STAT 502 Project proposal:

My suggestion for a project in STAT 502 is looking at confidence interval coverage probabilities for different sample sizes in a binomial setting. In STAT 216, we say that a sample size large enough that we see 10 successes and 10 failures for each group is enough to be able to use theory-based methods. I want to investigate where that number 10 comes from ($np = 10$, $n(1-p) = 10$), why it's the rule of thumb, and what happens for deviations from that value. I would like to compare the theory based methods to bootstrap confidence intervals, and possibly make a RShiny app to be able to mess around with the results.

If this takes less time than I anticipate, I could look at things like how outliers affect the coverage and other tests coverage probabilities for different sample sizes.