Anonymised CDM Coursework Dataset

Contained in this directory is a dataset that has been anonymised using various techniques including banding, perturbation, and redaction. The following is a description of each column header and a brief summary of the information contained.

Variable	Description	Type
Sample.ID	Unique ID for each individual	
Gender	M/F (Male/Female)	Binary categorical
Birthyear	Birth year in 20-year brackets	Ordinal
Location.of.Birth	Northern/Southern hemisphere	Binary categorical
UK.Region	England/Other (UK residence) Other includes: Wales, Scotland, Northern Ireland	Binary categorical
Height	Height (cm)	Continuous
Weight	Weight (kg)	Continuous
Education.Level	Higher/BasicOther Higher includes: bachelor, master's, and PhD BasicOther includes: primary, secondary, and other	Binary categorical
Countries.Visited	Number of countries visited	Discrete
Avg.Alcohol	Average number of alcohol units consumed per week	Continuous
Avg.Cigarettes	Average number of of cigarettes smoked per week	Continuous
CC.Status	Case-control status 1 = Wanderlust gene, 0 = no Wanderlust gene	Binary categorical