Hungry Hands: Food Security’s Role in Predicting Physical Security

Jared M. Moore

University of Texas, Austin

Case Studies in Machine Learning

Fall 2024

Abstract

Food insecurity represents a critical public health and social challenge, with emerging evidence linking it to increased rates of crime and violence. This study explores the intersection between food access, socioeconomic indicators, and crime in Connecticut, utilizing data from the USDA Food Access Research Atlas, FBI crime reports, and census tract-level demographic information. By aggregating food access and crime data at the town level, we examine the spatial and statistical relationships between these variables to assess how food insecurity influences physical security, and use food insecurity as a predictor for overall crime.

Our machine learning analysis reveals that poverty rates and SNAP enrollment are robust predictors of crime, with higher crime rates observed in areas with lower median family incomes and greater proportions of low-access, low-income populations. Contrary to expectations, the presence of food-insecure children demonstrated minimal predictive power, suggesting the influence of unobserved variables or limitations in dataset granularity. Additionally, machine learning models like KNN regression outperformed other models like Neural Networks when using geospatial demographic data to predict crime. These findings emphasize the complex interplay between food insecurity and crime, providing actionable insights for policymakers. Future work should explore broader datasets and incorporate sociological factors to develop comprehensive strategies addressing food insecurity and its role in community safety.

Introduction

Food security, defined as consistent access to sufficient, safe, and nutritious food, is a cornerstone of public health and well-being (Rabbitt et al., n.d.). Its absence - food insecurity - disproportionately affects vulnerable populations, particularly young people, where it is linked to developmental challenges, poor mental health outcomes, and heightened risk behaviors which can have negative community outcomes. Adolescents in food-insecure households are more likely to experience anxiety, depression, and even suicidal tendencies, underscoring the profound impact of food access on both physical and psychological health.

The dynamics of food security are deeply tied to geographic and socioeconomic factors. Across zip codes, disparities in food access are shaped by poverty, transportation barriers, and the uneven distribution of food retailers, creating pockets of "food deserts" where all food is scarce to buy, and "food swamps" where available food is of limited quality. These areas often coincide with neighborhoods experiencing elevated crime rates, suggesting a complex interplay between economic deprivation, limited access to nutritious food, and community safety. Exploring these intersections is critical to understanding how food insecurity influences broader societal outcomes, including public safety; by identifying effective interventions, public policy can be crafted to improve both food access and physical security for low access communities.

Background

Literature on Food Insecurity

To understand how food insecurity might influence crime, we need to first consider how it affects individuals, demographics, and locations. Understanding the health impacts helps us make the connection with health-based crime indicators, for which there are large bodies of research. Focusing on location helps model different population sizes and how their distances to key resources affect resource predation. Finally, we will look at how food insecurity is measured to more accurately understand the data being used as the basis of prediction.

Food Insecurities Affect on Health

Food insecurity has long been known to have an impact on lifetime health outcomes. It has been found that food insufficiency is strongly associated with depressive disorders and suicidal symptoms in U.S. adolescents (Alaimo et al., 2002). These depressive disorders are represented differently across food insecure populations from different countries, indicating the cultural relationship with food may also play a role in these symptoms (Koyanagi et al., 2019). In addition to creating mental health problems for low access children, food insecurity also affects psychosocial functioning between low-income American children and their peers, correlating with absenteeism from school, poor behavior, and low academic performance (Murphy et al., 1998). Interventions such as free school lunch programs, SNAP/EBT enrollment, and subsidized housing programs, many of which are implemented in Connecticut where this data was focused, lead to long term societal savings, but require long term financing (Thornton et al., 2016). SNAP especially, which is a food voucher program provided to low income individuals in the US, has been shown to improve nutritional outcomes and lower healthcare costs, providing a net benefit for every dollar spent (Carlson & Keith-Jennings, 2018). Exploring the relationship between these programs, children's mental health, absenteeism, and general hunger, especially in areas with varying access to food, may reveal some underlying drivers of crime as they relate to food insecurity.

Food Insecurities Impacts vary across Locations

Food insecurity is not solely affected by the benefits received or the resources available, but rather effective access to food. 20% of low-income Americans reported having no access to a personal vehicle, with one of the fifth highest low-income no-vehicle cities in the country existing in this state, Waterbury, Conn at 16% (*Car Ownership Statistics in 2024 | The Zebra*, n.d.). Additionally, crime has been found to have a negative relationship with small retail growth, and areas with high inequality and few SNAP recipients are negatively correlated with fast food growth, demonstrating the complex relationship between retail location selection and an area's population factors (Fuad et al., 2023). Socioeconomically-disadvantaged populations and minority populations may attract 'unhealthy' food outlets over time, indicating that the availability of 'less healthy' food outlet types may be more important than the potential lack of supermarkets or full service restaurants in determining food insecurity (Rummo et al., 2017).

Race and community support also plays a key factor in understanding the locational impacts of food insecurity. It was found that there are key avenues (side hustle, government benefits, emergency food assistance) which people utilize when food-insecure, and that local supply and demand dynamics were interlaced with stigma and community ownership of the food supply, especially in primarily black neighborhoods (Freedman et al., 2022). They show that social support and social cohesion have a greater impact in reducing the risk of food insecurity, which may indicate shifting our focus from improving food access of households to interventions for social cohesion at the neighborhood level to more adequately address food insecurity (King, 2017). This is further supported by the recommendations made by risk analysis of the causes of food deserts, which said that increasing access to supermarkets in low income areas is not enough alone, but must be supplemented by nutrition education, labor support through unemployment benefits, and increased access to farmers' markets in historically low access areas, which uniquely fit the needs of those areas (Wang et al., n.d.). There are many factors at play in food desert dynamics, and solving the problem cannot be done with one single change. For this reason, looking at the problem as a multi featured machine learning project may be helpful.

How Food Insecurity is Measured

Before modeling with the USDA Food Access Research Atlas data, it's important to understand how that data is measured and collected. Food insecurity is a long run measure of a more acute problem of food insufficiency. Food insufficiency asks if you had the kinds of foods you want to eat, and enough food, within the last 7 days, whereas food insecurity asks if you've been worried about running out of food in two ways, by reducing the quality of meals or reducing the frequency over the last 30 days or 12 months (Rabbitt et al., n.d.). This contrasts the ideas of food insecurity, being a long running struggle to ensure you have food, with food insufficiency being a state, more like hunger.

Additionally, research measuring the ways to influence the effects of vehicle ownership rates, public transit access, crime, and poor traffic safety, found adjusting for vehicle ownership and crime tended to increase measured disparities in access to supermarkets by neighborhood, race, and income adjusting for public transit and traffic safety tended to narrow these disparities, suggesting a much more fragmented and unique food landscape than previously thought (Bader et al., 2010). Another question related to measuring food insecurity is how much we can influence outcomes through public policy (Zohrabi et al., 2021). It seems like availability drives choices, meaning areas with high volumes of supermarkets and low volumes of convenience stores have healthier populations, and areas with low availability of fast food have lower rates of obesity (Larson et al., 2009). More research will be needed to fully understand the complexities of how public policy related to food and other social issues can be an avenue to reduce criminal behavior.

Literature on Crime

Criminal behavior is a complex phenomenon, shaped by psychological, social, environmental, and structural factors. Research investigating *why crime happens* emphasizes the effect of environmental stressors and systematic inequalities on the individual. Case studies on high-violence areas and gang involvement often are used as tools to observe these phenomena more closely. The spatial distribution of crime, or *where crime happens*, often underscores the differences or similarities of different populations and outcomes to determine causation. Structural determinants, such as income inequality and residential instability, drive disparities in personal victimization rates across communities. Proximity to supermarkets, liquor stores, convenience stores, and fast food all correlate with crime hotspots, which highlight how a combination of environmental and individual factors is likely to impact an area's predisposition to criminal behavior.

