Assignment 4 - ESM 244

Tasks 1, 2, and 3 due on Gauchospace Monday, March 20, 2023 by 5 pm PST. Task 4 is done once you’ve presented your Shiny App!

* Task 1: Spatial data visualization: **10 pts**; Submit to Gauchospace:
  + a4\_task1\_lastname\_firstname.html
* Task 2: Willamette Falls time series: **20 pts**; Submit to Gauchospace:
  + a4\_task2\_lastname\_firstname.html
* Task 3: Text wrangling and analysis: **10 pts**; Submit to Gauchospace:
  + a4\_task3\_lastname\_firstname.html
* Task 4: Prepare to present your Shiny app
  + Graded separately - see the [Term Project description here](https://docs.google.com/document/d/1L2413--sYZLijV10klH_U0kpN9mkAqITDAJK4rqaYMw/edit?usp=share_link)

# TASK 1: Spatial data visualization (interactive map & choropleth)

Download data from CA DFW Oil Spill Incident Tracking from 2008 (available [HERE](https://gis.data.ca.gov/datasets/7464e3d6f4924b50ad06e5a553d71086_0/explore?location=36.873195%2C-119.422009%2C6.74), download as .csv). You should also find and download shapefile data for borders of California counties (we have used this data in ESM 206 and 244, or you can find your own source for California county polygons from another source or R package).

For this task, read in the spatial data, and create a professionally formatted and prepared HTML (showing all of your code directly or making it available with code-folding) from an .Rmd in which you:

* Make an exploratory **interactive map** **in tmap** showing the location of oil spill events included in the data.
* Make a **finalized** static [choropleth map](https://www.caliper.com/glossary/what-is-a-choropleth-map.htm) **in ggplot** in which the fill color for each county depends on the **count** of **inland** oil spill events by county for the 2008 oil spill data
* Challenge yourself (OPTIONAL): perform a point pattern analysis to assess whether oil spills tend to be more clustered or more uniform than complete spatial randomness. Plot the G function (here, units are in degrees lat-long, use that to help decide on the r) and include a brief interpretation of the results.

**Include an overview section** that briefly summarizes the dataset and your analysis. This should include a well formatted data citation.

# TASK 2: Willamette Falls fish passage time series summary

The goal of this task is to create a single beautiful, customized RMarkdown document to explore time series data in 3 parts, based on [this Willamette Falls fish passage data](https://drive.google.com/file/d/1lkTWZTcy1nxyMZuYLXSh0FlRCmqOOv5j/view?usp=share_link).

**Your .Rmd must be customized as follows:**

* With a **non-default theme**, for example using [bootstrap themes](https://bootswatch.com/3/), with a custom theme, or with a theme from another [package](https://www.datadreaming.org/post/r-markdown-theme-gallery/) (see [this section](https://bookdown.org/yihui/rmarkdown/html-document.html#appearance_and_style) of [RMarkdown: The Definitive Guide](https://bookdown.org/yihui/rmarkdown/) by Yihui Xie, J. J. Allaire, and Garrett Grolemund for more details)
* With a **separate *tab* for each of the three sections** below the Overview (see [this section](https://bookdown.org/yihui/rmarkdown/html-document.html#tabbed-sections) on tabs in [RMarkdown: The Definitive Guide](https://bookdown.org/yihui/rmarkdown/) by Yihui Xie, J. J. Allaire, and Garrett Grolemund)
* With **code folding applied**. *All* code (including setup chunk and packages) used to wrangle the data and prepare your graphs & tables should be available for us to see if we click on the code buttons (but **hide any warnings & messages**)
* The document should be cohesive in presentation and communication (e.g. all graphs should use the same theming, consistent color schemes, etc.)

**In your RMarkdown document,** you will explore [adult fish passage recorded from 2001-01-01 to 2010-12-31 at the Willamette Falls fish ladder on the Willamette River (Oregon).](https://drive.google.com/file/d/1lkTWZTcy1nxyMZuYLXSh0FlRCmqOOv5j/view?usp=share_link) Data were shared by and accessed from [Columbia River DART](http://www.cbr.washington.edu/dart/query/adult_graph_text) (Data Access in Real Time).

