

Final Project

William DeForest, Daniel Krasemann, Sam Chotzen 12/16/2021

Project Overview

Background:

Recidivism is a fundamental concept of criminal justice. It refers to a person's relapse into criminal behavior after the person has served their prison sentence. It is measured as criminal acts that resulted in rearrest, re-conviction, or return to prison with or without a new sentence during a three-year period following a person's release. Historically, recidivism rates have been very high, with several studies reporting that more than 60% of former prisoners recidivate within three years of being released. As a result, recidivism has become known as the "revolving door" in and out of prison. It is closely related to and significantly affects other important criminal justice topics such as incapacitation, specific deterrence, rehabilitation, and desistance. Incapacitation refers to stopping people from committing a crime by removing them from the community. Specific deterrence measures whether a sanction prevents people from committing further crime, once the sanction has been imposed or completed.

Rehabilitation refers to the reduction of crime by "repairing" the individual in some way by addressing his or her needs or deficits. Desistance refers to the process by which a person arrives at a permanent state of non-offending. Ultimately, a person released from prison will either recidivate or desist. Of course, the desired outcome is always desistance, so it is important to consider what the crucial factors are in determining whether an individual will recidivate or not.

The Bureau of Justice Statistics (BJS) has been collecting recidivism data since the early 1980s. This data has been used in the past by the National Institute of Justice (NIJ) to inform its probation and parole policy. However, this practice is highly controversial because the use of recidivism as a criminal justice statistic is a double-edged sword.

On the one hand, analyzing trends in recidivism data can help identify which prisoners are the most likely to recidivate. This of course can be used to prevent more crimes from

occurring. To this point, the main way the NJI has been utilizing this data is to be able to create a “supervision risk score” scale with which they can rank and identify the individuals with the highest risk of recidivating.

However, this is an issue because recidivism data is skewed by inconsistencies and biases in policing, charging, and supervision. Incarceration and recidivism most directly affect non-white and poor individuals, reflecting the “disproportionate minority contact” of the criminal justice system, and the link between poverty and criminal justice system involvement. Poor communities of color are affected by recidivism most acutely, since recidivism has a negative impact on economic mobility with financial consequences not only for the formerly incarcerated individual, but for their family and community as well.

As a result, in recent years there has been a shift in focus towards the creation of programs that promote desistance rather than prevent recidivism as well as towards improving law enforcement techniques and training to reduce the number of people who end up incarcerated in the first place. This is driven by research that shows that imprisonment has little positive effect on criminal activity other than for the period when the individual is incarcerated. Furthermore, it shows that imprisonment has disruptive effects on an individual’s life trajectory. This is because a perfect outcome of reintegration into society is often difficult to achieve for people leaving incarceration, as evidenced by the overwhelming prevalence of homelessness, unemployment, and poverty among formerly incarcerated people.

Project Objectives:

With this project, we hope to contribute to the research on achieving desistance by identifying the characteristics that are most highly correlated with recidivism in order to create targeted programs to help incarcerated individuals have a smoother reentry into society. We will focus first on what factors are the best predictors of recidivism and then dive deeper into certain categories of characteristics to hopefully help craft specific recommendations for ways to promote desistance and lower recidivism. While we will look to create a model with the highest accuracy, precision, and recall, we will be more interested in the specific variables within the model. We will focus on accuracy measures and Gini coefficients primarily through a random forest predictive model when determining the best predictors of recidivism.

Data Used:

We have data from the National Institute of Justice on roughly 26,000 individuals from the State of Georgia released from Georgia prisons on discretionary parole to the custody of the Georgia Department of Community Supervision (GDCS) for the purpose of post-incarceration supervision between January 1, 2013 and December 31, 2015. This data was previously used earlier this year as a part of the NIJ's "Recidivism Challenge" which tasked teams of students and professionals to create predictive models with the hopes of increasing recidivism prediction accuracy.

Summary of Data

To get a sense of the data, we looked at the raw number of Recidivism counts in the years 1, 2, 3, and total. In order to run our predictive models, we needed to remove all NA observations from the data set. Since there were not any columns that had a significant number of NA observations, removing all rows with NA observations was necessary. While this was not ideal, it still left us with 9,838 rows of data.

```
## # A tibble: 2 x 2
##   Recidivism_Arrest_Year1      n
##   <lgl>                    <int>
## 1 FALSE                    10134
## 2 TRUE                      4036
```

```
## # A tibble: 2 x 2
##   Recidivism_Arrest_Year2      n
##   <lgl>                    <int>
## 1 FALSE                    11418
## 2 TRUE                      2752
```

```
## # A tibble: 2 x 2
##   Recidivism_Arrest_Year3      n
##   <lgl>                    <int>
## 1 FALSE                    12535
## 2 TRUE                      1635
```

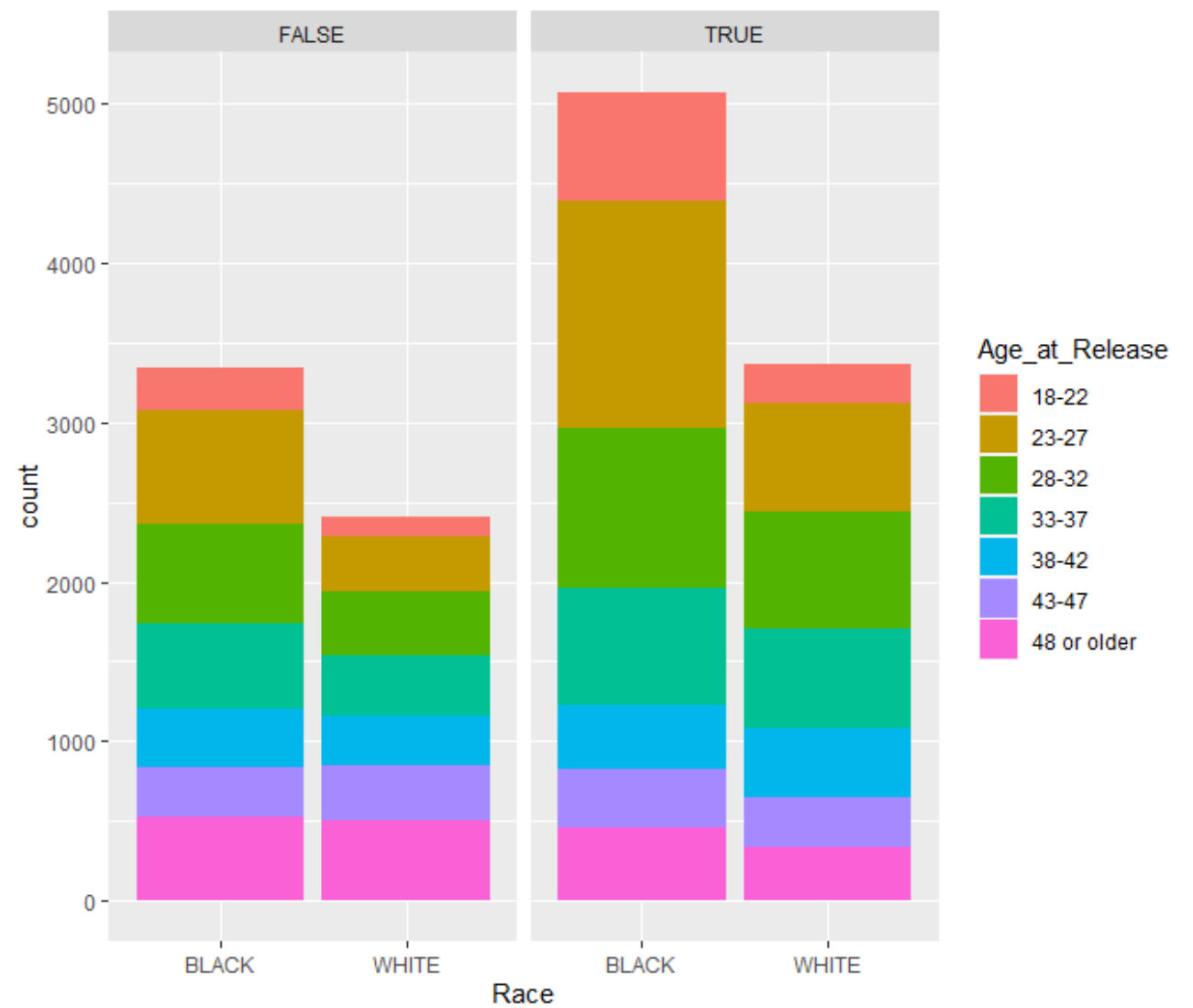
```
## # A tibble: 2 x 2
##   Recidivism_Within_3years      n
##   <lgl>                    <int>
## 1 FALSE                     5747
## 2 TRUE                      8423
```

As we can see, there were the most arrests in year 1, followed by year 2, and then year 3. Overall, there was recidivism for more than half of the prisoners within 3 years in the data set, as we can see by the number of TRUE observations in the `Recidivism_Within_3years`

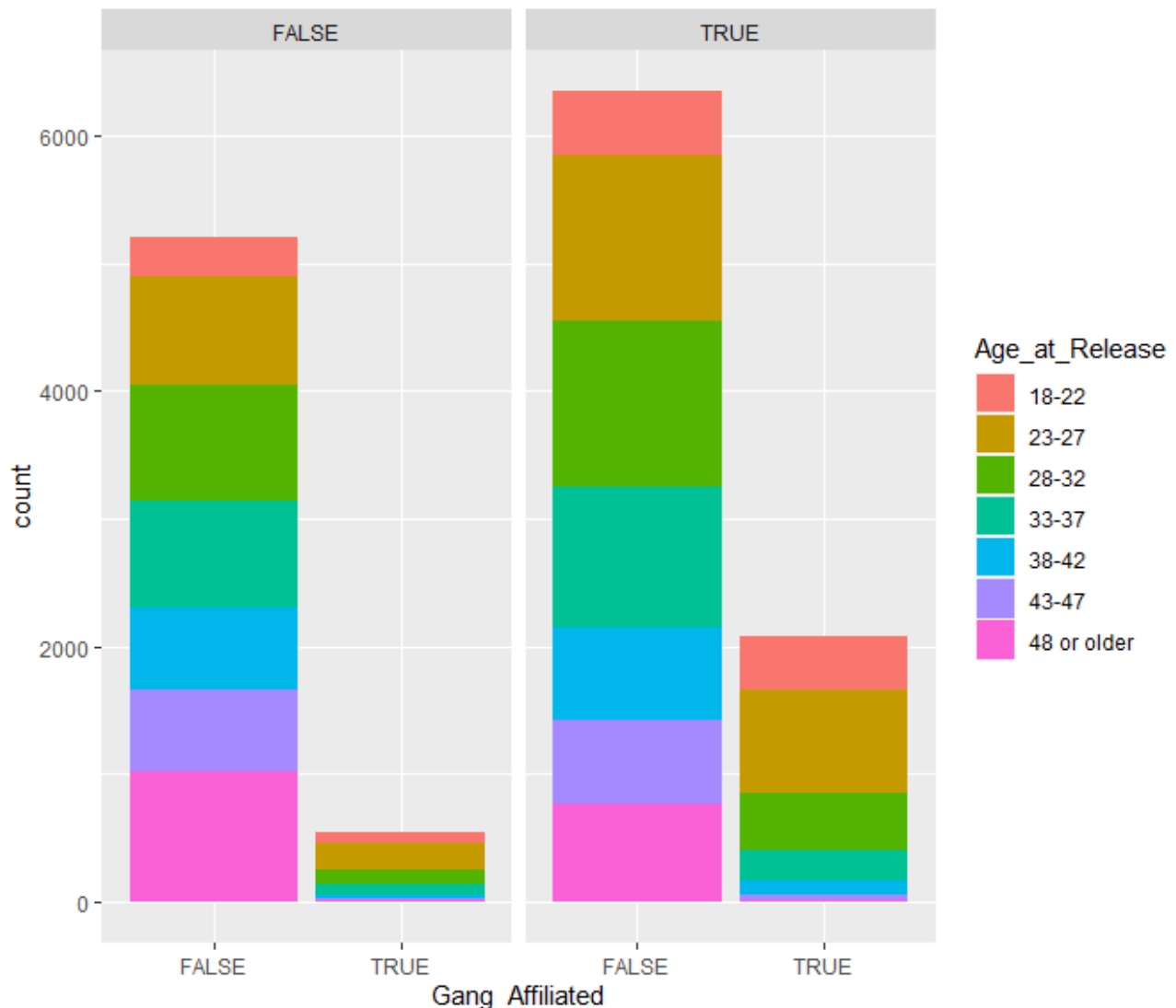
column. These initial findings related to the difficulty behind successful desistance and inspired our research to find the most accurate predictors for recidivism. We also observed that, as far as the predictive models go, if one were to predict that all cases of recidivism within 3 years were TRUE, the accuracy would be 59.44%.

Data Visualizations

Before diving into any predictive analysis, we created a couple of graphs to offer visualizations of clear data points in our data set. To begin, we created a bar graph measuring the number of recidivism counts within 3 years that were divided by race and age. As we can see, there were more arrests of younger individuals, the majority of whom were minorities. Reaching the youth and minorities when dealing with former prisoners will be an essential aspect of preventing recidivism. The graph below was also interesting because it showed us that there were more minorities in the data set than white individuals.



We also took a look at the effect of gang affiliation on recidivism rates. The visualization below shows us that having a gang affiliation does lead to higher rates of recidivism. Our predictive analysis further below showed how impactful the gang affiliation column truly was in predicting recidivism. Again, we also see that the younger ages were most prominent in the data set, since there were relatively few observations of individuals 48 or older.



Logistic Model

To begin our predictive analysis, we ran a basic logistic model to determine the accuracy, precision, and recall of the recidivism within 3 years. We chose the

`Recidivism_Within_3years` column to base our prediction off of because it looked slightly into the future and considered all three years after being released from prison rather than just one year. Before running the model, we needed to clean and organize the data to ensure it ran smoothly. This process included removing NA values to ensure the proper number of observations. As mentioned above, removing the NA values decreased the

number of rows in the data set, but it was necessary for the logistic model. We also removed two columns that included supervision data, because those columns were based on predictions that the researchers behind the data set made. If we were looking to make our own predictions, including previous predictions would only interfere with the objective analysis of our data. Another important element we had to complete before running the logistic model was removing any columns that only had 1 unique observation. Running the `sapply` command below showed us that the column `Gender` only had one unique observations. We blamed this on an error in the code since there were male and female observations. We removed the column to ensure our logistic model could run. The logistic model below contains all columns except for the ones mentioned above as well as the `ID` column.

