

# Bias Detection Report

## Language Distribution Bias Analysis Report

This report provides a detailed analysis of the distribution bias in the 'Language' feature within the COMPAS dataset. The analysis includes results from various statistical tests and visualizations to offer a comprehensive understanding of any potential biases.

### 1. Chi-Square Test Results

- **Chi-Square Statistic**: 39671.96
- **p-value**: 0.0
- **Interpretation**: The test indicates a highly significant deviation from a uniform distribution, suggesting a strong distribution bias in the 'Language' feature.

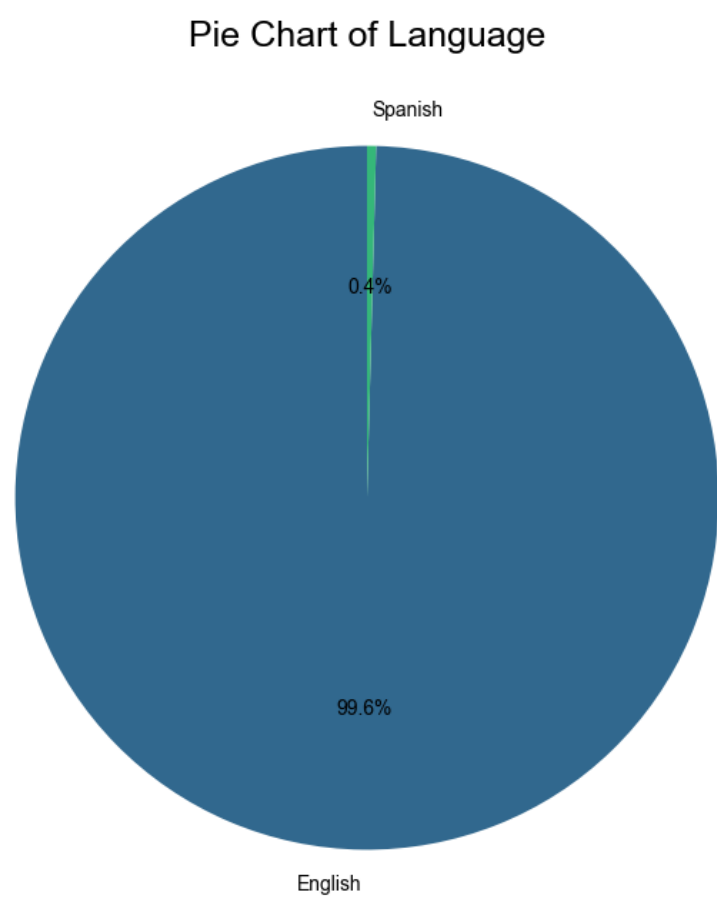
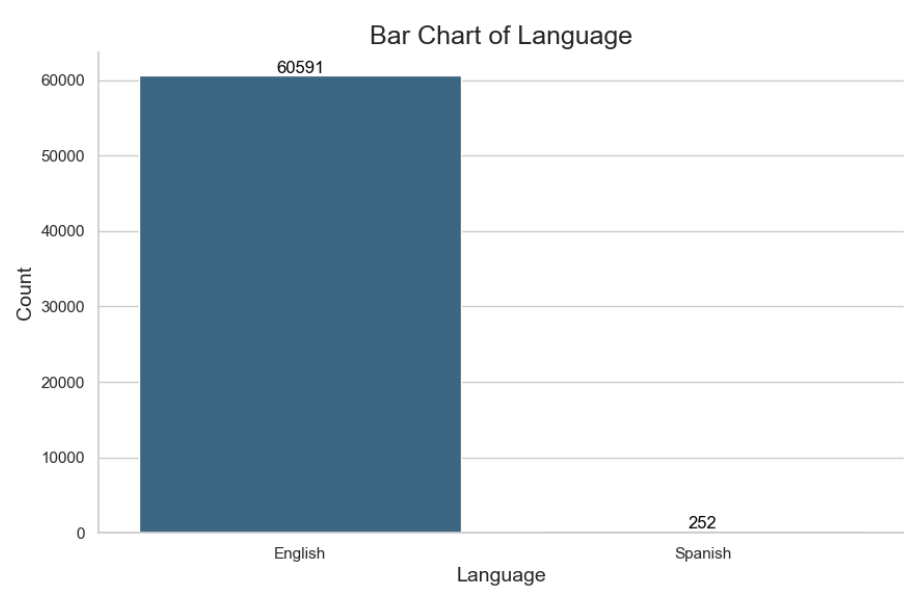
### 2. Kolmogorov-Smirnov Test Results

- **K-S Statistic**: 0.5
- **p-value**: 1.0
- **Interpretation**: The test shows no significant deviation from a uniform distribution, indicating no apparent bias according to this method.

### 3. Jensen-Shannon Divergence Results

- **Jensen-Shannon Divergence**: 0.452
- **Interpretation**: The divergence suggests a moderate distribution bias, indicating some deviation

from a uniform distribution.



4. Visualizations

The bar and pie charts provide a visual representation of the language distribution, highlighting any noticeable trends or dominant languages. These visuals complement the statistical analysis and offer an intuitive understanding of the data distribution.

## **Conclusion**

The analysis presents mixed results, with the Chi-Square Test indicating strong bias, while the Kolmogorov-Smirnov Test shows no bias, and the Jensen-Shannon Divergence indicates moderate bias. The visualizations further illustrate the distribution, providing context to the statistical findings.