Stat7350: Assignment Two

Due: March 28, 2019

Design a research study

For this assignment you will (surprise!) design an experimental research study. This is a thought project; you will not actually have to carry it out. What you propose should be realistic in terms of time and resources required—let's say, if you had a month to work on this and a small research budget of $5000. Let's throw in an undergraduate assistant working ~10 hours a week too (you are not required to use the undergrad, but they're available if you want them).

The motivation for this assignment comes from this blog post from Roger Peng on the simplystatisics blog: https://simplystatistics.org/2019/01/09/how-data-scientists-think-a-mini-case-study/

The goal is for you to think both creatively and logically about experimental design. This is also an opportunity to practice your written communication skills.

Your proposed project can be anything that involves data collection. If you want to think about biology, great. If you want to carry out an experiment on humans (or yourself), also fine. The sky is the limit (as long as it can be done in a month and for under $5000).

Start a new `Assignment2` folder in your `.Rproj` directory and push to gihub

You can submit the proposal in any format you like (e.g., pdf, rmd, doc), but the final document should be uploaded to github.

1. Frame the problem

Clearly state the research question and motivation for your project. What is already known about this topic? Why is it important to know more? I'm not looking for a full paper introduction, but you should provide enough background so that your reader knows why there is a gap in our knowledge and why this is important to fill. It is expected this will require reading, synthesizing, and citing source material.

2. State the testable question(s) and your hypotheses.

The main question should be a single sentence that should have an unambiguous answer. Then state your hypotheses. Depending on the question, this could be phrased as null and alternate hypotheses. If your project is like the one discussed in the blog post it could be very specific (e.g., "It will take me 45 minutes on average to get to work").

3. Project proposal

This is the main component of the project. *A complete proposal will include enough information that anyone with reasonable intellect could conduct your experiment given just your proposal.* It may be helpful to include a methods figure that outlines the design of your experiment (like the one I showed you for my proposed research in class on Tuesday).

Some topics that will likely be included in this section (note that this is not an exhaustive list): - What data will you collect? - What are the treatments? Is there a control treatment? - What are the experimental units (subjects)? If you are experimenting on people, how will you recruit volunteers? Is a randomization procedure or block design required? - How much replication will there be? Are there biological and technical replicates? How much variation among replicates do you expect? - Are there confounding factors/covariates? How will you control for these? - What are the experimental steps to take? (this could be provided as a flow chart, checklist, bullet-points, etc.)

4. Data collection and statistical analysis plan

In what format will you store the collected data? What will the spreadsheet look like (e.g., what are the column headers)? What type of analysis do you plan to carry out on the data you collect? What are the fixed and random effects?

Evaluation

Components of the assignment will be graded on a 3-point scale. Each aspect has different weights based on the amount of work expected to go into that section. The general rubric is here:

Weight	Topic	Excellent: 3	Satistfactory: 2	Needs Work: 1
3	Frame the problem	Establishes strong sense of purpose. Source materials are introduced and contextualized. Clear divide between ideas from student and sources.	Purpose is established. Evidence that source material was read and shaped student's writing. Materials are generally cited.	Fails to establish purpose or the purpose is too easy to attain. Does not contextualize source material. Repeats or summarizes source text without analyzing.
1	Question and hypotheses	Clearly stated question with logical, testable hypotheses.	A question statement and hypotheses are stated but could be better phrased.	Lacking either or both a problem statement and hypotheses or hypothees do not flow logically from the problem.
5	Project proposal	Research plan is comprehensive and doable. Ideas are arranged logically to fully address the problem statement and desired outcomes. Reader can easily follow the line of reasoning.	Research plan needs modification to be comprehensive or doable. Ideas flow and are usually linked to each other. The reader can follow the line of reasoning most of the time.	Research plan is insufficient. The writing is not logically organized and does not address the problem statement. The reader cannot identify a ine of reasoning.
2	Analysis plan	Data collection methods are fully described. An analysis plan is fully outlined.	Data collection methods are mostly described. An analysis plan is mostly outlined.	Data collection methods are poorly descrived. An analysis plan is not included or is inappropriate.
3	Organization & Style	Well planned structure, written in an engaging, iteresting style. Strong paragraph structure. No grammatic errors. Reference have a consistent format.	Some evidence of organization, most paragraphs have topic sentences with supporting details. Style is competent though not engaging or inventive. Few grammatical errors. References are mostly in a consistent format.	Organization is unpredictable, paragraphs poorly structured. Lacks control over sentence structure, difficult to follow. Many grammatical errors. Referencing is inconsistent.
1	Ease of access for instructor, compliance with course conventions for submitted work	Access as easy as possible.	Satisfactory.	Not an earnest effort to reduce friction and comply with conventions.