Demographic Exploration and Discovery: Final Projects

Tim Riffe

July 10, 2019

Project description

The only aspect of this course that will be explicitly graded is a single report, due in September. The report has two pieces.

1. A data visualization manuscript, following the guidelines used at the journal Socius. Base your structure on examples found in the collection: Socius Data Visualization Collection

Explicit limitations are given in the submission guidelines, here quoted in full:

Visualization Submissions

> Visualization submissions should include two parts. The main article will be a single figure (i.e. something that could be viewed on a single screen/page) and approximately 500 words (excluding caption & references) of expository text that highlights what is interesting and important about the figure. The reference list should be limited to no more than 5 citations. In addition, you should include a supplementary information file that contains details about the data and methods used to generate the figure. The supplementary file has no page limits and serves to provide methods and background for the work. At minimum, the supplemental file should provide methodological information of sufficient detail for readers to judge the veracity of the figure, reasonableness of the methods, and soundness of the conclusions drawn from the figure. To that end, while not required, we strongly encourage people to include code or post data that would allow others to replicate the work as well as discussion of aesthetic/display choices involved in producing the figure. The supplemental information file can also include background literature or other information author(s) think relevant to include. Socius will not copy-edit the supplementary information file."

Summary

1. The main manuscript should be ca **500** words and contain a **single** visualization, which is your main deliverable. It should have a clear message and be skillfully crafted. You'll give me this manuscript in html, pdf, or Word, as you please, but also the R markdown file that generated it. This R markdown file should contain all steps to get from the source data to the visualization, even though not all outputs will be visible in the knitted manuscript.
2. You'll give me another R markdown file, which contains all the iterations you went through to get to (1). That is, the code steps, annotated, including failures (accidental art, that sort of thing), and versions that you decided against. This will give a glimpse of your exploration and design process. This file can be big, and it doesn't have to be in English. You don't need to follow the journal guidelines exactly for supplementary material.

Due date

Send me these items by email (include any coauthors in cc) by 17:00 September 9. No later, no exceptions.

Assistance

If you want me to look over your progress and give advice / tell you whether you're on track for what is expected, then you should arrange a time to meet between August 5 and 16. I'll check my emails outside of those dates, but I will not be able to look at your work outside those dates. If there is demand I can reserve room 100 once or more in those dates. Otherwise, I expect you to help each other where necessary and to give credit where due to each other (see next subsection).

Policy on coauthorship

To receive credit for this course, each student must be **first author** on an assignment. To be first author means you take ultimate responsibility for coordinating, drafting, and conceiving the idea, although these and all other aspects of the assignment can be collaborative among students in this course. If anyone gives explicit assistance then this needs to be declared. This in no way deducts from the merit of the assignment, and indeed students can increase their overall merit by assisting others. But that's up to you to organize. Assignments with more than one author require an **author contributions** section, where you say who did what.

Formatting

Don't waste time trying to follow the formatting details of Socius. Just use the R markdown default aesthetics, and make use of `#`, `##` sectioning. Remember use `echo = FALSE, results = "hide", message = FALSE` if you want to run code without it or its results being seen, such as data preprocessing. Do be sure to include references, see how by reading here.

Five projects

Surprise surprise everyoe has decided to recycle a data source used in a different course (I explicitly said this was OK). We'll see whether this source really contains five gold nuggets.

First	Dataset	Question	Prep	Analysis	Concept	Implement	Draft	SuppMat
JD	NHIS	mig x mort/health	-	-	-	-	-	-
MK	NHIS	Decomp LE/HLE	-	-	-	-	-	-
MF	NHIS	insurance x mort	-	-	-	-	-	-
SK	NHIS	?	-	-	-	-	-	-
NU	NHIS	poverty x mort/health	-	-	-	-	-	-

Remember, the objective is to have an impactful or policy-relevant result distilled in a single visualization. This assignment is probably smaller than other similar things you've done, and it is designed to be a light-weight simulation of a real kind of research output. Data visualization manuscript categories are sprouting up here and there in different journals. This kind of manuscript is usually more distilled and than classical research output in many ways, but that does not mean it's less rigorous, and one presumes much more effort and skill will have been placed into the visualizations than in a typical research article.