OCEAN 215 | Autumn 2020

Final Project

Due Dates	
Topic check-in:	Friday, October 16th, 11:59 pm PT
Data check-in:	Saturday, November 7th, 11:59 pm PT
Piazza responses due:	Saturday, November 14th, 11:59 pm PT
Project presentations:	Tuesday 8th/ Thursday 10th, During class
Slides and code due:	Saturday, December 12th, 11:59 pm PT

Project Description

During this course, you will conduct a small scientific research project from start to finish. You will choose a topic, produce a scientific question related to your topic, suggest a hypothesis, locate data that will help support or reject your hypothesis, analyze/visualize this data using Python, and present your findings to the class. Along the way, you will be responsible for giving your input on other students' projects and you will receive input from other students as well. To further reflect the collaborative nature of scientific research, we also encourage you to post any questions or challenges you encounter during this project to the class on Piazza. **If you wish, you may work with a partner on this project. See the third page for important information if you choose this option.**

The majority of this project will involve writing Python code to analyze and visualize your chosen data. We will dedicate a substantial amount of class time for this work, during which instructors and peers will be present to help you work through coding challenges. Throughout the quarter, there are a number of due dates for different parts (see the table above) designed to guide you through completing your research. The expectations for the deliverables of this project are detailed below:

1. **Topic check-in:** Consider a topic of research that you would like to examine. If you are having trouble identifying a topic, contact the instructors privately on Piazza so they can help you find something that interests you. Once you have identified your topic, create a private note to the instructors on Piazza in the 'final_project' folder that answers the questions below.

   - What research topics or questions are you interested in?

   - What type(s) of data would help you look into those topics/questions?

2. **Data check-in:** Locate data that will help you look into your selected topic. We will set aside some class time for students to work on this. You can start your exploration by using an internet search engine to look up background information on your topic and find possible data sources. You can also use the oceanography data repositories (e.g. PO.DAAC, NASA Giovanni, BCO-DMO, etc.) listed in the Class #1 slides. As always, the instructors are also available to help you locate a fitting data source. In the 'final_project' folder of Piazza, respond to the data check-in post with answers to the questions below. Make sure that your response is visible to the whole class.

   - What data set(s) do you plan to use?
   - What is one scientific question that you might be able to answer using these data?
   - What is your hypothesis? What do you anticipate the answer to your scientific question is, and why? (try to bring in scientific knowledge from previous courses, published literature, and/or reliable internet sources)

3. **Piazza responses:** Respond to at least 3 other data check-in posts written by your classmates on Piazza with an additional question that they might be able to investigate using their data or about their topic. To reflect the collaborative nature of research, where colleagues often help to dictate research priorities, you will choose one question suggested by a classmate and one question of your own to investigate.

4. **Project presentations:** Present the results of your project to the class in a 5 minute presentation. Presentation schedules will be posted to Canvas later this quarter. Your presentation should include the following:

   - Scientific background on your topic [~1 slide]
   - Two scientific questions (yours and a classmate's from Piazza) with your hypotheses [~1 slide]
   - Information about your data (How/when/where was it collected? What instruments were used? Are there any limitations to your data?) [~1 slide]
   - Your process for obtaining, loading, cleaning, visualizing and analyzing the data. Describe your data file format(s) and structure(s) as well as any challenges you encountered [~1-2 slides]
   - Answers to your scientific questions with associated plots and an explanation of your analysis results [~2-3 slides, ~2-3 figures]

5. **Slides and code:** Submit the slides from your project presentation, your data files/folders, and the code you wrote to analyze your data and create your figures. Your code should follow proper coding etiquette and your figures should be formatted properly. To submit your code, data, and slides, save them and put them in your individual class Google Drive folder. There is no written essay required for this project.

NOTE: Piazza posts that are required for the final project do not count towards the required 5 Piazza posts detailed in the syllabus.

Pair programming option

If you wish, you may work with a partner on this project. This could be a valuable opportunity to experience a research collaboration, work through coding challenges together, and accomplish even more analysis! If you choose this option, the following expectations supersede (override) the requirements listed elsewhere in this document:

- Starting with the **data check-in**, you and your partner may choose a single data set together, and share this identical data set on Piazza. However, please each offer a **different scientific question and hypothesis** in your Piazza posts (i.e. the two of you will come up with two questions and two hypotheses in total).

- For the **Piazza responses**, you and your partner should **each respond to three classmates' posts** (not including your partner's post), for a total of six posts. You will jointly choose **one question from a classmate and two questions of your own** to investigate for your project, for a total of three questions to investigate.

- For the **final project presentations**, please prepare a single **8-10 minute slideshow**, instead of a 5 minute slideshow. Include at least the number of slides specified above for each category. Trade off roles when presenting (i.e. each person should be presenting for about 4-5 minutes).

