Exploring the Relationship Between Incarceration Rates and Reported Mental Health per Census Tract in Chicago (2022)

Duaa Alzouby & Layan Qaimari

Problem Statement & Objectives

Mass incarceration has emerged as a significant public health issue with documented consequences extending beyond incarcerated individuals to affect entire neighborhoods and communities. While prior research has established connections between incarceration and adverse health outcomes at the individual and family level, less is known about how incarceration rates at the spatial neighborhood level correlate with population health outcomes. This study investigates **whether neighborhood-level incarceration rates in Chicago are associated and disproportionately impact neighborhood mental health rates, and if so, to what extent?**

Understanding this relationship has important implications for public health policy and urban planning. Evidence of a neighborhood-level link between incarceration and mental health could inform targeted interventions and policy recommendations that address the interconnected challenges of criminal justice and public health in disproportionately affected communities, and can be used to address health disparities in further research.

Literature Review

Individual and Family-Level Effects

Existing research demonstrates that jail incarceration produces detrimental mental and physical health consequences extending beyond incarcerated individuals to their families and immediate social networks. The prevalence of jail incarceration is unequally distributed across race, sex, socio-economic levels, education attainment levels, and neighborhoods, disproportionately affecting Black men with low-income levels and less than a high-school degree (Turney & Conner, 2019). At the individual level, incarceration exacerbates pre-existing health conditions and is associated with increased risk of geriatric and chronic diseases (Garcia-Grossman et al., 2023).

The Spillover Effect Framework

The spillover effect describes how partners, children, and household members of incarcerated individuals experience health and social consequences (including elevated stress, anxiety, depression, and physical health deterioration) resulting from cumulative social and economic disruptions (Schnittker, Massoglia, & Uggen, 2011). This operates through three proposed pathways:

1. **Economic Disruption:** Incarceration removes wage earners while families incur substantial costs (bail, legal fees, commissary expenses), creating healthcare access barriers, nutritional insecurity, and housing instability.
2. **Psychosocial Stress:** Family separation generates chronic stress that activates allostatic

load, manifesting as elevated blood pressure, cardiovascular disease, metabolic dysfunction, and immune compromise.

3. **Social Isolation:** Incarceration stigma leads to social exclusion and damaged networks, reducing access to support systems and depleting neighborhood social capital.

Neighborhood-Level Effects

Multiple studies indicate jail incarceration has a relationship with increased mortality and morbidity at the county level (Kajeepeta et al. 2020). When incarceration becomes endemic to a neighborhood, the spillover effect operates cumulatively, generating health consequences that exceed individual-level predictions through three mechanisms:

Concentrated Economic Disadvantage: Multiple households simultaneously experience income loss and reduced employment prospects, creating neighborhood-level economic disinvestment—reduced consumer spending, business closures, deteriorating infrastructure—that affects all residents through reduced healthcare access and chronic environmental stress.

Family and Social Disruption at Scale: Widespread family disruption undermines neighborhood institutions (schools, religious organizations, community groups) and overwhelms informal support systems, reducing collective efficacy and increasing mortality from multiple causes.

Institutional Degradation: High-incarceration neighborhoods experience institutional disinvestment. Community trust deteriorates, public health organizations face increased demand with reduced resources, and schools experience increased trauma and resource strain, diminishing all residents' access to health-promoting institutions.

Census Tract Analysis & Methods Matters

County-level analysis masks geographic heterogeneity and obscures where incarceration concentrates. Census tract analysis (neighborhoods of 1,500–8,000 residents) provides critical advantages: it captures the spatial scale at which spillover effects operate directly, enables identification of incarceration and health "hotspots," allows comparison across neighborhoods with similar demographic characteristics, and yields findings with direct relevance for neighborhood-level intervention. Similar to the methods used in LeMasters et al. (2023) to explore the spatial autocorrelation of the relationship between probation and mental health at the county level, we aim to use Moran's I and a Spatial Regression in our census tract analysis. Critically, census tract analysis allows researchers to control for confounders at the same geographic scale as outcomes, substantially improving causal inference.

Research Gap

Most of the existing literature focuses on individual- or family-level impacts, while neighborhood-specific analyses remain limited. This research addresses that gap by examining the association between incarceration and mental health at the census tract level in Chicago—the precise geographic scale where spillover effects materialize and where policy operates.

Data Collection & Study Area Maps

Public Health Statistics (Source: Chicago Health Atlas): Measures the reported mental health per census tract as the percent of resident adults aged 18 and older who report 14 or more days in the past 30 days when their mental health was not good | Year: 2022 | *Format: Excel*

Incarceration Data (Source: Opportunity Atlas): The proportion of the tract population that is incarcerated as defined by the Decennial Census | Year: 2010 | *Format: CSV*

Census and Control Variable Data (Source: American Community Survey 5-year Estimate 2019-2023):

- ➢ Census Tract geographic boundaries | Year: 2019 | *Format: Shapefile*

Control variables | *Format: Shapefile*

- ➢ **Poverty:** The proportion of the population that falls under the poverty line (0-.99)
- ➢ **Wealth:** The median household income per census tract
- ➢ **Race:** The proportion of the population that identifies as black or African American
- ➢ **Unemploymen**t: The proportion of the aged 16+ population that is unemployed
- ➢ **Educational Attainment:** The proportion of the aged 25+ population that has a high school diploma or less
- ➢ **Age/Sex:** The proportion of the population that is male aged 18-34

For all the control variables, we normalized the data to be a percent of the total population (or relevant population), deleted all irrelevant fields, and clipped to the Chicago Geographic Boundaries

Control Variables

Control variables address established confounders from public health and incarceration literature:

Median Household Income & Poverty Rates: Income fundamentally determines health through healthcare access, housing quality, food security, and chronic stress. Both incarceration and mortality are stratified by income; without controlling for it, observed associations may reflect underlying socioeconomic disparities rather than incarceration effects.

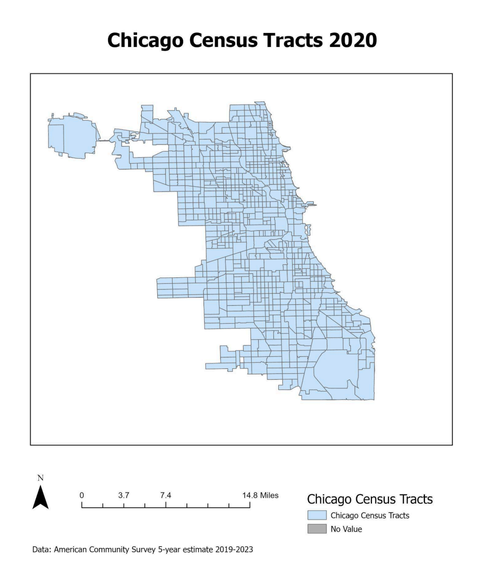
Unemployment Rate: Associated with both increased incarceration and increased mortality through economic stress and reduced healthcare access. Controlling for unemployment isolates incarceration's independent effect.

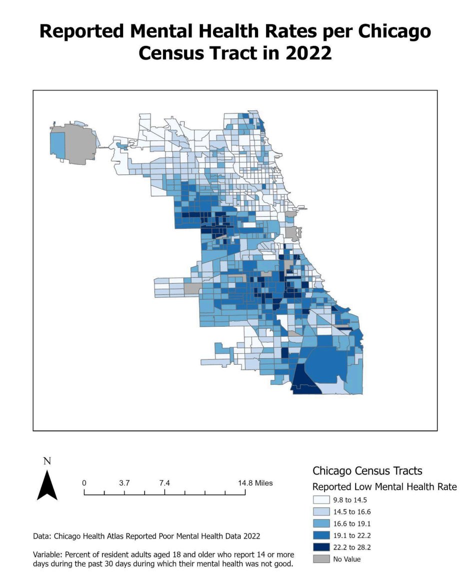
Educational Attainment: Independently associated with both incarceration risk and health outcomes through health literacy, occupational status, and social networks. Education operates through distinct mechanisms from income.

