Homework 5 **Soc 723**
Multilevel Models

Unless otherwise specified, all of these exercises use the `ess5000.rds` data. I have listwise deleted missing data for all variables except `incdecile` [income decile]. I left this in for those who want to experiment more with missing data.

For these exercises, we're going to focus on the outcome `ppltrst`, which is a 0–10 scale of how much "most people can be trusted." After you've finished these exercises, you can try experimenting with other outcome variables in the dataset (e.g., `euftf` [opinion about European unification] or `vote` [1 if the respondent voted in the last national election, 0 otherwise]).

Preliminaries

1. Describe the data using `psych::describe()`. What are the mean and variance of `ppltrst`?
2. Make a histogram of `ppltrst`. What is the modal value?
3. Plot the distribution of `ppltrst` by country and the distribution of country means. Pay attention to the number of observations by country. What do you expect will happen to the country means if we estimate them with partial pooling?

Single dimension

4. Estimate a variance components model (partially pooled intercepts) with `ppltrst` as the outcome and random intercepts at the country level.
5. Use the estimates from the partial pooling model to manually calculate the intraclass correlation.
6. Use `performance::icc()` to check your calculations. They should be very close.
7. Now interpret this ICC value. What does it mean?
8. Compare the partially pooled random intercepts model to a completely pooled OLS model that ignores the level-2 structure. Do we need the complexity of random intercepts?
9. Get the empirical Bayes estimates of the random effects from the model above. Which country has the highest estimated trust? Which has the lowest? Plot the random intercepts estimates with their error bars. Compare them to the no pooling estimates to see if the predictions you made in question 3 were right.

Two dimensions

10. Rescale the `eduyrs` variable so that it ranges from 0 to 1. Add the predictor `eduyrs01` to the model with partially pooled intercepts. What is the effect of education on trust?
11. Plot the predicted values of trust by education, including the random intercepts for each country.
12. Now estimate a model with a random slope on `eduyrs01`. Plot the predictions.
13. Does this improve model fit over the model without random slopes? What does this mean?

Multivariate model

14. Rescale the other predictors (`agea`, `attend`, `health`) and add them and `female` to the model. Use a quadratic specification for age. Use no random slopes. Plot the effect of age on trust holding other covariates at their means.