Breaking the Cycle: Reducing Recidivism in Iowa State Prisons

Lou Zhou, Amar Kanakamedala, Henry Yu, Jeffrey Liu, Evan Wu
Modeling the Future Challenge 2023

Contents

1 Executive Summary

The cycle of crime is a pervasive issue in American society. Nearly half of all crimes are committed by repeat offenders[51] because incarcerated individuals are often released from prison with insufficient support. The state of Iowa is not immune to this trend: over one in three inmates released from Iowa State Prisons found themselves back in prison within three years in 2022[28]. This high rate not only has obvious detrimental consequences for offenders, like spending years in confinement, but it also has financial costs for Iowa's citizens. To combat these issues, this work aims to identify the salient characteristics that lead to prison recidivism. We will also predict the monetary costs of recidivism for the current population of Iowa state prison and provide cost-effective, data-driven recommendations for reducing those costs.

Utilizing information provided by the Iowa Department of Corrections (IDOC), we trained a feed-forward neural network (FNN) to predict the likelihood that an inmate, upon release, will re-offend. We incorporated various inmate-specific parameters such as age, gender, and type of crime; parameters describing the state of Iowa Prisons (e.g. overpopulation) , as well as parameters related to the inmate's home county such as median personal income. Our model achieved an AUC-ROC score of 0.849 based on historical data.

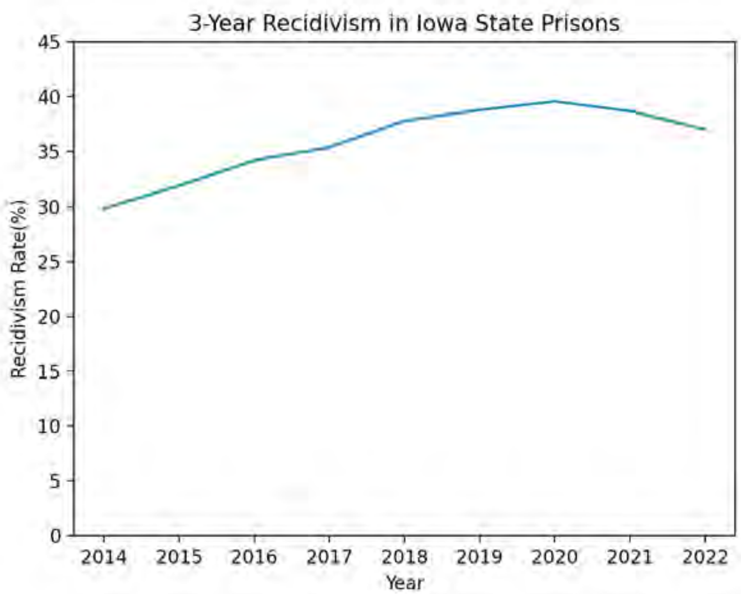
To evaluate the monetary costs associated with prison recidivism, we first used past studies to deduce the costs of each criminal classification[45][36][38][37]. Next, we created a heat-map to find the distribution of crimes committed by classification, which gave us information on whether Iowa crime was concentrated in a specific class of felony. Combining the probability that a given inmate will recidivate with financial estimates, we predicted that the cost of recidivism will be over $348 million. To make our recommendations more practical and cost-effective, we stratified costs by state prison and type of previous crime, targeting at-risk inmate populations and prisons.

After discerning which crimes had the greatest monetary cost, we analyzed our model's output with SHAP feature importance analysis to find which variables contribute most to prison recidivism. With this information, we formulated four major recommendations to reduce recidivism rates and associated monetary costs.

Given the significant correlation between prison overpopulation and recidivism rates, we propose that the Iowa Department of Justice increase its implementation of alternative correctional programs, such as probation and community service, and reduce the use of mandatory minimums and determinate sentencing in the Iowa state code. Furthermore, because of a high correlation between drug offenses and recidivism rates, we suggest that the IDOC allocate increased funding and support for drug and behavioral therapy programs. We also urge that the IDOC increase education and counseling opportunities, particularly in maximum security prisons where these programs are lacking the most and expected losses are highest. Finally, we advise that the IDOC provide employment opportunities and mentorship for at-risk populations that we identified in the paper to improve their chances of successful reintegration into society. By following these recommendations, we believe that Iowa can make significant progress in reducing recidivism rates and creating a safer environment for all its citizens.

2 Introduction and Background Information

Recidivism, or re-offense, is a critical matter in the United States, and Iowa is no exception. Too regularly, inmates are released from prison, fail to receive the support to properly re-integrate into society, fall back into a life of crime, and end up back behind bars, perpetuating a crime cycle where offenders are constantly in and out of prison. Unfortunately, in the state of Iowa, this cycle is worsening considerably. While improving within the last two years, the 3-year recidivism rate, which counts inmates who have been released from an Iowa state prison and are re-incarcerated into any state or federal prison, has increased by 7.3%, from 29.8% in 2014 to 37.0% in 2022[28], so it is still imperative that steps be taken to break the cycle.



This cycle of recidivism bore a significant economic cost to Iowa. In 2017, the annual loss per inmate in some Iowan prisons exceeded $50,000[21]. With inmates constantly moving between a life behind and beyond bars, this cycle of recidivism will continue to cost the Iowa government and Iowan taxpayers millions of dollars.

More importantly, recidivism has taken a social toll on Iowan society. In this cycle, offenders are unable to successfully re-integrate into the community, spending significant amounts of time behind bars and nearly eliminating any chance of contributing positively to society. With each re-offense, the cycle creates a new victim, directly damaging the lives of the everyday Iowan citizen.

This cycle is especially pressing to the state of Iowa as it is estimated that the Iowa prison population will increase by 39% by 2024, where a vast majority will be released from prison in the future[6]. With such a large number of offenders leaving the Iowan prison system, it is crucial that this cycle be broken, for the sake of both the offender and the Iowan citizen.

In an analysis of released Pennsylvania state prisoners[47], case-specific variables like age, ethnicity, gender, and type of crime committed were strong indicators in predicting whether an inmate will re-offend, with younger, male inmates having a history of crime at the highest risk for re-offense.

Previous works also indicate that prison and geographical information has a strong influence on an offender's likelihood of re-offending. While prison size has been found to influence re-offense rates, overcrowding was shown to have a strong negative relationship with recidivism rate[12]. In addition to economic opportunities[46], studies also point to the availability of vocational programs[2] and the level of security as key factors which may affect a prisoner's likelihood of recidivism[39].

This work seeks to analyze how the factors found above affect an offender's probability of re-offending and estimate the fiscal cost of state prison recidivism to the state of Iowa. We will also find certain at-risk groups with high re-offense rates to provide cost-effective, data-driven recommendations to help "break the cycle," and make Iowa a safer place. However, this work does not account for inmates who may have re-offended but were not caught.

3 Data Methodology

3.1 Data Collection

To analyze the factors contributing to prison recidivism in Iowa state prisons, we used open government datasets provided by the Iowa Department of Corrections, the latest annual reports and admissions from all nine Iowa state prisons, per capita unemployment insurance claims, and per capita personal income, provided by the Iowa Workforce Development. Because the datasets are published and updated regularly by Iowa government sources, allowing errors and duplicates to be revised and corrected, the data is reliable enough to be used for analysis.

By using these sources, we can establish the relationship between recidivism and other factors, helping us to develop a model which can predict the probability that a prisoner will re-offend once released.

Iowa Prisoner Release Data[20]

- **Type of Data:** Released Inmates from the Iowa state prison system since 2011.
- **Source:** Iowa Department of Corrections
- **Variables:** Sex, Age, Ethnicity, Type of Crime, Year Released, Prison, Jurisdiction, Type of Release.
- **Purpose:** This dataset gives us a per-person view of releases from the Iowa state prison system. We use this dataset of 55,595 different prison releases to count the number of prisoners who have and have not re-offended. By doing so, we can analyze whether a prisoner re-offends because of some combination of parameters that have been found below.

Iowa Current Prison Population Data[19]

- **Type of Data:** The current population of the Iowa state Prison Population
- **Source:** Iowa Department of Corrections
- **Variables:** Sex, Age, Ethnicity, Type of Crime, Year Released, Prison, Jurisdiction, Type of Release.
- **Purpose:** This dataset, similar to the dataset of released prisoners, will be used to predict recidivism in Iowa state prisons in the future.

State Prison Annual Reports[23]

- **Type of Data:** Prison-specific data for each Iowa state penitentiary
- **Source:** Iowa Department of Corrections
- **Variables:** Population, Capacity, Number of Correctional Staff, Treatment Staff, Type of Prison
- **Purpose:** These reports display data specific to the prisons, like the number of correctional staff, allowing us to consider differences in each prison that may influence an ex-inmate's likelihood of re-offending (e.g., varying emphasis on education and therapy).

County and State Level Unemployment Insurance Claims[25][24]

- **Type of Data:** Information on unemployment insurance claims for different Iowa counties
- **Source:** Iowa Workforce Development
- **Variables:** Number of Claims per County, County FIP, GNIS Feature ID, County Coordinates, County Latitude, County Longitude, Fiscal Value of All Claims Paid
- **Purpose:** Iowa Admissions Data contains the jurisdiction of where a crime was committed. This source lists the number of unemployment benefit claims in each Iowa county, which we can combine with the data on crimes to quantify the relationship between recidivism and job stability. To predict these parameters in the future, we rely on state-level unemployment claims starting from 2011 as inmates may relocate after release. While the unemployment rate is a conventional statistic, analyzing unemployment claims per capita allows us to accurately determine the severity of unemployment in a jurisdiction.

County and State Level Median Household Income[26][27]

- **Type of Data:** Average median income per household in each Iowa county.
- **Source:** Iowa Workforce Development

- **Variables:** Median Household Income
- **Purpose:** Being in a state of poverty correlates to a much higher possibility of committing a crime compared to those who are not for the simple reason of survival. The same complication can be extended to ex-offenders, who often are released with no connections or jobs and therefore end up below the poverty line. Hence, the economic status of the location a prisoner resides may significantly influence the odds of recidivism. Like unemployment insurance claims, we use state-level data starting from 2011 to predict these values in the future to account for the potential movement of inmates post-release.

3.2 Data Cleaning

Using the Python code in Appendix 1, we combine the household income, unemployment benefits, and prison datasets using Python with the Iowa Admissions Data so that a complete, unified dataset can be formed. To contain appropriate data, we remove all inmate releases either as a result of acquitting in prison as well as any inmates where portions of their data has been removed due to privacy concerns. The complete dataset of 51,097 releases will then have information about the inmate, the county in which the crime occurred, and the prison in which the inmate resides. We can determine whether an individual re-offended if the inmate's offender number appears multiple times within the dataset.

