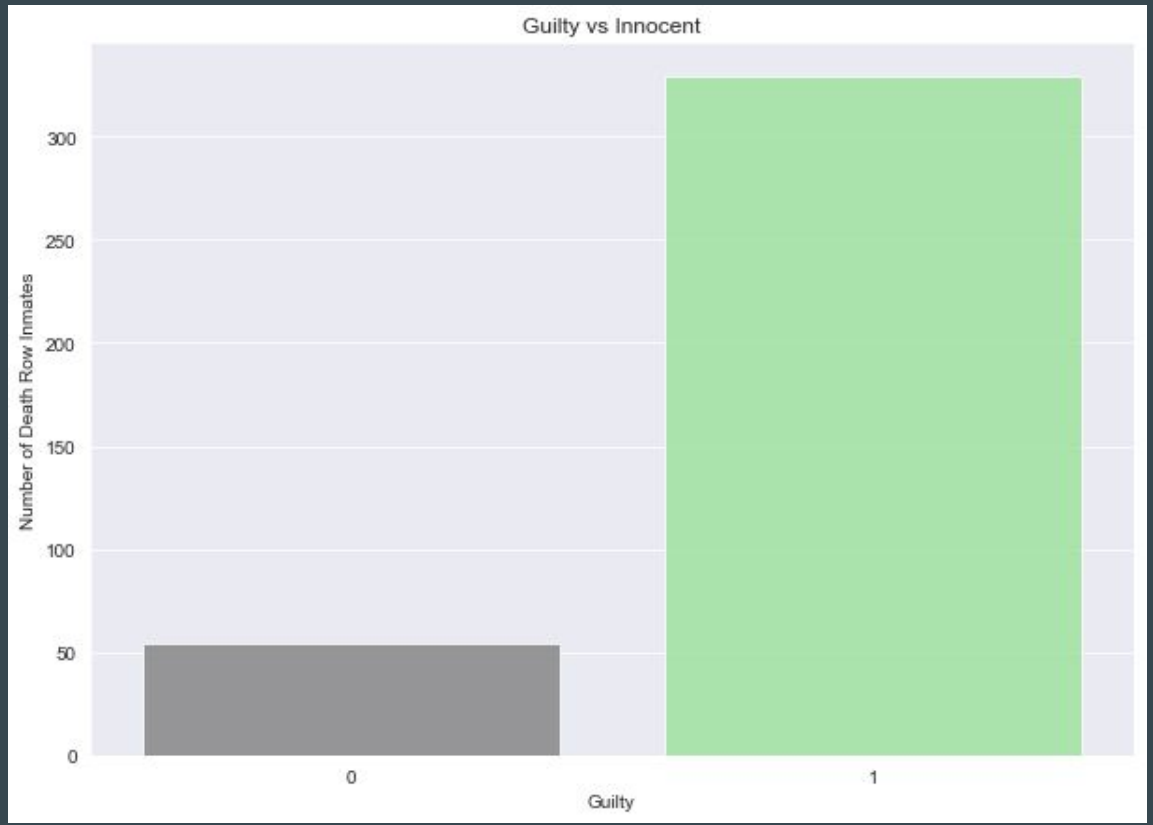
Death Row Inmates' Last Words

"I want the victim's family to know that I didn't commit this crime. I didn't kill your loved one. Sharon Wilson, y'all convicted an innocent man and you know it. There are some lawyers hired that is gonna prove that, and I hope you can live with it. To my family and loved ones, I love you. Thank you for supporting me. Y'all stay strong. Warden, take me home." - Richard Jones, 40 years old

By: Julia Chong

Declaring Innocence



- Innocent
- Did not do
- Truth
- Not guilty
- Didn’t kill
- Did not kill
- Did not murder
- Never killed
- Innocence

