NEMS Data Management

Designing accurate and user-friendly data management systems that work for *your* particular situation is very important. Below are some general guidelines and suggestions that can serve as a framework for developing your own data management plan. Please refer to the **Quality Assurance** document for additional data management tips.

Enumeration
1. The enumerated list of food outlets should constantly be updated as raters go out in the field and learn of additions or changes to the food outlets in the neighborhoods; therefore, it is extremely important that all project personnel understand the protocol for updating the master list. (i.e., raters may simply make a note on the cover page of a survey and the project coordinator makes the corrections or additions to the master list). A sample protocol for updating the master list is found in the NEMS-R and NEMS-S Enumeration Procedures.
2. Also, as with other data, always make sure you have a back-up copy, just in case.

Surveys

Fieldwork assignments
1. Create an accurate user-friendly system for **assigning food outlets to raters**.
   a. Create a master log and assign different sections to different raters to complete within a set time frame.
   b. Give out daily assignments to raters. Outlets not completed that day can be reassigned the following day.

2. We recommend performing periodic **reliability checks** between raters.
   a. Assign the same food outlets to two raters on the same day. They can ride together to the food outlets but must complete the surveys independently.
   b. Collect the surveys when the raters return and review them concurrently for discrepancies. Discuss the discrepancies with the raters, so that any issues are resolved.

3. Designate an area for **field supplies** (e.g., cover sheets, surveys, pencils, mileage logs, the pre-printed nutrition guides, etc.).
   a. Maintain a filing cabinet of the nutrition guides downloaded and printed from restaurants' websites, so that others can photocopy them when they are needed, rather than searching the Internet again. Be sure they have dates marked on them, and don't use website information that is more than a few weeks old.
   b. It may also be helpful to organize the collected take-away menus in binders by neighborhood.

Completed Surveys
1. **Establish a clear process for the "flow of data."** Create a system for tracking completed surveys. Again, develop a system which works best for you and your staff. A sample system is outlined in the Quality Assurance document.

    a. It may be best to have a data cover sheet to make sure that each survey was reviewed, entered, and filed.

    b. Alternatively, perhaps making notes on the cover sheet (top page) of each survey will work best for you.

Survey Flow Suggestions

1. Designate **where surveys should be placed** after each step has been completed.
   - The basic piles may be "To be reviewed," "Needs follow up", "To be entered", and "To be filed".
2. **Fieldwork tracking log**.
   - Raters initial next to the food outlets they were able to rate that day,
   - or, create an electronic file in MS Access or Excel.
3. Create a system for raters to '**flag**' surveys.
   - Surveys should be flagged when there are questions or problems about the information on the survey, and a supervisor should be notified of these problems.
4. Establish a **timeframe** for **checking surveys for quality**.
   - In the beginning of data collection or when personnel are new, surveys should be checked as soon as possible after being turned in, so that any protocol errors are caught and addressed before raters go out again to survey more food outlets.
   - Surveys should also be checked for completeness (i.e., all questions are answered, legibly within designated spaces, I.D. information is filled in on all pages, bubbles are filled in correctly, etc.).
5. Establish a clearly labeled **filing system,** so that surveys can be located quickly if needed.

Data Entry

Although few researchers will say that data entry is their favorite part of research, it is extremely important ~ The results of your study depend on the quality of data entry.

Scanning

1. If you are collecting information from a large number of outlets and you have access to a **TeleForm scanner and software**, you may want to use that system to scan/verify the data. If you would like TeleForm versions of the surveys/rating forms, please contact us at nems@sph.emory.edu .
   - a. Scan a few surveys from each rater early on, to make sure that they are darkening the bubbles and numbers enough for the machine to read.
   - b. Also, you may need to have raters enter the comments into the database by hand, because the writing may be too small or unclear for the machine to read and interpret properly.

Data Entry with *Access*

1. We have provided you with **user-friendly data entry screens** that replicate each page of the surveys.
   - a. Remember to share the *Data Entry Tips for Using the NEMS Access Database* document with the staff that will do the data entry.
   - b. Remember to modify the database if any items were added to the surveys.

Data Cleaning

As you would do with any data set, the NEMS data must be screened and cleaned prior to analysis and/or interpretation. Your data cleaning and screening plan will depend on the volume of data collected, the goals of your particular project, and the statistics that will be used.

1. Suggested Basic Steps
   a. Perform simple **minimum and maximum value ranges** for each variable to make sure that all data points fall within a reasonable/possible range.
      i. Example: If the maximum for Number of Healthy Salads was 78 for a restaurant, then someone needs to recheck the hardcopy survey to make sure this large number is actually correct.
   b. Perform **frequency counts** for restaurants and stores (overall, by type, by location, etc) to make sure the counts in the database match the master list.
      i. Example: If the master list indicates that 12 sit-down restaurants were surveyed, but the database only has 11, then there is a problem.
   c. For scale or ratio information (not frequencies or counts), compute the **mean and the standard deviation** for those variables. If there is a value for one of those variables that is +/- three standard deviations from the mean, it may be an outlier.
   d. Compute the **number of blank cells** for each variable and ascertain if this is a reasonable number of blanks or if something may be missing or may have been miscoded.