

Table of Contents

- [\[1\] Data Sources \(pg. 1\)](#)
- [\[2\] Prototyping Maps \(pg. 1\)](#)
- [\[3\] Scatter plots and statistics \(pg. 5\)](#)
- [\[4\] Other Areas Worth Researching \(pg. 7\)](#)

Data Sources

[Github repo w/ Jupyter Notebooks + data](#)

Police Employee Data ([via 2018 FBI UCR](#))

- “Per_1k” field is number of officers in city’s police department per 1,000 city residents

Known offenses to law enforcement ([via 2018 FBI UCR](#))

- [By county](#)
- [By city](#)

Racial makeup data ([via 2018 U.S. Census population estimates](#))

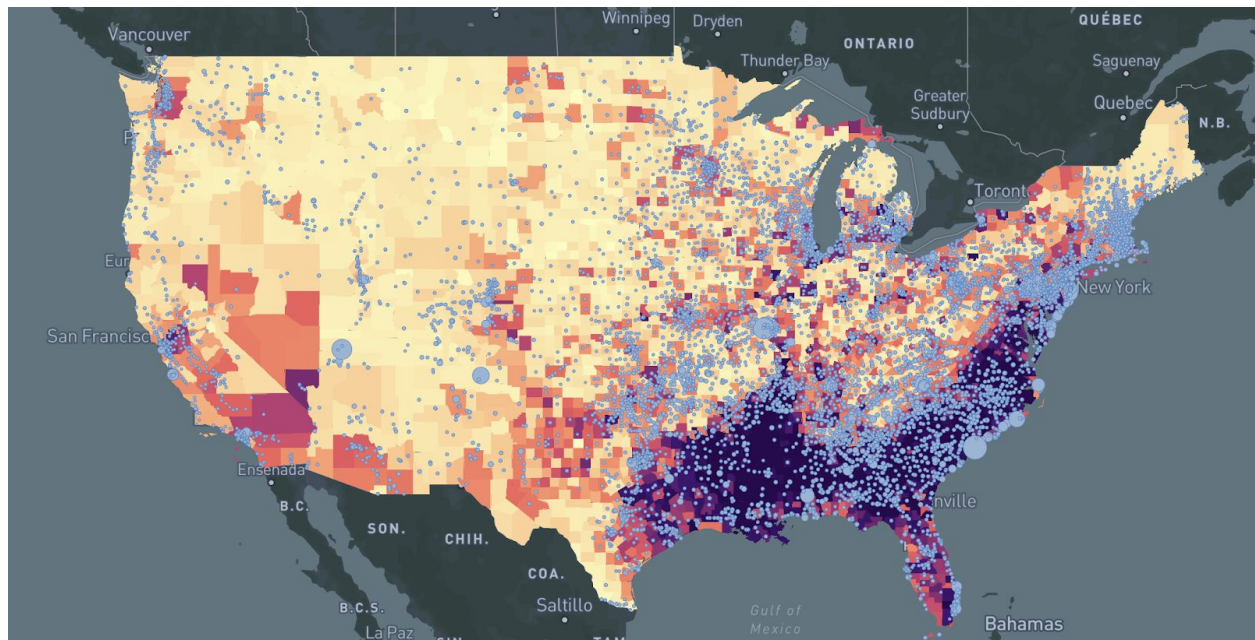
- Census data is recorded by “Metropolitan Statistical Areas” (eg. San Francisco-Oakland-Hayward); joined to FBI data by first city listed