Precision and Recall

Decision Tree

	Precision	Recall	F1 Score
Innocent	0.50	0.17	0.25
Guilty	0.89	0.98	0.93

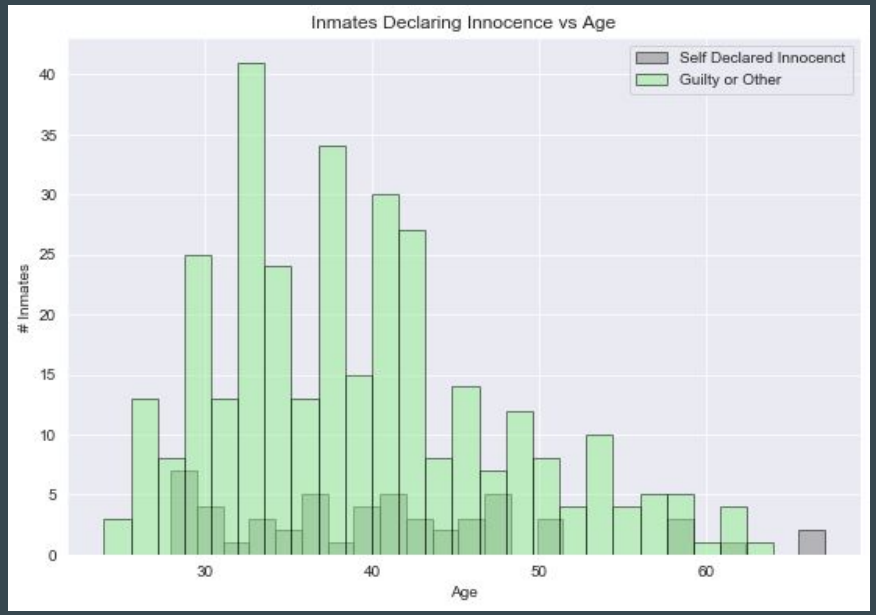
ADA Boost

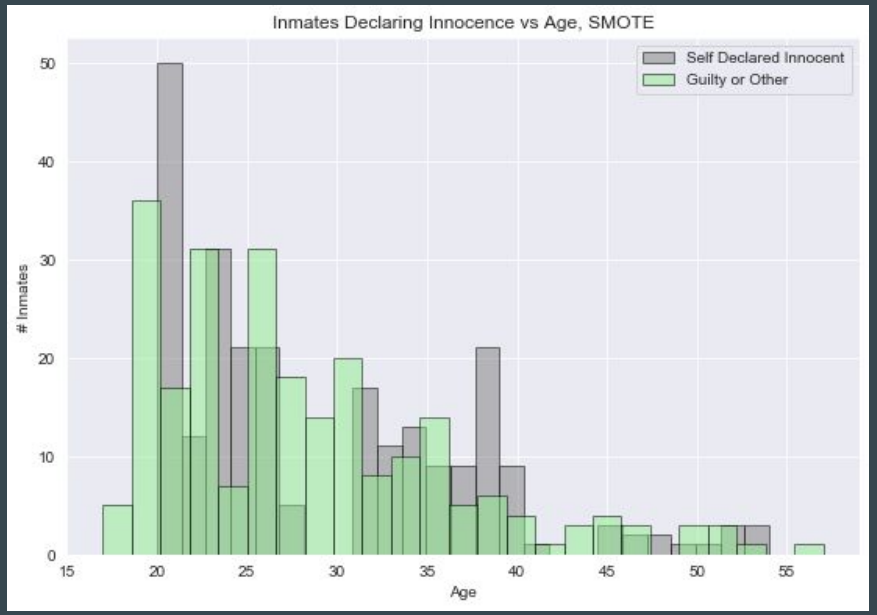
	Precision	Recall	F1 Score
Innocent	0.50	0.08	0.14
Guilty	0.90	0.99	0.94