## You should include the following content:

## Overview (above the tabs) should contain, in whatever order you choose:

* An engaging image (with caption, incl. photo credit) that is relevant to the dataset
* A brief summary (3 - 4 sentences) of the dataset, and what is included in this report
* A map of the fish ladder location (you can make this in R on your own, or include an existing map appropriately licensed, with attribution)
* A professionally formatted data citation
* **Remember:** All code that you used to wrangle the data and prepare the graphs should be available to see if we click on the Code buttons.

## Tab 1: Original time series

* A finalized, static graph of adult passage for **coho**, **jack coho**, and **steelhead** salmon (you decide if these all occupy the same panel, or if they are in separate panels). Add a figure caption. **Replace NA values with zero.**
* 2 - 3 bullet points summarizing major patterns / notable trends in the salmon passage time series data

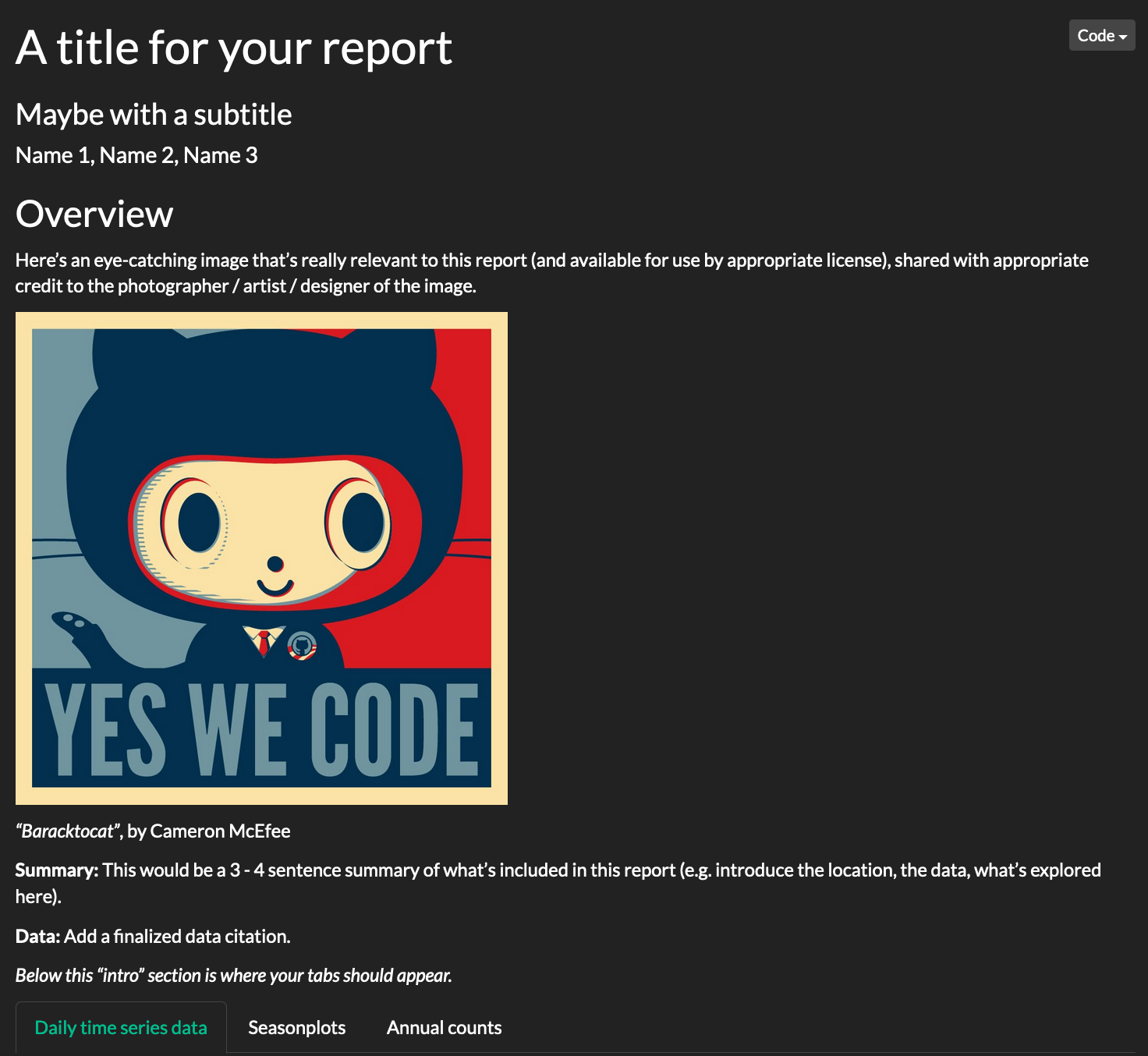
## Tab 2: Seasonplots

* A finalized seasonplot for each species (**coho, jack coho, steelhead)** in its own panel. You can either facet\_wrap/facet\_grid, or use patchwork or cowplot, to make the figure. Add a caption.
* Add 2 - 3 bullet points summarizing the major trends you see in the seasonplots.

## Tab 3: Annual counts by species

* A finalized figure of **annual totals** for fish passage, for each of the 3 species (coho, jack coho, steelhead). You decide if this is in a single panel, or multiple panels. Add a figure caption.
* Add 2 - 3 bullet points summarizing major trends you see in the annual totals by species from 2000 - 2010.

Here’s the general idea for the layout (top “Overview” section, with the three tabs containing your visualizations below):



# TASK 3: Text wrangling and analysis (coder’s choice)

For this task, prepare a professionally formatted HTML (showing all of your code directly or making it available with code-folding) from an .Rmd in which you:

* Import text of your choosing (from a PDF, text file, or otherwise) - I encourage you to find some text that is of interest to you, but some suggestions for places to look are listed below this task. You can also import two text files if you want to do a comparison of most frequent words or sentiment analysis.
