AssignmentB

Martijn Koster, William Schaafsma, Martijn van Dam, Victor Hovius

3/21/2022

Contents

- verhaaltje over wat we gaan doen
- daadwerkelijk naive imputatie doen
- imputatie beschrijven (de plots o.a.)
- data invullen n.a.v. imputatie
- data beschrijven
- aangeven wat we willen veranderen voor assignment C

1 Imputations

After observing the data for missing values, we now try to solve the missingness problem by doing multiple imputation with the package `mice`. In order to answer the research question with the imputed data we will follow the main steps in multiple imputation following van Buuren, 2018, shown in Figure 1. In first instance we will use the default settings to impute the missingness, this will be further elaborated in the Default Imputations section. After the the default imputations we will evaluate the quality of the imputations by examining multiple plots about the convergence and the distribution. This will be done in the Evaluating Default Imputations section. After evaluating the default imputations, we will make an outline of how the imputations can be improved, by using more sophisticated imputations. This will be discussed in the Improving the Imputations section.

1.1 Default Imputations

As mentioned before we will first use the default settings of `mice` to impute the missingness. In this section we will describe them. By default, `mice` will produce 5 imputations and 5 iterations. **Say here something about the imputations and iterations (what do they do and why do we use them?)**

Say something about the predictor matrix:

Table 1 gives an overview of the predictor matrix for the imputations.

