How we decide the model

Table 1. VIF Values

	feature	VIF
0	population_group_code	3.542983
1	offense_code	4.057491
2	offender_race	13.655167
3	offender_ethnicity	14.380545
4	offender_age	2.217070
5	offender_sex	14.020552
6	victim_type_code	3.816982
7	location_code	5.148715
8	prop_desc_code	2.538717
9	stolen_value	1.000263

Table 2. Logit Regression Model

```
Optimization terminated successfully.
        Current function value: 0.377012
        Iterations 6
```

Logit Regression Results

Dep. Variable:	recovered_flag	No. Observations:	92217
Model:	Logit	Df Residuals:	92206
Method:	MLE	Df Model:	10
Date:	Thu, 09 Feb 2023	Pseudo R-squ.:	0.07130
Time:	16:49:04	Log-Likelihood:	-34767.
converged:	True	LL-Null:	-37436.
Covariance Type:	nonrobust	LLR p-value:	0.000

	coef	std err	z	P>|z|	[0.025	0.975]
population_group_code	-0.0121	0.002	-5.724	0.000	-0.016	-0.008
offense_code	0.0620	0.003	24.659	0.000	0.057	0.067
offender_race	0.0187	0.009	2.136	0.033	0.002	0.036
offender_ethnicity	-0.3449	0.013	-25.602	0.000	-0.371	-0.319
offender_age	0.0065	0.001	9.813	0.000	0.005	0.008
offender_sex	-0.3664	0.020	-18.439	0.000	-0.405	-0.327
victim_type_code	-0.0344	0.007	-5.178	0.000	-0.047	-0.021
location_code	-0.0015	0.001	-1.383	0.167	-0.004	0.001
prop_desc_code	-0.0049	0.000	-13.194	0.000	-0.006	-0.004
stolen_value	4.621e-09	5.17e-09	0.894	0.371	-5.51e-09	1.48e-08
intercept	-0.8176	0.060	-13.702	0.000	-0.935	-0.701

Table 3. GLM Model

Generalized Linear Model Regression Results

Dep. Variable:	recovered_flag	No. Observations:	92217
Model:	GLM	Df Residuals:	92206
Model Family:	Binomial	Df Model:	10
Link Function:	Logit	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-34767.
Date:	Thu, 09 Feb 2023	Deviance:	69534.
Time:	16:49:05	Pearson chi2:	8.90e+04
No. Iterations:	5	Pseudo R-squ. (CS):	0.05625
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
population_group_code	-0.0121	0.002	-5.724	0.000	-0.016	-0.008
offense_code	0.0620	0.003	24.659	0.000	0.057	0.067
offender_race	0.0187	0.009	2.136	0.033	0.002	0.036
offender_ethnicity	-0.3449	0.013	-25.602	0.000	-0.371	-0.319
offender_age	0.0065	0.001	9.813	0.000	0.005	0.008
offender_sex	-0.3664	0.020	-18.439	0.000	-0.405	-0.327
victim_type_code	-0.0344	0.007	-5.178	0.000	-0.047	-0.021
location_code	-0.0015	0.001	-1.383	0.167	-0.004	0.001
prop_desc_code	-0.0049	0.000	-13.194	0.000	-0.006	-0.004
stolen_value	4.621e-09	5.17e-09	0.894	0.371	-5.51e-09	1.48e-08
intercept	-0.8176	0.060	-13.702	0.000	-0.935	-0.701

Table 4. Odds Ratio Values for Variables

	coefficients	p-value	odds_ratio
const	-8.165584e-01	2.021458e-45	0.441950
population_group_code	-1.466845e-02	8.167097e-13	0.985439
offense_code	6.359470e-02	7.985539e-150	1.065660
offender_race	1.848176e-02	2.911607e-02	1.018654
offender_ethnicity	-3.475229e-01	2.142245e-156	0.706436
offender_age	7.303784e-03	3.161318e-30	1.007331
offender_sex	-3.544591e-01	2.239136e-75	0.701553
victim_type_code	-3.749586e-02	6.468313e-09	0.963198
location_code	-2.616511e-03	1.495732e-02	0.997387
prop_desc_code	-4.629227e-03	1.062554e-37	0.995381
stolen_value	4.396706e-09	3.828072e-01	1.000000