

Elements of a Data Management Report

1. Introduction
 - a. define purpose of the report
 - b. reflect on study goals so that it is clear that you understand the purpose of data collection and intended use of the data
2. Data sources
 - a. Input files – include names and brief description of key variables
 - b. Output files – name(s) of formatted, labeled, edited data file(s)
3. Data cleaning
 - a. Duplicate records – include details; how resolved
 - b. Missing data – how handled, coded
 - c. Problem values – invalid values, data out of range or questionable; how handled
4. New variable definitions
 - a. Grouping
 - b. Computed
5. Data summaries – goal is to present distributions
 - a. Frequency tables / bar or pie charts for categorical factors
 - b. Descriptive Statistics for continuous factors / histograms or boxplots
 - c. Include cross-tabulations, scatterplots and/or stratification for descriptive statistics where this will help you find inconsistencies or problems in data

Appendices

- A. Contents of original and final data sets
- B. Program summary table – listing name of programs, purpose, data read in; data files created if any
- C. Program logs (included for homework assignment / not typically a part of the report, though it can be)