$

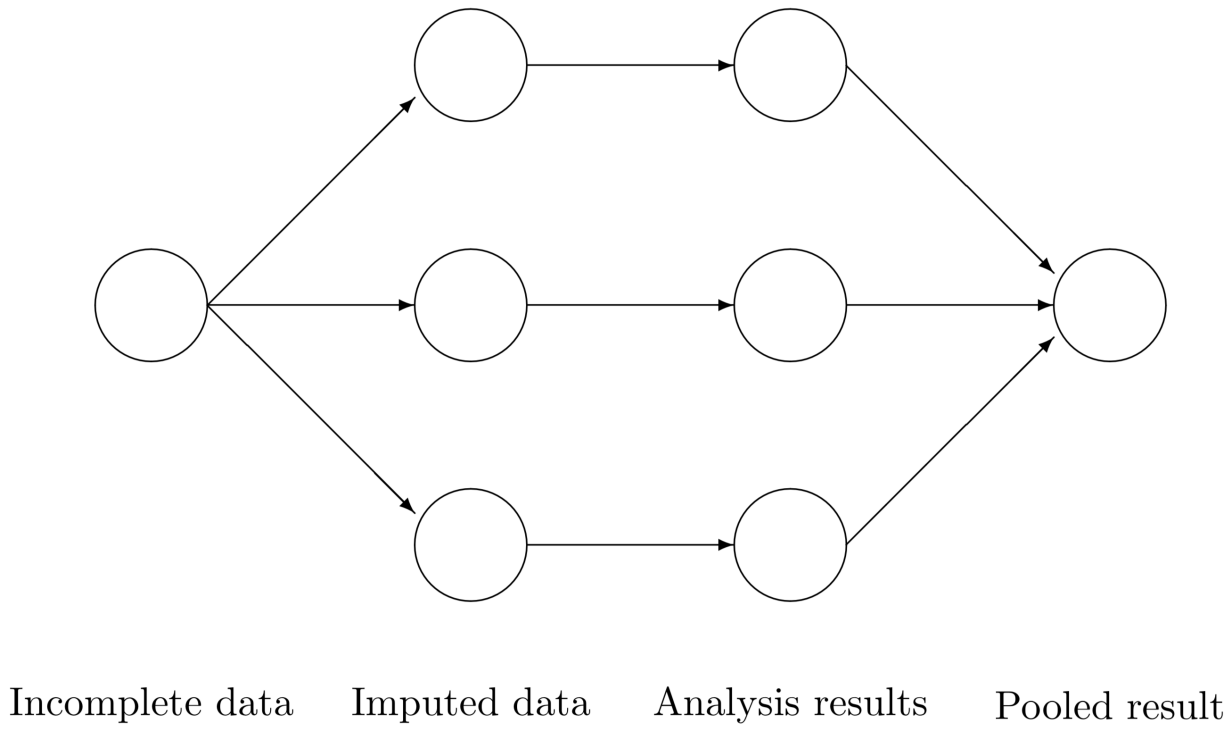


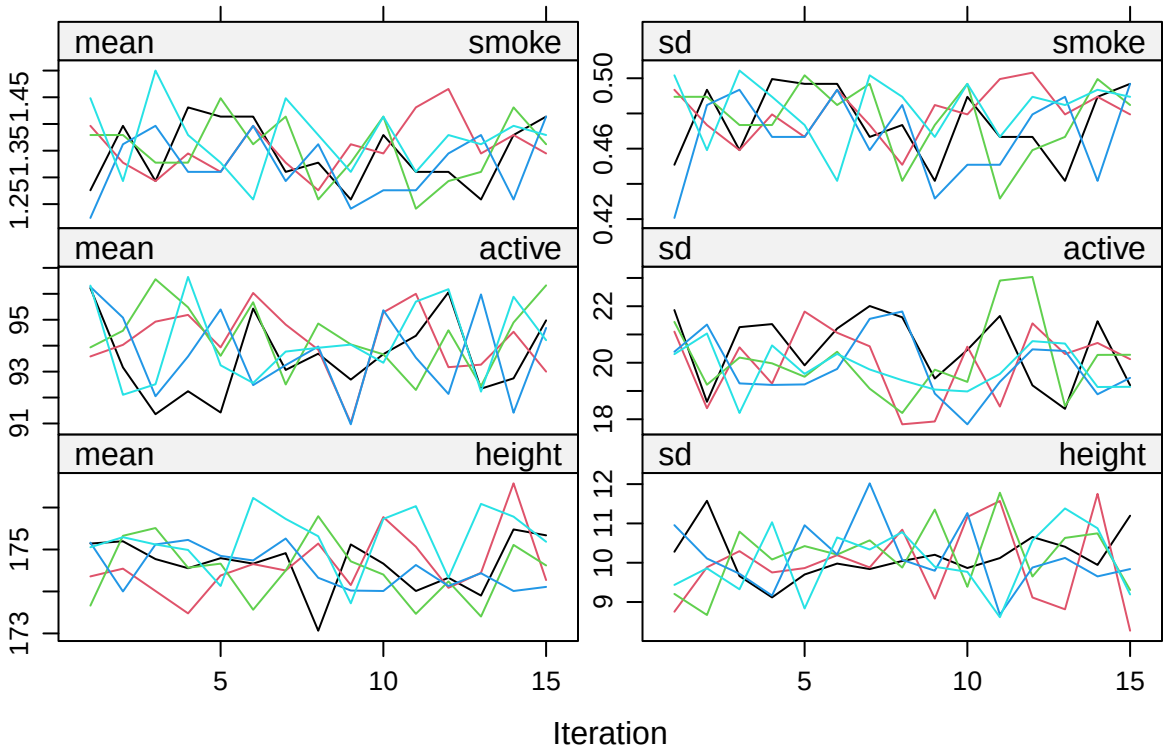
Figure 1: Scheme of main steps in multiple imputation

