UNITED STATES MILITARY ACADEMY
WEST POINT

Modeling the Influence of Socioeconomic and Temporal Factors on Burglary Rates in Chicago

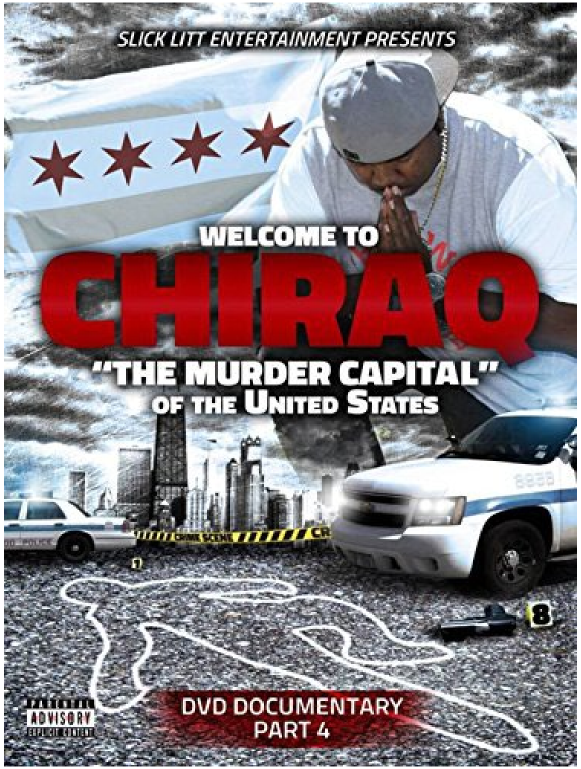
MA478: Final Project

CDT Garrett Chrisman
CDT Lucas Villanti

MATHEMATICAL ∞ SCIENCES ∞

Background

- "Chiraq": A Controversial Nickname
  - Chicago is often called "Chiraq" in hip-hop culture
  - Nickname reflects the city's crime and violence history
- Crime Statistics in Chicago
  - Burglary incidence has fluctuated historically
  - Poverty, inequality, limited education, and job opportunities are major contributing factors
- Neighborhood Disparities
  - Research indicates a spatial distribution of burglaries
  - Higher poverty and unemployment correlate with more burglaries



Problem Statement

Given the noticeable disparities in burglary rates across different Chicago neighborhoods, what model can accurately capture the impact of **seasonal changes**, **population**, and **wealth index** on **burglary counts**, and account for **spatial and temporal variability**?

Objectives

Research Questions:

1. What are the key socioeconomic and demographic factors that influence burglary rates in Chicago neighborhoods?

2. How do seasonal patterns, district socio-economic profiles, and wealth indices impact the spatial and temporal variability of robbery incidence in Chicago over a seven-year period?

Hypothesis:

1. Higher levels of poverty and population density in a Chicago neighborhood are positively associated with higher burglary rates.

2. The application of a mixed-effects model, accounting for both fixed (socioeconomic status, unemployment) and random (seasonal variations) effects, will provide a more accurate prediction of burglary rates across Chicago neighborhoods than models considering only fixed effects.

UNITED STATES MILITARY ACADEMY
WEST POINT

Methodology

1. Combine each indv. CSV into one combined one
2. Create a Total Crime (Each Loc.) Variable within Combined Df
3. Create a Seasonal Crime Count & Long Df to look at each Month more efficiently
4. Conduct Data Visualization
5. Model the Effects of Burglary Rates using Count Regressions
6. Use Model Selection Methods to decide which model fits problem best

