Chapter 10 exercises: Coder edition

Use the data from this chapter and the appropriate tests to examine additional predictors of library use.

1. Import the library data that have not been cleaned from this chapter: **pew\_libraries\_2016\_ch10.csv**
2. Create a small library data frame that contains variables for age, sex, parental status, education, and registered to vote (reg).
3. Follow the strategies in Box 10.2 to clean the variables in the small data frame. Write new code to clean the reg variable. The reg variable has the options listed below; recode 8 and 9 to be NA and make sure the other three categories have logical names:

* 1 - You are absolutely certain that you are registered to vote at your current address
* 2 - You are probably registered, but there is a chance your registration has lapsed
* 3 - You are not registered to vote at your current address
* 8 - Don’t know
* 9 - Refused

1. Use CreateTableOne() to create a table showing the bivariate relationships between library use and all of the variables in the data frame.
2. Write out the statistical form of the model explaining library use by age, sex, parental status, and education.
3. Use glm() to run the model corresponding to the formula you wrote out and odds.n.ends() to get model significance, model fit, and odds ratios with confidence intervals.
4. Discuss model significance and model fit.
5. Interpret the model odds ratios and confidence intervals.
6. Check the assumptions and conduct diagnostics. Interpret what you find, including examining any observations that appear problematic during diagnostics.
7. Add the voting variable to the model, run the model, interpret results, and compare the two models using the likelihood ratio test.
8. Decide which model is preferable, and explain why you selected the model.

Chapter 10 exercises: Hacker edition

Complete #1 and #2 from the coder edition and create a subset of the data so that the observations are removed when age is under 18 years old. Complete #3–#9 using *all* variables to explain library use, including voting. After estimating and interpreting the model, add an interaction between sex and voting to the model. Compare the model with and without the new interaction term using the likelihood ratio test. Based on the results of the LR test, choose a final model. Report and interpret model significance, model fit, and the odds ratios and confidence intervals for the predictors. Write at least one sentence explaining what you found.