K Nearest Neighbors

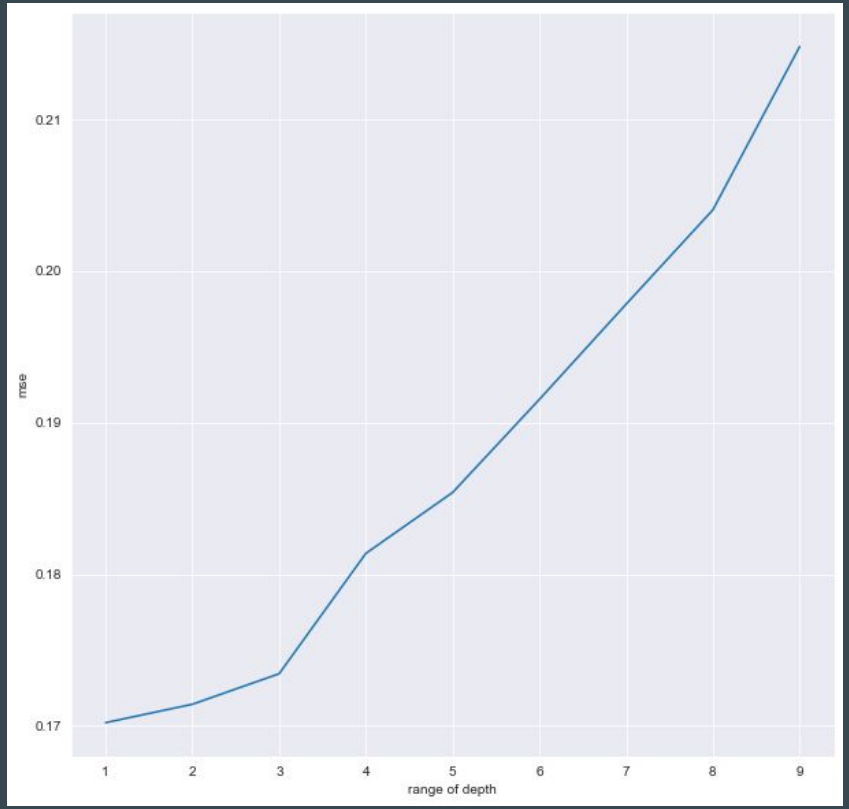
Precision	Recall	F1 Score	Accuracy
0.89583	0.89583	0.8615	0.89583

