



PREDICTING RECIDIVISM

Group 4



AGENDA

01 PROBLEM STATEMENT

02 DATA OVERVIEW

03 EXPLORATORY DATA ANALYSIS

04 MODELING

05 CONCLUSIONS AND RECOMMENDATIONS

PROBLEM STATEMENT

PROBLEM STATEMENT

We are a team of policy researchers trying to determine if an algorithm can accurately predict whether an inmate will be reincarcerated. We will build and evaluate a series of different classification models to determine what features should be included. Types of features to consider will include demographic, behavioral and criminal history.

Recidivism is measured by criminal acts that resulted in rearrest, reconviction or return to prison with or without a new sentence.



DATA OVERVIEW



DATA OVERVIEW

BASIC

New York: 188,000 observations with 5 features

CRIMINAL HISTORY

Florida: 11,000 observations with 34 features

BEHAVIORAL

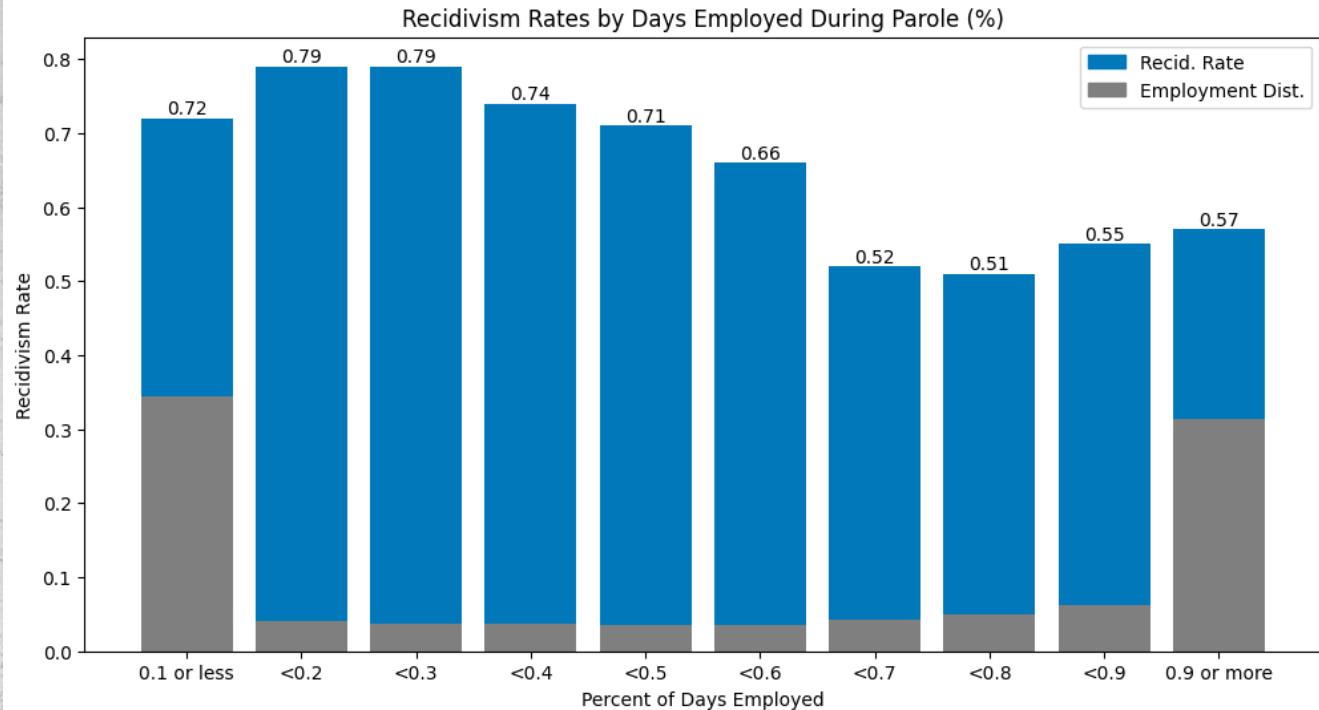
Georgia: 26,000 observations with 53 features

