Data Science Coursework

Kyriacos Xanthos
CID: 01389741

May 1, 2022

Abstract

In this report we dive into the US Communities and Crime Dataset to try and find meaningful predictors that increase violent crime in a community in the United States. We perform a multiverse analysis to analyse different partitions of the dataset in order to explore the inherent biases that the available dataset includes. We build models that predict the violence in a community with an R^2 score of up to 0.78, and look closely at the predictors that influenced that score. We explore how different states in the US have different predictors and how states like California are overrepresented in the dataset, influencing our results. We reserve judgment to make any final statements on exactly what factors affect crime in a community because we believe the bias in the dataset does not allow us to give a reliable conclusion.

1 Introduction

The BLM (Black Lives Matter) protests began in 2013 and have been going on to present day. The BLM movement has the objective to highlight racism, discrimination and the inequality felt by black people in the world [1]. The protests escalated after George Floyd's brutal murder by the Minneapolis police officer and that sparked the movement on an international level, against police brutality towards minorities. One of the key points that were vocal in the movement is that there is a strong bias from the police on black people and that there is an enormous discrimination when associating minorities with crime in the US. In this report we will be analyzing if this bias is well founded through the lens of data.

We are using the US Communities and Crime Data Set [2] [3] [4]. This dataset includes socio-economic data from the 1990 US Census [5], where sample data were weighted to represent the whole population of the US. This is combined with the 1990 US LEMAS survey from law enforcement [6] which is a survey that includes agency personnel, expenditures and pay, written policies and many more. Lastly, the Crime in the United States dataset [4] is included that represented almost 251 million US inhabitants with all the crime reports nationwide. Our dataset only includes violent crimes in the US: murder, rape, robbery and assault.

Our objective is to understand what factors affect the crime rate in a community in the US. We will use the total number of violent crimes per 100,000 population as our target and see how the rest of the variables act as predictors for the Violent Crimes. The diversity of our predictors allows us to check multiple inherent biases towards minorities in the US and understand how they actually affect the crime rate in many communities. We can consider variables like race, immigrants in a community, income, family structure, police personnel data and many more. Understanding the key predictors for violent crimes will mean that we need to understand the inherent biases about our collected data, and how they affect our results. This is a very difficult problem but it is central to correctly interpreting data as data scientists. We will need to use a multiverse analysis for reliable conclusions and pre-registering the hypothesis we need to test.

Although violent crime has been decreasing in the US for the last 30 years [7] it is apparent that most news outlets amplify the national-level fear through over-reporting violent crime with "buzz words" like "Violent Crime in the U.S. Is Surging" [8] and "US crime: Is America seeing a surge in violence?" [9]. Coupled with the use of Donald's Trump use of fear as a political weapon [10], US citizents regard crime on national level as an "immediate crisis"

[11]. The problem with US News Media reporting is that black Americans and Hispanic men are overrepresented as perpetrators of violent crime. It is undoubtable that news media not only affects, but controls public opinion so the aim of this study is to dive deep into what the data represent, and understand the real factors that affect violent crime. We hope that this study will give an insight into how we should interpret radical News Media Outlets and explore the intricacies that come with reporting numbers that have an inherent bias.

2 Background

A recent study [12] uses the 2018 FBI Uniiform Crime Reporting (UCR) program to compare with the Bureau of Justice Statistics' National Crime Victimisation Survey (NCVS) to address the issue of association between race and violent crime. The report measures the amount of violent crimes and the race and ethnicity of offenders. The Persons arrested data comes from UCR and the interviews from 151,055 U.S. households come from the NCVS. The study has an objective to find statistically significant variables that are correlated with (non-fatal) violent crimes. This is somewhat different from our dataset because our dataset includes fatal crimes (murders) but the rest of violent crimes (assault, robbery, rape) are included in both this study and our dataset.

The Bureau found that black people were overrepresented among people arrested for violent crimes, in comparison to their representation in the country's population. This was also the case for Hispanics, whereas white people were underrepresented. The study also found that there was a disparity between the race of the offenders and whether they were arrested by the police. For example, the authors report that 52% of white offenders were identified by their victims and only 45.9% were arrested, whereas out of the 28.9% black offenders, 33% were arrested for their crimes. This shows that the data at our disposal for violent crimes committed will have an over-representation of black offenders since they are both more reported, and arrested by the police. The report also found that there was no statistically significant difference with race and the people arrested. There is a large disparity of Hispanic offenders (21%) and how many violent crimes were reported (12%) but the authors explain that this might be because the victims could sometimes not identify the race of their offenders.

