# Regression Analysis for the South Bronx

The regression analysis focused on the South Bronx neighborhoods, analyzing the relationship between the Flood Social Hazard Risk Index (FSHRI) and racial demographic proportions. Key findings and insights are summarized below.

## 1. Model Overview

- Dependent Variable: Mean FSHRI (Flood Risk Index).

- R-squared = 1.000: The model perfectly explains the variation in FSHRI within the dataset. This is likely due to the small number of observations (5 NTA2020 values).

- Adjusted R-squared = NaN: The small sample size leads to undefined adjustments for the number of predictors.

- F-statistic and Prob (F-statistic): Both are undefined (NaN), suggesting the model lacks sufficient degrees of freedom for a robust statistical test.

## 2. Statistically Significant Variables

- None of the variables show standard errors or t-statistics due to the limited sample size and perfect collinearity.

However, the coefficients suggest possible relationships:

- White Alone (coef = 12.4580): Positive association with FSHRI.

- Black or African American Alone (coef = 5.2821): Moderate positive association.

- Some Other Race Alone (coef = 3.4343): Identified as the racial group most at risk.

## 3. Potential Issues

- Small Sample Size: The dataset includes only five neighborhoods, limiting the model’s ability to generalize findings.

- Multicollinearity: The condition number (Cond. No. = 283) indicates moderate multicollinearity, where predictors may be highly correlated.

- Perfect Fit (R-squared = 1.000): This is unrealistic in most real-world data and suggests potential overfitting due to the small dataset or highly correlated predictors.

## 4. Key Findings

- The Racial Group Most at Risk: 'Some Other Race Alone' is identified as the group most associated with higher flood risk in the South Bronx.

- High Positive Coefficients: Other groups, such as 'White Alone' and 'Black or African American Alone,' also show positive coefficients but are less prominent.