SPSS Complex Samples™ 12.0

SPSS®

Preface

SPSS 12.0 is a powerful software package for microcomputer data management and analysis. The Complex Samples option is an add-on enhancement that provides a set of procedures for selecting and analyzing samples according to complex designs.

The Complex Samples option includes procedures for:

- Specifying and drawing from a complex sample plan.
- Preparing a publicly available complex sample for analysis.
- Frequency tables for complex samples.
- Descriptive statistics for complex samples.
- Crosstabulation tables for complex samples.
- Ratio statistics for complex samples.

Installation

To install Complex Samples, follow the instructions for adding and removing features in the installation instructions supplied with the SPSS Base. (To start, double-click on the SPSS Setup icon.)

Compatibility

SPSS is designed to run on many computer systems. See the materials that came with your system for specific information on minimum and recommended requirements.

Serial Numbers

Your serial number is your identification number with SPSS Inc. You will need this serial number when you call SPSS Inc. for information regarding support, payment, or an upgraded system. The serial number was provided with your Base system.

Customer Service

If you have any questions concerning your shipment or account, contact your local office, listed on the SPSS Web site at *http://www.spss.com/worldwide*. Please have your serial number ready for identification.

Training Seminars

SPSS Inc. provides both public and onsite training seminars. All seminars feature hands-on workshops. Seminars will be offered in major cities on a regular basis. For more information on these seminars, contact your local office, listed on the SPSS Web site at *http://www.spss.com/worldwide*.

Technical Support

The services of SPSS Technical Support are available to registered customers. Customers may contact Technical Support for assistance in using SPSS or for installation help for one of the supported hardware environments. To reach Technical Support, see the SPSS Web site at *http://www.spss.com*, or contact your local office, listed on the SPSS Web site at *http://www.spss.com/worldwide*. Be prepared to identify yourself, your organization, and the serial number of your system.

Additional Publications

Additional copies of SPSS product manuals may be purchased directly from SPSS Inc. Visit our Web site at *http://www.spss.com/estore*, or contact your local SPSS office, listed on the SPSS Web site at *http://www.spss.com/worldwide*. For telephone orders in the United States and Canada, call SPSS Inc. at 800-543-2185. For telephone orders outside of North America, contact your local office, listed on the SPSS Web site.

The *SPSS Statistical Procedures Companion*, by Marija Norušis, is being prepared for publication by Prentice Hall. It contains overviews of the procedures in the SPSS Base, plus Logistic Regression, General Linear Models, and Linear Mixed Models. Further information will be available on the SPSS Web site at *http://www.spss.com* (click Store, select your country, and click Books).

Tell Us Your Thoughts

Your comments are important. Please send us a letter and let us know about your experiences with SPSS products. We especially like to hear about new and interesting applications using the SPSS system. Write to SPSS Inc., Attn.: Director of Product Planning, 233 South Wacker Drive, 11th Floor, Chicago, IL 60606-6412. You can also send e-mail to *suggest@spss.com*.

About This Manual

This manual documents the graphical user interface for the procedures included in the Complex Samples module. Illustrations of dialog boxes are taken from SPSS for Windows. Dialog boxes in other operating systems are similar. The Complex Samples command syntax is included in the SPSS 12.0 Syntax Reference Guide, available on the product CD-ROM.

Contacting SPSS

If you would like to be on our mailing list, contact one of our offices, listed on our Web site at *http://www.spss.com/worldwide*. We will send you a copy of our newsletter and let you know about SPSS Inc. activities in your area.

Contents

4 Complex Samples Plan   35

5 Complex Samples Frequencies   37

6 Complex Samples Descriptives   43

7 Complex Samples Crosstabs   49

8 Complex Samples Ratios 55

9 Complex Samples Sampling Wizard 61

10 Complex Samples Analysis Preparation Wizard 95

Introduction to SPSS Complex Samples Procedures

An inherent assumption of analytical procedures in traditional software packages is that the observations in a data file represent a simple random sample from the population of interest. This assumption is untenable for an increasing number of companies and researchers who find it both cost-effective and convenient to obtain samples in a more structured way.

The SPSS Complex Samples option allows you to select a sample according to a complex design and incorporate the design specifications into the data analysis, thus ensuring that your results are valid.

Properties of Complex Samples

A complex sample can differ from a simple random sample in many ways. In a simple random sample, individual sampling units are selected at random with equal probability and without replacement (WOR) directly from the entire population. By contrast, a given complex sample can have some or all of the following features:

Stratification. Stratified sampling involves selecting samples independently within non-overlapping subgroups of the population, or strata. For example, strata may be socioeconomic groups, job categories, age groups, or ethnic groups. With stratification, you can ensure adequate sample sizes for subgroups of interest, improve the precision of overall estimates, and use different sampling methods from stratum to stratum.

Clustering. Cluster sampling involves the selection of groups of sampling units, or clusters. For example, clusters may be schools, hospitals, or geographical areas, and sampling units may be students, patients, or citizens. Clustering is common in multistage designs and area (geographic) samples.

Multiple stages. In multistage sampling, you select a first-stage sample based on clusters. Then you create a second-stage sample by drawing subsamples from the selected clusters. If the second-stage sample is based on subclusters, you can then add a third stage to the sample. For example, in the first stage of a survey, a sample of cities could be drawn. Then from the selected cities, households could be sampled. Finally, from the selected households, individuals could be polled. The Sampling and Analysis Preparation wizards allow you to specify three stages in a design.

Nonrandom sampling. When selection at random is difficult to obtain, units can be sampled systematically (at a fixed interval) or sequentially.

Unequal selection probabilities. When sampling clusters that contain unequal numbers of units, you can use probability-proportional-to-size (PPS) sampling to make a cluster's selection probability equal to the proportion of units it contains. PPS sampling can also use more general weighting schemes to select units.

Unrestricted sampling. Unrestricted sampling selects units with replacement (WR), thus an individual unit can be selected for the sample more than once.

Sampling weights. Sampling weights are automatically computed while drawing a complex sample and ideally correspond to the "frequency" that each sampling unit represents in the target population. Therefore, the sum of the weights over the sample should estimate the population size. Complex Samples analysis procedures require sampling weights in order to properly analyze a complex sample. Note that these weights should be used entirely within the Complex Samples option and should not be used with other analytical procedures via the Weight Cases procedure, which treats weights as case replications.

Usage of Complex Samples Procedures

Your usage of Complex Samples procedures depends on your particular needs. The primary types of users are those who:

■ Plan and carry out surveys according to complex designs, possibly analyzing the sample later. The primary tool for surveyors is the Sampling Wizard.

■ Analyze sample data files previously obtained according to complex designs. Before using the Complex Samples analysis procedures, you may need to use the Analysis Preparation Wizard.

Regardless of which type of user you are, you need to supply design information to Complex Samples procedures. This information is stored in a **plan file** for easy reuse.

Plan Files

A plan file contains complex sample specifications. There are two types of plan files:

Sampling plan. The specifications given in the Sampling Wizard define a sample design that is used to draw a complex sample. The sampling plan file contains those specifications. The sampling plan file also contains a default analysis plan that uses estimation methods suitable for the specified sample design.

Analysis plan. This plan file contains information needed by Complex Samples analysis procedures to properly compute variance estimates for a complex sample. The plan includes the sample structure, estimation methods for each stage, and references to required variables, such as sample weights. The Analysis Preparation Wizard allows you to create and edit analysis plans.

There are several advantages to saving your specifications in a plan file, including:

■ A surveyor can specify the first stage of a multistage sampling plan and draw first-stage units now, collect information on sampling units for the second stage, and then modify the sampling plan to include the second stage.

- An analyst who doesn't have access to the sampling plan file can specify an analysis plan and refer to that plan from each Complex Samples analysis procedure.

- A designer of large-scale public use samples can publish the sampling plan file, thus simplifying the instructions for analysts and avoiding the need for each analyst to specify his or her own analysis plans.

Further Readings

For more information on sampling techniques, see the following texts:

Cochran, W. G. 1977. *Sampling Techniques.* New York: John Wiley and Sons.

Kish, L. 1965. *Survey Sampling.* New York: John Wiley and Sons.

Kish, L. 1987. *Statistical Design for Research.* New York: John Wiley and Sons.

Murthy, M. N. 1967. *Sampling Theory and Methods.* Calcutta, India: Statistical Publishing Society.

2

Sampling from a Complex Design

Figure 2-1
Sampling Wizard Welcome step
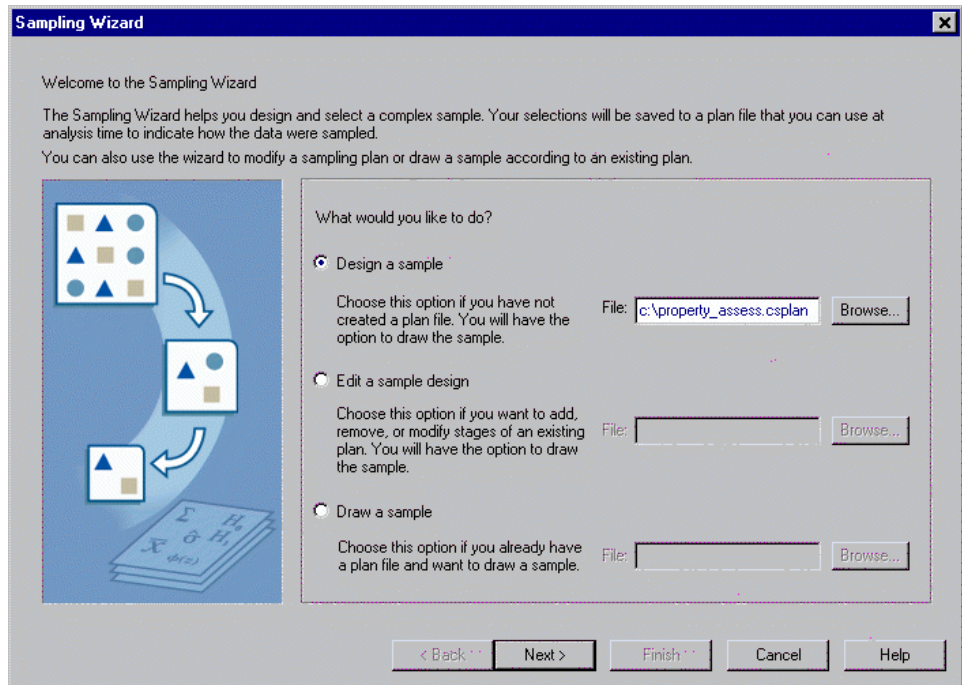
The Sampling Wizard guides you through the steps for creating, modifying, or executing a sampling plan file. Before using the Wizard, you should have a well-defined target population, a list of sampling units, and an appropriate sample design in mind.

Creating a New Sample Plan

▶ From the menus choose:
    Analyze
      Complex Samples
        Select a Sample...

▶ Select Design a sample and choose a plan filename to save the sample plan.

▶ Click Next to continue through the Wizard.

▶ Optionally, in the Define Variables step, you can define strata, clusters, and input sample weights. After you define these, click Next.

▶ Optionally, in the Sampling Method step, you can choose a method for selecting items.

   If you select PPS Brewer or PPS Murthy, you can click Finish to draw the sample. Otherwise, click Next and then:

▶ In the Sample Size step, specify the number or proportion of units to sample.

   You can now click Finish to draw the sample. Optionally, in further steps, you can:
   ■ Choose output variables to save.
   ■ Add a second or third stage to the design.
   ■ Set various selection options, including which stages to draw samples from, the random number seed, and whether to treat user-missing values as valid values of design variables.
   ■ Choose where to save output data.
   ■ Paste your selections as command syntax.

Sampling Wizard: Design Variables

Figure 2-2
Sampling Wizard Design Variables step
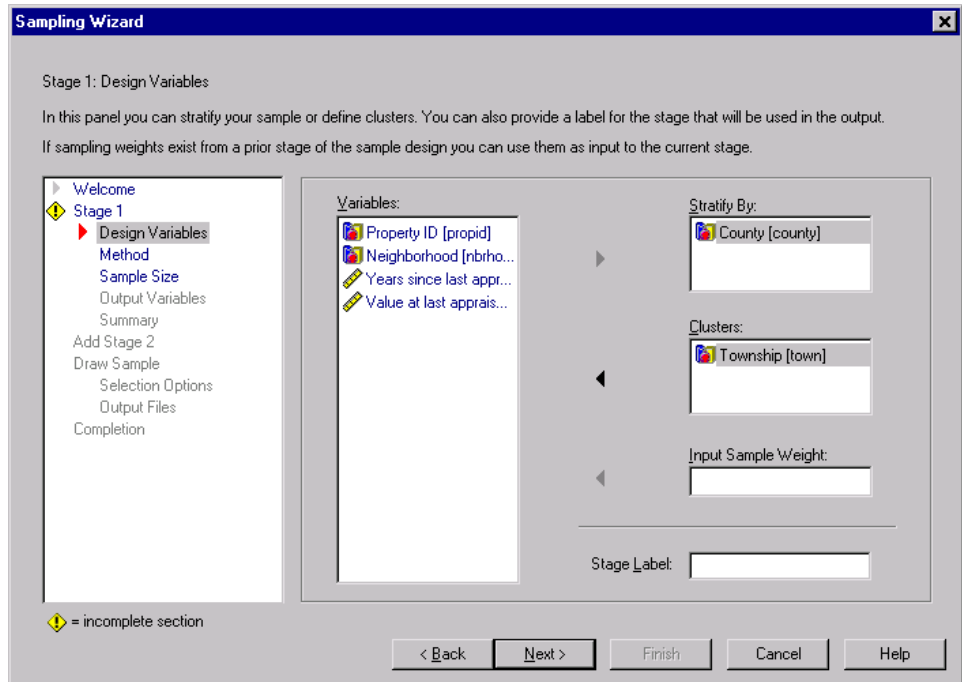
This step allows you to select stratification and clustering variables and to define input sample weights. You can also specify a label for the stage.

Stratify By. The cross-classification of stratification variables defines distinct subpopulations, or strata. Separate samples are obtained for each stratum. To improve the precision of your estimates, units within strata should be as homogeneous as possible for the characteristics of interest.

Clusters. Cluster variables define groups of observational units, or clusters. Clusters are useful when directly sampling observational units from the population is expensive or impossible; instead, you can sample clusters from the population and then sample observational units from the selected clusters. However, the use of clusters can introduce correlations among sampling units, resulting in a loss of

precision. To minimize this effect, units within clusters should be as heterogeneous as possible for the characteristics of interest. You must define at least one cluster variable in order to plan a multistage design. Clusters are also necessary in the use of several different sampling methods. For more information, see "Sampling Wizard: Sampling Method" on page 9.

Input Sample Weight. If the current sample design is part of a larger sample design, you may have sample weights from a previous stage of the larger design. You can specify a numeric variable containing these weights in the first stage of the current design. Sample weights are computed automatically for subsequent stages of the current design.

Stage Label. You can specify an optional string label for each stage. This is used in the output to help identify stagewise information.

Note: The source variable list has the same content across steps of the Wizard. In other words, variables removed from the source list in a particular step are removed from the list in all steps. Variables returned to the source list appear in the list in all steps.

Tree Controls for Navigating the Sampling Wizard

On the left side of each step in the Sampling Wizard is an outline of all the steps. You can navigate the Wizard by clicking on the name of an enabled step in the outline. Steps are enabled as long as all previous steps are valid—that is, if each previous step has been given the minimum required specifications for that step. See the Help for individual steps for more information on why a given step may be invalid.

