Data Sources
and
Data Entry
Overview

- Sources of Data
- Data Entry Concepts
- Data Entry Software
  - Example of MS Excel (demo in class)
  - Example of EpiInfo (demo in class)
  - Example of SAS/FSEDIT
- Example Programs
Sources of Data

- Written or electronic logs
- Surveys
  - Telephone
  - Personal interview
  - Mail
- Existing data files
Data Entry Concepts

- Problems
  - Out-of-range values
  - Transposed characters
  - Incorrect values
- Exceptions
  - Missing values
  - Inconsistent values
  - Ambiguous responses
Data Entry Concepts

- Error Checks
  - Up-front checking
  - Backend checking
- Verification
  - Re-entry of given field
  - Double entry of all fields
    - By the same operator
    - By two operators
Data Entry Concepts

- Some techniques to avoid problems
  - Scanning
  - Table look-up
  - Branching
  - Auto-entry
    - Default values
    - Repeated values
Data Entry Software

- Numerous packages on the market
  - Simple ones
    - EpiInfo
    - EasyEntry
  - Spread sheet (e.g. MS Excel)
  - Database (e.g. MS Access)
  - Statistical packages
    - SAS
    - SPSS
Data Entry Software

Examples

- In-class example using MS-Excel
  (See class notes.)

- In-class example using EpiInfo
  (See class notes.)

- Example using SAS/FSEDIT.
Goal: Create a data entry screen to accept

- DATE: The observation date
- IDCODE: 5-character subject identification code
- GENDER: Subject’s gender
  - Code M or m, F or f
- CLASS: The statistics class number and name in which the subject is enrolled.
  - Use a table lookup method to enter the class.
Simple Data Entry with Excel

- Enter variable names in the first row.
- Enter values in subsequent rows.
  - TAB to move from column to column.
  - ENTER to move to beginning of next row.
- Use Validation to perform upfront error checks.
- Use the LOOKUP function for table look-up.
Simple Data Entry with Epi Info

- Use either the DOS or Win2000 version.
- Create a project
- Create a data entry screen (view)
  - Use menus to create upfront error checks
  - Use menus to perform table lookups
SAS/FSEDIT Procedure: Example

- Create a custom informat for table lookup.
- Use PROC FORMAT. More about this later.

* Create an informat to be used in a table-lookup for data entry.;

proc format; * Procedure to create and manage formats/informsats;
    * Create a character informat, saved in the WORK library;
    invalue $class
    208="208: Statistical Thinking"
    210="210: Basic Practice of Statistics"
    212="212: Concepts of Statistics"
    321="321: Introduction to Statistical Computing"
    other="???";
run;
SAS/FSEDIT Procedure: Example

- Create new data set and data entry screen.

```sas
proc fsedit
  new=mysdl.datentry  /* Create Data Set */
  screen=mysdl.datentry.screen /* Entry Screen */;
```

- Enter attributes into the provided spaces.
SAS/FSEDIT Procedure:

Example

- Move the bottom slide-bar to the right to enter informat.

- Enter an informat for DATE and your custom informat for CLASS.
SAS/FSEDIT Procedure: Example

From the File pull-down menu, Close the screen to continue. This new screen will appear:

DATE: 
IDCODE: 
GENDER: 
CLASS: 

FSEDIT Menu
SAS/FSEEDIT Procedure: Example

- Click the FSEEDIT Menu icon
  - You may be asked for a password.
  - Modification of the data entry screen will be demonstrated in class.
Modifying the Entry Screen

- Choose option 2 to modify the screen
  - File/Close to close the screen
  - Answer *no* to the question about computational fields (for this example)
  - Identify fields. Read directions on the bottom of the screen.
  - File/Close to close the screen and return to the Menu.
SAS/FSEEDIT Procedure: Note

You may program many of the data entry checks by using SAS SCL (Screen Control Language). This is a very flexible control language that has many applications beyond data entry. Examples may be seen in the example programs mentioned on the last slide.

We will not cover SCL in this course, except via the given examples. However, lookup Help FSEdit or Help SCL in the command line if you wish to learn more.

- Choose option 3 to add SCL code.
  - File/Close to close and compile the code.
  - Click GOBACK to return to data entry screen.
**SAS/FSEDIT Procedure:**

**Program Statements (SCL)**

```sas
init:
return;

main:
if gender in ("m","M","f","F") then return;
else do;
   error on gender;
   _msg_="Please enter M for male or F for female. ";
   end;
return;

term
return;
```
SAS/FSEDIT Procedure: Example

You are now ready to enter data. You may go back to the modification screen at any time by clicking on the FSEDIT Menu icon.

Click to add new data
SAS/FSEDIT Procedure:

Example

- Next time you enter data into this data set, substitute `data=` for `new=` in the procedure statement.

  -OR-

- You may start entry from the command line by entering: `fsedit mysdl.datentry mysdl.datentry.screen`
Example Programs

- Three example programs are available on the class web site.
  - fsedit.sas (This is the example given in this lecture.)
  - SAS Sample Library.
    - fsedit-1.sas
    - fsedit-2.sas