Why Crime Happens

Understanding why crime happens requires an analysis of intersecting environmental, social, and economic stressors. Research indicates that high-crime areas often share common structural challenges, such as poverty, unemployment, and systemic disinvestment, which create conditions conducive to criminality (Smith et al., 2022). These stressors often amplify individual vulnerabilities, such as impulsivity or low academic performance, particularly in communities with limited access to economic resources and social support. These relationships are further illustrated by the complex dynamics of gang involvement amongst at-risk youth, where risk and protective factors differ significantly among gang-involved, pressured-to-join, and non-gang involved youth (Merrin et al., 2015). Low economic opportunities brought about by large outside economic forces such as the expansion of low-wage retail chains like Walmart, can also play a role in the proliferation of crime by disrupting local economies and weakening community cohesion (Wolfe & Pyrooz, 2014). These findings emphasize criminal behavior emerges not in isolation but as a product of interrelated systemic and situational influences, necessitating more comprehensive approaches to prevention and intervention.

Where Crime Happens

The location of criminal activity is closely tied to neighborhood characteristics and environmental factors. Structural determinants, such as income inequality, residential insecurity, and social disorganization, significantly influence where crime is mostly likely to occur (SAMPSON, 1985). Specific types of establishments, including liquor stores, convenience stores, and drug treatment centers, are often associated with higher crime rates in their immediate surroundings due to their role as healthcare facilities and economic hubs (Furr-Holden et al., 2016). The spatial patterns emphasize how environment and community context shapes the changing geography of crime, and provides insight into which businesses and features attract or repel crime to support targeted crime prevention strategies.

Overlap of Crime and Food Insecurity

Now that we have established how food insecurity is quantified and explored the structural and environmental factors that contribute to the causes and locations of criminal activity, we can examine and understand literature focusing where these issues overlap. Research reveals a direct link between food insecurity and violence, where the stress and instability associated with limited access to food can heighten the risk of criminal behavior.

Additionally, studies highlight how disparities in food access and neighborhood safety shape individual actions and community dynamics. By investigating these overlaps, we gain deeper insights into the interconnectedness of nutrition equity, social cohesion, and public safety, offering pathways for addressing the systemic roots of food insecurity and its relationship to crime.

Direct Relationship between Violence and Food Insecurity

The interconnection between food insecurity and violence emerges as a critical public health and social issue, with multiple studies revealing a complex and multifaceted relationship. Research has demonstrated that communities experiencing high levels of food insecurity are more likely to experience elevated rates of violent crime (Blankenberger, 2016), while county-level analysis reveals a much more granular correlation between limited food retailer availability and increased violent crime rates (Singleton et al., 2022). Studies exploring the nutritional equity landscape highlight the complex interplay between economic and nutritional barriers that can contribute to hightented community tension and criminal behavior (Singleton, 2024). Targeted research in urban environments, specifically examining gun violence in major metropolitan areas, has found a direct association between food insecurity and increased violent incidents (Ali et al., 2022; Caughron, 2016). These findings collectively highlight the urgent need for integrated approaches that address both food access and community safety, recognizing that nutritional security is a crucial factor in maintaining social cohesion and reducing violent interactions.

Conflict and Food Insecurity

Food security and conflict are intrinsically linked, with emerging research revealing critical connections between nutritional resources and social stability. Research emphasizes the complex empirical challenges in understanding how food insecurity contributes to conflict dynamics, citing that more and better data is needed from conflict zones to understand the true nature of local food systems (Martin-Shields & Stojetz, 2019). The research is further marred by the diversity of experience of food insecurity and conflict, and how the specific practices and policies adopted may reduce the negative effects of conflict via food security (Brück & d’Errico, 2019). Targeted nutrition assistance programs like SNAP can potentially mitigate social tensions, suggesting that food security interventions may serve as a strategic approach to reducing conflict risks (Carr & Packham, 2019). These studies provide evidence of the relationships between food security and conflict on a micro level, which may allow us to reduce conflict recorded as crime on a macro level.

Juvenile Crime and Food Insecurity

Food insecurity significantly impacts juvenile delinquency through complex social mechanisms. Research reveals associations between food insecurity and reduced neighborhood safety, compromised social cohesion, and diminished social control, particularly among mothers of preschool aged children (DiFiore et al., 2022). Moreover, studies indicate that food insecurity correlates with low self control, increasing the likelihood of early delinquent behaviors among youth and adolescents (Jackson et al., 2018). This suggests nutritional deprivation may fundamentally alter children's behavioral trajectories and social circles, inducing criminality.

Relation of Health Conditions to Crime

Research exploring the intersection of health and crime reveals significant correlations between physiological conditions and criminal behavior. A Chicago-based study documented clear links between police-recorded crime and health disparities, specifically examining obesity and blood pressure status (Tung et al., 2018). These associations are so strong that the American

Dietetic Association suggests that nutritional factors play a crucial role in social interactions and potential criminal activities (ADA Reports, 1985), highlighting how intrinsically linked healthcare outcomes and legal outcomes can be. These studies showcase how a direct descendent of food insecurity, health conditions, has been found to be directly correlated with criminal behavior as early as 1985.

Data

Data Sources

This research investigates the relationship between food insecurity and crime across Connecticut towns using comprehensive data sources from the USDA Food Access Research Atlas, FBI, and CT Data Collaborative. By understanding, cleaning, and integrating tract and town level data, this study provides a nuanced analysis of the spatial and statistical connections between food access and criminal activities.

USDA Food Access Research Atlas Data 2019 is a collection of comprehensive geospatial data on food access and food insecurity across the United States. It is publicly available https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/, and offers a plethora of information about food desert indicators for each census tract in the U.S. such as number of residents with limited food access, distance to nearest supermarket or grocery store, and population percentage with low food access. The data also contains demographic information such as income levels, population density, vehicle availability, proximity to food retailers, race, sex, and child demographic information. This rich information would need to be trimmed for analysis to focus on Connecticut and Austin, TX, but provides a great basis of indicators for food security throughout the country.

FBI CT Offense Type by Agency is a collection of the count of reported offenses by type for all of the towns with a law enforcement agency in Connecticut. This data is publicly available https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/tables/table-8/table-8-state-cuts/connecticut.xls, and gives descriptive information on the number of reported incidents in each town as well as town population totals. These population sizes will allow us to normalize the crime statistics to per capita, and will provide a robust foundation for analysis.

CT Data Collaborative Tract-2-Town Data is a dataset of matching census fips codes to corresponding town and county names for all of the census tracts in Connecticut. This dataset is maintained by the CT Data Collaborative, which works to improve data equity by publicizing and open sourcing critical public data about the state of Connecticut. Their many datasets and more information about their organization can be found at https://www.ctdata.org/.

Data Pre-Processing

Before I can use the selected data to train a model, a few key challenges need to be addressed. Firstly, there are towns contained in the FBI Agency data which are not represented in the CT Data Collaborative's data, meaning those towns will need to be added manually. Additionally, the USDA data is currently by census tract, not by town, making it difficult to combine with crime statistics which are quoted on a town by town basis. To solve this, the census tracts will be replaced with the towns these tracts reside in, and the rows from similar towns will be aggregated and combined to give an overall picture of the USDA data on a town by town basis.

CT Data Collaborative Tract-2-Town Data was missing two towns in Connecticut, Willimantic and Grotton Town. To remedy this, tracts which were mislabeled with their county designation were relabeled with the proper town name to enrich the data and ensure those towns were not dropped from the analysis.

FBI CT Offense Type by Agency had a few key areas needing data cleanup. The first was the inclusion of State College data, and state police records, which are separate from the town-by-town records. This data was omitted since its location could not be attributed to any particular town. The next issue was the unintuitive labeling of the crime columns, which were replaced for clarity. Finally, any columns not containing data were removed and the statistics were converted from crime by type per town, to crimes per capita by type per town. This makes the data easily scalable, which facilitates combination with the USDA data.