The authors also reported that the average number of offenders per incident, ie. the number of people involved in the crime did not vary significantly with the victim's race. The authors report that an average of 1.3 offenders were reported for each incident and the proportion of single-offender crimes was lower for black victims (78%) in comparison with the white victims (90%).

Overall this study shows that the assumption of association between race and ethnicity of violent crime in the US is unfounded. It specifically concluded that "White and Black people were arrested proportionate to their involvement in SNVC overall and proportionate to their involvement in SNVC reported to the police." The report also highlights the fact that the reported crimes also depend on the race of the victims which is something affecting our own dataset and will need further exploration.

The original paper published by Redmond (creator of the Communities and Crimes dataset) builds a software tool that can be used by police departments based on the published dataset [3]. The aim of the paper is to provide insights to police departments on how they operate. The authors want to help their departments beyond the departments' own experiences and instead allow them to benefit from the cumulative experiences of other police departments. The fact that the feedback from two police departments was positive shows how a data-driven approach can have real impact in law enforcement. However, this makes even more important to explore biases in the data, because while police departments might give positive feedback, the biasness of the program can sometimes discriminate and flag people of minor racial backgrounds.

One suggestion by the authors was to use the program to set goals for reducing crime within communities and check if the goal is feasible by looking at other similar communities. One problem that might arise though is that each community has very different demographics. In fact, this software was criticised because the biasness of the data lead to increased

police surveillance of minorities.

There are a lot of papers that tackle the fairness of data and algorithms. Mehrabi *et al* [13] review many approaches with identifying and combating bias in machine learning models. Looking specifically at discrimination bias, they use the example of COMPAS which is a decision tool used by US courts to assess the likelihood of a defendant recommitting a crime [14]. It was found to have higher false positive rates for black offenders than white offenders by falsely predicting that they have a higher risk of recommitting a crime. This is directly related to the data analysis that we perform here, where we are trying to find if race is an indicator for violent crime in communities in an unbiased way.

One main criticism of COMPAS is that it is data-dependent and therefore biased if the data gathered are biased. In fact the authors of the review paper talk about the two sources of unfairness that a model can have, those coming from biases in the data and those that come from biases in the algorithms. They explain that " Algorithms can even amplify and perpetuate existing biases in the data. In addition, algorithms themselves can display biased behavior due to certain design choices, even if the data itself is not biased.". [13]

The biases in the data appeared in the COMPAS tool where that if friends and family previously offended, they would be flagged as more high risk than others. Also prior arrests were a very key factor that played a role in the tool but this has the inherent bias that minority communities are controlled by the police more frequently, which means their arrest rates are higher. This was actually also proven in Bureau's report [12]. Other forms of bias include when some variables are left out of the model, when the data do not represent the diversity of a population, aggregation bias and many more.

Algorithms can also be very biased having algorithmic bias, user interaction bias, popularity bias etc as mentioned in [13] but they will not be very relevant for our own analysis since we will only be using simple linear regression models.

The violent crime report by Saridakis [15] explores the socio-economic and demographic variables on violent crime in the US. The study contains crime data over the period of 1960-2000 from similar sources to our dataset like the UCR. It also uses statistics from NCVS, and the Bureau of Economic Analysis. The results of the study found that income inequality had a significant positive effect with murder (not necessarily with other crimes). Moreover, alcohol consumption was very significant with murder and rape (not assault), and this is backed up by psycho-pharmacological theories that support the idea that alcohol consumption is actually associated with violent behaviour. Unemployment showed that it was not a large predictor for violent behaviour but the proportion of black males had a positive and significant effect on all types of crime.

The author suggests that "the three strikes"[1] policy adopted in 1993 might be the reason the violent crime has experienced a decrease for the years 1993 onwards.

Overall, these studies show different approaches for using violent crime data to arrive at conclusions. Saridakis [15] finds a conflicting result to the Bureou [12] on the correlation of violent crimes and race. This shows how instigate and interesting the datasets with socio-economic factors are. Mehrabi *et al* [13] introduce all the biases that we need to be mindful and how careful we need to be in drawing conclusions from data.

