Bias Detection Report

Gender and Income Bias Analysis Report

Introduction

This report explores the potential pay gap between men and women using the Adult dataset. The analysis focuses on the correlation between the 'sex' and 'income' features.

Data Understanding and Preparation

The dataset includes features such as age, workclass, fnlwgt, education, marital-status, occupation, relationship, race, sex, capital-gain, capital-loss, hours-per-week, native-country, and income. The analysis specifically focuses on 'sex' and 'income'.

Statistical Analysis

- **Chi-Square Test**: Revealed a statistically significant association between 'sex' and 'income' with a Chi-Square statistic of 1517.81 and a p-value of 0.0.

- **Cramér's V**: Resulted in a value of 0.216, indicating a weak to moderate association.

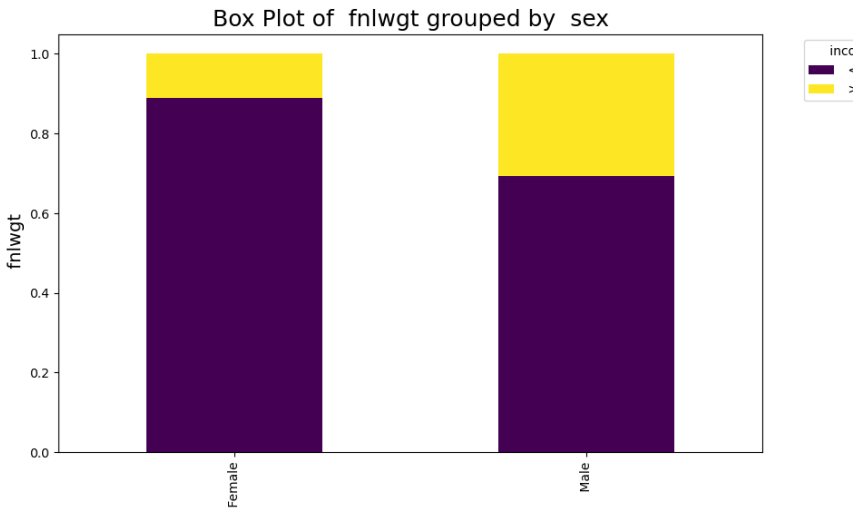
- **Elift**: Calculated as 1.270, suggesting a minor bias.

- **Statistical Parity**: Z-value of 0.307 indicates moderate statistical parity.

Visualizations

- **Stacked Bar Chart**: Demonstrates the distribution of income categories across genders.

- **Box Plot**: Shows the distribution of 'fnlwgt' across genders.



Conclusion

The analysis indicates a weak to moderate level of bias in income distribution across genders. Statistical tests and visualizations suggest a disparity, although not extreme.

Bias Level

Based on statistical findings, the bias level is categorized as 'Moderate Bias'.

Recommendations

Further investigation may be warranted to explore underlying causes of income disparity and potential corrective actions.