USDA Food Access Research Atlas Data 2019 was the most challenging of the data preprocessing tasks. First, all nulls, missing values, and blanks needed to be normalized for the matrix calculations. Since many of the blanks were indicative of no residents, they could be replaced with 0's. Next, columns containing all 0's, such as the column for all LALOWI (Low Access Low Income) individuals from 10-20 miles from the nearest grocery store, were dropped from the table, as there are no areas in Connecticut that far from a grocery store. Additionally, the USDA data had population numbers, raw numbers of individuals in certain conditions, and percentage-based data. To compress this data, columns containing percentage-based data were dropped. Utilizing the FIPS code to match the corresponding towns in the Tract-2-Town data, towns were added to the USDA data. Rows with cells from the same town were combined based on data aggregation techniques. Metrics like Median Family Income, Flags, and poverty rate were aggregated with averages, while metrics like population and demographic information were summed.

Combining Food & Crime Data

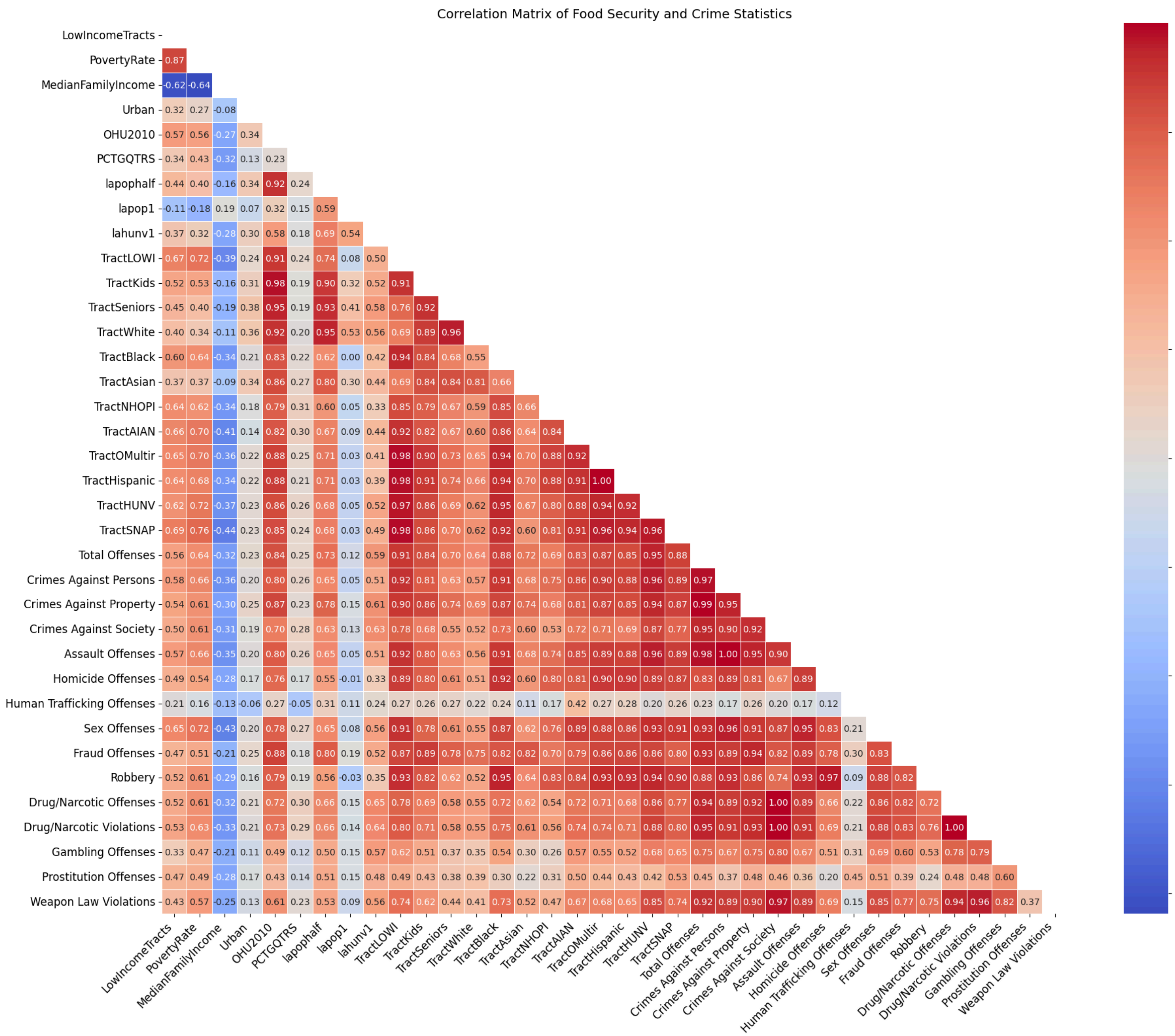
Combining the prepared FBI and USDA data was simple after preprocessing steps. Columns which had a matching town in both the USDA and FBI data were added to the same row, and redundant columns such as population were dropped. USDA data was only kept if there was a corresponding town law enforcement agency to report data (some small towns in Connecticut do not have their own law enforcement agencies). The combined and aggregated data is prepared for analysis and training a machine learning model.

Methods

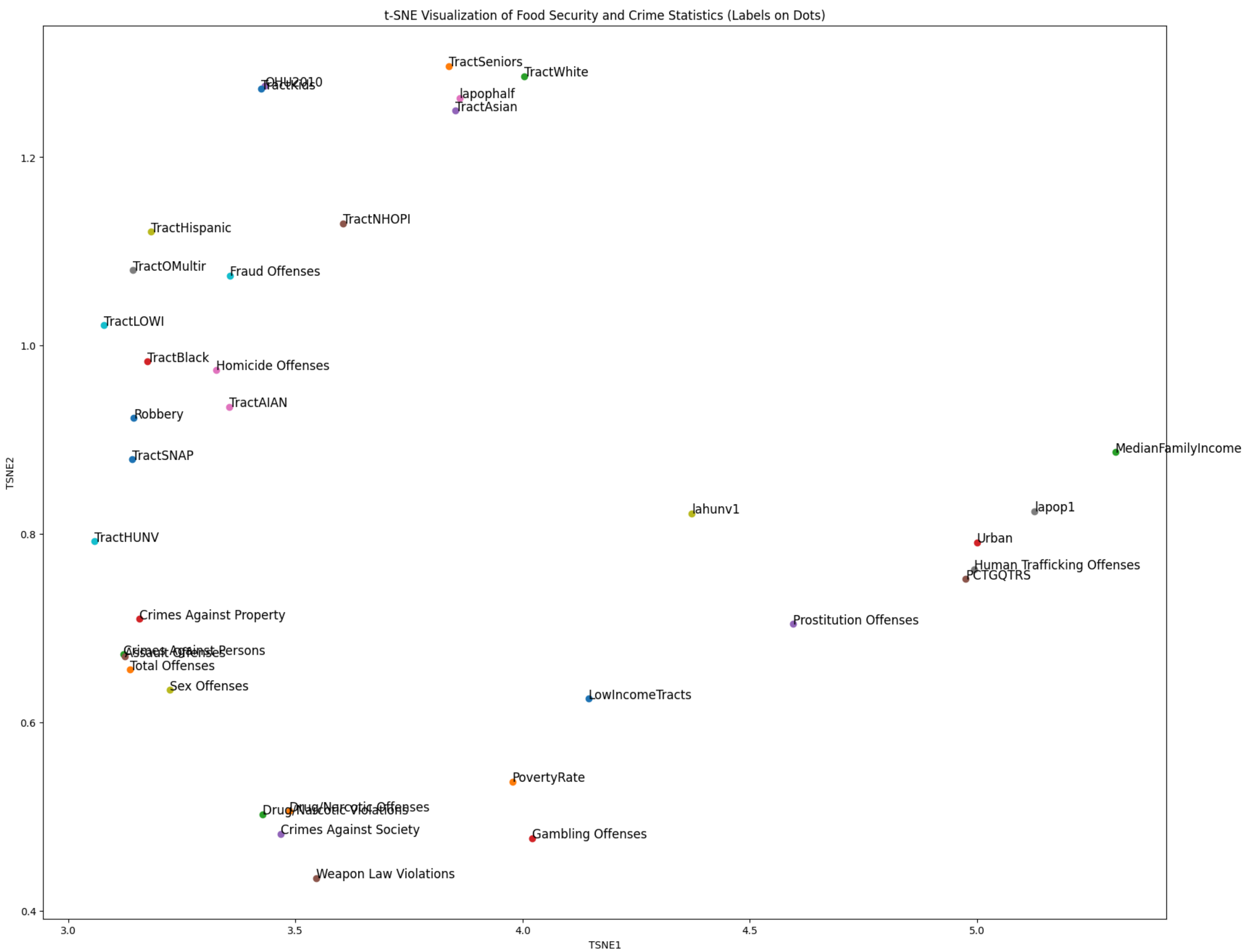
Data Analysis

Before we choose a model and begin using our data as a basis of prediction, we should ensure we understand our specific data. The methods below explore relationships in the data we can then utilize to predict other key factors about our data. By clustering our data, plotting key relationships, and comparing how each variable correlates with each other variable, it gives valuable information we can use to decide on the best model for our machine learning task.

