For your final project, you must contribute a Jupyter notebook or R Markdown file containing your analysis of a domain of your choice. You should implement concepts from three high-level topics we've covered or will cover in class: 1-data wrangling, 2-analysis and modeling, and 3-visualization.

Your notebook should cover the following:

1. The data you are proposing to analyze. I suggest you begin with a description of the variables in your data and the number of observations. Explain briefly how this data is typically used in the domain.

2. Your analysis should include.
    a. Data Wrangling: Common issues encountered when dealing with this type of data. Consider missing values; can you impute or should you drop them? Data may not be in the correct format, such as site names that need to be standardized or data that needs to be combined from two different files.
    b. Exploratory data analysis: Useful visualizations for gaining a quick overview of the data. Such visualizations are typically perfect for detecting any residual anomalies in the data, or for simply observing any trends or differences between groups or samples.
    c. Analysis and/or modeling: what types of analyses are common in this domain? Asking someone who specializes in this type of analysis is the easiest way to answer this question.
    d. Insight communication: Using appropriate data visualization to summarize your results. Once again, I suggest consulting a colleague who is familiar with the types of visualization used with the data you are using.

2. You will be graded on the following:
    - Data complexity and tidiness after processing
    - The difficulty of the proposed analysis, code correctness, and complexity of the solution proposed.
    - Code and style, and clarity
    - Appropriateness of the visualization, the model or statistical tests to explain the overall objective of the project.

Note: please email me your team project title on before or on Nov 10 at the latest