

Project Description

Research Question

Is it possible to predict the outcome of a police stop-and-search based on geographic and temporal data combined with personal characteristics of the suspect in question?

In order to answer this question, we use data gathered through the API of the UK police to train a logistic regression model. The objective is to predict whether or not a given search is “successful” in the sense that the suspect is, indeed, found to be in possession of illegal goods or exhibiting other unlawful behaviour.

Discussion of ethical implications

We discuss a variation of the conundrum posed by the movie [Minority Report \(2002\)](#): is it ethically defensible to use prediction models in criminal investigations?

A simple prediction model leaves us unable to uncover the underlying causal mechanisms at work in the relation between personal characteristics and crime. Implementing biased models as neutral instruments in police work may compromise the rights of individual citizens.

Data

The majority of our data is collected through the UK police API, found [here](#). By transforming geospatial information found in this main dataset, we merge it with external data on UK local authorities in order to add a mean income variable for all areas.