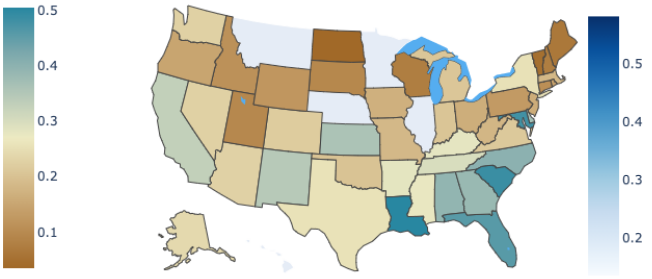
3 Data and Methods

The Communities and Crime dataset includes 122 predictive normalized features (scaled between 0 and 1), 5 non-predictive that have to do with the geographical location of the community and one target, the total number of crimes per 100,000 population. Out of the predictive features, we observed that 22 of them had exactly 1,675 missing values. Looking to this further we understood that the features corresponded to the responses of the US LEMAS survey. We noticed that there was no particular geographical pattern for the missing data, since most of the states had at least one response to the survey which just meant that not all of the communities included in the 1990 US Census where given the LEMAS survey. We removed these features from our dataset since they affected 84% of our observations (total observations are 1,992). That left us with just 1 observation containing missing values which

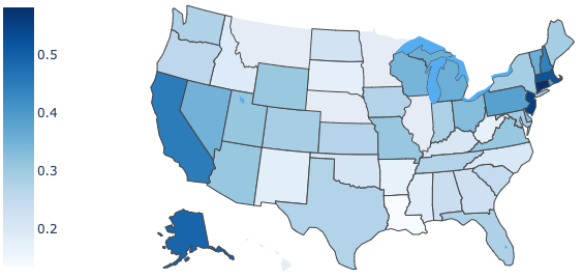
[1]The three strikes policy increases the prison term for the felons who convicted 3 or more serious crimes.

we just removed from the dataset because it only consisted of a community (Natchez city) with a small population (0.02) so it would not affect our model.

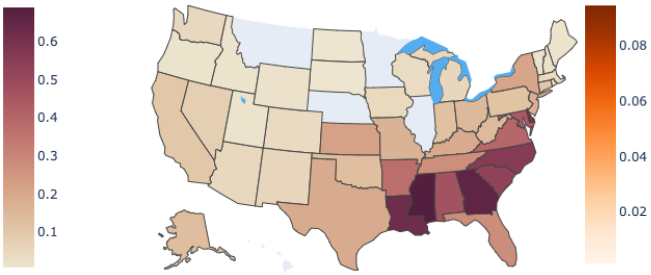
The US-States were labeled with a number which did not map to any of the official numbering of states found online, so we manually mapped the number to the correct state by looking at the communities names. This manual step was important for understanding the geographical representation for some of our socio-economic variables. In figure 1a we can visualize how the average crime levels vary across the states. It seems that the largest amount of crime is concentrated in the south-eastern part of the US with some notably high levels in California, New Mexico and Arizona. From 1c we can see that most of the African-American population in the US is also concentrated in the south-eastern part. Figure 1b shows that the average median income is very high in states like Alaska, California, New Jersey, and Connecticut with the rest of the states having very similar average household income. Finally, we can see from figure 1d that most immigrants are concentrated in the western part of the US with the exception of Kansas.



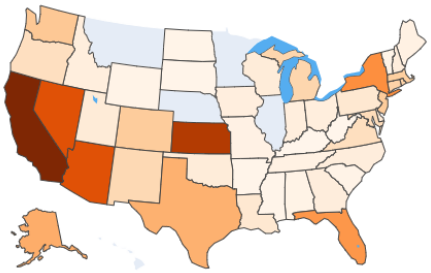
(a) Average Violent Crimes Per 100K Population



(b) Average Median Household Income



(c) Average percentage of African-American population



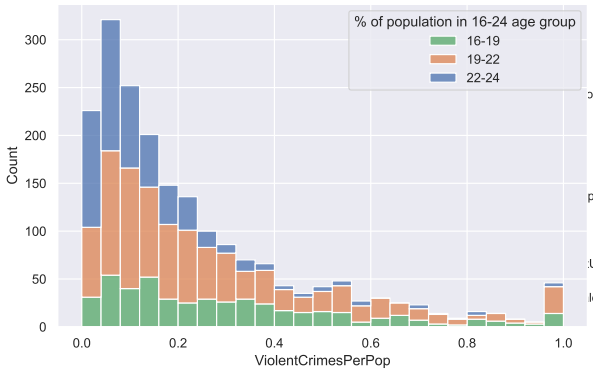
(d) Average total number of Immigrants

Figure 1: Chloropleth Maps of variables across the US. Grey states represent states that we do not have enough (or any) data on the states and light blue colors represent the lakes.