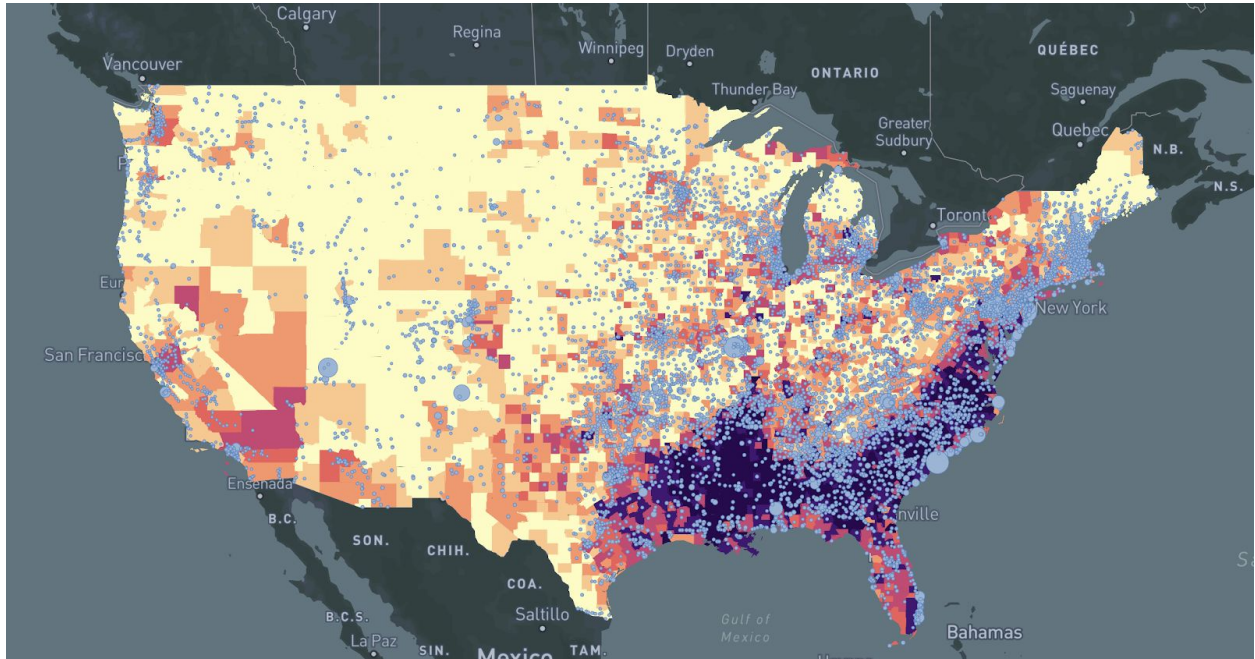
Median income (via 2018 U.S. Census estimates)

- [By county](#)
- [By MSA](#)