Sampling Wizard: Sampling Method

Figure 2-3
Sampling Wizard Method step
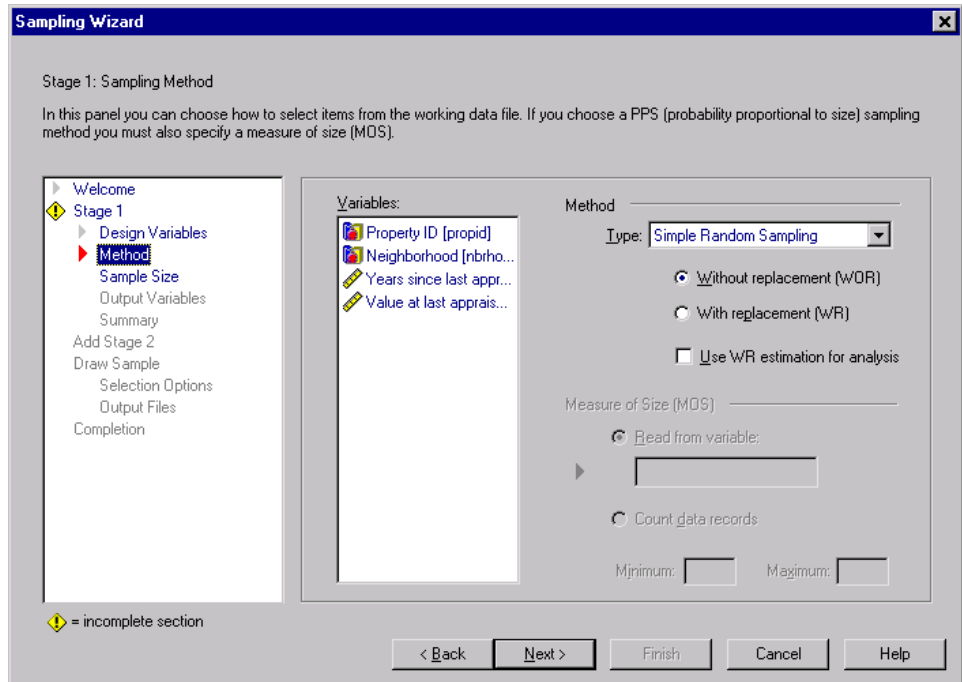
This step allows you to specify how to select cases from the working data file.

Method. Controls in this group are used to choose a selection method. Some sampling types allow you to choose whether to sample with replacement (WR) or without replacement (WOR). See the type descriptions for more information. Note that some probability-proportional-to-size (PPS) types are available only when clusters have been defined and that all PPS types are available only in the first stage of a design. Moreover, WR methods are available only in the last stage of a design.

- **Simple Random Sampling.** Units are selected with equal probability. They can be selected with or without replacement.

- **Simple Systematic.** Units are selected at a fixed interval throughout the sampling frame (or strata, if they have been specified) and extracted without replacement. A randomly selected unit within the first interval is chosen as the starting point.

- **Simple Sequential.** Units are selected sequentially with equal probability and without replacement.

- **PPS.** This is a first-stage method that selects units at random with probability proportional to size. Any units can be selected with replacement; only clusters can be sampled without replacement.

- **PPS Systematic.** This is a first-stage method that systematically selects units with probability proportional to size. They are selected without replacement.

- **PPS Sequential.** This is a first-stage method that sequentially selects units with probability proportional to cluster size and without replacement.

- **PPS Brewer.** This is a first-stage method that selects two clusters from each stratum with probability proportional to cluster size and without replacement. A cluster variable must be specified to use this method.

- **PPS Murthy.** This is a first-stage method that selects two clusters from each stratum with probability proportional to cluster size and without replacement. A cluster variable must be specified to use this method.

- **PPS Sampford.** This is a first-stage method that selects more than two clusters from each stratum with probability proportional to cluster size and without replacement. It is an extension of Brewer's method. A cluster variable must be specified to use this method.

- **Use WR estimation for analysis.** By default, an estimation method is specified in the plan file that is consistent with the selected sampling method. This allows you to use with-replacement estimation even if the sampling method implies WOR estimation. This option is available only in stage 1.

Measure of Size (MOS). If a PPS method is selected, you must specify a measure of size that defines the size of each unit. These sizes can be explicitly defined in a variable or they can be computed from the data. Optionally, you can set lower and upper bounds on the MOS, overriding any values found in the MOS variable or computed from the data. These options are available only in stage 1.

Sampling Wizard: Sample Size

Figure 2-4
Sampling Wizard Sample Size step
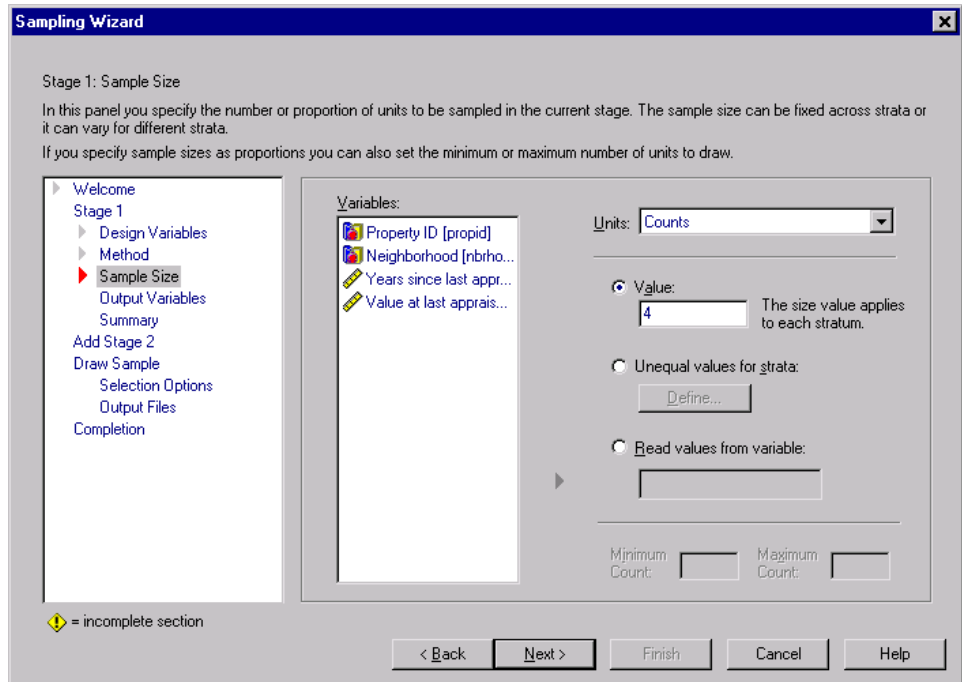
This step allows you to specify the number or proportion of units to sample within the current stage. The sample size can be fixed or it can vary across strata. For the purpose of specifying sample size, clusters chosen in previous stages can be used to define strata.

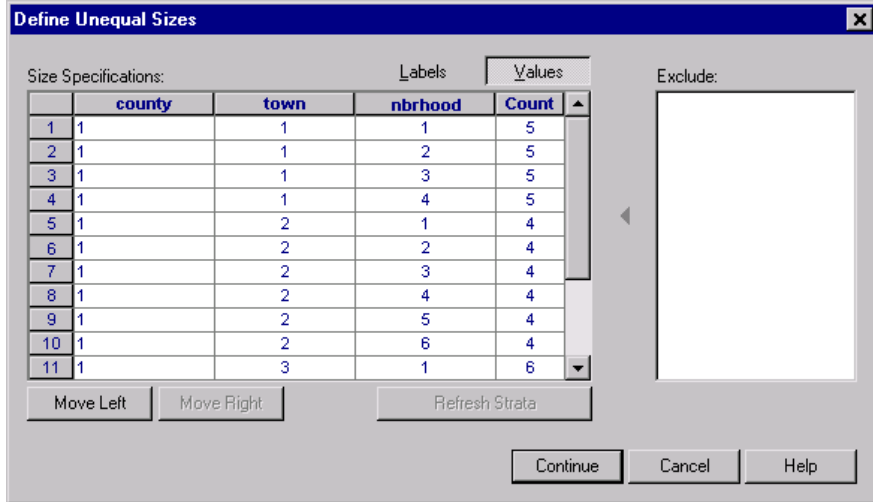**Units.** You can specify an exact sample size or a proportion of units to sample.

- **Value.** A single value is applied to all strata. If Counts is selected as the unit metric, you should enter a positive integer. If Proportions is selected, you should enter a non-negative value. Unless sampling with replacement, proportion values should also be no greater than 1.

■ **Unequal values for strata.** Allows you to enter size values on a per-stratum basis via the Define Unequal Sizes dialog box.

■ **Read values from variable.** Allows you to select a numeric variable that contains size values for strata.

If Proportions is selected, you have the option to set lower and upper bounds on the number of units sampled.

Define Unequal Sizes

Figure 2-5
Define Unequal Sizes dialog box

The Define Unequal Sizes dialog box allows you to enter sizes on a per-stratum basis.

Size Specifications grid. The grid displays the cross-classifications of up to five strata or cluster variables, one stratum/cluster combination per row. Eligible grid variables include all stratification variables from the current and previous stages and all cluster variables from previous stages. Variables can be reordered within the grid or moved to the Exclude list. Enter sizes in the rightmost column. Click Labels or Values to toggle the display of value labels and data values for stratification and cluster variables in the grid cells. Cells that contain unlabeled values always show values.

Click Refresh Strata to repopulate the grid with each combination of labeled data values for variables in the grid.

Exclude. To specify sizes for a subset of stratum/cluster combinations, move one or more variables to the Exclude list. These variables are not used to define sample sizes.

Sampling Wizard: Output Variables

Figure 2-6
Sampling Wizard Output Variables step
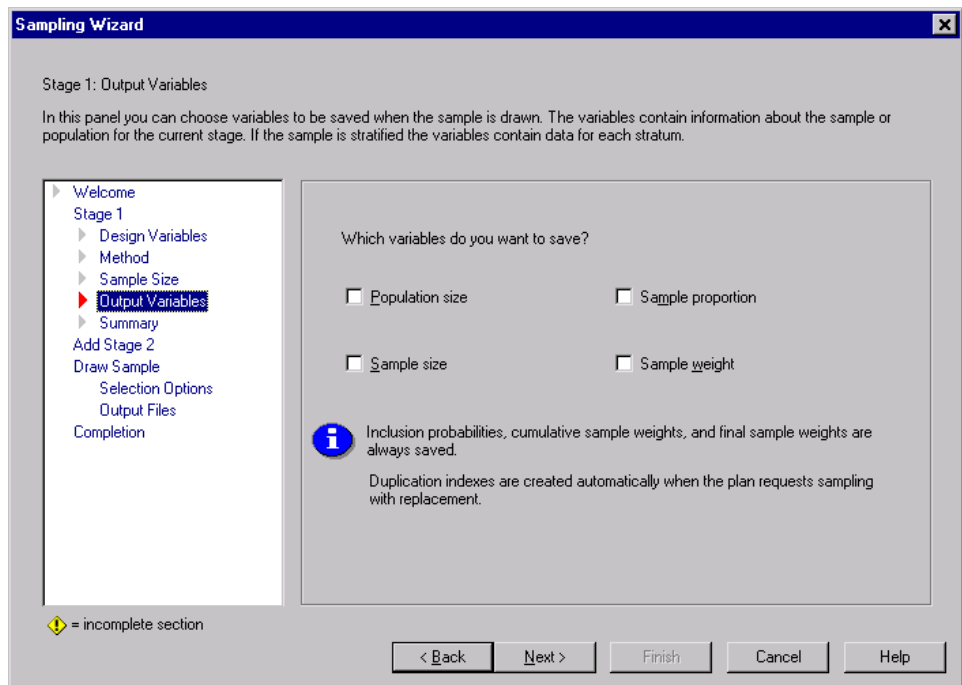
This step allows you to choose variables to save when the sample is drawn.

Population size. The estimated number of units in the population for a given stage. The root name for the saved variable is *PopulationSize_*.

Sample proportion. The sampling rate at a given stage. The root name for the saved variable is *SamplingRate_*.

Sample size. The number of units drawn at a given stage. The root name for the saved variable is *SampleSize_*.

Sample weight. This is the inverse of the inclusion probabilities. The root name for the saved variable is *SampleWeight_*.

Some stagewise variables are generated automatically. These include:

Inclusion probabilities. The proportion of units drawn at a given stage. The root name for the saved variable is *InclusionProbability_*.

Cumulative weight. The cumulative sample weight over stages previous to and including the current one. The root name for the saved variable is *SampleWeightCumulative_*.

Index. Identifies units selected multiple times within a given stage. The root name for the saved variable is *Index_*.

Note: Saved variable root names include an integer suffix that reflects the stage number—for example, *PopulationSize_1_* for the saved population size for stage 1.

Sampling Wizard: Plan Summary

Figure 2-7
Sampling Wizard Plan Summary step
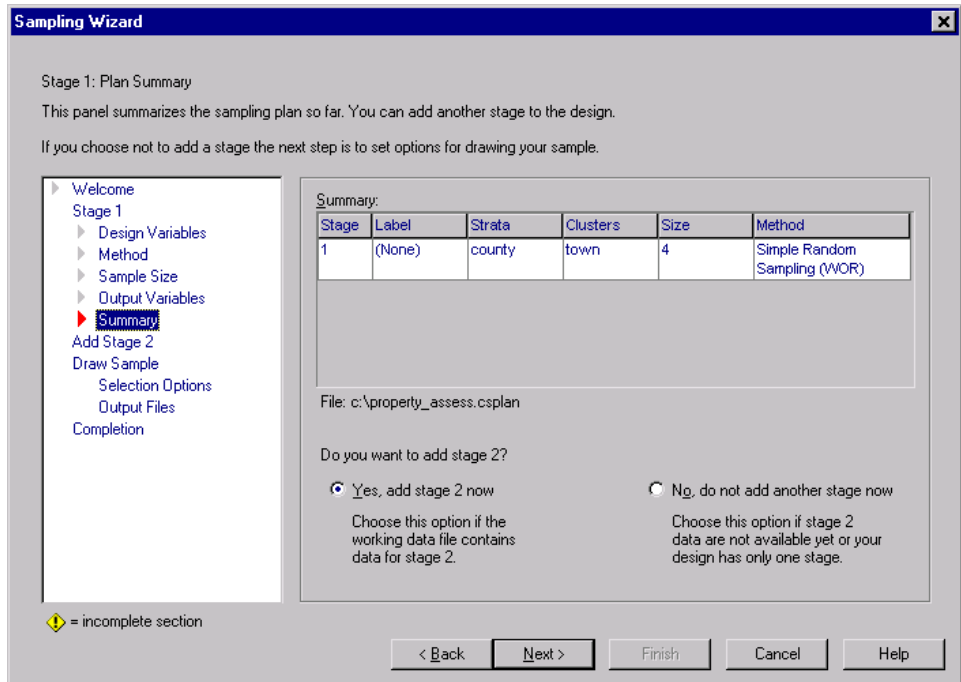
This is the last step within each stage, providing a summary of the sample design specifications through the current stage. From here, you can either proceed to the next stage (creating it, if necessary) or set options for drawing the sample.