Before and After SMOTE

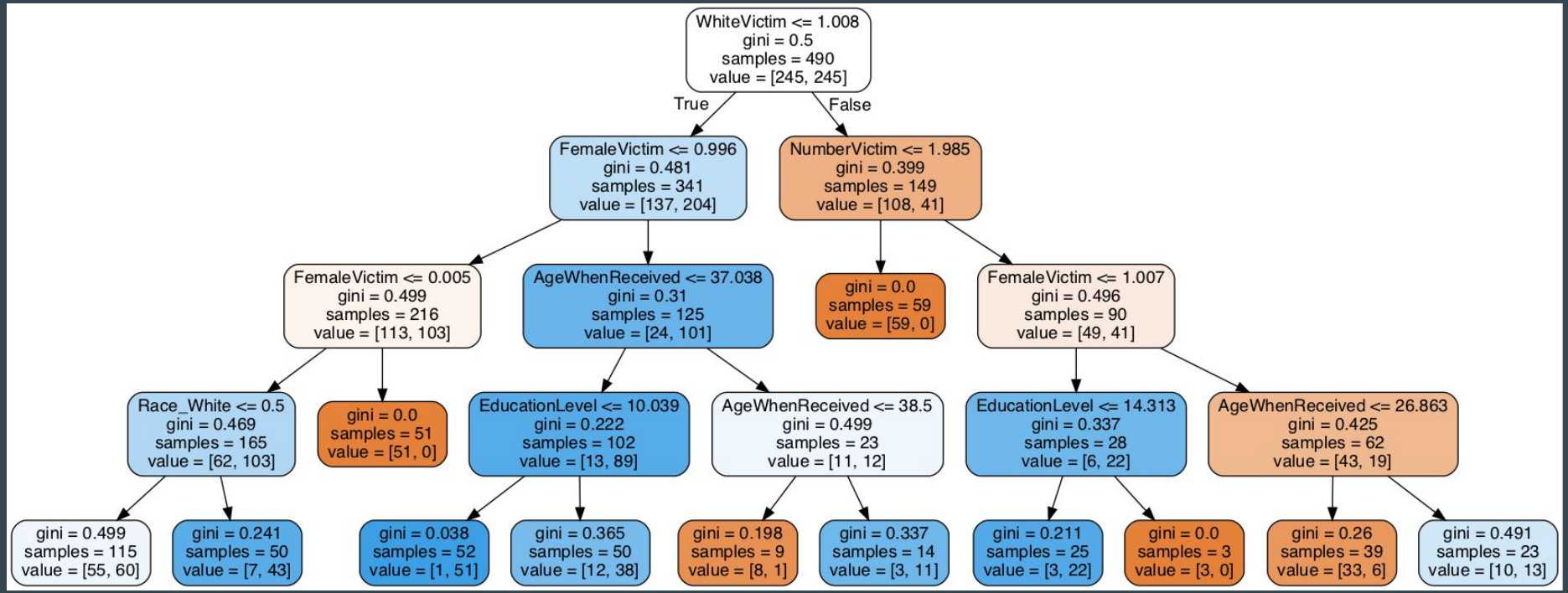




Validation Curve



Pruned & Tuned Decision Tree



Strongest Predictors of Innocence

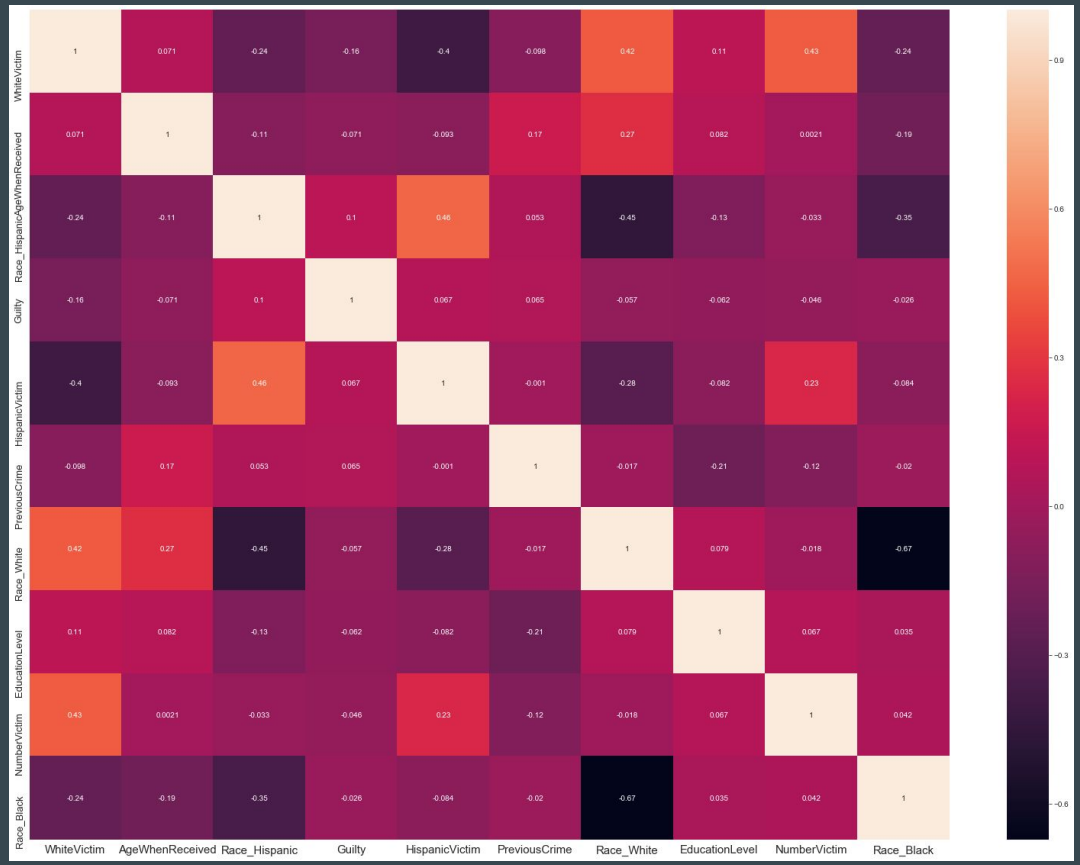
Correlation Matrix

White Victim	-0.156095
Age	-0.070505
Native County	-0.066625
Years of Education	-0.062124
White Inmate	-0.056575
Number of Victims	-0.046472
Black Inmate	-0.026476