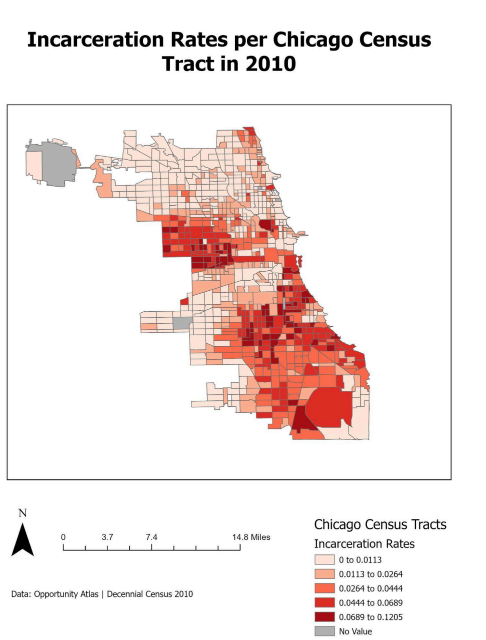
Race/Ethnicity Composition: Black individuals experience disproportionately high incarceration rates and mortality from cardiovascular disease, diabetes, and violence, driven by systemic inequities. Controlling for racial composition allows estimation of the incarceration-mortality relationship within neighborhoods of comparable composition.

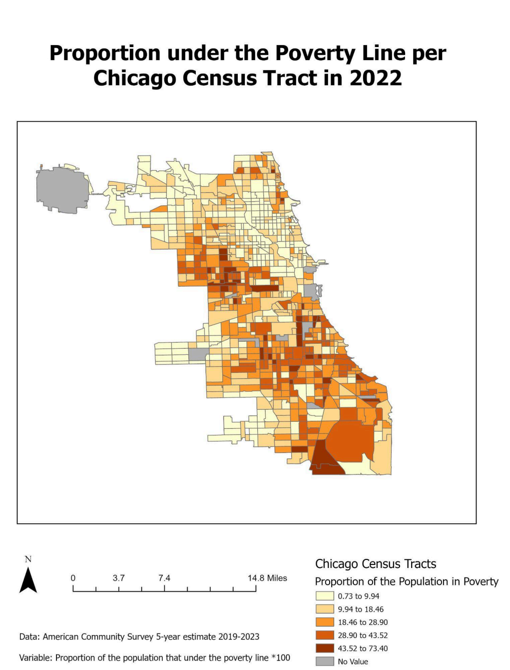
Age & Sex Composition: Age structure affects incarceration patterns (young adults have higher incarceration rates) and ensures observed associations don't reflect demographic composition.

Study Area Map



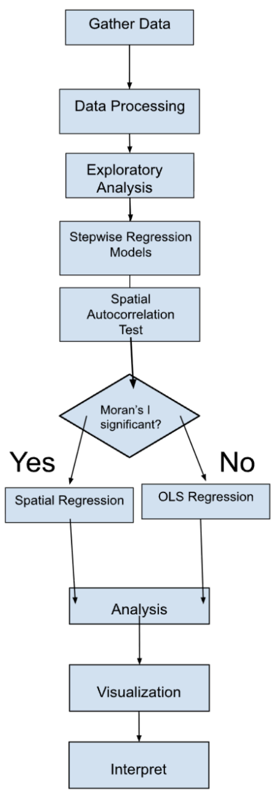






Our study area consists of 831 census tracts within Chicago city boundaries (see Figure D.1 in Appendix D). Figure D.2 displays the spatial distribution of reported mental health outcomes across Chicago showing our dependent variable with rates ranging from 9.8% to 28.2%. Figure D.3 shows incarceration *rates from 2010, our key independent variable, with rates ranging from 0 to 120.5 per 1,000 residents. Control variables including poverty rates (Figure D.4), racial composition (Figure D.5), and educational attainment (Figure D.6) show similar geographic clustering on Chicago's South and West sides. The bivariate choropleth map (Figure D.7) visually demonstrates the spatial co-occurrence of high incarceration and poor mental health outcomes in these same neighborhoods.*

Methodology



Preliminary Processing

To complete the data processing and preliminary steps before analysis, we normalized and clipped the data to our study area of interest, joined the variables to the census tract shapefile, and reprojected the layers and map to NAD83 Chicago (feet). We created exploratory maps and tables of our variables, explaining the spatial and numeric distribution of each variable of interest.

Stepwise Regression Approach

We used a Stepwise Regression Approach to analyze the correlation and strength of the relationship between incarceration and mental health rates while controlling for various factors. We use 3 models to run this analysis:

Model 1 — Bivariate: $MR_t = \beta_0 + \beta_1(IR_t) + \epsilon_t$

Estimates the unadjusted association between incarceration and mortality.

Model 2 — Socioeconomic Adjustment:

$$MR_t = \beta_0 + \beta_1(IR_t) + \beta_2(MedianIncome_t) + \beta_3(UnemploymentRate_t) + \epsilon_t$$

Tests whether the relationship persists after accounting for economic conditions.

Model 3 — Full Demographic Adjustment:

$$MR_t = \beta_0 + \beta_1(IR_t) + \beta_2(MedianIncome_t) + \beta_3(UnemploymentRate_t) + \beta_4(ControlVariables_t) + \epsilon_t$$

Controls for all major confounders. Comparing β_1 across models reveals attenuation patterns and identifies which factors confound the relationship.

Spatial Analysis

Moran's I Test: A Global Moran's I Test was used to calculate the spatial autocorrelation on Model 3 residuals. Significant results ($p < 0.05$) indicate neighboring tracts have similar residuals, suggesting omitted spatial variables or spillover effects.

Spatial Regression: As Moran's I was found to be significant, we employed Geographically Weighted Regression (GWR) to produce unbiased estimates and correct standard errors.

Residual Mapping: Visualizes spatial patterns in prediction errors, identifying neighborhoods requiring additional investigation.

Spatial Analysis Parameters and Settings:

Data Preprocessing:

- Projection: NAD 1983 StatePlane Illinois East FIPS 1201 (US Feet)
    - Rationale: Maintains accurate distance measurements for Chicago metropolitan area and ensures compatibility with local planning datasets.
- Spatial Join Method: FIPS field matching between census tract boundaries and demographic/incarceration data tables.
- Geographic Clipping: Census tracts clipped to Chicago city boundaries to focus analysis on municipal jurisdiction.

Ordinary Least Squares Regression:

- Input Feature Class: Chicago Census Tracts
- Unique ID: FIPS code
- Dependent Variable: Reported Mental Health (percentage reporting 14+ poor mental health days)
- Model 1 Explanatory Variables: Incarceration rate only
- Model 2 Explanatory Variables: Incarceration rate, median income, unemployment rate
- Model 3 Explanatory Variables: Incarceration rate, median income, unemployment rate, proportion under BA (education), proportion male aged 18-34, proportion Black/African American, total population
- Rationale: Stepwise inclusion of control variables reveals whether the incarceration-mental health relationship persists after accounting for confounders.

Global Moran's I (Spatial Autocorrelation):

- Input Feature Class: Model 3 output
- Input Field: Standard Residuals
- Conceptualization of Spatial Relationships: Contiguity Edges Only
- Standardization: Row standardization
- Rationale: We chose contiguity edges because it captures spillover effects between adjacent census tracts. Row standardization accounts for varying numbers of neighbors across tracts. Testing residuals rather than raw values identifies spatial patterns unexplained by demographic controls.

