INCONSISTENCY

of Machine Learning in Recidivism

Donny Fitrado
Data Science Student
Career Semester

COMPAS

definition

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) is a popular commercial algorithm, designed by the Software Company Northpointe Inc, used by judges and parole officers for scoring criminal defendant's likelihood of reoffending (**recidivism**).

Machine Bias

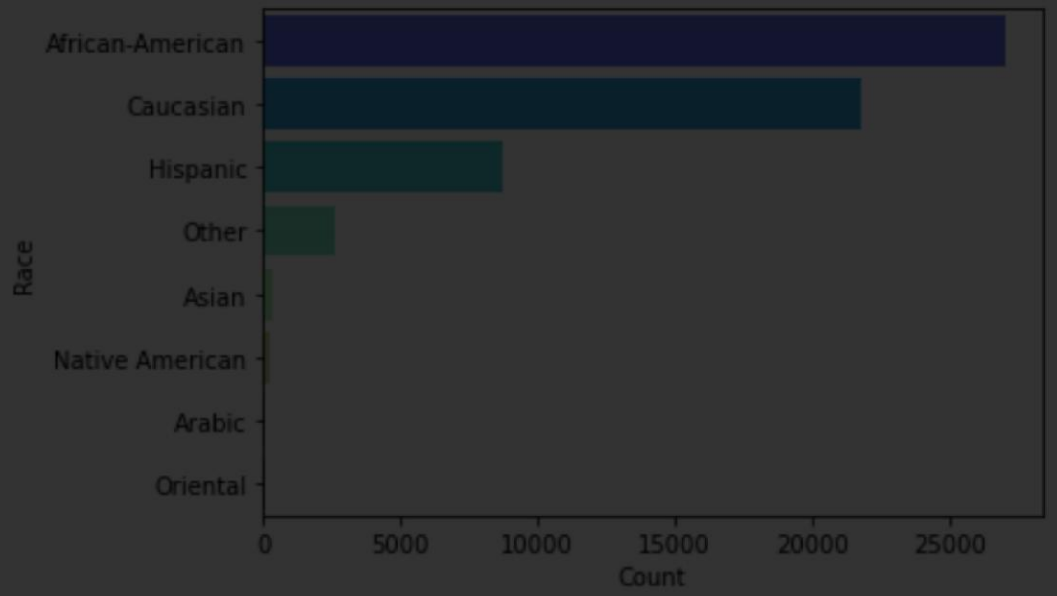
With technological developments in recent years, people have started to use AI algorithms, past statistics, and data science to reduce crime in this world.

The formula was particularly likely to falsely flag black defendants as future criminals, wrongly labeling them this way at almost **twice** the rate as white defendants. White defendants were also mislabeled as low risk more often than black defendants. The algorithm essentially segregates based on race and fails to serve equal justice.

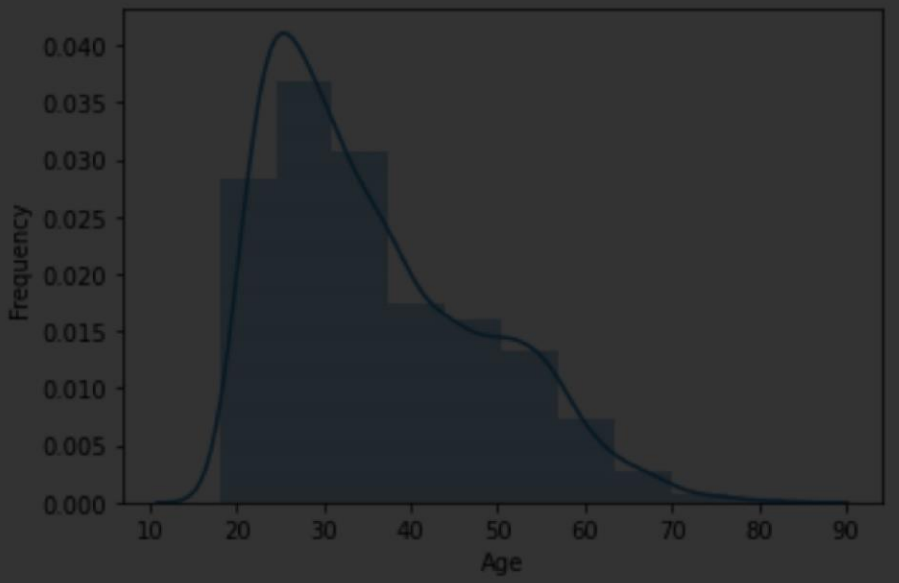
(**Prorepublica**, 2016)