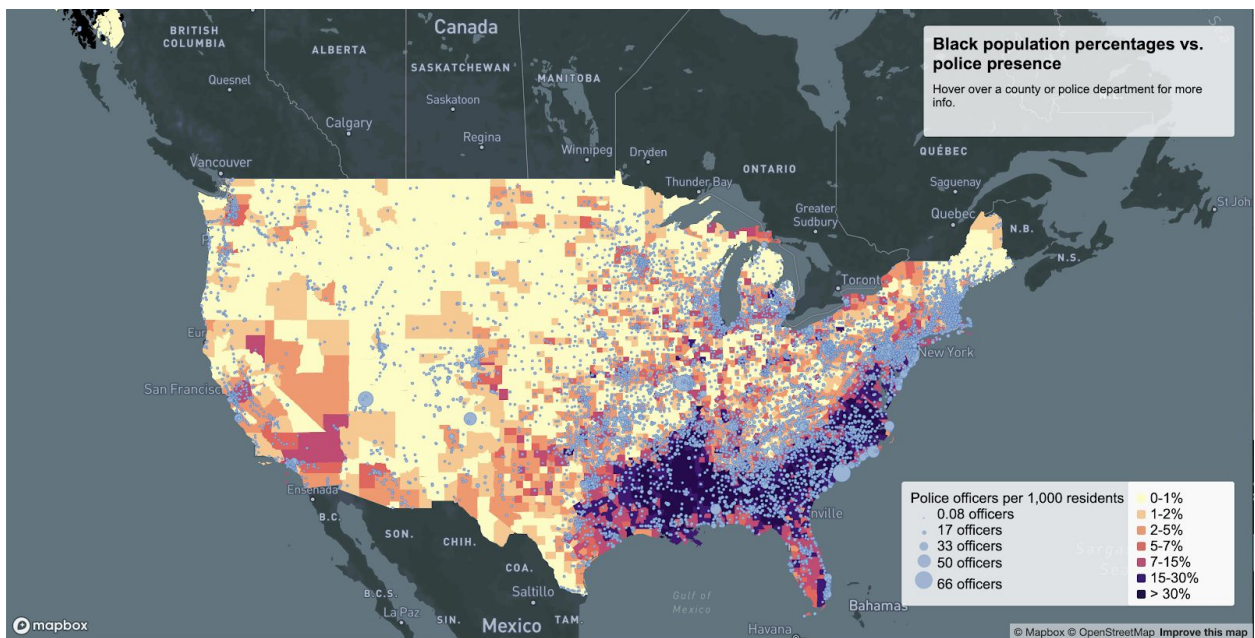
Prototyping Maps

By Race