To prepare categorical data for the model, we use a process called one-hot encoding, which uses dummy variables to convert each possible value into a binary representation, using 0 or 1 to indicate the absence or presence of a value respectively.

4 Mathematics Methodology

4.1 Assumptions

Assumption: Offenders spend a significant amount of time in the jurisdiction they commit the crime.

- **Justification:** In a work by Curtis-Ham et. al.[10], it was found that offenders were generally more likely to commit crimes closer to places where they spent a large amount of time during their day (e.g. their home or place of work). In such places, offenders are familiar with the area, have access to local resources, and therefore are more likely to commit crimes.

Assumption: The ratio of treatment staff to prisoners is correlated with the quality of treatment or education an inmate would receive in prison.

- **Justification:** The quality of treatment and education received by inmates in prison is positively correlated with a higher ratio of treatment staff to prisoners. This is because a higher ratio enables staff members to have more frequent meetings with inmates, providing them with individualized attention. On the other hand, a smaller ratio can cause staff members to be overworked, leading to a lower quality of treatment programs.

Assumption: Changes month-by-month in per capita personal income are negligible.

- **Justification:** We utilize per capita personal income as a metric to account for the potential influence of wealth on an individual's likelihood of re-offending. Given the limitations of data, we can assume that an individual's monthly per capita income can be determined from their annual income, as provided by Iowa. This is a reasonable assumption because it is unlikely that there will be significant fluctuations in a person's income over a year. Moreover, this approach enables us to assess the income disparities across various counties in Iowa and their effects on recidivism.

Assumption: Prison statistics remain relatively constant throughout the years

- **Justification:** Based on reports from the Congressional Research Service and the Bureau of Justice Statistics, there was little change in the number of incarcerated inmates from 2000 to 2014[7]. Similarly, the age and sex of the arrested individual also stay relatively consistent. It is also impossible to predict whether treatment or correctional programs will change in the future, so we assume it remains constant.

4.2 Variables

Personal Variables

Variable	Description
A	Age of Inmate at Release
G	Gender of the Inmate
E	Ethnicity of the Inmate
T	Type of Crime Committed
N	Number of Re-offenses

Prison and County Variables

Variable	Description
O	Amount in which a Prison is Overpopulated
R_C	Ratio of Correctional Officers to Inmate
R_T	Ratio of Treatment Staff to Inmate
U	Dollar Amount of Unemployment Insurance Benefits
I	Personal Income per Capita

4.3 Binary Classification Using a Neural Network

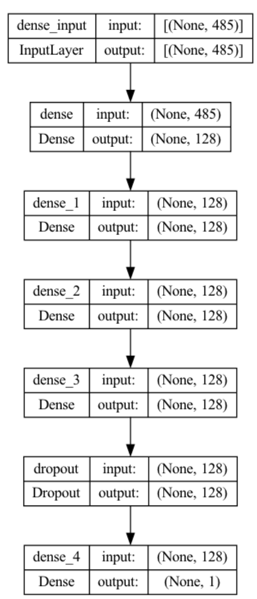
Using the code in Appendix 2, we build a binary classifier with a feed-forward neural network (FNN) from the Keras Python framework[49] to predict the chance that a released prison will re-offend. A FNN receives input, weighs them respectively, applies a mathematical function, and then passes the output to the next layer of nodes. The final layer will produce the neural network's output.

The weights applied by each neuron are learned during the training of the dataset using a supervised learning algorithm. We use an algorithm called backpropagation [55], which attempts to minimize the loss of a model through the adjustment of these weights using gradient descent. The scale of adjustment is then found by determining the gradient of the loss of the model with respect to each parameter. In the case of finding probabilities, it is common to look at binary cross-entropy loss, which can be defined as:

$$L = -\frac{1}{m}\sum_{i=1}^{m}(a_i \cdot \log(p_i) + (1 - a_i) \cdot \log(1 - p_i)) \tag{1}$$

where,
m = number of training samples
a = actual probability
p = predicted probability

In our case, we trained a five-layer neural network using a combination of dense layers with sigmoid activation functions, layers that take input from every neuron of the preceding layer and apply a mathematical function, and dropout layers, which randomly drop out a certain percentage of neurons in the previous layer to prevent overfitting. A visualization of the model can be found below:



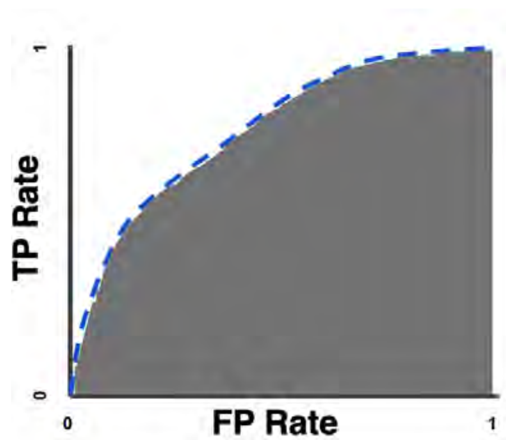
4.4 Evaluating the Model

4.4.1 AUC-ROC Score

Traditionally in literature[1][18], the AUC-ROC Score is used to evaluate the strength of a model's prediction of probabilities. The AUC-ROC Score is defined as the area under the curve (AUC) of the ROC curve, which plots the true positive rate against the false positive rate. A true positive occurs when the model correctly predicts that an ex-offender re-offends or doesn't, whereas a false positive occurs when the model predicts recidivism when it does not occur, and vice versa. They are defined as follows,

$$TPR = \frac{TP}{TP + FN} \quad \text{and} \quad FPR = \frac{FP}{FP + TN}$$

Ideally, the probability distributions for True and False rates do not intersect, which would yield an AUC-ROC graph with area 1. If they do intersect, there is the existence of Type I and Type II errors, and the AUC-ROC score decreases.



For our purposes, the AUC-ROC Score is a more suitable statistical measure than other absolute measures such as accuracy. This is because the AUC-ROC Score takes into account the probability of the predictions, rather than just the classification of the predictions. In our case, the accuracy of the model in estimating the true probability is more significant than its ability to predict whether a prisoner will re-offend.

The AUC-ROC Score ranges from 0 to 1, with 0 meaning the model is giving opposite predictions for each sample, 1 meaning the model is giving correct predictions for each sample, and 0.5 meaning that the model is unable to distinguish between correct and incorrect. In our case, we received an AUC-ROC score of 0.84928, which is strong, especially with the highly chaotic nature of crime and recidivism.

4.4.2 Comparison to Logistic Regression

In addition to an AUC-ROC Score, we also seek to evaluate our neural network against a base model which serves as a benchmark. In our case, we use a logistic regression model, which is a simple method for binary classification. The regression models the relationship between the dependent and independent variables using a sigmoid function and is especially effective at modeling the likelihood of an event taking place. In our case, this corresponds to predicting the likelihood of an ex-offender recidivating based on a range of parameters. Traditionally, a multivariate logistic regression is given by the equation

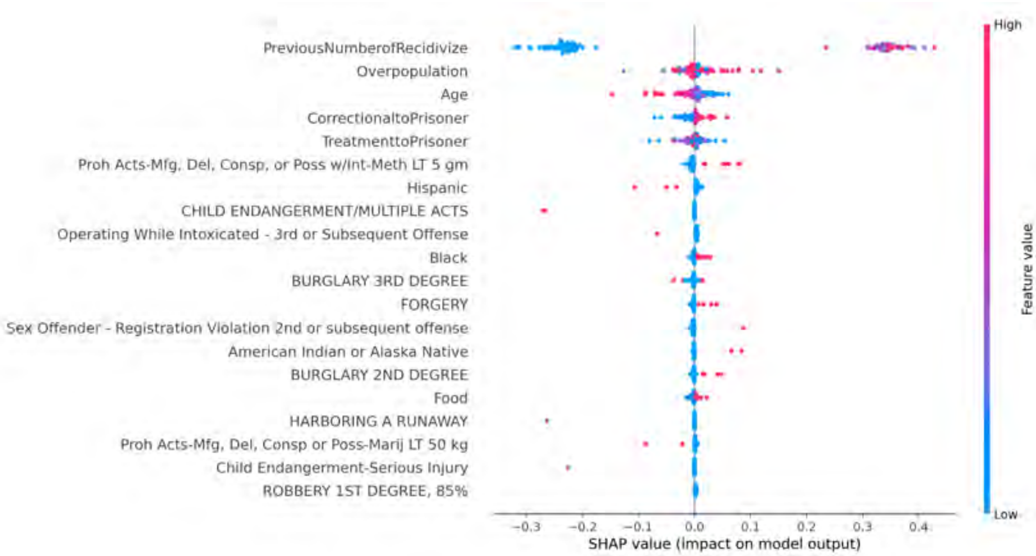
$$P(Y = 1|X_1, X_2, \ldots, X_p) = \frac{1}{1 + e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_p X_p)}} \tag{2}$$

where
X_i = independent variable
β_i = associated weight

The logistic regression yielded an AUC-ROC Score of 0.64, which is outclassed by the FNN's score of 0.85.

4.5 Feature Importance

Due to the lack of feature importance for FNNs in the Keras framework, we instead use the SHAP (SHapley Additive exPlantations) methodology, which utilizes Shapley values to determine the contribution of each feature to the model's output for a given input[33]. The SHAP value is a metric that indicates the effect of a feature on the model's prediction, while feature importance measures the relative significance of each factor in the model's fit.
The SHAP analysis for 100 predictions is presented below:



Consistent with traditional literature[31], the SHAP analysis found that having a history of crime and recidivism (N) are the most critical factors in predicting recidivism. Additionally, our work agrees with the work of Farrington and Nuttall[12], which shows that the status of overpopulation (O) in prisons significantly influences recidivism rates.

The ratio of prisoners per correctional officer (R_C) has a significant positive SHAP value, showing that punishment and prison sentences may exacerbate existing issues and increase the likelihood of recidivism. In contrast, the ratio of prisoners to treatment staff (R_T) has a significant negative SHAP value, implying that effective treatment can improve an ex-offender's situation and reduce their chances of re-offending.

Moreover, the type of crime (T) was found to be a significant feature in predicting the probability of recidivism. In accordance with literature[47][11], it was found that property and drug crimes were noteworthy predictors which increased an inmate's likelihood of re-offending.

