

**MA 207 – Spring 2018**  
*Lab #10: Conduct your own statistical research!*

---

---

You will undertake a data science project on a topic of your choice. This project is an opportunity to show off what you've learned about statistics. Your task is to use data to tell us something interesting. This project is deliberately open-ended to allow you to fully explore your creativity. Here are the main rules that must be followed:

- **Use the materials learned in this class.** (e.g., descriptive statistics, hypothesis tests, confidence intervals, etc.).
  - **Your project must be centered around a data set.** You will need to find a data set that interests you. (Your instructor will give you some ideas about where to look for these). Be sure to cite your sources.
  - **Your project must tell us something.** Make sure that your project is thought-provoking and has some underlying meaning!
  - **Cleverness and creativity will be rewarded.** Going above and beyond what we did in class will be rewarded.
  - **Collaboration.** You may work with a partner, but each of you needs to contribute meaningfully to the outcome. You may consult other sources, but you should credit these sources in your report. Feel free to consult with me.
- 
- 

**Part 1 – Executive Summary:**

Your 1-page summary should contain the following content:

Title: The title of your project

Purpose: Describe the general topic/phenomenon you want to explore, as well some carefully considered questions that you hope to address. You should make an argument motivating your work. Why should someone be interested in what you are doing? What do you hope people will learn from your project?

Data: Describe where you found your data, and what kind of data it is. Be as specific as you can, listing URLs and file formats if possible.

Variables: List, and briefly describe, each variable that you plan to incorporate. If you can, be specific about units, scale, etc.

Honor Code: Indicate if any component of this project overlaps with work for another class/thesis. If this is the case, please speak to your professor/advisor and have them email me their consent by Wednesday April 12th.

---

---

**Part 2 – Full Project Write-up:**

Your write-up should tell a statistically savvy audience about your project, why they should care about it, and what you have discovered. Your audience will be people like you. Your write-up should make it clear to your instructor and any student in the class what methods and techniques you have used to produce your finished product.

**Motivation.** Be sure to motivate your topic at the beginning of your write-up. You should try to hook the reader early on. Assume that your audience is a skeptical researcher who has stumbled across your report but has very little time to read it. Can you give them a reason to continue reading? A cool visualization or result can help.

**Format.** You don't need to follow a specific format in the write-up, but you should start with an introductory paragraph and finish with a conclusion. Your write-up should address the following questions:

Why should anyone care about this?

What is this about? Do not assume that your readers have any domain knowledge! The burden of explanation as to what you are talking about is on you! For example, if your project involves phylogenetic trees, do not assume that your audience has anything other than a basic, lay understanding of genetics.

Where did your data come from? What kind of data was it? Is there a link to the data or some other way for the reader to follow up on your work?

What are your findings? What kind of statistical computations have you done to support those conclusions? When using graphics, do not forget about units, axis labels, etc.

What are the limitations of your work? Be clear so that others do not misinterpret your findings. To what population do your results apply? Do they generalize? How could your study be improved? Suggesting plausible extensions don't weaken your work, they strengthen it by connecting it to future work.

**Style.** Your write-up should include a bibliography, and your references should be embedded via links.

Visualizations will be well-received. That said, do not overuse visualizations. Any plots should be well-thought out, properly labelled, informative, and visually appealing!