# Quantitative Methodology and Findings

## Methodological Summary

This study investigates the relationship between racially motivated zoning, heavy industrial activity, and contemporary climate-change-driven flood risks in the South Bronx and New York City. A multifaceted approach was employed, combining data acquisition, integration, statistical modeling, and visual analysis to elucidate the intersection of environmental vulnerability and systemic inequities.

## Data Acquisition and Preparation

1. Data Sources:  
- Flood Risk: Flood Hazard Surface Risk Index (FSHRI) and storm surge/tidal risk scores for present, 2050s, and 2080s.  
- Demographics: Neighborhood Tabulation Areas (NTA) data, including race/ethnicity composition, age, and poverty indicators.  
- Land Use: NYC PLUTO dataset for zoning and land-use characteristics.

2. Data Cleaning:  
- Removal of duplicate rows and unnecessary columns (e.g., "Unnamed" values).  
- Harmonization of column names to ensure alignment (e.g., NTA2020 identifiers).  
- Flood risk scores consolidated to focus on the FSHRI metric.  
- Demographic data, specifically racial proportions, merged with flood risk scores at the neighborhood level.

3. Data Standardization and Filtering:  
- Normalized industrial zoning (zonedist1 starting with “M”) into binary variables.  
- Aggregated data by NTA codes, including flood risk, racial proportions, and industrial zoning metrics.  
- Subset the data into South Bronx neighborhoods and the rest of New York City for comparative analysis.

## Analytical Methods

1. Descriptive Statistics:  
- Computed average flood risk and racial composition across NTAs, with a focus on the South Bronx.

2. Visual Analysis:  
- Stacked Bar Charts: Visualized racial group exposure to flood risk by neighborhoods (e.g., Hunts Point and Mott Haven).  
- Correlation Matrices: Illustrated relationships between industrial zoning, racial composition, and flood risks.

3. Statistical Modeling:  
- Multiple Linear Regression:  
 - Dependent Variable: Flood Risk (FSHRI).  
 - Independent Variables: Proportions of major racial groups and socioeconomic indicators (% White, % Black/African American, % Asian, % Other, % 65+ years, poverty rates).  
- Regression models tested relationships between zoning, racial composition, and flood risks, providing insights into systemic inequities.

## Regression Results

1. Negative Correlation for White Populations:  
- Coefficient: -1.0644 (p = 0.003).  
- Higher proportions of White residents were associated with lower flood risk scores, reflecting historical patterns of environmental privilege.

2. Positive Correlation for Asian and Other Populations:  
- Asian alone: Coefficient = 3.0629 (p < 0.001).  
- Some Other Race alone: Coefficient = 2.6895 (p < 0.001).  
- Neighborhoods with larger Asian and Other race populations exhibited significantly higher flood risks, often near industrial zones and waterfronts.

3. Insignificant Relationship for Black Populations:  
- Coefficient: 0.2583 (p = 0.478).  
- The relationship between flood risk and Black/African American populations was not statistically significant, possibly due to heterogeneous spatial distributions.

4. Zoning and Industrial Activity:  
- Preliminary analysis showed a strong positive correlation between industrial zoning and flood risk, particularly in marginalized neighborhoods.

## Python Code Summary

Below is the Python code snippet used for data cleaning, aggregation, regression analysis, and visualization:

import pandas as pd  
import seaborn as sns  
import matplotlib.pyplot as plt  
import statsmodels.api as sm  
  
# Load data  
data = pd.read\_csv("/Users/ermaswartz/Documents/methods\_data/CSVs/plutoanddemographic.csv")  
  
# Normalize zoning  
data['industrial'] = data['zonedist1'].str.startswith('M').astype(int)  
  
# Aggregate by NTA  
nta\_analysis = data.groupby('ntacode').agg(  
 industrial\_proportion=('industrial', 'mean'),  
 mean\_flood\_risk=('FSHRI', 'mean'),  
 total\_population=('Estimate Total:', 'sum'),  
 \*\*{col: (col, 'sum') for col in [  
 'Estimate Total: White alone',  
 'Estimate Total: Black or African American alone',  
 'Estimate Total: Asian alone',  
 ]}  
).reset\_index()  
  
# Racial proportions  
racial\_cols = ['Estimate Total: White alone', 'Estimate Total: Black or African American alone', 'Estimate Total: Asian alone']  
for col in racial\_cols:  
 nta\_analysis[col + '\_proportion'] = nta\_analysis[col] / nta\_analysis['total\_population']  
  
# Regressions and Correlation Heatmaps  
def regression\_analysis(data, label):  
 print(f"=== {label} ===")  
 for col in racial\_cols:  
 X = sm.add\_constant(data[['industrial\_proportion']])  
 y = data[col + '\_proportion']  
 model = sm.OLS(y, X).fit()  
 print(model.summary())  
   
 plt.figure(figsize=(10, 8))  
 sns.heatmap(data[['industrial\_proportion'] + [col + '\_proportion' for col in racial\_cols]].corr(), annot=True)  
 plt.title(f"Correlation Matrix: {label}")  
 plt.show()  
  
regression\_analysis(nta\_analysis, "South Bronx")

## Conclusion and Policy Implications

The study reveals critical disparities in flood risk exposure driven by zoning policies and racial demographics. Key takeaways include:  
1. Mitigation Strategies: Target investments in high-risk neighborhoods, particularly in the South Bronx.  
2. Environmental Justice: Address systemic inequities by integrating racial and socioeconomic equity into resilience planning.  
3. Policy Interventions: Restrict industrial zoning near vulnerable communities and prioritize flood mitigation infrastructure.  
  
By combining robust quantitative analysis with clear visualizations, this research highlights actionable pathways for equitable urban resilience.