4.6 Modeling Length of Sentence

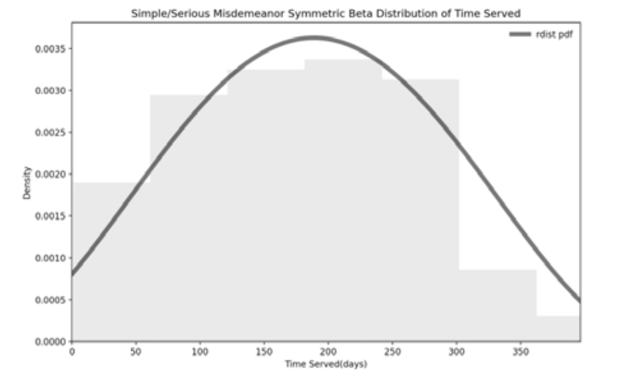
Because the Monte Carlo simulation requires time-sensitive geographical data, we must be able to predict release dates by modeling the length of the sentence based on the classification of the offender's crime. To effectively model this, we use probability density functions for each classification using the *fitter* Python package, which fits a dataset with 106 common probability distributions, returning its maximum likelihood estimations (MLE) and sum-square error (SSE).

To ensure the distributions fit the data, we employ the Kolomogorov-Smirnov (KS) test, which is a nonparametric, distribution-free test that conservatively estimates how likely samples chosen randomly would be similar to the dataset tested; the higher the test p-value, the more consistent the random samples are to the distribution being tested. To select a model, we find the distribution with the lowest sum square error whose KS test p-value was above $\alpha = .05$.

Because the p-value will naturally decrease as sample size increases[50], thus making every p-value statistically significant, we use the *fitter* algorithm with a random sample of size $n = 200$ for each crime classification.

Graphs for each classification of crime, the calculated maximum likelihood estimations (MLE), KS test p-values, the SSE, and the equations, with x being the number of days served, for each distribution can be found below:

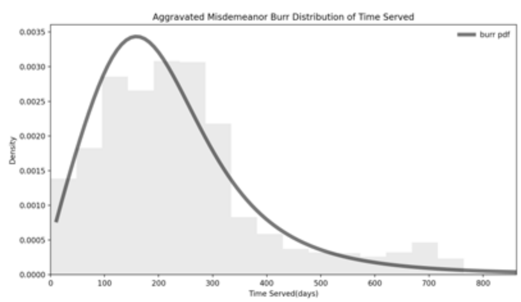
Simple / Serious Misdemeanor - Symmetrical Beta Distribution[1]



$$f(x) = \frac{(1-x^2)^{\frac{c}{2}-1}}{\beta(\frac{1}{2}, \frac{c}{2})}$$

	Parameter	Value
MLE	c	6.5689
Evaluation	KS-test p-value	0.606725
	SSE	0.000077

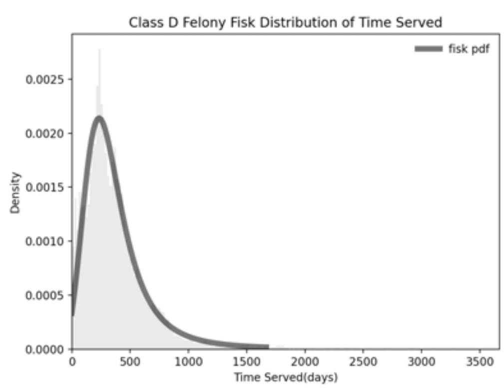
Aggravated Misdemeanor - t-distribution



$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

	Parameter	Value
MLE	ν	1.996570
Evaluation	KS-test p-value	0.25190
	SSE	0.00002

Class D Felony - Fisk Distribution

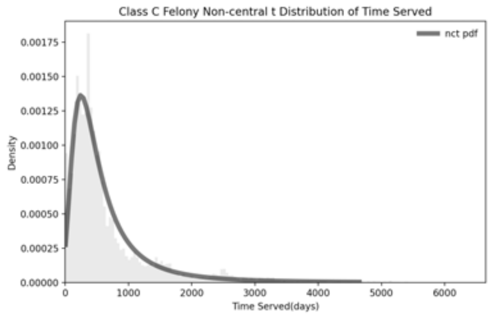


$$f(x) = \frac{cx^{c-1}}{(1+x^c)^2}$$

	Parameter	Value
MLE	c	3.12507
Evaluation	KS-test p-value	0.957
	SSE	0.000015

Class C Felony - Non-Central t-distribution

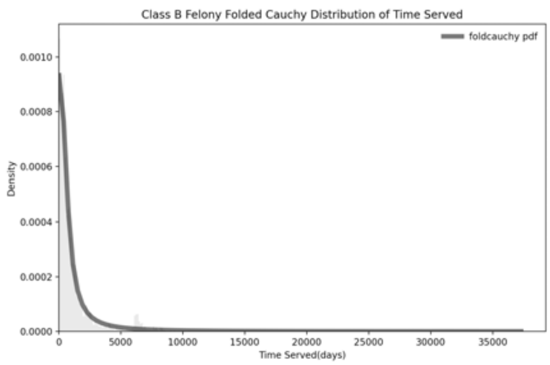
[1]Because of the small number of inmates (6) with simple misdemeanors in Iowa state prisons and the similarity of crimes, we combined the simple and serious misdemeanors in one distribution.



$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{\delta^2}{n}\right)^{-\frac{n+1}{2}} \cdot \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} \cdot \frac{{}_2F_1\left(\frac{n+1}{2}, \frac{1}{2}; \frac{n+2}{2}; \frac{\delta^2}{\delta^2+nx^2}\right)}{\sqrt{n}\left(1+\frac{\delta^2}{nx^2}\right)^{\frac{n+1}{2}}}$$

	Parameter	**Value**
MLE	n	2.22062
	δ	5.342343
Evaluation	KS-test p-value	0.882
	SSE	0.000002

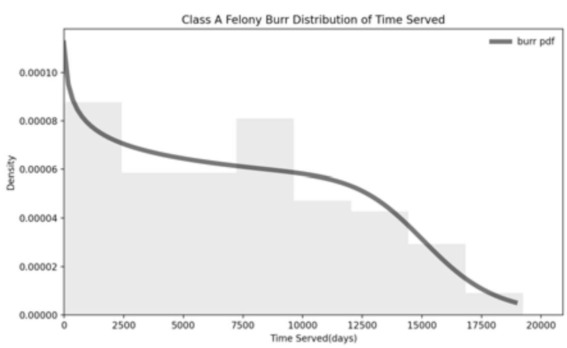
Class B Felony - Folded Cauchy Distribution



$$f(x) = \frac{1}{\pi(1+(x-\gamma)^2)} + \frac{1}{\pi(1+(x+\gamma)^2)} \quad \text{for} \quad x \geq 0.$$

	Parameter	**Value**
MLE	γ	0.4065447
Evaluation	KS-test p-value	0.879
	SSE	5.44×10^{-7}

Class A Felony - Burr Distribution



$$f(x) = c \cdot d \frac{x^{-c-1}}{(1+x^{-c})^{d+1}}$$

	Parameter	**Value**
MLE	c	10.43455
	d	0.083109
Evaluation	KS-test p-value	0.889577
	SSE	8.12×10^{-8}

4.7 Predicting County Parameters in the Future

In order for our model to effectively predict trends in recidivism rates, it is important that we find and use reasonable and relatively accurate values for county parameters in the future. To predict these values, we look to use a combination of regressions and distributions to accurately model each parameter in the future.

Unemployment Benefits Paid per Capita

Because unemployment benefits paid remained relatively constant, with any variation independent of time, we fit a Skewed-Cauchy distribution, with a KS p-value of 0.9848 and a sum-square error of 0.0162, of values of the value of unemployment benefits paid per capita(U) from 2000 to 2022, dropping the notable outlier of the 2020 COVID Year, generating a distribution where x is defined as the per capita dollar value of unemployment benefits paid, with $a = -0.04439$ where:

$$f(x) = \frac{1-a}{2} + \frac{1-a}{\pi} \arctan(\frac{x}{1-a}), x < 0. \tag{3}$$

We then find projections of U by randomly generating numbers within the Skewed Cauchy distribution for each year.

Changes in Median Household Income over Time in Iowa

Using State Median Household Income in Iowa, there exists an increasing linear relationship existed between Time and Median Household Income in Iowa.

Because of the linearity in the data, we fit a simple linear regression model to predict future income values from historical data. Using the StatsModel and sklearn libraries, we trained a model with an r^2 of 0.985, and we found that the annual increase is represented by the equation, where t is the calendar year:

$$1355.6136(t-1985) + 18030. \tag{4}$$

4.8 Strengths and Weaknesses

The strength of our model stems from its ability to handle diverse information; it is able to represent complex relationships between its parameters and recidivism, and accurately predict, with an AUC-ROC score of 0.849 the probability that an inmate will re-offend even with the often chaotic nature of recidivism and crime.

Another strength of our model is its scalability and flexibility. Our FNN can easily be scaled up to include other important variables and can inherit new information after it has been built with minimal performance changes in a relatively cost-effective manner. Because the neural network is not limited, the dataset is not restricted.

Because our model operates on a person-by-person basis using input data on documented offenders from Iowa, including information on sex, race, age, type of crime, and jurisdiction, we can generate a personal prediction of re-offending. This enables us to create a targeted intervention program based on relevant background information, rather than a broad approach.

A major weakness of the model is the neural network's failure to account for other important personal factors like the number of living family members, potential gang affiliations, or mental health issues. Because of privacy laws, it is often difficult for the public to find information on these factors for each ex-offender, so our model cannot train on and therefore account for these variables.

Furthermore, the FNN is a black-box model, meaning the internal computations and logic of the model are not interpretable to the user. We are only able to discern the input and output, but the steps it took to achieve these results are not easily understood. By using the SHAP mathematical framework, we can mitigate these effects by gaining a per-prediction understanding of the neural network, but we cannot generate real feature importance, as found in other algorithms like the Decision Tree or XGBoost.

