##### **LAS 6292: WEEK 1 IN-CLASS ASSIGNMENT**

One of the goals this semester is to take a “messy” dataset, clean it, organize it, and – assuming you have permission and if appropriate at this time – archive it in a repository. This assignment will introduce me to your data set so we can evaluate if it is suitable for the assignment (i.e., is it messy enough, is it big enough?). The data can be from previous research, even if it is unpublished or not part of a thesis.

It is important to remember that the ‘dataset’ can be more than one file. For instance, your dataset might be four files (or collections of files): (1) A file with the responses to surveys of community members, (2) Census data downloaded from a government website, (3) The transcripts of interviews with key informants, and (4) A collection of photographs for ethnographic research.

**This week’s assignment is to briefly answer the following questions and submit the responses via the course e-learning site.**

1. What degree are you pursuing?
2. What is your department or degree program?
3. Please provide a brief summary - 1 paragraph max - of your research interests.
4. Do You Have a Messy Data you would like to clean up and organize as your course project? This can be data from your current or prior research, including side projects, Master’s thesis, undergraduate thesis, etc. *If you don’t have a dataset of your own, THAT’S NO PROBLEM - I will help you find one.*
5. What is the specific research question you addressed with these data?
6. What is the status of this dataset? *Leave only the most relevant answer.*
   1. I have finished collecting the data
   2. I have started data collection but *will not* be done before the semester ends
   3. I have started data collection and *will* be done before the semester ends
   4. Other (please elaborate)
7. Tell me about the dataset(s):
   1. What are they and what kind of information do they contain?
   2. What tools and techniques did you use to collect it?
   3. What format are they in (e.g., sheets of paper, notebooks, spreadsheet, text files, photographs)
   4. How much data is it? *(estimated number of files, surveys, photographs, etc).*
8. Do you use any third-party data? If so what is the source?
9. Are the data from research conducted with human subjects?
10. Are any collaborators currently working with the data?
11. Will anyone (e.g., collaborators, advisor) be working with the data in the future?
12. What instruments did you use for **collecting and recording** raw data *(e.g., paper survey or data forms, online surveys, notebooks/forms in which observations/responses were recorded, specimens (pressed plants, animal collection, blood/tissue), audio recordings, automated data loggers, photographs, video)*
13. What instruments and software do you use for **processing and organizing** data? *(e.g., MS Word, Google Docs, other word processor, MS Excel, Google Sheets, other spreadsheets, Scrivener, OSF, Evernote, Flickr, R, nVivo)*
14. Are you an R user? If so,how would you rate your proficiency? *Note: No prior programming experience is required for this class.*
15. Will you be gathering data this summer or fall?
16. Do you have any favorite data-collection-or-management-related tools?
17. Is there anything you’d like to add? Are there particular topics you’d like to see covered or tools you’d like to learn?

**Grading Rubric:**  1) Survey completed with thorough answers: 50  
2) Most questions answered completely; some require instructor follow-up: 40  
3) Many questions missing answers or answers are cursory: 30 4) Instructor follow-up required for survey submission: 20