13. *Writing about Data and Methods*

SUGGESTED COURSE EXTENSIONS

A. Reviewing

1. In a one- or two-page article in the health or science section of a newspaper or magazine, circle the information on data and methods.

    a. Critique the presentation of that information, using the guidelines in chapter 13 of *Writing about Multivariate Analysis, 2nd Edition* regarding writing about data and methods in general-interest articles for a lay audience.
    b. Assess whether additional information would be helpful for people seeking information to compare with findings from another study.
    c. Evaluate the authors' discussion of how the data and methods affect interpretation of the findings.

2. Read the data and methods section from an article about an application of OLS regression in a journal from your field.

    a. Critique it, using the guidelines in chapter 13 regarding writing about data and methods for scientific articles.
    b. List additional information needed by researchers seeking to replicate the data collection protocol.
    c. List additional information needed by researchers seeking to replicate the statistical analysis.
    d. Assess how well the article discusses how the data and methods affect interpretation of the findings.
    e. Indicate whether the authors suggest directions for future research.
    f. Rewrite the description of data and methods in the discussion to rectify problems you identified in parts d and e.

3. Read the methods section of an article about an application of logistic regression in a journal from your field.

    a. Evaluate whether the categories of the dependent variable were defined in the raw data or calculated by the authors. If the latter, indicate whether the authors specified the criteria or cutoffs used to perform the classification.
    b. Indicate whether the authors identify the omitted category of the dependent variable in the text and the tables.

4. Go to a data website such as the US Census Bureau, National Center for Health Statistics, or the Bureau of Labor Statistics and identify a topic of interest involving two or three variables. Evaluate the website in terms of how easy it is to find information about

    a. the type of study design (e.g., cross-sectional sample survey, retrospective, prospective);
    b. the data sources (e.g., vital registration forms, questionnaires, administrative records);
    c. the wording of questions used to collect the variables of interest to you;
    d. the units or coding of those variables;
    e. sampling weights, if applicable;
    f. the response rate;
    g. loss to follow-up (for longitudinal studies only).

B. Writing

1. Outline the data section for a scientific paper about a multivariate analysis you are conducting, using the checklist in chapter 13 of *Writing about Multivariate Analysis, 2nd Edition*.

2. Write an equation to convey your final model specification.

3. Write an explanation of why you chose the type of statistical model used in your analysis given your research question and data, following the guidelines in chapter 13.

4. Write an explanation of how you arrived at your final model specification, including the following topics:

    a. The criteria used to determine which variables were included in the model, with reference to your specific research question.
    b. Whether and why nonlinear specifications were used for any of the independent variables.
    c. Whether interactions were included among two or more independent variables, and if so, which ones and why; see also chapter 16.

5. Write a discussion of the strengths and limitations of your data and methods for a scientific audience.

6. Exchange your answers to questions B.1 through B.5 with someone studying a different topic or data. Peer-edit each other's work and revise according to the feedback you receive.

7. Write a short discussion for a lay audience about how strengths and limitations of your data and methods affect how your study's findings should be interpreted and applied in a real-world context, following the guidelines in chapters 13 and 20.

C. Revising

1. Critique a data and methods section of a scientific paper you have written previously, using the criteria in chapter 13 of *Writing about Multivariate Analysis, 2nd Edition*.

    a. Identify elements you have omitted.
    b. Track down that information in data documentation or other publications on the same data.
    c. Identify material that could be organized better or explained more clearly.
    d. Revise your data and methods section to fix the problems you identified in parts a and c.

2. Critique the discussion of data and methods in the discussion section of a scientific paper you have written previously, using the guidelines in chapter 13.

    a. Identify implications of strengths or limitations of the data that were omitted or explained poorly.
    b. Identify directions for future research related to your data and methods that were omitted or explained poorly.
    c. Revise your discussion section to correct the problems you identified in parts a and b.

3. Exchange your answers to questions C.1 and C.2 with someone studying a different topic or data. Peer-edit each other's work and revise according to the feedback you receive.

4. Exchange data and methods sections with someone who is analyzing different data and a different research question. Using only the information in that section (e.g., without reference to their computer output or data documentation, and without asking them any questions),

    a. Write an equation to express their final model specification (or a selected model if several models are presented in the paper). If some of the information needed to write an equation is missing information, list it.
    b. Identify the units or coding and omitted categories for each variable in the final model specification (or a selected model) based on the data section and tables of descriptive statistics. If any of this information is missing, unclear, or inconsistent between the tables and prose, list it.
    c. Rewrite your data and methods section to correct the problems identified by your peer-editor.