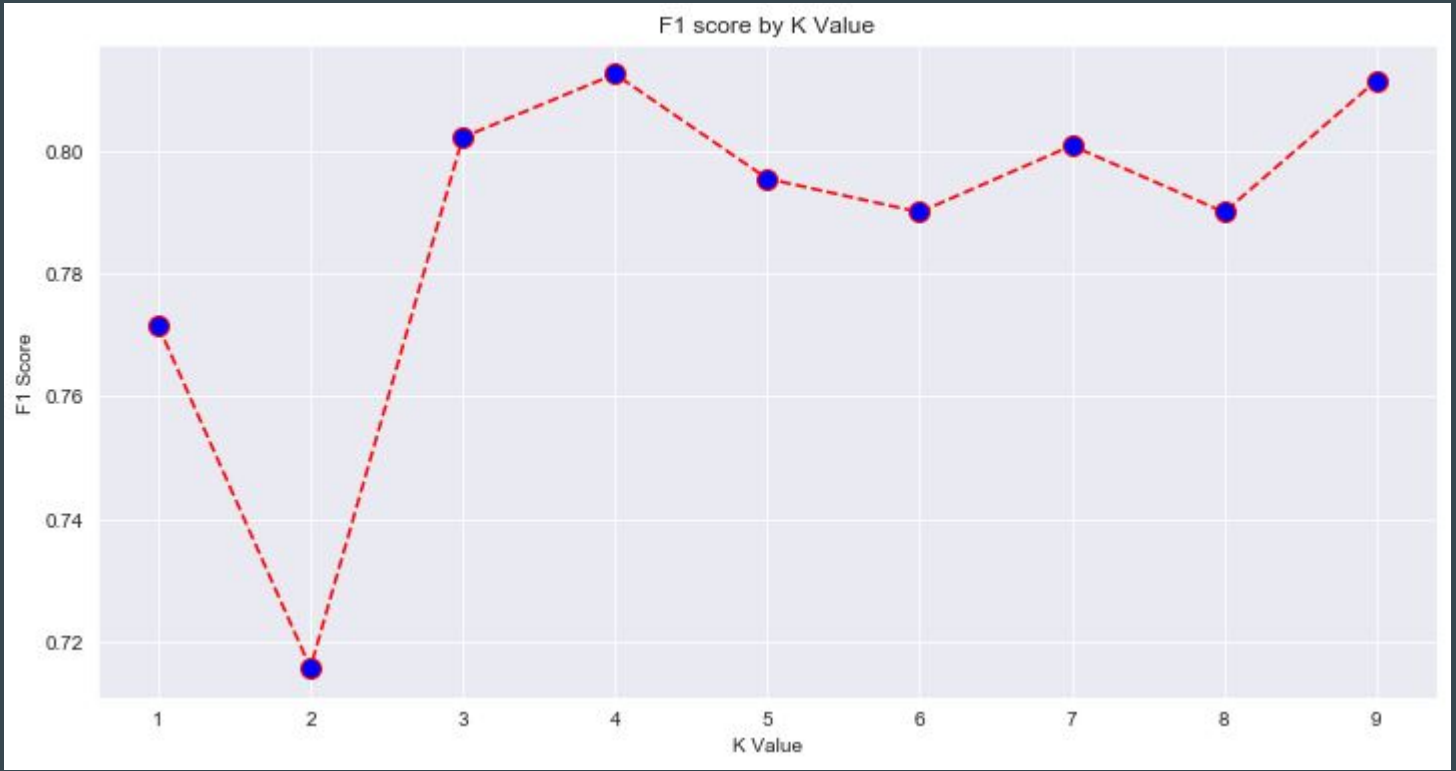
Feature Importance

White Victim	7.213729
Age	4.459612
Hispanic Inmate	3.168072
Hispanic Victim	2.528992
Native County	1.381020
Previous Crime	0.806269
White Inmate	0.652958

Feature Importance Heat Map



K Nearest Neighbors



Conclusion

- KNN and Decision Tree > ADA Boost
- White Victim was the strongest indicator among all tests
- Race of Victim and Race of Inmate were both the strongest indicators

Future Work

- Try classification using another variable, such as White vs Non White
- Do further EDA to drop more variables and improve accuracy
- Use the Z Score for non-dummy variables to reduce variance within model