Comments on homework, Simple Linear Regression

- The null for the statistical test is $\beta = 0$ or flat line, be careful when phrasing it as difference (only difference from zero).
- The purpose of doing three-way contingency tables is to look for potential confounding.
- Check sample answers.

Multiple Linear Regression

- Allows to control for other independent (explanatory) variables.
- Reflects partial association between the explanatory variables and the response (dependent) variable, holding other variables constant.
- $Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i$
- $\hat{Y} = \hat{\alpha} + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}$

Multiple Linear Regression

Coefficient interpretation

- $\hat{Y} = \hat{\alpha} + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}$
- $\hat{\alpha}$ - intercept, constant
- $\hat{\beta}_1$ - coefficient of X_1
  Expected change in the response variable for a one-unit increase in the explanatory variable, **controlling for all other explanatory variables in the model**.
- $\hat{\beta}_2$ - coefficient of X_2

Multiple Linear Regression

Dummy variables

- $\hat{Y} = \hat{\alpha} + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}$
  where X_2 is e.g. gender and is represented by either 0 or 1
- 0 - baseline category
- 1 - category that you are focusing on
- $\hat{\beta}_2$ - coefficient of X_2
  Difference in expected value of the response variable between a case coded 1 and a case coded 0, **controlling for all other explanatory variables in the model**.

Multiple Linear Regression

International Human Development

- $\hat{Y}_{school\_years} = \hat{\alpha} + \hat{\beta}_{urban\_pop} X_{urban\_pop} + \hat{\beta}_{governance} X_{governance} + \hat{\beta}_{middle\_income} X_{middle\_income} + \hat{\beta}_{high\_income} X_{high\_income}$