Sampling Wizard: Draw Sample Selection Options

Figure 2-8
Sampling Wizard Draw Sample Selection Options step
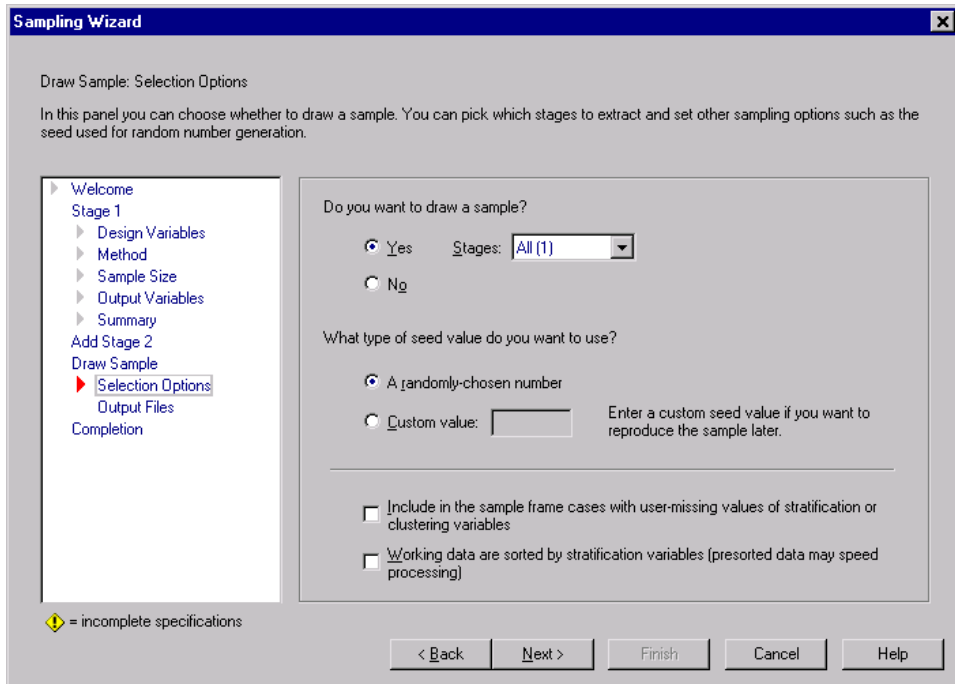
This step allows you to choose whether to draw a sample. You can also control other sampling options, such as the random seed and missing value handling.

Draw sample. In addition to choosing whether to draw a sample, you can also choose to execute part of the sampling design. Stages must be drawn in order—that is, stage 2 cannot be drawn unless stage 1 is also drawn. When editing or executing a plan, you cannot resample locked stages.

Seed. This allows you to choose a seed value for random number generation.

Include user-missing values. This determines whether user-missing values are valid. If so, user-missing values are treated as a separate category.

Data already sorted. If your sample frame is presorted by the values of the stratification variables, this option allows you to speed the selection process.

Sampling Wizard: Draw Sample Output Files

Sampling Wizard Draw Sample Output Files step
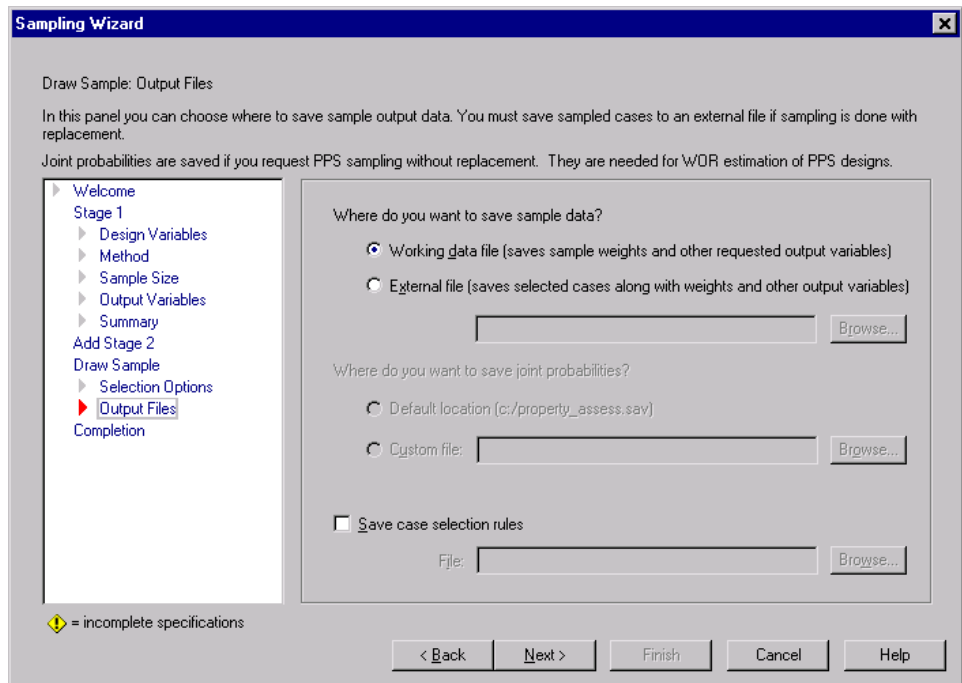
This step allows you to choose where to direct sampled cases, weight variables, joint probabilities, and case selection rules.

Sample data. These options let you determine where sample output is written. It can be added to the working data file or saved to an external file. If an external file is specified, the sampling output variables and variables in the working data file for the selected cases are saved to the file.

Joint probabilities. These options let you determine where joint probabilities are written. Joint probabilities are produced if the PPS WOR, PPS Brewer, PPS Sampford, or PSS Murthy method is selected and WR estimation is not specified.

Case selection rules. If you are constructing your sample one stage at a time, you may want to save the case selection rules to a text file. They are useful for constructing the subframe for subsequent stages.

Sampling Wizard: Finish

Figure 2-10
Sampling Wizard Finish step
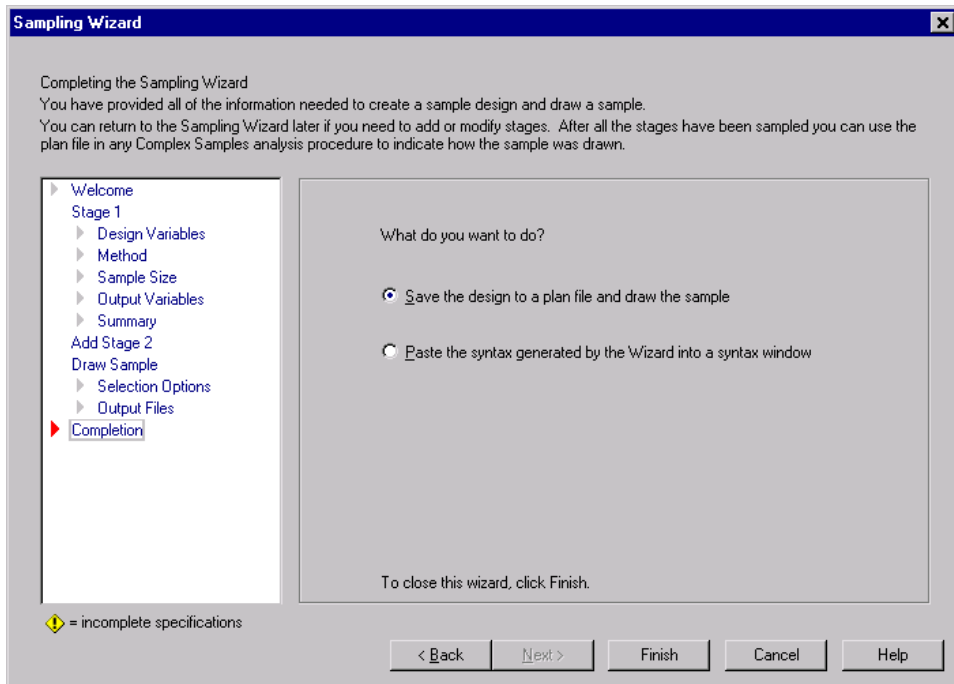
This is the final step. You can save the plan file and draw the sample now or paste your selections into a syntax window.

When editing a plan, you can save the edited plan to a new file or overwrite the existing plan file.

Modifying an Existing Sample Plan

▶ From the menus choose:

Analyze
  Complex Samples
    Select a Sample...

▶ Select Edit a sample design and choose a plan file to edit.

▶ Click Next to continue through the Wizard.

▶ Review the sampling plan in the Plan Summary step, and then click Next.

Subsequent steps are largely the same as for a new design. See the Help for individual steps for more information.

▶ Navigate to the Finish step, and specify a new name for the edited plan file or choose to overwrite the existing plan file.

Optionally, you can:

- Specify stages that have already been sampled.
- Remove stages from the plan.

Sampling Wizard: Plan Summary

Figure 2-11
Sampling Wizard Plan Summary step
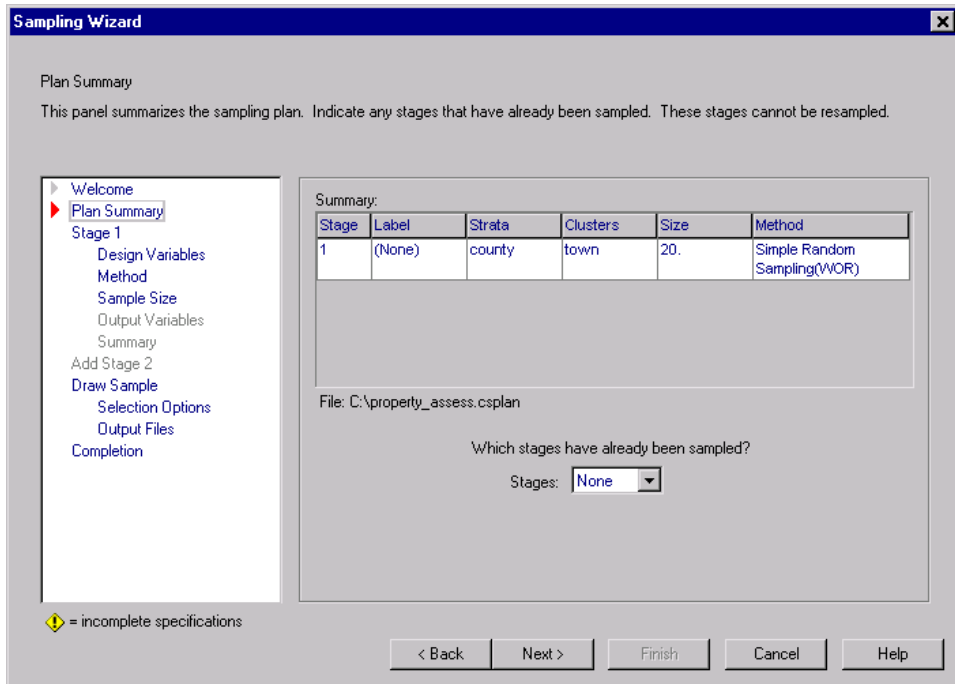
This step allows you to review the sampling plan and indicate stages that have already been sampled. If editing a plan, you can also remove stages from the plan.

Previously sampled stages. If an extended sampling frame is not available, you will have to execute a multistage sampling design one stage at a time. Select which stages have already been sampled from the drop-down list. Any stages that have been executed are locked; they are not available in the Draw Sample Selection Options step, and they cannot be altered when editing a plan.

Remove stages. You can remove stages 2 and 3 from a multistage design.

Running an Existing Sample Plan

▶ From the menus choose:

Analyze
 Complex Samples
  Select a Sample...

▶ Select Draw a sample and choose a plan file to run.

▶ Click Next to continue through the Wizard.

▶ Review the sampling plan in the Plan Summary step, and then click Next.

▶ The individual steps containing stage information are skipped when executing a sample plan. You can now go on to the Finish step at any time.

Optionally, you can:

■ Specify stages that have already been sampled.

# CSPLAN and CSSELECT Commands Additional Features

The SPSS command language also allows you to:

■ Specify custom names for output variables.

■ Control the output in the Viewer. For example, you can suppress the stagewise summary of the plan that is displayed if a sample is designed or modified, suppress the summary of the distribution of sampled cases by strata that is shown if the sample design is executed, and request a case processing summary.

■ Choose a subset of variables in the working data file to write to an external sample file.

See the *SPSS Command Syntax Reference* for complete syntax information.

Preparing a Complex Sample for Analysis

Figure 3-1
Analysis Preparation Wizard Welcome step
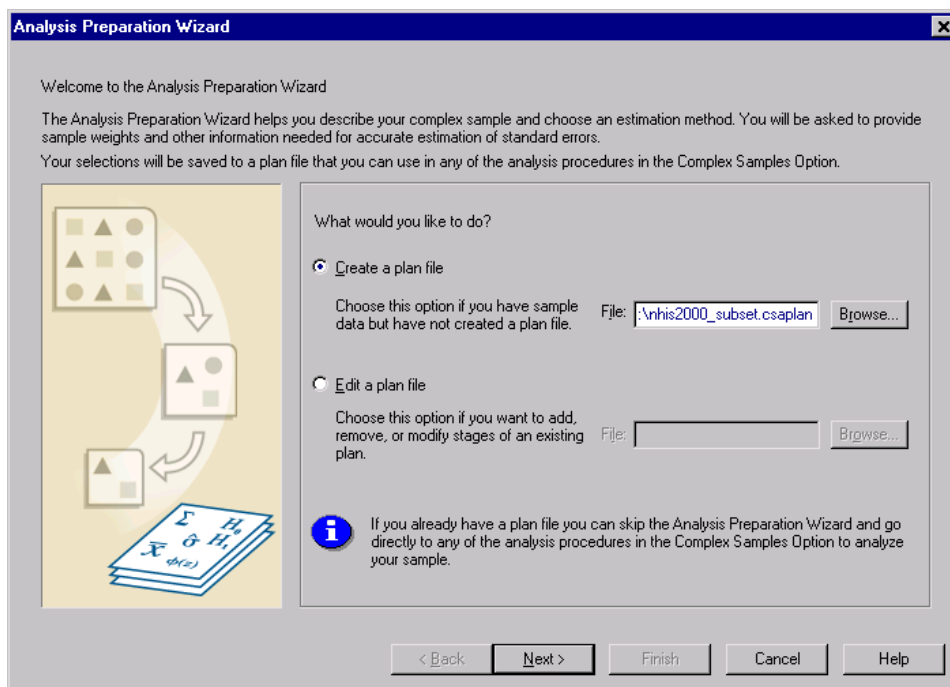
The Analysis Preparation Wizard guides you through the steps for creating or modifying an analysis plan for use with the various Complex Samples analysis procedures. Before using the Wizard, you should have a sample drawn according to a complex design.

Creating a new plan is most useful when you do not have access to the sampling plan file used to draw the sample (recall that the sampling plan contains a default analysis plan). If you do have access to the sampling plan file used to draw the sample, you can use the default analysis plan contained in the sampling plan file or override the default analysis specifications and save your changes to a new file.

Creating a New Analysis Plan

▶ From the menus choose:
Analyze
  Complex Samples
    Prepare for Analysis...

▶ Select Create a plan file, and choose a plan filename to which you will save the analysis plan.

▶ Click Next to continue through the Wizard.

▶ Specify the variable containing sample weights in the Design Variables step, optionally defining strata and clusters.

You can now click Finish to save the plan. Optionally, in further steps you can:

■ Select the method for estimating standard errors in the Estimation Method step.

■ Specify the number of units sampled or the inclusion probability per unit in the Size step.

■ Add a second or third stage to the design.

■ Paste your selections as command syntax.

Analysis Preparation Wizard: Design Variables

Figure 3-2
Analysis Preparation Wizard Design Variables step
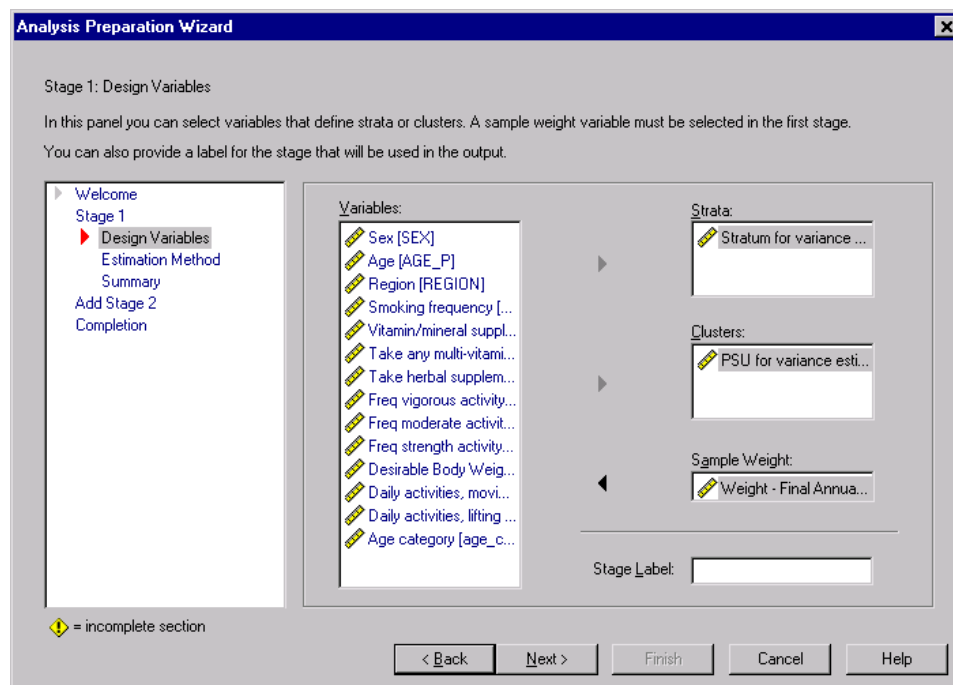
This step allows you to identify the stratification and clustering variables and define sample weights. You can also provide a label for the stage.