Along with calculating the accuracy, precision, and recall, we also plotted a double density and ROC curve.

```
##
## Call:
## glm(formula = Recidivism_Within_3years ~ Race + Age_at_Release +
##   Residence_PUMA + Gang_Affiliated + Education_Level + Dependents +
##   Prison_Offense + Prison_Years + Prior_Arrest_Episodes_Felony +
##   Prior_Arrest_Episodes_Misd + Prior_Arrest_Episodes_Violent +
##   Prior_Arrest_Episodes_Property + Prior_Arrest_Episodes_Drug +
##   Prior_Arrest_Episodes_PPViolationCharges + Prior_Arrest_Episodes_DVCharges +
##   Prior_Arrest_Episodes_GunCharges + Prior_Conviction_Episodes_Felony +
##   Prior_Conviction_Episodes_Misd + Prior_Conviction_Episodes_Viol +
##   Prior_Conviction_Episodes_Prop + Prior_Conviction_Episodes_Drug +
##   Prior_Conviction_Episodes_PPViolationCharges + Prior_Conviction_Episodes_Domes
##   Prior_Conviction_Episodes_GunCharges + Prior_Revocations_Parole +
##
##   Prior_Revocations_Probation + Condition_MH_SA + Condition_Cog_Ed +
##   Condition_Other + Violations_ElectronicMonitoring + Violations_Instruction +
##   Violations_FailToReport + Violations_MoveWithoutPermission +
##   Delinquency_Reports + Program_Attendances + Program_UnexcusedAbsences +
##   Residence_Changes + Avg_Days_per_DrugTest + DrugTests_THC_Positive +
##   DrugTests_Cocaine_Positive + DrugTests_Meth_Positive + DrugTests_Other_Positiv
##   Percent_Days_Employed + Jobs_Per_Year + Employment_Exempt,
##   family = "binomial", data = NJ)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0716  -0.9364   0.4883   0.8633   2.5741
##
## Coefficients:
##
##                                     Estimate Std. Error
## (Intercept)                      -0.4119775   0.3199757
## RaceWHITE                        0.1230427   0.0463242
```

## Age_at_Release23-27	-0.3091897	0.0826018
## Age_at_Release28-32	-0.5888404	0.0900444
## Age_at_Release33-37	-0.8725993	0.0959510
## Age_at_Release38-42	-1.0795892	0.1045294
## Age_at_Release43-47	-1.2767855	0.1084786
## Age_at_Release48 or older	-1.5720556	0.1064209
## Residence_PUMA	0.0106635	0.0028148
## Gang_AffiliatedTRUE	0.8088443	0.0605411
## Education_LevelHigh School Diploma	0.1297466	0.0591647
## Education_LevelLess than HS diploma	-0.0445859	0.0612261
## Dependents1	0.1160899	0.0555818
## Dependents2	0.0439394	0.0599969
## Dependents3 or more	0.0507122	0.0531130
## Prison_OffenseOther	0.1278261	0.0750589
## Prison_OffenseProperty	0.1620540	0.0674883
## Prison_OffenseViolent/Non-Sex	0.1873827	0.0766021
## Prison_OffenseViolent/Sex	-0.0708039	0.1223468
## Prison_YearsGreater than 2 to 3 years	-0.1673202	0.0585706
## Prison_YearsLess than 1 year	0.1376789	0.0572469
## Prison_YearsMore than 3 years	0.0117338	0.0597129
## Prior_Arrest_Episodes_Felony1	0.0894407	0.3036644
## Prior_Arrest_Episodes_Felony10 or more	1.1768178	0.3254827
## Prior_Arrest_Episodes_Felony2	0.4517800	0.3037936
## Prior_Arrest_Episodes_Felony3	0.6800029	0.3055459
## Prior_Arrest_Episodes_Felony4	0.7333111	0.3084061
## Prior_Arrest_Episodes_Felony5	0.7753657	0.3113786
## Prior_Arrest_Episodes_Felony6	0.8561642	0.3148681
## Prior_Arrest_Episodes_Felony7	0.6425710	0.3179072
## Prior_Arrest_Episodes_Felony8	0.8683457	0.3237783
## Prior_Arrest_Episodes_Felony9	1.0000167	0.3275764
## Prior_Arrest_Episodes_Misd1	0.1521341	0.0768664
## Prior_Arrest_Episodes_Misd2	0.1405437	0.0859211
## Prior_Arrest_Episodes_Misd3	0.1090519	0.0962574
## Prior_Arrest_Episodes_Misd4	0.3265198	0.1033259
## Prior_Arrest_Episodes_Misd5	0.2189643	0.1111844
## Prior_Arrest_Episodes_Misd6 or more	0.3713560	0.1079376
## Prior_Arrest_Episodes_Violent1	-0.0213224	0.0546491
## Prior_Arrest_Episodes_Violent2	-0.0028733	0.0723117
## Prior_Arrest_Episodes_Violent3 or more	-0.0126292	0.0845229
## Prior_Arrest_Episodes_Property1	0.1132862	0.0655846
## Prior_Arrest_Episodes_Property2	0.0973675	0.0797502
## Prior_Arrest_Episodes_Property3	0.0691387	0.0957443
## Prior_Arrest_Episodes_Property4	0.1028270	0.1123362
## Prior_Arrest_Episodes_Property5 or more	0.2718468	0.1179899
## Prior_Arrest_Episodes_Drug1	-0.0706636	0.0624381
## Prior_Arrest_Episodes_Drug2	0.0142817	0.0771755
## Prior_Arrest_Episodes_Drug3	0.0586924	0.0922341
## Prior_Arrest_Episodes_Drug4	-0.0061928	0.1083412
## Prior Arrest Episodes Drug5 or more	-0.0247642	0.1108809

## Prior_Arrest_Episodes_PPViolationCharges1	0.0704736	0.0669888
## Prior_Arrest_Episodes_PPViolationCharges2	0.1296930	0.0778022
## Prior_Arrest_Episodes_PPViolationCharges3	0.1202981	0.0883694
## Prior_Arrest_Episodes_PPViolationCharges4	0.2398370	0.0989311
## Prior_Arrest_Episodes_PPViolationCharges5 or more	0.3273994	0.0960828
## Prior_Arrest_Episodes_DVChargesTRUE	0.0478258	0.0649996
## Prior_Arrest_Episodes_GunChargesTRUE	0.0406669	0.0541754
## Prior_Conviction_Episodes_Felony1	0.0501685	0.0558392
## Prior_Conviction_Episodes_Felony2	0.0864788	0.0705704
## Prior_Conviction_Episodes_Felony3 or more	0.0830855	0.0810670
## Prior_Conviction_Episodes_Misd1	0.1116025	0.0647137
## Prior_Conviction_Episodes_Misd2	0.2086023	0.0784870
## Prior_Conviction_Episodes_Misd3	0.1498255	0.0912589
## Prior_Conviction_Episodes_Misd4 or more	0.2802647	0.0947537
## Prior_Conviction_Episodes_ViolTRUE	0.0042449	0.0564433
## Prior_Conviction_Episodes_Prop1	0.0033990	0.0610757
## Prior_Conviction_Episodes_Prop2	-0.0599872	0.0833501
## Prior_Conviction_Episodes_Prop3 or more	0.0634091	0.1002294
## Prior_Conviction_Episodes_Drug1	0.0835074	0.0611687
## Prior_Conviction_Episodes_Drug2 or more	0.0054477	0.0800510
## Prior_Conviction_Episodes_PPViolationChargesTRUE	-0.1274452	0.0555429
## Prior_Conviction_Episodes_DomesticViolenceChargesTRUE	0.0593816	0.0851404
## Prior_Conviction_Episodes_GunChargesTRUE	0.0073048	0.0682026
## Prior_Revocations_ParoleTRUE	0.4512750	0.0724646
## Prior_Revocations_ProbationTRUE	-0.1260069	0.0589663
## Condition_MH_SATrue	0.2776651	0.0460356
## Condition_Cog_EdTRUE	-0.0175501	0.0439183
## Condition_OtherTRUE	-0.0267434	0.0485135
## Violations_ElectronicMonitoringTRUE	0.3703555	0.0721249
## Violations_InstructionTRUE	0.1372815	0.0556496
## Violations_FailToReportTRUE	-0.0816881	0.0755551
## Violations_MoveWithoutPermissionTRUE	-0.0571210	0.0640830
## Delinquency_Reports1	0.5851988	0.1026609
## Delinquency_Reports2	-0.0446136	0.0926021
## Delinquency_Reports3	-0.3051479	0.0920176
## Delinquency_Reports4 or more	-0.5537548	0.0604168
## Program_Attendances1	0.0113974	0.1074735
## Program_Attendances10 or more	-0.3514788	0.0693656
## Program_Attendances2	-0.0072058	0.1143056
## Program_Attendances3	0.0511861	0.1293025
## Program_Attendances4	-0.1486280	0.1161470
## Program_Attendances5	0.0460651	0.0933221
## Program_Attendances6	-0.1411223	0.0646424
## Program_Attendances7	0.0582753	0.1095943
## Program_Attendances8	-0.0271090	0.1435678
## Program_Attendances9	-0.0855129	0.1516405
## Program_UnexcusedAbsences1	0.1187916	0.0802186
## Program_UnexcusedAbsences2	0.1089087	0.0959006
## Program_UnexcusedAbsences3 or more	0.0804345	0.0761227


```

## Residence_Changes1 0.1331387 0.0490379
## Residence_Changes2 0.1301336 0.0614483
## Residence_Changes3 or more 0.3852803 0.0669360
## Avg_Days_per_DrugTest 0.0002583 0.0001777
## DrugTests_THC_Positive 1.1064365 0.1733839
## DrugTests_Cocaine_Positive 0.9815050 0.3663180
## DrugTests_Meth_Positive 3.0101723 0.5300668
## DrugTests_Other_Positive 1.0508702 0.5454142
## Percent_Days_Employed -1.9439963 0.0641524
## Jobs_Per_Year 0.4864227 0.0327098
## Employment_ExemptTRUE -0.1210072 0.0612078
##
z value Pr(>|z|)
## (Intercept) -1.288 0.197911
## RaceWHITE 2.656 0.007904 **
## Age_at_Release23-27 -3.743 0.000182 ***
## Age_at_Release28-32 -6.539 6.17e-11 ***
## Age_at_Release33-37 -9.094 < 2e-16 ***
## Age_at_Release38-42 -10.328 < 2e-16 ***
## Age_at_Release43-47 -11.770 < 2e-16 ***
## Age_at_Release48 or older -14.772 < 2e-16 ***
## Residence_PUMA 3.788 0.000152 ***
## Gang_AffiliatedTRUE 13.360 < 2e-16 ***
## Education_LevelHigh School Diploma 2.193 0.028309 *
## Education_LevelLess than HS diploma -0.728 0.466480
## Dependents1 2.089 0.036741 *
## Dependents2 0.732 0.463948
## Dependents3 or more 0.955 0.339679
## Prison_OffenseOther 1.703 0.088566 .
## Prison_OffenseProperty 2.401 0.016341 *
## Prison_OffenseViolent/Non-Sex 2.446 0.014438 *
## Prison_OffenseViolent/Sex -0.579 0.562782
## Prison_YearsGreater than 2 to 3 years -2.857 0.004280 **

## Prison_YearsLess than 1 year 2.405 0.016172 *
## Prison_YearsMore than 3 years 0.197 0.844216
## Prior_Arrest_Episodes_Felony1 0.295 0.768347
## Prior_Arrest_Episodes_Felony10 or more 3.616 0.000300 ***
## Prior_Arrest_Episodes_Felony2 1.487 0.136981
## Prior_Arrest_Episodes_Felony3 2.226 0.026045 *
## Prior_Arrest_Episodes_Felony4 2.378 0.017419 *
## Prior_Arrest_Episodes_Felony5 2.490 0.012770 *
## Prior_Arrest_Episodes_Felony6 2.719 0.006546 **
## Prior_Arrest_Episodes_Felony7 2.021 0.043254 *
## Prior_Arrest_Episodes_Felony8 2.682 0.007320 **
## Prior_Arrest_Episodes_Felony9 3.053 0.002267 **
## Prior_Arrest_Episodes_Misd1 1.979 0.047793 *
## Prior_Arrest_Episodes_Misd2 1.636 0.101896
## Prior_Arrest_Episodes_Misd3 1.133 0.257248
## Prior_Arrest_Episodes_Misd4 3.160 0.001577 **
## Prior Arrest Episodes Misd5 1.969 0.048910 *

```

```

... ..
## Prior_Arrest_Episodes_Misd6 or more 3.440 0.000581 ***
## Prior_Arrest_Episodes_Violent1 -0.390 0.696412
## Prior_Arrest_Episodes_Violent2 -0.040 0.968304
## Prior_Arrest_Episodes_Violent3 or more -0.149 0.881224
## Prior_Arrest_Episodes_Property1 1.727 0.084109 .
## Prior_Arrest_Episodes_Property2 1.221 0.222121
## Prior_Arrest_Episodes_Property3 0.722 0.470222
## Prior_Arrest_Episodes_Property4 0.915 0.360008
## Prior_Arrest_Episodes_Property5 or more 2.304 0.021224 *
## Prior_Arrest_Episodes_Drug1 -1.132 0.257745
## Prior_Arrest_Episodes_Drug2 0.185 0.853186
## Prior_Arrest_Episodes_Drug3 0.636 0.524554
## Prior_Arrest_Episodes_Drug4 -0.057 0.954418
## Prior_Arrest_Episodes_Drug5 or more -0.223 0.823270
## Prior_Arrest_Episodes_PPViolationCharges1 1.052 0.292791
## Prior_Arrest_Episodes_PPViolationCharges2 1.667 0.095523 .
## Prior_Arrest_Episodes_PPViolationCharges3 1.361 0.173416
## Prior_Arrest_Episodes_PPViolationCharges4 2.424 0.015339 *
## Prior_Arrest_Episodes_PPViolationCharges5 or more 3.407 0.000656 ***
## Prior_Arrest_Episodes_DVChargesTRUE 0.736 0.461861
## Prior_Arrest_Episodes_GunChargesTRUE 0.751 0.452862
## Prior_Conviction_Episodes_Felony1 0.898 0.368948
## Prior_Conviction_Episodes_Felony2 1.225 0.220415
## Prior_Conviction_Episodes_Felony3 or more 1.025 0.305411
## Prior_Conviction_Episodes_Misd1 1.725 0.084607 .
## Prior_Conviction_Episodes_Misd2 2.658 0.007865 **
## Prior_Conviction_Episodes_Misd3 1.642 0.100639
## Prior_Conviction_Episodes_Misd4 or more 2.958 0.003098 **
## Prior_Conviction_Episodes_ViolTRUE 0.075 0.940050
## Prior_Conviction_Episodes_Prop1 0.056 0.955619
## Prior_Conviction_Episodes_Prop2 -0.720 0.471709
## Prior_Conviction_Episodes_Prop3 or more 0.633 0.526969

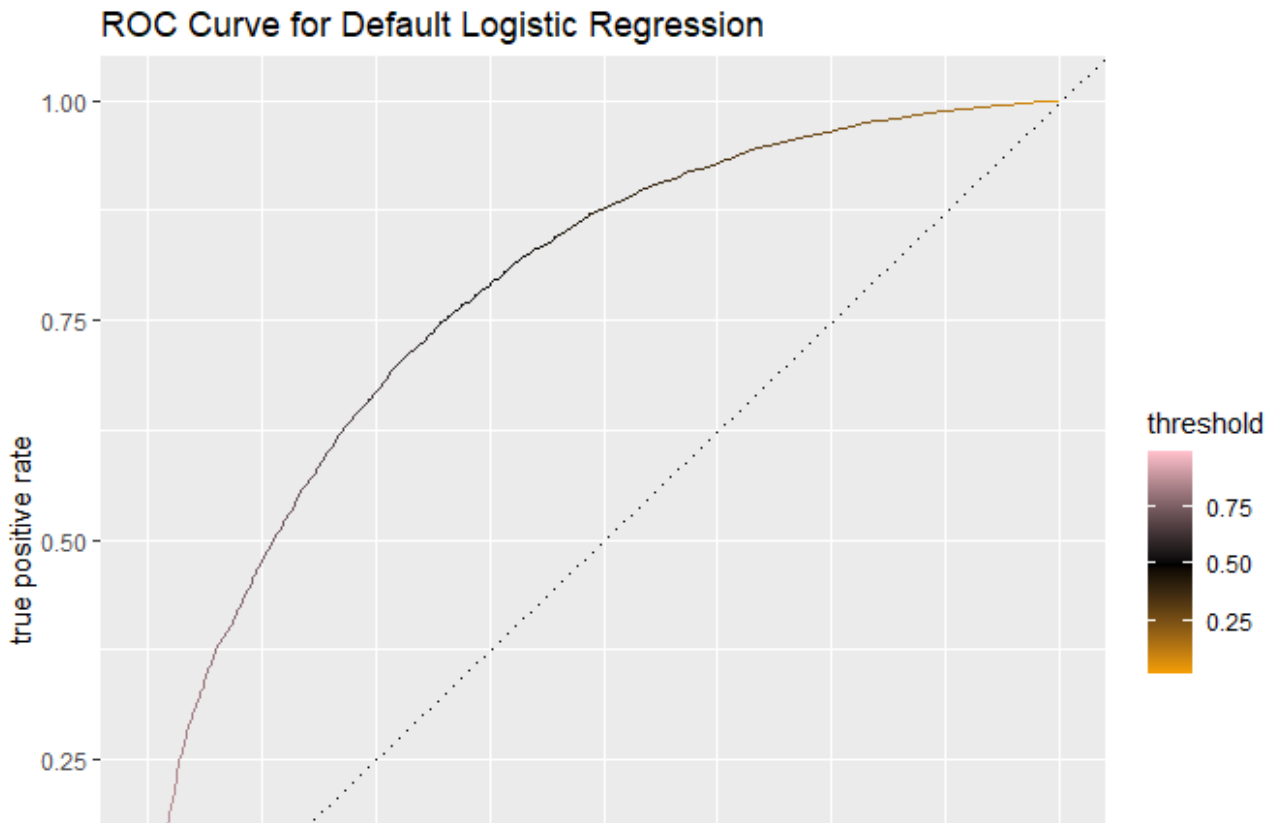
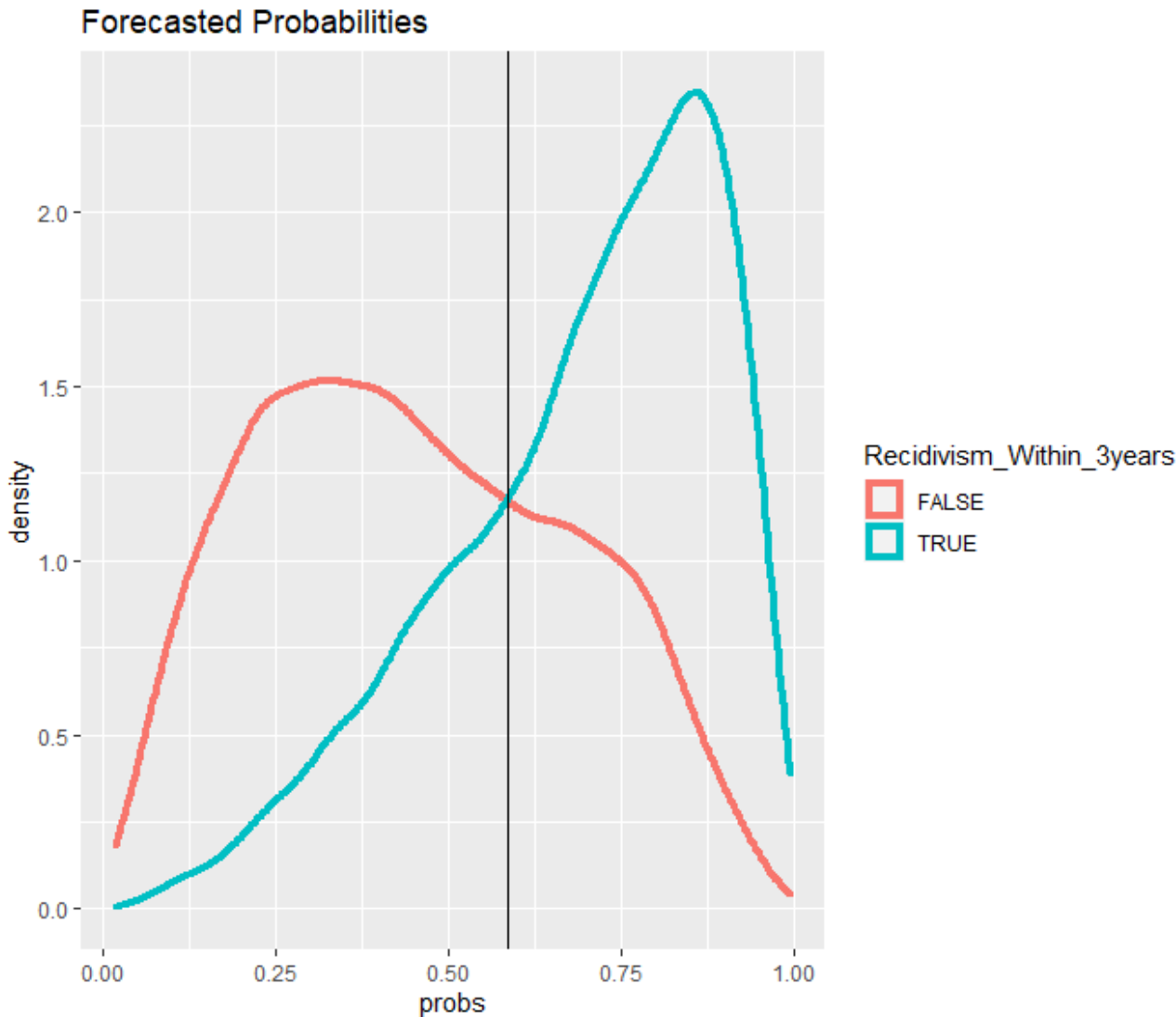
## Prior_Conviction_Episodes_Drug1 1.365 0.172191
## Prior_Conviction_Episodes_Drug2 or more 0.068 0.945743
## Prior_Conviction_Episodes_PPViolationChargesTRUE -2.295 0.021760 *
## Prior_Conviction_Episodes_DomesticViolenceChargesTRUE 0.697 0.485518
## Prior_Conviction_Episodes_GunChargesTRUE 0.107 0.914706
## Prior_Revocations_ParoleTRUE 6.228 4.74e-10 ***
## Prior_Revocations_ProbationTRUE -2.137 0.032603 *
## Condition_MH_SATTRUE 6.032 1.62e-09 ***
## Condition_Cog_EdTRUE -0.400 0.689445
## Condition_OtherTRUE -0.551 0.581457
## Violations_ElectronicMonitoringTRUE 5.135 2.82e-07 ***
## Violations_InstructionTRUE 2.467 0.013629 *
## Violations_FailToReportTRUE -1.081 0.279620
## Violations_MoveWithoutPermissionTRUE -0.891 0.372736
## Delinquency_Reports1 5.700 1.20e-08 ***
## Delinquency_Reports2 -0.482 0.629964
## Delinquency_Reports3 -3.316 0.000913 ***

