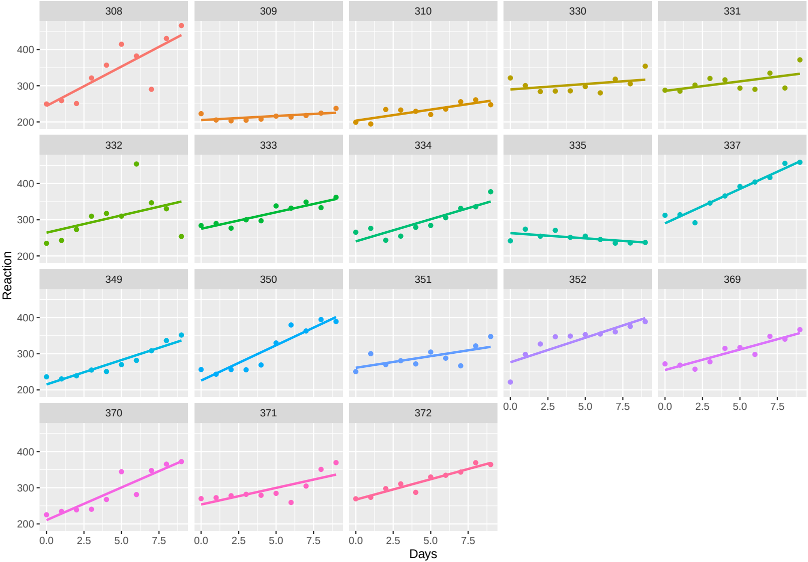
Multilevel linear Models

Mark Andrews
Psychology Department, Nottingham Trent University

mark.andrews@ntu.ac.uk

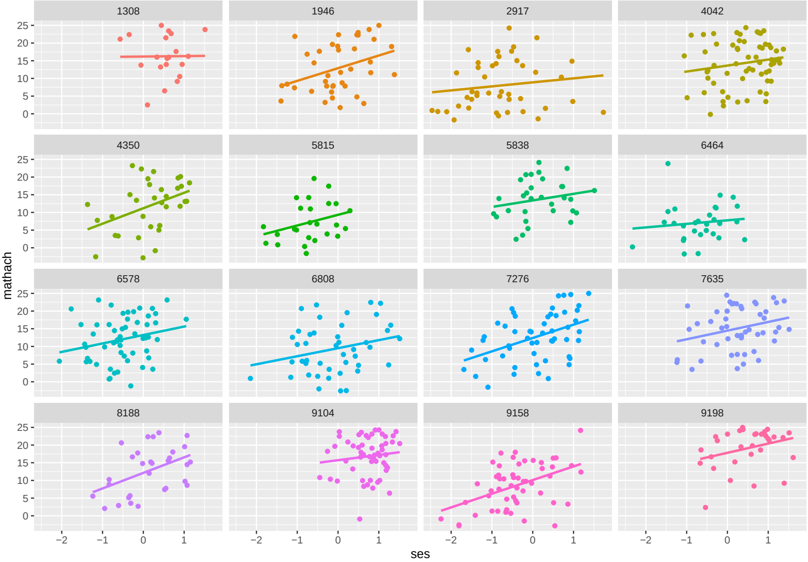
Multilevel data: Example 1

Reaction time as a function of sleep deprivation.



Multilevel data: Example 2

Mathematical achievement as function of socio-economic status.



Example: Reaction time and math achievement

- In this problem, we have J subject. For subject j, we have n_j data points.
- In observation i from subject j, their number of days without sleep is x_{ji} and the reaction time is y_{ji}.
- A multilevel model for this data is

$$
\begin{aligned}
y_{ji} &\sim N(\alpha_j + \beta_j x_{ji}, \sigma^2), \\
\alpha_j &\sim N(a, \tau_a^2), \\
\beta_j &\sim N(b, \tau_b^2).
\end{aligned}
$$

Example: Reaction time and math achievement

- The model

$$
\begin{aligned}
y_{ji} &\sim N(\alpha_j + \beta_j x_{ji}, \sigma^2), \\
\alpha_j &\sim N(a, \tau_a^2), \\
\beta_j &\sim N(b, \tau_b^2),
\end{aligned}
$$

can be re-written

$$y_{ji} = \underbrace{(a + \eta_j)}_{\alpha_j} + \underbrace{(b + \zeta_j)}_{\beta_j} x_{ji} + \epsilon_{ji},$$

or

$$y_{ji} = \underbrace{a + bx_{ji}}_{\text{Fixed effect}} + \underbrace{\eta_j + \zeta_j x_{ji}}_{\text{Random effect}} + \epsilon_{ji},$$

where

$$\eta_j \sim N(0, \tau_a^2),\ \zeta_j \sim N(0, \tau_b^2),\ \epsilon_j \sim N(0, \sigma^2).$$

Example: Reaction time and math achievement

- In the model just described, a and b are the general regression coefficients.
- The variance τ_a^2 tells us how much variation in the intercept term there is across schools. The variance τ_b^2 tells us how much variation in the slope term there is across schools.
- For example, 95% and 99% of the intercepts for individual schools will be in the ranges

$$a \pm 1.96 \times \tau_a, \quad a \pm 2.56 \times \tau_a,$$

respectively. Likewise, 95% and 99% of the slope terms for schools will be in the ranges

$$b \pm 1.96 \times \tau_b, \quad b \pm 2.56 \times \tau_b.$$