

# Lab 5

1. Read in the following three datasets, using {here} and {rio}:

- *sesame13.sav*
- *star.csv*
- *ais.xlsx*

*Hint:* For the *ais.xlsx* dataset look at the `skip` argument within the help documentation.

1. Using the *ais* data, compute the average red blood cell count and average bmi by sport. Output these data as SPSS and EXCEL files.
2. Use the *sesame* dataset to answer the following question: Was the average female age higher in schools or at home?
3. How many rows and columns are in the *star* dataset? Remove outliers using a really poor method, just for practice, by eliminating students whose math (*tmathss*) scores were more than three standard deviations above or below the corresponding mean. How many rows are in the dataset now?
4. Use the *star* dataset to compute standardized versions of the math and reading scores, i.e., for each variable, subtract the mean from each observation and divide by the standard deviation:  $x_s = \frac{x_i - \bar{X}}{sd(X)}$ .
  - Check that your computations were correct by computing the mean and standard deviation of each variable (they should be 0 and 1).
  - Compute the mean of the standardized variable for all sex/frl combinations (I'm asking you to extend what you know, we haven't talked explicitly about how to do this yet).
  - What do you make of these findings? Do you see a gender effect? FRL effect? Is there evidence of an interaction? (i.e., that the effect of FRL is greater for boys versus girls).