```

```

## Delinquency_Reports4 or more -9.166 < 2e-16 ***
## Program_Attendances1 0.106 0.915544
## Program_Attendances10 or more -5.067 4.04e-07 ***
## Program_Attendances2 -0.063 0.949735
## Program_Attendances3 0.396 0.692206
## Program_Attendances4 -1.280 0.200667
## Program_Attendances5 0.494 0.621579
## Program_Attendances6 -2.183 0.029027 *
## Program_Attendances7 0.532 0.594908
## Program_Attendances8 -0.189 0.850231
## Program_Attendances9 -0.564 0.572809
## Program_UnexcusedAbsences1 1.481 0.138647
## Program_UnexcusedAbsences2 1.136 0.256107
## Program_UnexcusedAbsences3 or more 1.057 0.290675
## Residence_Changes1 2.715 0.006627 **
## Residence_Changes2 2.118 0.034194 *
## Residence_Changes3 or more 5.756 8.62e-09 ***
## Avg_Days_per_DrugTest 1.454 0.145917
## DrugTests_THC_Positive 6.381 1.75e-10 ***
## DrugTests_Cocaine_Positive 2.679 0.007376 **
## DrugTests_Meth_Positive 5.679 1.36e-08 ***
## DrugTests_Other_Positive 1.927 0.054012 .
## Percent_Days_Employed -30.303 < 2e-16 ***
## Jobs_Per_Year 14.871 < 2e-16 ***
## Employment_ExemptTRUE -1.977 0.048043 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 19135 on 14169 degrees of freedom
## Residual deviance: 15428 on 14059 degrees of freedom
##
## AIC: 15650
##
## Number of Fisher Scoring iterations: 5
##
## prediction_default
## Recidivism_Within_3years FALSE TRUE
## FALSE 3412 2335
## TRUE 1541 6882
##
## # A tibble: 1 x 3
## accuracy precision recall
## <dbl> <dbl> <dbl>
## 1 0.726 0.747 0.817

```



The results of the logistic model were promising. The most important variable for us was the accuracy, which was 72.6%. The double density curves and ROC curve further demonstrated the accuracy of the model. The double density curves had some (but not a ton of) overlap, and the area under the ROC curve was well above 0.5.

In terms of comparing accuracy to assess the strongest predictive model, more important to us than the logistic model was the K-fold cross validation method of assessing accuracy. This method considered over-fitting, which we needed in order to produce an accurate prediction and forecast of recidivism in our data set. The K-fold data below shows that the accuracy of the model using all of the variables was 71.9%, which was slightly lower than the logistic model alone.

```
## [1] 0.2800988 0.2790049
```

```
## [1] 0.7199012 0.7209951
```

We next wanted to run a logistic model using the significant variables that were identified in the previous logistic model and other variables we believed would be important to include. We looked at the variables with the lowest significance codes, or the variables with the most stars (*) next to their name. We subjectively chose other variables (Race , Program_Attendances) that we were interested in exploring further.

```
##
## Call:
## glm(formula = Recidivism_Within_3years ~ Race + Percent_Days_Employed +
##      Jobs_Per_Year + Avg_Days_per_DrugTest + Program_Attendances +
##      DrugTests_THC_Positive, family = "binomial", data = NJ)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8940  -1.0429   0.6805   0.9225   1.7287
##
## Coefficients:
##                                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)                        0.9220278  0.0456016  20.219  < 2e-16 ***
```

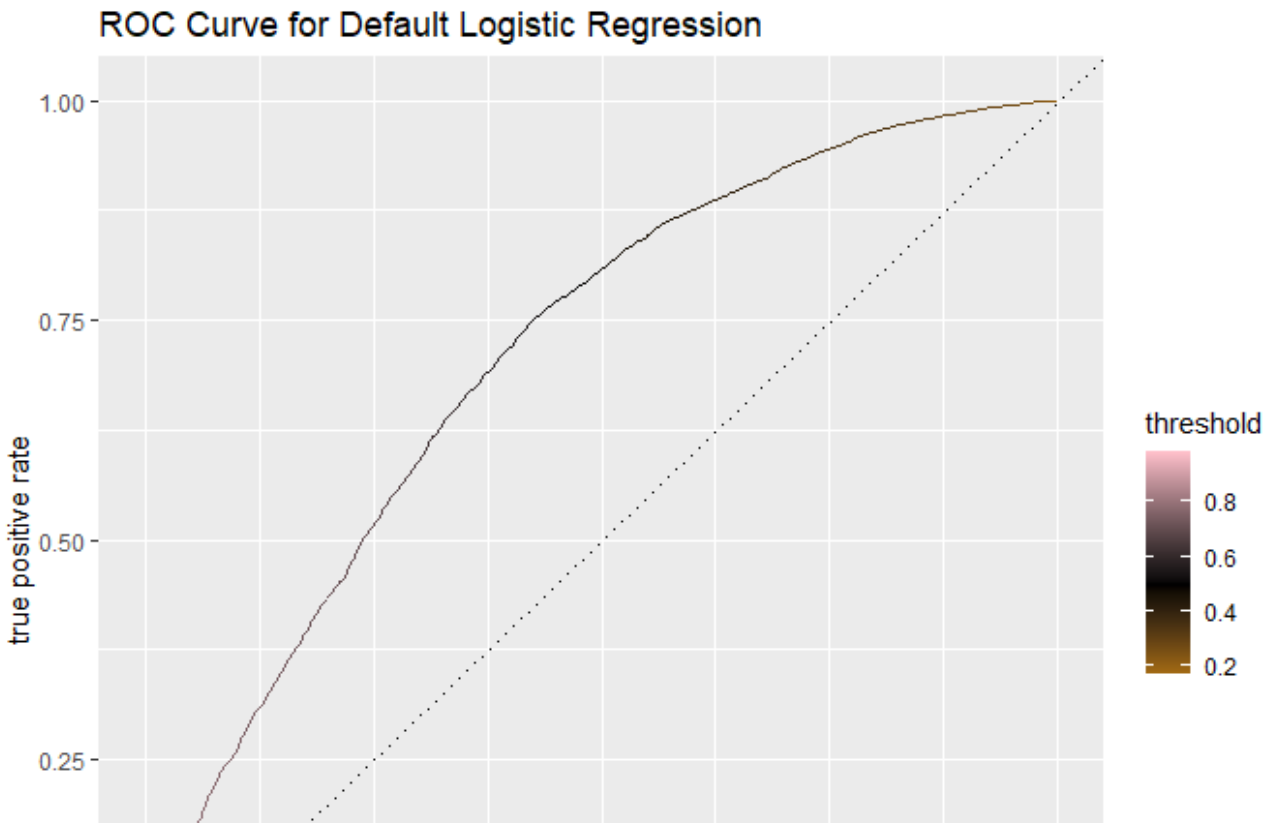
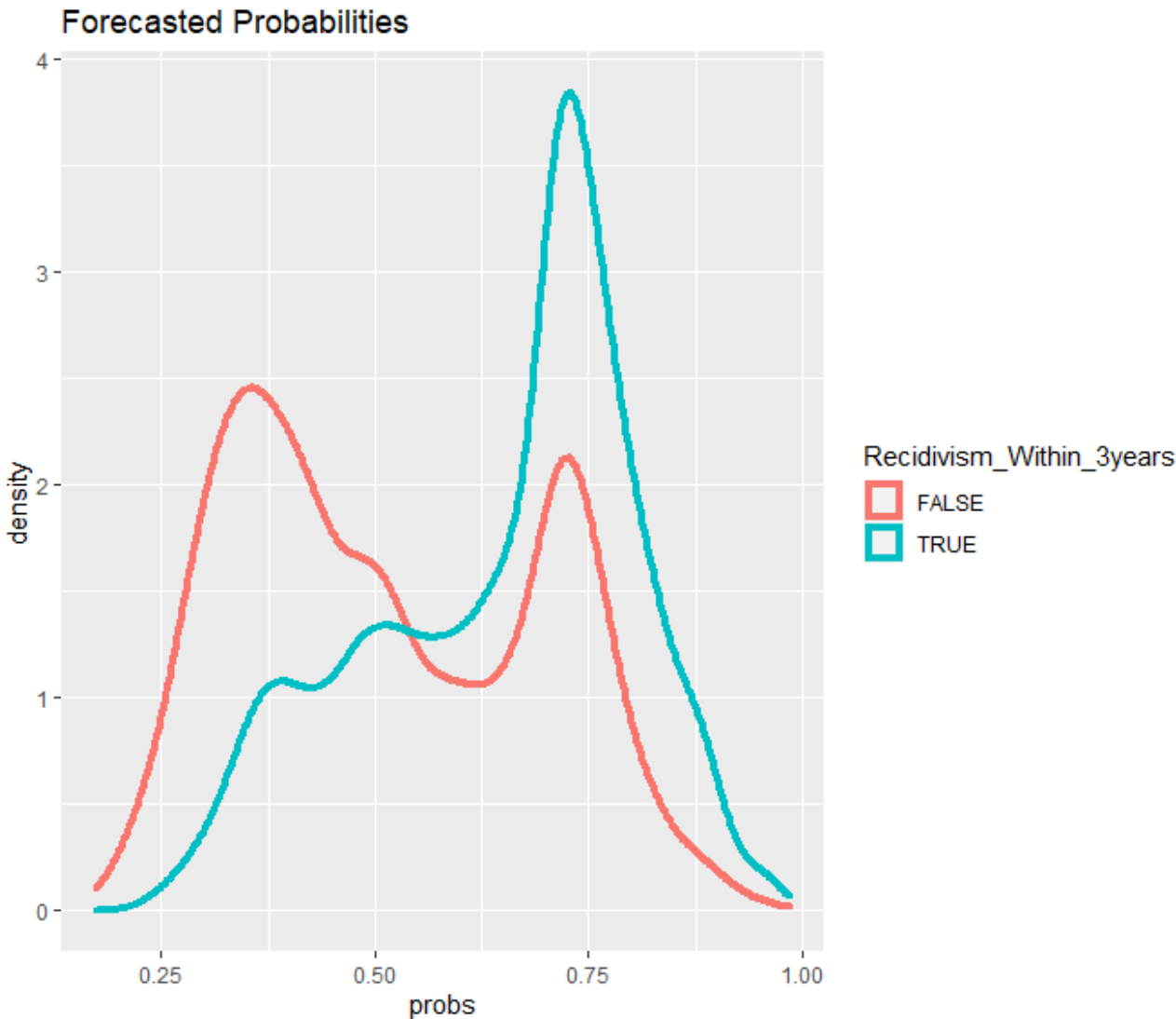
```

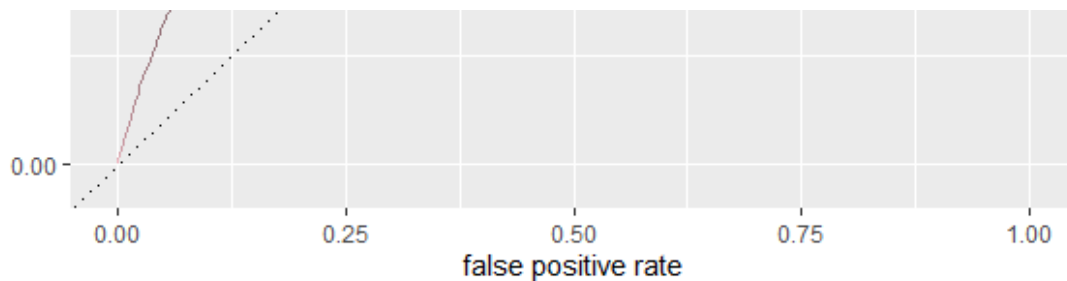
## RaceWHITE          0.1575044  0.0384652   4.095 4.23e-05 ***
## Percent_Days_Employed -2.1195521  0.0590089 -35.919 < 2e-16 ***
## Jobs_Per_Year        0.5869793  0.0303809  19.321 < 2e-16 ***
## Avg_Days_per_DrugTest -0.0002140  0.0001608  -1.330  0.1834
## Program_Attendances1  0.1574209  0.0987807   1.594  0.1110
## Program_Attendances10 or more -0.4419571  0.0542982  -8.139 3.97e-16 ***
## Program_Attendances2  0.1156074  0.1054958   1.096  0.2731
## Program_Attendances3  0.1372923  0.1189861   1.154  0.2486
## Program_Attendances4  0.0890358  0.1060517   0.840  0.4012
## Program_Attendances5  0.1474813  0.0860259   1.714  0.0865 .
## Program_Attendances6 -0.0832895  0.0585867  -1.422  0.1551
## Program_Attendances7  0.1440952  0.1002229   1.438  0.1505
## Program_Attendances8  0.0797526  0.1311432   0.608  0.5431
## Program_Attendances9  0.0332133  0.1380489   0.241  0.8099
## DrugTests_THC_Positive 2.1443370  0.1638187  13.090 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 19135  on 14169  degrees of freedom
## Residual deviance: 17199  on 14154  degrees of freedom
## AIC: 17231
##
## Number of Fisher Scoring iterations: 4

##               prediction_default
## Recidivism_Within_3years FALSE TRUE
##               FALSE  3014 2733
##               TRUE   1782 6641

## # A tibble: 1 x 3
##   accuracy precision recall
##   <dbl>      <dbl> <dbl>
## 1    0.681    0.708 0.788

```





Examining the results of the logistic model using more specific variables, we can determine which variables had positive or negative effects on the probability of recidivism. The variables with a negative coefficient, such as `Percent_Days_Employed`, `Avg._Days_per_DrugTest`, and `Program_Attendance`, of 10 or more decrease the probability of recidivism. These all make intuitive sense except possibly for the avg. days per drug test. We would expect that a former prisoner who got tested for drugs less frequently might be more likely to get away with using drugs. However, that does not appear to be the case. One theory might be that former prisoners who had no history of drug use would be tested less frequently and this distance from drugs could also contribute to them not being as likely to recidivate.