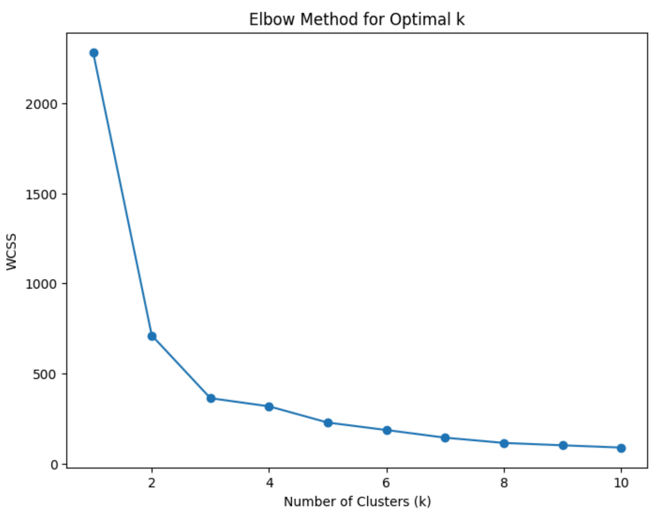
Correlation Matrix. A correlation matrix systematically quantifies the strength and direction of linear relationships between multiple variables, presenting a comprehensive visual representation of interconnectedness through computed correlation coefficients ranging from -1 to 1, where positive correlated items often increase or decrease together, negative correlated items have a directly inverse relationship, and items with 0 correlation are unrelated to each other and independent. We can see metrics like median family income correlate negatively with poverty rate, robberies and homicides seem to be positively correlated, and metrics like if a tract is urban or not seem to have very little correlation with other metrics. Analyzing these relationships helps us better understand the data, such as noticing the lack of correlation of human trafficking offenses. These types of charges are very rare in this dataset, and as such seem to happen "randomly" compared to other data which has more predictable patterns over time.

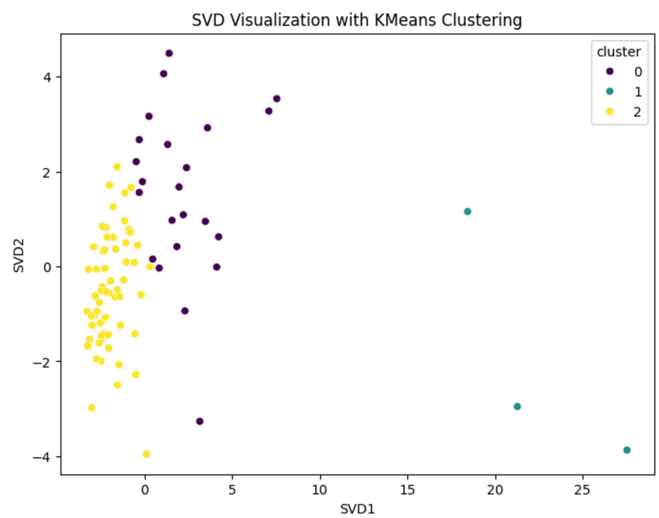


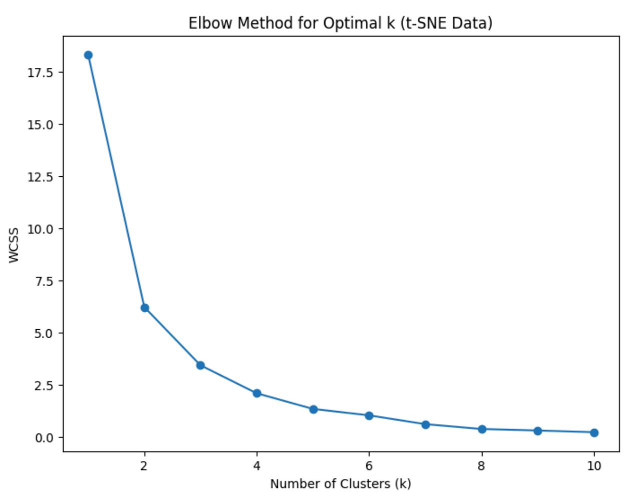
T-SNE is a statistical method for visualizing high-dimensional data by giving each piece of data a location on a 2D plane. This allows comparison of the relationships of these points, points that are near each other are alike, although it does not necessarily mean they are correlated or differ in the same way. It seems like Total Offenses is a good indicator of crimes against property, persons, and sex offenses, making this a relevant piece of data to predict.

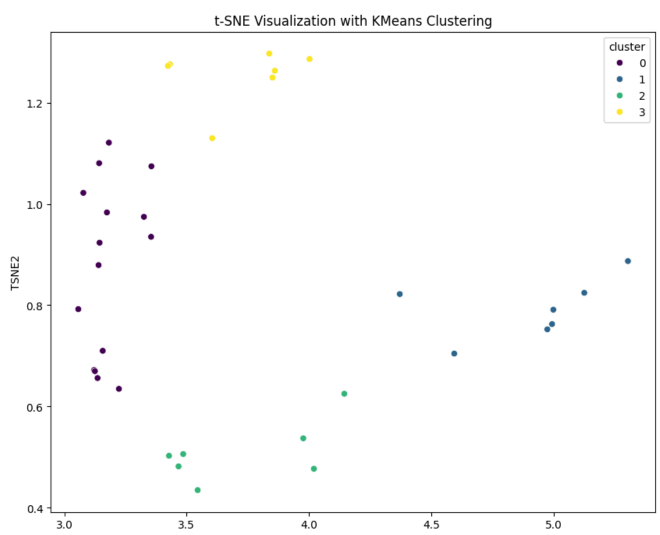


SVD and T-SNE with K-Means Clustering was used to group the points after reducing their dimensionality, with the hopes that the optimal number of points and types and size of clusters would provide valuable information for classifying the data. Here using the elbow method to find an optimal k, we can find 3 seems to be ideal for SVD, and 4 for t-SNE. We can also see clustering the dimensionality reduced data results in one tightly correlated group, one loosely correlated group, and one group of outliers, which may indicate that if the data we want to predict is in group 2, it may be more intuitive to predict than if it is in group 0 or 1.

