**CMPINF 2140**

Evaluation of Data Management Plan

**Plan Title:**

**Plan Source:** [Indicate whether your plan is from the University of Pittsburgh, University of California San Diego, University of Minnesota]

1. Examine your plan. Does it address the components outlined in the table?[[1]](#footnote-1)

|  |  |  |
| --- | --- | --- |
| **DMP Component** | Is component addressed in plan (Yes, No, Unclear) | Comments on clarity, level of detail, and usefulness |
| **Data Description**: Does the plan identify and describe the data or information objects that will be generated? Are the methods for collection described? | Yes | De-identified data, dataset with a list of variables, but more information will be documented when the actual data is acquired |
| **Data Formats**: Are the file formats of expected data specified (e.g. file extensions, name of data collection software)? | Yes | Data will be in STATA format, export will be in STATA, but other formats can be used and/or requested |
| **Metadata and documentation**: Does the plan describe specific metadata standards and/or other description methods (e.g. readme files, codebooks, and lab notebooks) that will be used? | Yes | Data use log will document data-related activities. |
| **Storage and Backup**: Is a method of storage and data backup (e.g. could storage, remote backup, external hard drive) specified in the plan? | Yes | There will be a secured network drive controlled by the University of Minnesota. |
| **Data Security and Access:** Does the plan specify if there is sensitive data and, if so, how security and access will be handled? | Yes | For use, a user must log in and STATA log function will keep track of who does log in to use the data. |
| **Description of Data to be Shared:** Will any data and/or code be made accessible after the study? | Yes | The data will be available and managed by the Data Repository of UM (DRUM). There will also be two DUAs, one of pre-analysis and one for post-analysis. |
| **Data Archiving:** Will data be deposited in a dedicated data repository/archive? | Yes | The data will be available and managed by the Data Repository of UM (DRUM). |
| **Responsible Party and Roles**: Are the individuals responsible for data management identified? Are roles clear? | No |  |

2. Next, imagine you were invited to join the team as a data scientist or data curator. Record notes on:

* Would you be able to follow the plan?
  + I think this specific DMP is very vague and relies on documentation to occur once data has been acquired. But to me, this could lead to messy documentation and overall inadequate onboarding.
* Is there anything that should be made more clear?
  + I think that the data description and metadata could be way more fleshed out.
* Would you know what your responsibilities are for data management?
  + No I wouldn’t. And I wouldn’t know who to contact or who to interact with it with regards to collection, entry, and analysis.

3. Finally, summarize what you learned about the purpose, coverage, organization, and potential value of data management plans based on your examination. Post a 200 word summary on the Discussion Board.

A Data Management Plan (DMP) is crucial for any research project, outlining roles, responsibilities, methods, and goals for handling data. Examining different DMPs highlights the importance of specificity in detailing protocols and processes. The University of Minnesota’s (UM) School of Public Health (SPH) example provides only a vague description of how data will be documented once acquired, leaving room for interpretation. Their metadata section simply states that a Data Use Log will track activities when users access a secured network. This lack of detail could lead to inconsistencies or errors.

A well-structured DMP improves transparency, reproducibility, and data integrity by covering key areas such as data collection, storage, sharing, and preservation. It ensures compliance with institutional and funding requirements while helping research teams stay organized. Beyond compliance, a DMP is a tool for responsible research management, reducing inefficiencies and misunderstandings while ensuring future researchers can build on the data.

If I were a new scientist at UM-SPH, my main concerns would be the lack of project-specific details and unclear roles. Not knowing whom to ask or where to find resources creates knowledge gaps, leading to errors in communication, data handling, and quality.

1. This table was adapted from: James E. Van Loon, Katherine G. Akers, Cole Hudson and Alexandra Sarkozy, “Quality evaluation of data management plans at a research university,” *International Federation of Library Associations and Institutions* 43, no. 1 (2017): 98-104.

   [↑](#footnote-ref-1)