MATHEMATICAL
SCIENCES

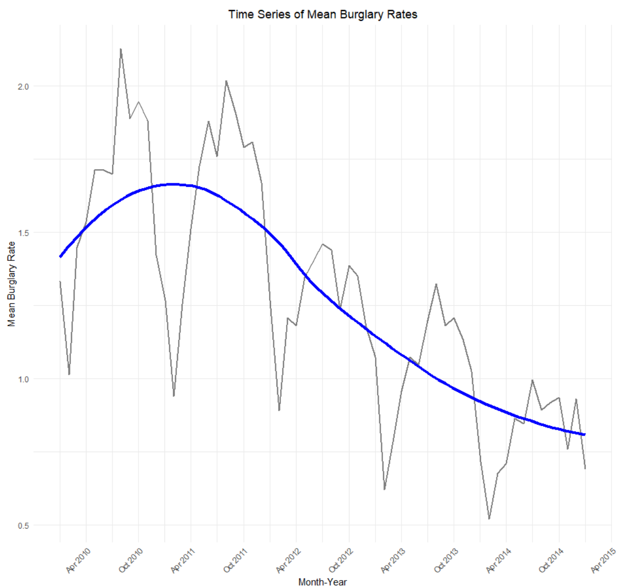
Data Description

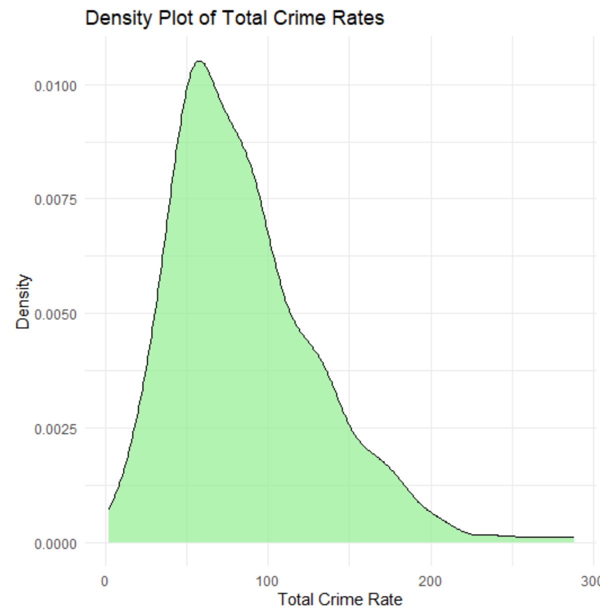
Chicago Burglary Dataset Overview
Source: Extracted from a dedicated GitHub repository.
Scope: Encompasses 552 unique locations within Chicago.
Variables:

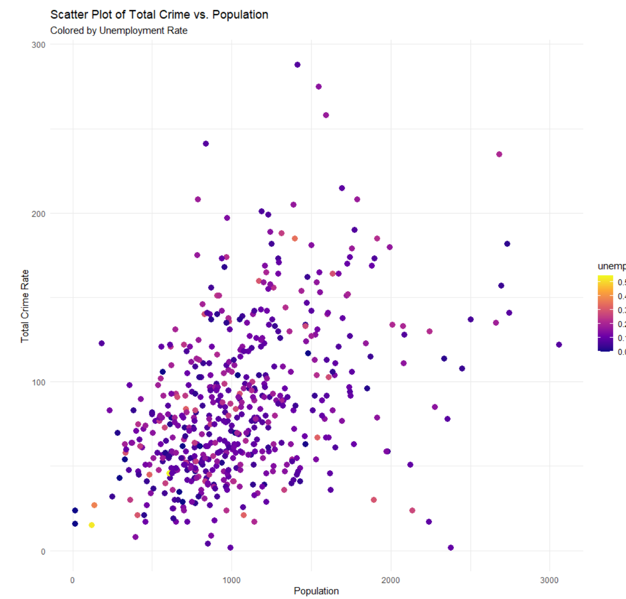
- **Population**: Headcount within each block.
- **Wealth Index**: A metric representing the economic status of the block's residents.
- **Young Male Demographic**: Count of young male individuals, indicative of the age and gender distribution.
- **Unemployment Rate**: Percentage or metric indicating the joblessness within each block
- **Total Crime:** The number of burglaries in each location block
  - **Counts of Seasonal Crime:** The number of burglaries in each season based on location
  - **Month/Year of Counted Crime:** The number of crime in each Month from Jan 2010 to Jan 2015

UNITED STATES MILITARY ACADEMY
WEST POINT

Data Exploration







MATHEMATICAL
∞ SCIENCES ∞

Three Models

1. **Poisson Regression Model**

2. **Mixed Effects Model**

3. **Chiraq Model (Bayesian Hierarchical Poisson)**

Poisson Regression

$$\log(\lambda_i) = \beta_0 + \beta_1 * Unemp_i + \beta_2 * young\ Males_i + \beta_3 * Total\ Summer\ Count_i + \beta_4 * wealth_i$$

- λ_i is the expected count of total crimes for the i^{th} Census Block Group.
- $Unemp_i$ is the unemployment rate for the i^{th} Census Block Group.
- $young\ Males_i$ is the young male population for the i^{th} Census Block Group.
- $Total\ Summer\ Count_i$ is the total number of crimes in the summer for the i^{th} Census Block Group.
- $wealth_i$ is the wealth index for the i^{th} Census Block Group.

