Exercise 1: Height vs. age in the `Juul`-data

We want to model height as a function of age, with the aim of comparing boys and girls (and maybe also for constructing reference charts, i.e. prediction intervals for each age).

We once again use the data in `juul2.txt` and restrict ourselves to look at the ages 5 to 20.

Questions from earlier exercise:

1. Fit a straight line for male height as a function of age, and look at the fit graphically.
   What is wrong (not surprisingly) with this model?
2. Look at fits of polynomials of second and third order, perhaps after transformation of height. How can we compare these models to similar models for females?
3. You may also want to campare the above fits with the automatic *smoothing*-possibilities in SAS, using '`i=smxx`' in the `symbol`-statement (`xx` denoting a number between 1 and 99, indicating the degree of smoothness).

   Instead of polynomials, we will now fit a piecewise straight line, i.e. a **linear spline**.
4. Introduce cutpoints for a linear spline, e.g. at the ages 10, 12, 13 and 15. Define variables denoting number of years above these respective thresholds and fit a linear spline for boys.
5. At which age do we first see a deviation from a simple straight line?
6. Compare with similar results for the girls.
7. Fit a model for both genders simultaneously and find out at which age we first see a difference between girls and boys.
   Describe this difference, in numbers as well as in words.

turn the page!

New questions:

8. The **Gompertz growth model** takes the following form:

$$\texttt{height} = \alpha \times e^{-\beta \times e^{-\gamma \times \texttt{age}}}$$

How can we obtain starting values for this model? (Hint: start with a guess of α – what happens if `age` is large, assuming that the three parameters are positive? Then divide both sides of the equation with α and take logs twice.)

9. Fit the nonlinear model for each sex separately, and try to evaluate the adequacy of the model (if you got lost in the previous question, just use $\alpha = 200$, $\beta = 1.5$, and $\gamma = 0.15$ for starting values).

10. Any suggestions for comparing males and females?

Exercise 2: Biochemical Oxygen Consumption (BOC)

An earlier exercise was concerned with the oxygen loss (`boc`, in mg/l) as a function of 'number of days sealed' (`days`).

We tried out the formula

`boc`=γ exp(-β/`days`)

by transforming `boc` with the natural logarithm, and we obtained an estimate with confidence limits for the asymptotic oxygen loss, when the number of days of sealing gets very large.

New questions:

1. Perform a nonlinear regression directly based on the exponential expression above and compare the results with those obtained from the linear regression.
2. Which type of analysis is to be preferred in this situation?