

## Dataset Statistical Analysis Report

### Introduction

This report provides a comprehensive statistical analysis of a dataset that includes various traits and demographic information. The goal is to identify potential biases in the perception of traits based on race and gender. We applied various statistical tests, including ANOVA, t-tests, and the four-fifth rule, to determine significant biases.

### Methodology

**One-Way ANOVA Test (Race and Traits):** The one-way ANOVA test was used to compare the means of various traits across different racial groups. The test checks if there are any statistically significant differences between the means of three or more independent (unrelated) groups. Our analysis included racial groups identified as White, Black, and East Asian.

**T-Test (Gender and Traits):** The t-test was used to compare the means of various traits between two groups determined by gender. This test assesses whether the means of two groups are statistically different from each other.

**Identification of Top 5 Biases:** We ranked the biases by their p-values to identify the most significant biases in trait perception based on race and gender. The lower the p-value, the more statistically significant the bias is, indicating a stronger evidence against the null hypothesis of no effect.

**Adverse Impact Analysis (Four-Fifth Rule):** The four-fifth rule was applied to identify any adverse impacts on specific races. It determines if the selection rate for any disadvantaged group is less than four-fifths (80%) of that for the group with the highest rate, suggesting potential discrimination.

### Results

**ANOVA Results (Race and Traits):** The ANOVA test revealed that the perception of the trait 'Intelligent' had the most statistically significant difference across races, with a p-value of  $1.67 \times 10^{-15}$ . This indicates a strong bias in the perception of intelligence across different races.

**T-Test Results (Gender and Traits):** The t-test showed an extremely significant bias in the perception of 'Emotional', with a p-value of  $1.03 \times 10^{-164}$ , indicating a substantial difference in how emotional traits are perceived between genders.

#### Top 5 Significant Racial Biases:

1. Intelligent (p-value:  $1.67 \times 10^{-15}$ )
2. Aggressive (p-value:  $2.71 \times 10^{-11}$ )
3. Responsible (p-value:  $8.01 \times 10^{-11}$ )
4. Egotistic (p-value:  $1.44 \times 10^{-7}$ )
5. Trustworthy (p-value:  $3.67 \times 10^{-6}$ )

**Top 5 Significant Gender Biases:**

1. Emotional (p-value:  $1.03 \times 10^{-164}$ )
2. Caring (p-value:  $6.75 \times 10^{-105}$ )
3. Egotistic (p-value:  $5.72 \times 10^{-103}$ )
4. Aggressive (p-value:  $6.29 \times 10^{-101}$ )
5. Kind (p-value:  $3.68 \times 10^{-84}$ )

**Races with Adverse Impact:** Racial groups 0, 2, 4, and 5 were found to have adverse impacts, as determined by the four-fifth rule.

**Interpretation**

The analysis suggests that there are significant biases in the perception of traits based on both race and gender. The 'Intelligent' trait is perceived very differently across racial lines, which may reflect stereotypes or systemic biases. Similarly, 'Emotional' and 'Caring' traits showed massive gender biases, possibly reflecting traditional gender roles. It's critical to recognize these biases to address them effectively, promoting fairness and equality in social perceptions.

**Conclusion**

The statistical analyses reveal deeply ingrained biases that can have profound implications in various sectors, including employment, education, and social services. By identifying and understanding these biases, we can work towards mitigating their effects and fostering a more equitable society.