COMPAS DATASET

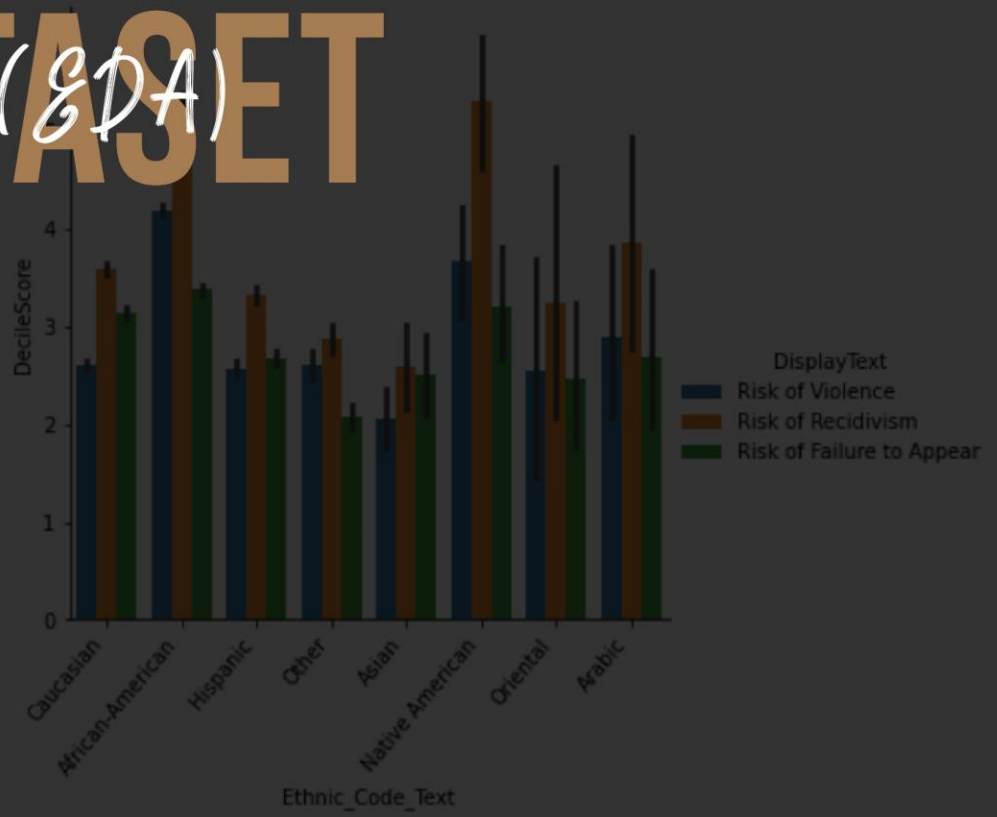
exploratory data analysis (EDA)



	Ethnic_Code_Text	count
0	African-American	27069
1	Arabic	75
2	Asian	324
3	Caucasian	21783
4	Hispanic	8742
5	Native American	219
6	Oriental	39
7	Other	2592



RecSupervisionLevelText	Low	Medium		
Ethnic_Code_Text				
African-American	12831	5976	5232	3030
Arabic	51	15	6	3
Asian	288	18	9	9
Caucasian	16155	3075	1881	672
Hispanic	6903	1026	537	276
Native American	108	51	42	18
Oriental	36	0	0	3
Other	2100	315	120	57