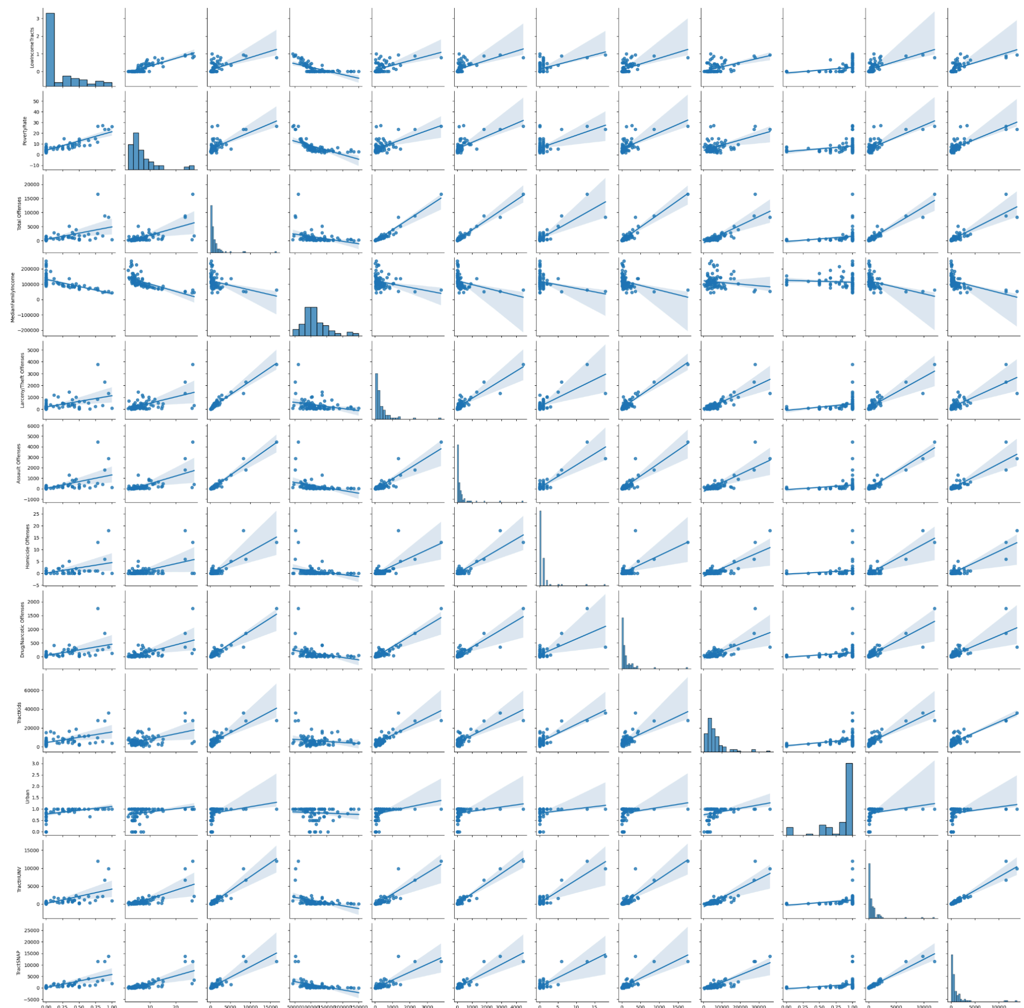


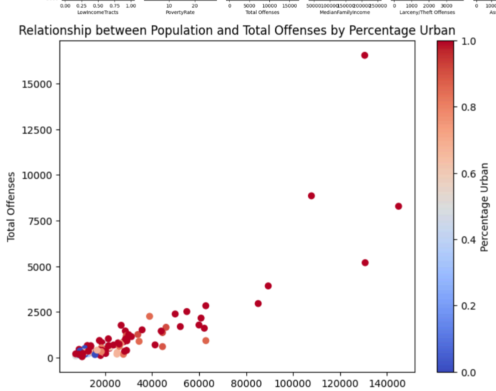


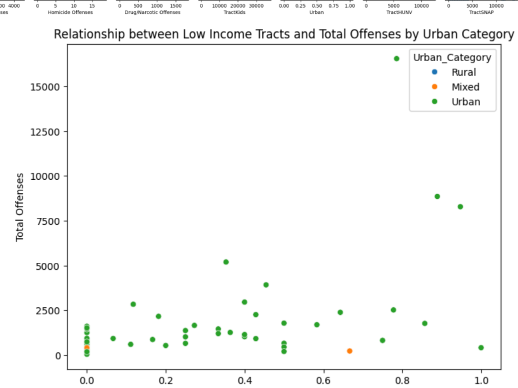




Pairplot of Key Relationships allows exploration of how specific variables in the data are related. Here we can make inferences from the data which we can use to predict why factors may be correlated. It is important when looking at pairplot relationships to remember correlation does not imply causation. Here, graphs with points which form lines with slopes of positive 1 are correlated, and slopes near negative 1 are negatively correlated. Additionally other relationships between points, such as completely scattered points or grouping of points, can provide valuable non-numeric insights into the ways different data points are correlated.







Model Analysis

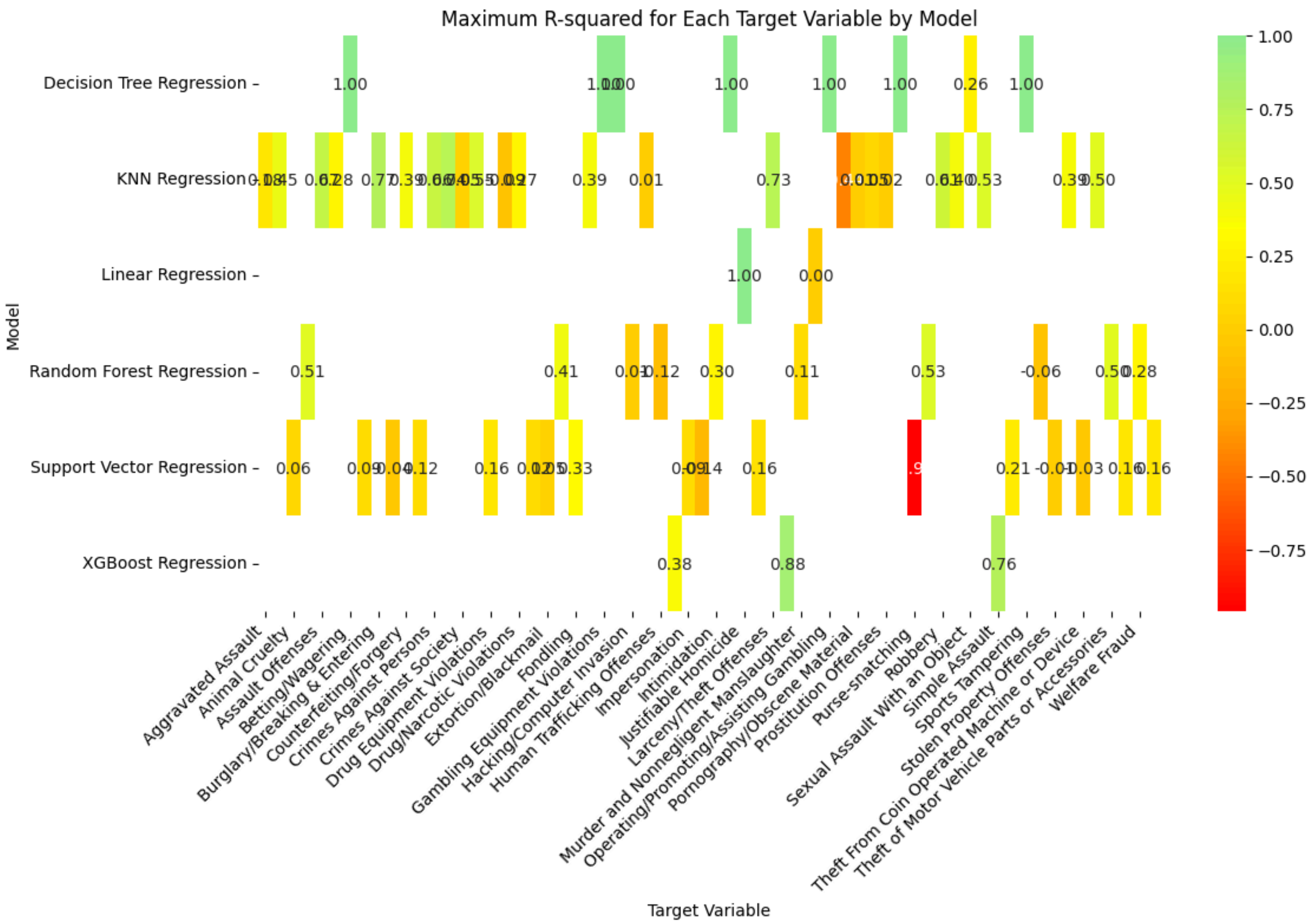
The model analysis evaluates the predictive accuracy and performance of various regression models in understanding the relationship between food insecurity and crime rates. A variety of models, including decision trees, KNN, and neural networks, were assessed to determine the predictive power of each model across different crime categories. Insights from prior t-SNE clustering revealed strong correlations between predictors of property crime, violent offenses, and sex crimes with overall crime rates. Therefore, models demonstrating high accuracy in these specific categories warrant closer examination for their potential for predicting total crime within a given area.

