Intro to Data Visualization and Communication Final Projects

A key part of many branches of science is the analysis and communication of data. As budding scientists, these are basic skills that you should take away from CPSC 229. Thus, the final project in CPSC 229 seeks to assess your ability to analyze and communicate data.

For the final project, you will be assessed on your ability to communicate data in two ways: written and oral. The choice of how you choose to communicate your data is up to you within each category. For the written section, you may choose to write a short scientific paper or a grant proposal. For the oral section, you may choose to present a talk or a poster.

The data set is also up to you to choose. You may pick a data set from any field of science that you find interesting and want to learn more about. It must be sufficiently complex to ask and answer interesting questions. It must be approved by me prior to you beginning your project. **The deadline to have an approved data set is Friday, November 8, 2019.**

Part 1: Written Formats

Option 1: Standard scientific paper manuscript. Scientific papers are a fundamental aspect of science: communicating your data and having those results and your study be reviewed by your peers. The purpose of a scientific paper is to report what you've done and put those results in the context of the other science that has been done. Many papers also try to convince their readers that their work is important and novel as well.

Your scientific paper manuscript should have the following sections:

- **Title page**: containing a descriptive title of your paper, your name and ID number, the date, and an abstract of the manuscript.
- **Abstract**: a 250-word (maximum) summary of the contents of your manuscript. It should include a brief background, your hypothesis/question, a brief overview of the methods used, a summary of the most important findings, and a short interpretation of those findings. The abstract should not contain references.
- **Introduction.** The introduction or background of a paper gives the reader just enough information in the field so that they understand the work reported in the paper. It is important to explain to the reader key information they may not know in order to understand why the work is important. Additionally, all of science is built upon the work of others, and a key part of the introduction is to give credit to those scientists whose work laid the basis for the current work through citing relevant papers. It also sets up the hypotheses that the current work is going to test. After reading the introduction, the reader should understand: the background of the work (with appropriate credit given), why the work is worth doing, and what the hypotheses are being tested.
- **Materials and Methods.** The materials and methods section covers the details of the procedure and tools used to complete the work in enough detail that the results could be reproduced by another, independent scientist. Typically written as a narrative (not a list!), it

includes details like: what tools, equipment, and reagents were used (including brand names and manufacturer); what important points of the procedure in what order; and how statistics were done and with what methods/software. For papers reporting computational results, the basics of the mathematical model and the method of how it was solved are reported, the type of software used, and what values of parameters were swept.

- **Results.** The results section serves to *only report results, not give interpretation.* This includes any figures or tables including things produced during the course of the experiment, and text explaining the patterns observed in the raw data. For example, a parameter sweep of a computational model would produce a graph of the output of the model versus the values of the parameter tested. The text would reference this figure (the graph) and could report things like the range of output values found from running the model, the relationship (is it linear? non-linear?) between the output and the parameter, any odd-ball outliers that might be present, but would not interpret what those findings mean.
- **Discussion.** The discussion section serves to *connect* the introduction with the results sections by interpreting the hypotheses presented in the introduction with information reported in the results. Typically, the hypotheses will be evaluated (supported or not supported) by the results. The results will then be interpreted in the greater context of background information: what do these results mean? why are they important? what do they tell us about the greater system that is being studied? This is also the place where scientists will outline the weaknesses of their experiments and models, as well as the next steps that they would take to further study their system.
- **Acknowledgements**: Acknowledge people who helped you with this project (excluding authors).
- **References**: List of full references included in the paper. See style guidelines below. **You must include at least three peer-reviewed literature sources in your final paper.** Be sure to cite any source that you use with a full reference. You are allowed to use non-peer reviewed sources, but know they will not count toward the required three.

Option 2: Grant proposal manuscript. Grant proposals are very similar to scientific papers in scholarship and somewhat similar in format, but they have a very different aim. The aim of a grant proposal is to convince a funding body (whether it is NSF, NIH, or a private organization) to give you money to do some science! Therefore, you should aim to show your readers that: what you want to do is important, and what you want to do is feasible.

Your grant proposal manuscript should have the following sections:

- **Title page**: containing a descriptive title of your paper, the names and Banner IDs of your group members, the date, and an abstract of the manuscript.
- **Abstract**: a 250-word (maximum) summary of the contents of your manuscript. It should include a brief background, your hypothesis/question, a brief overview of the methods used, a summary of the most important findings, and a short interpretation of those findings. The abstract should not contain references.

- **Introduction.** Similar to the scientific paper, the introduction or background of a paper gives the reader just enough information in the field so that they understand the work to be done in the proposal. It is important to explain to the reader key information they may not know in order to understand why the work is important. Additionally, all of science is built upon the work of others, and a key part of the introduction is to give credit to those scientists whose work laid the basis for the current work through citing relevant papers. It also sets up the hypotheses that the current work is going to test. After reading the introduction, the reader should understand: the background of the work (with appropriate credit given), why the work is worth doing, and what the hypotheses are being tested.
- **Section titles as appropriate**: The organization of section headings typically depends on the content.
- **Conclusions**: The take-away message of the proposal, summarize what you want to do and why it is important.
- **Acknowledgements**: Acknowledge people who helped you with this project (excluding authors).
- **References**: List of full references included in the paper. See style guidelines below. **You must include at least three peer-reviewed literature sources in your final paper.** Be sure to cite any source that you use with a full reference. You are allowed to use non-peer reviewed sources, but know they will not count toward the required three.