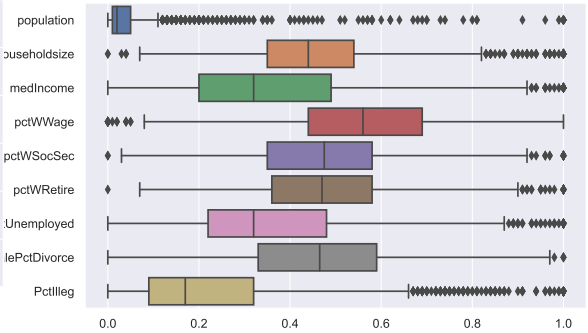
In figure 2a we can observe the age group that is most responsible for the crime. We use the upper, lower and median quantiles to separate the % of population variable to three levels and inspect the amount of committed crime by each age group. The population with ages 19-22 seem to be responsible for most of the crime since they are consistently larger in proportion than the other two age groups. We explored the distributions of many socioeconomic features and figure 2b visualizes some of them. We can see that the population and the percentage of kids born and never married are both right-skewed with heavy tail. Most of the other variables seem to be well distributed, having few outliers.

We needed to address the right-skewness of our target variable since that would introduce heteroscedasticity in our model. One way to address this is by performing a log transformation and it is clear from figure 3 that our target variable is now centered with balanced values.

After removing the most of the LEMAS data we were left with 100 predictive features. These features attained many correlations between them, so we decided to remove the most in-between correlated variables (and only keep 1 of them) in order to decrease multi-colinearity in our model. First we check the ethnic and racial composition of the communities and notice that there is a -0.79 correlation between the percentage of black people



(a) Distribution of violent crimes in the 16-24 age group



(b) Socioeconomic variables visualisation. Horizontal lines represent the 95% confidence intervals and the points represent outliers.

Figure 2: EDA of features in the dataset

in a community with percentage of Caucasians in the community so we can remove one of the two variables. This shows that communities with a lot of Caucasians in the US do not usually have a lot of black people. This might be interesting to explore since it shows that the two races do not mix together well and this was also shown from the map in 1c. Moving on, we remove all the features that have more than 85% correlation between them, which accounts to ≈ 50 features. Finally we remove any variables that have less than 20% correlation with the response (since they would not serve as good predictors in our model). After all the data cleaning we are left with m =38 features, 1 target and n =1,992 observations.

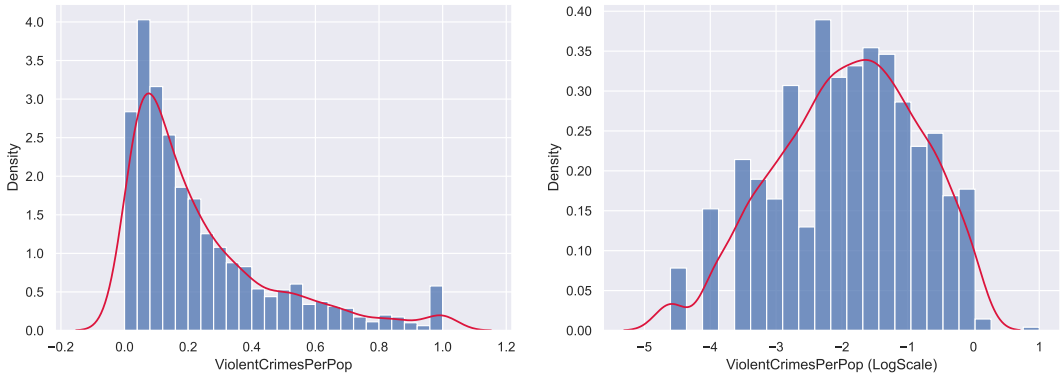


Figure 3: Violent Crime Per 100k Population distribution and its log transformation.