5 Risk Analysis

5.1 Risk Overview

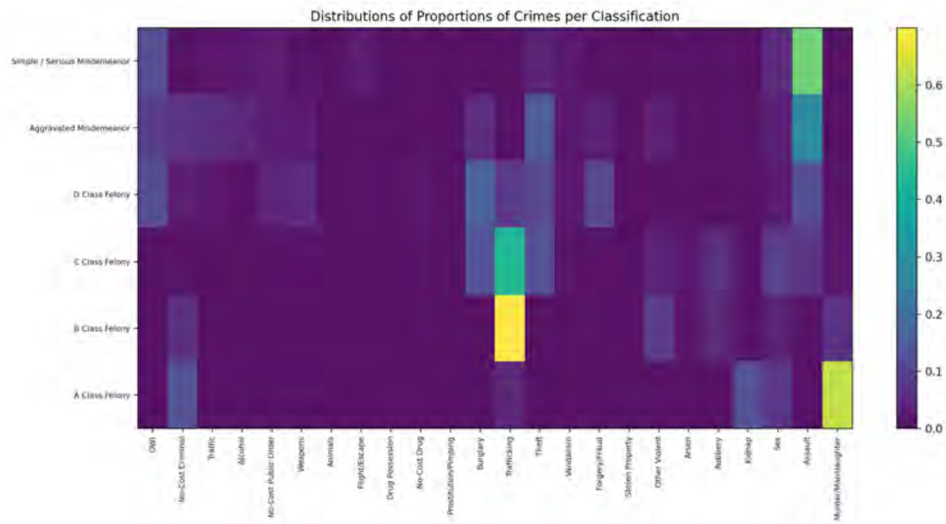
To properly evaluate the loss due to recidivism in Iowa in the future, we apply the model as detailed in 4.4.1 to a dataset containing current inmates within the Iowa state prison system. Because future crimes and therefore prison intake are impossible to predict, we instead evaluate, using a Monte Carlo Simulation, the expected loss that recidivism will induce for the current population of Iowa state prisoners.

Additionally, we use the Monte Carlo simulation to determine certain high-risk groups which have the highest expected loss, by finding the expected losses by an inmate's current subcrime (the crime they have committed) as well as expected losses for all nine Iowa state prisons. By finding these certain at-risk groups, we can create our recommendations to be as cost-effective and efficient as possible.

5.2 Quantifying Financial Loss to Crimes

Because crime is not specialized (i.e. re-offenders are not more likely to commit a crime they have committed before)[30], we find the average loss caused by a crime categorized by criminal classification level (e.g. A Class Felony, B Class Felony, etc.) to properly and conveniently quantify the financial loss as a direct result of recidivism.

To quantify the expected loss for each class of crime, we separate each classification into a ratio of criminal sub-types. A heatmap of crime sub-type proportions for each crime classification can be found below, with exact percentages found in Appendix 4:



We then found and assigned all possible tangible and intangible losses (e.g., victim costs, justice system costs, career costs, grief) for each crime sub-type according to the works of Rajkumar, Miller and McCollister, et.al[45][36][38][37]. The costs for each sub-type can be found below[2], with inflation being accounted for:

Violent Crimes

Subcrime	Cost
Assault[36]	$331,914.50
Sexual Crimes[36]	$331,911.75
Arson[36]	$29,090.00
Manslaughter / Murder[36]	$12,383,013.00
Kidnapping[38]	$124,222.84
Other Violent Crimes[38]	$17,781.34

Larceny Crimes

Subcrime	Cost
Burglary[45]	$2,758.54
Theft[36]	$4,868.89
Fraud[36]	$6,936.65
Stolen Property[36]	$10,992.23
Robbery[36]	$58,324.69

[2] In situations where the cost of the exact crime is unknown, the cost of the closest crime in type and severity is used.

Illicit Drug Crimes

Subcrime	Cost
Drug Trafficking[37]	$48,855.75
Drug Possession[45]	$28.95
OWI[38]	$5,148.24
Alcohol[37]	$123.91
No-Cost Drug Offenses[3]	$0

[3] In our case,"no-cost crimes" are those which have no victims and have negligible cost to the economy and government.

Public Order Crimes

Subcrime	Cost
Vandalism[36]	$6,699.55
Weapons[36]	$28.95
Prostitution[45]	$90.96
Traffic[4]	$0
Flight[5]	$0
No-Cost Other Offenses	$0

[4] Negligible costs; traffic violations warranting state prison are carried out in conjunction with a more serious crime (e.g. speeding while committing a felony).
[5] This mostly involves crimes like failure to appear, which have a negligible cost. Other crimes, like prison escapes, are impossible to calculate the cost of and are so infrequent that their costs are disregarded.

The expected loss for each criminal classification is then found by using the traditional expected value formula,

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \cdot p_i \tag{5}$$

where $\mathbb{E}[X]$= expected economic loss, x_i= loss due to event, and p_i=probability of event.

Classification	Cost
A Class Felony	$7,755,510.88
B Class Felony	$768,295.65
C Class Felony	$232,772.64
D Class Felony	$148,201.24
Aggravated Misdemeanor	$112,363.64
Simple / Serious Misdemeanor[6]	$194,238.19

[6] Despite being a less serious classification,"Simple / Serious Misdemeanor" outranks both "Aggravated Misdemeanor" and "D Class Felony" in cost because the latter two involve a substantial amount of repeat offenses of relatively low-cost crimes (e.g. OWI, drug possession), while "Serious Misdemeanor" has a higher number of first-time assault offenses, which are high-cost crimes.

5.3 Change of Crime Classification

It is also quite obvious that not all re-offending convicts will commit a crime that falls under the same classification as their previous conviction. To account for this change, we use previous re-offending data to determine the probability that a convict, given that they re-offend, will commit a crime with differing severity, with positive values meaning more severe and negative values meaning less severe. Exact percentages can be found in Appendix 5.

These proportions will be used for the Monte Carlo Simulation in 5.4 to account for potential changes in a re-offender's crime type. Using these values, we can apply them to the dataset of the present population of Iowa state inmates for our Monte Carlo Simulation.

5.4 Monte Carlo Simulation

To accurately predict the expected costs of recidivism, we run a Monte Carlo simulation, which allows us to use the law of large numbers to account for randomness in the criminal justice system, like changes in the classification of crime upon re-offending (5.3) and distributions of sentence length (4.6).

For each trial, the simulation applies the FNN on the current Iowa prison population dataset, using relevant county data and release dates for each prisoner from the distributions and regressions found in 4.6 and 4.7, generating a probability that the given inmate would re-offend.

Future crime costs, given that a prisoner would re-offend, are generated by predicting the classification of crime, using the probabilities and associated costs from 5.2 and 5.3. The model then finds the product of the cost and the probability, generating an expected cost of recidivism for each prisoner. The expected loss for the entire present population is then found using Equation (3).

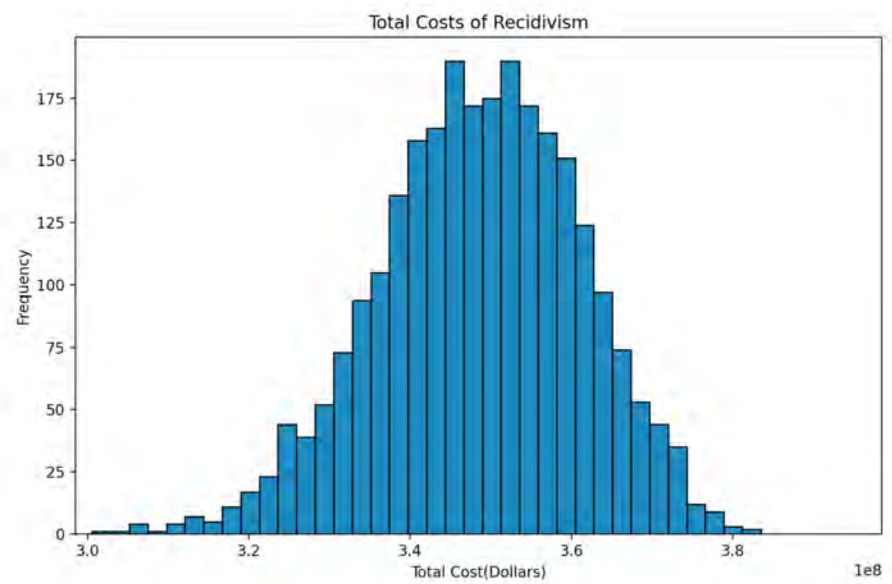
Because of the computationally expensive characteristics of our simulation, it is wise to find a mathematically-efficient number of trials. Using the work of Liu[32], the number of trials can be found where:

$$c = \frac{\phi}{s \cdot \sqrt{n}} \tag{6}$$

where:
c = confidence level z-statistic
s = sample standard deviation
n = number of trials
ϕ = precision

Using a 95% confidence level and a precision of $\phi < \$500,000$, we find the z-statistic one standard deviation from the mean, with $c = 1.96$. Additionally, s is found by calculating the standard deviation of a 500 trial Monte Carlo Simulation, where $s = \$1.301 \times 10^7$. Utilizing Equation (6), we determine the appropriate number of simulations to be $n = 2603$.

The results from of the 2603 trial simulation, with a precision of $\phi < \$500,000$ and a 95% confidence level are shown below:



Projection	Total Cost ($)
$\mu - \sigma$	$\$3.358 \times 10^8$
μ	$\$3.484 \times 10^8$
$\mu + \sigma$	$\$3.610 \times 10^8$

5.5 At-Risk Groups

To better direct our recommendations, we look to find which state prisons and previous offenses lead to a higher expected loss. We look to analyze certain at-risk groups as opposed to trends of recidivism in the future to better understand the effects of recidivism.

5.5.1 State Prisons

Since an inmate's experience and therefore their re-offending chance will differ depending on the state prison they are sentenced to, it is noteworthy to find the estimated loss for each prison. Since inmates can and regularly transfer between institutions, we stratify our Monte Carlo Simulation by the inmate's intaking state prison, finding the total cost and cost per prisoner for the intaking prison populations of all nine state prisons.

Prison	Total Cost ($)	Cost per Inmate
Iowa State Penitentiary	$60,634,532.10	$279,421.81
Anamosa State Penitentiary	$43,346,642.20	$228,140.22
Mt. Pleasant Correctional	$1,733,770.19	$216,721.27
Fort Dodge Correctional	$1,361,735.45	$170,216.93
North Central Correctional	$939,820.36	$156,636.73
Newton Correctional	$5,135,303.39	$131,674.45
Clarinda Correctional	$2,205,968.37	$129,762.85
Iowa Medical & Classification Center	$215,970,772.71	$31,565.44
Iowa Correctional for Women	$17,579,453.35	$27,727.84

According to the Monte Carlo analysis, maximum security prisons like Iowa State Penitentiary and Anamosa State Penitentiary yielded the highest expected loss. Exemplified by the murder of two staff members in an escape attempt in 2021[44], these prisons often contain high-risk, violent inmates who are naturally larger threats to society if released.

Because of their size, maximum security prisons inherently focus more on securing the inmate population than rehabilitation and release programs[34], with higher numbers of Correctional Officers and smaller numbers of treatment staff. This combination of high-risk inmates and therefore a focus on prison security naturally yields a higher risk of recidivism.