Testing Methods

To evaluate the performance of each model, we implemented a program that systematically trains decision tree regression, KNN regression, neural networks, and other models against the dataset, while leaving out a portion to test accuracy. The program calculates key evaluation metrics, including R-squared, root mean squared error (RMSE), and cross validated R-squared, to assess both the explanatory power and predictive accuracy of each model. Applying these metrics to predictions across different crime categories (e.g. property crimes, violent offenses, and sex offenses), allows identification of the models best suited for predicting total crime rates. This approach ensures the model selection is data-driven, and highlights models with the strongest generalization capabilities.

Model Analysis

The model analysis evaluates decision tree regression, KNN regression, neural networks, random forest regression, support vector regression, and XGBoost regression to identify the most effective approach for predicting crime rates. These models were chosen for their diverse strengths: Decision trees and random forests excel at handling non-linear relationships, KNN captures local patterns, neural networks leverage their ability to model complex structures in data, support vector regression is robust to overfitting in high-dimensional spaces, and XGBoost is known for its accuracy and computational efficiency in gradient boosted frameworks.



Decision Tree Regression seems to work really well on crimes like betting and wagering, or hacking and computer invasions, likely because these types of crimes happen very infrequently so rules about always predicting 0 or a low number have higher degrees of accuracy. While decision tree regression does perform well for some types of data, it fails to capture some of the complex interdependencies and relationships of the USDA and Crime data.

KNN Regression seems to be an optimal candidate for model selection. Compared to all the other models it was consistently the most accurate, while also performing best with property crimes, violent crimes, and sex offenses. This model seems to effectively rely on locally relevant patterns from data to make proximity based predictions, which align well with the spatial and socioeconomic characteristics of this dataset.

Linear Regression seems to have performed poorly compared to the other models tested, only modeling the simplest of relationships such as justifiable homicides and operating or promoting gambling, which had few or no data points. Because of linear regression's inability to capture complex relationships, it is not useful for predicting total crime.

Random Forest Regression seems to have done better than linear regression, but still performed poorly overall. Random forests can struggle with datasets where key predictors are sparse or unevenly distributed, which may have led to this poor performance.

Support Vector Regression performs well, notably capturing the loosely clustered crimes which were less commonly represented in the data. This makes sense, as SVR finds an optimal hyperplane that minimizes prediction errors while retaining robustness to outliers. SVR would be a good choice for prediction of more rare types of crimes that are less commonly represented in our data.

XGBoost Regression performs adequately on simple data with defined structures, but is likely not going to be useful for this analysis. The gradient-boosted decision trees fail to capture the complexities of the data, and have high computational overhead.

Final Chosen Models for Evaluation

This study's goal is to predict the total crime rate in an area by leveraging USDA food insecurity data alongside other socioeconomic factors. The analysis aims to identify how food security, specifically in areas with limited access to nutritious food, correlates to crime, and use that correlation to predict crime in areas where only food access information is known. Based on the results from the model evaluation, KNN regression seems to be the clear leader for predicting general crime data, and Support Vector regression appears to be the clear leader for predicting crime with more sparse data points. Because we are interested in predicting total crime, KNN is the clear choice. Additionally, while it was not tested in model evaluation due to computational overhead, a neural network is included in the analysis out of curiosity to explore how well it can learn the complex interactions between highly predictable variables like general socioeconomic factors, and sparse data, such as the occurrence of rare crimes.

Neural Networks may perform well because of their ability to capture complex, non-linear relationships between variables with different frequencies in a dataset, making it effective in modeling complex relationships, and a common avenue of study for current machine learning projects.

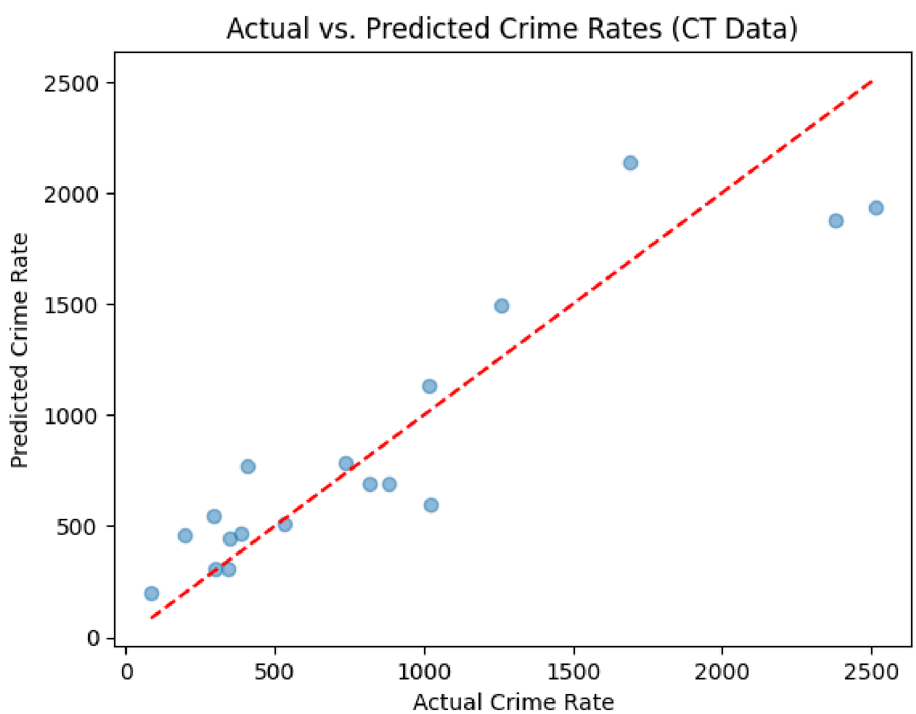
KNN Regression may perform well as a non-parametric model that is able to capture local patterns in the data without assuming a specific functional form. The ability to account for spatial or contextual relationships in neighborhoods, such as the proximity of areas with similar food insecurity levels or crime rates, enables KNN to effectively predict crime rates in areas with comparable characteristics.

Results

The results of the model prediction and training provide valuable insights into the relationship between food insecurity and crime rates in Connecticut. KNN had a high degree of accuracy, while generalizing well to new datasets. The Neural Network struggled to capture some of the complexities of the data, and was unable to outperform KNN. KNN appears to be an optimal way of modeling spatial relationships found in census data, and using them to predict other characteristics of your dataset.