EDA

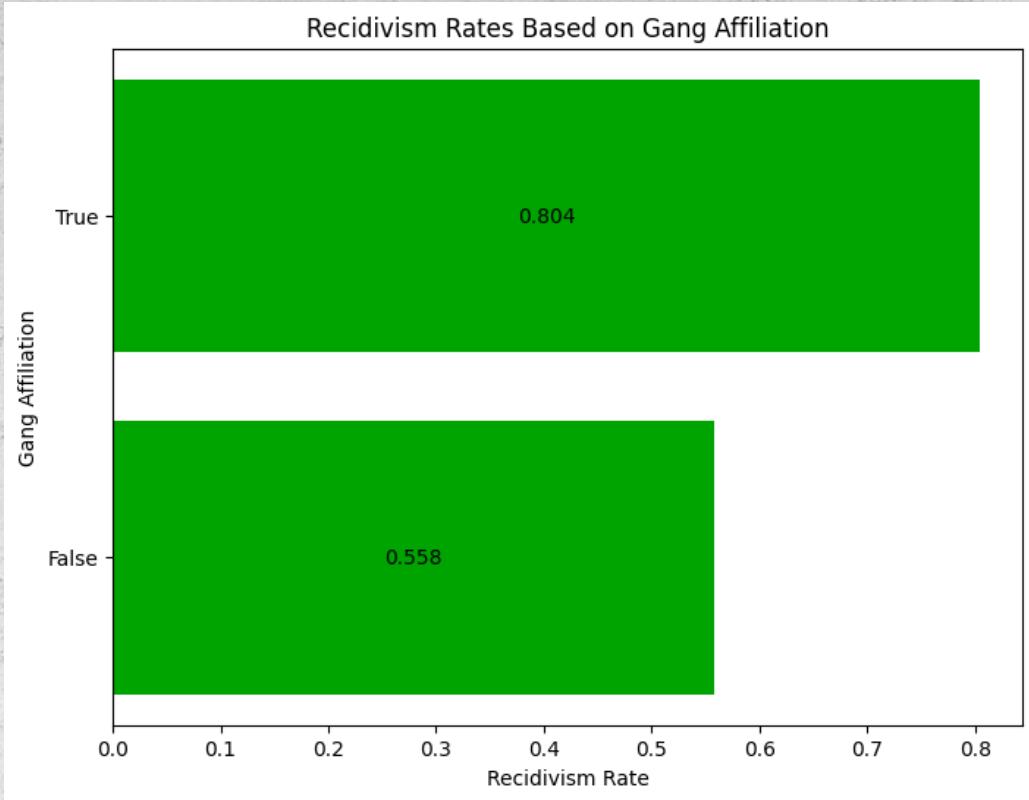
Percent of Days Employed

Percent days employed represents the percentage of days an individual was employed over the course of their parole.

Strong indication that increasing employment opportunities for those recently released from prison will decrease recidivism rates.