The method we will be using to understand the factors that affect the crime rate in a community will be a Linear Regression model where the solution can be derived using least squares represented as:

$$\underset{\boldsymbol{\beta}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \tag{1}$$

where $\boldsymbol{y}$ is the $n \times 1$ target variable (violent crimes per 100k population), $\boldsymbol{X}$ is the $n \times m$ feature matrix (including a vector of ones as the intercept) and $\boldsymbol{\beta}$ is the $m \times 1$ coefficient matrix corresponding to each feature. The solution to the problem is given by:

$$\boldsymbol{\beta} = \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \boldsymbol{y} \tag{2}$$

We chose this simple model for our analysis because with sensitive topics including race we want to be able to interpret how our model arrives to its conclusions. It might not be the most accurate model, but we can explore exactly how the weights are assigned and understand the solution to our problem. This method is also widely accepted in existing literature for this kind of dataset like for example in [15].

After solving the least squares problem we will look at a few metrics to assess the quality of our trained model. The metrics include the Root Mean Squared Error (rmse) and the adjusted R^2 defined as:

$$\text{rmse} = \left(\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \right)^{1/2} \qquad R^2_{\text{adj}} = 1 - \frac{\text{RSS}/(n-p)}{\left(\sum_{i=1}^{n} (y_i - \bar{y})^2 \right)/(n-1)} \tag{3}$$

where $\hat{\boldsymbol{y}}$ represents the predicted values from the model, RSS the residual sum of squares and $\bar{y}$ the average value of $\boldsymbol{y}$. The value for R^2_{adj} lies between 0 and 1 and shows how much of the variation in the model can be explained. We use the adjusted version because it takes into consideration the amount of features we include in the model.

To explore the features that have the largest impact to our model we will look at the absolute value of the coefficients $\boldsymbol{\beta}$ and their corresponding p-values when computed using the t-test [16] with the Null Hypothesis that they should not be included in the model.

To assess whether our models are biased we will split our dataset into training and testing random splits consisting of 70% and 30% respectively. When computing our metrics on our test dataset, we make sure that any overfitting that might happen in the training dataset will be reflected with a much lower RMSE and R^2_{adj} scores in the test set. We also repeat the train-test split, compute the metrics 100 times and take their average, in order to decrease the bias in the model as much as possible.

We will use a multiverse analysis approach as suggested by Steegen *et al* [17] in order to explore the different conclusions from each approach. We will first consider a model that only includes ethnic and racial features and see how the conclusions compare with the full model that includes all variables mentioned above. We will also explore the possibility of including the initially removed missing values from the LEMAS survey.

4 Results

Our initial model (RACE) only included ethnic and racial features from the dataset. These were the percentage of population of the different races (Caucasian, African-American, Asian, Hispanic, native American) aswell as their per capita income (total of 10 features). The results are shown in table 1. The target was the log-transformed violent crimes per 100K population. Both train and test splits have similar R^2 and rmse values, which shows that the models have not overfitted in training. R^2 is very low, where intuitively it means that only 48% of the variation in the dataset is explained by the model. This shows a very low performance and suggests that the model including only these variables is not powerful enough to predict the violent crime in a community. The most important features were found to be the % of black and hispanic people in the community with p-values $\approx 10^{-12}$ that mean that the null hypothesis of not including them in the model is rejected. Overall this model can explain some of the variation in the dataset but not enough to conclude that race/ethnicity plays an important role in the amount of violent crime in a community.

Then we fitted our BASE model which includes all 38 features and the target variable (not transformed). The base model performed significantly better giving a test R^2 of 0.64. This is still a relatively low R^2 score, but as expected more features in our model matrix gave a more accurate model. Interestingly, none of the predictors from the RACE model were present in the top 3 predictors of BASE. In fact, the number of homeless people in the street was the highest followed by kids born to never married and percentage of households with investment as their income. These are predictors that all have to do with the economic state of the community rather than race or ethnicity.

Performing a White test [18] which is a statistical test to check the heteroscedasticity of the errors we get a p-value of 10^{-8} indicating our model suffers from heavy heteroscedasticity. To resolve this we log-transform our target variable as seen in figure 3 and create a new model called LBASE. This model has a slightly lower R^2 than the BASE but performing a White's test we observe a p-value of 0.18 suggesting that there is no sufficient evidence to reject the null hypothesis that our model suffers from heteroscedasticity. Figure 4 shows a visualisation

of our test set predictions, showing that the model is able to capture the general trend of values, but still suffers for high and low levels of crime.

