Bias Detection Report

Bias Detection Report

Dataset Analyzed:

- **Path**: `source_files/Statlog.csv`

- **Features Examined**:

  - Number of people being liable to provide maintenance for (Dependents)

  - Credit risk (Good/Bad)

Types of Bias Detected:

- **Correlation Bias**: Between two categorical features.

Tools and Methods Used:

- **Chi-Square Test**: Utilized Python's `scipy.stats.chi2_contingency()` function to assess the correlation.

Results Obtained:

- **Chi-Square Statistic**: 0.0

- **p-value**: 1.0

- **Degrees of Freedom**: 1

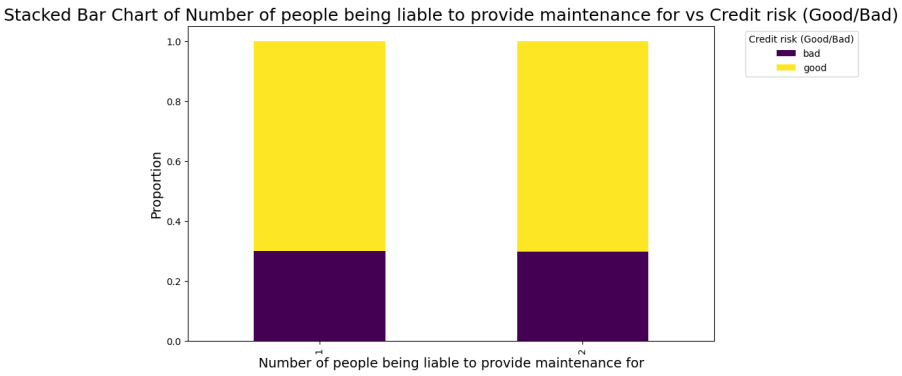
- **Expected Frequencies**:

  - For Dependents = 1, Credit Risk (Good): 253.5, (Bad): 591.5

  - For Dependents = 2, Credit Risk (Good): 46.5, (Bad): 108.5

Visualization:

- A stacked bar chart was created to visualize the distribution of credit risk outcomes based on the number of dependents.



Additional Explanation:

The analysis using the Chi-Square test indicates that there is no significant correlation between the number of dependents and credit risk in the dataset. The p-value of 1.0 shows that the null hypothesis (no association) cannot be rejected. Thus, the number of dependents does not appear to influence credit risk outcomes based on this dataset.