

# Advanced Computational Econometrics

## Final problem

The `Communities and crime` dataset contains socio-economic data from the 1990 US Census, law enforcement data, and crime for 1994 communities in the United States. The objective is to predict the `crime` variable (total number of violent crimes per 100K population) using the other variables, and to understand the influence of the most important predictors.

1. Preprocess the data : remove non-predictive variables and variables with missing data.
2. Using all the methods studied in this course, find a prediction function with minimal test error. What are the most important predictors? What is their influence on the response?