Strata. The cross-classification of stratification variables define distinct subpopulations, or strata. Your total sample represents the combination of independent samples from each stratum.

Clusters. Cluster variables define groups of observational units, or clusters. Samples drawn in multiple stages select clusters in the earlier stages and then subsample units from the selected clusters. When analyzing a data file obtained by sampling clusters with replacement, you should include the duplication index as a cluster variable.

Sample Weights. You must provide sample weights in the first stage. Sample weights are computed automatically for subsequent stages of the current design.

Stage Label. You can specify an optional string label for each stage. This is used in the output to help identify stagewise information.

Note: The source variable list has the same contents across steps of the Wizard. In other words, variables removed from the source list in a particular step are removed from the list in all steps. Variables returned to the source list show up in all steps.

Tree Controls for Navigating the Analysis Wizard

At the left side of each step of the Analysis Wizard is an outline of all the steps. You can navigate the Wizard by clicking on the name of an enabled step in the outline. Steps are enabled as long as all previous steps are valid–that is, as long as each previous step has been given the minimum required specifications for that step. For more information on why a given step may be invalid, see the Help for individual steps.

Analysis Preparation Wizard: Estimation Method

Figure 3-3
Analysis Preparation Wizard Estimation Method step
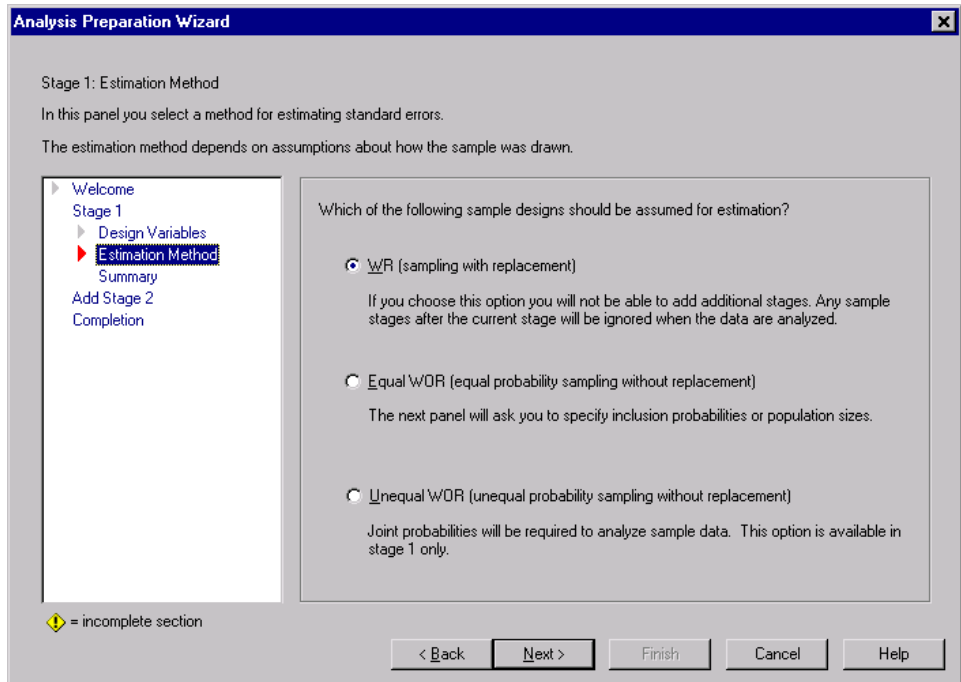
This step allows you to specify an estimation method for the stage.

WR (sampling with replacement). WR estimation does not include a correction for sampling from a finite population, since it assumes that the sample was taken from an infinite population. When the population for the stage is much larger than the sample, this is a reasonable assumption. WR estimation can be specified only in the final stage of a design; the Wizard will not allow you to add another stage if you select WR estimation.

Equal WOR (equal probability sampling without replacement). Equal WOR estimation includes the finite population correction and assumes that units are sampled with equal probability. Equal WOR can be specified in any stage of a design.

Unequal WOR (unequal probability sampling without replacement). In addition to using the finite population correction, Unequal WOR accounts for sampling units (usually clusters) selected with unequal probability. This estimation method is available only in the first stage.

Analysis Preparation Wizard: Size

Figure 3-4
Analysis Preparation Wizard Size step
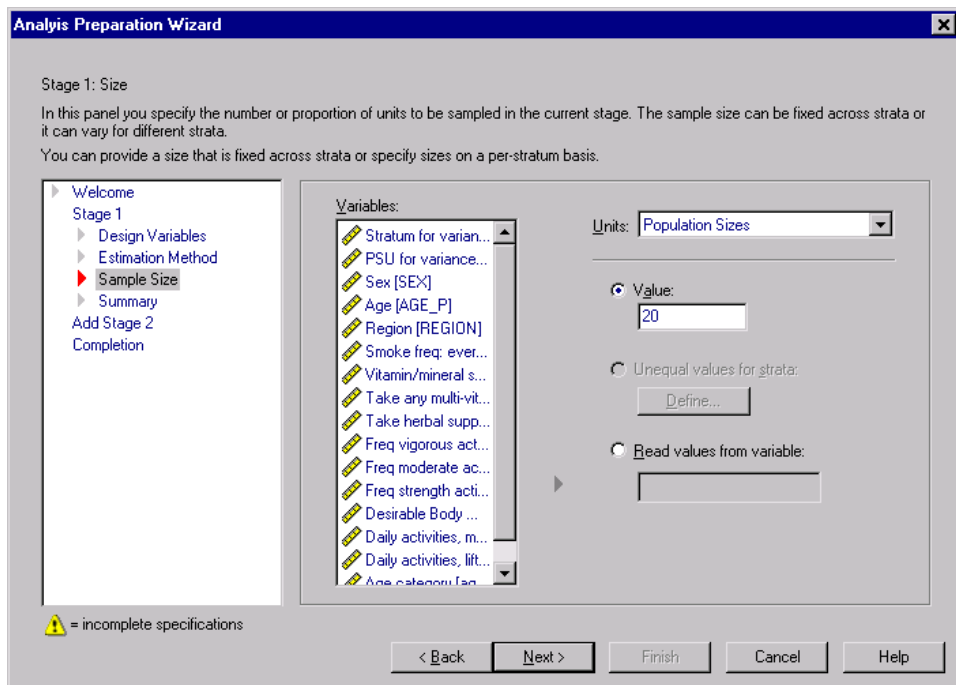
This step is used to specify inclusion probabilities or population sizes for the current stage. Sizes can be fixed or can vary across strata. For the purpose of specifying sizes, clusters specified in previous stages can be used to define strata.

Units. You can specify exact population sizes or the probabilities with which units were sampled.

■ **Value.** A single value is applied to all strata. If Population Sizes is selected as the unit metric, you should enter a non-negative integer. If Inclusion Probabilities is selected, you should enter a value between 0 and 1, inclusive.

■ **Unequal values for strata.** Allows you to enter size values on a per-stratum basis via the Define Unequal Sizes dialog box.

■ **Read values from variable.** Allows you to select a numeric variable that contains size values for strata.

Define Unequal Sizes

Figure 3-5
Define Unequal Sizes dialog box

The Define Unequal Sizes dialog box allows you to enter sizes on a per-stratum basis.

Size Specifications grid. The grid displays the cross-classifications of up to five strata or cluster variables, one stratum/cluster combination per row. Eligible grid variables include all stratification variables from the current and previous stages and all cluster variables from previous stages. Variables can be reordered within the grid or moved to the Exclude list. Enter sizes in the rightmost column. Click Labels or Values to toggle the display of value labels and data values for stratification and cluster variables in the grid cells. Cells that contain unlabeled values always show values.

Click Refresh Strata to repopulate the grid with each combination of labeled data values for variables in the grid.

Exclude. To specify sizes for a subset of stratum/cluster combinations, move one or more variables to the Exclude list. These variables are not used to define sample sizes.

Analysis Preparation Wizard: Stage Summary

Figure 3-6
Analysis Preparation Wizard Stage Summary step
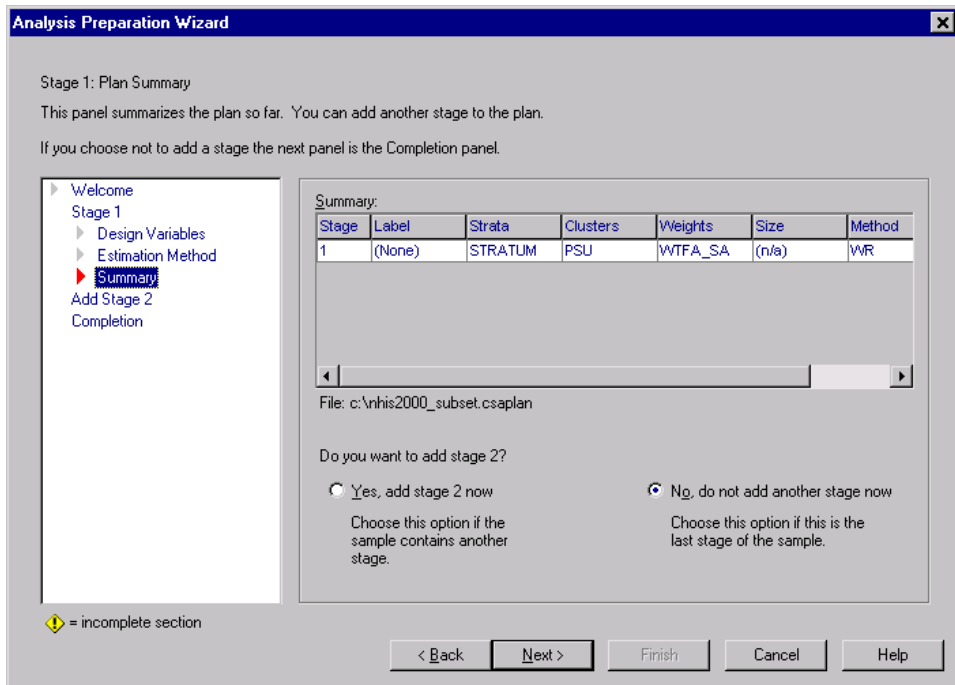
This is the last step within each stage, providing a summary of the analysis design specifications through the current stage. From here, you can either proceed to the next stage (creating it if necessary) or save the analysis specifications.
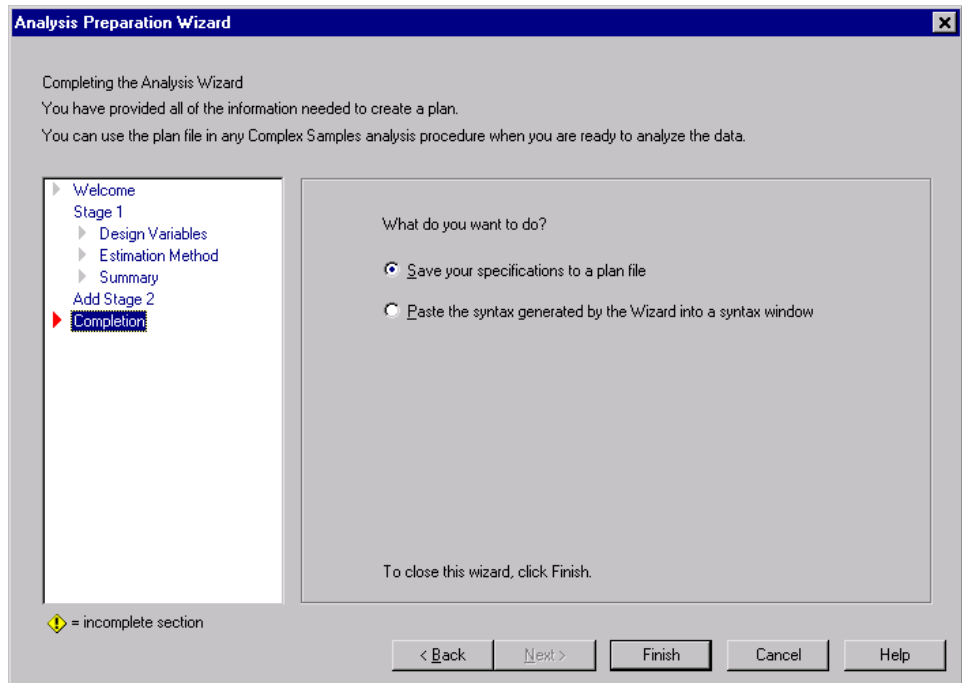
If you cannot add another stage, it is likely because:

■ No cluster variable was specified in the Design Variables step.

- You selected WR estimation in the Estimation Method step.

- This is the third stage of the analysis, and the Wizard supports a maximum of three stages.

Analysis Preparation Wizard: Finish

Figure 3-7
Analysis Preparation Wizard Finish step

This is the final step. You can save the plan file now or paste your selections to a syntax window.

When editing a plan, you can save the edited plan to a new file or overwrite the existing plan file.

Modifying an Existing Analysis Plan

▶ From the menus choose:

Analyze
  Complex Samples
    Prepare for Analysis...

▶ Select Edit a plan file, and choose a plan filename to which you will save the analysis plan.

▶ Click Next to continue through the Wizard.

▶ Review the analysis plan in the Plan Summary step, and then click Next.

Subsequent steps are largely the same as for a new design. See the help for individual steps for more information.

▶ Navigate to the Finish step, and specify a new name for the edited plan file or choose to overwrite the existing plan file.

Optionally, you can:

■ Remove stages from the plan.

Analysis Preparation Wizard: Plan Summary

Figure 3-8
*Analysis Preparation Wizard Plan Summary step*
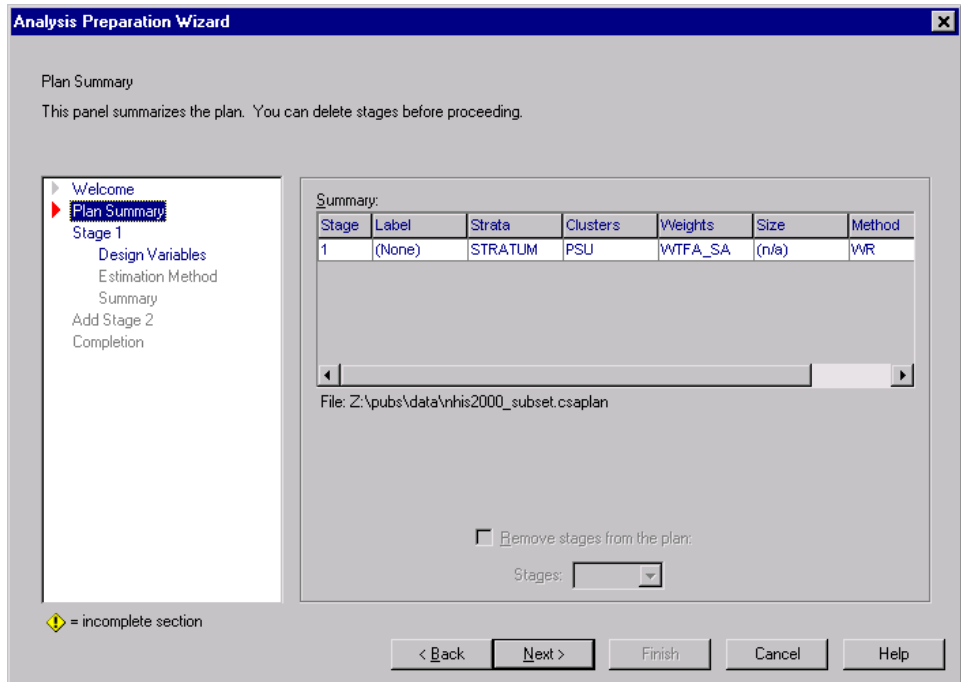
This step allows you to review the analysis plan and remove stages from the plan.

