

# ESM 244 Assignment 2

**Due 2pm on Tuesday 2/19.** Prepare your individual efforts for Tasks 1, 2, 3 and 4 in a single well-formatted Rmarkdown document. Submit the **final knitted HTML and your .Rmd** to the GauchoSpace link. **Include all code (but not messages/warnings/etc.) in your knitted HTML.** We want to see the code and the outputs requested, but nothing else.

## Task 1: Data wrangling and visualization - cetaceans in captivity

Download the data **captive\_cetacean.csv** from GauchoSpace.

Visit the TidyTuesday ReadMe for data and variable information:

<https://github.com/rfordatascience/tidytuesday/tree/master/data/2018/2018-12-18>

For Task 1, wrangle and graph the data to create a finalized visualization of the data showing comparisons or trends for cetaceans in captivity that are interesting to you. **You** should make decisions about the type of visualization, comparisons, variables and observations included, and decisions about aesthetics and presentation.

### **For Task 1:**

In your beautifully organized .Rmd:

- a) Include any code used to wrangle and visualize the data
- b) Make sure that the code and your final graph appears in your knitted HTML
- c) Include a figure caption

## Task 2: Parameter Estimation – Wild Fish Catch

**Source:** Global wild fish catch and aquaculture production, compiled by Earth Policy Institute with 1950-2010 from U.N. Food and Agriculture Organization (FAO), *Global Capture Production* and *Global Aquaculture Production*, electronic databases, at [www.fao.org/fishery/topic/16140/en](http://www.fao.org/fishery/topic/16140/en).

Download the data **fish\_catch.csv** from GauchoSpace.

For Task 2, you will find an equation with parameters estimated by nonlinear least squares for the increase in global wild fish catch from 1950 – 2012. ***\*\*Hint: You will want to set 1950 = 0 (i.e., create a new column with years starting at 0, instead of value 1950, and use those values for your model...)\*\****

## For Task 2:

In your beautifully organized .Rmd:

- Create an exploratory graph of wild catch over time (does not need to be finalized). Include the exploratory graph in your knitted HTML.
- In text below the exploratory graph: What type of relationship describes the trend? What does that look like mathematically (include an equation, possibly using LaTeX)? What are your initial estimates for the parameters in the model?
- Use nonlinear least squares to find parameters for your model describing wild catch. Report the parameter outcomes (with units) in text or a table.
- Prepare a **finalized** (publication quality) graph showing both the original data *and* your model output. No figure caption required. The code and graph should appear in your knitted HTML.

## Task 3: Bootstrapped Confidence Interval for Proportions

The following data are from the 2014 UCSB Campus Climate Project Final Report (prepared by Rankin & Associates Consulting, available at [http://campusclimate.ucop.edu/\\_common/files/pdf-climate/ucsb-full-report.pdf](http://campusclimate.ucop.edu/_common/files/pdf-climate/ucsb-full-report.pdf)).

In the study, 22 out of 36 surveyed UCSB community members (61%) identifying as nonbinary/genderqueer responded that they had personally experienced “exclusionary, offensive, hostile or intimidating conduct” (compared to 26% and 19% for those identifying as women and men, respectively).

Your goal is to find a confidence interval for the proportion of nonbinary/genderqueer students experiencing exclusionary, hostile or intimidating experience at UCSB using bootstrapping. Create a vector reflecting the collected survey data ( $n = 36$ ), then find the bootstrapped 95% confidence interval for the *proportion* of genderqueer individuals experiencing exclusionary conduct based on the 2014 UCSB survey data.

***Hint: this task will require you to recreate the data **and** create a function that calculates the proportion of a specified outcome for each bootstrap sample.***

### **For Task 3:**

In your beautifully organized .Rmd:

- a) Any code you created to make the original sample vector, create the “proportions function,” and find bootstrap samples (use at least 10,000 here).
- b) A histogram of bootstrapped proportions (does not need to be finalized). The histogram should show up in your knitted HTML.
- c) A final statement (in text) describing the bootstrapped CI in the context of the data (suitable for publication).

### **Task 4: Watch 3 RStudio::conf talks**

Watch 3 RStudio::conf 2019 talks (~ 20 min each) of your choice (<https://resources.rstudio.com/rstudio-conf-2019>).

For each talk, include the following as text (in your .Rmd, that show up in your HTML):

- a) Talk title and speaker name
- b) Two bulletpoints (for each talk), interesting things that you learned
- c) One question that you would have liked to asked the speaker

**SUBMIT YOUR SINGLE WELL-ORGANIZED AND THOUGHTFULLY FORMATTED .RMD AND KNITTED HTML TO THE LINK ON GAUCHOSPACE BY 2pm TUESDAY 2/19.**