect crime rates. By comparing Chicago and Houston—one with the death penalty and the other without—we can learn more about how capital punishment might influence crime. Expanding the analysis to include all U.S. states will help us see if the presence or absence of the death penalty affects crime trends across the country. Additionally, adding unemployment data will allow us to explore how economic factors may also impact crime rates, giving a broader understanding of the issue.*

**Data Question**

*The central question of this project is: How does the presence of the death penalty in Houston affect crime rates, and are there noticeable differences in crime patterns compared to Chicago? Additionally, does the death penalty's presence at the state level impact crime trends in other U.S. states? Finally, how does unemployment correlate with crime rates across the U.S.? The aim is to identify any patterns or insights that could inform discussions on the effectiveness of the death penalty and the role of economic conditions in shaping crime.*

**Minimum Viable Product (MVP)**

*The final capstone will include cleaned and merged datasets for Chicago and Houston crimes using Python, along with merged unemployment and crime data at the state level. PowerPoint presentation will explain briefly what to expect, will go over crimes that end with death penalty, go over population numbers for Chicago and Houston and overall crime rates. Power BI visualizations will provide an in-depth analysis of crime patterns in both cities.*

*A separate map will show crimes per state, color-coded to indicate whether the state has the death penalty. The presentation is aimed at policymakers and law enforcement, offering insights into the impact of the death penalty and unemployment on crime rates.*

**Schedule (through <date of demo day>)**

1. Get the Data (03/29/2025)
2. Clean & Explore the Data (04/11/2025)
3. Create Presentation of your Analysis (04/18/2025)

* Should be a presentation, but could include a Jupyter Notebook or dashboard in Excel, Tableau, or PowerBI

1. Internal demos (<04/18/2025)
2. Demo Day!! <04/25/2025

**Data Sources**

<https://www.houstontx.gov/police/cs/Monthly_Crime_Data_by_Street_and_Police_Beat.htm> - houston dataset (needs to be cleaned up)

<https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2/about_data> - chicago dataset (needs to be cleaned up)

<https://www.kaggle.com/datasets/justin2028/unemployment-in-america-per-us-state?utm_source=chatgpt.com> – unemployment data (clean, need to merge)

<https://corgis-edu.github.io//corgis/csv/state_crime/> - crimes per state (clean already)

**Known Issues and Challenges**

*Several challenges are anticipated during the project:*

1. ***Categorizing Data:*** *Both Chicago and Houston crime datasets contain many categories that will need to be simplified into broader, meaningful groups for analysis.*
2. ***Data Cleaning and Merging:*** *The datasets from Chicago and Houston will require extensive cleaning, including renaming columns for consistency, and merging them into a single, unified dataset. Additionally, I will need to clean up and format the unemployment data for integration with the crime data.*
3. ***ZIP Code to County Mapping:*** *Since the datasets contain ZIP codes but not county information, I will need to map ZIP codes to their respective counties to enable county-level visualizations in Power BI.*
4. ***Time Extraction:*** *Extracting time of day and month from the date columns will be necessary to analyze crime patterns by time.*
5. ***Visualization Complexity:*** *Ensuring that the data visualizations in Power BI are clear and effectively represent the relationships between crime rates, the death penalty, and unemployment across multiple dimensions will be a challenge, especially with a large dataset.*
6. ***Crime Rates per Capita:*** *I have chosen Chicago and Houston as comparison since their population numbers are the most similar compared to overall crimes. Although, to present crime data accurately, I will need to account for population size. Crime numbers alone may not provide a clear picture without adjusting for the population in each area. To do this, I will likely need to add a new table that includes population data for each city or state. This will allow me to calculate crime rates per capita (e.g., crimes per 100,000 people) to ensure the data is presented in a meaningful way.*

*03/29/2025,*

*Karolina Nikolajew*