On the other hand, minimum security prisons like the Iowa Correctional Institution for Women naturally have a lower expected loss for the same reasons. With a lower-risk prison population, the prison will focus more on rehabilitation and release programs, leading to a lower recidivism rate and expected loss.

However, a notable exception is the Iowa Medical and Classification Center (IMCC), a medium-security facility with a smaller expected loss than some minimum-security prisons, like the North Central Correctional Facility. Although IMCC may contain high-risk criminals, the facility also has the highest treatment staff-to-inmate ratio, which implies that rehabilitation-focused programs can succeed in higher security facilities.

5.5.2 Type of Crime

Because previous crime is a strong predictor of the probabilities of recidivism[31][47], we also found the highest estimated loss by stratifying the Monte Carlo Simulation by current offense. The five current offenses with the highest expected losses are found below.

Offense Type	Total Cost ($)	Cost per Inmate
Kidnapping	$32,719,213.00	$177,821.80
Murder / Manslaughter	$191,069,486.37	$161,376.26
Sex Crimes	$40,500,628.97	$35,651.96
Robbery	$16,716,781.70	$27,137.63
Drug Trafficking	$22,525,719.45	$21,742.97

It is not surprising that Class A felonies (kidnapping and murder) hold the place for the top two most expensive crimes. Based on our analysis in 5.3, A-Class felons who re-offend are most likely to commit high-severity, high-cost A-Class felonies. Because of the high cost of these crimes, the expected loss for current Class A felonies is notably higher. However, A Class felons are also extremely unlikely to re-offend, as, according to Iowa State Law, A Class felons are punishable only by life sentences[35] with no probation, meaning that most inmates will not be released from prison once arrested, so cost per prisoner is still minimal compared to the actual cost of a Class A felony ($7,000,000+).

This same reasoning can be applied to Sexual Crimes, as although they have an extremely high tangible and intangible cost, their recidivism rates are much lower compared to general criminals. However, it is unlikely that this difference is due to these offenders successfully reintegrating into society. According to Przybylski, while lower than general criminals, those recidivism rates for sex offenders are most likely a gross underestimate because of police under-reporting[42], so it is probable that our cost per prisoner estimates for sexual criminals is a gross underestimate.

Additionally, routine crimes like Robbery and Drug Trafficking have high estimated costs, because offenders who committed these types of crimes are much more likely to re-offend because of external factors associated with these crimes (gang affiliation, poverty, substance dependency)[31]. These crimes also usually have high economic losses attached to them, yielding a high expected cost.

However, we must consider some of the technicalities associated with the nature of the sentences imposed on the convicts. Individuals who continually re-offend are subject to Felony Sentencing Enhancements, which may extend their sentence up to three times the normal amount[35]. Since many ex-offenders cannot re-offend due to the length of their sentence, these figures may be overestimations since some felons, who may commit lower-classification crimes, will never be released from prison.

6 Recommendations

Based on our analysis of prison recidivism, we identified four major recommendations for the Iowa Department of Corrections aimed at reducing recidivism rates and associated costs. The recommendations focus on two aspects of incarceration: the Iowa Department of Justice's sentencing and punishment of offenders, and the Iowa Department of Corrections' management of incarceration facilities and policy.

6.1 Overpopulation of Prisons

In accordance with existing research[12], our model indicates that prison overpopulation is a significant factor in high recidivism rates. This relationship can be explained by the adverse effects of overcrowding, such as poor living conditions, inadequate medical care, and an increased risk of violence[40]. Furthermore, correctional and treatment staff are overworked, leading to decreased morale and lower quality treatment and rehabilitation programs. This issue is particularly pressing in Iowa, where more than half of the state prisons are currently operating above maximum capacity. Given these connections, it is imperative to implement measures to alleviate overcrowding in prisons.

We recommend that the Iowa Department of Justice and Department of Corrections increase resources in implementing alternative correctional programs, like community service, probation, and Residential Correctional Facilities, especially for less serious misdemeanors (e.g., public order crimes and vandalism). These programs have demonstrated effectiveness in combating recidivism in low-severity offenders while also being less expensive than traditional incarceration[48][56].

Expanding on the previous recommendation, we also recommend the implementation of sentencing reforms in the Iowa Sentencing Code, specifically reducing mandatory minimums and determinate sentencing. This will provide judges and Department of Justice officials with greater freedom to consider both the circumstances of the crime and the needs of the offender when determining sentences, which will reduce admittance rates and length of sentences, leading to a decrease in overall state prison populations.

6.2 Drug and Behavioral Therapy

According to the U.S. Department of Justice, over one-third of inmates in state prisons have been diagnosed with mental illnesses, while over half have substance abuse issues[52]. These problems exacerbate issues faced by ex-inmates upon release, as they can lead to erratic behavior, impaired decision-making, decreased self-control, and social stigmatization, hindering opportunities such as employment, healthcare, and housing opportunities. Not surprisingly, both traditional literature[11] and our model agree that inmates with these issues are at a greater risk of re-offense. Considering the extensive and harmful impact of these problems, it is imperative for the Iowa Department of Corrections to implement measures aimed at resolving and minimizing them.

A reliable solution to substance abuse rehabilitation is drug and behavioral therapy, both during and after incarceration. Past research indicates that these treatments can help offenders manage their addictions and mental health[9], which can in turn mitigate their chance of relapse and facilitate their return to society.

As such, we recommend that the Iowa Department of Corrections increase funding and implementation of drug and behavioral therapies. Extending implementation of the Cognitive-Behavioral Interventions for Substance Abuse and the Seeking Safety programs to all state prisons, and increasing funding and support for the Moral Reconation Therapy program, which has been shown to reduce recidivism rates by 30% to 50%[22], would greatly benefit inmates struggling with these issues[3]. This will not only improve the lives of inmates, but also has the potential to save Iowan taxpayers millions of dollars. It has been estimated that $12 is saved for every $1 invested in drug treatment[29].

Along with treatment during incarceration, we recommend that the Iowa Department of Public Health prioritize the Jail-Based Substance Abuse Treatment program, which provides services during and after incarceration. This program, along with other services such as employment assistance and overdose education upon release, provides substance abuse and mental health support for ex-offenders, mitigating substance and mental health issues and aiding their reintegration into society.

6.3 Education and Counseling Opportunities at Max Security Prisons

Because many inmates lack the knowledge, training, and skills to support a successful return to society, they are forced back into a life of crime, and they end up back behind bars [4]. To reduce the likelihood that inmates leave prison without marketable skills, we must increase the availability of educational and counseling staff.

According to Feature Importance Analysis (4.5), the ratio of treatment staff to prisoners is inversely proportional to the probability of re-offending. The ratio of treatment staff to prisoners has a negative SHAP value, which means that an effective rehabilitation staff reduces the likelihood of recidivism. The implementation of these programs has also been found to be highly cost-effective, as a 2014 Rand Report on Correctional Education found that $4.55 - $5.26 are saved in incarceration costs for every $1 investment in prison education programs[43].

While programs exist in low-security prisons, our analysis in 5.5.1 revealed a lack of adequate rehabilitation resources in maximum-security prisons in Iowa. To address this issue, we recommend that a significant portion of the state budget be allocated to improving rehabilitation resources inside maximum-security prisons. While we understand that these facilities must prioritize correctional staff, we want to note that the case of the Iowa Medical and Classification Center proves a balance can be achieved. Despite it being a medium-security facility, it yielded lower expected loss values than some minimum-security prisons due to its ability to have the highest treatment staff-to-inmate ratio.

Within the maximum-security prisons (e.g., Anamosa State Penitentiary and Iowa State Penitentiary), the budget should be allocated to literacy classes, English as a Second Language (ESL), parenting classes, wellness education, adult-continuing education, library services, and instruction in leisure-time activities according to the standards set forth by the Federal Bureau of Prisons[14].

These resources will help prepare inmates for the GED and HiSET—two high school equivalency tests. The latter should be administered yearly at each prison because it can be taken on paper or online. Implementing the HiSET curriculum will be cost-effective because the organization offers free sample problems, online math tutorials, and a prep book.

To improve counseling services in maximum-security prisons, we look to the standards set by the Federal Bureau of Prisons[15]. This organization offers mental health programs for inmates through its Psychology Services Branch and Health Services Division.

6.4 Post-Release Programming with a Special Focus on At-Risk Inmates

Substance dependence increases the probability of recidivism by reducing the likelihood of employment, adding financial needs, and augmenting the possibility of parole violation detection [31]. Indeed, our model found this offense type to be an important factor in determining the probability of recidivism for a given incarcerated individual in Section 4.5. It discerned that conviction of a drug offense increased an inmate's likelihood of re-offending. Our analysis in 5.5.2 found "Drug Trafficking" offenses to yield the 4th highest total expected loss and the 5th highest expected loss per inmate.

In addition to substance abuse rehabilitation (6.2), we urge that the Iowa Department of Corrections implement community re-entry programs for incarcerated individuals convicted of drug-related crimes because punishment alone does not address criminal behavior that is related to drug use [8].

The Federal Bureau of Prisons contracts with Residential Reentry Centers (RRC), which provide inmates with housing, employment, and treatment resources. The directory only lists one Iowa RRC contracted by the Federal Bureau of Prisons, and it is at the western border in Sioux Falls. Centers should be put under contract in other parts of the state because drug-addicted individuals need a safe, structured, and supervised environment.

In addition, we recommend that Iowa State Prisons connect drug-related offenders post-release with the Center for Health and Justice at the TASC, an organization that provides individuals afflicted by substance abuse with health recovery management services to reduce recidivism.

Another at-risk group comprises individuals convicted of property crimes like theft and burglary (see Section 4.5). The model also found that these individuals yield the 4th highest estimated loss per inmate and the 5th greatest total estimated loss.

A conviction of theft or burglary can indicate a lack of economic stability, as poverty is a potential motive for stealing. To address this issue, work and vocational programs that target these offenders should be implemented in Iowa State Prisons.

A recent brief prepared by the Vera Institute of Justice delineated five successful work programs targeting incarcerated individuals. We advise that the State of Iowa keep these notes in mind when designing these types of programs [54]:

1. Create strong partnerships between DOC and education providers alongside other relevant stakeholders. Labor unions, community colleges, and nonprofit organizations all played major roles in these endeavors.

2. These training programs should develop pathways for students to develop degree certification.

3. Connections with employers are vital to decreasing barriers to reentry.

Acknowledgments

The authors of this work would like to thank our coach and mentor, Dr. Steve Gadbois, for his unwavering support throughout this process. His insights and guidance throughout this project have been invaluable.