Gang Affiliation



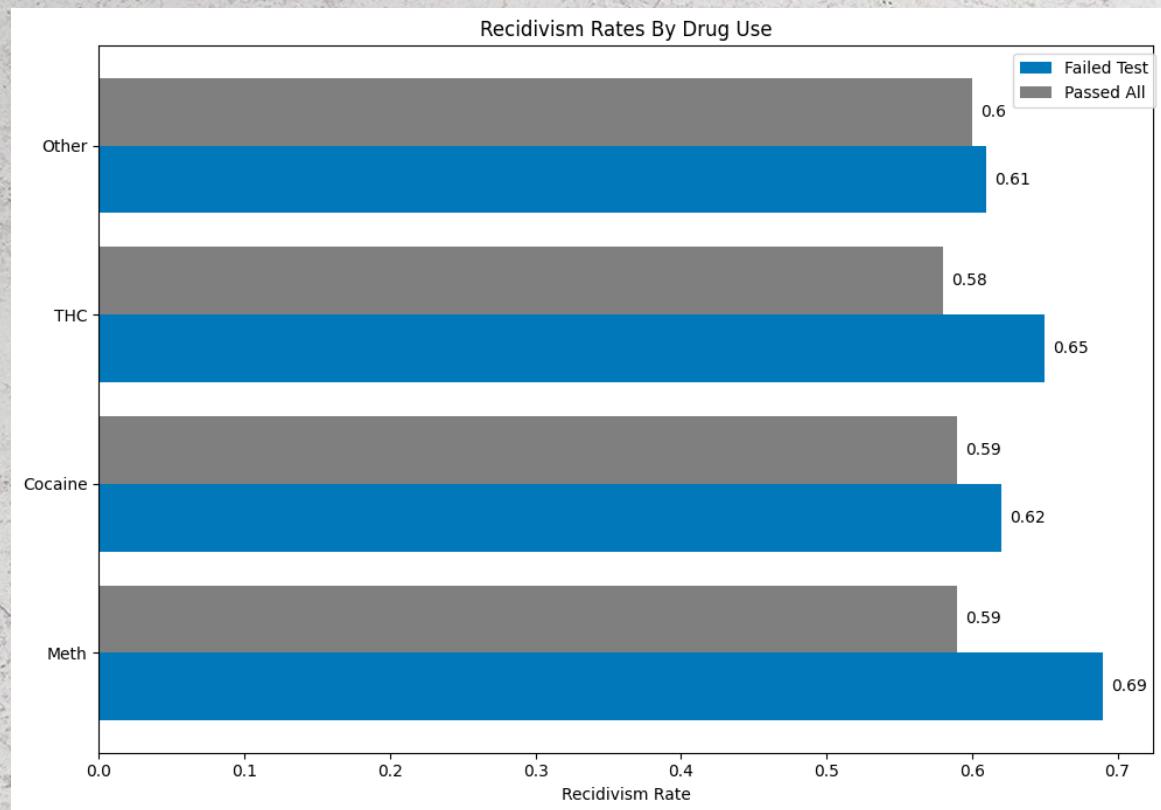
Those who are gang-affiliated have a 25% greater mean recidivism rate than those who are not.

During feature importance, we found that those who are gang-affiliated are twice as likely to reoffend compared to their non-affiliated counterparts.

Substance Abuse

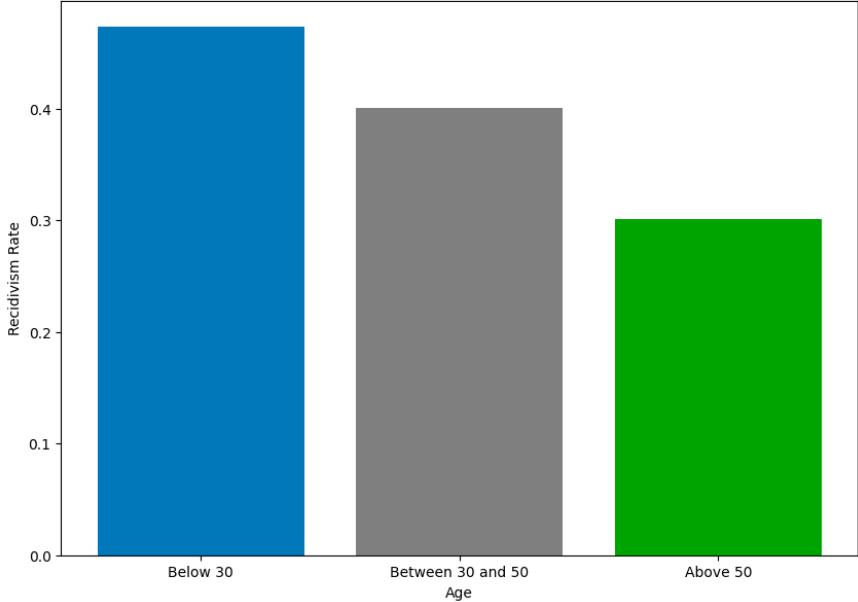
Those who tested positive at least once for THC, cocaine, or methamphetamine during their parole were more likely to recidivate than those who did not test positive at all.

We see the largest discrepancy among meth users and non-meth users.

