**Anonymised CDM Coursework Dataset**

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

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