Troubled Families A Synthetic dataset

January 2026

Introduction

This file describes the synthetic data created using the output from the [National evaluation of the Troubled Families Programme 2015 - 2020: family outcomes – national and local datasets: part 1](#) report.

The experiment and creation of the synthetic data is described [here](#) in Little et al. (2025). The variables are all binary and there are 10586 records

The synthetic data is designed so that it would solely reproduce the information found in Charts 2 and 3 of the report. It is unlikely to be useful for any other analysis.

Data Dictionary

To indicate that the data is synthetic, all of the variable names are prefixed with "synthetic_"

Variable name	Description	Label	Meaning

References

Little, C., Elliot, M., Allmendinger, R. (2025). Producing synthetic teaching datasets using evolutionary algorithms. *UNECE Expert Meeting on Statistical Data Confidentiality.* https://unece.org/sites/default/files/2025-12/SDC2025_Sd_UnivManchester_SyntheticTeachingDatasets_Elliot_D.pdf

Department for Communities and Local Government (2017). National evaluation of the troubled families programme 2015 - 2020: family outcomes - national and local datasets: part 1. Technical report. https://assets.publishing.service.gov.uk/media/5a82d38a40f0b6230269ce8f/Family_outcomes_-_national_and_local_datasets.pdf