It is interesting to see that the percentage of households with investment as their income appears again in the top three indicators, followed by the percentage of households with wage as their income and the percentage of males who are divorced. Again race and ethnicity do not appear in the model and the p-values of these features are $\approx 10^{-5}$ showing that they are very significant for our model.

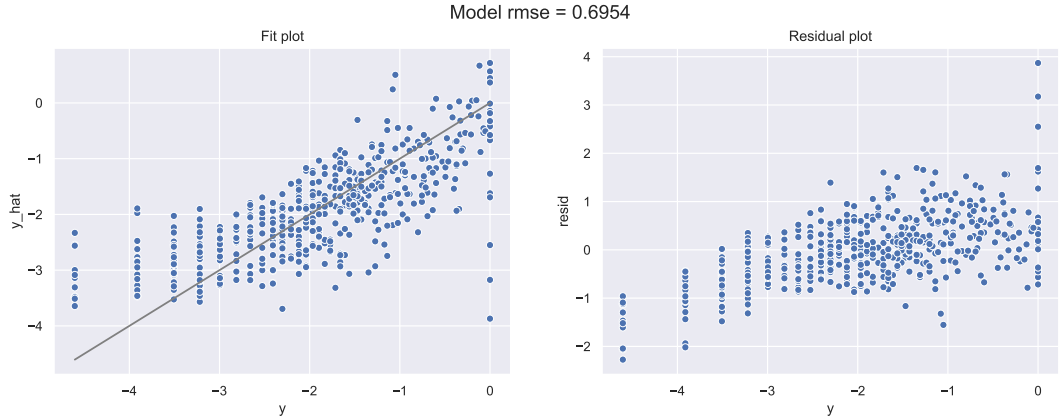


Figure 4: Results of linear regression from LBASE model. Plot on the left shows how close our predictions (`y hat`) were with the true values (`y`) in our test (30%) dataset. The plot on the left shows the distribution of the residuals for the fit.

We also created a LEMA model which only includes the 319 observations that do not have missing values for the features collected from the LEMA survey. Cleaning the very correlated features ($> 85\%$) we are left with 73 predictors. Table 1 shows that this model gave the highest train R^2 of 0.80. The test R^2 is significantly lower however and we believe this is mainly because the dataset only has 319 observations so the model must have overfitted on the small amount of data available. Again the main predictors were mainly economic ones with `pctWInvInc` coming to the top.

model	train rmse	test rmse	train R^2	test R^2	most important features
RACE	0.77	0.78	0.48	0.47	`racepctblack, racePctHisp, whitePerCap`
BASE	0.13	0.14	0.67	0.64	`NumStreet, PctIlleg, pctWInvInc`
LBASE	0.65	0.70	0.63	0.61	`pctWInvInc, pctWage, MalePctDivorce`
LEMA	0.34	0.61	0.80	0.24	`pctWInvInc, PctFam2Par, PctHousLess3BR`

Table 1: Summary of results from all tested models. Note that the datasets for each model are different but they always consist of 70% training and 30% random splits. The most important features of each model are in the order of absolute value of their coefficient. Details of what each feature represents can be found here.

Finally we wanted to see if we arrive at the same conclusions for different states in the US. To address this we chose the states that had at least 50 observations (to make sure we have enough data for a reasonable study). We then fitted a linear model on each state separately using the log-transformed target and recorded the rmse, R^2 values and top predictors. The results are shown in table 2.