RISK BY ETHNICITY

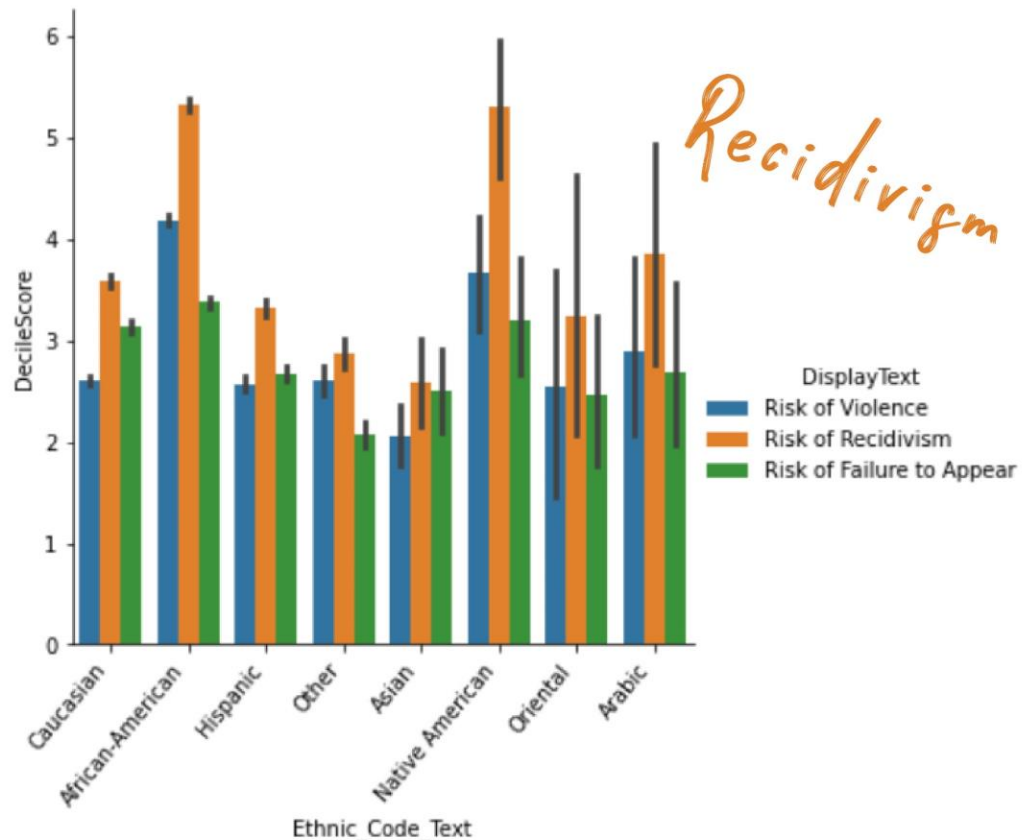
Assessment

RecSupervisionLevelText Ethnic_Code_Text	Low	Medium	Medium with Override Consideration	High
African-American	12831	5976	5232	3030
Arabic	51	15	6	3
Asian	288	18	9	9
Caucasian	16155	3075	1881	672
Hispanic	6903	1026	537	276
Native American	108	51	42	18
Oriental	36	0	0	3
Other	2100	315	120	57

Using some basic EDAs, we compare the risk of **violence**, **recidivism**, and risk of **failure to appear**. With this result, we can provide a similar output to Prorepublica's analysis.

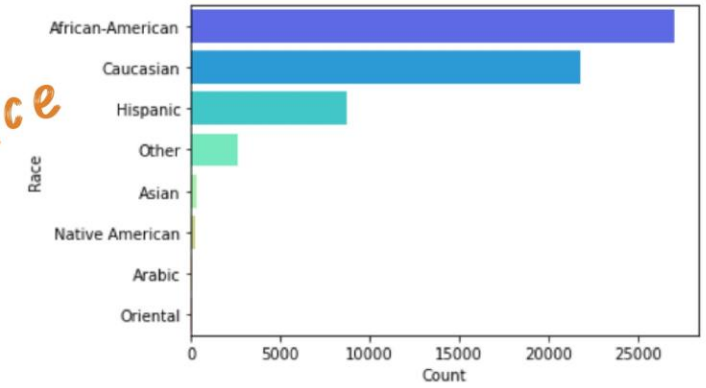
The table above represents the **assessment** risk by each ethnicity, while the graph below represents the **recidivism** risks by each ethnicity, starting from African-American, Arabic, Asian, Caucasian, Hispanic, Native American, Oriental, and Other.

Recidivism



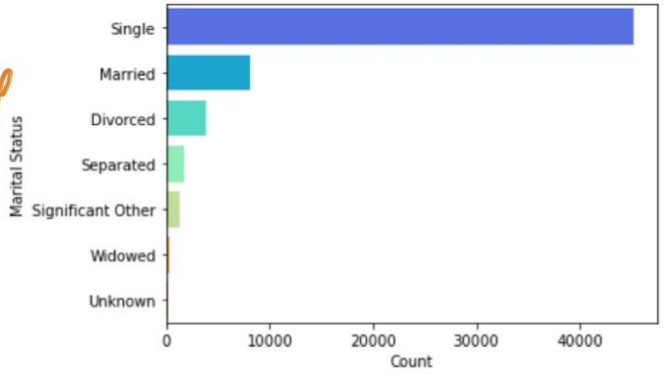
	Ethnic_Code_Text	count
0	African-American	27069
1	Arabic	75
2	Asian	324
3	Caucasian	21783
4	Hispanic	8742
5	Native American	219
6	Oriental	39
7	Other	2592