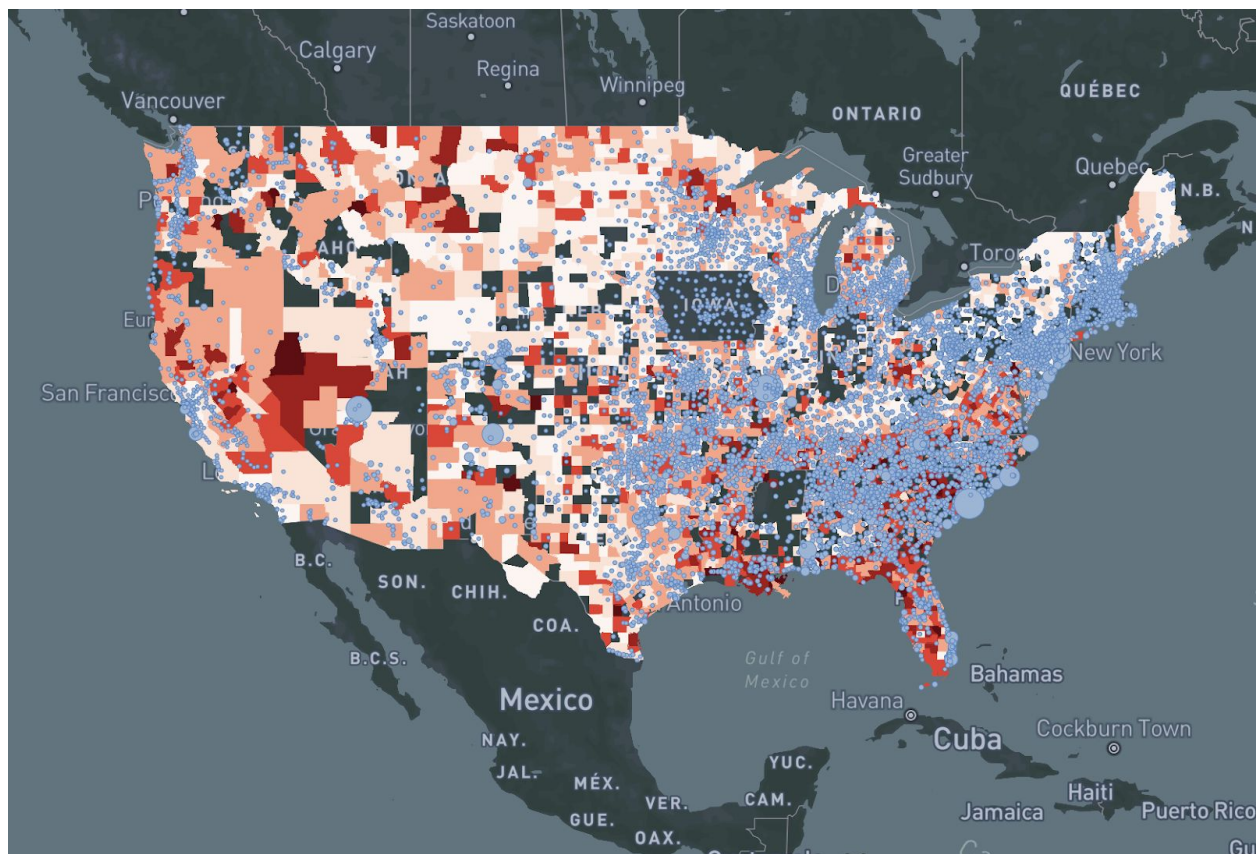
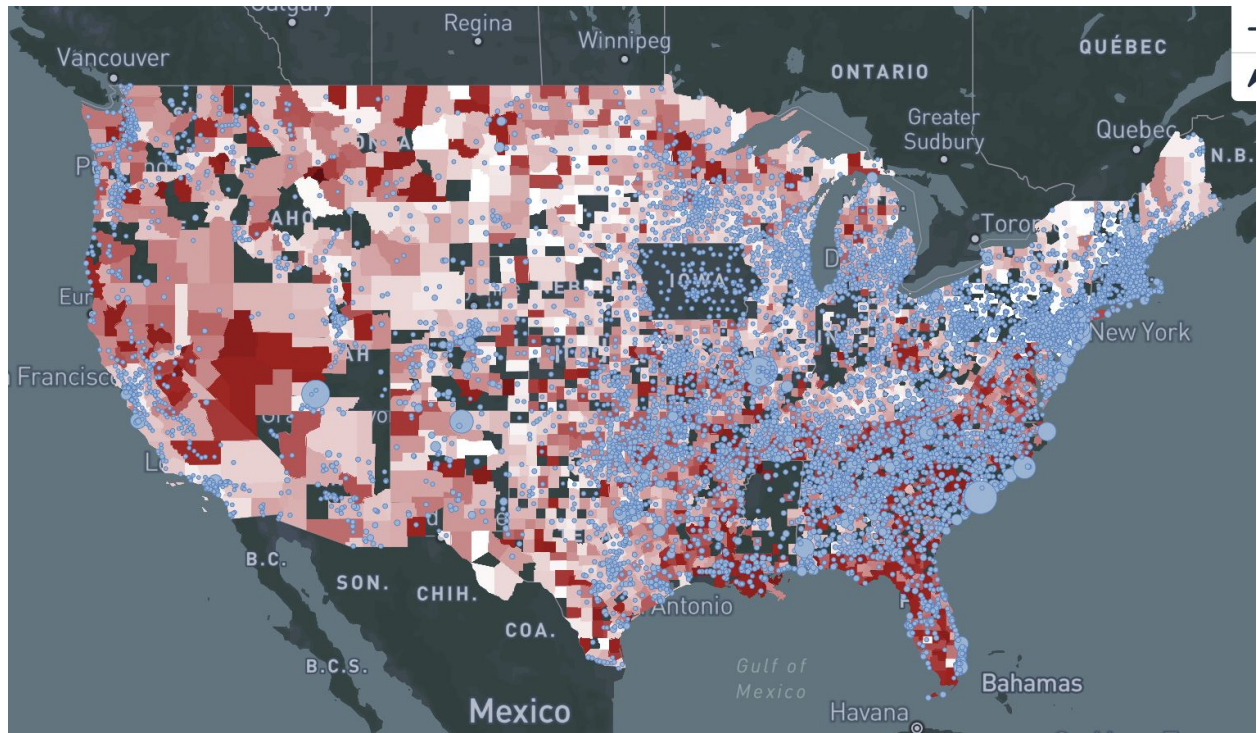