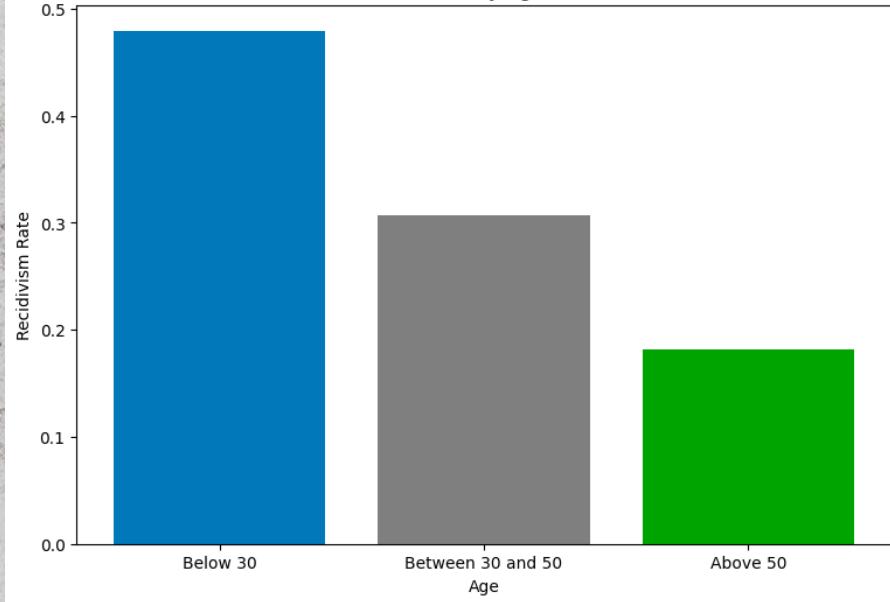


Age at Release

NY- Recidivism Rates by Age Group

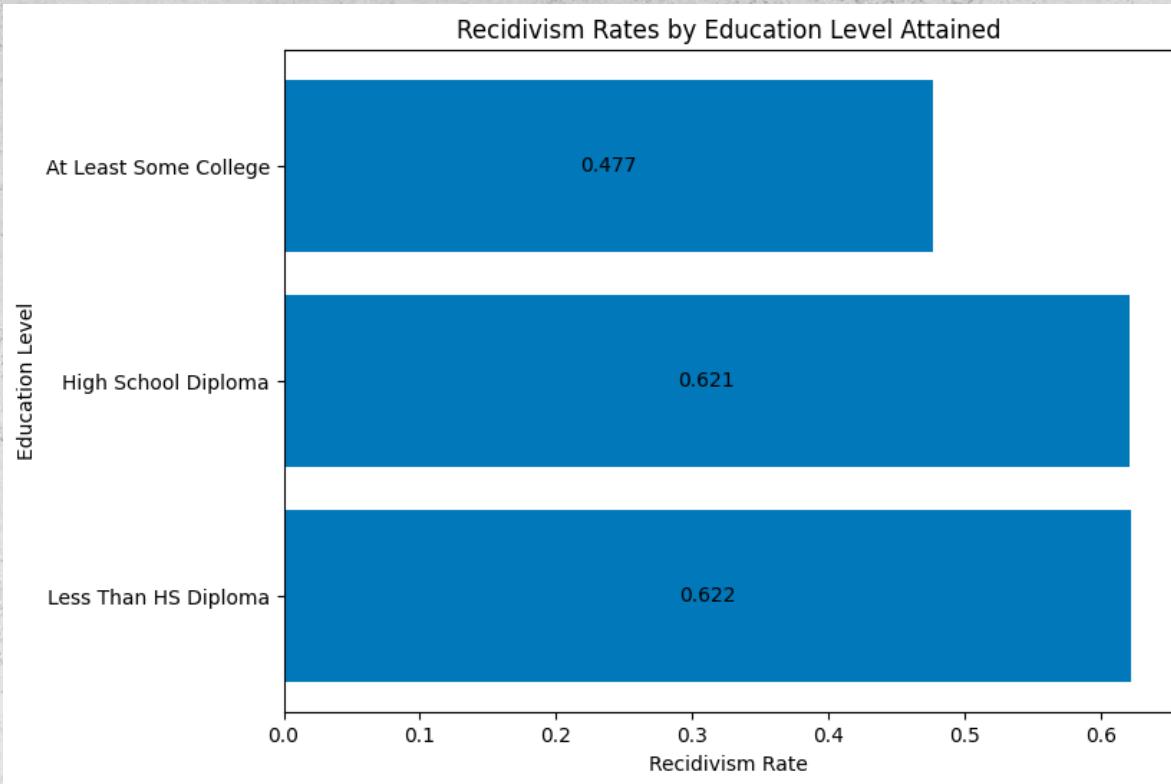


FL- Recidivism Rates by Age at First Release



There is a clear relationship between an individual's age at their initial prison release and their likelihood of reincarceration. In all three datasets, we found that **those who are younger were significantly more likely to be arrested again.**

Education Level



Those who did not graduate from high school and those who do not have any education beyond a high school diploma have similar average recidivism rates of about 62%.