Race

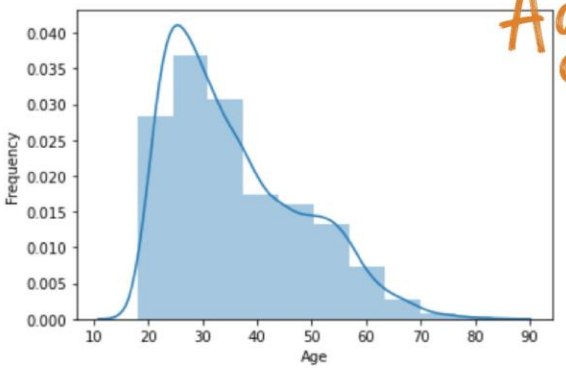


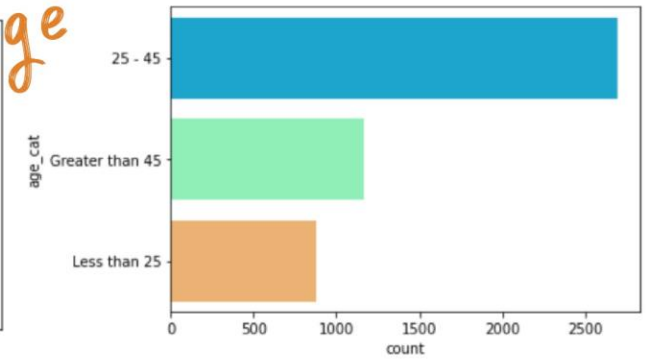
	MaritalStatus	count
0	Divorced	3879
1	Married	8172
2	Separated	1803
3	Significant Other	1260
4	Single	45126
5	Unknown	213
6	Widowed	390