While the logistic models were interesting and informative, especially due to the fact that we can tell the direction of the effect of each variable, we wanted to perform further analysis and try and get a more accurate prediction by using a random forest model.

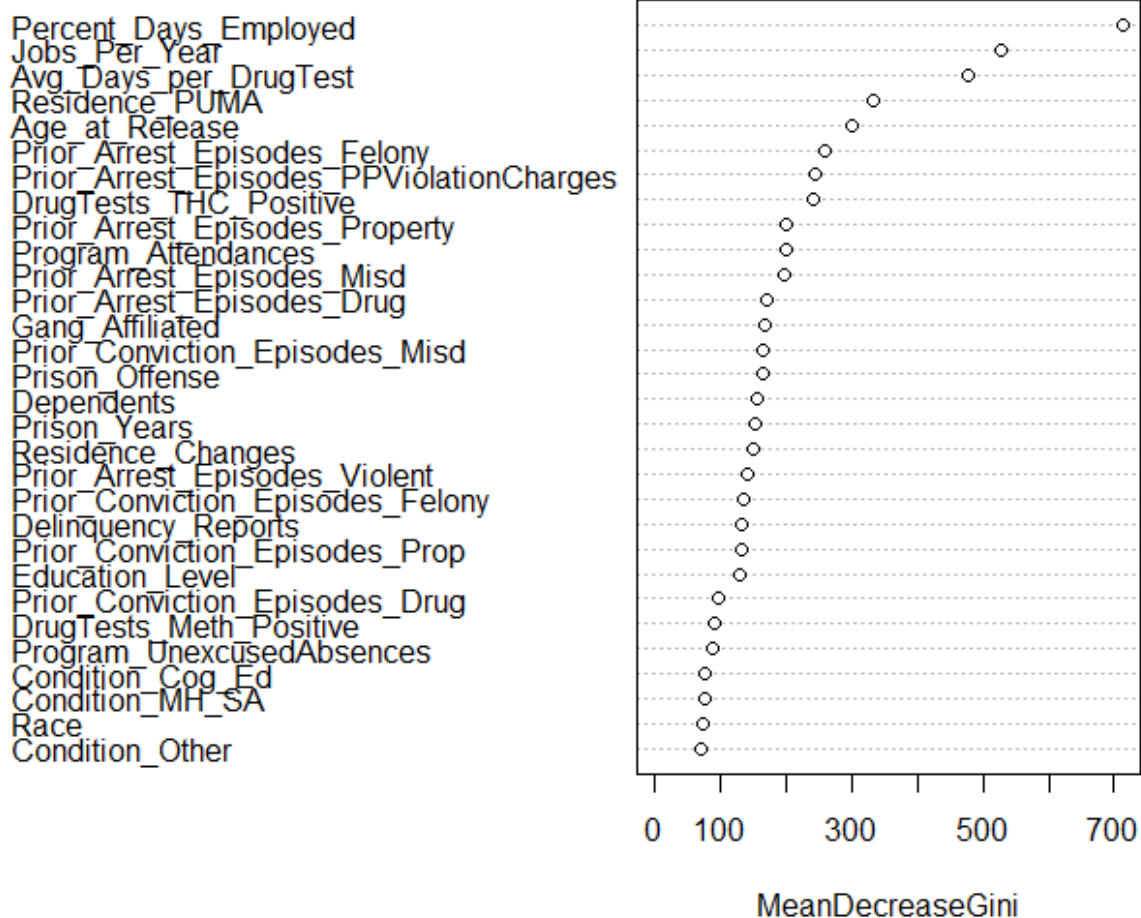
Random Forest Model

We anticipated that to produce the highest level of accuracy, we would need to use a random forest prediction model. This model would also provide analysis on which variables were the most significant in predicting recidivism, which we were interested in. To run the random forest model, we needed to convert the recidivism data to factors, which we did using the `as.factor` command. We used the same variables that we used for the logistic model, and found the accuracy to be 72.58%, which is slightly higher than the K-fold model.

##	MeanDecreaseGini
## Race	71.85845
## Age_at_Release	299.96800
## Residence_PUMA	331.59255
## Gang_Affiliated	168.41652
## Education_Level	127.35713
## Dependents	156.62664
## Prison_Offense	162.92739
## Prison_Years	152.74028
## Prior_Arrest_Episodes_Felony	258.29446

## Prior_Arrest_Episodes_Misd	196.16159
## Prior_Arrest_Episodes_Violent	139.92681
## Prior_Arrest_Episodes_Property	199.52402
## Prior_Arrest_Episodes_Drug	171.32246
## Prior_Arrest_Episodes_PPViolationCharges	242.44079
## Prior_Arrest_Episodes_DVCharges	49.43178
## Prior_Arrest_Episodes_GunCharges	64.96398
## Prior_Conviction_Episodes_Felony	134.73904
## Prior_Conviction_Episodes_Misd	164.92402
## Prior_Conviction_Episodes_Viol	65.74422
## Prior_Conviction_Episodes_Prop	130.63890
## Prior_Conviction_Episodes_Drug	97.08147
## Prior_Conviction_Episodes_PPViolationCharges	54.27880
## Prior_Conviction_Episodes_DomesticViolenceCharges	33.21190
## Prior_Conviction_Episodes_GunCharges	48.69047
## Prior_Revocations_Parole	44.03652
## Prior_Revocations_Probation	48.05805
## Condition_MH_SA	75.19074
## Condition_Cog_Ed	76.39191
## Condition_Other	69.20717
## Violations_ElectronicMonitoring	44.29851
## Violations_Instruction	57.25658
## Violations_FailToReport	37.99011
## Violations_MoveWithoutPermission	47.34712
## Delinquency_Reports	132.75209
## Program_Attendances	199.34340
## Program_UnexcusedAbsences	87.05784
## Residence_Changes	148.15519
## Avg_Days_per_DrugTest	476.00628
## DrugTests_THC_Positive	241.03153
## DrugTests_Cocaine_Positive	66.78018
## DrugTests_Meth_Positive	88.82059
## DrugTests_Other_Positive	50.82590
## Percent_Days_Employed	711.94948
## Jobs_Per_Year	526.19680
## Employment_Exempt	54.37716

rforest_all



```
## # A tibble: 10 x 3
```

```
##   Recidivism_Within_3years probs_rf prediction_rf
```

```
##   <lgl>                      <dbl> <fct>
```

```
## 1 FALSE                     0.753 TRUE
```

```
## 2 TRUE                      0.619 TRUE
```

```
## 3 TRUE                      0.576 TRUE
```

```
## 4 FALSE                     0.420 FALSE
```

```
## 5 TRUE                      0.846 TRUE
```

```
## 6 FALSE                     0.203 FALSE
```

```
## 7 FALSE                     0.677 TRUE
```

```
## 8 TRUE                      0.868 TRUE
```

```
## 9 TRUE                      0.595 TRUE
```

```
## 10 TRUE                     0.591 TRUE
```

```
##
```

```
## Call:
```

```
##   randomForest(formula = as.factor(Recidivism_Within_3years) ~ Race + Age_at_R
```

```
##               Type of random forest: classification
```

```
##               Number of trees: 200
```

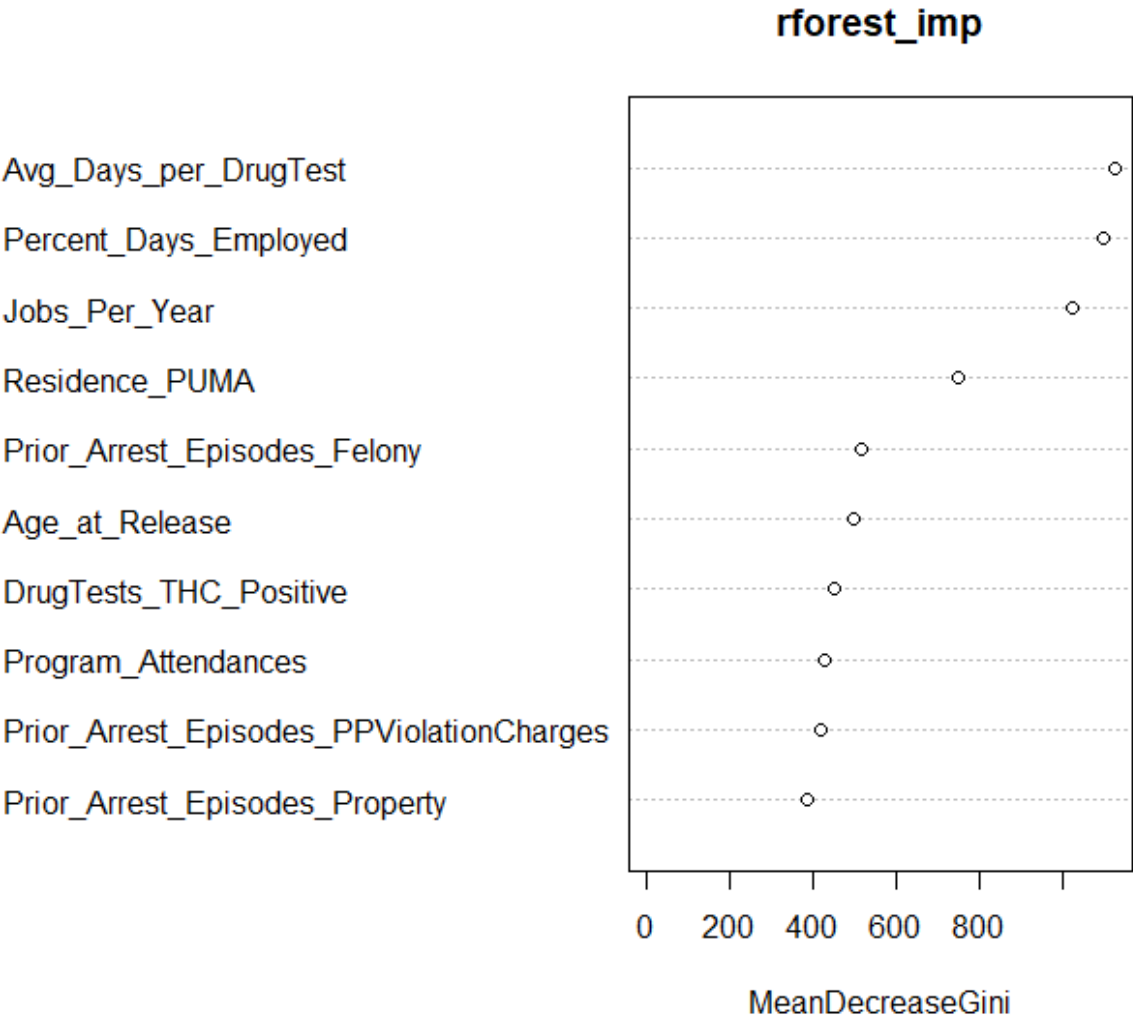
```
## No. of variables tried at each split: 5
```

```
## no. of variables tried at each split: 6
##
##          OOB estimate of  error rate: 27.42%
## Confusion matrix:
##          FALSE TRUE class.error
## FALSE   3139 2608    0.4538020
## TRUE    1277 7146    0.1516087

## # A tibble: 1 x 3
##   accuracy precision recall
##   <dbl>      <dbl> <dbl>
## 1    0.726    0.733 0.848
```

Next, we chose specific variables from the random forest model to run a further model on. We decided to use the 10 most significant variables as measured by the mean decrease in Gini coefficient.

```
##                               MeanDecreaseGini
## Age_at_Release                498.1771
## Residence_PUMA                 746.5437
## Prior_Arrest_Episodes_Felony   513.2167
## Prior_Arrest_Episodes_Property 382.5469
## Prior_Arrest_Episodes_PPViolationCharges 416.5583
## Program_Attendances            426.9954
## Avg_Days_per_DrugTest          1122.7288
## DrugTests_THC_Positive         451.8340
## Percent_Days_Employed          1096.3855
## Jobs_Per_Year                  1024.4575
```



```
##
## Call:
## randomForest(formula = as.factor(Recidivism_Within_3years) ~      Age_at_Release
##               Type of random forest: classification
##               Number of trees: 200
## No. of variables tried at each split: 3
##
## OOB estimate of  error rate: 29.33%
## Confusion matrix:
##      FALSE TRUE class.error
## FALSE  3118 2629   0.4574561
## TRUE   1527 6896   0.1812893

## # A tibble: 1 x 3
##   accuracy precision recall
##   <dbl>      <dbl> <dbl>
## 1    0.707    0.724  0.819
```

Looking at the importance of the variables based on the mean decrease in Gini, we can see that the most significant variable, or the variable with the highest mean decrease in Gini, is the `Avg_Days_per_DrugTest`. Other significant variables include the `Jobs_Per_Year`, `Percent_Days_Employed`, and `Residence_PUMA`. This is interesting in comparison with the results from the logistic model. Both models identify similar variables as being particularly effective at predicting recidivism. The random forest model suggests that the Avg days per drug test variable is the most significant. In addition, it further supports the takeaway that the variables measured after release are the best for predicting recidivism. This is important because it tells us that it is difficult to predict recidivism before releasing someone. This does not support the NIJ's use of recidivism statistics to inform their parole and probation policies because those are decided before the prisoner is released. As a result, those decisions of parole and probation policies will likely be more heavily influenced by characteristics like race and where the prisoner's home was. These don't take into account the prisoner's actual personality.

Digging Deeper

With the most important predictors identified, we decided to dig deeper into the data and see how the variables would change if we conditioned the dataset on certain characteristics.

We decided to look at race first. There is an immediate issue and limitation with the race variable in the dataset in that it only categorizes prisoners as black or white. Of course there are a myriad of other races and many very prevalent in the United States. However, Census data does tell us that white (57.8%) and black (31.9%) constitute nearly 90% of Georgia's racial makeup. As such, we believe the variable is usable and can still provide insight.

```
## # A tibble: 2 x 2
##   Race      N
##   <chr> <int>
## 1 BLACK  8400
## 2 WHITE  5770
```

As we can see by the above breakdown, black people are disproportionately represented in the data. This points towards the pervasive and disturbing issue of minorities being incarcerated more than white people. To gain more insight into the characteristics of these two groups, we ran two separate random forest models - one conditioned on race being black and another with race being white.