Additionally, we would like to thank the Actuarial Foundation and the Institute of Competition Sciences for organizing the Modeling the Future Challenge and giving us the resources and opportunity to create this work.

References

[1] Anzer, G.; Bauer, P., 2021 *A Goal Scoring Probability Model for Shots Based on Synchronized Positional and Event Data in Football (Soccer)*, URL. https://www.frontiersin.org/articles/10.3389/fspor.2021.624475/full

[2] Bales, et. al; 2016 *An Assessment of the Effectiveness of Prison Work Release Programs on Post-Release Recidivism and Employment*, URL. https://www.ojp.gov/pdffiles1/nij/grants/249845.pdf

[3] Belenko, S; et al. 2014 *Treating Substance Use Disorders in the Criminal Justice System*, URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3859122/

[4] Bender, K. 2018 *Education Opportunities in Prison Are Key to Reducing Crime*, URL. https://www.americanprogress.org/article/education-opportunities-prison-key-reducing-crime/

[5] The Harvard University Institute of Politics Criminal Justice Policy Group. 2019 *SUCCESSFUL REENTRY: A COMMUNITY-LEVEL ANALYSIS*, URL. https://iop.harvard.edu/sites/default/files/sources/program/IOP_Policy_Program_2019_Reentry_Policy.pdf

[6] Boshart, R. 2014 *Report: Iowa prison population set to grow 39 percent by 2024*, URL. https://www.thegazette.com/crime-courts/report-iowa-prison-population-set-to-grow-39-percent-by-2024/

[7] Carson, E. 2015 *Prisoners in 2014*, URL. https://bjs.ojp.gov/content/pub/pdf/p14.pdf

[8] Chandler, Fletcher, and Volkow, 2009 *Treating Drug Abuse and Addiction in the Criminal Justice System: Improving Public Health and Safety*, URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2681083/

[9] Chandler, R. et al. 2010 *Treating Drug Abuse and Addiction in the Criminal Justice System: Improving Public Health and Safety* , URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2681083/

[10] Curtis-Ham, S.; Bernasco, W.; Medvedev, O.N.; Polaschek, D.L.L., 2021 *A National Examination of the Spatial Extent and Similarity of Offenders' Activity Spaces Using Police Data*, URL. https://www.mdpi.com/2220-9964/10/2/47

[11] Eisenberg, M; 1985 *Factors Associated With Recidivism*, URL. https://www.ojp.gov/ncjrs/virtual-library/abstracts/factors-associated-recidivism

[12] Farrington, D.P.; Nutall, C.P. 1980 *Prison Size, Overcrowding, Prison Violence, and Recidivism*, URL. https://www.ojp.gov/ncjrs/virtual-library/abstracts/prison-size-overcrowding-prison-violence-and-recidivism

[13] Federal Bureau of Prisons. 2023 *About Our Facilities*, URL. https://www.bop.gov/about/facilities/federal_prisons.jsp

[14] Federal Bureau of Prisons *Education Programs*, URL. https://www.bop.gov/inmates/custody_and_care/education.jsp

[15] Federal Bureau of Prisons *Inmate Mental Health*, URL. https://www.bop.gov/inmates/custody_and_care/education.jsp

[16] Federal Bureau of Prisons, *Inmate Offenses*, URL.https://www.bop.gov/about/statistics/statistics_inmate_offenses.jsp

[17] Federal Bureau of Prisons. *Sentences Imposed*, URL. https://www.bop.gov/about/statistics/statistics_inmate_sentences.jsp

[18] Hajian-Tilaki, K., 2013 *Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation*, URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3755824/

[19] Iowa Data, 2023 *Current Iowa Correctional System Prison Population*, URL. https://data.iowa.gov/Correctional-System/Current-Iowa-Correctional-System-Prison-Population/xbcv-c6t2

[20] Iowa Data, 2023 *Offenders Released from Iowa Prisons* , URL. https://data.iowa.gov/Correctional-System/Offenders-Released-from-Iowa-Prisons/runv-jsix

[21] Iowa Department of Corrections. 2017 *Average Annual Cost Per Inmate by Iowa Prison*, URL. https://www.legis.iowa.gov/docs/publications/FCTA/861052.pdf

[22] Iowa Department of Corrections. 2021 *Institutional Core Programming*, URL. https://doc.iowa.gov/sites/default/files/is-cl-03_attachment_1_idoc_core_programming.pdf

[23] Iowa Department of Corrections, 2023 *About Institutions*, URL. https://doc.iowa.gov/about-us/about-institutions

[24] Iowa Department of Corrections, 2023 *Iowa Unemployment Insurance Benefit Payments and Claims (Annual))*, URL. https://data.iowa.gov/Economic-Statistics/Iowa-Unemployment-Insurance-Benefit-Payments-and-C/rmcb-sifx

[25] Iowa Department of Corrections, 2023 *Iowa Unemployment Insurance Benefit Payments and Recipients by County (Monthly)*, URL. https://data.iowa.gov/Economic-Statistics/Iowa-Unemployment-Insurance-Benefit-Payments-and-R/aeyn-twxp

[26] Iowa Department of Corrections, *2023 Annual Personal Income for State of Iowa by County*, URL. https://data.iowa.gov/Economic-Statistics/Annual-Personal-Income-for-State-of-Iowa-by-County/st2k-2ti2

[27] Iowa Department of Corrections, *2023 Annual Personal Income for State of Iowa, URL.* https://data.iowa.gov/Economic-Statistics/Annual-Personal-Income-for-State-of-Iowa/dxzz-fkf8

[28] Iowa Department of Corrections. *Prison Recidivism FY2022*, URL. https://doc.iowa.gov/data/prison-recidivism-fy2022

[29] Kubic, M.; Pendergrass, T. 2017 *Diversion Programs Are Cheaper and More Effective Than Incarceration. Prosecutors Should Embrace Them.*, URL. https://www.aclu.org/news/smart-justice/diversion-programs-are-cheaper-and-more-effective-incarceration-prosecutors

[30] Kempf-Leonard, K. 2014 *Offense Specialization/Expertise*, URL. https://www.oxfordbibliographies.com/display/document/obo-9780195396607/obo-9780195396607-0104.xml

[31] Kyckelhahn, T.; Cooper, T. 2017 *The Past Predicts the Future: Criminal History and Recidivism of Federal Offenders*, URL. https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2017/20170309_Recidivism-CH.pdf

[32] Liu, Marco. *Optimal Number of Trials for Monte Carlo Simulation*, URL. https://www.valuationresearch.com/wp-content/uploads/2019/11/SpecialReport_MonteCarloSimulationTrials-11-19.pdf

[33] Lundberg, S.; Lee, S. 2017 *A Unified Approach to Interpreting Model Predictions*, URL. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

[34] Mears, D. 2006 *Evaluating the Effectiveness of Supermax Prisons* , URL. https://www.ojp.gov/pdffiles1/nij/grants/211971.pdf

[35] McEniry, J. 2023 *Criminal Law Overview*, URL. https://www.legis.iowa.gov/docs/publications/LG/14969.pdf

[36] McCollister KE, French MT, Fang H. , 2010 *The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation*, URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2835847/

[37] Miller, et al., 2006 *Costs of alcohol and drug-involved crime*, URL. https://pubmed.ncbi.nlm.nih.gov/16845591/

[38] Miller, et. al.,1996 *Victim Costs and Consequences: A New Look*, URL. https://www.ojp.gov/pdffiles/victcost.pdf

[39] National Institute of Justice. 2011 *Impact of Prison Experience on Recidivism*, URL. https://nij.ojp.gov/topics/articles/impact-prison-experience-recidivism

[40] Penal Reform International, *Overcrowding*, URL. https://www.penalreform.org/issues/prison-conditions/key-facts/overcrowding/

[41] Prison Policy Initiative. *Iowa Portfolio*, URL. https://www.prisonpolicy.org/profiles/IA.html

[42] Przybylski, R. 2015 *Adult Sex Offender Recidivism.*, URL. https://smart.ojp.gov/somapi/chapter-5-adult-sex-offender-recidivism

[43] RAND Research Group, 2014 *How Effective Is Correctional Education, and Where Do We Go from Here?* , URL. https://www.rand.org/pubs/research_reports/RR564.html

[44] Rosario, et al. 2021 *What to know about the Iowa prison attack and the Anamosa State Penitentiary victims*, URL. https://www.desmoinesregister.com/story/news/crime-and-courts/2021/03/23/anamosa-state-penitentiary-attack-history-what-know-iowa-prison/6969432002/

[45] Rajkumar, A.; French, M. 1997 *Drug Abuse, Crime Costs, and the Economic Benefits of Treatment*, URL. https://www.jstor.org/stable/23365794

[46] Schnepel, K. 2018 *Good Jobs and Recidivism Economic Journal*, URL. https://www.ojp.gov/ncjrs/virtual-library/abstracts/good-jobs-and-recidivism-economic-journal

[47] Stahler, et. al; 2013 *Predicting Recidivism for Released State Offenders*, URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3891510/

[48] Stemen, D. 2017 *The Prison Paradox: More Incarceration Will Not Make Us Safer*, URL. https://www.vera.org/downloads/publications/for-the-record-prison-paradox_02.pdf

[49] The Keras Team; *Keras*, URL. https://github.com/keras-team/keras

[50] Thiese, MS; Ronna, B; Ott U. P; 2016 *P value interpretations and considerations*, URL. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5059270/

[51] Tilly, Nick, 2016 *Analyzing and Responding to Repeat Offending*, URL. https://popcenter.asu.edu/content/analyzing-and-responding-repeat-offending

[52] US Department of Justice. 2022 *Mental Health and Reentry: How Court Services Offender Agency Meets the Challenge of Mental Health Community Supervision*, URL. https://cops.usdoj.gov/html/dispatch/05-2022/mental_health_reentry.html

[53] U.S. Department of Justice. (July 2021). *Recidivism of Prisoners Released in 34 States in 2012: A 5-Year Follow-Up Period (2012-2017)*. URL . https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/rpr34s125yfup1217.pdf

[54] Vera Institute of Justice *Work Programs*, URL. https://www.vera.org/downloads/publications/how-corrections-departments-are-preparing-people-for-in-demand-careers.pdf

[55] Wood, T.; *Backpropagation*, URL. https://deepai.org/machine-learning-glossary-and-terms/backpropagation

[56] Zarkin, G. et al. 2012 *Lifetime Benefits and Costs of Diverting Substance-Abusing Offenders From State Prison*, URL. https://journals.sagepub.com/doi/abs/10.1177/0011128712461904