Hot Spot Analysis (Getis-Ord Gi*):

- Input Feature Class: Model 3 output
- Input Field: Standard Residuals
- Conceptualization of Spatial Relationships: Contiguity Edges Only
- False Discovery Rate (FDR) Correction: Applied
- Rationale: FDR correction controls for multiple testing when identifying statistically significant clusters. Analyzing residuals identifies where the model under-predicts or over-predicts mental health outcomes, highlighting localized spillover effects.

Geographically Weighted Regression Settings:

- Input Feature Class: Chicago Census Tracts (Model 3 variables)
- Kernel Type: Adaptive (Gaussian)
- Bandwidth Method: AICc (corrected Akaike Information Criterion) - this optimized automatically
- Dependent Variable: Reported Mental Health
- Explanatory Variables: Same as Model 3
- Neighborhood type: Number of neighbors
- Neighborhood selection: Golden Search
- Rationale: Adaptive kernel accounts for varying census tract densities across Chicago. GWR tests whether the incarceration-mental health relationship varies geographically, identifying neighborhoods where spillover effects are strongest.

Descriptive Mapping:

- Classification Method for Univariate Maps: Jenks Natural Breaks
    - Rationale: Natural breaks classification identifies inherent groupings in the data by minimizing within-class variance and maximizing between-class variance. This shows natural clusters in incarceration rates, mental health outcomes, and demographic variables across Chicago neighborhoods.
- Classification Method for Bivariate Map (Figure D.7): Quantile (equal count per class)
    - Rationale: Quantile breaks for the bivariate choropleth make sure that there is a balanced visual representation, with each class containing equal numbers of census tracts. This facilitates the identification of co-occurrence patterns between high or low incarceration and high or low mental health outcomes.

Hypotheses

Null Hypothesis: No significant relationship between neighborhood incarceration rates and mental health rates at the census tract level, after accounting for confounders.

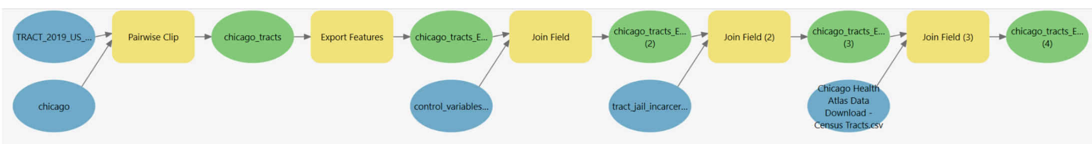
Alternative Hypothesis: Higher neighborhood incarceration rates are associated with higher mental health rates, independent of income, unemployment, education, race/ethnicity, and age composition.

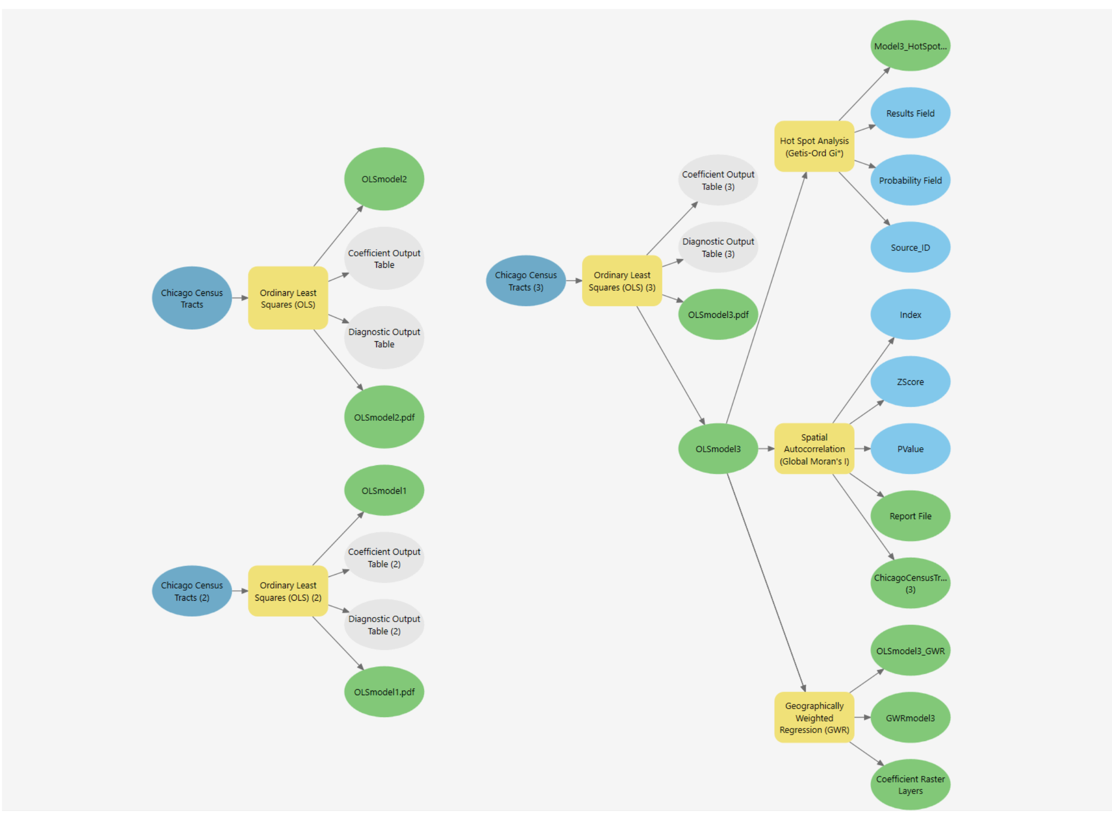
We anticipated stronger relationships in economically disinvested neighborhoods and communities of color, reflecting cumulative systemic disadvantage.

Key Assumptions:

- The relationship between 2010 incarceration and 2022 mental health reflects long-term community-level impacts rather than immediate effects
- Self-reported mental health, while being subjective, captures meaningful variation in community well-being
- Census tracts are appropriate units for measuring neighborhood-level spillover effects
- Control variables address major confounders, though unmeasured factors (ex., Policing intensity, historical trauma) may remain
- Spatial relationships are stationary within contiguous neighborhoods but may vary across the city

Project Implementation





This shows our complete workflow in Model Builder, FIgures F.1-2: three regression models feeding into spatial autocorrelation tests, hot spot analysis, and GWR. It is fully automated and reproducible for other cities.

Results & Interpretation

Descriptive Statistics & Maps

Alias	Nulls	Chart Preview	Min	Max	Mean	Std. Dev.	Median	Count	Unique	Outliers	Sum	Range	IQR	Q1	Q3	CV	Skewness	Kurtosis
Proportion Black	23 (2.7%)		0	100	32.689781	37.354689	8.919406	831 (97.3%)	801	0	27,165.20781	100	71.674905	2.613823	74.288728	1.142702	0.734743	1.795815
Proportion under the poverty line	23 (2.7%)		0.732535	73.401878	17.769778	12.53505	14.637904	831 (97.3%)	831	18	14,766.685657	72.669342	16.735708	8.014162	24.74987	0.705414	1.19738	4.462021
Proportion without a BA	23 (2.7%)		5.548345	98.068077	60.243793	26.210507	67.021277	831 (97.3%)	831	0	50,062.592093	92.519732	43.769594	39.545823	83.315416	0.435074	-0.534502	1.996624
Proportion Unemployed	23 (2.7%)		0	48.909657	9.213234	7.369508	6.86747	831 (97.3%)	822	36	7,656.197081	48.909657	8.559323	3.954329	12.513652	0.799883	1.492062	5.531166
Proportion of Residents that are Male 18-34	23 (2.7%)		1.158717	53.5183	13.774737	6.556304	12.479871	831 (97.3%)	829	32	11,446.806376	52.359583	7.620596	9.201314	16.821909	0.475966	1.348174	6.066562
Incarceration Rate	6 (0.7%)		0	0.1205	0.025718	0.024846	0.0166	848 (99.3%)	454	13	21.8091	0.1205	0.035175	0.0058	0.040975	0.966082	1.153668	3.741424
Rerported Mental Health	65 (7.6%)		9.8	28.2	17.135995	3.130507	16.6	789 (92.4%)	145	8	13,520.3	18.4	4.4	14.7	19.1	0.182686	0.680679	3.109297
median income	33 (3.9%)		13,489	244,286	78,057.113276	40,724.127121	70,714	821 (96.1%)	811	33	64,084,890	230,797	47,347	49,017	96,364	0.521722	1.150928	4.494718

Table C.1 presents descriptive statistics for all variables across Chicago census tracts. Mental health rates range from 9.8% to 28.2% (mean = 16.4%), while incarceration rates vary from 0 to 120.5 per 1,000 residents (mean = 21.8). The variation in both outcome and predictor variables, combined with geographic clustering patterns (see Figures D.2-D.3), provides an appropriate context for examining neighborhood-level relationships.