##	MeanDecreaseGini
## Age_at_Release	173.14835
## Residence_PUMA	207.14849
## Gang_Affiliated	93.73200
## Education_Level	81.06296
## Dependents	96.40381
## Prison_Offense	91.68201
## Prison_Years	95.79581
## Prior_Arrest_Episodes_Felony	157.21805
## Prior_Arrest_Episodes_Misd	118.78131
## Prior_Arrest_Episodes_Violent	89.73484
## Prior_Arrest_Episodes_Property	116.88968
## Prior_Arrest_Episodes_Drug	103.75051
## Prior_Arrest_Episodes_PPViolationCharges	126.89481
## Prior_Arrest_Episodes_DVCharges	29.53126
## Prior_Arrest_Episodes_GunCharges	43.95368
## Prior_Conviction_Episodes_Felony	82.29656
## Prior_Conviction_Episodes_Misd	98.53787
## Prior_Conviction_Episodes_Viol	42.49739
## Prior_Conviction_Episodes_Prop	75.09380
## Prior_Conviction_Episodes_Drug	61.19679
## Prior_Conviction_Episodes_PPViolationCharges	33.92692
## Prior_Conviction_Episodes_DomesticViolenceCharges	20.71586
## Prior_Conviction_Episodes_GunCharges	32.46799
## Prior_Revocations_Parole	28.76565
## Prior_Revocations_Probation	28.57760
## Condition_MH_SA	47.69966
## Condition_Cog_Ed	49.06725
## Condition_Other	43.27081
## Violations_ElectronicMonitoring	30.31096
## Violations_Instruction	38.36824
## Violations_FailToReport	23.59898
## Violations_MoveWithoutPermission	30.37780
## Delinquency_Reports	84.32916
## Program_Attendances	119.20601
## Program_UnexcusedAbsences	58.41871
## Residence_Changes	93.37815
## Avg_Days_per_DrugTest	278.55292
## DrugTests_THC_Positive	170.07057
## DrugTests_Cocaine_Positive	53.31678
## DrugTests_Meth_Positive	12.09869
## DrugTests_Other_Positive	23.03574
## Percent_Days_Employed	382.73065
## Jobs_Per_Year	284.47601
## Employment_Exempt	31.81762

##

```
## Call:
## randomForest(formula = as.factor(Recidivism_Within_3years) ~      Age_at_Release
##               Type of random forest: classification
##               Number of trees: 200
## No. of variables tried at each split: 5
##
## OOB estimate of error rate: 28.94%
## Confusion matrix:
##      FALSE TRUE class.error
## FALSE 1694 1642  0.4922062
## TRUE   789 4275  0.1558057

## # A tibble: 1 x 3
##   accuracy precision recall
##   <dbl>      <dbl> <dbl>
## 1    0.711    0.722  0.844

##
##                               MeanDecreaseGini
## Age_at_Release                133.00770
## Residence_PUMA                124.09726
## Gang_Affiliated                57.93742
## Education_Level                46.59053
## Dependents                    59.47148
## Prison_Offense                 76.47981
## Prison_Years                   59.77927
## Prior_Arrest_Episodes_Felony   113.16739
## Prior_Arrest_Episodes_Misd     92.96566
## Prior_Arrest_Episodes_Violent  52.30378
## Prior_Arrest_Episodes_Property 86.88817
## Prior_Arrest_Episodes_Drug     74.16283
## Prior_Arrest_Episodes_PPViolationCharges 133.14758
## Prior_Arrest_Episodes_DVCharges 20.33633
## Prior_Arrest_Episodes_GunCharges 21.66330
## Prior_Conviction_Episodes_Felony 57.22115
## Prior_Conviction_Episodes_Misd  76.42671
## Prior_Conviction_Episodes_Viol  23.91376
## Prior_Conviction_Episodes_Prop  58.58686
## Prior_Conviction_Episodes_Drug  39.62642
## Prior_Conviction_Episodes_PPViolationCharges 22.70578
## Prior_Conviction_Episodes_DomesticViolenceCharges 13.77145
## Prior_Conviction_Episodes_GunCharges 17.25261
## Prior_Revocations_Parole       17.35347
## Prior_Revocations_Probation    22.09771
## Condition_MH_SA                30.69587
## Condition_Cog_Ed               28.36505
## Condition_Other                 25.24882
## Violations_ElectronicMonitoring 13.10342
## Violations_Instruction          22.45938
## Violations_FailToReport        14.14769
```



```

## Violations_MoveWithoutPermission      18.60778
## Delinquency_Reports                   49.34553
## Program_Attendances                   81.44825
## Program_UnexcusedAbsences             33.27634
## Residence_Changes                     60.24793
## Avg_Days_per_DrugTest                  186.97862
## DrugTests_THC_Positive                 78.48243
## DrugTests_Cocaine_Positive             13.75437
## DrugTests_Meth_Positive                75.38479
## DrugTests_Other_Positive               27.55803
## Percent_Days_Employed                 310.10843
## Jobs_Per_Year                         208.78272
## Employment_Exempt                     22.85583

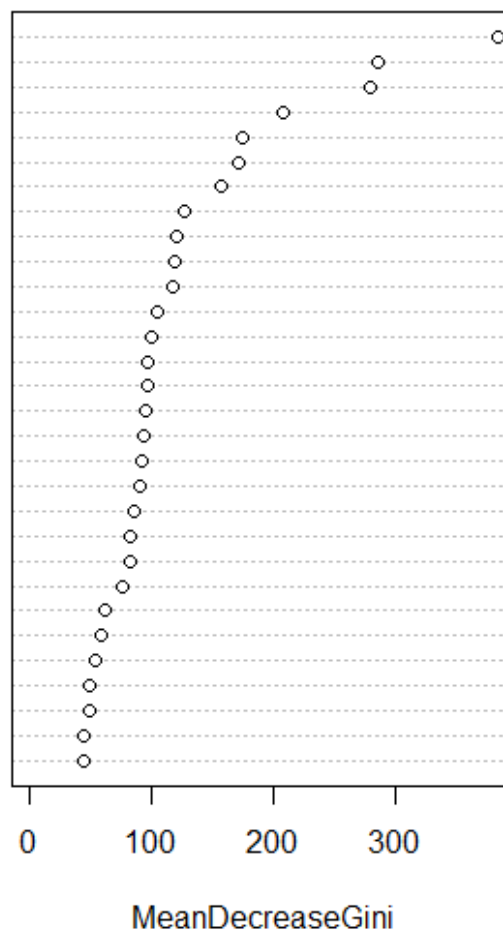
##
## Call:
##  randomForest(formula = as.factor(Recidivism_Within_3years) ~      Age_at_Release
##              Type of random forest: classification
##              Number of trees: 200
## No. of variables tried at each split: 7
##
##      OOB estimate of  error rate: 26.53%
## Confusion matrix:
##      FALSE TRUE class.error
## FALSE  1428  983   0.4077146
## TRUE    548 2811   0.1631438

## # A tibble: 1 x 3
##   accuracy precision recall
##   <dbl>      <dbl> <dbl>
## 1    0.735    0.741 0.837

```

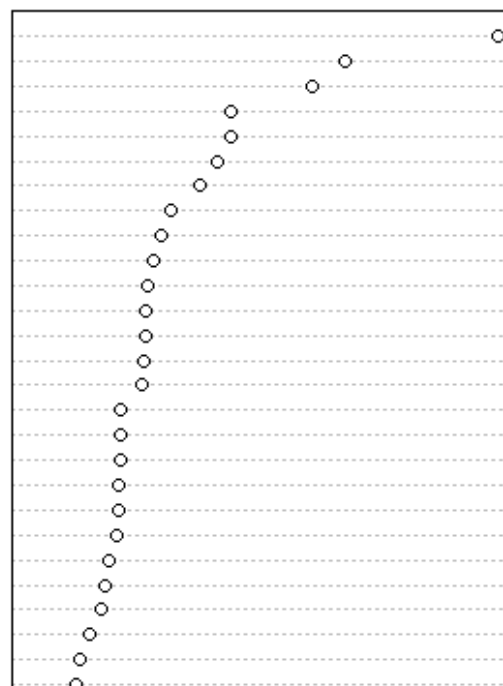
rforest_Black

Percent_Days_Employed
 Jobs_Per_Year
 Avg_Days_per_DrugTest
 Residence_PUMA
 Age_at_Release
 DrugTests_THC_Positive
 Prior_Arrest_Episodes_Felony
 Prior_Arrest_Episodes_PPViolationCharges
 Program_Attendances
 Prior_Arrest_Episodes_Misd
 Prior_Arrest_Episodes_Property
 Prior_Arrest_Episodes_Drug
 Prior_Conviction_Episodes_Misd
 Dependents
 Prison_Years
 Gang_Affiliated
 Residence_Changes
 Prison_Offense
 Prior_Arrest_Episodes_Violent
 Delinquency_Reports
 Prior_Conviction_Episodes_Felony
 Education_Level
 Prior_Conviction_Episodes_Prop
 Prior_Conviction_Episodes_Drug
 Program_UnexcusedAbsences
 DrugTests_Cocaine_Positive
 Condition_Cog_Ed
 Condition_MH_SA
 Prior_Arrest_Episodes_GunCharges
 Condition_Other

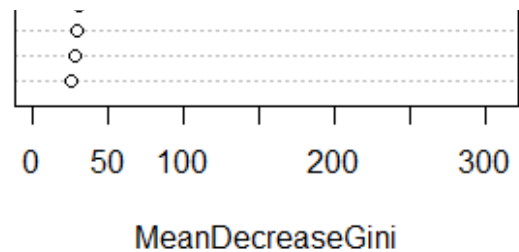


rforest_White

Percent_Days_Employed
 Jobs_Per_Year
 Avg_Days_per_DrugTest
 Prior_Arrest_Episodes_PPViolationCharges
 Age_at_Release
 Residence_PUMA
 Prior_Arrest_Episodes_Felony
 Prior_Arrest_Episodes_Misd
 Prior_Arrest_Episodes_Property
 Program_Attendances
 DrugTests_THC_Positive
 Prison_Offense
 Prior_Conviction_Episodes_Misd
 DrugTests_Meth_Positive
 Prior_Arrest_Episodes_Drug
 Residence_Changes
 Prison_Years
 Dependents
 Prior_Conviction_Episodes_Prop
 Gang_Affiliated
 Prior_Conviction_Episodes_Felony
 Prior_Arrest_Episodes_Violent
 Delinquency_Reports
 Education_Level
 Prior_Conviction_Episodes_Drug
 Program_UnexcusedAbsences
 Condition_MH_SA



Condition_With_3Y
 Condition_Cog_Ed
 DrugTests_Other_Positive
 Condition_Other



Comparing the results from the random forest models, a few interesting similarities and differences jump out. First, regardless of race, it appears that `Percent_Days_Employed`, `Jobs_Per_Year`, and `Avg._Days_per_DrugTest` were the most significant predictors of recidivism. All three of these variables were measured after release during the period of parole supervision. It makes sense that these would be the best predictors of recidivism because they showcase how the former prisoner is acting now that they are out of jail. If they are employed fewer days, bouncing around jobs, and not being tested for drugs frequently, the chances that they get up to no good are much higher. We can confirm this intuition regarding the positive or negative affects an increase in these variables has on the probability of recidivism by running a logistic regression.

The signs on the coefficients of the `percent_days_employed`, `jobs_per_year`, and `Avg._Days_per_DrugTest` variables confirm our intuition on how they affect the probability of recidivism for both races.

The next interesting takeaway is that the next most significant variable differs for the groups. For black prisoners, where they lived after release (`residence_PUMA`) was next most important and for white prisoners it was whether they had a prior arrest for violating probation or parole. These findings are extremely interesting because it seems to support the claim that the neighborhood/community one lives in has a large effect on the person's life and that black people come from neighborhoods that have higher rates of recidivism and crime than white people. This once again goes back to the severe racial inequality that is prevalent in the United States. Black communities and neighborhoods are historically less economically well-off as a result of past national policies such as redlining and this data suggests that this inequality extends to increasing the likelihood of ending up in jail and recidivating too.

```
##
## Call:
## glm(formula = Recidivism_Within_3years ~ Age_at_Release + Residence_PUMA +
##      Gang_Affiliated + Education_Level + Dependents + Prison_Offense +
##      Prison_Years + Prior_Arrest_Episodes_Felony + Prior_Arrest_Episodes_Misd +
##      Prior_Arrest_Episodes_Violent + Prior_Arrest_Episodes_Property +
##      Prior_Arrest_Episodes_Drug + Prior_Arrest_Episodes_PPViolationCharges +
```

```

##      Prior_Arrest_Episodes_DVCharges + Prior_Arrest_Episodes_GunCharges +
##      Prior_Conviction_Episodes_Felony + Prior_Conviction_Episodes_Misd +
##      Prior_Conviction_Episodes_Viol + Prior_Conviction_Episodes_Prop +
##      Prior_Conviction_Episodes_Drug + Prior_Conviction_Episodes_PPViolationCharges
##      Prior_Conviction_Episodes_DomesticViolenceCharges + Prior_Conviction_Episodes_
##      Prior_Revocations_Parole + Prior_Revocations_Probation +
##      Condition_MH_SA + Condition_Cog_Ed + Condition_Other + Violations_ElectronicMo
##      Violations_Instruction + Violations_FailToReport + Violations_MoveWithoutPermi
##      Delinquency_Reports + Program_Attendances + Program_UnexcusedAbsences +

##      Residence_Changes + Avg_Days_per_DrugTest + DrugTests_THC_Positive +
##      DrugTests_Cocaine_Positive + DrugTests_Meth_Positive + DrugTests_Other_Positiv
##      Percent_Days_Employed + Jobs_Per_Year + Employment_Exempt,
##      family = "binomial", data = NJ_Black)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q        Max
## -2.5383  -0.9555   0.5131   0.8709   2.5587
##
## Coefficients:
##                                     Estimate Std. Error
## (Intercept)                       -4.446e-01  3.932e-01
## Age_at_Release23-27                 -3.454e-01  9.990e-02
## Age_at_Release28-32                 -6.584e-01  1.121e-01
## Age_at_Release33-37                 -9.593e-01  1.212e-01
## Age_at_Release38-42                 -1.126e+00  1.354e-01
## Age_at_Release43-47                 -1.182e+00  1.432e-01
## Age_at_Release48 or older           -1.532e+00  1.383e-01
## Residence_PUMA                      1.516e-02  3.637e-03
## Gang_AffiliatedTRUE                  7.537e-01  7.237e-02
## Education_LevelHigh School Diploma  1.562e-01  7.655e-02
## Education_LevelLess than HS diploma  7.798e-02  7.870e-02
## Dependents1                         9.984e-02  7.374e-02
## Dependents2                         1.162e-01  8.003e-02
## Dependents3 or more                 6.958e-02  6.800e-02
## Prison_OffenseOther                  4.374e-02  9.635e-02
## Prison_OffenseProperty               1.318e-01  9.190e-02
## Prison_OffenseViolent/Non-Sex        1.794e-01  9.865e-02
## Prison_OffenseViolent/Sex           1.856e-01  1.707e-01
## Prison_YearsGreater than 2 to 3 years -4.654e-02  7.624e-02
## Prison_YearsLess than 1 year         1.007e-01  7.597e-02
## Prison_YearsMore than 3 years        4.557e-02  7.557e-02
## Prior_Arrest_Episodes_Felony1       1.533e-01  3.725e-01
## Prior_Arrest_Episodes_Felony10 or more 1.130e+00  4.040e-01
## Prior_Arrest_Episodes_Felony2       4.944e-01  3.734e-01
## Prior_Arrest_Episodes_Felony3       8.091e-01  3.758e-01
## Prior_Arrest_Episodes_Felony4       8.363e-01  3.798e-01
## Prior_Arrest_Episodes_Felony5       8.279e-01  3.844e-01
## Prior_Arrest_Episodes_Felony6       9.371e-01  3.894e-01
## Prior_Arrest_Episodes_Felony7       7.235e-01  3.936e-01

```

## Prior_Arrest_Episodes_Felony8	9.875e-01	4.010e-01
## Prior_Arrest_Episodes_Felony9	9.837e-01	4.067e-01
## Prior_Arrest_Episodes_Misd1	8.508e-02	9.451e-02
## Prior_Arrest_Episodes_Misd2	9.232e-02	1.082e-01
## Prior_Arrest_Episodes_Misd3	-2.145e-02	1.227e-01
## Prior_Arrest_Episodes_Misd4	3.036e-01	1.333e-01
## Prior_Arrest_Episodes_Misd5	1.704e-01	1.447e-01
## Prior_Arrest_Episodes_Misd6 or more	2.584e-01	1.392e-01
## Prior_Arrest_Episodes_Violent1	-6.940e-02	7.242e-02
## Prior_Arrest_Episodes_Violent2	-5.797e-02	9.243e-02
## Prior_Arrest_Episodes_Violent3 or more	-1.146e-01	1.075e-01
## Prior_Arrest_Episodes_Property1	1.706e-01	8.178e-02
## Prior_Arrest_Episodes_Property2	1.492e-01	1.017e-01
## Prior_Arrest_Episodes_Property3	1.049e-01	1.243e-01
## Prior_Arrest_Episodes_Property4	1.531e-01	1.485e-01
## Prior_Arrest_Episodes_Property5 or more	3.530e-01	1.570e-01
## Prior_Arrest_Episodes_Drug1	-1.903e-01	8.131e-02
## Prior_Arrest_Episodes_Drug2	-7.932e-02	1.017e-01
## Prior_Arrest_Episodes_Drug3	-1.889e-01	1.213e-01
## Prior_Arrest_Episodes_Drug4	-1.534e-01	1.413e-01
## Prior_Arrest_Episodes_Drug5 or more	-1.105e-01	1.450e-01
## Prior_Arrest_Episodes_PPViolationCharges1	5.764e-02	8.567e-02
## Prior_Arrest_Episodes_PPViolationCharges2	1.159e-01	1.009e-01
## Prior_Arrest_Episodes_PPViolationCharges3	-5.981e-03	1.153e-01
## Prior_Arrest_Episodes_PPViolationCharges4	2.011e-01	1.287e-01
## Prior_Arrest_Episodes_PPViolationCharges5 or more	2.901e-01	1.247e-01
## Prior_Arrest_Episodes_DVChargesTRUE	1.032e-01	8.700e-02
## Prior_Arrest_Episodes_GunChargesTRUE	1.108e-01	6.672e-02
## Prior_Conviction_Episodes_Felony1	3.967e-02	7.156e-02
## Prior_Conviction_Episodes_Felony2	4.028e-02	9.237e-02
## Prior_Conviction_Episodes_Felony3 or more	1.176e-01	1.066e-01
## Prior_Conviction_Episodes_Misd1	1.260e-01	8.317e-02
## Prior_Conviction_Episodes_Misd2	1.852e-01	1.014e-01
## Prior_Conviction_Episodes_Misd3	5.314e-02	1.197e-01
## Prior_Conviction_Episodes_Misd4 or more	3.740e-01	1.260e-01
## Prior_Conviction_Episodes_ViolTRUE	2.671e-02	7.071e-02
## Prior_Conviction_Episodes_Prop1	4.191e-05	7.853e-02
## Prior_Conviction_Episodes_Prop2	-4.604e-02	1.106e-01
## Prior_Conviction_Episodes_Prop3 or more	7.791e-02	1.357e-01
## Prior_Conviction_Episodes_Drug1	1.947e-01	8.187e-02
## Prior_Conviction_Episodes_Drug2 or more	2.925e-02	1.067e-01
## Prior_Conviction_Episodes_PPViolationChargesTRUE	-1.018e-01	7.300e-02
## Prior_Conviction_Episodes_DomesticViolenceChargesTRUE	-2.019e-02	1.126e-01
## Prior_Conviction_Episodes_GunChargesTRUE	-1.641e-02	8.265e-02
## Prior_Revocations_ParoleTRUE	5.047e-01	9.075e-02
## Prior_Revocations_ProbationTRUE	-8.907e-02	8.124e-02
## Condition_MH_SATTRUE	2.668e-01	5.678e-02
## Condition_Cog_EdTRUE	-4.916e-02	5.677e-02
## Condition_OtherTRUE	3.504e-02	6.378e-02