The R^2 and rmse values are not very informative since each linear regression only included a few data points (50-278). Florida has the highest predictive accuracy, with $R^2 = 0.87$. What is interesting to look at is the predictors which affect the linear regression the most. Firstly, we observe that California which has the most observations in the dataset (278) has the `pctWInvInc` as the top predictor. This predictor only appears once more in Wisconsin, which shows that our results from the LBASE and LEMA model were heavily influenced by the single state of California. This shows how inbalanced data from each state can heavily affect our models. The main predictor that appears in 7 out of the 9 states is the total number

state	R^2	rmse	most important features
CA	0.70	0.41	`pctWInvInc, NumImmig, population`
NJ	0.78	0.47	`racePctWhite, NumImmig, medIncome`
TX	0.63	0.49	`NumInShelters, MalePctNevMarr, PctFam2Par`
MA	0.68	0.58	`NumImmig, population, racePctWhite`
OH	0.69	0.60	`NumImmig, racePctWhite, racepctblack`
PA	0.74	0.53	`NumImmig, NumStreet, racePctWhite`
FL	0.87	0.26	`racepctblack, racePctHisp, racePctWhite`
CT	0.82	0.45	`population, racePctWhite, NumImmig`
WI	0.67	0.48	`NumStreet, NumImmig, pctWInvInc`

Table 2: Summary of results from Linear regression for top 9 states in the US in terms of observations present in the dataset. There was no train-test split for the calculation of R^2 and rmse. Details of what each feature represents can be found here.

of immigrants (`NumImmig`). This is interesting because this predictor does not appear at all in our 4 models in table 1. Looking at 1d this predictor is well spread across the country, and it appears like it is a heavily right skewed predictor. Racial predictors appear in 5 states, with Florida having all of its top predictors having to do with race. Florida has the largest African-American population out of the states appearing in this table (although not the largest overall), which means it played a large role in predicting a high violent crime level. The analysis of the states has showed that clearly different states have different predictors that affect their violent crime.

5 Conclusion

Our multiverse analysis has shown that there are multiple conclusions that the reader can adhere to. We have shown that 4 different models on the same data arrive at similar conclusions with different methods. We have seen that just looking at the race/ethnicity of individuals is not enough to explain violence in a community from our RACE model. The BASE and LBASE models have shown that reasonable conclusions can be drawn from data, with each model using different predictors to arrive at conclusions. It is clear that there is no universal predictor for predicting violence in the US. Each state had different important features and the predictive accuracy for each state was different. This shows how important equal representation between states is when collecting socio-economic data.

The cleaning of the data has allowed us to avoid creating a multi-collinear model, with many predictors meaning the same thing. We have tried to address the potential problems that come with biased algorithms and models, as suggested by [13], by splitting the data into training and testing sets multiple times and averaging the results. One extension to this would be permuting the data to make sure our predictions using the most important variables are better than a random permutation of the covariates. Another extension could be to look at color-coding figure 4 with some predictors like race or age. If the model was biased towards any of those predictors, we would clearly see patterns arising between our predictions for the target variable and the actual values.

Most of the LEMA survey was excluded from the analysis because of the amount of missing values, but if the survey was completed for all the communities more insightful predictors could have come to our attention. Running a model that only included the observations without missing data did not reveal any variables from the LEMA survey that were very significant. However, a 14% sample from the dataset is very small to arrive a reliable conclusions on this so we reserve our judgment as to whether the LEMA survey could bring an immediate impact to this analysis.

Bias in the data was something that we had little control over. Unfortunately when conducting such analysis we are not able to know the exact methods that data is collected and what biases could be present. For example, Saridakis [15] noted that the violent crimes of rape were an underrepresented sample in the census, because victims felt less comfortable giving details about their abusers than other forms of violent crime. Moreover, the racial

composition of the police in each state might have an inherent bias against other races and the amount of arrests affect that. The amount of reported crimes does depend on race which means states with more prominent racial compositions will have different proportions of reported crimes. We tried to combat this using a multiverse analysis and taking different subsets with different assumptions in our analysis, but since this a systematic bias in the dataset, it is not something that we can easily eliminate.

Overall, our analysis has shown that the racial and ethnic composition of a community does not affect the amount of violent crime in the US. Instead, economic aspects of a community have a greater impact to the amount of violence. For example, the communities that earned their salary from investments and rents are likely to be wealthier individuals, living in more secure areas with well structured and well funded law enforcement units, which therefore leads to less violent crime. Homelessness in a community also proved to be an important predictor, which also makes intuitive sense, since homeless people live under rough circumstances that sometimes mean crime is part of their survival. Another possible cause of violent crime in the US that was not explored in this study was suggested by Nevin to be lead exposure [19]. This is one of many other predictors that could be correlated with increased crime rates, which shows how broad and intricate is the problem we try to address.