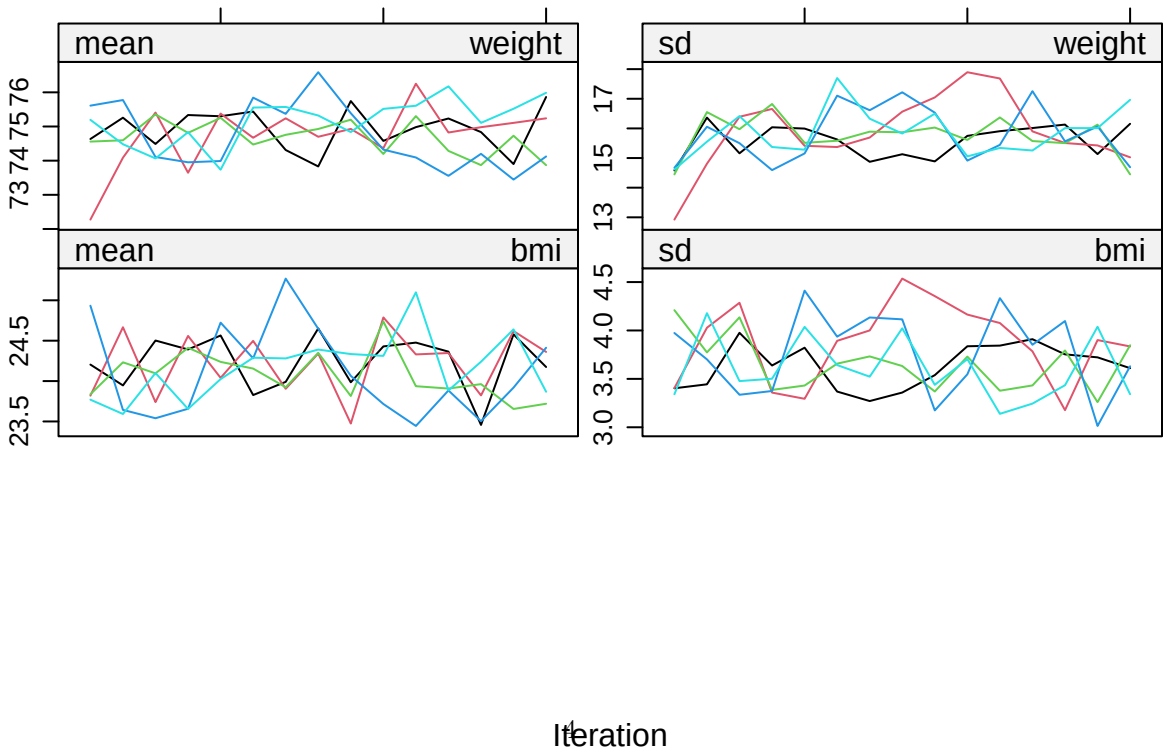
Table 1: Predictor Matrix of Default Imputations

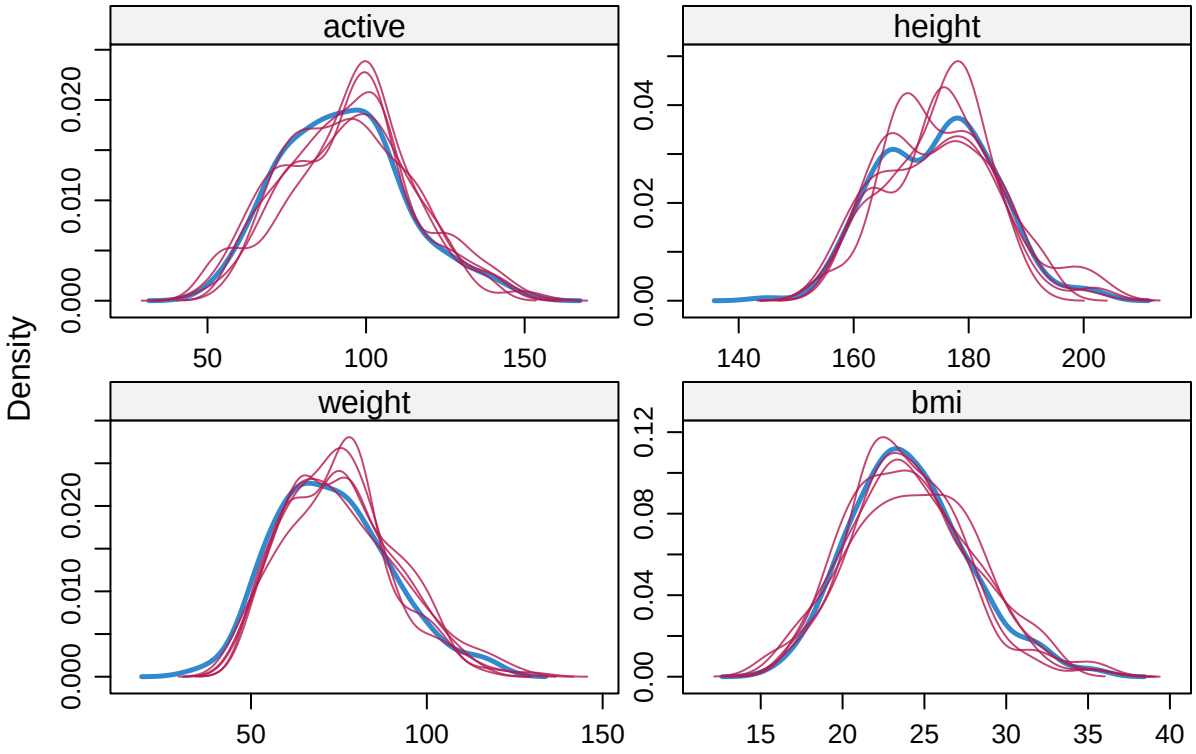
	age	smoke	sex	intensity	active	rest	height	weight	bmi
age	0	1	1	1	1	1	1	1	1
smoke	1	0	1	1	1	1	1	1	1
sex	1	1	0	1	1	1	1	1	1
intensity	1	1	1	0	1	1	1	1	1
active	1	1	1	1	0	1	1	1	1
rest	1	1	1	1	1	0	1	1	1
height	1	1	1	1	1	1	0	1	1
weight	1	1	1	1	1	1	1	0	1
bmi	1	1	1	1	1	1	1	1	0

