

Data Ingestion & Wrangling – Expectations

Goals:

- Ingest your data for use with R, Python, Tableau, etc.
- Clean your data so that it is usable in the analysis process.

Expectations:

- Read in your data live.
 - For example, dataMontgomery updates its datasets frequently. Make sure you are reading data from the dataMontgomery url each time in your code.
 - Do not save a CSV of the data and then import that CSV in your code.
- Clean your data:
 - Fix typos
 - Consistent missing value notation
 - Appropriate column types
 - Appropriate Date/Time variables
 - Column splitting or merging
 - Etc.

Deliverables:

- Code in Python, R, etc.
 - Code should be commented appropriately and readable
 - These can be scripts, R Markdown files, Jupyter Notebook files, etc.
- If you use a different tool to clean your data, for example OpenRefine, give a summary of what you did to clean the data. Include:
 - The tool you used
 - Summary in a bulleted list
 - If there are screen shots or other images you would like to include, add those to the summary list
- Post your deliverables on your GitHub page in your capstone project repository.