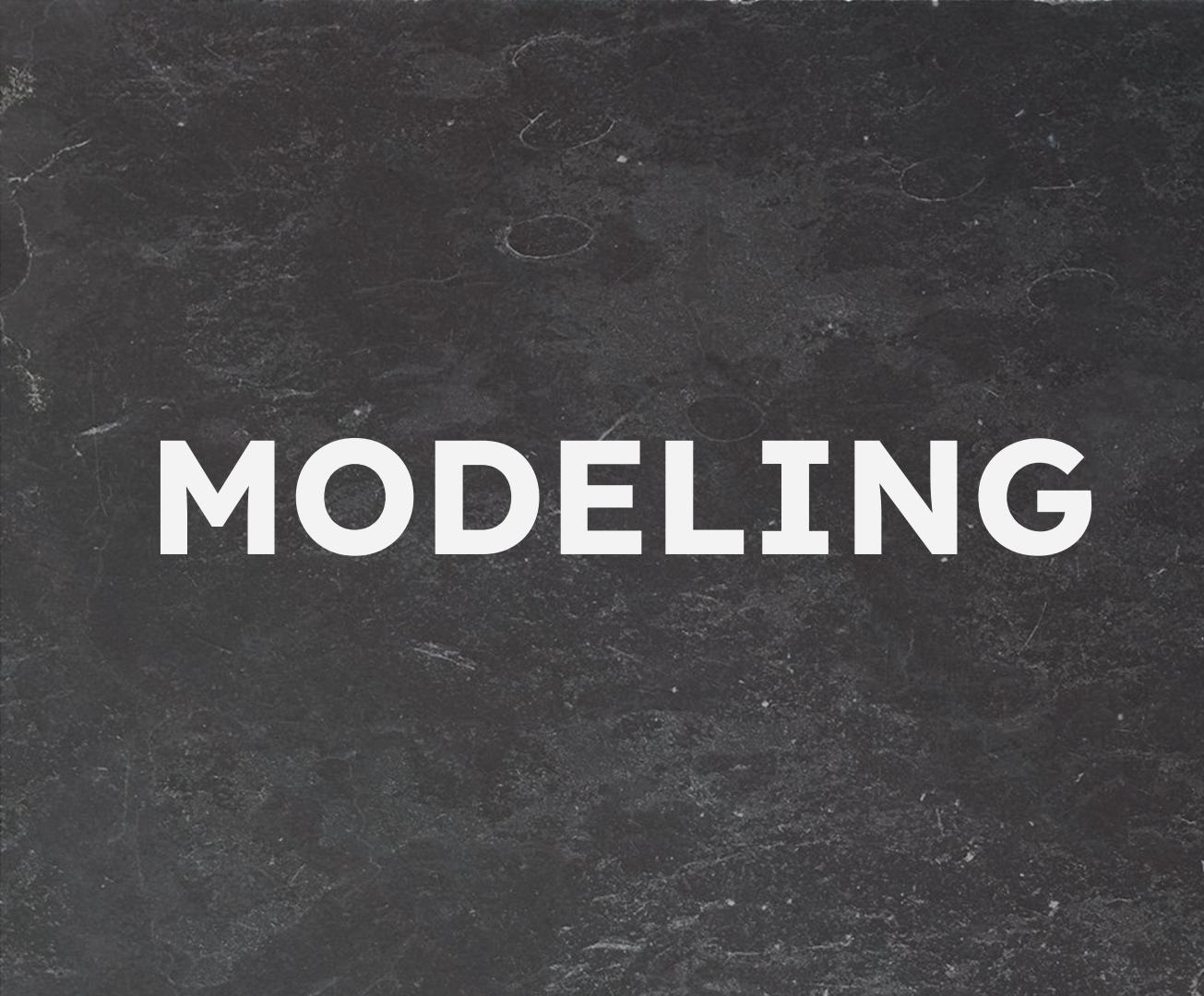
The lowest recidivism rate is seen in those that have at least some college experience.

