Assignment II for
EP16: Missing Values in Clinical Research
Multiple Imputation
13 − 17 May, 2019

Data

The **MIdat13** data comprise 625 observations of 625 mothers measured repeatedly before and during pregnancy. All women were scheduled to have their weight measured once each trimester and were asked for their pre-pregnancy weight and BMI.

The dataset contains the following variables:

variable	explanation
ID	subject identifier
gender	gender of the child
birthwgt	birthweight in kg
leanfrac	proportion of child's lean mass (lean mass/total mass; lean mass = total mass - fat mass)
parity	number of pregnancies of more than 20 weeks the mother had (nulliparity: this was the first pregnancy, multiparity: mother had previous pregnancies)
sun	average sun light duration in hours/day in the month before blood sampling
birth	date of birth
BMC	bone mineral content of the child in grams, determined by DXA scan
vitD	mother's serum vitamin D concentrations in 10 nmol/L
length	child's length at time of DXA scan in centimeters
ethn	child's ethnicity
season	season of blood sampling
lean	lean mass in kg
weight	child's total weight in kg at DXA scan
sun_birth	average sun light duration in the month before birth in hours/day
sports	does the child do sports regularily?

Analysis model of interest

The analysis model of interest is a linear regression with outcome BMC, and covariates `vitD`, `ethn`, `gender`, `leanfrac`, `sports`, `sun`, `season`, `length` and `weight`.

We assume that `vitD` has a non-linear (quadratic) effect.