- You may submit separate Colab code notebooks, or a single joint Colab notebook. However, in all notebooks submitted, please indicate which student wrote each section of code using Python comments. **We expect that both partners will contribute approximately equally to writing code for the project.**

- You will be graded jointly and **will receive the same grade for the project**, except under extenuating circumstances to be determined on a case-by-case basis.

Grading breakdown

Project Part	Grading	Points
Topic check-in:	Complete/Incomplete	10 points
Data check-in:	Complete/Incomplete	15 points
Piazza responses:	Complete/Incomplete	15 points
Project presentation:	See rubric below	20 points
Code:	See rubric below	40 points
		(100 points total)

Presentation Content					
	Limited (0-50%)	**Good (50-75%)**	**Exceptional (75-100%)**		
Background	Topic background is missing or severely lacking in detail.	Topic background is sufficient, but missing some details or lacks coherency.	Topic background is clear, complete, and relevant.	3 points	
Questions/ Hypotheses	Questions are not well-defined. Hypotheses are not substantiated.	Questions are well-defined. Hypotheses draw on prior knowledge.	Questions are well-defined and pertinent for the topic. Hypotheses draw on prior knowledge and have clear explanations for why they are expected.	2 points	
Data Information	Information about the data collection process is missing key details or is inaccurate. The limitations of the data are missing or not realistic.	Information about the data collection process is accurate, but missing some minor details. The limitations of the data are explained.	Information about the data collection process is complete and accurate. Underlying problems and limitations of the data are explained. Use of these data to answer the project questions is justified.	3 points	
Data Processing	The student has made errors in processing their data. The student is missing steps.	The student has processed the data correctly. Steps for obtaining, loading, cleaning, and analyzing the data are well-defined.	The student has processed the data correctly and taken precautions to ensure that their results are appropriate. Steps for obtaining, loading, cleaning, and analyzing the data are well-defined.	3 points	
Results	Results of the project do not attempt to answer the scientific questions. The data visualizations are not relevant.	Results of the project somewhat answer the scientific questions. Data visualizations are mostly appropriate for the data.	Results of the project answer, or earnestly attempt to answer, the scientific questions. Data visualizations are entirely appropriate for the data.	3 points	
Presentation Skills					
Organization	The presentation is not in a logical order and the student makes no effort to guide the audience.	The presentation is organized in a logical order and takes some care to guide the audience.	The presentation is organized in a logical order and shows exceptional attention to guiding the audience.	2 points	
Timing	The student far exceeds their allotted time and/or has not made an effort to practice.	The student completes the presentation somewhat over 5 minutes.	The student completes the presentation within 5 minutes and it is clear that they have practiced.	1 point	
Explanation of Ideas/ Information	The ideas and information explained in the presentation were not clear and were not relevant.	The ideas and information explained in the presentation were clear and relevant.	The ideas and information explained in the presentation were exceptionally clear, relevant, and coherent.	3 points	Presentation: /20 points

Code					
	Limited (0-50%)	**Good (50-75%)**	**Exceptional (75-100%)**		
Correctness	The student misuses code and does not produce reasonable results.	The student uses some coding techniques/tools learned throughout the quarter. The analysis produces reasonable answers that can be replicated with some effort.	The student properly and efficiently uses the coding techniques/tools learned throughout the quarter. The analysis produces reasonable answers that can be replicated easily.	8 points	
Functionality	The code does not run and has egregious errors.	The code is mostly able to run, but has some (small) errors.	The code runs efficiently with no errors.	5 points	
Tidiness	The code breaks proper etiquette and should not be shared with others.	The code mostly follows proper coding etiquette. The organization is somewhat lacking and would need review before sharing.	The code follows proper coding etiquette. It is organized and commented effectively so that it can easily be shared with another person.	6 points	
Perseverance	The student has made no effort to work through problems and hurdles.	The student has made some effort to work through problems.	The student has made a gallant effort to work through problems and documented in their code their best understanding of the problems they are facing.	5 points	
Plots					
Plot Clarity	The plots are unclear and do not make sense in the context of the project.	The plots are mostly clear and show some thought from the students about ways to present their data.	The plots are extremely clear and are effective tools to help the audience understand the results/analysis.	5 points	
Colormaps	The colormaps are not appropriate for the data being shown.	The colormaps are appropriate for the data being shown.	The colormaps are appropriate for the data being shown and take into account colorblindness, and perceptual accuracy.	3 points	
Proper Labels	The plots are missing most/all labels or have improper labels.	The plots are labeled with general accuracy and completion.	The plots are labeled extremely accurately in a way that guides the audience through the figure.	5 points	Code:
Creativity	The student made no effort to create original plots.	The student has made some effort to create original plots.	The student has created original plots that show the data/analysis in an extremely effective manner.	3 points	/40 points