Remove Stages. You can remove stages 2 and 3 from a multistage design. Since a plan must have at least one stage, you can edit but not remove stage 1 from the design.

Complex Samples Plan

Complex Samples analysis procedures require analysis specifications from an analysis or sample plan file in order to provide valid results.

Figure 4-1
Complex Samples Plan dialog box
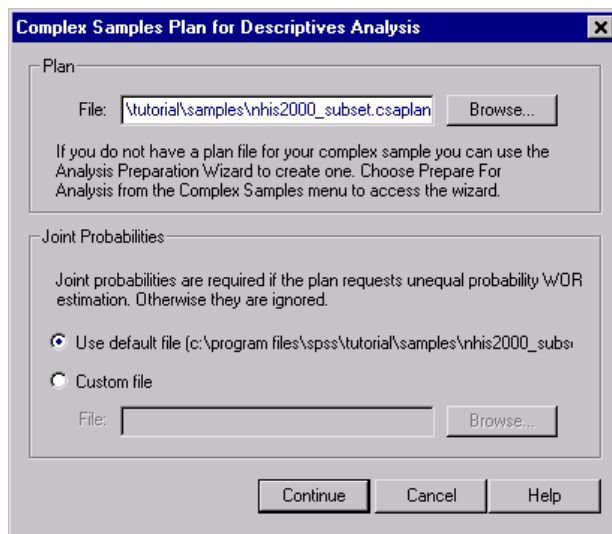
Plan. Specify the path of an analysis or sample plan file.

Joint Probabilities. In order to use Unequal WOR estimation for clusters drawn using a PPS WOR method, you need to specify a separate file containing the joint probabilities. This file is created by the Sampling Wizard during sampling.

Complex Samples Frequencies

The Complex Samples Frequencies procedure produces frequency tables for selected variables and displays univariate statistics. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

Example. Using the Complex Samples Frequencies procedure, you can obtain univariate tabular statistics for vitamin usage among U.S. citizens, based on the results of the National Health Interview Survey (NHIS) and with an appropriate analysis plan for this public use data.

Statistics. The procedure produces estimates of cell population sizes and table percentages, plus standard errors, confidence intervals, coefficients of variation, design effects, square roots of design effects, cumulative values, and unweighted counts for each estimate. Additionally, chi-square and likelihood ratio statistics are computed for the test of equal cell proportions.

## *Complex Samples Frequencies Data Considerations*

Data. Variables for which frequency tables are produced should be categorical. Subpopulation variables can be string or numeric, but should be categorical.

Assumptions. The cases in the data file represent a sample from a complex design that should be analyzed according to the specifications in the file selected in the Plan dialog box.

Obtaining Complex Samples Frequencies

▶ From the menus choose:

Analyze
  Complex Samples
    Frequencies...

▶ Select a plan file and optionally select a custom joint probabilities file.
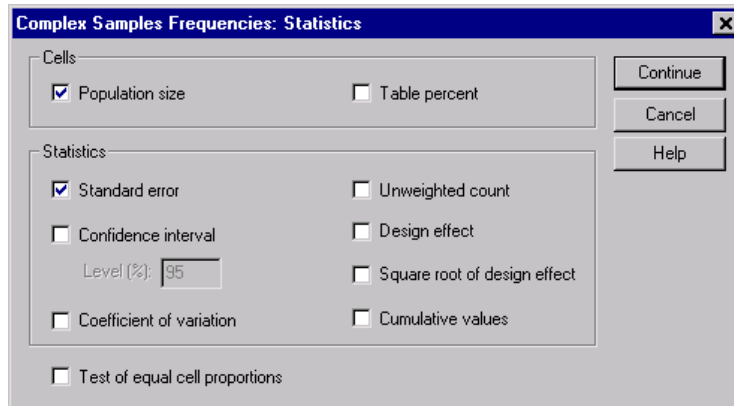
▶ Click Continue.

Figure 5-1
Frequencies dialog box
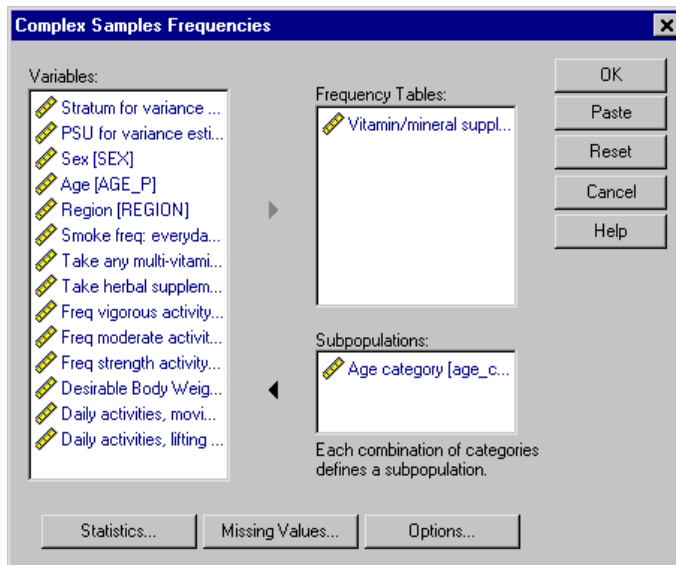
▶ Select at least one frequency variable.

Optionally, you can:

■ Specify variables to define subpopulations. Statistics are computed separately for
  each subpopulation.

Complex Samples Frequencies Statistics

Figure 5-2
Complex Samples Frequencies Statistics dialog box

Cells. This group allows you to request estimates of the cell population sizes and table percentages.

Statistics. This group produces statistics associated with the population size or table percentage.

- **Standard error.** The standard error of the estimate.
- **Confidence interval.** A confidence interval for the estimate, using the specified level.
- **Coeffcient of variation.** The ratio of the standard error of the estimate to the estimate.
- **Unweighted count.** The number of units used to compute the estimate.
- **Design effect.** The ratio of the variance of the estimate to the variance obtained by assuming that the sample is a simple random sample. This is a measure of the effect of specifying a complex design, where values further from 1 indicate greater effects.
- **Square root of design effect.** This is a measure of the effect of specifying a complex design, where smaller values indicate greater effects.
- **Cumulative values.** The cumulative estimate through each value of the variable.

Test of equal cell proportions. This produces chi-square and likelihood ratio tests of the hypothesis that the categories of a variable have equal frequencies. Separate tests are performed for each variable.

Complex Samples Missing Values

Figure 5-3
Missing Values dialog box
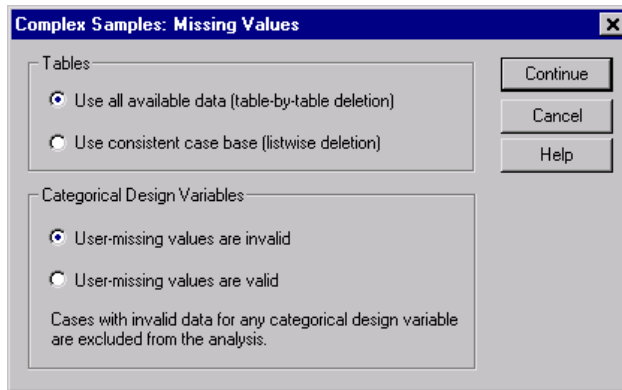
Tables. This group determines which cases are used in the analysis.

- **Use all available data.** Missing values are determined on a table-by-table basis, thus the cases used to compute statistics may vary across frequency or crosstabulation tables.
- **Ensure consistent case base.** Missing values are determined across all variables, thus the cases used to compute statistics are consistent across tables.

Categorical Design Variables. This group determines whether user-missing values are valid or invalid.

Complex Samples Options

Figure 5-4
Options dialog box
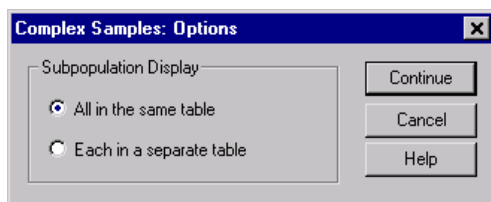
Display subpopulations. You can choose to have subpopulations displayed in the same table or in separate tables.

Complex Samples Descriptives

The Complex Samples Descriptives procedure displays univariate summary statistics for several variables. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

Example. Using the Complex Samples Descriptives procedure, you can obtain univariate descriptive statistics for the activity levels of U.S. citizens, based on the results of the National Health Interview Survey (NHIS) and with an appropriate analysis plan for this public use data.

Statistics. The procedure produces means and sums, plus *t*-tests, standard errors, confidence intervals, coefficients of variation, unweighted counts, population sizes, design effects, and square roots of design effects for each estimate.

Complex Samples Descriptives Data Considerations

Data. Measures should be scale variables. Subpopulation variables can be string or numeric, but should be categorical.

Assumptions. The cases in the data file represent a sample from a complex design that should be analyzed according to the specifications in the file selected in the Plan dialog box.

Obtaining Complex Samples Descriptives

▶ From the menus choose:

Analyze
 Complex Samples
  Descriptives...

▶ Select a plan file and optionally select a custom joint probabilities file.

▶ Click Continue.

Figure 6-1
Descriptives dialog box
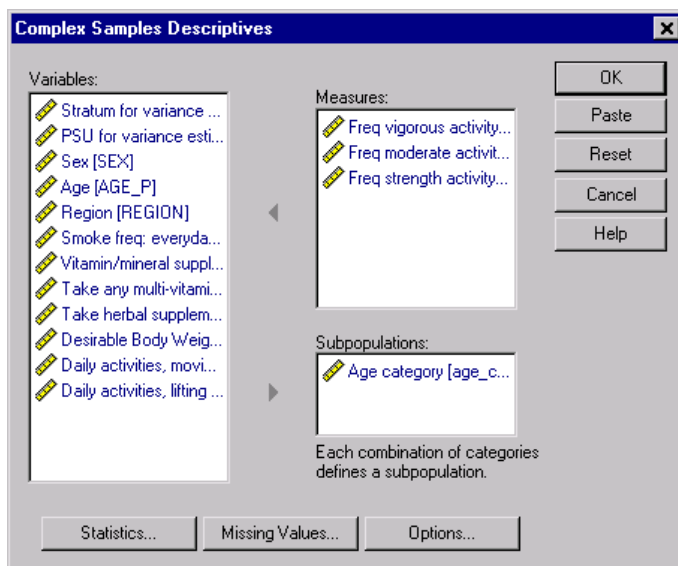
▶ Select at least one measure variable.

Optionally, you can:

■ Specify variables to define subpopulations. Statistics are computed separately for each subpopulation.

Complex Samples Descriptives Statistics

Figure 6-2
Descriptives Statistics dialog box

Summaries. This group allows you to request estimates of the means and sums of the measure variables. Additionally, you can request *t*-tests of the estimates against a specified value.

Statistics. This group produces statistics associated with the mean or sum.

- **Standard error.** The standard error of the estimate.

- **Confidence interval.** A confidence interval for the estimate, using the specified level.

- **Coeffcient of variation.** The ratio of the standard error of the estimate to the estimate.

- **Unweighted count.** The number of units used to compute the estimate.

- **Population size.** The estimated number of units in the population.

- **Design effect.** The ratio of the variance of the estimate to the variance obtained by assuming the sample is a simple random sample. This is a measure of the effect of specifying a complex design, where values further from 1 indicate greater effects.

- **Square root of design effect.** This is a measure of the effect of specifying a complex design, where smaller values indicate greater effects.

Complex Samples Descriptives Missing Values

Figure 6-3
Descriptives Missing Values dialog box
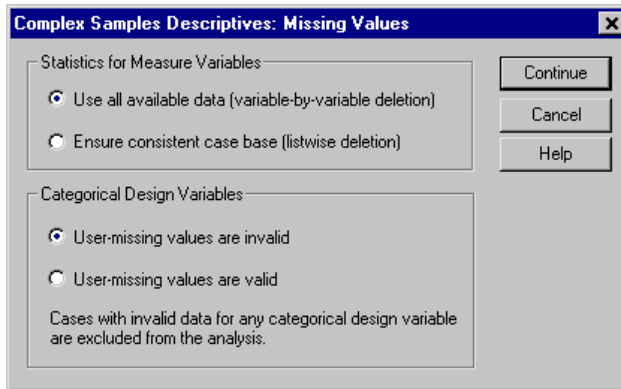
Statistics for Measure Variables. This group determines which cases are used in the analysis.

- **Use all available data.** Missing values are determined on a variable-by-variable basis, thus the cases used to compute statistics may vary across measure variables.

- **Ensure consistent case base.** Missing values are determined across all variables, thus the cases used to compute statistics are consistent.

Categorical Design Variables. This group determines whether user-missing values are valid or invalid.

Complex Samples Options

Figure 6-4
Options dialog box
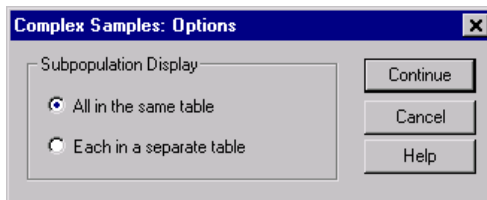
Display subpopulations. You can choose to have subpopulations displayed in the same table or in separate tables.

Complex Samples Crosstabs

The Complex Samples Crosstabs procedure produces crosstabulation tables for pairs of selected variables and displays two-way statistics. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

Example. Using the Complex Samples Crosstabs procedure, you can obtain crossclassification statistics for smoking frequency by vitamin usage of U.S. citizens, based on the results of the National Health Interview Survey (NHIS) and with an appropriate analysis plan for this public use data.

Statistics. The procedure produces estimates of cell population sizes and row, column, and table percentages, plus standard errors, confidence intervals, coefficients of variation, expected values, design effects, square roots of design effects, residuals, adjusted residuals, and unweighted counts for each estimate. The odds ratio, relative risk, and risk difference are computed for 2-by-2 tables. Additionally, Pearson and likelihood ratio statistics are computed for the test of independence of the row and column variables.

## Complex Samples Crosstabs Data Considerations

Data. Row and column variables should be categorical. Subpopulation variables can be string or numeric, but should be categorical.

Assumptions. The cases in the data file represent a sample from a complex design that should be analyzed according to the specifications in the file selected in the Plan dialog box.

Obtaining Complex Samples Crosstabs

▶ From the menus choose:

Analyze
 Complex Samples
  Crosstabs...

