

## Math 222 - Project 4

Due Friday, April 3

1. The file <https://bclins.github.io/spring26/math222/Examples/CalculusGrades.csv> contains grade data from my calculus classes (Math 141, 142, and 242) from 2013 to 2020.
  - (a) Make side-by-side boxplots to show the grade distribution for each of the three courses.
  - (b) Which course looks like it has the highest grades? Which has the lowest grades?
  - (c) Use one-way ANOVA to see if there is significant evidence of a difference in average grades for the three calculus courses. Clearly explain your conclusions.
  - (d) Careful analyze whether the conditions for an ANOVA F-test are satisfied. Include any additional graphs and tables that you need, and clearly explain your findings.
  - (e) What are the results if you just use a two-sample t-test to compare MATH 141 grades versus MATH 242? Is there a significant difference? Is that a big deal? Explain.
2. The link <http://people.hsc.edu/faculty-staff/blins/classes/spring17/math222/data/bat10.txt> contains data from 327 Major League baseball players from one season. The variable HR is the number of home runs hit by each player.
  - (a) Make a graph to visualize how the number of home runs differs for the four different types of position (catchers, designated hitters, in-field, and out-field).
  - (b) Does it appear that there are significant differences in the average number of home runs for the different positions? Do an ANOVA test and explain your results.
  - (c) Use the Tukey's Honest Significant Differences command in R to make confidence intervals for the differences between the average number of home runs for each pair of positions. Which pairs of positions are clearly better? How much better are they?
  - (d) Careful analyze whether the conditions for an ANOVA F-test are satisfied. Include any additional graphs and tables that you need, and clearly explain your findings.