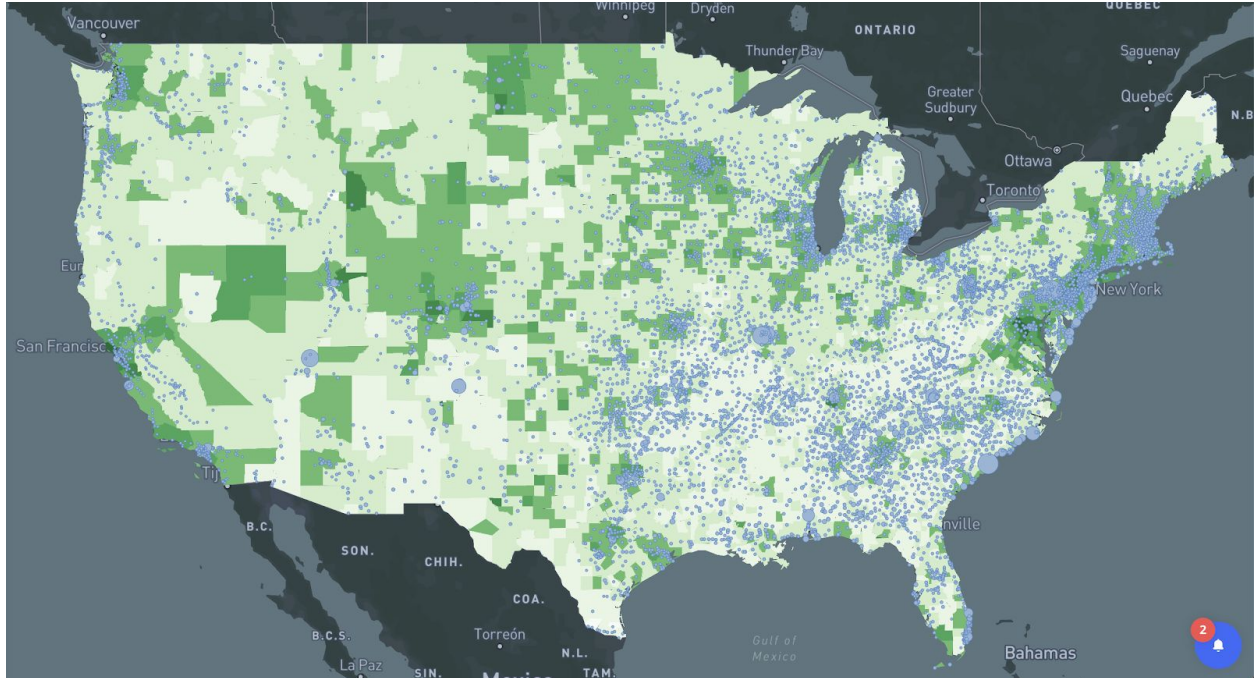
Final race:



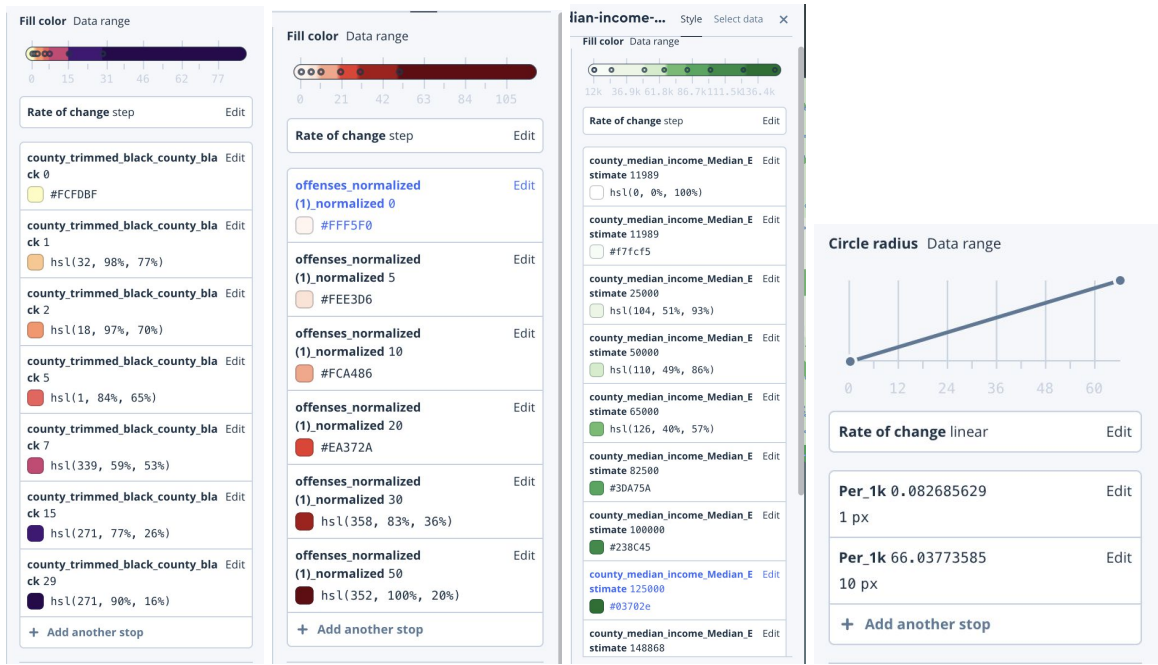
By Crime Rate



Income

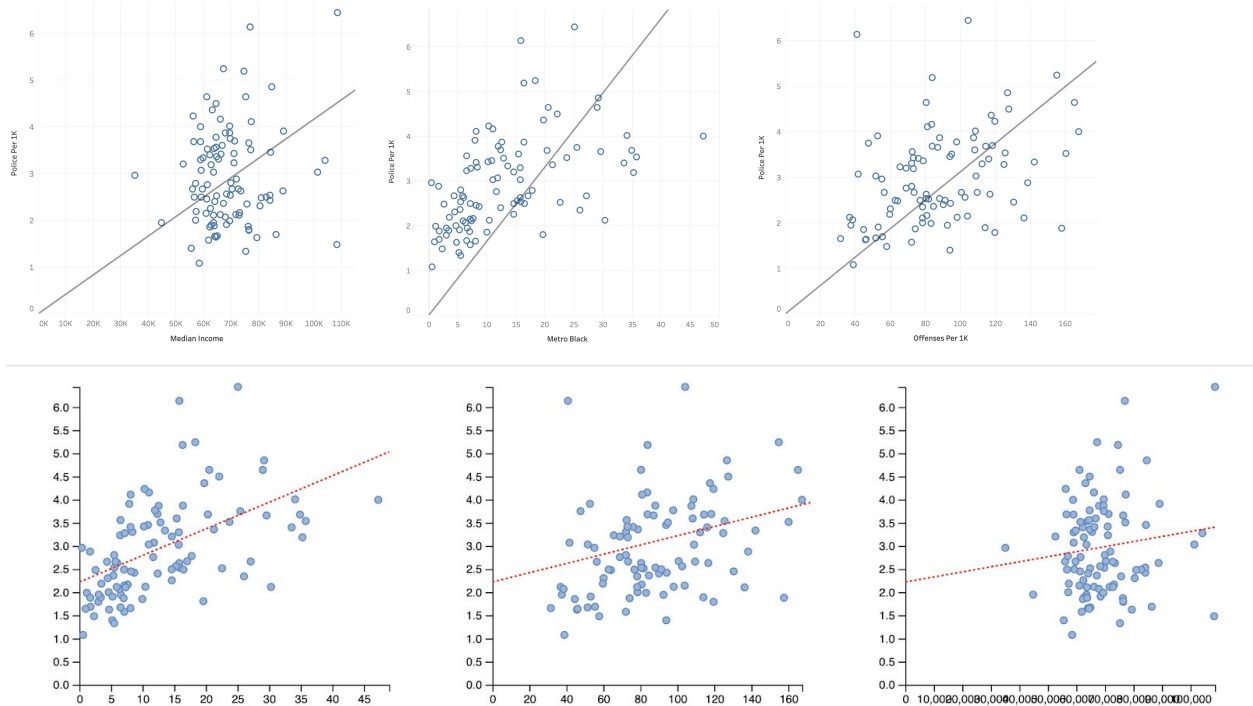


Legends

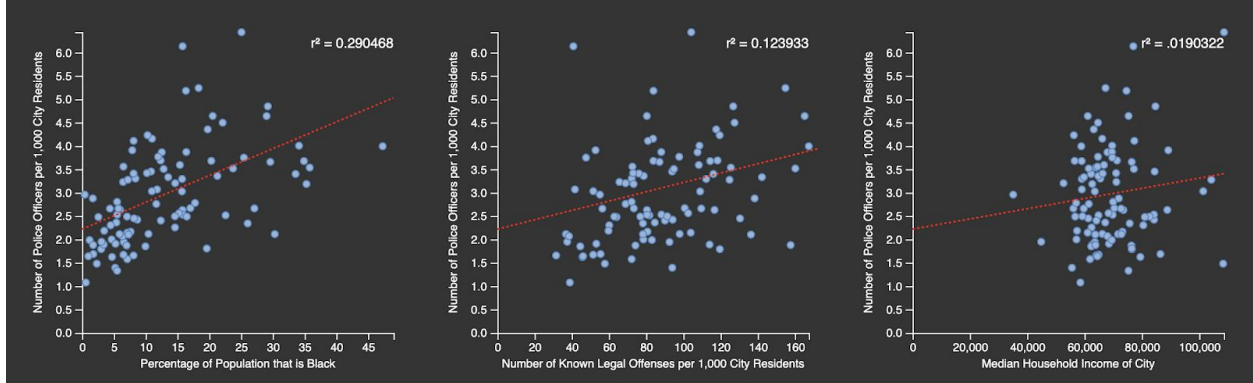


Scatter plots and statistics (Data limited to 100 most populous cities)

Simple OLS regression:

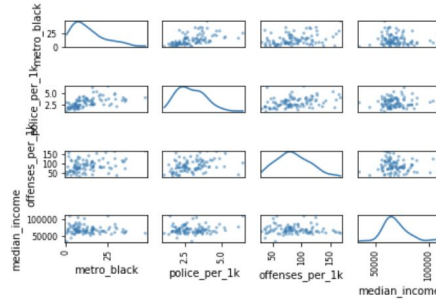
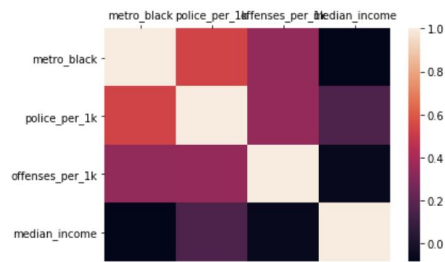


Correlation between Police Presence and Race, Crime rates, Income



Calculating Correlation between Race, Income, Known Offenses, Police Presence in 100 most populous cities

Drop null values -- dataset has 94 entries.



	Police Officers	Black Population	Known Offenses	Median Income
Police Officers	1	0.54	0.35	0.15
Black Population		1	0.35	-0.08
Known Offenses			1	-0.06
Median Income				1

OLS Multiple Regression:

OLS Regression Results					
Dep. Variable:	police_per_1k	R-squared:	0.930		
Model:	OLS	Adj. R-squared:	0.927		
Method:	Least Squares	F-statistic:	401.6		
Date:	Thu, 04 Jun 2020	Prob (F-statistic):	2.43e-52		
Time:	10:37:20	Log-Likelihood:	-116.10		
No. Observations:	94	AIC:	238.2		
Df Residuals:	91	BIC:	245.8		
Df Model:	3				
Covariance Type:	nonrobust				
	coef	std err	t	P> t	[0.025 0.975]
metro_black	0.0529	0.010	5.570	0.000	0.034 0.072
offenses_per_1k	0.0074	0.003	2.736	0.007	0.002 0.013
median_income	2.339e-05	3.27e-06	7.160	0.000	1.69e-05 2.99e-05
Omnibus:	15.686	Durbin-Watson:	1.562		
Prob(Omnibus):	0.000	Jarque-Bera (JB):	18.153		
Skew:	0.911	Prob(JB):	0.000114		
Kurtosis:	4.146	Cond. No.	7.70e+03		