Appendix

Appendix 1: Data Cleaning

```
import pandas as pd
df=pd.read_csv ("/Users/amark/Desktop/Py4e/Puzzles/Prison (1).csv",
dtype={"Jurisdiction":"string"},parse_dates= ["Admission Date"])
df.dropna (inplace=True,axis=0,subset= ["Admission Date","Jurisdiction"])
df.dropna (inplace=True,axis=1,how="all")
start_date = "2001-01-01"
end_date ="2022-01-01"
mask = (df ["Admission Date"] >= start_date) & (df ["Admission Date"] < end_date)
df = df [mask]
df_insurance=pd.read_csv ("/Users/amark/Desktop/Py4e/Puzzles/
Iowa_Unemployment_Insurance_Benefit_Payments_and_Recipients_by_County__Monthly_.csv",
dtype={"Jurisdiction":'string'},parse_dates= ["Admission Date"])
df_insurance.dropna (inplace=True,axis=1,how="all")
df_insurance.dropna (inplace=True,axis=0,subset= ["Admission Date","Jurisdiction"])
mask = (df_insurance ["Admission Date"] >= start_date) & (df_insurance ["Admission
Date"] < end_date)
df_insurance = df_insurance [mask]
df2=pd.merge (df,df_insurance,how="left",
left_on= ["Admission Date","Jurisdiction"],right_on= ["Admission Date","Jurisdiction"])
df2.insert (len (df2.columns), "Admission Year", df2 ["Admission Date"], True)
df2 ["Admission Year"]=df2 ["Admission Year"].astype ("string")
df2 ["Admission Year"]=df2 ["Admission Year"].str.split ("-").str.get (0)
df2 ["Admission Year"]=df2 ["Admission Year"].astype ("int64")
df1=pd.read_csv ("/Users/amark/Desktop/Py4e/Puzzles
/Annual_Personal_Income_for_State_of_Iowa_by_County.csv",dtype={"Jurisdiction":
"string","Admission Date":"int64","MedianFamilyIncome":"int64"})
df1.dropna (inplace=True,axis=1,how="all")
```

```
dfAdditionToIncome=pd.read_csv ("/Users/amark/Desktop/Py4e/Puzzles
/Annual_Personal_Income_for_State_of_Iowa_by_County__Recent_Year.csv")
dfAdditionToIncome.dropna (inplace=True,axis=1,how="all")
dfAdditionToIncome.dropna (inplace=True,axis=0,how="all")
dfAdditionToIncome ["Jurisdiction"]=dfAdditionToIncome ["Jurisdiction"].
str.removesuffix (", IA")
df1=df1.append (dfAdditionToIncome,ignore_index=True)
df1 ["Admission Date"]=df1 ["Admission Date"].astype ("int64")
df1 ["MedianFamilyIncome"]=df1 ["MedianFamilyIncome"].astype ("int64")
df3=pd.merge (df2,df1,how="left",left_on= ["Admission Year","Jurisdiction"],right_on=
["Admission Date","Jurisdiction"],validate="many_to_one")
dfPopulation=pd.read_csv ("/Users/amark/Desktop/Py4e/Puzzles/
County_Population_in_Iowa_by_Year.csv",dtype={"Admission Date":"int64",
"Jurisdiction":"string"})
start_date = 2001
end_date =2022
mask = (dfPopulation ["Admission Date"] >= start_date) & (dfPopulation ["Admission
Date"] < end_date)
dfPopulation = dfPopulation [mask]
dfFinal=pd.merge (df3,dfPopulation,how="left",left_on= ["Admission Year","Jurisdiction"]
,right_on= ["Admission Date","Jurisdiction"])
dfFinal.dropna (axis=0,subset= ["MedianFamilyIncome","Population","Benefits Paid"],
inplace=True)
dfFinal.to_csv ("/Users/amark/Desktop/Py4e/Puzzles/Combined_Prison_Final.csv")
```

Appendix 2: Feedforward Neural Network

```
#FeedForwardNeural Net
import pandas as pd
import numpy as np
import tensorflow as tf
from tensorflow import feature_column
from tensorflow.keras import layers
from sklearn.model_selection import train_test_split
import tensorflow_datasets as tfds
import shap
import tensorflow as tf
from sklearn.metrics import roc_auc_score
import matplotlib.pyplot as plt
import time
df = pd.read_csv ("/Users/lzhou/FINALPrisonDatasetRelease.csv")

df = df [ ['PreviousNumberofRecidivize','Record ID','Offender Number','Release
Date',
'Fiscal Year Released','Closure Type','Supervision Status','Institution Name ','Sex',
```

```
'Race & Ethnicity','Age at Release','Offense Code','Offense Classification','Offense
Description',
'Offense Type','Offense Subtype','Jurisdiction','Admission Date_x','Months Served',
'Population Prison','Capacity','Correctional Staff','Treatment','Type','Cost of
Food_x','DidRecidivize','MedianFamilyIncome','Benefits Paid Per Capita']]
df = df.rename (columns={"Offense Description": "Offense_Description","Closure
Type":"Closure_Type","Age at Release":"Age","Cost of Food_x":"Food",
"Race & Ethnicity":"Race",
"Population Prison":"Population"})
df.replace ( [np.inf, -np.inf], np.nan, inplace=True)
df.dropna (inplace=True)
df ['Overpopulation'] = df ['Capacity'] - df ['Population']
df ['CorrectionaltoPrisoner'] = df ['Population'] / df ['Correctional Staff']
df ['TreatmenttoPrisoner'] = df ['Population'] / df ['Treatment']
df_interest = df [ ['PreviousNumberofRecidivize','DidRecidivize','Age','Food',
'Overpopulation','CorrectionaltoPrisoner',
'TreatmenttoPrisoner','Sex', 'Race','Offense_Description',''MedianFamilyIncome',
'Benefits Paid Per Capita']]
#Getting Variables of Interest


sexvariabledummy = pd.get_dummies (df_interest ['Sex'])
df_interest = pd.concat ( [df_interest, sexvariabledummy], axis=1)
racedummy = pd.get_dummies (df_interest ['Race'])
df_interest = pd.concat ( [df_interest, racedummy], axis=1)
offensedescriptiondummy = pd.get_dummies (df_interest ['Offense_Description'])
df_interest = pd.concat ( [df_interest, offensedescriptiondummy], axis=1)
df_interest = df_interest.drop ( ['Sex', 'Race','Offense_Description'], axis=1)
#One-hot encoding

train, test = train_test_split (df_interest, test_size=0.2)
x_train = train.drop ( ['DidRecidivize'], axis=1)
y_train = train ['DidRecidivize']
x_test = test.drop ( ['DidRecidivize'], axis=1)
y_test = test ['DidRecidivize']
#Splitting into test and train datasets


batch_size = 128
model = tf.keras.Sequential ( [
  layers.Dense (128, activation='sigmoid'),
  layers.Dense (128, activation='sigmoid'),
  layers.Dense (128, activation='sigmoid'),
  layers.Dense (128, activation='sigmoid'),
  layers.Dropout (.2),
```

```
  layers.Dense (1, activation ='sigmoid')
])

model.compile (optimizer='adam',
              loss=tf.keras.losses.BinaryCrossentropy (from_logits=True),
              metrics= [tf.keras.metrics.BinaryCrossentropy ()])
model.fit (x_train, y_train,
          epochs=50)
#Building and Fitting the Model
prediction = model.predict (x_test)
print (prediction.flatten ())
target = y_test
print (roc_auc_score (target.to_numpy (),prediction.flatten ()))
print (model.summary ())
#Evaluating the Model


shap.initjs ()
explainer = shap.explainers.Permutation (model.predict, x_test)
shap_values = explainer (x_test [:100])
shap.summary_plot (shap_values, plot_type = 'dot', feature_names = x_train.columns,
max_display = 20, plot_size = [13,5])

model.save ('FNN')
#Getting SHAP Values and Saving the Model
```