Other formatting guidelines for written communications (BOTH Options):

- The length limit will be no less than 3 pages and no more than 6 pages single-spaced and in 12 point Times or Times New Roman font.
- Include page numbers on each page.
- Double-space and include line numbers when you turn in your first draft.
- You should have at least 2 custom figures. Include figures and tables at the end of the manuscript. *These will not count toward your page limit.*
- Reference figures and tables in the order in which they occur in the text.
- Figures and tables must have legible font when printed.
- Figures, tables, values, and references should be formatted in the style of the *Journal of the Royal Society Proceedings B*: https://royalsociety.org/journals/authors/author-guidelines/
- Submit your manuscript as a single PDF.
- Upload the data set and code for reproducing your analysis and figures to your Github account and mark the final version that you'd like me to grade. (Be sure to have separate R files for your paper and presentation and label them so I can tell which is which.)

I will require a rough draft of your written component due at the time you give your oral component. I will give you feedback and return it to you. **The final written assignment is due Thursday, December 12, 2019 at 1:30 pm (24 hours after the scheduled final exam time).**

Part 2: Oral Presentation Formats

Option 1: a Poster Presentation. Poster presentations are a common form of presentations at scientific conferences where there is not enough time for everyone to give a talk. Researchers will prepare a poster that communicates their basic findings and pin it up in an exhibition hall the day before their session. This gives other attendees a chance to read their poster before the session. There is then a session where researchers stand by their posters to give short presentations and answer questions people might have. Many people (including me!) prefer to give poster presentations because of the conversations they invoke and the one-on-one nature of the interaction.

The poster itself consists of sections that are very similar to the scientific paper: usually a title and authors section, intro/background, materials & methods, results, conclusions, acknowledgements, and references. It's typical to locate the results and conclusions centrally to draw the viewer in and materials and methods, acknowledgements, and references towards the periphery.

The oral part of the presentation is usually fairly simple: they walk their viewers through the poster, pointing out important pieces of background, hypotheses, results, and conclusions. Poster presenters often have a "short" version of their presentation for someone who only has two or three minutes before moving on, and then a "long" version that is about five minutes.

You will present your poster to both me and your classmates during a poster session. Have a "short" and "long" version of your presentation ready. Posters (in PDF format) must be emailed to Cheryl Stack (cstack@chapman.edu) no later than Tuesday, December 10 at 9 am for printing. **Poster presentations will occur on Wednesday, December 11, 2019 during final exam time slot (1:30-4 pm).**

Option 2: a Powerpoint Talk. Talks are the other form of presentations at scientific conferences. Researchers will prepare a slide show in Powerpoint or a similar software and give prepared remarks along with their slides. There is typically a question/answer session immediately following the talk so that people in the audience can interact with the speaker.

Talks are usually organized very similarly to scientific papers (in that they have an introduction/background section, a materials and methods, results and discussion/conclusions are sometimes mixed together). The presentation should be well practiced, fit within the allotted time (neither too short nor too long), and well organized to allow for fast information transfer to the audience. Typically the types of figures/graphs that are appropriate for papers are not appropriate for talks, as the audience will have very little time in which to understand the main point before the speaker must move on.

You will have to give a **15 minute presentation** (10 minutes minimum, 15 minutes maximum) to the class on your topic with a few additional minutes for taking questions from the audience. **Powerpoint talks will occur on Wednesday, December 11, 2019 during final exam time slot (1:30-4pm).**

Other guidelines for oral communications (BOTH Options):

- Your presentation should be organized roughly the same as your paper, it is advisable to work on both at the same time. The presentation should be a less detailed version of your paper. Note how the first draft of the paper is due right at the start of presentations. *This is not an accident.*
- Attendance is required for everyone on those days (not attending will hurt your final project grade.) If you have any conflicts, please notify me as soon as possible!

Grading Scheme for Final Projects (with Deadlines)

Each assignment/deadline within the final project series is weighted as follows:

- Fri Nov 8, 5 pm: Data set approval (2.5%)
- Mon Nov 18 – Fri Nov 22: Individual progress conferences (2.5%)
- Mon Dec 2, 5 pm: Deadline for first drafts of oral presentations (2.5%)
- Tue Dec 3 & Thu Dec 5: Individual conferences for feedback on oral presentations (2.5%)
- Fri Dec 6, 5 pm: Deadline for first drafts of written assignment (5%)
- Wed Dec 11, final exam time: Participation in presentations (5%)
- Wed Dec 11, final exam time: Oral Presentation Score (40%)
- Thu Dec 12, 1:30pm: Written Score (40%)

Methods of Assessment:

- Many of these are milestones which will be graded based on meeting the milestone in by the deadline listed (full credit or no credit). (Examples: data set approval, turning in first drafts)
- Some will be participation based, which typically will be all or nothing based on participating in the event (full credit or no credit). (Examples: Individual progress conferences, final project participation)
- The final oral presentation will be based on a rubric and will be assessed by both the instructor and other students.
- The final written assignment will be based on a rubric that will be assessed by the instructor.