

International Journal of Population Data Science

Journal Website: www.ijpds.org



Integrating Supermarket Behavioural Data with Established Indicators to Improve Alcohol-Related Mortality Prediction in England

Raphael Derecki¹, James Goulding¹, Elizabeth Dolan¹, and Brian O'Shea²

¹University of Nottingham, N-LAB, United Kingdom

²University of Nottingham, Department of Psychology, United Kingdom

Introduction & Background

Within behavioural sciences, there has been a resurgence in not only explaining factors but also predicting them with precision that exceeds classic social science standards. In England (UK), alcohol is a ubiquitously used substance with cultural significance but also many potential issues related to consumption. High levels of alcohol usage are associated with a range of psychosocial issues including addiction, criminality, poverty, and ultimately death.

often criticized for their potentially low interpretability, particularly with more complex model classes, we implement a range of explainable AI techniques to elucidate variable importance.

Keywords

Alcohol misuse, Consumer Behaviour, Public Health, Epidemiology, Geospatial Analysis

Objectives & Approach

While our ultimate aim is to examine and predict problematic alcohol usage (i.e., consumption of more than 35 units per week), our preliminary results examine alcohol-related deaths as a simpler dependent variable to assess the robustness of our approach.

Using modern data scientific techniques, we provide a machine learning workflow for predicting alcohol deaths at the lower-tier local authority level using a range of literature-derived variables including psychosocial and environmental demographics, trade sales data, and novel supermarket behavioural customer data (i.e., alcohol purchases). Within this workflow, we discuss the importance of feature creation, selection, and engineering; correct sampling standards (i.e., cross-validation); and model class selection, implementation, and validation.

Relevance to Digital Footprints

This study uses data from a large UK supermarket to quantify alcohol purchasing rates within lower-tier local authority levels.

Conclusions & Implications

Overall, we find that incorporating a range of variables, including novel supermarket behavioural data, enhances the prediction of alcohol deaths. While machine learning methods are