Poisson Regression

Coefficients	Estimate	Std. Error	Z Value	P-value
Intercept	3.6887	0.0127	290.97	< 2e-16
Unemp	-0.1789	0.0659	-2.731	0.00667
Young Males	0.0044	0.0001	3.991	6.59e-0.5
Total Summer	0.0274	0.0003	91.878	<2e-16
Wealth	0.0364	0.0050	7.281	3.31e-13

Null Deviance: 12461
Residual Deviance: 2708.5
AIC: 6108.5

Mixed Effects

$$\log(\lambda_{ij}) = \beta_0 + \beta_1 \cdot unemp_{ij} + \beta_2 \cdot ym_{ij} + \beta_3 \cdot Total.Summer_{ij} + \beta_4 \cdot wealth_{ij} + \upsilon_i + \varphi_i + \varepsilon_{ij}$$

- $Y_{it} \sim Poisson(\lambda_{ij})$
- λ_{ij} is the expected count of Total Crime, for the i^{th} Census block group
- $\beta_1, \beta_2, \beta_3, \beta_4$ are the fixed effects coefficients
- υ_i is the random effect for Census block location
- φ_i is the random effect for month

**The random effects are assumed to be normally distributed with a mean of zero and their own variances

Mixed Effects

Coefficients	Estimate	Std. Error	Z Value	P-value
Intercept	3.472424	0.035122	98.784	< 2e-16
Unemp	-0.311086	0.1828	-1.702	0.08880
Total Summer	0.035432	0.001038	34.134	<2e-16
Wealth	0.039593	0.014310	2.767	0.00566

Random Effects	Variance	Std. Dev
Location	1.027e-01	3.205e-01
Month	1.455e-42	1.206e-21

Chiraq Model

$$\eta_{it} = B_o + B_1 \cdot unemployment_{it} + B_2 \cdot wealth_{2i} + B_3 \cdot total.summer_{it_{year}} + u_i + v_{t_{year}} + w_{t_{month}}$$

$Y_{it} \sim Poisson(\lambda_{it})$
$\eta_{it} = \log(\lambda_{it})$

u_i is the random effect for district i – iid

$v_{t_{year}}$ is the random effect for year t – rw1

$w_{t_{month}}$ is the random effect for month t – *rw1*

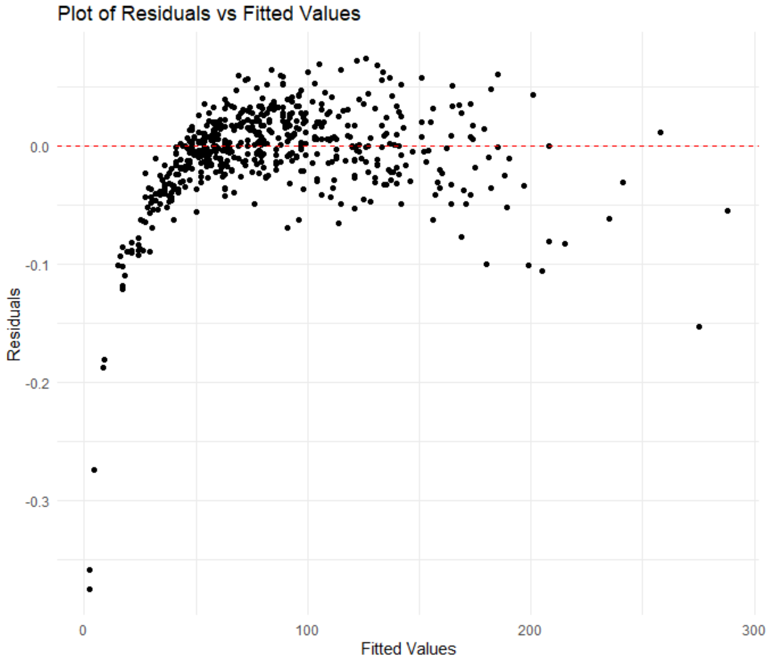
B_o is the intercept

B_1 is the coefficient for the unemployment rate

B_2 is the coefficient for the wealth index

B_3 is the coefficient for the total summer crime counts for each district, i, and each year

