Troubled Families A Synthetic dataset

January 2026

Introduction

This file describes the synthetic dataset 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 19 variables are all binary and there are 10586 records. Each row represents a family, and the variables represent the various events or problems that the families have.

The synthetic data is designed so that it would solely reproduce the information found in Charts 2 and 3 of the [National Evaluation](#) 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
synthetic_domestic_abuse	Domestic abuse	0	No
		1	Yes
synthetic_financial_exclusion	Financial exclusion	0	No
		1	Yes
synthetic_children_help	Children needing help	0	No
		1	Yes
synthetic_crime_asb	Crime and anti-social behaviour	0	No
		1	Yes
synthetic_health	Health	0	No
		1	Yes
synthetic_education_attendance	Education and attendance	0	No
		1	Yes
synthetic_police_callouts	Police call-outs	0	No
		1	Yes
synthetic_mental_health	Mental-health issue	0	No
		1	Yes
synthetic_rent_arrears	Rent arrears	0	No
		1	Yes
synthetic_alcohol_drug_depend	Alcohol or drug dependency	0	No
		1	Yes
synthetic_drug_depend	Drug dependency	0	No
		1	Yes

synthetic_asb_no_inv	Anti-social behaviour incident with no further investigation	0	No
		1	Yes
synthetic_alcohol_depend	Alcohol dependency	0	No
		1	Yes
synthetic_asb_inv	Anti-social behaviour incident with further investigation	0	No
		1	Yes
synthetic_neet	Young person not in employment, education or training	0	No
		1	Yes
synthetic_homeless_app	Homelessness applications	0	No
		1	Yes
synthetic_homeless_week	A week or more homeless	0	No
		1	Yes
synthetic_missing_education	Missing from education	0	No
		1	Yes
synthetic_evictions	Evictions	0	No
		1	Yes

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