

COMPAS Disparity for Hispanic Individuals

Data

Data Set Acquired From ProPublica:

- compas.db
- compas-scores-raw.csv
- compas-scores-two-years-violent.csv
- compas-scores-two-years.csv
- compas-scores.csv

from the COMPAS Recidivism Risk Score Data and Analysis.

Introduction

The data set I chose is a compilation of COMPAS scores from Broward County Sheriff's Office in Florida. The data set is from 2013 and 2014 and includes data for 18,610 individuals.

COMPAS, which stands for Correctional Offender Management Profiling for Alternative Sanctions, is used to assess a defendant's likelihood of becoming a recidivist, or in other words, their likelihood to reoffend. Created by a Michigan-based for-profit company Northpointe, COMPAS gives defendant's three scores "Risk of Recidivism," "Risk of Violence" and "Risk of Failure to Appear." The scores are on a 10-point scale, where 1-4 are considered "low", 5-7 are "medium" and 8-10 are "high". The scores are used by the criminal justice system to decide factors such as trial time, bail, and sentencing. It encompasses factors such as education levels and occupation. It asks questions such as: "Was one of your parents ever sent to jail or prison?" "How many of your friends/acquaintances are taking drugs illegally?" and "How often did you get in fights while at school?" The questionnaire also asks people to agree or disagree with statements such as "A hungry person has a right to steal" and "If people make me angry or lose my temper, I can be dangerous," [Larson Angwin 2016]. Northpointe refuses to release their exact algorithm.

Prediction

There are several interesting things about this data set. The first is that Broward County is the 17th most populated county in the US with nearly 2 million residents. The large population gives more diverse and expansive data. Another interesting thing about this data set is that it has been reviewed several times for similar reasons. Each have been successful in proving that COMPAS falsely predicts likelihood to reoffend more often than for Black individuals. Disproportionately more frequent than for White individuals. What has yet to be studied so in depth is if there is a disparity between Hispanic and White individuals in receiving inaccurate COMPAS results. I predict that there will be. I predict that it will be a less aggressive disparity between inaccurate results for Hispanics than Whites. The references I have compiled elude to my prediction being true.

References

- GitHub. propublica/compas-analysis. Retrieved October 1, 2020 from <https://github.com/propublica/compas-analysis/>
- Jeff Larson Julia Angwin. 2016. Machine Bias. (May 2016). Retrieved October 1, 2020 from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Julia Angwin Jeff Larson. 2016. How We Analyzed the COMPAS Recidivism Algorithm. (May 2016). Retrieved October 1, 2020 from <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>
- Justice served? Discrimination in algorithmic risk assessment. (October 2019). Retrieved October 1, 2020 from <https://researchoutreach.org/articles/justice-served-discrimination-in-algorithmic-risk-assessment/>
- ProPublica. 2019. COMPAS Recidivism Risk Score Data and Analysis. (March 2019). Retrieved October 1, 2020 from <https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis>
- The Biased Algorithm: Evidence of Disparate Impact on Hispanics. Retrieved October 1, 2020 from <https://www.law.georgetown.edu/american-criminal-law-review/in-print/volume-56-number-4-fall-2019/the-biased-algorithm-evidence-of-disparate-impact-on-hispanics/>