▶ Select a plan file and optionally select a custom joint probabilities file.
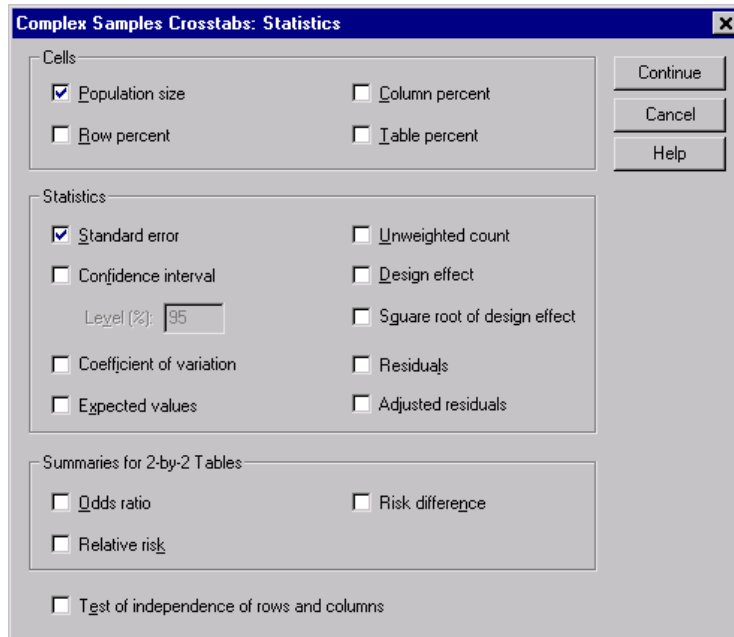
▶ Click Continue.

Figure 7-1
Crosstabs dialog box
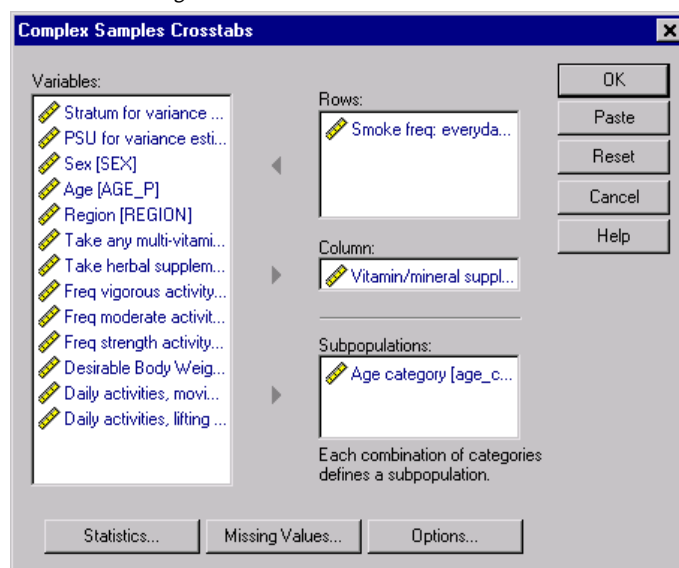
▶ Select at least one row variable and one column variable.

Optionally, you can:

■ Specify variables to define subpopulations. Statistics are computed separately for
  each subpopulation.

Complex Samples Crosstabs Statistics

Figure 7-2
Crosstabs Statistics dialog box

Cells. This group allows you to request estimates of the cell population sizes and row, column, and table percentages.

Statistics. This group produces statistics associated with the population size and row, column, and table percentages.

- **Standard error.** The standard error of the estimate.
- **Confidence interval.** A confidence interval for the estimate, using the specified level.
- **Coefficient of variation.** The ratio of the standard error of the estimate to the estimate.
- **Expected values.** The expected value of the estimate, under the hypothesis of independence of the row and column variable.
- **Unweighted count.** The number of units used to compute the estimate.

- ■ **Design effect.** The ratio of the variance of the estimate to the variance obtained by assuming the sample is a simple random sample. This is a measure of the effect of specifying a complex design, where values further from 1 indicate greater effects.

- ■ **Square root of design effect.** This is a measure of the effect of specifying a complex design, where smaller values indicate greater effects.

- ■ **Residuals.** The expected value is the number of cases you would expect in the cell if there were no relationship between the two variables. A positive residual indicates that there are more cases in the cell than there would be if the row and column variables were independent.

- ■ **Adjusted residuals.** The residual for a cell (observed minus expected value) divided by an estimate of its standard error. The resulting standardized residual is expressed in standard deviation units above or below the mean.

Summaries for 2-by-2 Tables. This group produces statistics for tables in which the row and column variable each have two categories. Each is a measure of the strength of the association between the presence of a factor and the occurrence of an event.

- ■ **Odds ratio.** The odds ratio can be used as an estimate of relative risk when the occurrence of the factor is rare.

- ■ **Relative risk.** The ratio of the risk of an event in the presence of the factor to the risk of the event in the absence of the factor.

- ■ **Risk difference.** The difference between the risk of an event in the presence of the factor and the risk of the event in the absence of the factor.

Test of independence of rows and columns. This produces chi-square and likelihood ratio tests of the hypothesis that a row and column variable are independent. Separate tests are performed for each pair of variables.

Complex Samples Missing Values

Figure 7-3
Missing Values dialog box
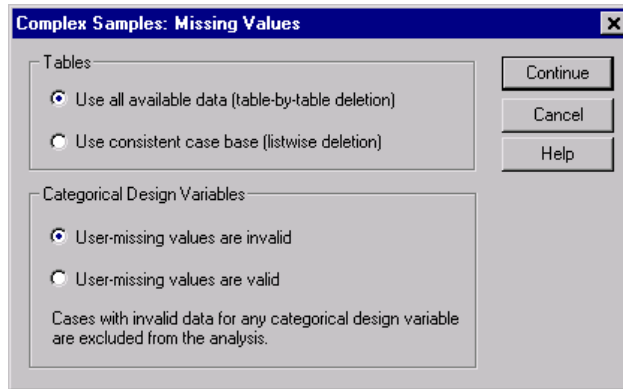
Tables. This group determines which cases are used in the analysis.

- **Use all available data.** Missing values are determined on a table-by-table basis, thus the cases used to compute statistics may vary across frequency or crosstabulation tables.

- **Ensure consistent case base.** Missing values are determined across all variables, thus the cases used to compute statistics are consistent across tables.

Categorical Design Variables. This group determines whether user-missing values are valid or invalid.

Complex Samples Options

Figure 7-4
Options dialog box
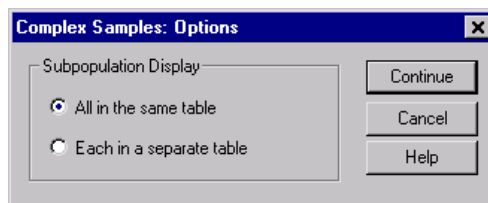
Display subpopulations. You can choose to have subpopulations displayed in the same table or in separate tables.

Complex Samples Ratios

The Complex Samples Ratios procedure displays univariate summary statistics for ratios of variables. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

Example. Using the Complex Samples Ratios procedure, you can obtain descriptive statistics for the ratio of current property value to last assessed value, based on the results of a statewide survey carried out according to a complex design and with an appropriate analysis plan for the data.

Statistics. The procedure produces ratio estimates, *t*-tests, standard errors, confidence intervals, coefficients of variation, unweighted counts, population sizes, design effects, and square roots of design effects.

Complex Samples Ratios Data Considerations

Data. Numerators and denominators should be positive-valued scale variables. Subpopulation variables can be string or numeric but should be categorical.

Assumptions. The cases in the data file represent a sample from a complex design that should be analyzed according to the specifications in the file selected in the Plan dialog box.

Obtaining Complex Samples Ratios

▶ From the menus choose:

Analyze
  Complex Samples
    Ratios...

▶ Select a plan file and, optionally, select a custom joint probabilities file.

▶ Click Continue.

Figure 8-1
Complex Samples Ratios dialog box
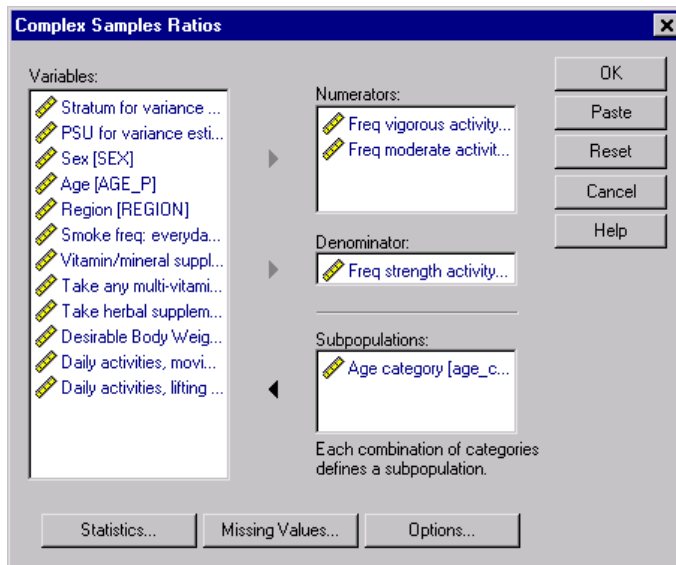
▶ Select at least one numerator variable and denominator variable.

Optionally, you can:

■ Specify variables to define subgroups for which statistics are produced.

Complex Samples Ratios Statistics

Figure 8-2
Statistics dialog box
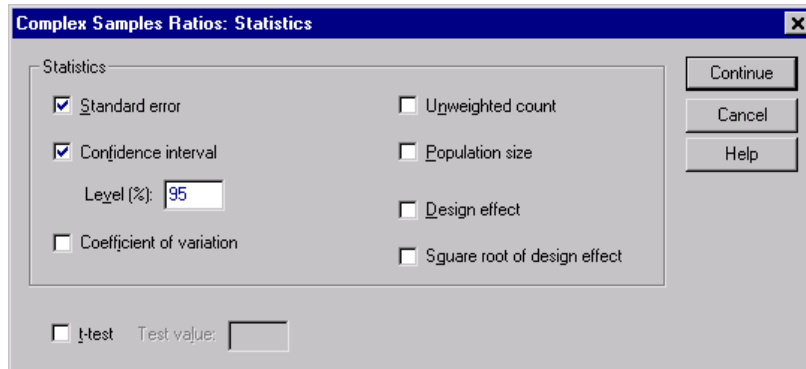
Statistics. This group produces statistics associated with the ratio estimate.

- **Standard error.** The standard error of the estimate.

- **Confidence interval.** A confidence interval for the estimate, using the specified level.

- **Coefficient of variation.** The ratio of the standard error of the estimate to the estimate.

- **Unweighted count.** The number of units used to compute the estimate.

- **Population size.** The estimated number of units in the population.

- **Design effect.** The ratio of the variance of the estimate to the variance obtained by assuming the sample is a simple random sample. This is a measure of the effect of specifying a complex design, where values further from 1 indicate greater effects.

- **Square root of design effect.** This is a measure of the effect of specifying a complex design, where smaller values indicate greater effects.

t-test. You can request *t*-tests of the estimates against a specified value.

Complex Samples Ratios Missing Values

Figure 8-3
Missing Values dialog box
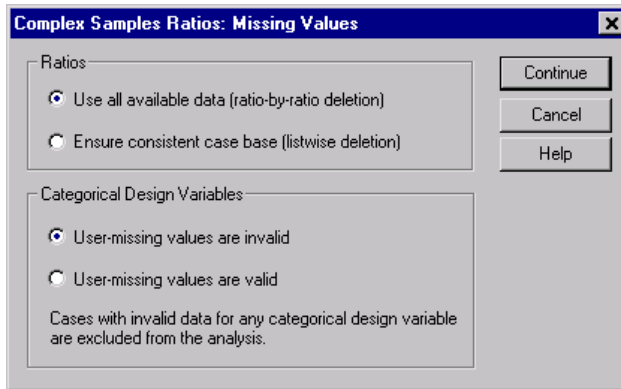
Ratios. This group determines which cases are used in the analysis.

- **Use all available data.** Missing values are determined on a ratio-by-ratio basis; thus, the cases used to compute statistics may vary across numerator-denominator pairs.

- **Ensure consistent case base.** Missing values are determined across all variables; thus, the cases used to compute statistics are consistent.

Categorical Design Variables. This group determines whether user-missing values are valid or invalid.

Complex Samples Options

Figure 8-4
Options dialog box
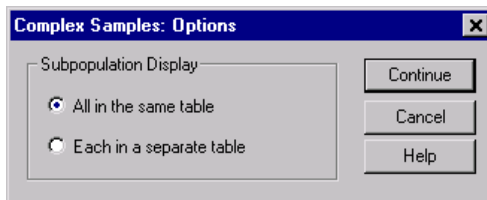
Display subpopulations. You can choose to have subpopulations displayed in the same table or in separate tables.

Complex Samples Sampling Wizard

The Sampling Wizard guides you through the steps for creating, modifying, or
executing a sampling plan file. Before using the Wizard, you should have a
well-defined target population, a list of sampling units, and an appropriate sample
design in mind.

Obtaining a Sample from a Full Sampling Frame

A state agency is charged with ensuring fair property taxes from county to county.
Taxes are based on the appraised value of the property, so the agency wants to survey
a sample of properties by county to be sure that each county's records are equally
up-to-date. However, resources for obtaining current appraisals are limited, so it's
important that what is available is used wisely. The agency decides to employ
complex sampling methodology to select a sample of properties.

A listing of properties is collected in *property_assess_cs.sav*. Use the Complex
Samples Sampling Wizard to select a sample.

Using the Wizard

▶ To run a Complex Samples Sampling Wizard analysis, from the menus choose:
Analyze
  Complex Samples
    Select a Sample...

62

Figure 9-1
Welcome step
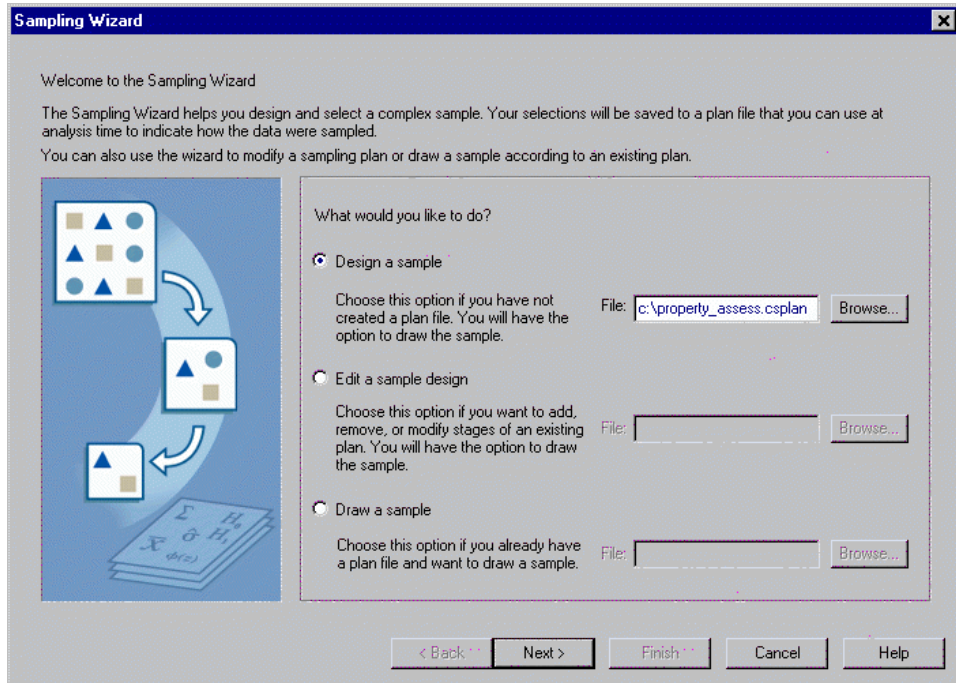
▶ Select Design a sample and type c:\property_assess.csplan as the name of the plan file.

▶ Click Next.

