

Assignment 5: Data Science 2024 (10% written of overall credit score)

Due: Tuesday December 10, 2024 (by 8 pm ET)

Submission method: email (eleisa2@rpi.edu)

Please use the following file naming for electronic submission for any individual documents: DataScience_2024_A5_YOURFIRSTNAME_YOURLASTNAME.xxx

Late submission policy: If you are more than 5 days late it is likely that you will not have your grade for this assignment included in your final grade before they need to be submitted.

Note: Your report for this assignment should be the result of individual work. Take care to avoid plagiarism (“copying”), including all web resources, texts, and class presentations. You may meet to discuss the tasks for this assignment with other members of the class but must produce the materials for the final assignment themselves, citing all outside sources.

General assignment: Stewardship: Workflow construction for preservation. Use the data workflow(s) from your group assignment. The weighting score for each question is included below. Please use the question numbering(1: 3) below for your written assignment.

1. Construct a scientific (data-oriented) workflow. Present the workflow in a diagram form with suitable annotations (documentation) for someone else to review. The workflow should be an accurate depiction of the data flow from the chosen project for your specific instance, i.e. it need not be generalized. The diagram should be included in what you submit. Provide a minimum of a 5-6-sentence description of what was required to carry out the construction of the workflow. (6%)
2. Describe each major stage of the workflow and indicate how well (or poorly) data and information preservation is enabled or accommodated. Minimum of 3-5 sentences (2%)
3. Describe how the workflow and your assessment of existing documentation such as provenance and other contextual information (e.g. related to the dataset, code libraries, etc.) enabled (or not) the data stewardship that was needed in your project. Minimum of 3-5 sentences (2%)