* Wrangle the data to get tokens into tidy format, removing stop words
* Find and make a finalized visualization of counts for the most frequently used words in the text (this can be split up by chapter / section for comparison, or for the entire document), for example in a column graph or wordcloud, or both.
* Perform sentiment analysis using one of the lexicons introduced in Lab Week 9 (your choice), and present in a final visualization.
* **Include an overview section** that briefly summarizes the dataset and your analysis. This should as always include a well formatted data citation. Consider including an image (e.g., book cover or illustration) for visual appeal.

Some places where you can find text:

* [Internet Archive (texts)](https://archive.org/details/texts)
* [CA digital archive](https://archive.org/details/cdl)
* You can also just copy & paste text from other sources online (like websites, news/journal articles, transcripts, etc.) and save as a .txt file to read in

# 

# TASK 4: Prepare to present your Shiny app (by your session)

Here is the [rubric for your Shiny app](https://docs.google.com/document/d/1l5oGmEtWY12l_x_9iM83_3QRphFpkuzfjru0W-5NPKk/edit) term project. Each group should prepare a ~10 minute presentation on your app that you will give during the [window that you’ve signed up for](https://docs.google.com/spreadsheets/d/1mFRLRN80Slj0q92hk0mQyRIvec44nYJBE9Cht986-9I/edit?usp=share_link), leaving several minutes for Q&A from the audience.

Each member of the app team is expected to participate in the presentation, though you can decide how to divide the presentation (e.g., each member describes a whole tab of the app, or one member describes the widgets while another explains the outputs, etc).

One team member can plug their laptop into the podium in the front of Bren Hall 1424 (if your laptop has an odd video output, check with me so we can make sure we have the proper adapter!).  
  
In your presentation, you should:

* Describe the purpose of the app
* Walk through each tab of your app
* Demonstrate the widgets and reactive outputs
* Describe a challenge you had in building the app
* Describe remaining problems/what you’d like to include but couldn’t
* Each member should share what you’re most proud of

End of Assignment 4!