Figure 9-2
Design Variables step
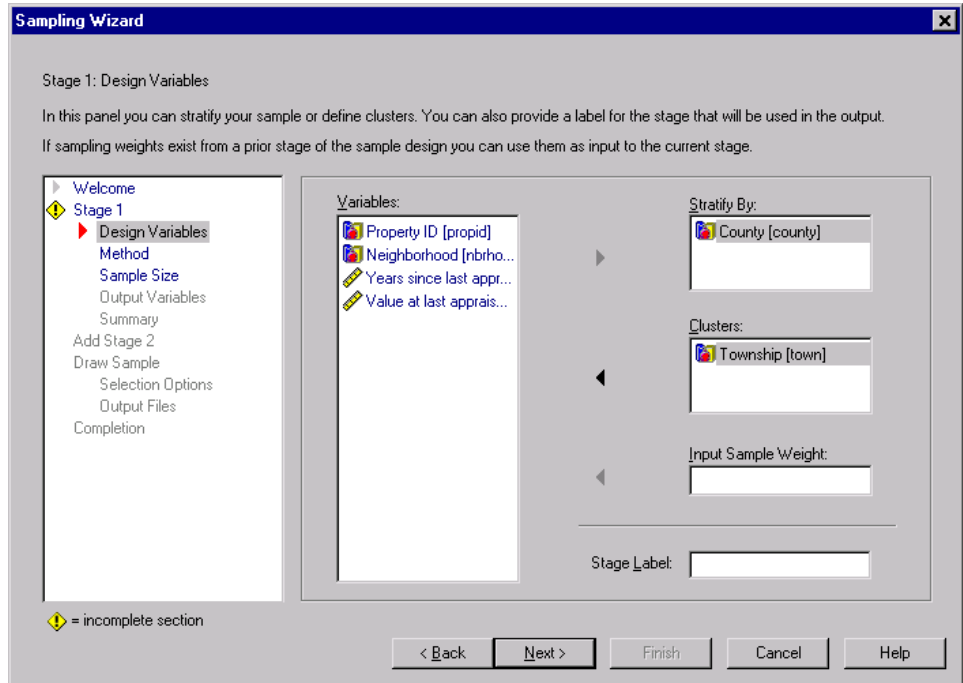
▶ Select *County* as a stratification variable.

▶ Select *Township* as a cluster variable.
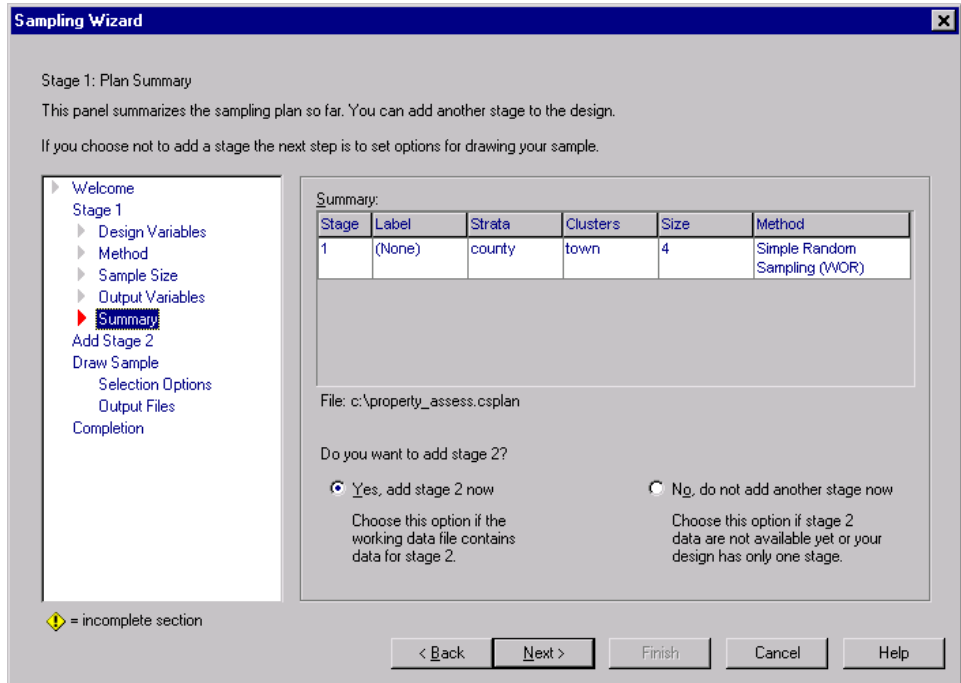
▶ Click Next, then click Next in the Method step.

This design structure means that independent samples are drawn for each county. In this stage, townships are drawn as the primary sampling unit using the default method, Simple Random Sampling.

Figure 9-3
Sample Size step
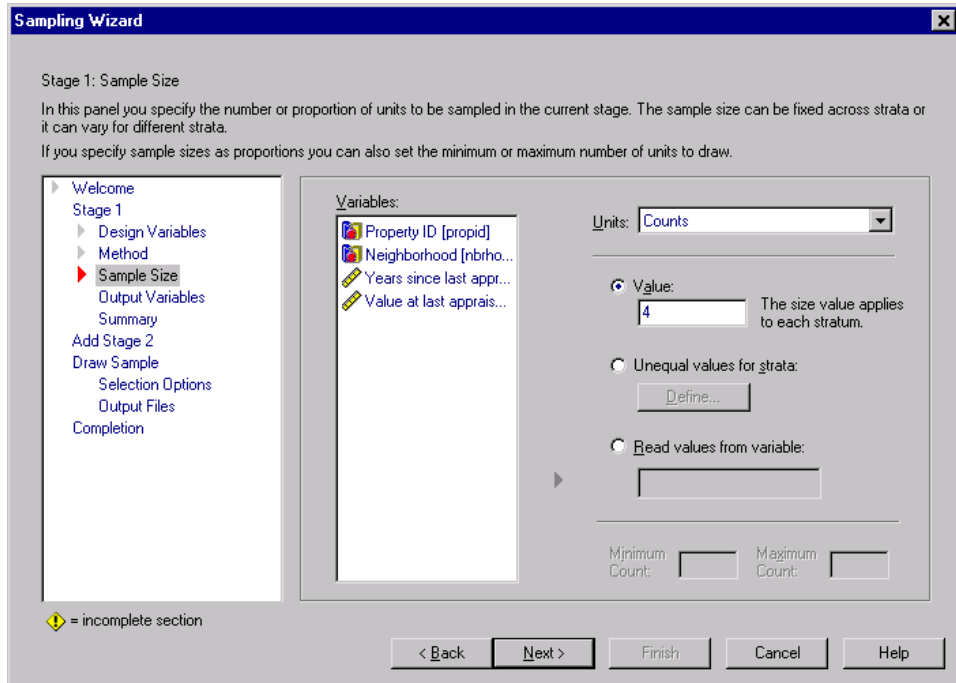
▶ Type 4 as the value for the number of clusters to select in this stage.

▶ Click Next, then click Next in the Output Variables step.

Figure 9-4
Stage Summary step

▶ Select Yes, add Stage 2 now.

▶ Click Next.

Figure 9-5
Design Variables step, stage 2
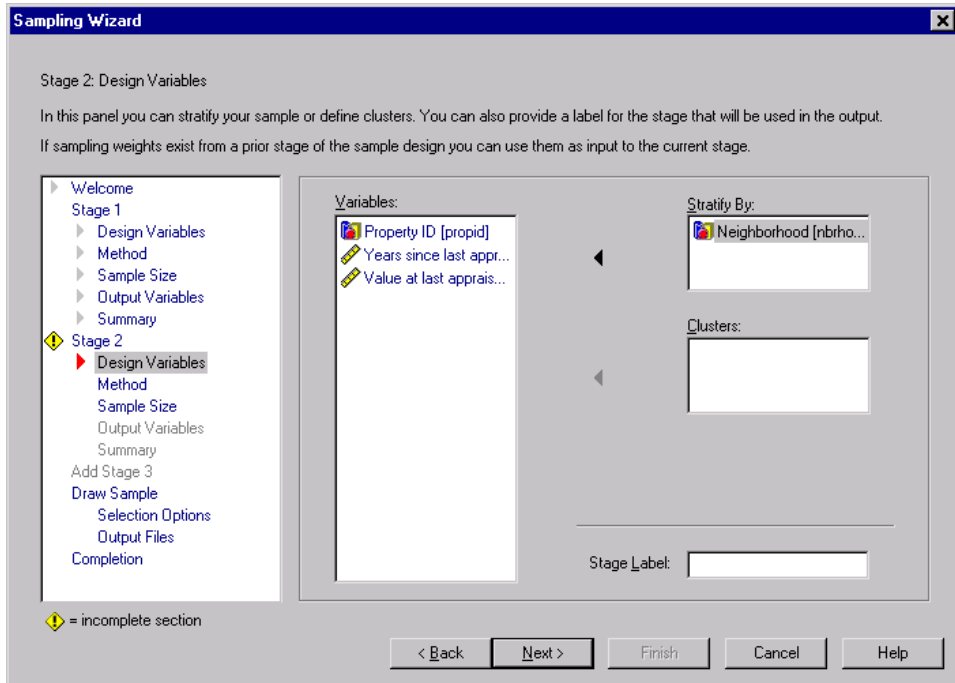
▶ Select *Neighborhood* as a stratification variable.

▶ Click Next, then click Next in the Method step.

This design structure means that independent samples are drawn for each neighborhood of the townships drawn in stage 1. In this stage, properties are drawn as the primary sampling unit using Simple Random Sampling.

Figure 9-6
Sample Size step
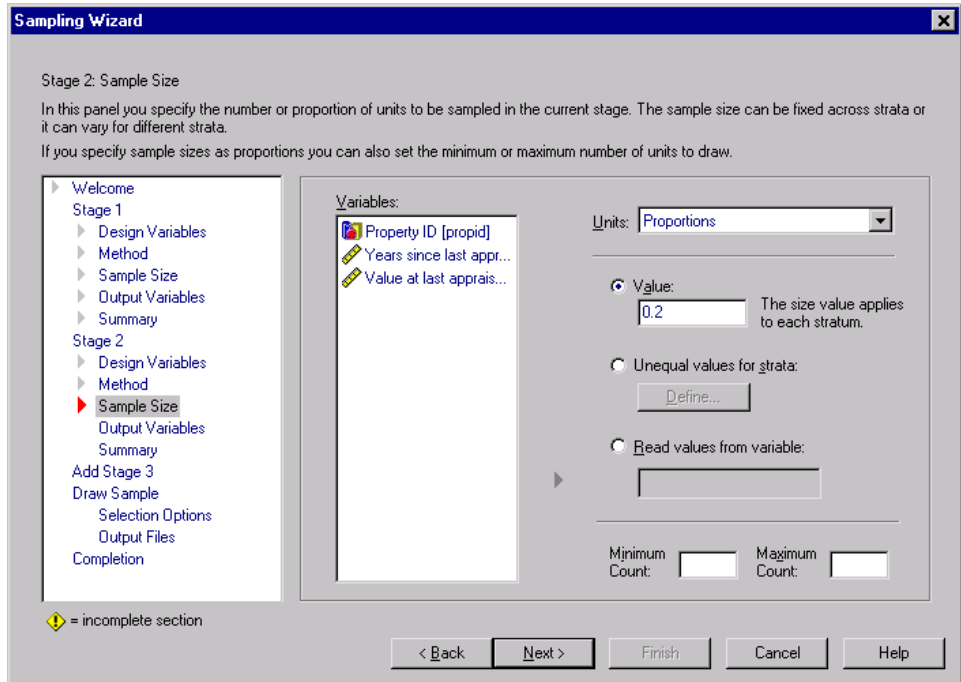
▶ Select Proportions from the Units drop-down list.

▶ Type 0.2 as the value of the proportion of units to sample from each strata.

▶ Click Next, then click Next in the Output Variables step.

Figure 9-7
Stage Summary step
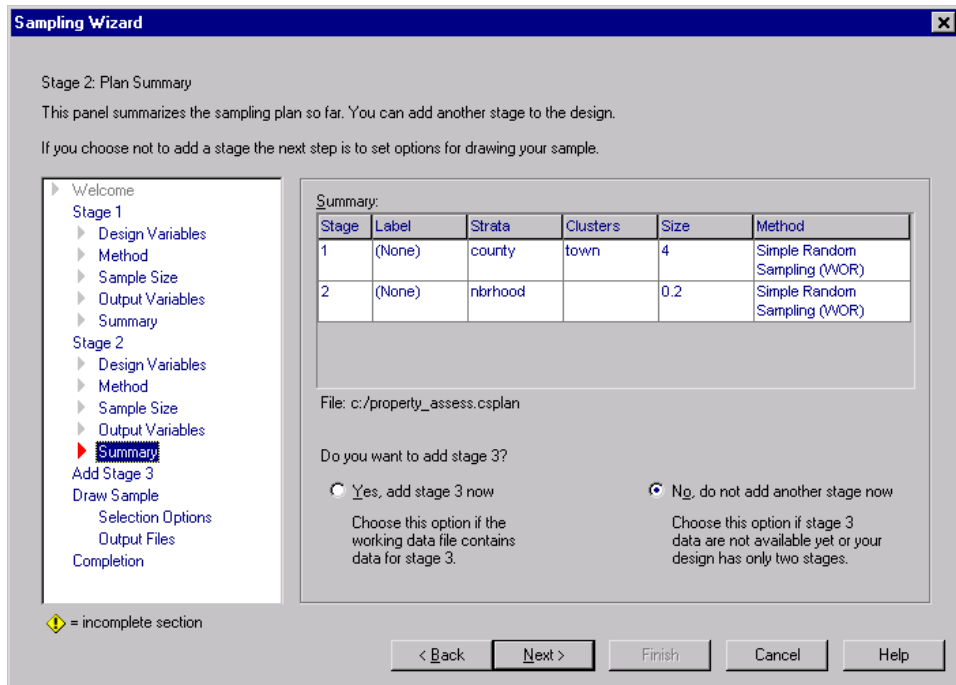
▶ Look over the sampling design, then click Next.

Figure 9-8
Draw Sample:Selection Options step
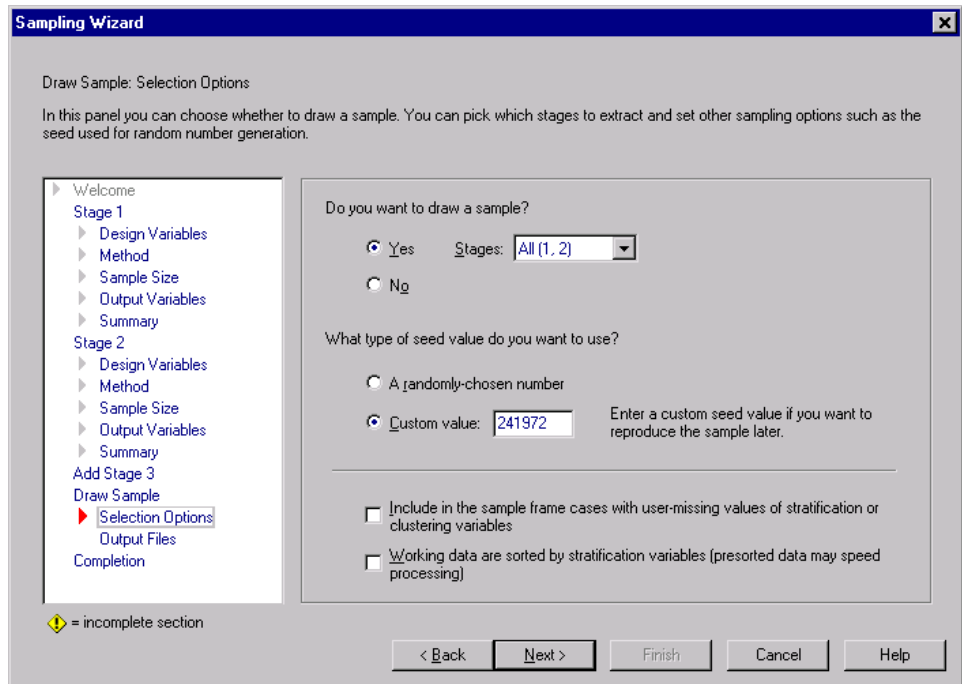
▶ Select Custom value for the type of random seed to use and type 241972 as the value.

  Using a custom value allows you to replicate the results of this example exactly.

▶ Click Next, then click Next in the Draw Sample:Output Files step.