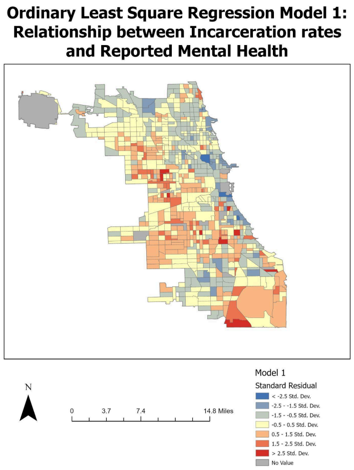
High incarceration rates and high mental health rates are clustered in the west and south sides of Chicago, as indicated in Figures D.2 & D.3, and the bivariate map Figure D.7. The areas with high incarceration rates and high mental health rates are not distributed randomly but instead are heavily correlated to the spatial makeup of Chicago. Through observational analysis of other variables in Figures D.4-D.6, these clusters correspond to a higher concentration of poverty, a high proportion of residents who identify as Black and African American, and a high proportion of residents with lower educational attainment.

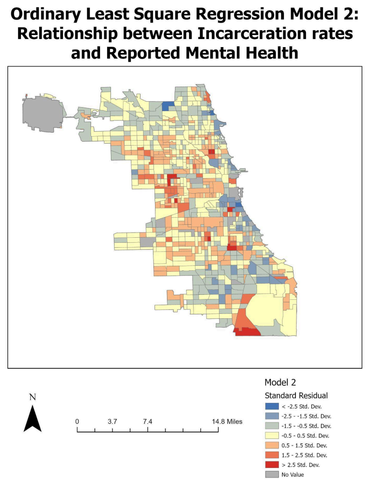
Stepwise Regressions

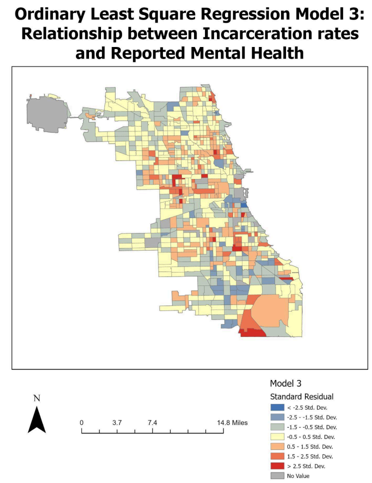
Table A.1 presents the bivariate regression results showing the unadjusted relationship between incarceration and mental health. Model 1 demonstrates a significant positive relationship ($\beta = 131.71$, $p < 0.001$, $R^2 = 0.38$), indicating that incarceration alone explains 38% of variation in mental health outcomes.

Table A.2 shows that after adding socioeconomic controls (median income, unemployment), the incarceration coefficient remains highly significant ($\beta = 37.62$, $p < 0.001$) and model fit improves substantially ($R^2 = 0.61$). This 28% increase in explained variance suggests that economic factors are important confounders.

Table A.3 presents the full demographic model. With all controls included (race, education, poverty, age/sex), the incarceration coefficient remains stable ($\beta = 37.40$, $p < 0.001$, $R^2 = 0.69$), demonstrating the relationship is independent of demographic composition. The minimal reduction of the incarceration coefficient across models indicates a stable association.







Figures E.2-E.4 display the spatial distribution of standardization residuals for all three models. The OLS Regression maps show the standard residual based on census tract for each of the three models, with darker blue and red tracts indicating a tract with a stronger relationship between incarceration rates and mental health rates. However, there are a few differences seen in the maps per tract and the spatial distribution of standard residuals. However, this does not particularly encompass the whole relationship as it does not include the changes and distribution of the coefficient with the changes of each model. There is no obvious clustering or patterns within the maps, except on the West Side of Chicago; however, a stepwise regression does not encompass the influence of neighboring tracts on the relationship of the variables in a specific tract, as does the Geographic Weighted Regression.

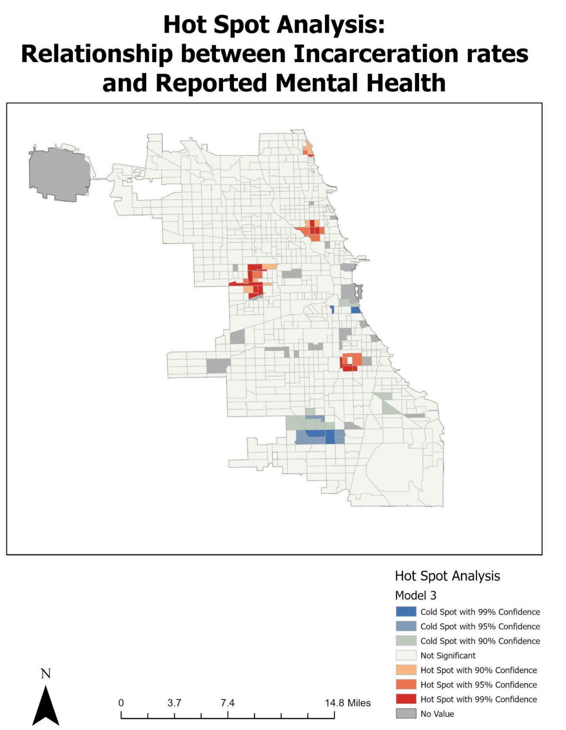
Spatial Autocorrelation Results:

Figure B.1: *Global Moran's I results showing spatial clustering*

Global Moran's I Summary

Moran's Index	0.235677
Expected Index	-0.001292
Variance	0.000428
z-score	11.452283
p-value	0.000000

We tested for spatial clustering using Moran's I on Model 3 residuals. The results were highly significant - Moran's I of 0.236 with p-value less than 0.001.



In Figure E.1, hot spot areas show clustering of statistically significant areas of high incarceration and high mental health proportions, while cold spots show clustering of low incarceration and low mental health reports. The map indicates that the hot spot areas are on the South and West sides of Chicago, where both mental health and incarceration rates are high. However, the hot spot clustering in the north side of Chicago does not correspond with the

Geographic Weighted Regression Results

The maps, Figure E.5 and Figure E.6, show that the incarceration-mental health relationship is strongest on the South and West sides, where incarceration rates are already highest. The relationship had a higher t-statistic and had a higher coefficient in the areas on the south and west side. This indicates that the relationship between incarceration and mental health was statistically more significant. Figure E.6 shows that a higher coefficient indicates that with each additional incarcerated person per 1,000 residents, the coefficient predicts percentage point increases in adults with poor mental health. In the north side of Chicago, 1 additional incarcerated person per 1,000 residents predicts a -0.20 to 10.47 percentage point change in the proportion of adults reporting poor mental health, while in the south side, this number is 28.70 to 38.61.

