Bias Detection Report

Housing Distribution Bias Analysis Report

Dataset Analyzed:

- Statlog.csv

Feature of Interest:

- Housing

Type of Bias Detected:

- Distribution Bias

Analysis Methods and Results:

Toolset Methods:

1. Shannon Entropy and Balance Metric:

   - Balance: 0.7186

- Shannon Entropy: 1.139

- Bias Level: Moderate Bias

2. Max/Min Ratio:

- Ratio: 6.602

- Bias Level: Moderate Bias

3. Gini Index:

- Corrected Gini Index: 0.4488

- Adjusted Gini Index: 0.6732

- Bias Level: Moderate Bias

4. Relative Risk:

- Normalized Risk Score: 6.602

- Bias Level: Moderate Bias

Reference Literature Methods:

1. Chi-Square Test:

- Chi-Square Statistic: 299.1655

- p-value: 1.089e-65

- Conclusion: Extremely significant deviation from uniform distribution.

- Bias Level: Extreme Bias

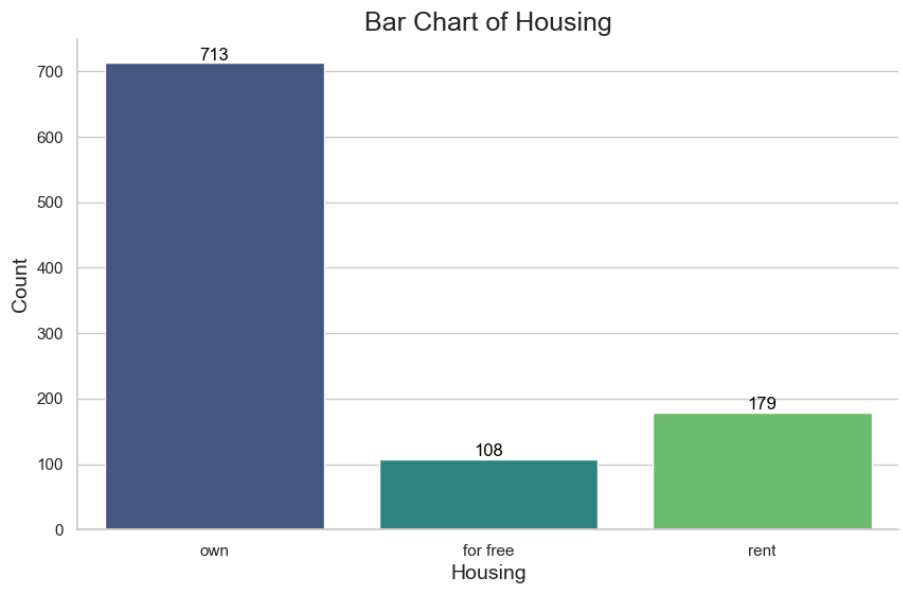
2. Kolmogorov-Smirnov Test:

- KS Statistic: 0.3333

- p-value: 1.0

- Conclusion: No significant deviation from reference CDF.

- Bias Level: No Bias



Conclusion and Recommendations:

The Housing feature shows moderate distribution bias. The results from multiple tools indicate that while the bias does not fully compromise the dataset's usability, it should be considered in certain applications. The Chi-Square test indicates extreme bias, which might affect the dataset's reliability. However, the Kolmogorov-Smirnov test suggests no significant bias. Therefore, it is advised to give these results additional thought before proceeding with analysis or modeling involving the Housing feature.