

# Bias Detection Report

## Bias Detection Report: Loan Purpose Distribution

**Dataset Analyzed:** Statlog.csv

**Feature Examined:** Purpose

**Bias Type:** Distribution

The purpose of this analysis was to determine if one loan purpose dominates the dataset, thereby introducing distribution bias.

### Analysis Results:

#### 1. Shannon Entropy and Balance Metric Analysis:

- Balance: 0.803
- Shannon Entropy: 2.667
- **\*\*Interpretation:\*\*** Indicates minimal bias with a fairly balanced and uncertain distribution.

## **2. Max/Min Ratio Analysis:**

- Ratio: 31.111
- **Interpretation:** Shows extreme bias, with significant disparity between the most and least frequent categories.

## **3. Entropy and Normalized Entropy Analysis:**

- Entropy: 2.667
- Normalized Entropy: 0.803
- **Interpretation:** Moderate bias, indicating some unevenness in category distribution.

## **4. Gini Index Analysis:**

- Adjusted Gini Index: 0.902
- **Interpretation:** Minimal bias, suggesting relatively even distribution.

## **5. Relative Risk Analysis:**

- Normalized Bias Score: 31.111
- **Interpretation:** Extreme bias, highlighting significant differences in observed versus expected frequencies.

## 6. Kolmogorov-Smirnov Test:

- K-S Statistic: 0.987
- P-Value: 0.0
- **\*\*Interpretation:\*\*** Extreme bias, strong indication of deviation from expected distribution.

## Visualizations:

- Visualizations such as bar charts and pie charts could further illustrate the distribution biases observed.

## Recommendations:

Given the results, the distribution of loan purposes is highly skewed, with certain purposes dominating the dataset. To address this bias, consider using re-sampling techniques or stratification to ensure a more balanced feature representation. This is crucial for analyses or models that rely on the 'Purpose' feature to prevent skewed results and ensure fair predictive performance.