Discussions & Conclusion

Our findings strongly support the hypothesis. Incarceration significantly predicts mental health, while controlling for poverty, race, education, age, sex, and income. The spatial analysis shows clear clustering on Chicago's South and West sides, with the strongest effects where incarceration is already highest. This shows that this is a public health crisis with neighborhood-level spillover effects.

Project Evaluation & Future Work

What Worked Well:

- *Using stepwise regressions successfully demonstrated the stability of the incarceration-mental health relationship. The coefficient remained stable (*$\beta \approx 37.6$*) across all three models despite adding multiple control variables, providing strong evidence that the relationship is not a coincidence.*
- *The model builder will allow us to automate data preprocessing and analysis steps, ensuring reproducibility in the future.*
- *The combination of global (Moran's I) and local (Hot Spot Analysis, GWR) spatial statistics showed the geographic variance in the relationship. GWR results also showed that spillover effects concentrate in already disadvantaged neighborhoods, directly supporting our hypothesis.*
- *Using census tract level for analysis captured the neighborhood-level spillover effects better than county-level studies while maintaining sufficient sample size (N=775) for a reliable statistical inference.*

Limitations

One of our limitations is the structure and values of the reported mental health data. Some tracts have missing data, which may affect the GWR analysis and the neighbors of certain tracts. The data is also reported mental health, which may lose many nuances as it is not fully representative. In addition, there may be various confounding variables we have not accounted for that impact both incarceration rates and mental health rates, and the relationship between them. We are also limited to the Chicago Census Tracts and, therefore, cannot generalize our results.

Further research

In further research, generalizability and robustness should be tested thoroughly to establish the association.

1. Robustness Checks: Outlier analysis using Cook's distance, sensitivity testing by sequentially removing controls, alternative outcome measures (mental health diagnosis), and sensitivity analysis by using a different neighbor metric in our analysis
2. Functional Form Testing: Tests whether the relationship is linear or polynomial using residual plots and AIC/BIC criteria.
3. Interaction Effects: Tests whether incarceration affects mental health differently by income level (IR x) or racial composition (x), identifying whether effects concentrate in vulnerable areas.

References

Garcia-Grossman, I. R., Cenzer, I., Steinman, M. A., & Williams, B. A. (2023). History of incarceration and its association with geriatric and chronic health outcomes in older adulthood. *JAMA Network Open, 6*(1), e2249785. https://doi.org/10.1001/jamanetworkopen.2022.49785

Kajeepeta, S., Rutherford, C. G., Keyes, K. M., El-Sayed, A. M., & Prins, S. J. (2020). County jail incarceration rates and county mortality rates in the United States, 1987–2016. *American Journal of Public Health, 110*(Suppl. 1), S109–S115.

LeMasters, K., Delamater, P., Brinkley-Rubinstein, L., Edwards, J. K., Robinson, W. R., & Pence, B. (2023). Mass probation: Temporal and geographic correlation of county-level probation rates & mental health in North Carolina. SSM. Mental health, 3, 100189. https://doi.org/10.1016/j.ssmmh.2023.100189

Schnittker, J., Massoglia, M., & Uggen, C. (2011). Incarceration and the health of the African American community. *Du Bois Review: Social Science Research on Race, 8*(1), 133–141. https://doi.org/10.1017/S1742058X11000026

Turney, K., & Conner, E. (2019). Jail incarceration: A common and consequential form of criminal justice contact. *Annual Review of Criminology, 2*, 265–290.

Appendix

Appendix A: Complete Regression Model Outputs

Summary of OLS Results

Variable	Coefficient[a]	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_Pr[b]
Intercept	15.119481	0.127174	118.888193	0.000000*	0.113066	133.722249	0.000000*
TRACT_JAIL_INCARCERATION_RATES.CSV.INCARCERATION_RATE_RP_GP_PALL	77.353312	3.505791	22.064442	0.000000*	3.859171	20.044025	0.000000*

OLS Diagnostics

Input Features	Chicago Census Tracts	Dependent Variable	CHICAGO HEALTH ATLAS DATA DOWNLOAD - CENSUS TRACTS.CSV.PMH_2022
Number of Observations	785	Akaike's Information Criterion (AICc)[d]	3642.415629
Multiple R-Squared[d]	0.383387	Adjusted R-Squared[d]	0.382599
Joint F-Statistic[e]	486.839614	Prob(>F), (1,783) degrees of freedom	0.000000*
Joint Wald Statistic[e]	401.762935	Prob(>chi-squared), (1) degrees of freedom	0.000000*
Koenker (BP) Statistic[f]	28.676139	Prob(>chi-squared), (1) degrees of freedom	0.000000*
Jarque-Bera Statistic[g]	37.858850	Prob(>chi-squared), (2) degrees of freedom	0.000000*

Notes on Interpretation

*	An asterisk next to a number indicates a statistically significant p-value (p < 0.01).
a	Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.
b	Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.
c	Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.
d	R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.
e	Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Wald Statistic to determine overall model significance.
f	Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.
g	Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are biased (the residuals are not normally distributed).

Table A.1: OLS Model 1 Output Table
Bivariate regression results showing the unadjusted relationship between incarceration rates and mental health outcomes. R squared = 0.38, indicating incarceration alone explains 38% of variation in mental health (N = 785).

Summary of OLS Results

Variable	Coefficient[a]	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_Pr[b]	VIF[c]
Intercept	17.570337	0.265569	66.161079	0.000000*	0.297782	59.004012	0.000000*	--------
CHICAGO_TRACTS_EXPORTFEATURES.PROP_UNEMPLOYMENT1	0.099666	0.012521	7.960151	0.000000*	0.015412	6.466651	0.000000*	1.774363
CHICAGO_TRACTS_EXPORTFEATURES.MEDIAN_INCOME_1	-0.000030	0.000002	-14.059226	0.000000*	0.000003	-12.001124	0.000000*	1.588871
TRACT_JAIL_INCARCERATION_RATES.CSV.INCARCERATION_RATE_RP_GP_PALL	37.621184	3.455251	10.888119	0.000000*	3.900520	9.645170	0.000000*	1.518531

OLS Diagnostics

Input Features	Chicago Census Tracts	Dependent Variable	CHICAGO HEALTH ATLAS DATA DOWNLOAD - CENSUS TRACTS.CSV.PMH_2022
Number of Observations	775	Akaike's Information Criterion (AICc)[d]	3239.522550
Multiple R-Squared[d]	0.610327	Adjusted R-Squared[d]	0.608810
Joint F-Statistic[e]	402.526876	Prob(>F), (3,771) degrees of freedom	0.000000*
Joint Wald Statistic[e]	1015.516997	Prob(>chi-squared), (3) degrees of freedom	0.000000*
Koenker (BP) Statistic[f]	23.156887	Prob(>chi-squared), (3) degrees of freedom	0.000037*
Jarque-Bera Statistic[g]	24.132974	Prob(>chi-squared), (2) degrees of freedom	0.000006*

Notes on Interpretation

*	An asterisk next to a number indicates a statistically significant p-value (p < 0.01).
a	Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.
b	Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.
c	Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.
d	R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.
e	Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Wald Statistic to determine overall model significance.
f	Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.
g	Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are biased (the residuals are not normally distributed).