## Violations_ElectronicMonitoringTRUE	2.329e-01	8.490e-02
## Violations_InstructionTRUE	5.127e-02	6.816e-02
## Violations_FailToReportTRUE	-8.113e-02	9.426e-02
## Violations_MoveWithoutPermissionTRUE	-1.777e-01	8.125e-02
## Delinquency_Reports1	7.255e-01	1.334e-01
## Delinquency_Reports2	-1.176e-01	1.174e-01
## Delinquency_Reports3	-3.428e-01	1.179e-01
## Delinquency_Reports4 or more	-5.273e-01	7.634e-02
## Program_Attendances1	-6.272e-02	1.348e-01
## Program_Attendances10 or more	-3.180e-01	9.186e-02
## Program_Attendances2	-2.423e-02	1.425e-01
## Program_Attendances3	1.027e-01	1.610e-01
## Program_Attendances4	-2.593e-01	1.473e-01
## Program_Attendances5	-3.682e-02	1.226e-01
## Program_Attendances6	-1.332e-01	8.843e-02
## Program_Attendances7	1.041e-02	1.534e-01
## Program_Attendances8	-3.232e-02	1.871e-01
## Program_Attendances9	-1.567e-01	1.920e-01
## Program_UnexcusedAbsences1	7.860e-02	1.030e-01
## Program_UnexcusedAbsences2	-8.308e-03	1.176e-01
## Program_UnexcusedAbsences3 or more	7.463e-02	9.514e-02
## Residence_Changes1	2.014e-01	6.329e-02
## Residence_Changes2	1.162e-01	7.961e-02
## Residence_Changes3 or more	4.375e-01	8.839e-02
## Avg_Days_per_DrugTest	2.487e-04	2.120e-04
## DrugTests_THC_Positive	1.232e+00	2.004e-01
## DrugTests_Cocaine_Positive	7.312e-01	3.872e-01
## DrugTests_Meth_Positive	1.895e+00	1.606e+00
## DrugTests_Other_Positive	3.821e-01	8.291e-01
## Percent_Days_Employed	-1.971e+00	8.486e-02
## Jobs_Per_Year	5.334e-01	4.532e-02
## Employment_ExemptTRUE	-1.623e-01	8.003e-02
##	z value Pr(> z)	
## (Intercept)	-1.131	0.258211
## Age_at_Release23-27	-3.458	0.000545 ***
## Age_at_Release28-32	-5.876	4.21e-09 ***
## Age_at_Release33-37	-7.917	2.43e-15 ***
## Age_at_Release38-42	-8.314	< 2e-16 ***
## Age_at_Release43-47	-8.251	< 2e-16 ***
## Age_at_Release48 or older	-11.082	< 2e-16 ***
## Residence_PUMA	4.169	3.06e-05 ***
## Gang_AffiliatedTRUE	10.415	< 2e-16 ***
## Education_LevelHigh School Diploma	2.041	0.041271 *
## Education_LevelLess than HS diploma	0.991	0.321757
## Dependents1	1.354	0.175724
## Dependents2	1.451	0.146648
## Dependents3 or more	1.023	0.306234
## Prison_OffenseOther	0.454	0.649876
## Prison_OffenseProperty	1.434	0.151648

## Prison_OffenseViolent/Non-Sex	1.819	0.068961	.
## Prison_OffenseViolent/Sex	1.087	0.276876	
## Prison_YearsGreater than 2 to 3 years	-0.610	0.541546	
## Prison_YearsLess than 1 year	1.325	0.185176	
## Prison_YearsMore than 3 years	0.603	0.546497	
## Prior_Arrest_Episodes_Felony1	0.411	0.680741	
## Prior_Arrest_Episodes_Felony10 or more	2.797	0.005161	**
## Prior_Arrest_Episodes_Felony2	1.324	0.185509	
## Prior_Arrest_Episodes_Felony3	2.153	0.031332	*
## Prior_Arrest_Episodes_Felony4	2.202	0.027668	*
## Prior_Arrest_Episodes_Felony5	2.154	0.031248	*
## Prior_Arrest_Episodes_Felony6	2.407	0.016101	*
## Prior_Arrest_Episodes_Felony7	1.838	0.066063	.
## Prior_Arrest_Episodes_Felony8	2.462	0.013803	*
## Prior_Arrest_Episodes_Felony9	2.419	0.015570	*
## Prior_Arrest_Episodes_Misd1	0.900	0.367977	
## Prior_Arrest_Episodes_Misd2	0.853	0.393666	
## Prior_Arrest_Episodes_Misd3	-0.175	0.861214	
## Prior_Arrest_Episodes_Misd4	2.277	0.022812	*
## Prior_Arrest_Episodes_Misd5	1.178	0.238964	
## Prior_Arrest_Episodes_Misd6 or more	1.856	0.063413	.
## Prior_Arrest_Episodes_Violent1	-0.958	0.337881	
## Prior_Arrest_Episodes_Violent2	-0.627	0.530499	
## Prior_Arrest_Episodes_Violent3 or more	-1.066	0.286570	
## Prior_Arrest_Episodes_Property1	2.087	0.036923	*
## Prior_Arrest_Episodes_Property2	1.467	0.142465	
## Prior_Arrest_Episodes_Property3	0.844	0.398479	
## Prior_Arrest_Episodes_Property4	1.031	0.302475	
## Prior_Arrest_Episodes_Property5 or more	2.248	0.024595	*
## Prior_Arrest_Episodes_Drug1	-2.341	0.019252	*
## Prior_Arrest_Episodes_Drug2	-0.780	0.435197	
## Prior_Arrest_Episodes_Drug3	-1.558	0.119244	
## Prior_Arrest_Episodes_Drug4	-1.086	0.277597	
## Prior_Arrest_Episodes_Drug5 or more	-0.762	0.446017	
## Prior_Arrest_Episodes_PPViolationCharges1	0.673	0.501089	
## Prior_Arrest_Episodes_PPViolationCharges2	1.148	0.250766	
## Prior_Arrest_Episodes_PPViolationCharges3	-0.052	0.958627	
## Prior_Arrest_Episodes_PPViolationCharges4	1.563	0.118152	
## Prior_Arrest_Episodes_PPViolationCharges5 or more	2.326	0.020009	*
## Prior_Arrest_Episodes_DVChargesTRUE	1.187	0.235416	
## Prior_Arrest_Episodes_GunChargesTRUE	1.660	0.096929	.
## Prior_Conviction_Episodes_Felony1	0.554	0.579308	
## Prior_Conviction_Episodes_Felony2	0.436	0.662755	
## Prior_Conviction_Episodes_Felony3 or more	1.103	0.270078	
## Prior_Conviction_Episodes_Misd1	1.514	0.129918	
## Prior_Conviction_Episodes_Misd2	1.826	0.067817	.
## Prior_Conviction_Episodes_Misd3	0.444	0.656990	
## Prior_Conviction_Episodes_Misd4 or more	2.968	0.003002	**
## Prior_Conviction_Episodes_ViolTRUE	0.378	0.705589	

```

## Prior_Conviction_Episodes_Prop1          0.001 0.999574
## Prior_Conviction_Episodes_Prop2          -0.416 0.677277
## Prior_Conviction_Episodes_Prop3 or more    0.574 0.565720
## Prior_Conviction_Episodes_Drug1           2.378 0.017420 *
## Prior_Conviction_Episodes_Drug2 or more    0.274 0.784047
## Prior_Conviction_Episodes_PPViolationChargesTRUE -1.395 0.162979
## Prior_Conviction_Episodes_DomesticViolenceChargesTRUE -0.179 0.857740
## Prior_Conviction_Episodes_GunChargesTRUE  -0.199 0.842651
## Prior_Revocations_ParoleTRUE              5.561 2.68e-08 ***

## Prior_Revocations_ProbationTRUE           -1.096 0.272929
## Condition_MH_SATTRUE                     4.699 2.61e-06 ***
## Condition_Cog_EdTRUE                     -0.866 0.386520
## Condition_OtherTRUE                      0.549 0.582740
## Violations_ElectronicMonitoringTRUE       2.743 0.006084 **
## Violations_InstructionTRUE                0.752 0.451950
## Violations_FailToReportTRUE              -0.861 0.389390
## Violations_MoveWithoutPermissionTRUE      -2.187 0.028729 *
## Delinquency_Reports1                     5.438 5.38e-08 ***
## Delinquency_Reports2                     -1.002 0.316520
## Delinquency_Reports3                     -2.906 0.003658 **
## Delinquency_Reports4 or more              -6.907 4.96e-12 ***
## Program_Attendances1                     -0.465 0.641707
## Program_Attendances10 or more             -3.462 0.000537 ***
## Program_Attendances2                     -0.170 0.864962
## Program_Attendances3                     0.638 0.523646
## Program_Attendances4                     -1.761 0.078295 .
## Program_Attendances5                     -0.300 0.764035
## Program_Attendances6                     -1.507 0.131889
## Program_Attendances7                     0.068 0.945889
## Program_Attendances8                     -0.173 0.862814
## Program_Attendances9                     -0.816 0.414364
## Program_UnexcusedAbsences1                0.763 0.445272
## Program_UnexcusedAbsences2               -0.071 0.943655
## Program_UnexcusedAbsences3 or more        0.784 0.432790
## Residence_Changes1                       3.182 0.001461 **
## Residence_Changes2                       1.459 0.144495
## Residence_Changes3 or more                4.950 7.41e-07 ***
## Avg_Days_per_DrugTest                    1.173 0.240797
## DrugTests_THC_Positive                    6.147 7.92e-10 ***
## DrugTests_Cocaine_Positive                1.888 0.058987 .
## DrugTests_Meth_Positive                   1.180 0.237961
## DrugTests_Other_Positive                  0.461 0.644897
## Percent_Days_Employed                     -23.230 < 2e-16 ***
## Jobs_Per_Year                             11.769 < 2e-16 ***
## Employment_ExemptTRUE                    -2.028 0.042569 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)

```



```

##
##      Null deviance: 11286.8  on 8399  degrees of freedom
## Residual deviance:  9285.1  on 8290  degrees of freedom
## AIC: 9505.1
##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = Recidivism_Within_3years ~ Age_at_Release + Residence_PUMA +
##      Gang_Affiliated + Education_Level + Dependents + Prison_Offense +
##      Prison_Years + Prior_Arrest_Episodes_Felony + Prior_Arrest_Episodes_Misd +
##      Prior_Arrest_Episodes_Violent + Prior_Arrest_Episodes_Property +
##      Prior_Arrest_Episodes_Drug + Prior_Arrest_Episodes_PPViolationCharges +
##      Prior_Arrest_Episodes_DVCharges + Prior_Arrest_Episodes_GunCharges +
##      Prior_Conviction_Episodes_Felony + Prior_Conviction_Episodes_Misd +
##      Prior_Conviction_Episodes_Viol + Prior_Conviction_Episodes_Prop +
##      Prior_Conviction_Episodes_Drug + Prior_Conviction_Episodes_PPViolationCharges
##      Prior_Conviction_Episodes_DomesticViolenceCharges + Prior_Conviction_Episodes_
##      Prior_Revocations_Parole + Prior_Revocations_Probation +
##      Condition_MH_SA + Condition_Cog_Ed + Condition_Other + Violations_ElectronicMo
##      Violations_Instruction + Violations_FailToReport + Violations_MoveWithoutPermi
##      Delinquency_Reports + Program_Attendances + Program_UnexcusedAbsences +
##      Residence_Changes + Avg_Days_per_DrugTest + DrugTests_THC_Positive +
##      DrugTests_Cocaine_Positive + DrugTests_Meth_Positive + DrugTests_Other_Positiv
##      Percent_Days_Employed + Jobs_Per_Year + Employment_Exempt,
##      family = "binomial", data = NJ_White)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0423  -0.8726   0.4042   0.8201   2.6258
##
## Coefficients:
##                                     Estimate Std. Error
## (Intercept)                       -0.3475864   0.5498335
## Age_at_Release23-27                 -0.1588940   0.1523468
## Age_at_Release28-32                 -0.3457088   0.1578599
## Age_at_Release33-37                 -0.6284519   0.1642462
## Age_at_Release38-42                 -0.9281812   0.1738995
## Age_at_Release43-47                 -1.2861355   0.1775259
## Age_at_Release48 or older           -1.5161579   0.1780841
## Residence_PUMA                      0.0040107   0.0046116
## Gang_AffiliatedTRUE                 0.9364344   0.1158478
## Education_LevelHigh School Diploma  0.1374215   0.0961743
## Education_LevelLess than HS diploma -0.1774074   0.1008712
## Dependents1                        0.1285478   0.0873014
## Dependents2                       -0.0668084   0.0939566
## Dependents3 or more                0.0559206   0.0891131
## Prison_OffenseOther                 0.2988591   0.1242374

```

## Prison_OffenseProperty	0.2290741	0.1026917
## Prison_OffenseViolent/Non-Sex	0.1797899	0.1272619
## Prison_OffenseViolent/Sex	-0.2690844	0.1885131
## Prison_YearsGreater than 2 to 3 years	-0.3827060	0.0946494
## Prison_YearsLess than 1 year	0.1800349	0.0896125
## Prison_YearsMore than 3 years	-0.0275091	0.1014174
## Prior_Arrest_Episodes_Felony1	-0.1101733	0.5150294
## Prior_Arrest_Episodes_Felony10 or more	1.1641837	0.5453155
## Prior_Arrest_Episodes_Felony2	0.2862880	0.5125402
## Prior_Arrest_Episodes_Felony3	0.3915702	0.5153432
## Prior_Arrest_Episodes_Felony4	0.4577383	0.5202925
## Prior_Arrest_Episodes_Felony5	0.5881024	0.5234604
## Prior_Arrest_Episodes_Felony6	0.6723837	0.5286324
## Prior_Arrest_Episodes_Felony7	0.4432580	0.5332687
## Prior_Arrest_Episodes_Felony8	0.6197684	0.5429909
## Prior_Arrest_Episodes_Felony9	0.9369303	0.5482624
## Prior_Arrest_Episodes_Misd1	0.3258558	0.1370584
## Prior_Arrest_Episodes_Misd2	0.2916014	0.1470885
## Prior_Arrest_Episodes_Misd3	0.3928002	0.1609884
## Prior_Arrest_Episodes_Misd4	0.4079707	0.1699699
## Prior_Arrest_Episodes_Misd5	0.3665903	0.1813001
## Prior_Arrest_Episodes_Misd6 or more	0.5587687	0.1775612
## Prior_Arrest_Episodes_Violent1	0.0272810	0.0862064
## Prior_Arrest_Episodes_Violent2	0.0499790	0.1212178
## Prior_Arrest_Episodes_Violent3 or more	0.1746365	0.1436818
## Prior_Arrest_Episodes_Property1	0.0254650	0.1138040
## Prior_Arrest_Episodes_Property2	0.0170671	0.1337260
## Prior_Arrest_Episodes_Property3	0.0323195	0.1570106
## Prior_Arrest_Episodes_Property4	0.0276033	0.1793768
## Prior_Arrest_Episodes_Property5 or more	0.1323466	0.1875833
## Prior_Arrest_Episodes_Drug1	0.1222463	0.1003289
## Prior_Arrest_Episodes_Drug2	0.1859765	0.1222072
## Prior_Arrest_Episodes_Drug3	0.4468161	0.1472220
## Prior_Arrest_Episodes_Drug4	0.2529007	0.1735311
## Prior_Arrest_Episodes_Drug5 or more	0.1829601	0.1788221
## Prior_Arrest_Episodes_PPViolationCharges1	0.1314711	0.1118192
## Prior_Arrest_Episodes_PPViolationCharges2	0.2044845	0.1261970
## Prior_Arrest_Episodes_PPViolationCharges3	0.3287601	0.1419804
## Prior_Arrest_Episodes_PPViolationCharges4	0.3293295	0.1596693
## Prior_Arrest_Episodes_PPViolationCharges5 or more	0.4277899	0.1555333
## Prior_Arrest_Episodes_DVChargesTRUE	0.0038785	0.1015598
## Prior_Arrest_Episodes_GunChargesTRUE	-0.0788409	0.0965860
## Prior_Conviction_Episodes_Felony1	0.1036321	0.0921376
## Prior_Conviction_Episodes_Felony2	0.1650503	0.1123428
## Prior_Conviction_Episodes_Felony3 or more	0.0665121	0.1286838
## Prior_Conviction_Episodes_Misd1	0.0792829	0.1056418
## Prior_Conviction_Episodes_Misd2	0.2637392	0.1269320
## Prior_Conviction_Episodes_Misd3	0.2932885	0.1449225
## Prior_Conviction_Episodes_Misd4 or more	0.1820084	0.1481826

