

Find any data relevant to a topic I am interested in?

It is suggested, **in this order**, to:

Look at the 'list of data collection forms.pdf' CD 2: is it very likely/standard to collect the relevant data in this examination or test?

Go to the 'definition of codes' (data dictionary & codebook) in the original in the original (CD 1) or edited (CD 2 subdirectory 4) pdf documentation for that form and look for variables/codes that are relevant to the topic.

Go to the copy of the form, if it is legible, in the original or edited pdf documentation and examine its specific items.

And/or/then

Read the basic contents of the variable file in vface.pdf (CD 2 subdirectory 3)

Examine the relevant section of 1972 documentation + ...pdf (CD 2.\3\ 1972 documentation +\)

Read the rest of 1972 documentation + ...pdf .

And/or/then

Read COMMON INTRODUCTION.pdf (CD 2 subdirectory 2) which describes and explain the terms and classifications used in the CPP and look for the commonly used terms for your topic.

Look in the original 41-44.pdf (CD 1) or edited III.A ...pdf (CD 2 subdirectory 2) variable file index(ices).

Look in the original 57-60.pdf (CD 1) or edited V...pdf (CD 2 subdirectory 2) master index(ices).

Find the electronic data or variable?

It is suggested, **in this order**, to

Check if that form's data has been input into SAS (XXXXXXX.pdf) and use the variable naming convention (XXXXXXX.pdf) to find it in the SAS dataset.

Or

If it is not input into SAS check if the SAS syntax has been written but no SAS dataset created, most likely with obstetric data and forms,

Or

write your own SAS syntax, perhaps using the syntax supplied as a guide.

Use the SAS syntax as a guide in writing SAS input?

Substitute! But, for understanding look up

Key words: column input conditional input @ missover
 logical record length match-merging

Use other than SAS software - SPSS or STATA or a database like ACCESS or Paradox?

SPSS datasets [will be supplied] in CD 2 subdirectory 4 SPSS 11.0 data

STATA datasets [will be supplied] in CD 2 subdirectory 5 STATA 7 data

or

write your own syntax to input the data into any package by modifying the SAS syntax using 'find and replace' or a text editor with regular expression capability like textpad (txpeng...exe in CD 2 subdirectory 6 'miscellaneous').

SPSS users: punch card records are '**group nested files**'

See sample SPSS syntax (SPSS read in example.pdf).

SPSS has had the capability of reading SAS files: including .sas7 .sas7bdat and transport .xpt since v10,0,5. These SAS datasets must be created with SAS compression off (compress=no).

However, in practice, this is not always successful with these large datasets.

STATA users: use the '**infix**' statement and a data dictionary. See STATA VARFILE.pdf.

or

use data transfer software like DBMS/COPY/ENGINES or STAT TRANSFER, but, again, there have been practical technical difficulties, especially with STATA.

NOTE: BECAUSE OF THE POTENTIAL SIZE OF THE DATASETS THERE ARE PROBLEMS WITH DIRECT TRANSFER FROM SAS DATASETS AND DIRECT INPUT FROM ASCII MAY BE UNAVOIDABLE IN PRODUCING LARGE DATASETS.

Know if the data have been read into the system dataset properly from ASCII?

Run/generate a one-way frequency table for the variable to see if the values are legitimate. For example, if the valid values for sex are 1=male 2=female 9=missing or system missing and none of these are present, or if there are other undocumented values like 3-8, check the input syntax first.