Table A.2: OLS Model 2 Output Table

Regression results after adding socioeconomic controls (median income, unemployment). R squared increases to 0.61, and incarceration coefficient remains highly significant ($\beta = 37.62, \rho < 0.001$*), controlling for economic factors (N = 775).*

Summary of OLS Results

Variable	Coefficient[a]	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_Pr[b]	VIF[c]
Intercept	11.597560	0.492256	23.560030	0.000000*	0.536061	21.634784	0.000000*	--------
CHICAGO_TRACTS_EXPORTFEATURES.PROP_BLACK1	0.000193	0.002930	0.066012	0.947370	0.003490	0.055417	0.955806	3.104389
CHICAGO_TRACTS_EXPORTFEATURES.PROP_UNDERBA1	0.054407	0.004030	13.500101	0.000000*	0.003973	13.692697	0.000000*	2.870330
CHICAGO_TRACTS_EXPORTFEATURES.PROP_UNEMPLOYMENT1	0.081070	0.012030	6.738902	0.000000*	0.015593	5.199176	0.000001*	2.049633
CHICAGO_TRACTS_EXPORTFEATURES.PROP_MALE18TO341	0.099997	0.011356	8.805474	0.000000*	0.011760	8.503483	0.000000*	1.397430
CHICAGO_TRACTS_EXPORTFEATURES.MEDIAN_INCOME_1	-0.000011	0.000003	-4.347162	0.000019*	0.000003	-3.791679	0.000172*	2.791541
TRACT_JAIL_INCARCERATION_RATES.CSV.INCARCERATION_RATE_RP_GP_PALL	37.604393	3.867946	9.722058	0.000000*	4.486820	8.381079	0.000000*	2.381072

OLS Diagnostics

Input Features	Chicago Census Tracts	Dependent Variable	CHICAGO HEALTH ATLAS DATA DOWNLOAD - CENSUS TRACTS.CSV.PMH_2022
Number of Observations	775	Akaike's Information Criterion (AICc)[d]	3068.894191
Multiple R-Squared[d]	0.689787	Adjusted R-Squared[d]	0.687363
Joint F-Statistic[e]	284.619704	Prob(>F), (6,768) degrees of freedom	0.000000*
Joint Wald Statistic[e]	1582.584913	Prob(>chi-squared), (6) degrees of freedom	0.000000*
Koenker (BP) Statistic[f]	55.020204	Prob(>chi-squared), (6) degrees of freedom	0.000000*
Jarque-Bera Statistic[g]	95.190077	Prob(>chi-squared), (2) degrees of freedom	0.000000*

Notes on Interpretation

*	An asterisk next to a number indicates a statistically significant p-value (p < 0.01).
a	Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.
b	Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.
c	Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.
d	R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.
e	Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Wald Statistic to determine overall model significance.
f	Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.
g	Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are biased (the residuals are not normally distributed).

Table A.3: OLS Model 3 Output Table

Full demographic model controlling for all confounders including race, education, poverty, age/sex, and economic variables. Final R squared = 0.69, with stable incarceration coefficient ($\beta = 37.40, \rho < 0.001$*), demonstrating a relationship independent of demographic composition (N = 775).*

Appendix B: Spatial Analysis Results

Global Moran's I Summary

Moran's Index	0.235677
Expected Index	-0.001292
Variance	0.000428
z-score	11.452283
p-value	0.000000

Table B.1: Global Moran's I (Spatial Autocorrelation) Table

Spatial autocorrelation test results for Model 3 residuals. Moran's I = 0.236 (p < 0.001) indicates significant positive spatial clustering, suggesting spillover effects beyond individual tract characteristics.

Model Diagnostics

R2	0.8065
AdjR2	0.7772
AICc	2849.1356
Sigma-Squared	2.1597
Sigma-Squared MLE	1.8763
Effective Degrees of Freedom	673.3265
Adjusted Critical Value of Pseudo-t Statistics	2.8254

Succeeded at Monday, December 1, 2025 11:49:58 AM (Elapsed Time: 1 minutes 58 seconds)

Table B.2: Geographically Weighted Regression Table

GWR model diagnostics showing improvements over global OLS. AICc comparison indicates better model fit when allowing coefficients to vary spatially, confirming geographic variability in the incarceration-mental health relationship.

Appendix C: Descriptive Statistics & Summary Tables

Alias	Nulls	Chart Preview	Min	Max	Mean	Std. Dev.	Median	Count	Unique	Outliers	Sum	Range	IQR	Q1	Q3	CV	Skewness	Kurtosis
Proportion Black	23 (2.7%)		0	100	32.689781	37.354689	8.919406	831 (97.3%)	801	0	27,165.20781	100	71.674905	2.613823	74.288728	1.142702	0.734743	1.795815
Proportion under the poverty line	23 (2.7%)		0.732535	73.401878	17.769778	12.53505	14.637904	831 (97.3%)	831	18	14,766.685657	72.669342	16.735708	8.014162	24.74987	0.705414	1.19738	4.462021
Proportion without a BA	23 (2.7%)		5.548345	98.068077	60.243793	26.210507	67.021277	831 (97.3%)	831	0	50,062.592093	92.519732	43.769594	39.545823	83.315416	0.435074	-0.534502	1.996624
Proportion Unemployed	23 (2.7%)		0	48.909657	9.213234	7.369508	6.86747	831 (97.3%)	822	36	7,656.197081	48.909657	8.559323	3.954329	12.513652	0.799883	1.492062	5.531166
Proportion of Residents that are Male 18-34	23 (2.7%)		1.158717	53.5183	13.774737	6.556304	12.479871	831 (97.3%)	829	32	11,446.806376	52.359583	7.620596	9.201314	16.821909	0.475966	1.348174	6.066562
Incarceration Rate	6 (0.7%)		0	0.1205	0.025718	0.024846	0.0166	848 (99.3%)	454	13	21.8091	0.1205	0.035175	0.0058	0.040975	0.966082	1.153668	3.741424
Rerported Mental Health	65 (7.6%)		9.8	28.2	17.135995	3.130507	16.6	789 (92.4%)	145	8	13,520.3	18.4	4.4	14.7	19.1	0.182686	0.680679	3.109297
median income	33 (3.9%)		13,489	244,286	78,057.113276	40,724.127121	70,714	821 (96.1%)	811	33	64,084,890	230,797	47,347	49,017	96,364	0.521722	1.150928	4.494718

Table C.1: Summary Statistic Table

Descriptive Statistics for all variables across 775-848 Chicago census tracts. Shows variation: mental health ranges 9.8%-28.2%, incarceration rates 0-120.5 per 1,000 residents, and poverty 0.7%-73.4%.

Appendix D: Descriptive Maps

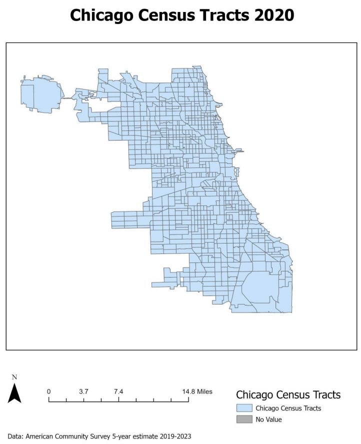


Figure D.1: Chicago Census Tracts (2020)
Base map that shows the area of study (Chicago boundaries) and unit of study (Census Tract) in the most recent geographic boundaries.

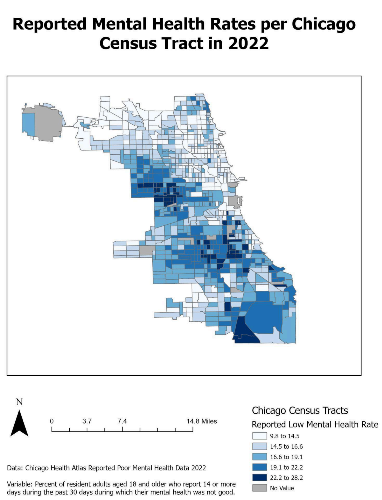


Figure D.2: Reported Mental Health Rates per Chicago Census Tract in 2022
Spatial distribution of self-reported poor mental health (14+ days in past month). Darker blue indicates higher rates, concentrated on the South and West sides.