Charge Type

We ran NLP on the Florida dataset and found these to be the most commonly-used bigrams in the charge descriptions.

As we might expect, the majority of charges indicated by these top 10 are for minor offenses such as 3rd-degree crimes, suspended driver's licenses and cannabis possession.

1. possession of
2. fail to
3. license suspended
4. grand theft
5. in the
6. theft in
7. 3rd degree
8. the 3rd
9. or less
10. possess cannabis



MODELING



Basic Dataset

Age at Release/Gender/County of Indictment

model_used	train_score	test_score	specificity
fnn gen/age/cty	0.599808	0.600451	0.884096
ada gen/age/cty	0.599569	0.600027	0.870779
stack gen/age/cty	0.600981	0.599867	0.876233
logreg gen/age/cty	0.598734	0.599019	0.876778
grad gen/age/cty	0.598940	0.598834	0.918276

Modeling with only age, gender, and county of indictment was unsuccessful.

Baseline accuracy: 0.58

Best performing model:
FNN with layers of 84, 64, and 32 neurons, 60% accuracy.

Behavioral Dataset

Comprehensive Features Model

model_used	train_score	test_score	specificity
gboost2	0.753582	0.717560	0.595335
gboost	0.739048	0.714695	0.567444
ada3	0.728454	0.713058	0.574544
ada	0.727175	0.711420	0.574037
rf3	0.745599	0.704871	0.480223

Significant jump in accuracy from the basic model, as various behavioral features are factored in. Specificity score leaves room for improvement.

Baseline accuracy: 0.59

**Best performing model:
Gradient Boost model,
71% accuracy.**

Criminal History Dataset

Non-NLP Model

model_used	train_score	test_score	specificity
grad	0.886566	0.881635	0.871616
rf	0.891616	0.877877	0.860789
ada	0.863786	0.866134	0.859242
bag	0.996477	0.863786	0.878577
fnn1	0.852513	0.854392	0.861562

Significant jump in accuracy from the basic model, as various behavioral features are factored in. Specificity score leaves room for improvement.

Baseline accuracy: 0.60

**Best performing model:
Gradient Boost model,
88% accuracy.**

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Can we predict recidivism?

- In theory, yes. Our criminal history model scored 88% in test accuracy, a 27% increase from the baseline.

Should we predict recidivism?

- It's critical to maintain objectivity throughout, as there are many opportunities for implicit bias to leak into modeling. We recommend excluding features that are outside of the subject's control e.g. age, race, gender, etc.

What factors should be included in predictive models?

- We found that certain behavioral and demographic features like gang affiliation, employment during parole, previous charges, and education level are strong determinants of one's likelihood of recidivism.

Recommendations

1. Increase employment opportunities for prisoners and parolees.

- Strong indication in our analysis that being employed during parole will reduce probability of recidivism.
- There can be limited job opportunities for those who have been released from prison, but involvement in employment programs both during and after prison have been shown to reduce the likelihood of recidivism by up to 60%.

2. Target gang affiliates to break the cycle of recidivism.

- Those who are gang-affiliated were twice as likely to be rearrested compared to those who are not.
- Nonprofits like Filthy Rags Outreach focus on reducing gang violence by targeting current and former gang members in prison and equipping them with the skills needed to reenter their communities.

Recommendations

3. Increase access to educational programs and support for the incarcerated.

- Our analysis showed a 20% decrease in the likelihood of recidivism for those who had more than a high school education.
- According to the Bureau of Justice Statistics, there is a 43% reduction in recidivism rates for individuals who participate in prison education programs.
- Programs like The Last Mile focus on education for workforce reentry, and have significantly reduced the recidivism rates of their participants.

4. Improve rehabilitation services, such as mental health and substance abuse programs.

- Individuals with a substance abuse and/or mental health disorder are about twice as likely to be rearrested as opposed to those who have neither.
- More than 60% of individuals with a history of mental illness do not receive treatment while incarcerated.