The data is from the Amsterdam Growth and Health Study (Kemper, 1995), which examines the relationship between lifestyle and health in adolescence and young adulthood. **The first column** is the subject id. **The second column** is the continuous response variable \( Y \) of our interest, which is the total serum cholesterol expressed in mmol/l. The columns 3-6 are four predictors: \( X_1 \) (fitness level at baseline measured as maximal oxygen uptake on a treadmill, **the third column**), \( X_2 \) (body fat estimated by the sum of the thickness of four skinfolds, **the fourth column**), \( X_3 \) (smoking behavior dichotomized as smoking versus non-smoking, 0 = no; 1 = yes, **the fifth column**), and \( X_4 \) (gender, 1 = male; 2 = female, the six column). All variables were measured at \( n = 6 \) time occasions, which is coded from 1 to 6. The time of measurement is taken as an additional predictor (\( X_5 \), **the seventh column**). The columns 8 and 9 are other available dichotomous and categorical response variables. For this project, you can ignore these two columns. Here a total of \( m = 147 \) subjects participated in the study, and the overall sample size is \( N = 147 \times 6 = 882 \).