Marital Status



Age





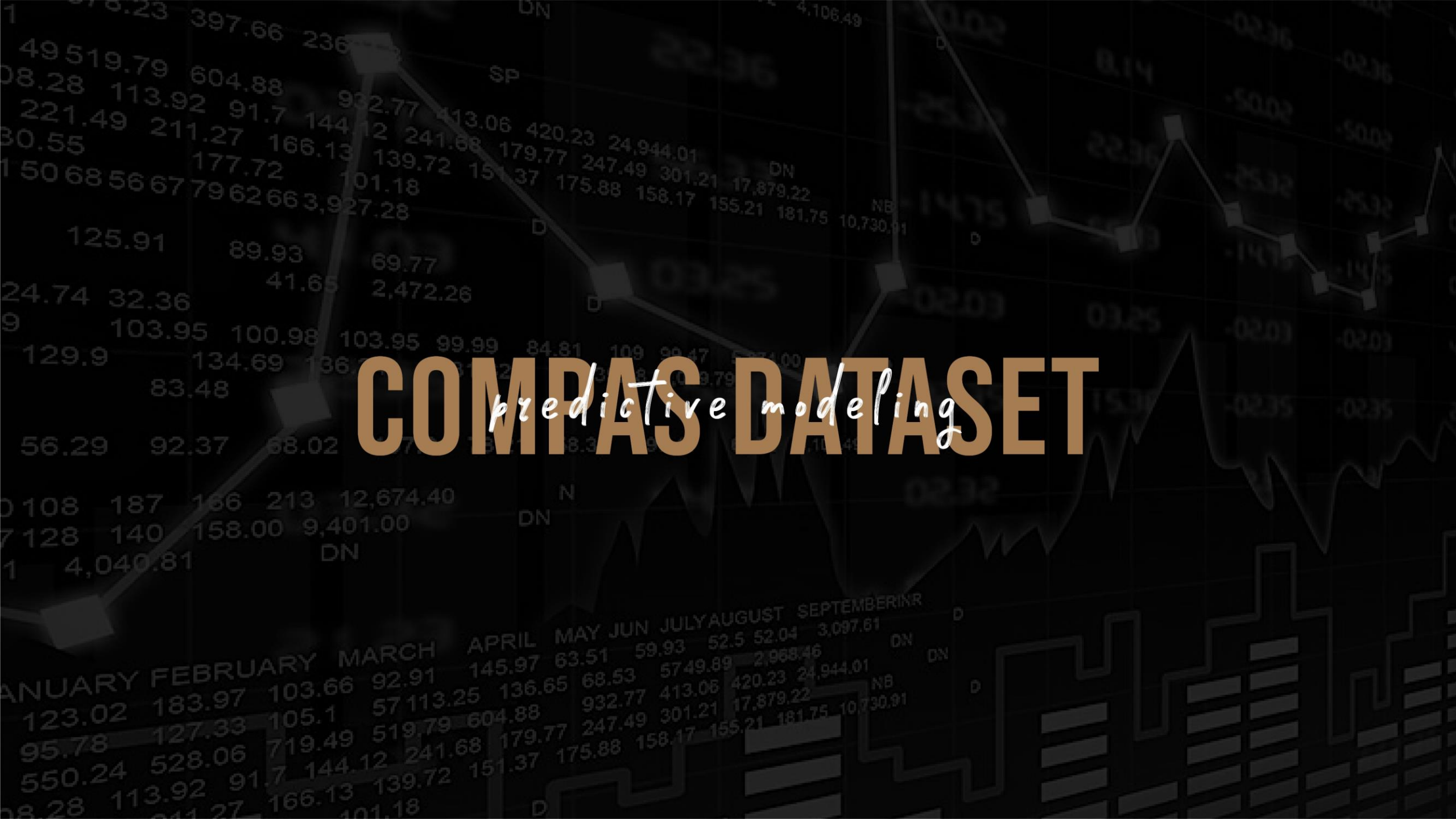
MORE

EDA

I also provide more exploratory data analysis (EDA) on the compas dataset. The EDA mainly discuss about the count difference between **race**, **marital status**, and **age**. We can clearly see that African-American race has the most number of count, comparing to the other race.

While on the marital status graph, single status has the most count, followed by **married**, **divorced**, and **seperated**.

Last, there's a graph that shows the frequency between age, starting from age 0 to 100. Besides that, there's a number count of **age group** of 25 to 45, then greather than 45, and less than 25.


COMPAS DATASET
predictive modeling

THE OUTPUT

Based on my predictive model code, the **output** of the program are the number in nearest neighbors, linear support vector machine (SVM), Radial Basis Function SVM, Decision Tree, Random Forest, Neural Network, Adaboost, Naive Bayes, QDA, and Logistic Regression.

```
Nearest Neighbors 0.6123110151187905
Linear SVM 0.652267818574514
RBF SVM 0.6641468682505399
Decision Tree 0.6760259179265659
Random Forest 0.6711663066954644
Neural Net 0.673866090712743
AdaBoost 0.66792656587473
Naive Bayes 0.6241900647948164
QDA 0.6452483801295896
LogisticRegression 0.6652267818574514
```

Best Score:

0.719840807891299

Best Estimator:

```
DecisionTreeClassifier(criterion='entropy', max_depth=3, max_features=8,
                       min_samples_leaf=4)
```

Best Param:

```
{'criterion': 'entropy',
 'max_depth': 3,
 'max_features': 8,
 'min_samples_leaf': 4}
```

From the code, we can obtain the best score, best estimator, and best parameter.

The best estimator is Decision Tree Classifier, and the DT's parameter which consists of the criterion, max_depth, max_features and min_sample_leaf, also become this predictive model's best parameter.

	precision	recall	f1-score	support
0	0.68	0.78	0.73	1020
1	0.67	0.54	0.60	832
accuracy			0.67	1852
macro avg	0.67	0.66	0.66	1852
weighted avg	0.67	0.67	0.67	1852

FINAL REPORT

predictive model

On this final report, we can obtain a table that consists of precision, recall, f1-score, and support. While on the left side, there are accuracy elements, such as macro average, and weighted average. Each of the elements have different value that represents the predictive model that was built.

Conclusions

With the EDA and predictive model I built, the COMPAS model really has a bias in terms of race and tends to incorrecly predicts the recidivism risk of **African-American race**. I think COMPAS dataset needs to be rechecked, because it seems to also affect the algorithm judgement on some specific race. Besides that, I do believe that COMPAS will become the future of crime prevention if the algorithm is becoming better and better everyday.

Most importantly, we should **remember** that there must be a justice and equity in terms of on each race's crime prediction.

THE

SOURCES

https://www.kaggle.com/danofer/compass

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

http://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html

http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

http://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html

https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html

http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostRegressor.html

http://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html

http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html


THANK YOU
Hope you enjoy the presentation!