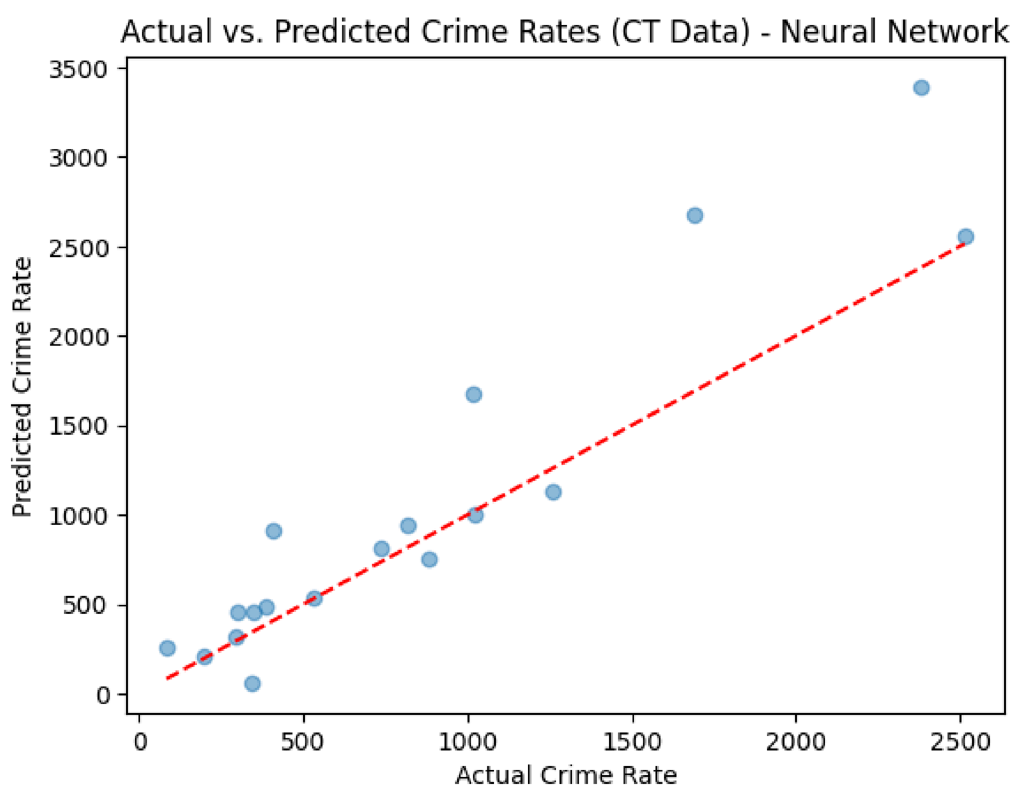
KNN Patterns

KNN predicted crime from USDA food data remarkably well, capturing approximately 84% of the variance in the data after hyperparameter tuning. When using a value of 4 for the number of neighbors referenced at each datapoint, and using cross validated scores from the Connecticut dataset, this model has a Mean Squared Error of 76844, and R-Squared score of 0.841213, much higher than that of the neural network. Additionally this model seemed to generalize well when testing on the Austin, TX dataset, predicting the total crime in each tract.



Neural Network Patterns

The Neural Network did not perform as well as expected on the USDA Connecticut data. Perhaps due to the hyperparameters or layers chosen, the network was unable to capture the complex relationships existent throughout the data. The Neural Network trained with 3 hidden layers, 64 neurons per layer, and 100 epochs for training had a mean squared error of 161054, and an R-Squared score of .667205. This means we were only capturing 66% of the variance in the data. This shows that while neural networks can be extremely powerful tools, careful data analysis can often lead to models which are better at performing a specific task. While it is a current trend to utilize Neural Networks to solve all manner of machine learning problems, it is important to look closely at your data and choose a model that makes the most sense for your needs to lead to a high chance of success. With more hyperparameter tuning or a different network structure, it would be reasonable to believe a neural network could outperform KNN, however that is not what was observed in this study.



Conclusions

In conclusion, the results of this model analysis highlight the importance of selecting the appropriate machine learning algorithm for the task at hand. The KNN model demonstrated strong predictive performance, effectively capturing and exploiting relationships between food insecurity and crime rates in Connecticut. The success of KNN for this task emphasizes the utility of KNN for modeling spatial relationships within other types of aggregated census data. On the other hand, the Neural Network, despite its potential, struggled to capture the complexity of the data, underperforming relative to KNN. These findings reinforce the notion that while advanced models like neural networks offer powerful capabilities, careful model selection and tuning are critical to achieving the best results for specific datasets and predictive goals. Based on the current analysis, KNN is the most effective model for predicting crime based on food security data, though future work may explore further refinements or alternative models to improve prediction accuracy.

Discussion

The results from this analysis suggest that food insecurity is closely linked to variations in crime rates, particularly in urban areas with high poverty rates and low levels of access. This section delves into the broader implications of these findings, integrating them with the existing literature to explore potential underlying mechanisms.

Takeaways from Study Conclusions

This study provides compelling evidence that factors such as food insecurity, socio-economic disparities, and local economic conditions are key contributors to crime rates in an area. This highlights the importance of addressing food insecurity as a part of targeted strategies to mitigate crime, and may be an important metric to track alongside crime.

Agreement / Disagreement with Other Studies. This study aligns with existing research by confirming that food insecurity is linked to higher crime rates, particularly violent crime, as shown by (Blankenberger, 2016) and (Singleton et al., 2022). However, it expands on this by emphasizing the role of socioeconomic and environmental factors, such as access to unhealthy food outlets, local economic conditions, and community cohesion, which play a significant role in shaping the food insecurity-crime relationship (Freedman et al., 2022; Rummo et al., 2017). Unlike much of the literature focused on urban areas, this study highlights how both rural and suburban regions face unique challenges due to transportation barriers and limited vehicle access (Bader et al., 2010). This study also reinforces the importance of programs like SNAP in improving health and reducing crime, but suggests that addressing food insecurity requires broader interventions, including social cohesion strategies (Thornton et al., 2016; Wang et al., n.d.).

Possible Explanations for Variance seen in Results Variance observed could be explained by the specific socioeconomic factors and programs in Connecticut, which do not exist in other parts of the country, as well as other differences in regional socioeconomic conditions, data quality, and the granularity of the datasets used. Local assistance programs and the variety of data reporting, as well as choices made during the data aggregation and data cleaning process, may have contributed to the variance seen in the resulting KNN models performance.

Why KNN may have been a Better Predictor than a Neural Network K-Nearest Neighbors (KNN) Regression performed particularly well in predicting crime rates in Connecticut, likely due to its ability to capture the local spatial dependencies and patterns in the data. Neural networks, while powerful, require more data and careful tuning for good results, which may have limited their ability to learn the complex relationships in this dataset effectively.

Future Research Opportunities

This study revealed new understandings and potential techniques for exploring the problem of crime and food security. Through the process of researching this topic, further avenues of research from expanding the training set to include more of the country, exploring some of the sociological reasons for the results we see, or exploring how public policy influences these variables, may be necessary next steps.

Expanding the Training Set Incorporating data from other regions and diverse socio-economic contexts would enhance the generalizability of this model. While there are areas of the US with individuals 10 miles or more from a grocery store, small states like Connecticut do not have enough samples of that type of data to effectively use it for analysis. Utilizing the whole USA's FBI data and USDA data may provide an even more detailed analysis with differing conclusions.

Exploring Sociological Reasons for the Results Further exploration into the sociological mechanisms linking food insecurity to crime is needed. This could include examining other data aside from USDA statistics, such as the role of social networks, family structure, community cohesion, and other factors in shaping both food security and criminal behavior.

Exploring Public Policies Influence on these Variables Oftentimes studies like these are used as the basis for public policy as it relates to food assistance, housing, and social services impact on crime. Further research into the impact of public policy changes on crime, and how changing underlying socioeconomic factors correlates to changes in criminal behavior, are necessary to understand the impact of these programs. Understanding these connections can guide future policy design to address both food security and crime reduction in vulnerable communities.

References

ADA Reports. (1985). Position Paper of The American Dietetic Association on diet and criminal behavior1. *Journal of the American Dietetic Association*, *85*(3), 361–362. https://doi.org/10.1016/S0002-8223(21)03617-8

Alaimo, K., Olson, C. M., & Frongillo, E. A. (2002). Family Food Insufficiency, but Not Low Family Income, Is Positively Associated with Dysthymia and Suicide Symptoms in Adolescents. *The Journal of Nutrition*, *132*(4), 719–725. https://doi.org/10.1093/jn/132.4.719

Ali, A., Broome, J., Tatum, D., Fleckman, J., Theall, K., Chaparro, M. P., Duchesne, J., & Taghavi, S. (2022). The association between food insecurity and gun violence in a major metropolitan city. *Journal of Trauma and Acute Care Surgery*, *93*(1), 91. https://doi.org/10.1097/TA.0000000000003578

Bader, M. D. M., Purciel, M., Yousefzadeh, P., & Neckerman, K. M. (2010). Disparities in neighborhood food environments: Implications of measurement strategies. *Economic Geography*, *86*(4), 409–430. https://doi.org/10.1111/j.1944-8287.2010.01084.x

Blankenberger, E. (2016). Feeding Peace: An Investigation of the Relationship between Food Insecurity and Violence. *Stevenson Center for Community and Economic Development—Student Research*. https://ir.library.illinoisstate.edu/scced/20

Brück, T., & d'Errico, M. (2019). Food security and violent conflict: Introduction to the special issue. *World Development*, *117*, 167–171. https://doi.org/10.1016/j.worlddev.2019.01.007

Car Ownership Statistics in 2024 | The Zebra. (n.d.). Retrieved November 27, 2024, from https://www.thezebra.com/resources/research/car-ownership-statistics/

Carlson, S., & Keith-Jennings, B. (2018). SNAP is linked with improved nutritional outcomes and lower health care costs. *Washington, DC: Center on Budget and Policy Priorities*, 1–19.