Appendix 3: Monte Carlo Simulations

```
import pandas as pd
import numpy as np
import time
import matplotlib.pyplot as plt
import seaborn as sns
import scipy
from scipy.stats import burr, foldcauchy, rdist, nct, fisk, skewcauchy
import time
import random
from datetime import datetime
from tensorflow import keras



changelst = [-4,-3,-2,-1,0,1,2,3,4]
simpleseriousweights = [0,0,0,0,0.509933775,0.218543046,0.205298013,0.052980132,
0.013245033]
```

```
aggravatedweights = [0,0,0.000167983,0.003359651,0.656139761,0.183604905,
0.149840417,0.006719301,0.000167983]
dclassweights = [0,.0000661,0.001387329,0.047697694,0.907775649,
0.033692277,0.009380987,.0000661,0]
cclassweights = [0, 0.000994926, 0.038304646, 0.057307731, 0.890657646, 0.012735051,
0, 0, 0]
bclassweights = [0.001258653, 0.006922593, 0.021397105,
0.009439899, 0.96098175, 0, 0, 0, 0]
aclassweights = [0.076923077, 0, 0, 0.076923077, 0.846153846, 0, 0, 0, 0]
costofcrimes = {
0: 275.85,
1: 194238.193,
2: 112363.641,
3: 148201.244,
4: 232772.637,
5: 768295.654,
6: 7755510.88,
}#declaring constants
present = datetime.now ().date ()
model = keras.models.load_model ('/Users/lzhou/FNN')
df = pd.read_csv ("/Users/lzhou/MonteCarloFinal.csv")
UniqueNames = df ['Offense Subtype'].unique ()
ListNameDict = {elem : [] for elem in UniqueNames}

UniqueNamesPrisons = df ['Prison'].unique ()
PrisonNameDict = {elem : [] for elem in UniqueNamesPrisons}

dfkeep = pd.read_csv ("/Users/lzhou/MonteCarloFinal.csv")
df_interest = df [ ['MedianFamilyIncome','Benefits Paid Per Capita','Sex', 'Race
& Ethnicity','Offense Description',"PreviousNumberofRecidivize",'Cost of Food',
'Overpopulation','CorrectionaltoPrisoner','TreatmenttoPrisoner']]
df_interest.rename (columns={'Cost of Food': "Food", "Race & Ethnicity": "Race",
"Offense Description":"Offense_Description"})
sexvariabledummy = pd.get_dummies (df_interest ['Sex'])
df_interest = pd.concat ( [df_interest, sexvariabledummy], axis=1)
racedummy = pd.get_dummies (df_interest ['Race & Ethnicity'])
df_interest = pd.concat ( [df_interest, racedummy], axis=1)
offensedescriptiondummy = pd.get_dummies (df_interest ['Offense Description'])
df_interest = pd.concat ( [df_interest, offensedescriptiondummy], axis=1)
df_interest = df_interest.drop ( ['Cost of Food','Sex', 'Race & Ethnicity','Offense
Description'], axis=1)
#getting dummies and cleaning data
df1 = pd.DataFrame (df [df ['Offense Classification'] == 1])
df2 = pd.DataFrame (df [df ['Offense Classification'] == 2])
df3 = pd.DataFrame (df [df ['Offense Classification'] == 3])
```

```
df4 = pd.DataFrame (df [df ['Offense Classification'] == 4])
df5 = pd.DataFrame (df [df ['Offense Classification'] == 5])
df6 = pd.DataFrame (df [df ['Offense Classification'] == 6])
#separation by offense classification for computational efficiency
for i in range (2603):
    start_time = time.time ()

    df1 ["changeinclass"] = random.choices (changelst, weights=simpleseriousweights,
    k=len (df1 ["Offense Classification"]))
    df2 ["changeinclass"] = random.choices (changelst, weights=aggravatedweights,
    k=len (df2 ["Offense Classification"]))
    df3 ["changeinclass"] = random.choices (changelst, weights=dclassweights,
    k=len (df3 ["Offense Classification"]))
    df4 ["changeinclass"] = random.choices (changelst, weights=cclassweights,
    k=len (df4 ["Offense Classification"]))
    df5 ["changeinclass"] = random.choices (changelst, weights=bclassweights,
    k=len (df5 ["Offense Classification"]))
    df6 ["changeinclass"] = random.choices (changelst, weights=aclassweights,
    k=len (df6 ["Offense Classification"])) #modeling change in class

    df1 ['Estimated Time'] = rdist.rvs (6.56980821775444, 188.67155647951017,
    271.39501665147156, size=len (df1 ["changeinclass"]))
    df2 ['Estimated Time'] = burr.rvs (3.4390415311332028, 0.5539524934124328,
    -13.352161944958677, 273.475702118251,
    size=len (df2 ["changeinclass"]))
    df3 ['Estimated Time'] = fisk.rvs (3.1250724774007135, -94.19921007350189,
    405.11714597252353, size=len (df3 ["changeinclass"]))
    df4 ['Estimated Time'] = nct.rvs (2.220622981521773, 5.342343243604949,
    -230.23142805524765,109.70853045404898, size=len (df4 ["changeinclass"]))
    df5 ['Estimated Time'] = foldcauchy.rvs (0.4065447443936276, 9.999999997470635,
    585.6671047093056, size=len (df5 ["changeinclass"]))
    df6 ['Estimated Time'] = burr.rvs (10.434553629473438, 0.08310923509449561,
    -94.87644125914952,15619.217751236414, size=len (df6 ["changeinclass"]))
    df = pd.concat ( [df1,df2,df3,df4,df5,df6]) #modeling estimated time served
    df ['newclassification'] = df ['changeinclass'] + df ['Offense Classification']
    df ['CostofCrime'] = df ['newclassification'].map (costofcrimes)
    df ['Prison Start Date']= pd.to_datetime (df ['Prison Start Date']).dt.date
    df ["Estimated Release_DATE"] = df ['Prison Start Date'] +
    pd.to_timedelta (df ['Estimated Time'], unit='D')
    dfchange = pd.DataFrame (df [df ['Estimated Release_DATE'] <= present])
    dfnochange = pd.DataFrame (df [df ['Estimated Release_DATE'] > present]) #prep
for prediction
    dfchange ['Year of Estimated Release'] = 2023
    dfnochange ['Year of Estimated Release'] =
    dfnochange ['Estimated Release_DATE'].astype (str).str [:4]
```

```
    df = pd.concat ( [dfchange, dfnochange])
    df ['TimeElapsed'] =
    df ['Year of Estimated Release'].astype (int) - 2023
    df ['EstimatedReleaseAge'] = df ['TimeElapsed'] + df ['Age']
    df_interest ['MedianFamilyIncome'] = 1355.6136 * df ['Year of Estimated Release']
+ 18030
    df_interest ['Benefits Paid Per Capita'] = skewcauchy.rvs (-0.04439145386124918,
127.08134552493124,12.915796802226604, size=len (df ["changeinclass"]))
    df_interest = pd.concat ( [df_interest, df ['EstimatedReleaseAge']], axis=1)
    df_interest = df_interest.rename (columns={"EstimatedReleaseAge": "Age"})
    df ['Model Probability Prediction'] = model.predict (df_interest)#make prediction
    df ['Expected Cost'] = df ['Model Probability Prediction'] * df ['CostofCrime']
#find expected cost for each prisoner
    DataFrameDict = {elem : pd.DataFrame () for elem in UniqueNames}
    for key in DataFrameDict.keys ():
        DataFrameDict [key] = df [:] [df ['Offense Subtype'] == key]
    for names in UniqueNames:
        ListNameDict [names].append (DataFrameDict [names] ["Expected Cost"].sum
())
        uniquesums = uniquesums + DataFrameDict [names] ["Expected Cost"].sum
()

    PrisonFrameDict = {elem : pd.DataFrame () for elem in UniqueNamesPrisons}
    for key in PrisonFrameDict.keys ():
        PrisonFrameDict [key] = df [:] [df ['Prison'] == key]
    for names in UniqueNamesPrisons:
        PrisonNameDict [names].append (PrisonFrameDict [names] ["Expected Cost"].sum
())
        uniquesums = uniquesums + PrisonFrameDict [names] ["Expected Cost"].sum
()  #saving expected cost for offense subtype and prisons

    print ("--- %s seconds ---" % (time.time () - start_time))
    print (i)

crimenames = []
crimemeans = []
crimeamounts = []
crimestanddevs = []
crimepercriminal = []
dfcrime = pd.DataFrame ()
for names in UniqueNames:
    series = pd.Series (ListNameDict [names])
    mean = series.mean ()
    standdev = series.standdev ()
    numberofcriminals = len (DataFrameDict [names] ['Offense Classification'])
```

```
    costpercriminal = mean / numberofcriminals
    crimenames.append (names)
    crimemeans.append (mean)
    crimeamounts.append (numberofcriminals)
    crimestanddevs.append (standded)
    crimepercriminal.append (costpercriminal)
dfcrime ["NameofCrime"] = crimenames
dfcrime ["AverageMCCost"] = crimemeans
dfcrime ['STANDDEV'] = crimestanddevs
dfcrime ["NumberofCriminals"] = crimeamounts
dfcrime ["CostperCriminal"] = crimepercriminal
prisonnames = []
prisonmeans = []
prisonamounts = []
prisonstanddevs = []
prisonpercriminal = []
dfprison = pd.DataFrame ()
for names in UniqueNamesPrisons:
    series = pd.Series (PrisonNameDict [names])
    mean = series.mean ()
    standdev = series.standdev ()
    numberofcriminals = len (PrisonFrameDict [names] ['Offense Classification'])
    costpercriminal = mean / numberofcriminals
    prisonnames.append (names)
    prisonmeans.append (mean)
    prisonamounts.append (numberofcriminals)
    prisonstanddevs.append (standdev)
    prisonpercriminal.append (costpercriminal)
dfprison ["NameofCrime"] = prisonnames
dfprison ["AverageMCCost"] = prisonmeans
dfprison ['STANDDEV'] = prisonstanddevs
dfprison ["NumberofCriminals"] = prisonamounts
dfprison ["CostperCriminal"] = prisonpercriminal
dfprison.to_csv ("PrisonResultsMonteCarlo.csv")
dfcrime.to_csv ("CrimeResultsMonteCarlo.csv")
#exporting to csv
```

Appendix 4: Exact Crime Sub-Type Proportions for each Classification

Crime Subtype	Simple/Serious Misdemeanor	Aggravated Misdemeanor	D Class	C Class	B Class	A Class
OWI	0.1221374	0.08416703	0.118407919	0	0	0
No-Cost Criminal	0	0.06807308	0.025120612	0.002273487	0.058551618	0.12903226
Traffic	0.01145038	0.05687255	0.018299784	0.000334336	0	0
Alcohol	0.01908397	0.04338843	0	0	0	0
No-Cost Public Order	0.02671756	0.02729448	0.038221594	0.0036777	0	0
Weapons	0.00381679	0.02446716	0.046581268	0.001604814	0	0
Animals	0	0.00086994	0	0	0	0
Flight/Escape	0.02671756	0.00065246	0.005531526	0	0	0
Drug Possession	0.14122137	0.0777512	0.107012144	0.00220662	0	0
No-Cost Drug	0.00381679	0.00065246	0.013849609	0.003543965	0	0
Prostitution/Pimping	0	0.00184863	0.000790218	0	0	0
Burglary	0	0.05284906	0.161661953	0.13293213	0	0
Trafficking	0	0.00315355	0.08164199	0.437178201	0.699922958	0.04301075
Theft	0.03435115	0.1499565	0.105930794	0.125309261	0	0
Vandalism	0.01908397	0.0184863	0.017093662	0.007689736	0	0
Forgery/Fraud	0	0.03251414	0.101355848	0.003209629	0	0
Stolen Property	0	0	0.000166362	0.000267469	0	0
Other Violent	0	0.02816442	0.003867909	0.034971581	0.080315871	0
Arson	0.00381679	0.00260983	0.001247713	0.019458375	0.011941448	0
Robbery	0	0.00293606	0	0.055031762	0.046032357	0
Kidnap	0	0.0030448	0.001829978	0.004680709	0.003659476	0.12365591
Sex	0.04961832	0.03207916	0.02291632	0.085055165	0.033127889	0.07526882
Assault	0.53435115	0.2877338	0.12048744	0.06633233	0.005200308	0
Murder/Manslaughter	0	0.00032623	0.00794377	0.014175861	0.060477658	0.62365591

Appendix 5: Exact Change of Crime Classification Proportions

Change	Simple / Serious Misdemeanor	Aggravated Misdemeanor	D Class Felony	C Class Felony	B Class Felony	A Class Felony
-4	0	0	0	0	0.001258653	0.076923077
-3	0	0	0	0.000994926	0.006922593	0
-2	0	0	0.001387329	0.038304646	0.021397105	0
-1	0	0.003359651	0.047697694	0.057307731	0.009439899	0.076923077
0	0.509933775	0.656139761	0.907775649	0.890657646	0.96098175	0.846153846
1	0.218543046	0.183604905	0.033692277	0.012735051	0	0
2	0.205298013	0.149840417	0.009380987	0	0	0
3	0.052980132	0.006719301	0.0000661	0	0	0
4	0.013245033	0.000167983	0	0	0	0