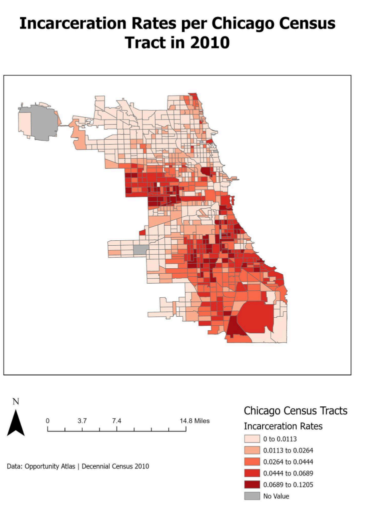


Figure D.3: Incarceration Rates per Chicago Census Tract in 2010

Incarceration rates per 1000 residents. Darker red indicates higher rates, showing similar geographic concentration to mental health outcomes.

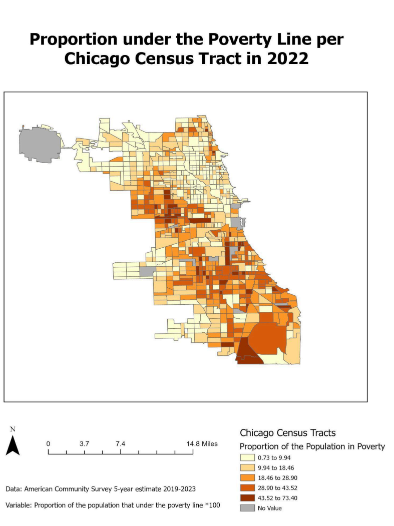


Figure D.4: Proportion under the Poverty Line per Chicago Census Tract in 2022

Percentage of population below federal poverty line. Note spatial overlap with incarceration and mental health patterns.

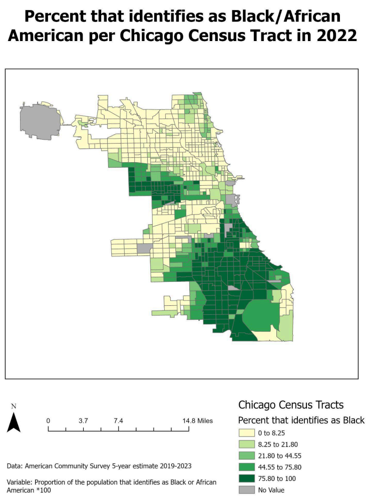


Figure D.5: Percent that identifies as Black/African American per Chicago Census Tract in 2022
Racial composition showing concentration of Black residents on South and West sides, overlapping with areas of high incarceration and poor mental health.

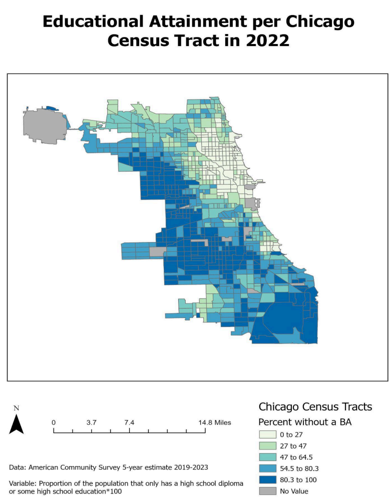


Figure D.6: Educational Attainment per Chicago Census Tract in 2022
Proportion without a bachelor’s degree. Lower educational attainment clusters geographically with other disadvantage indicators.

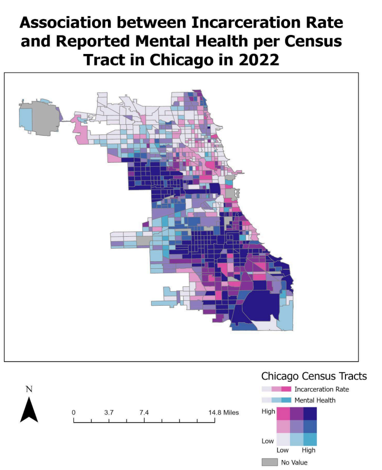


Figure D.7: Association between Incarceration Rate and Reported Mental Health per Census Tract in Chicago in 2022

Bivariate choropleth showing simultaneous spatial patterns. Dark purple areas indicate co-location of high incarceration and poor mental health.

Appendix E: Analysis Maps

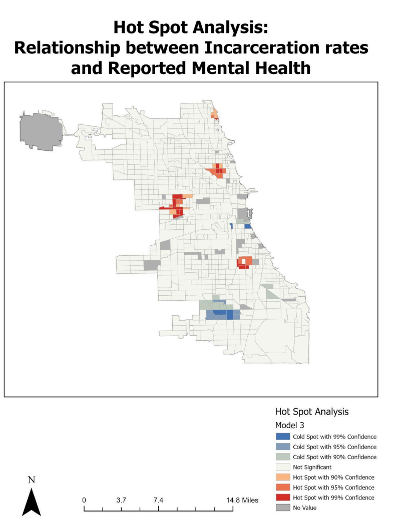


Figure E.1: Hot Spot Analysis of Incarceration and Mental Health Relationship

Getis-Ord Gi statistic identifies statistically significant spatial clusters of Model 3 residuals. Red hot spots (South and West sides) indicate neighborhoods where both incarceration and mental health outcomes cluster at higher levels than predicted by demographic controls, suggesting localized spillover effects. Blue cold spots (South suburbs) show areas where outcomes are better than predicted. The cream areas show no significant clustering (N=775 tracts).

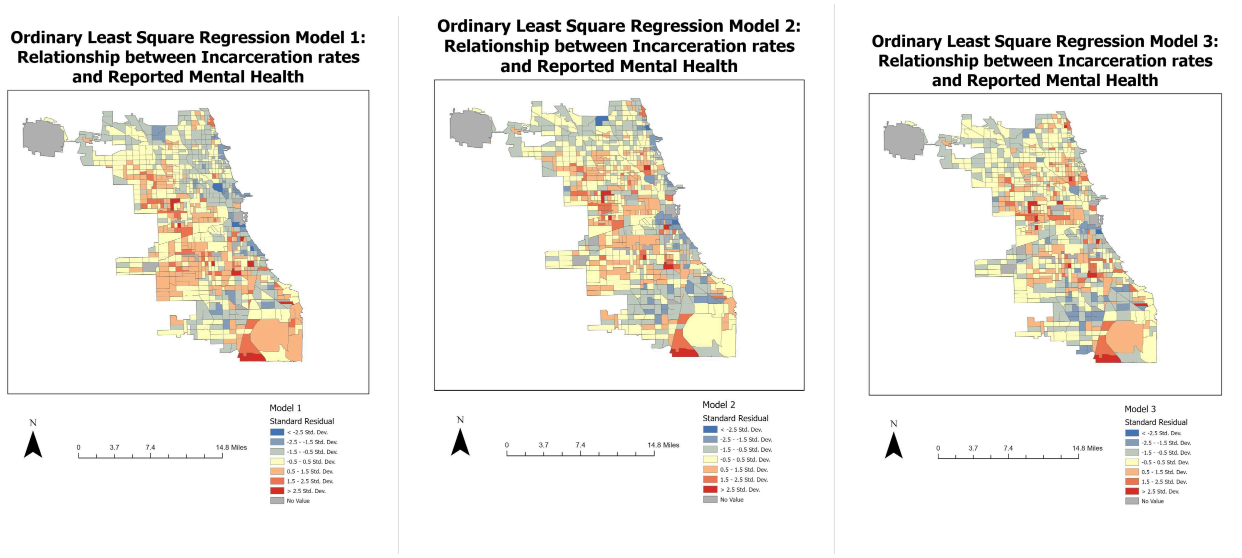


Figure E.2- E.4: Spatial Distribution of Standardized Residuals Across Three OLS Models

Standardized residuals from Models 1 (bivariate), 2 (socioeconomic controls), and 3 (full demographic controls) show where each model systemically over-predicts (blue, residual < -0.5 SD) or under-predicts (red, residual > +0.5 SD) mental health outcomes. Comparison across models reveals how adding control variables changes geographic patterns of prediction error. Persistent spatial clustering in Model 3 residuals motivated our GWR analysis.