## Prior_Conviction_Episodes_ViolTRUE	-0.0424240	0.0973187
## Prior_Conviction_Episodes_Prop1	-0.0294089	0.1001739
## Prior_Conviction_Episodes_Prop2	-0.1512448	0.1310691
## Prior_Conviction_Episodes_Prop3 or more	-0.0181756	0.1549361
## Prior_Conviction_Episodes_Drug1	-0.0956670	0.0947992
## Prior_Conviction_Episodes_Drug2 or more	-0.0971280	0.1243868
## Prior_Conviction_Episodes_PPViolationChargesTRUE	-0.1748832	0.0880294
## Prior_Conviction_Episodes_DomesticViolenceChargesTRUE	0.1789560	0.1344077
## Prior_Conviction_Episodes_GunChargesTRUE	0.0965741	0.1247863
## Prior_Revocations_ParoleTRUE	0.3561200	0.1251082
## Prior_Revocations_ProbationTRUE	-0.1909776	0.0885609
## Condition_MH_SATTRUE	0.3072081	0.0811607
## Condition_Cog_EdTRUE	0.0087312	0.0715047
## Condition_OtherTRUE	-0.0922027	0.0771132
## Violations_ElectronicMonitoringTRUE	0.7091514	0.1433742
## Violations_InstructionTRUE	0.2933301	0.0993517
## Violations_FailToReportTRUE	-0.1098324	0.1309538
## Violations_MoveWithoutPermissionTRUE	0.1499195	0.1081777
## Delinquency_Reports1	0.3302983	0.1667297
## Delinquency_Reports2	0.0724192	0.1542895
## Delinquency_Reports3	-0.2391075	0.1515272
## Delinquency_Reports4 or more	-0.6360243	0.1018259
## Program_Attendances1	0.1652788	0.1853565
## Program_Attendances10 or more	-0.3541100	0.1099033
## Program_Attendances2	0.0155398	0.1976649
## Program_Attendances3	-0.0127134	0.2224841
## Program_Attendances4	0.0453047	0.1947609
## Program_Attendances5	0.2167416	0.1477051
## Program_Attendances6	-0.1065716	0.0984439
## Program_Attendances7	0.1676421	0.1608971
## Program_Attendances8	0.0043853	0.2319922
## Program_Attendances9	0.0535555	0.2516118
## Program_UnexcusedAbsences1	0.1892073	0.1325871
## Program_UnexcusedAbsences2	0.3492989	0.1709909
## Program_UnexcusedAbsences3 or more	0.0556158	0.1322677
## Residence_Changes1	0.0371232	0.0797828
## Residence_Changes2	0.1442213	0.0996638
## Residence_Changes3 or more	0.3026579	0.1058429
## Avg_Days_per_DrugTest	0.0003460	0.0003397
## DrugTests_THC_Positive	0.9437061	0.3666712
## DrugTests_Cocaine_Positive	4.9911022	1.4057372
## DrugTests_Meth_Positive	2.5758112	0.5798820
## DrugTests_Other_Positive	1.1328054	0.7126051
## Percent_Days_Employed	-1.9686531	0.1025395
## Jobs_Per_Year	0.4221381	0.0481862
## Employment_ExemptTRUE	-0.0841998	0.0983935
##	z value Pr(> z)	
## (Intercept)	-0.632	0.527278
## Age_at_Release23-27	-1.043	0.296959

## Age_at_Release28-32	-2.190	0.028526	*
## Age_at_Release33-37	-3.826	0.000130	***
## Age_at_Release38-42	-5.337	9.43e-08	***
## Age_at_Release43-47	-7.245	4.33e-13	***
## Age_at_Release48 or older	-8.514	< 2e-16	***
## Residence_PUMA	0.870	0.384475	
## Gang_AffiliatedTRUE	8.083	6.30e-16	***
## Education_LevelHigh School Diploma	1.429	0.153039	
## Education_LevelLess than HS diploma	-1.759	0.078620	.
## Dependents1	1.472	0.140897	
## Dependents2	-0.711	0.477050	
## Dependents3 or more	0.628	0.530315	
## Prison_OffenseOther	2.406	0.016148	*
## Prison_OffenseProperty	2.231	0.025701	*
## Prison_OffenseViolent/Non-Sex	1.413	0.157728	
## Prison_OffenseViolent/Sex	-1.427	0.153463	
## Prison_YearsGreater than 2 to 3 years	-4.043	5.27e-05	***
## Prison_YearsLess than 1 year	2.009	0.044533	*
## Prison_YearsMore than 3 years	-0.271	0.786202	
## Prior_Arrest_Episodes_Felony1	-0.214	0.830612	
## Prior_Arrest_Episodes_Felony10 or more	2.135	0.032771	*
## Prior_Arrest_Episodes_Felony2	0.559	0.576457	
## Prior_Arrest_Episodes_Felony3	0.760	0.447360	
## Prior_Arrest_Episodes_Felony4	0.880	0.378983	
## Prior_Arrest_Episodes_Felony5	1.123	0.261230	
## Prior_Arrest_Episodes_Felony6	1.272	0.203398	
## Prior_Arrest_Episodes_Felony7	0.831	0.405855	
## Prior_Arrest_Episodes_Felony8	1.141	0.253705	
## Prior_Arrest_Episodes_Felony9	1.709	0.087468	.
## Prior_Arrest_Episodes_Misd1	2.377	0.017431	*
## Prior_Arrest_Episodes_Misd2	1.982	0.047424	*
## Prior_Arrest_Episodes_Misd3	2.440	0.014690	*
## Prior_Arrest_Episodes_Misd4	2.400	0.016384	*
## Prior_Arrest_Episodes_Misd5	2.022	0.043175	*
## Prior_Arrest_Episodes_Misd6 or more	3.147	0.001650	**
## Prior_Arrest_Episodes_Violent1	0.316	0.751653	
## Prior_Arrest_Episodes_Violent2	0.412	0.680114	
## Prior_Arrest_Episodes_Violent3 or more	1.215	0.224199	
## Prior_Arrest_Episodes_Property1	0.224	0.822942	
## Prior_Arrest_Episodes_Property2	0.128	0.898444	
## Prior_Arrest_Episodes_Property3	0.206	0.836914	
## Prior_Arrest_Episodes_Property4	0.154	0.877701	
## Prior_Arrest_Episodes_Property5 or more	0.706	0.480478	
## Prior_Arrest_Episodes_Drug1	1.218	0.223051	
## Prior_Arrest_Episodes_Drug2	1.522	0.128056	
## Prior_Arrest_Episodes_Drug3	3.035	0.002405	**
## Prior_Arrest_Episodes_Drug4	1.457	0.145012	
## Prior_Arrest_Episodes_Drug5 or more	1.023	0.306242	
## Prior_Arrest_Episodes_PPViolationCharges1	1.176	0.239696	

## Prior_Arrest_Episodes_PPViolationCharges2	1.620	0.105155	
## Prior_Arrest_Episodes_PPViolationCharges3	2.316	0.020584	*
## Prior_Arrest_Episodes_PPViolationCharges4	2.063	0.039153	*
## Prior_Arrest_Episodes_PPViolationCharges5 or more	2.750	0.005951	**
## Prior_Arrest_Episodes_DVChargesTRUE	0.038	0.969537	
## Prior_Arrest_Episodes_GunChargesTRUE	-0.816	0.414342	
## Prior_Conviction_Episodes_Felony1	1.125	0.260693	
## Prior_Conviction_Episodes_Felony2	1.469	0.141788	
## Prior_Conviction_Episodes_Felony3 or more	0.517	0.605251	
## Prior_Conviction_Episodes_Misd1	0.750	0.452961	
## Prior_Conviction_Episodes_Misd2	2.078	0.037728	*
## Prior_Conviction_Episodes_Misd3	2.024	0.042995	*
## Prior_Conviction_Episodes_Misd4 or more	1.228	0.219345	
## Prior_Conviction_Episodes_ViolTRUE	-0.436	0.662889	
## Prior_Conviction_Episodes_Prop1	-0.294	0.769080	
## Prior_Conviction_Episodes_Prop2	-1.154	0.248528	
## Prior_Conviction_Episodes_Prop3 or more	-0.117	0.906614	
## Prior_Conviction_Episodes_Drug1	-1.009	0.312900	
## Prior_Conviction_Episodes_Drug2 or more	-0.781	0.434888	
## Prior_Conviction_Episodes_PPViolationChargesTRUE	-1.987	0.046962	*
## Prior_Conviction_Episodes_DomesticViolenceChargesTRUE	1.331	0.183044	
## Prior_Conviction_Episodes_GunChargesTRUE	0.774	0.438980	
## Prior_Revocations_ParoleTRUE	2.846	0.004420	**
## Prior_Revocations_ProbationTRUE	-2.156	0.031048	*
## Condition_MH_SATTRUE	3.785	0.000154	***
## Condition_Cog_EdTRUE	0.122	0.902815	
## Condition_OtherTRUE	-1.196	0.231822	
## Violations_ElectronicMonitoringTRUE	4.946	7.57e-07	***
## Violations_InstructionTRUE	2.952	0.003153	**
## Violations_FailToReportTRUE	-0.839	0.401631	
## Violations_MoveWithoutPermissionTRUE	1.386	0.165789	
## Delinquency_Reports1	1.981	0.047587	*
## Delinquency_Reports2	0.469	0.638804	
## Delinquency_Reports3	-1.578	0.114569	
## Delinquency_Reports4 or more	-6.246	4.21e-10	***
## Program_Attendances1	0.892	0.372564	
## Program_Attendances10 or more	-3.222	0.001273	**
## Program_Attendances2	0.079	0.937337	
## Program_Attendances3	-0.057	0.954431	
## Program_Attendances4	0.233	0.816059	
## Program_Attendances5	1.467	0.142269	
## Program_Attendances6	-1.083	0.279003	
## Program_Attendances7	1.042	0.297448	
## Program_Attendances8	0.019	0.984919	
## Program_Attendances9	0.213	0.831444	
## Program_UnexcusedAbsences1	1.427	0.153568	
## Program_UnexcusedAbsences2	2.043	0.041073	*
## Program_UnexcusedAbsences3 or more	0.420	0.674136	
## Residence_Changes1	0.465	0.641714	

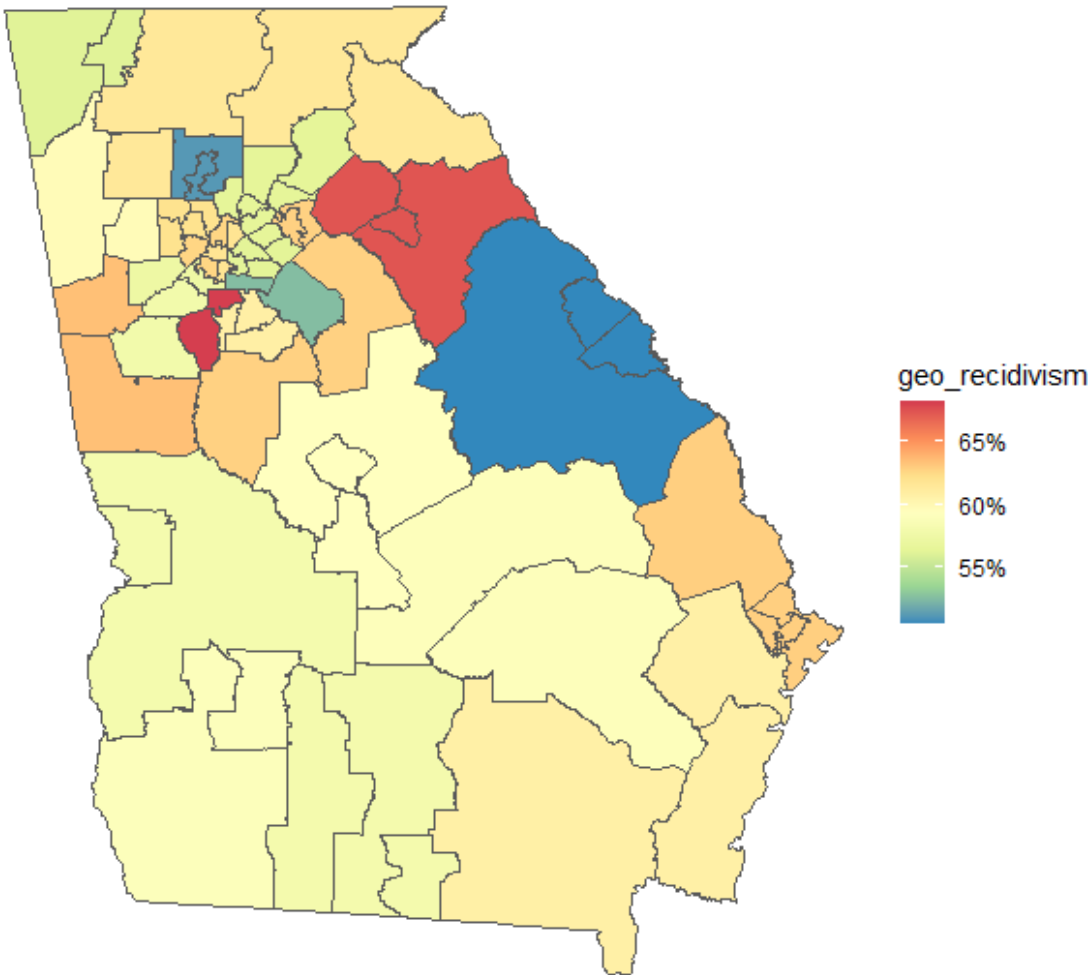
## Residence_Changes2	1.447	0.147875	
## Residence_Changes3 or more	2.860	0.004243	**
## Avg_Days_per_DrugTest	1.019	0.308353	
## DrugTests_THC_Positive	2.574	0.010061	*
## DrugTests_Cocaine_Positive	3.551	0.000384	***
## DrugTests_Meth_Positive	4.442	8.91e-06	***
## DrugTests_Other_Positive	1.590	0.111910	
## Percent_Days_Employed	-19.199	< 2e-16	***
## Jobs_Per_Year	8.761	< 2e-16	***
## Employment_ExemptTRUE	-0.856	0.392139	
## ---			
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			
##			
## (Dispersion parameter for binomial family taken to be 1)			
##			
## Null deviance: 7842.5 on 5769 degrees of freedom			
## Residual deviance: 5957.6 on 5660 degrees of freedom			
## AIC: 6177.6			
##			
## Number of Fisher Scoring iterations: 5			

This interesting result that geographic location plays an important role in determining whether a prisoner recidivates or not motivated us to create a visual mapping the proportion of individuals from each geographic area who recidivated.