The Chiraq Model



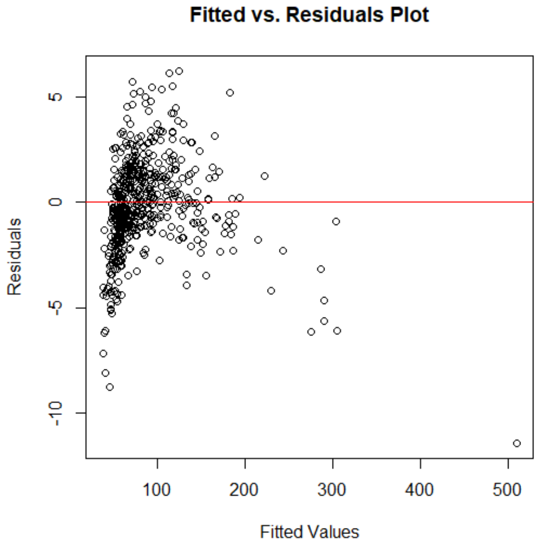
Term	Estimate
(Intercept)	3.472
unemp	-0.311
wealth	0.040
Total.Summer	0.035

DIC: 245211.55 – whoa

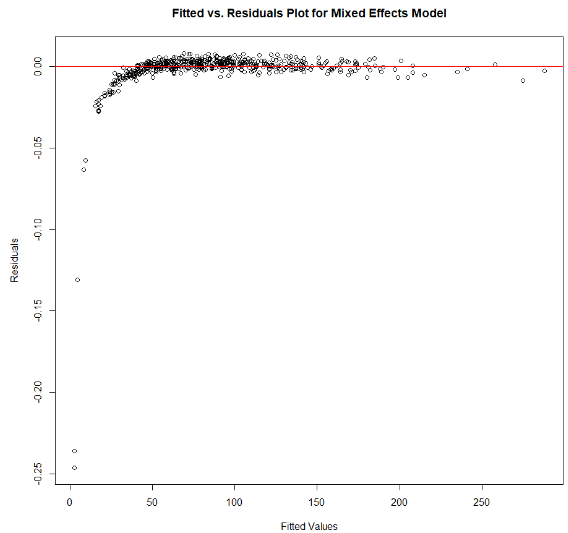
Name	Mean
Precision for ID	9.70
Precision for Year	63975.36
Precision for Month	94964.16

$$precision = \frac{1}{variance}$$

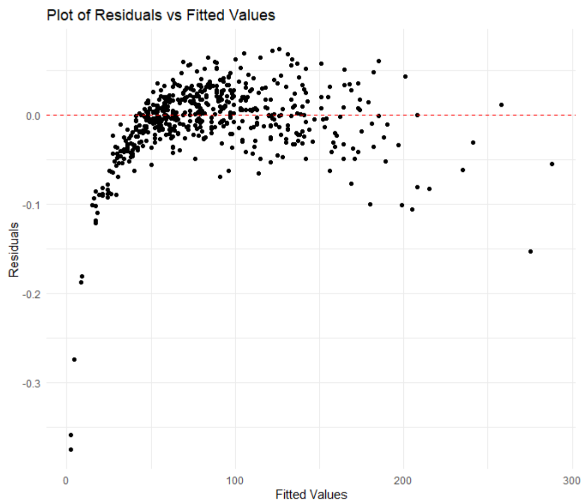
Model Selection



Poisson Regression



The Mixed Effects Model



The Chiraq Model

Importance

Trend Analysis: A year-over-year decrease in burglary rates. Over time, possible factors such as improvements in law enforcement, community initiatives, or socioeconomic changes might be contributing to a decline in burglary occurrences.

Seasonal Effects: Seasonal pattern, with summer months, showing higher burglary rates. How could we mitigate the increase in instances?

Influential Factors: The model highlights that unemployment and wealth are significant predictors of burglary rates, implying that economic factors play a crucial role in crime dynamics.

Spatial Variation: Spatial random effects accounts for unobserved heterogeneity across different census block groups, acknowledging that some areas may be inherently more prone to burglaries due to factors not explicitly included in the model.

Conclusions

Model Limitations:

- Poisson assumes mean and variance are equal - Overdispersion
- Random Walk Assumption in year and month assumes smooth temporal effects
- Predictive Power: The model is doubtful in its predictive abilities for burglary rates

Future Work

- Look at other model families to address above limitations
- The effect of school vacation periods