Correlation matrix:

	metro_black	police_per_1k	offenses_per_1k	median_income
metro_black	1.000000	0.538951	0.347633	-0.081102
police_per_1k	0.538951	1.000000	0.352041	0.152693
offenses_per_1k	0.347633	0.352041	1.000000	-0.057537
median_income	-0.081102	0.152693	-0.057537	1.000000

Race, number of known offenses and income are all correlated with heavier police presence, but the percentage of population that is Black is by far the most statistically significant variable.

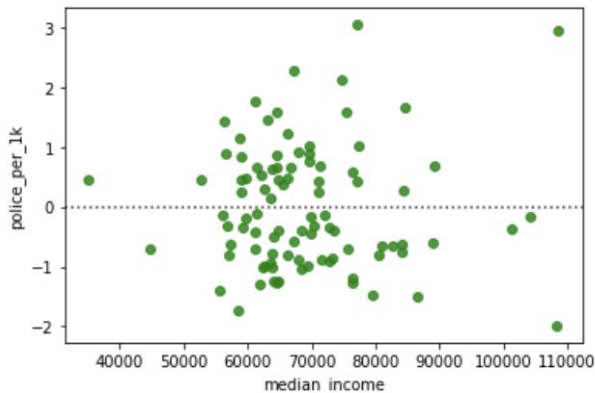
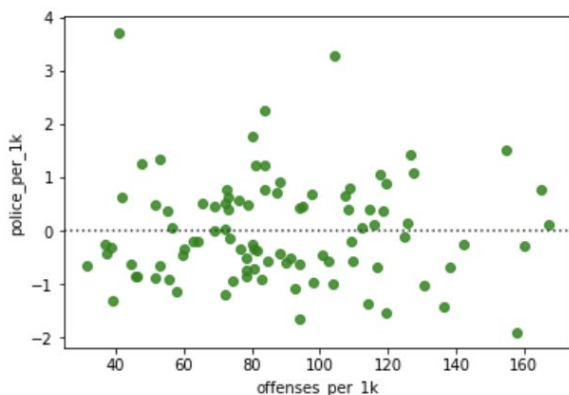
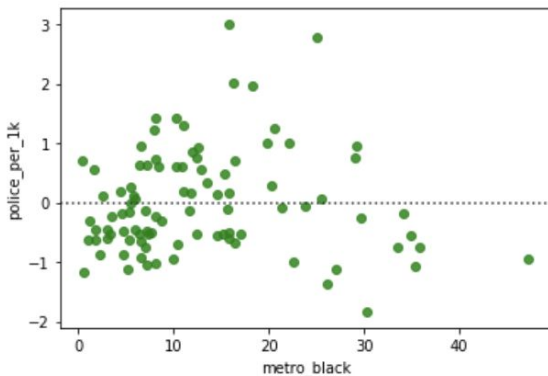
I also tested for multicollinearity. The variable importance factor (VIF) value for each of the variables was 1.142, 1.138 and 1.008 respectively — which is to say, some degree of collinearity exists between the variables, but nothing to be concerned about.

```
> vif(testlm)
```

metro_black	offenses_per_1k	median_income
1.142328	1.138583	1.007614

```
> |
```

Residuals:



Other Areas Worth Researching

- Analyzing use of force
 - The collection of law enforcement [use of force statistics](#) has been mandated as a responsibility of the Attorney General since the passage of the Violent Crime Control and Law Enforcement Act of 1994 (Biden) — worth seeing how these statistics have been used by social science research institutions and/or accountability orgs
- How many cops are charged, convicted, for on-duty offenses?
- Visualizing outcome of high-profile arrests/lawsuits regarding use of excessive force by police