Bias Detection Report

Bias Detection Report

Dataset Analyzed:

COMPAS Dataset

Features Examined:

- Categorical Feature: Ethnic_Code_Text

- Numerical Feature: DecileScore

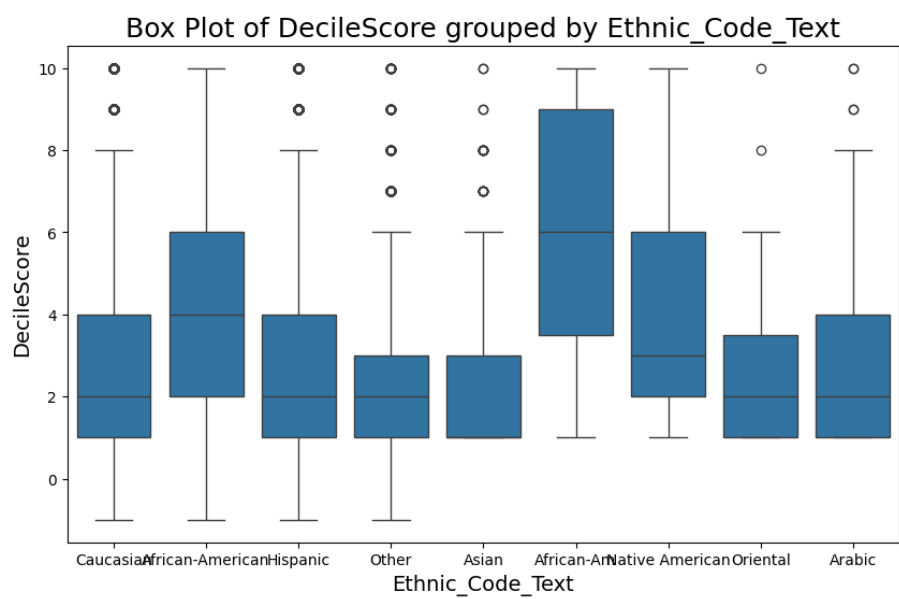### Bias Type:

Correlation

Tools Used and Values Obtained:

- **Hilbert-Schmidt Independence Criterion (HSIC)**: HSIC value = 3.82e-11

- **Max Absolute Mean**: Max N value = 0.995

- **Standardized Difference**: Standardized difference = 1.302

- **Cohen's d**: Not applicable (more than two categories)

Extent of Bias:

- HSIC indicates no significant correlation bias.

- Max Absolute Mean suggests moderate bias.

- Standardized Difference indicates significant bias.

Visualization:

A box plot has been created to show the distribution of decile scores across ethnic groups. It highlights variations among different ethnicities.

Box Plot of DecileScore grouped by Ethnic_Code_Text

Conclusion and Recommendations:

The analysis shows significant bias levels in the correlation between ethnicity and decile scores. It's advised to take into account these biases when using this dataset for decision-making. Further

investigation or data adjustment may be necessary to ensure fairness.