Carr, J. B., & Packham, A. (2019). SNAP Benefits and Crime: Evidence from Changing Disbursement Schedules. *The Review of Economics and Statistics*, *101*(2), 310–325. https://doi.org/10.1162/rest_a_00757

Caughron, J. (2016). An Examination of Food Insecurity and Its Impact on Violent Crime in American Communities. *All Theses*. https://open.clemson.edu/all_theses/2565

DiFiore, G., Hannan, C., Fiks, A. G., Virudachalam, S., Glanz, K., & Mayne, S. L. (2022). Associations Between Food Insecurity and Neighborhood Safety, Social Cohesion, Social Control, and Crime Among Mothers of Preschool-Aged Children. *Journal of Health Care for the Poor and Underserved*, *33*(3), 1258–1274. https://doi.org/10.1353/hpu.2022.0111

Freedman, D. A., Clark, J. K., Lounsbury, D. W., Boswell, L., Burns, M., Jackson, M. B., Mikelbank, K., Donley, G., Worley-Bell, L. Q., Mitchell, J., Ciesielski, T. H., Embaye, M., Lee, E. K., Roche, A., Gill, I., & Yamoah, O. (2022). Food system dynamics structuring nutrition equity in racialized urban neighborhoods. *The American Journal of Clinical Nutrition*, *115*(4), 1027–1038. https://doi.org/10.1093/ajcn/nqab380

Fuad, S., Badruddoza, S., & Amin, M. D. (2023). *Determinants of the presence, density, and popularity of US food retailers*. https://ageconsearch.umn.edu/record/335799/files/26949.pdf

Furr-Holden, C. D. M., Milam, A. J., Nesoff, E. D., Johnson, R. M., Fakunle, D. O., Jennings, J. M., & Thorpe, R. J. (2016). Not in My Back Yard: A Comparative Analysis of Crime Around Publicly Funded Drug Treatment Centers, Liquor Stores, Convenience Stores, and Corner Stores in One Mid-Atlantic City. *Journal of Studies on Alcohol and Drugs*, *77*(1), 17–24. https://doi.org/10.15288/jsad.2016.77.17

Jackson, D. B., Newsome, J., Vaughn, M. G., & Johnson, K. R. (2018). Considering the role of food insecurity in low self-control and early delinquency. *Journal of Criminal Justice*, *56*, 127–139. https://doi.org/10.1016/j.jcrimjus.2017.07.002

King, C. (2017). Informal assistance to urban families and the risk of household food insecurity. *Social Science & Medicine (1982)*, *189*, 105–113. https://doi.org/10.1016/j.socscimed.2017.07.030

Koyanagi, A., Stubbs, B., Oh, H., Veronese, N., Smith, L., Haro, J. M., & Vancampfort, D. (2019). Food insecurity (hunger) and suicide attempts among 179,771 adolescents attending school from 9 high-income, 31 middle-income, and 4 low-income countries: A cross-sectional study. *Journal of Affective Disorders*, *248*, 91–98. https://doi.org/10.1016/j.jad.2019.01.033

Larson, N. I., Story, M. T., & Nelson, M. C. (2009). Neighborhood environments: Disparities in access to healthy foods in the U.S. *American Journal of Preventive Medicine*, *36*(1), 74–81. https://doi.org/10.1016/j.amepre.2008.09.025

Martin-Shields, C. P., & Stojetz, W. (2019). Food security and conflict: Empirical challenges and future opportunities for research and policy making on food security and conflict. *World Development*, *119*, 150–164. https://doi.org/10.1016/j.worlddev.2018.07.011

Merrin, G. J., Hong, J. S., & Espelage, D. L. (2015). Are the risk and protective factors similar for gang-involved, pressured-to-join, and non-gang-involved youth? A social-ecological analysis. *American Journal of Orthopsychiatry*, *85*(6), 522–535. https://doi.org/10.1037/ort0000094

Murphy, J. M., Wehler, C. A., Pagano, M. E., Little, M., Kleinman, R. E., & Jellinek, M. S.

(1998). Relationship Between Hunger and Psychosocial Functioning in Low-Income American Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*(2), 163–170. https://doi.org/10.1097/00004583-199802000-00008

Rabbitt, M., Hales, L., & Reed-Jones, M. (n.d.). *USDA ERS - Definitions of Food Security*. Retrieved November 27, 2024, from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/definitions-of-food-security/

Rummo, P. E., Guilkey, D. K., Ng, S. W., Popkin, B. M., Evenson, K. R., & Gordon-Larsen, P. (2017). Beyond Supermarkets: Food Outlet Location Selection in Four U.S. Cities Over Time. *American Journal of Preventive Medicine*, *52*(3), 300–310. https://doi.org/10.1016/j.amepre.2016.08.042

SAMPSON, R. J. (1985). Neighborhood and Crime: The Structural Determinants of Personal Victimization. *Journal of Research in Crime and Delinquency*, *22*(1), 7–40. https://doi.org/10.1177/0022427885022001002

Singleton, C. R. (2024). Exploring the Interconnectedness of Crime and Nutrition: Current Evidence and Recommendations to Advance Nutrition Equity Research. *Journal of the Academy of Nutrition and Dietetics*, *124*(10), 1249–1254. https://doi.org/10.1016/j.jand.2024.03.008

Singleton, C. R., Winata, F., Adams, A. M., McLafferty, S. L., Sheehan, K. M., & Zenk, S. N. (2022). County-level associations between food retailer availability and violent crime rate. *BMC Public Health*, *22*(1), 2002. https://doi.org/10.1186/s12889-022-14415-y

Smith, S., Ferguson, C. J., & Henderson, H. (2022). An Exploratory Study of Environmental Stress in Four High Violent Crime Cities: What Sets Them Apart? *Crime & Delinquency*, *68*(11), 2092–2114. https://doi.org/10.1177/00111287211057858

Thornton, R. L. J., Glover, C. M., Cené, C. W., Glik, D. C., Henderson, J. A., & Williams, D. R. (2016). Evaluating Strategies For Reducing Health Disparities By Addressing The Social Determinants Of Health. *Health Affairs*, *35*(8), 1416–1423. https://doi.org/10.1377/hlthaff.2015.1357

Tung, E. L., Wroblewski, K. E., Boyd, K., Makelarski, J. A., Peek, M. E., & Lindau, S. T. (2018). Police‐Recorded Crime and Disparities in Obesity and Blood Pressure Status in Chicago. *Journal of the American Heart Association*, *7*(7), e008030. https://doi.org/10.1161/JAHA.117.008030

Wang, L., Hong, A., Chen, J., & Tak, A. (n.d.). *Risk Analysis of the Causes of Food Deserts*. Retrieved November 18, 2024, from https://dr.soa.org/49ce3a/globalassets/assets/files/static-pages/research/arch/2024/arch-2024-2-wang.pdf

Wolfe, S. E., & Pyrooz, D. C. (2014). Rolling Back Prices and Raising Crime Rates? The Walmart Effect on Crime in the United States. *The British Journal of Criminology*, *54*(2), 199–221. https://doi.org/10.1093/bjc/azt071

Zohrabi, N., Linkous, L., Eini, R., Adhikari, S., Keegan, B., Jones, J. C., Gooden, B., Verrelli, B. C., & Abdelwahed, S. (2021). Towards sustainable food security: An interdisciplinary approach. *2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI)*, 463–470. https://ieeexplore.ieee.org/abstract/document/9604394/