

Assignment 4: CSCI-6961-01 (written - 20% and 20% presentation of overall credit score)

Due: Wednesday November 9, 2009 (by 12pm ET, day of the class)

Submission method: email to pfox@cs.rpi.edu. Please use the following file naming for electronic submission for any individual documents:

CSCI-6961-01_A4_YOURGROUPNAME_part1.xxx, etc.

Late submission policy: first time with valid reason – no penalty, otherwise 20% of score deducted each late day

Office hours: Tuesday 10-11, Winslow 2120 or by appointment

Note: Your report for this assignment should be the result of group work. Take care to avoid plagiarism (“copying”), including all web resources, texts, and class presentations. You are expected to work within a group setting contributing equally and with complementary skills and are encouraged to discuss your ideas and the tasks for this assignment with other students in the group and with other groups in the class.

General assignment: Working with someone else’s data; group chooses an investigation, finds, accesses, analyzes and visualizes the data and manages the resulting products. The weighting score for each question is included below. Please use the question numbering (1-4) below for your written assignment.

1. Choose an investigation and identify a pre-existing source of data that can address the particular goal: (5%)
 - a. Choose and state the goal and reasons why the data set or sets were chosen and how they were found and managed, 3-4 sentences
 - b. Document and discuss the data format and any metadata standards, conventions in use and the method(s) of discovery and access and how they helped or hindered the process, 3-4 sentences
2. Data Analysis (6%)
 - a. Design an analysis study to answer two particular questions, stating these questions and the analysis design, 3-4 sentences (1%)
 - b. Provide a description of the choice of tools/ methods used or a description of any code or scripts written, and describe how your results were stored and managed, 3-4 sentences (2%)
 - c. Perform the analysis in a form that can be validated and describe the steps and results your group took to ensure this validation, 3-4 sentences. (3%)
3. Visualization (6%)
 - a. Prepare visualizations of both the data (and any metadata, information) and the results of the analysis and describe them, 2-3 sentences. (2%)
 - b. Document the management of the visualization products and any associated metadata, etc. 2-3 sentences (2%)
 - c. Describe how your visualization meets the goal of the investigation and highlight any value that was gained, 3-4 sentences (2%)

4. Describe your overall data management plan for the results of questions 1, 2 and 3 using the 9 categories from assignment 1, 1-2 sentences for each category (3%)
5. Oral presentation of the results for questions 1, 2, 3, and 4 (20%). Plan to present for 25 minutes, followed by up to 10 minutes for questions. Your grade will be assessed on demonstration of learned skills in data management. All members of the team should present part of the presentation. Please submit your presentation (ppt, pdf or similar), using the same naming scheme, after the class.