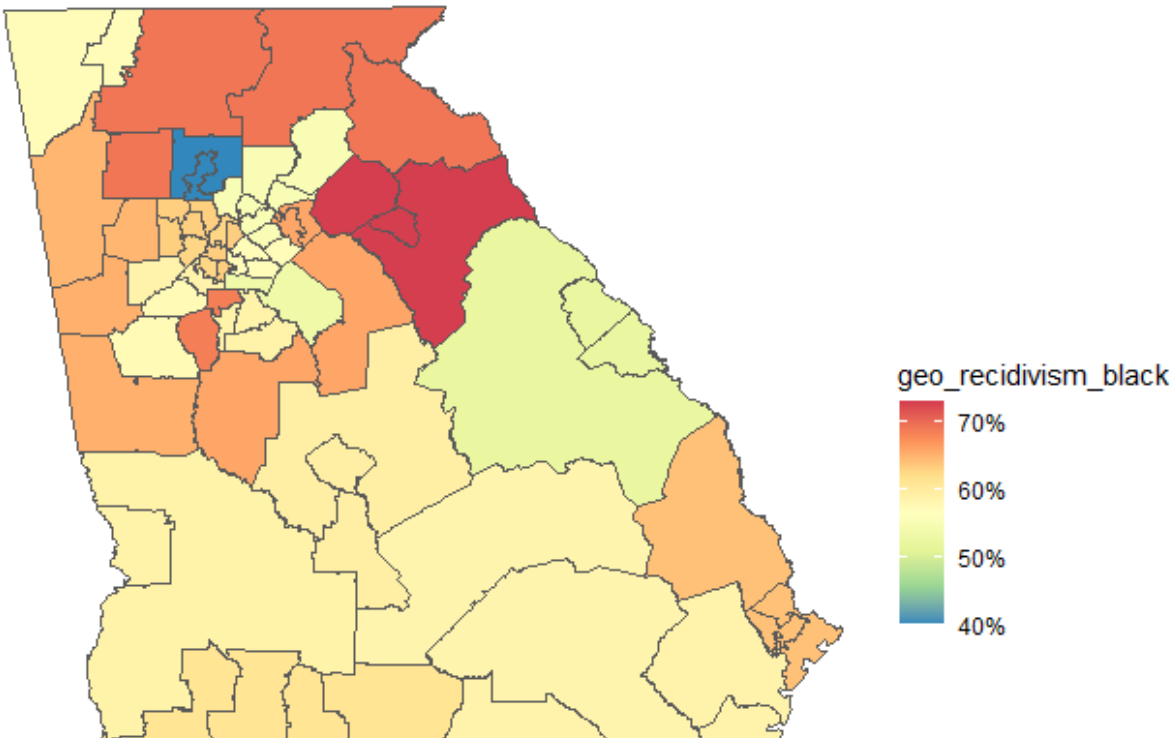
Mapping Recidivism in Georgia

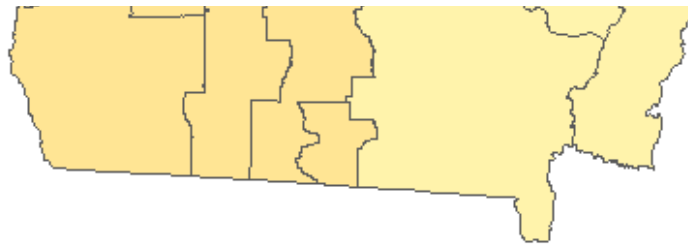
##		
----	--	--

Percentage of all prisoners who recidivated

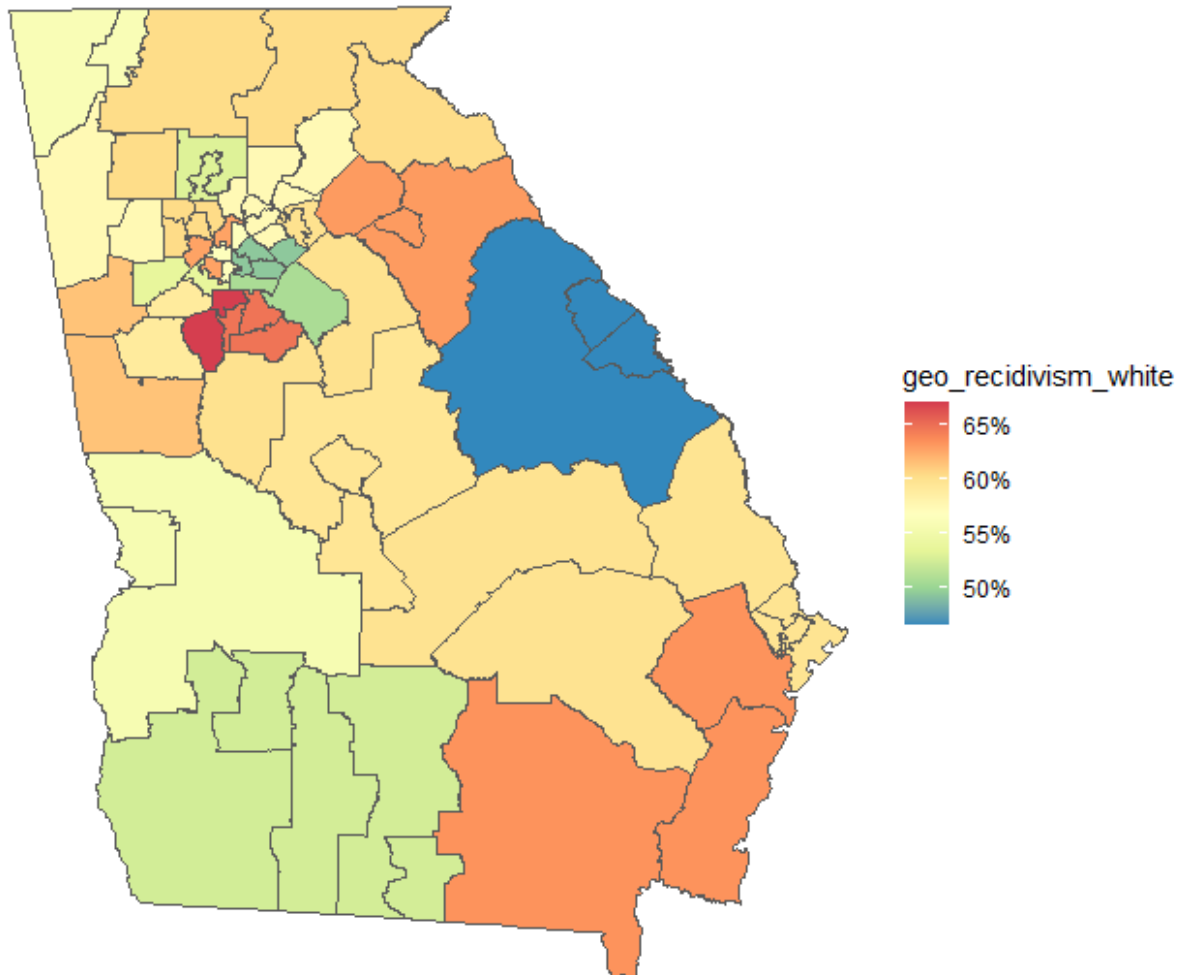


Percentage of black prisoners who recidivated





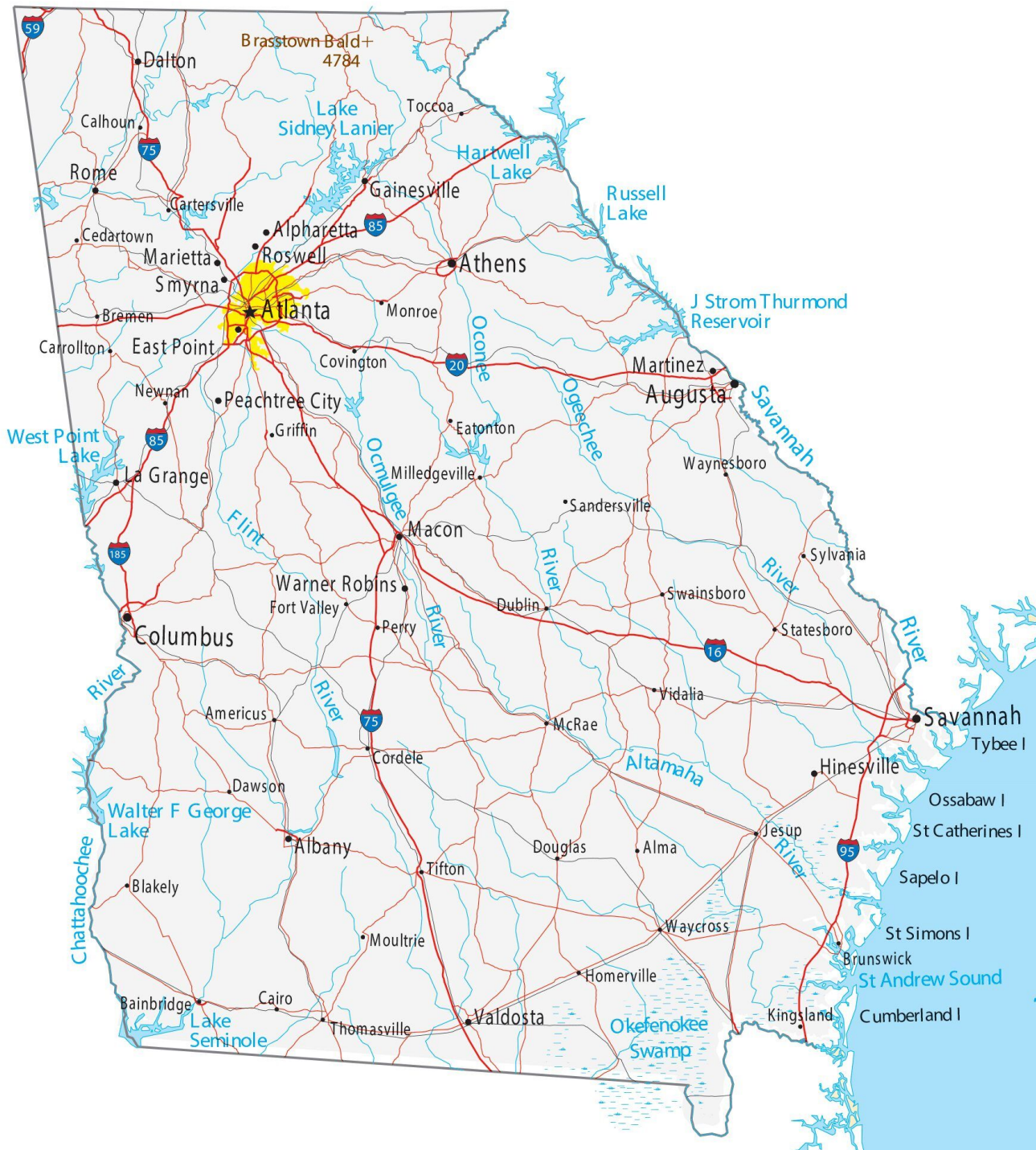
Percentage of white prisoners who recidivated



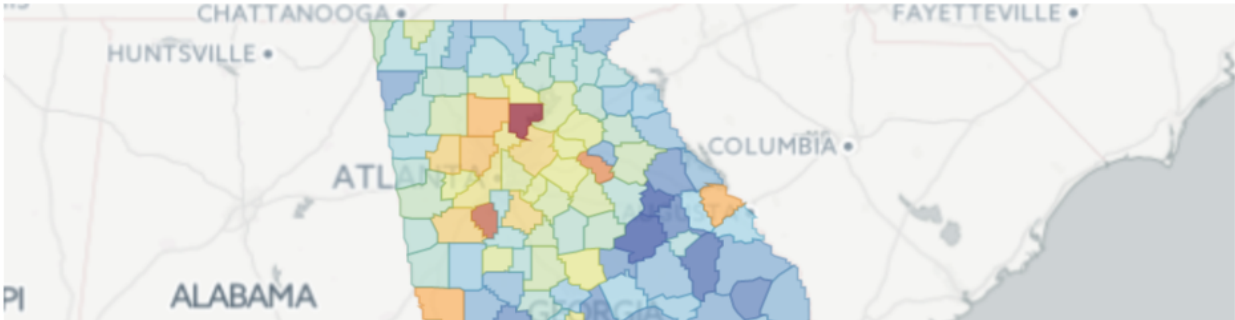
These graphs show the percentage of all, black, and white prisoners from clusters of PUMAs (public use microdata areas) in Georgia who recidivated. The graphs paint a very interesting story. First, we saw that there was a much wider range of recidivism rates for black prisoners depending on where they lived after release. There was a small area in the northwest part of the state where recidivism rates were 40%, while in the northeast, rates reached as high as 72%. This was significantly larger than the range of rates when looking at all and white prisoners who respectively ranged from 50%-68% and 46%-67%. This wider range likely explains why the Residence_PUMA variable was more significant for

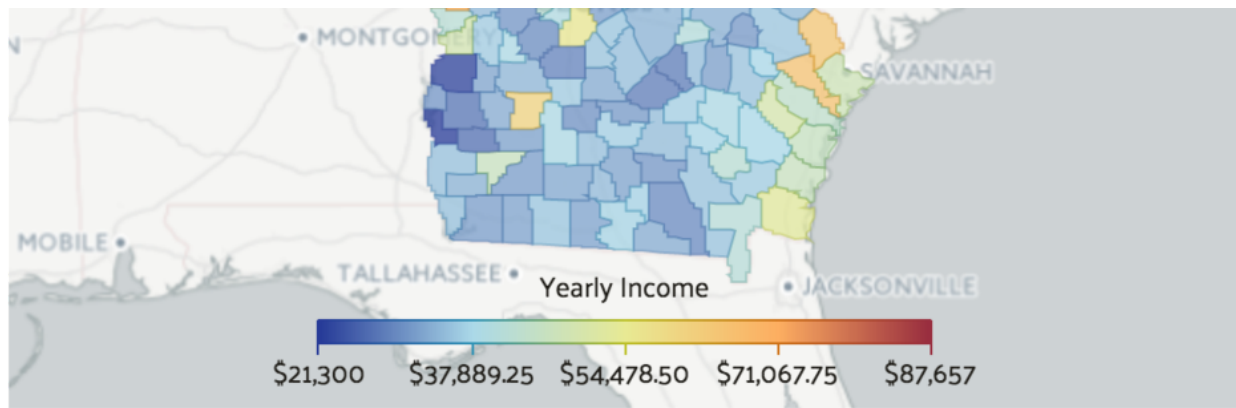
black prisoners. We could also see from the graphs that recidivism rates overall were higher for black prisoners than white prisoners.

Next, looking at the specific PUMA clusters, we could see that for the most part the areas with high recidivism rates were consistent for both black and white prisoners. The main difference being that the rates of recidivism for black prisoners are pretty much higher across the board. With little familiarity with counties in Georgia, we were curious to compare these graphs with a regular map of Georgia.



Income by Location in Georgia





Dataset: 2014 ACS 5-year Estimate
Source: Census Bureau

DATAUSA

From these maps we could see that the area around Athens, Georgia had the highest rates of recidivism while Augusta, Georgia had the lowest rates. Looking at the median household income of the different areas in 2014, it's interesting that there didn't really appear to be a huge connection between lower income and higher recidivism. What was even more notable is that Athens is about 63% white and 28% black, while Augusta is 57% black and 36% white (worldpopulationreview.com). These findings suggest that black prisoners are more successful at desisting when living among a community with a higher black population. It also shows that black prisoners had a harder time desisting when living in an area like Athens that had a higher white population. The reasons for these findings could be numerous and much more research needs to be done in this area to confirm and build upon these findings. However, one potential hypothesis could be that black people living in an area with a higher white population might stick out more and be subject to more racial biases and as a result be more prone to being re-arrested.

Conclusion

Our project shed light on the important variables in predicting recidivism. Although the accuracy of none of models were impressively high, the closer analysis of the significant variables can be useful to promote higher rates of desistance. We went into the project with the ambition of describing potential programs or areas worth investing man power or capital in to lower rates of recidivism. We came away with the prominent findings that employment, drug use, and geographic location are among the most significant variables in predicting recidivism. Rather than use a predictive model such as the one we created to closely monitor individuals with higher chances of recidivism, states and counties ought to consider the key variables we found when predicting recidivism in order to promote desistance. For employment, perhaps individuals newly-released from prison should receive some sort of

capital or social job opportunity. For drug use, individuals should be placed in rehab programs and supported in their battle with drugs. Considering the incarceration system holistically, we found distinctions between white and black prisoners. Racial equality will continue to remain an important ambition of states and societies, and that especially applies to prisons. For geographic location, perhaps the state can work toward equality of opportunity across communities. Working with, rather than against, newly-released prisoners can facilitate a positive feedback loop. Promoting the good rather than preparing for the bad is the solution to lower recidivism rates.

Works Cited

“The Problem: Recidivism & Mass Incarceration.” Prison Scholar Fund, 20 July 2019, <https://www.prisonerscholars.org/what-we-do/222-2/>.

“Recidivism.” National Institute of Justice, U.S. Department of Justice, 31 July 2019, <https://nij.ojp.gov/topics/corrections/recidivism>.

Wildeman, Christopher. “The Impact of Incarceration on the Desistance Process among Individuals Who Chronically Engage in Criminal Activity.” National Institute of Justice, 14 Oct. 2021, <https://nij.ojp.gov/topics/articles/impact-incarceration-desistance-process-among-individuals-who-chronically-engage>.

[https://en.wikipedia.org/wiki/Demographics_of_Georgia_\(U.S._state\)](https://en.wikipedia.org/wiki/Demographics_of_Georgia_(U.S._state))

<https://worldpopulationreview.com/us-cities/athens-ga-population>

<https://worldpopulationreview.com/us-cities/augusta-ga-population>

Link to data

<https://data.ojp.usdoj.gov/stories/s/daxx-hznc>

Appendices: <https://nij.ojp.gov/funding/recidivism-forecasting-challenge#appendices>

Appendix 2 was used to understand the variables and Appendix 3 was used to create the map visualizations