STAT 213 (SUMMER 2021): HW3 (SLOS A1, B1)

DUE ELECTRONICALLY BY THE CLASS TIME, FRIDAY 6/18/21

Instructions and Technical Tips

Write up your solutions and save/upload the file(s) to the RStudioPro server in the folder `stat213/turnin/hw3/` inside your Home directory. Include `hw3` (all lowercase) in your filenames.

Using an RMarkdown document is recommended, but not required for the non-lab portion of the assignment. If using Markdown you can either put your answers to the non-lab problems in a separate document, or just add them to the end of the lab.

RMarkdown Format. If you do use RMarkdown, R code should go in code chunks, and verbal commentary (and any math equations) should go outside code chunks.

Use section headings (a line starting with one or more `#` symbols) to demarcate the start of a problem. You can nest headings by using an additional `#` symbol for each level of nesting: one for a top-level heading, two for the next level within that, etc.

Periodically "Knit" your file to verify that it is working correctly.

"Knitting" Troubleshooting. If your code runs chunk by chunk but won't Knit, try clearing your environment (broom icon in the upper right) and running chunk by chunk from the start again. The most common cause is an undefined variable. This can happen if you change your variable names some places but not others, but a very common reason is that you read in the data from a file using a menu button instead of using the `read.file()` command, which means the "reading in the data" step is not recorded in your document. If you can't Knit, I won't be able to run your code either!

If you aren't able to Knit directly to `.pdf`, it may be because you are using special characters (such as $\neq$) in your `.Rmd`. Replace these with plain text and try again.

If you have done this and are still unable to Knit directly to `.pdf`, convert your Knitted `.html` or `.docx` into a `.pdf` (possibly by "printing" the `html` from your browser to a file) and save that there. A uniform file format across students will make grading 64 problem sets much more streamlined. Thanks!

Date: June 13, 2021.

What to turn in. You should turn in the source file (ending in `.Rmd`) and the Knitted output file for the `reading-in-data.Rmd` lab (in the `R-orientation` folder), and if you are using Markdown for the rest of the assigment, either a separate `.Rmd` and output file for that, or just add your responses to the other problems to the end of the lab document (clearly labeling what's what).

Verifying receipt of your work. When I run the script that collects your work after the due date, you will see a "receipt" file in the `~/stat213/receipts/hw2/` directory. If this does not appear within 24 hours or so of you submitting your assignment, let me know.

Problems

1. **Sparrow residuals.** Consider a model for the weight (in grams) of sparrows (`Weight`) using the wing length in millimeters (`WingLength`) as a predictor, using the dataset `Sparrows` (available in the `Stat2Data` package).

   (a) (SLO A1) Fit the model, plot the data with the fitted regression line, and write the regression equation in the form $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot X$, replacing Y and X with specific variable names, and $\hat{\beta}_0$ and $\hat{\beta}_1$ with fitted coefficients.

   (b) (SLO A1) Interpret the numerical value of slope of the regression model in the specific context.

   (c) (SLO B1) How well are the standard conditions for linear regression satisfied? Produce relevant plots, citing specific information from them to support your answer.

2. **Textbook prices.** The dataset `TextPrices` consists of data collected by two undergraduate students at CalPoly from the campus bookstore. They took a random sample of 30 textbooks from the CalPoly campus bookstore, collected in the fall of 2006. The variables of interest are `Price`, the price of each textbook, and `Pages`, the number of pages in the book. The students were interested in the question of whether the number of pages can be used to predict prices.

   (a) Produce the relevant scatterplot to investigate the students' question. Comment on what the scatterplot reveals about the question.

   (b) (SLO A1) Fit the model, write out the equation of the regression line for predicting price from number of pages, and interpret the numerical value of the slope in context.

   (c) (SLO B1) Produce and examine relevant residual plots, and comment on what they reveal about whether the conditions for inference are met with this data.