Problem Statement Worksheet (Hypothesis Formation)

How can Houston law enforcement enhance public safety through resource allocation or legislation by integrating weather information and optimizing the daily total and specific crime number prediction to a percentage error of <20% by a month?

H

1 Context

Houston faces significant challenges with crime and public safety as a bustling metropolitan area. Global warming adds further complexity and unpredictable results. The project seeks to provide data-backed insights for law enforcement to address weather-related crime spikes and enhance public safety during extreme weather events. The project looks to optimize the crime rate prediction to a percentage error of <20% by integrating weather information.

2 Criteria for success

- Identify the weather-related factors contributing to the crime rates in a month-long project, and be able to predict the daily crime numbers in a percentage error of <20%.

3 Scope of solution space

- Collect historical weather and crime data from 2010 to 2023
- Optimize the prediction of daily total crime number and specific crime numbers
- Interpret the model to public or implement causal inference on the incremental crime rate due to weather

4 Constraints within solution space

- Numerous record numbers (230M +) of crime
- Data inconsistency in formats and records

5 Stakeholders to provide key insight

- Houston police department

6 Key data sources

- Historical weather data from visualcrossing API
- Monthly crime reports from Houston Police Department
- [Optional] Unemployment rate and demographic data