Exploring Race in Dating: Is Dating Racist?

Team Members
- Olivia Mintz
- Aiden Jolibois
- Oliver Green

Problem Description
We are attempting to predict whether or not participants in this speed dating event were a match. Using the AI Fairness tool we will also explore the role race plays in our model and discuss the bias/accuracy tradeoff as it applies to our model. With the evolving dating landscape and the emergence of dating apps in our society, we'd like to explore how race plays into people's dating preferences.

Dataset
https://www.openml.org/search?type=data&sort=nr_of_downloads&status=active&id=40536
Record of race, gender, dating interests, personality traits, and preferences of participants in experimental speed dating events from 2002 - 2004.
- 8378 Records
- 121 features

Approach and Methodology
The dataset includes columns with the differences in ratings between the participants. We will not be using these. In terms of normalization, all ratings columns are on the same 1-10 scale and we have categorical data like race. So we will be normalizing the numerical data to a 0-1 scale and one hot encoding the categorical data.

Models
Decision Tree, Logistic Regression, and Support Vector Machine

Language and Packages
Python, AI Fairness 360, Scikit-learn, Pandas, Numpy

Evaluation and Metrics
Precision, Recall, Accuracy, F1 score

Outcome
The outcome of the project will give us more insight into what characteristics are most important to someone choosing a partner to match with. We will determine if race is a more prominent factor as opposed to personality traits and interests.

Plan

Aiden will be doing the majority of the data cleaning. Olivia will focus on data exploration and visualization tasks. Olivia will also be working on the decision tree model. Aiden will be coding the logistic regression model. Oliver will be doing the SVM model. Oliver will also be focusing on fairness evaluation at the end.