# Multivariate Patterns in Homicide Cases: Demographics, Weapons, and Crime Solvability

**Parth Maniar**

## Research Question

Through the use of multivariate statistical tools, this research examines trends in homicide incidence throughout the states of the United States. It supports law enforcement, policymakers, and social scientists by offering insights into crime resolution through data-driven approaches. Investigating the correlations between victim-perpetrator relationships, weapon types, and demographic factors (such as age, sex, and race) and the solvability of homicide cases is the aim.

Key questions include:  
- What demographic and weapon-related factors are most predictive of whether a homicide is solved?  
- Can homicide incidents be grouped into meaningful clusters based on their characteristics?  
- Are there latent variables underlying certain patterns in solved vs. unsolved homicide cases?

## Dataset Description

The dataset includes detailed incident-level homicide data reported by municipal and state law enforcement agencies across the U.S. Sample data includes attributes such as:  
- Year, Month, State, City  
- Crime Type, Crime Solved, Weapon Used, Relationship between victim and perpetrator  
- Victim and Perpetrator demographics: Sex, Age, Race, Ethnicity  
  
**Data Source:** <https://www.kaggle.com/datasets/murderaccountability/homicide-reports/data>

## Planned Statistical Methods

The following multivariate techniques will be applied:  
- **Principal Component Analysis (PCA)** to reduce dimensionality and identify dominant crime features  
- **Clustering (K-means or hierarchical)** to categorize cases by type and demographic/weapon patterns  
- **Factor Analysis** to identify latent behavioral or environmental patterns  
- **Classification** (e.g., logistic regression, decision trees) to predict solvability based on input variables  
- **Correlation and hypothesis testing** to analyze key variable relationships (e.g., gender vs. weapon type)