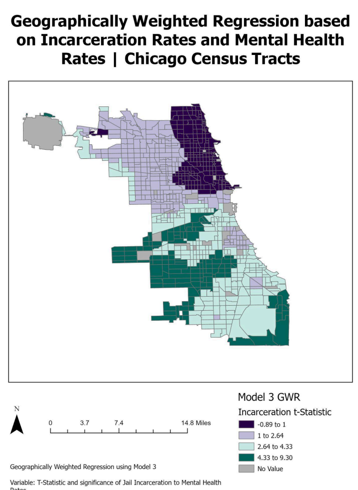


Figure E.5: Geographic Variation in Statistical Significance of Incarceration Effect

Local t-statistics from Geographically Weighted Regression showing where the incarceration-mental health relationship is statistically significant. Dark purple (Far North side) and dark teal (South and Southwest sides) indicate areas with t-statistics > 4.33 (p < 0.001), meaning the relationship is highly significant in these neighborhoods. Lighter areas show weaker or non-significant relationships. Geographic variation confirms the incarceration effect varies across Chicago.

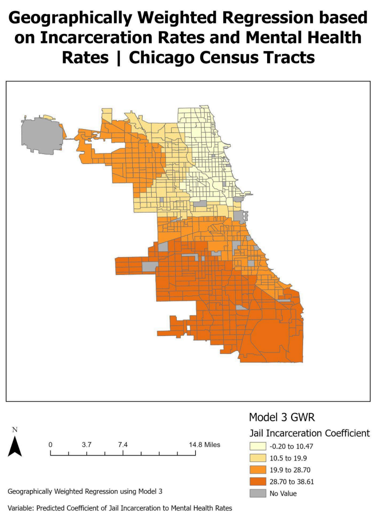
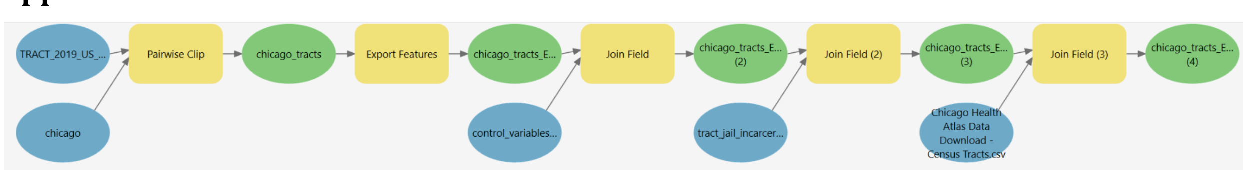


Figure E.6: Geographic Variation in Magnitude of Incarceration Effect

Local regression coefficients showing how much mental health deteriorates per unit increase in incarceration rate in each neighborhood. Dark orange areas (South and Southwest sides, coefficients 28.70-38.61) experience the largest mental health impacts from incarceration, while lighter areas show smaller effects. This geographic pattern demonstrates that spillover effects are strongest where incarceration rates are already highest, supporting the hypothesis of disproportionate impacts in economically disinvested communities.

Appendix F: Model Builder



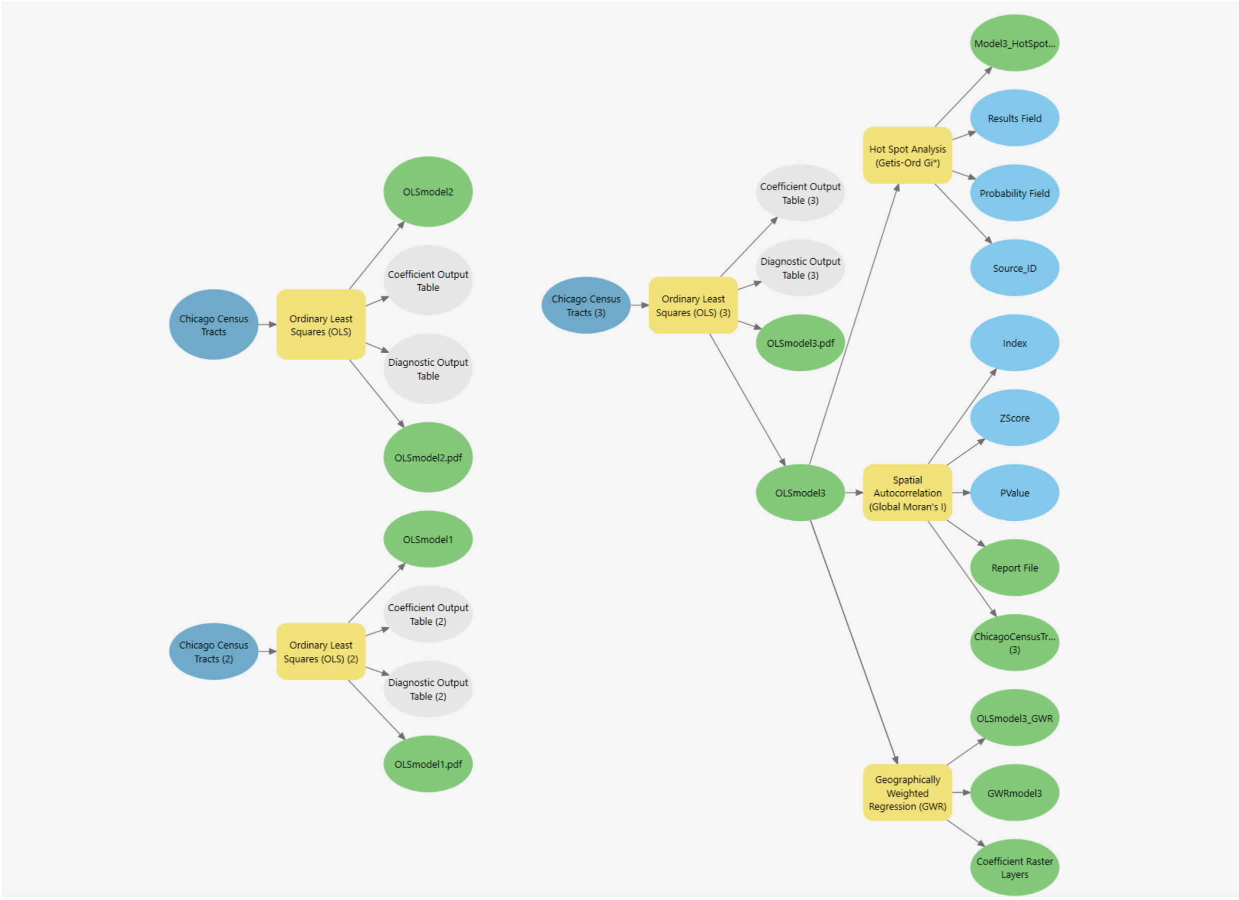


Figure F.1-F.2: Complete ArcGIS Model Builder Workflow

Automated, reproducible workflow for the complete analysis pipeline. Top figure shows data preprocessing: clipping census tract boundaries to Chicago, joining control variables, incarceration data, and mental health data. The bottom figure shows the analytical workflow: three parallel OLS regression models feeding into spatial autocorrelation testing (Global Moran's I), hot spot analysis (Getis-Ord Gi), and conditional Geographically Weighted Regression. Yellow boxes indicate geoprocessing tools, blue boxes show input data, and green boxes show outputs.