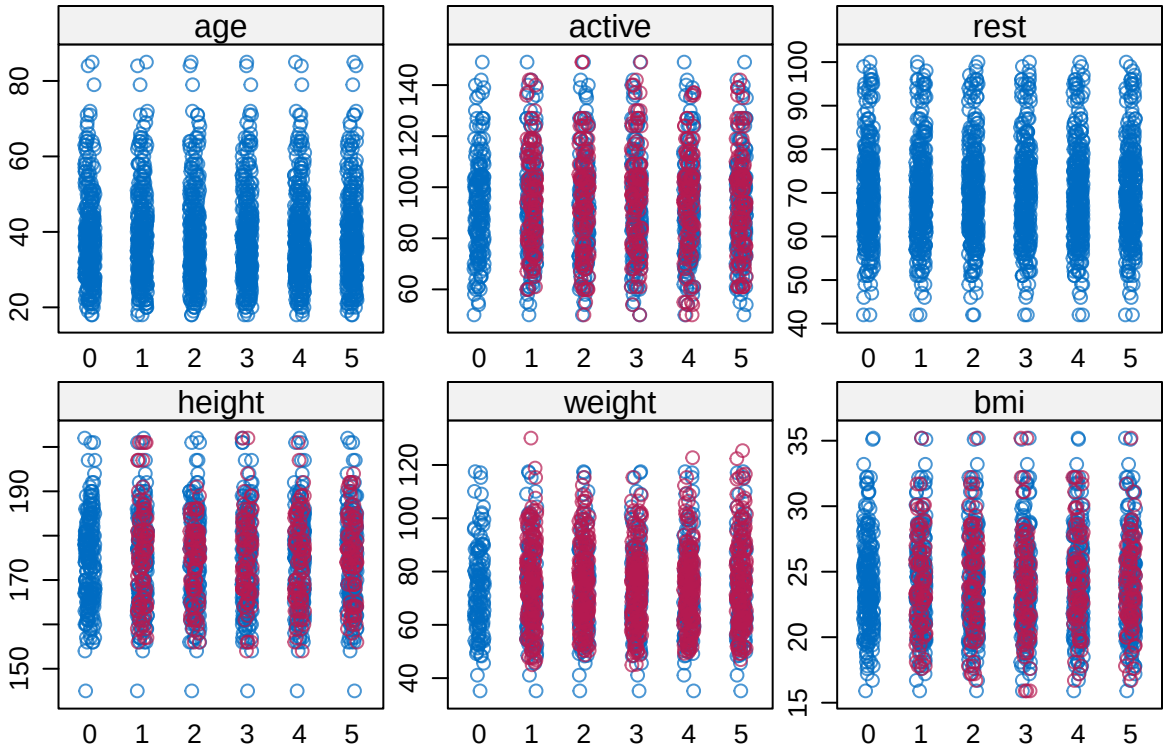
1.2 Evaluating Default Imputations

1.3 Improving the Imputations






active
height
weight
bmi
Density
0.000
0.010
0.020
0.00
0.02
0.04
0.04
0.08
0.12
50
100
150
140
160
180
200
15
20
25
30
35
40


age
active
rest
height
weight
bmi
0 1 2 3 4 5