

In response to concerns about racial bias in criminal sentencing, several court systems have adopted automated systems that are intended to be "unbiased", and to look at a defendant's case objectively. One such system is COMPAS, which is intended to help judges determine whether to grant bail to defendants awaiting trial.

However, it has been shown that COMPAS produces results that are biased against Black defendants, **even though the law forbids them from using race in their calculations**. How could that be?

To begin, read [this article](#) and try out the simulations that are included.

Then, prepare a reflection, according to the class rubric, that addresses the following questions:

1. What are the two different kinds of fairness that the article discusses regarding COMPAS?
2. Why is it not possible to reconcile them?
3. The article quotes sociologist Ruha Benjamin, who warns us that machine learning algorithms are often trained with "data produced through histories of exclusion and discrimination." What specific data in this case is rooted in discriminatory practices?
4. COMPAS was developed by a private company, rather than by the government itself. Why is that a problem in this instance?

Work Referenced: Can you make AI fairer than a judge? Play our courtroom algorithm game. Karen Hao and Jonathan Stray. MIT Technology Review. October 17, 2019.

<https://www.technologyreview.com/2019/10/17/75285/ai-fairer-than-judge-criminal-risk-assessment-algorithm/>