The racial and ethnic compositions of a society appear as predictors of violent crime because it is easy to give blame to something people have an inherent bias against. Unfortunately, there is still a lot of racism in the US, and although it might seem like the country has "woken" there are some embedded biases that people adhere to when it comes to race. This study has shown us that depending on what narrative one wants to believe, they can make the data show it. It is easy to focus the study in states with large African-American communities like Florida and Ohio while excluding most of the most important predictors like economic state of a community.

In conclusion, we believe that the data have so many inherent biases that it is very difficult to arrive to a reliable conclusion out of it. This is something that Steegen *et al* [17] encourage: "reserve judgment and and acknowledge that the data are not strong enough to draw a conclusion". With sensitive topics like this one, more unbiased data are needed to find reliable conclusions. We hope that this report has encouraged the reader to reserve judgment when looking at the next headline from the media saying that certain ethnic groups are responsible for violent crime in the US, and look further into how the study has been conducted.

References

[1] Wikipedia. Black Lives Matter — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Black%20Lives%20Matter&oldid=1079181171, 2022. [Online; accessed 21-April-2022].

[2] Michael Redmond. UCI machine learning repository, 2002. URL http://archive.ics.uci.edu/ml.

[3] Michael Redmond and Alok Baveja. A data-driven software tool for enabling cooperative information sharing among police departments. *European Journal of Operational Research*, 141(3):660–678, 2002. URL https://EconPapers.repec.org/RePEc:eee:ejores:v:141:y:2002:i:3:p:660-678.

[4] FBI. Crime in the united states 1995, 1995. URL https://ucr.fbi.gov/crime-in-the-u.s/1995.

[5] United States Census Bureau. 1990 census, 1990. URL https://www.census.gov/data/datasets/1990/dec/summary-file-3.html.

[6] United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. Law enforcement management and administrative statistics (lemas), 2007, 2011.

[7] Jeff Sessions. Violent crime is up some, but still well off historical highs, 2017. URL https://www.politifact.com/factchecks/2017/dec/04/jeff-sessions/violent-crime-some-still-well-historical-highs/.

[8] Thomas Abt, Eddie Bocanegra, Emaada Tingirides. Violent crime in the u.s. is surging. but we know what to do about it, 2022. URL https://time.com/6138650/violent-crime-us-surging-what-to-do/.

[9] Jake Horton. Us crime: Is america seeing a surge in violence?, 2021. URL https://www.bbc.co.uk/news/57581270.

[10] Alex Altman. No president has spread fear like donald trump, 2017. URL https://time.com/4665755/donald-trump-fear/.

[11] Elizabeth Sun. The dangerous racialization of crime in u.s. news media, 2017. URL https://www.americanprogress.org/article/dangerous-racialization-crime-u-s-news-media/.

[12] BJS Statistician Allen J. Beck, Ph.D. Race and ethnicity of violent crime offenders and arrestees, 2018. *Bureau of Justice Statistics*, 2021. URL https://bjs.ojp.gov/library/publications/race-and-ethnicity-violent-crime-offenders-and-arrestees-2018#additional-details-0.

[13] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning, 2019. URL https://arxiv.org/abs/1908.09635.

[14] Wikipedia. COMPAS (software) — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=COMPAS%20(software)&oldid=1078441070, 2022. [Online; accessed 21-April-2022].

[15] George Saridakis. Violent Crime in the United States of America: A Time-Series Analysis Between 1960-2000. Discussion Papers in Economics 03/14, Division of Economics, School of Business, University of Leicester, October 2003. URL https://ideas.repec.org/p/lec/leecon/03-14.html.

[16] Wikipedia. Student's t-test — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Student's%20t-test&oldid=1083567706, 2022. [Online; accessed 22-April-2022].

[17] Sara Steegen, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11 (5):702–712, 2016. doi: 10.1177/1745691616658637. URL https://doi.org/10.1177/1745691616658637. PMID: 27694465.

[18] Wikipedia. White test — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=White%20test&oldid=1078273400, 2022. [Online; accessed 22-April-2022].

[19] Rick Nevin. Understanding international crime trends: The legacy of preschool lead exposure. *Environmental Research*, 104(3):315–336, 2007. ISSN 0013-9351. doi: https://doi.org/10.1016/j.envres.2007.02.008. URL https://www.sciencedirect.com/science/article/pii/S0013935107000503.