Figure 9-9
Finish step
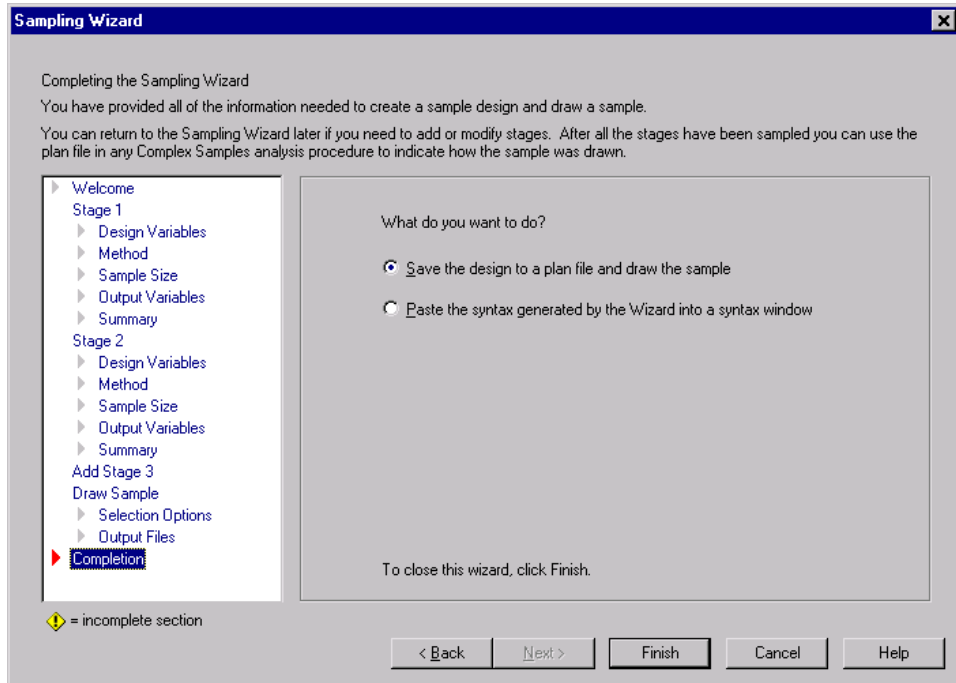
▶ Click Finish.

These selections produce the sampling plan file *property_assess.csplan* and draw a
sample according to that plan.

Plan Summary

Figure 9-10
Plan Summary

			Stage 1	Stage 2
Design Variables	Stratification	1	County	Neighborhood
	Cluster	1	Township	
Sample Information	Selection Method		Simple random sampling without replacement	Simple random sampling without replacement
	Number of Units Sampled		4	
	Variables Created or Modified	Stagewise Inclusion (Selection) Probability	Inclusion Probability_1_	Inclusion Probability_2_
		Stagewise Cumulative Sample Weight	Sample Weight Cumulative_1_	Sample Weight Cumulative_2_
		Stagewise Duplication Index	Index_1_	Index_2_
	Proportion of Units Sampled			.2
Analysis Information	Estimator Assumption		Equal probability sampling without replacement	Equal probability sampling without replacement
	Inclusion Probability		Obtained from variable Inclusion Probability_1_	Obtained from variable Inclusion Probability_2_

Plan File: c:\property_assess.csplan
Weight Variable: SampleWeight_Final_

The summary table reviews your sampling plan, and it is useful for making sure the plan represents your intentions.

Sampling Summary

Figure 9-11
Stage Summary

County	Number of Units Sampled		Proportion of Units Sampled	
	Requested	Actual	Requested	Actual
Eastern	4	4	44.4%	44.4%
Central	4	4	57.1%	57.1%
Western	4	4	25.0%	25.0%
Northern	4	4	44.4%	44.4%
Southern	4	4	50.0%	50.0%

This summary table reviews the first stage of sampling, and it is useful for checking that the sampling went according to plan. Four townships were sampled from each county, as requested.

Figure 9-12

Stage Summary

County	Township	Neighborhood	Number of Units Sampled		Proportion of Units Sampled	
			Requested	Actual	Requested	Actual
Southern	88	613	14	14	20.0%	20.6%
		612	15	15	20.0%	19.5%
		611	11	11	20.0%	20.4%
		610	16	16	20.0%	20.3%
	82	571	15	15	20.0%	20.3%
		570	14	14	20.0%	19.7%
		569	15	15	20.0%	19.5%
		568	8	8	20.0%	20.5%
	84	585	6	6	20.0%	20.7%
		582	11	11	20.0%	19.3%
		584	11	11	20.0%	19.3%
		583	14	14	20.0%	19.7%
		587	11	11	20.0%	20.4%
		586	8	8	20.0%	21.1%
	81	564	6	6	20.0%	21.4%
		563	5	5	20.0%	20.8%
		562	12	12	20.0%	19.4%
		561	15	15	20.0%	19.5%
		565	13	13	20.0%	19.7%
Eastern	6	36	13	13	20.0%	20.3%
		38	13	13	20.0%	20.6%
		37	14	14	20.0%	20.6%
	7	44	11	11	20.0%	19.6%
		45	11	11	20.0%	20.0%

This summary table (the top part of which is shown here) reviews the second stage of sampling. It is also useful for checking that the sampling went according to plan. Approximately 20% of the properties were sampled from each neighborhood from each township sampled in the first stage, as requested.

Sample Results

Figure 9-13
Data Editor with sample results

You can see the sampling results in the Data Editor. Five new variables were saved to the working file, representing the inclusion probabilities and cumulative sampling weights for each stage, plus the final sampling weights.

■ Cases with values for these variables were selected to the sample.

■ Cases with system-missing values for the variables were not selected.

The agency will now use its resources to collect current valuations for the properties selected in the sample. Once those valuations are available, you can process the sample with Complex Samples analysis procedures, using the sampling plan *property_assess.csplan* to provide the sampling specifications.

Obtaining a Sample from a Partial Sampling Frame

A company is interested in compiling and selling a database of high-quality survey information. The survey sample should be representative, but efficiently carried out, so complex sampling methods are used. The full sampling design calls for the following structure:

Stage	Strata	Clusters
1	Region	Province
2	District	City
3	Subdivision	

In the third stage, households are the primary sampling unit, and selected households will be surveyed. However, information is only easily available to the city level, so the company plans to execute the first two stages of the design now, then collect information on the numbers of subdivisions and households from the sampled cities. The available information to the city level is collected in *demo_cs_1.sav*. Use the Complex Samples Sampling Wizard to draw the first two stages.

Using the Wizard to Sample from the First Partial Frame

▶ To run the Complex Samples Sampling Wizard, from the menus choose:

Analyze
 Complex Samples
  Select a Sample...

Figure 9-14
Welcome step
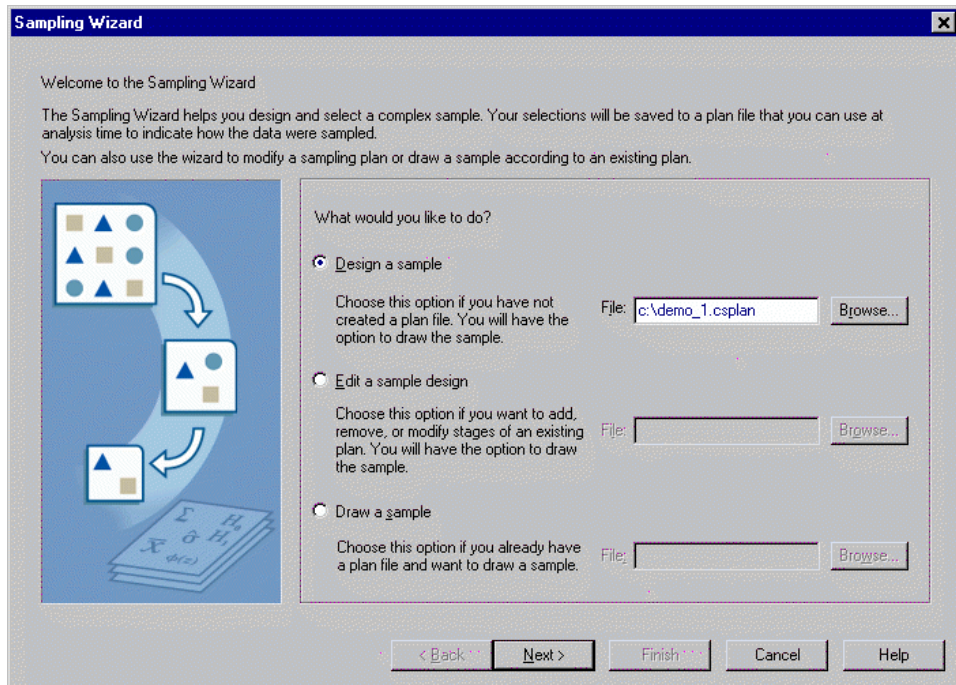
▶ Select Design a sample and type c:\demo_1.csplan as the name of the plan file.

▶ Click Next.

Figure 9-15
Design Variables step
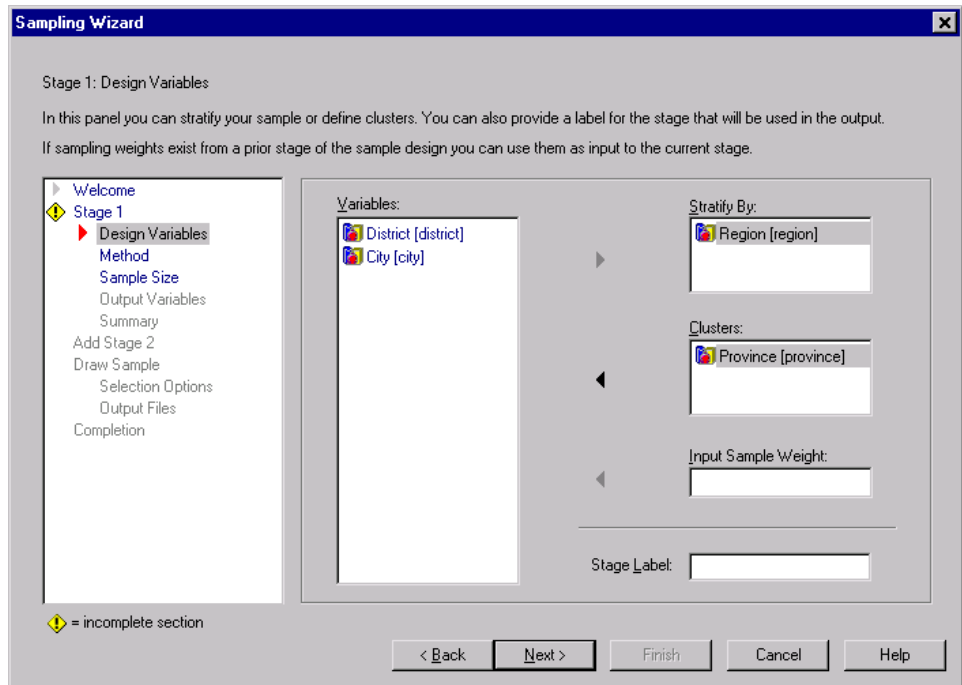
▶ Select *Region* as a stratification variable.

▶ Select *Province* as a cluster variable.
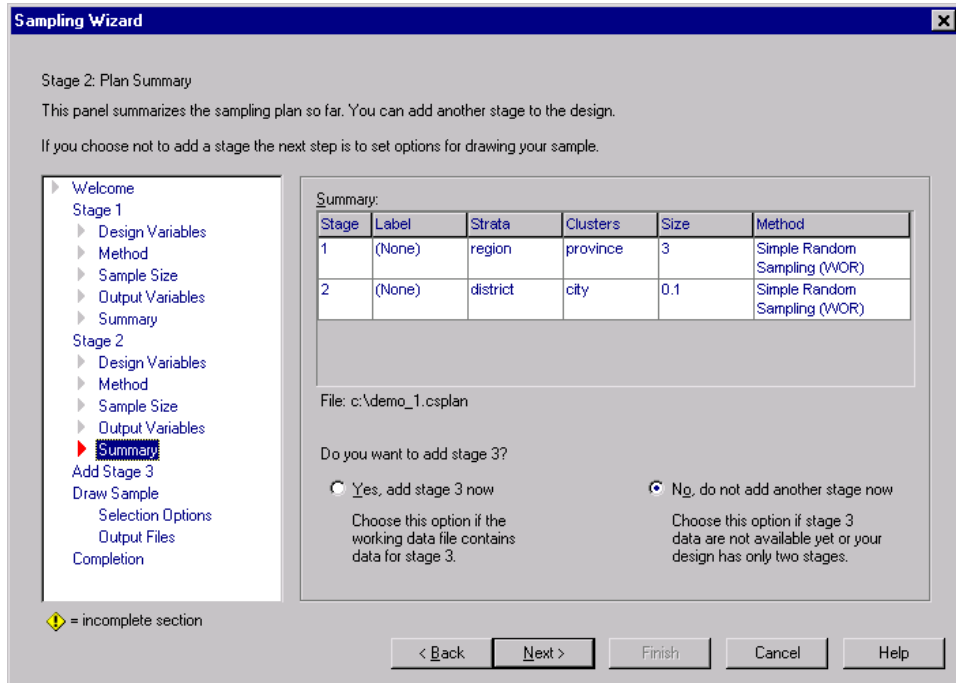
▶ Click Next, then click Next in the Method step.

This design structure means that independent samples are drawn for each region. In this stage, provinces are drawn as the primary sampling unit using the default method, Simple Random Sampling.

Chapter 9

Figure 9-16
Sample Size step
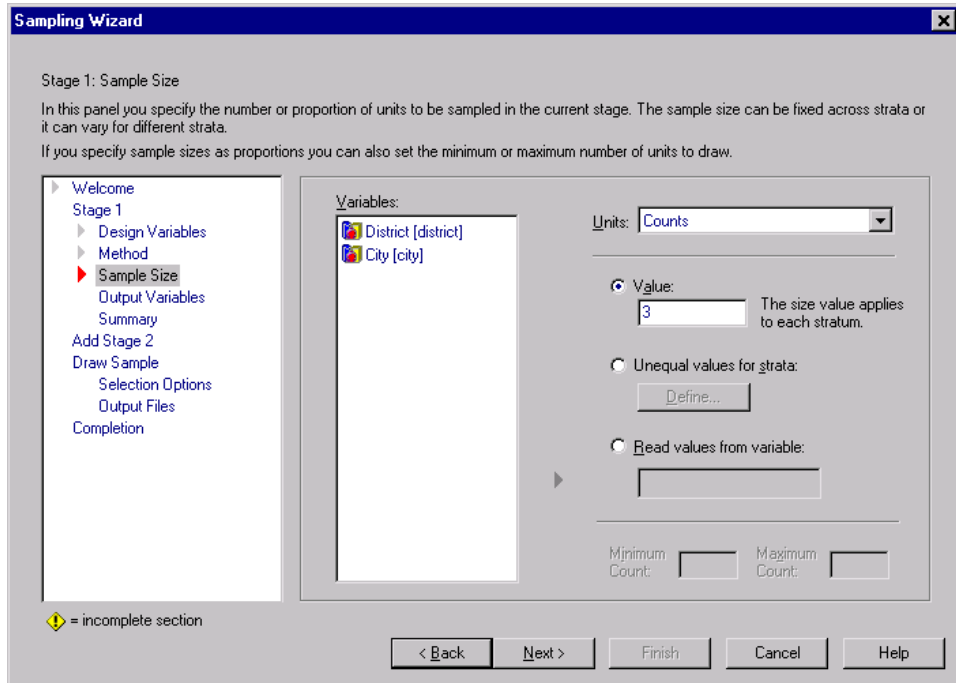
▶ Type 3 as the value for the number of clusters to select in this stage.

▶ Click Next, then click Next in the Output Variables step.

Figure 9-17
Stage Summary step
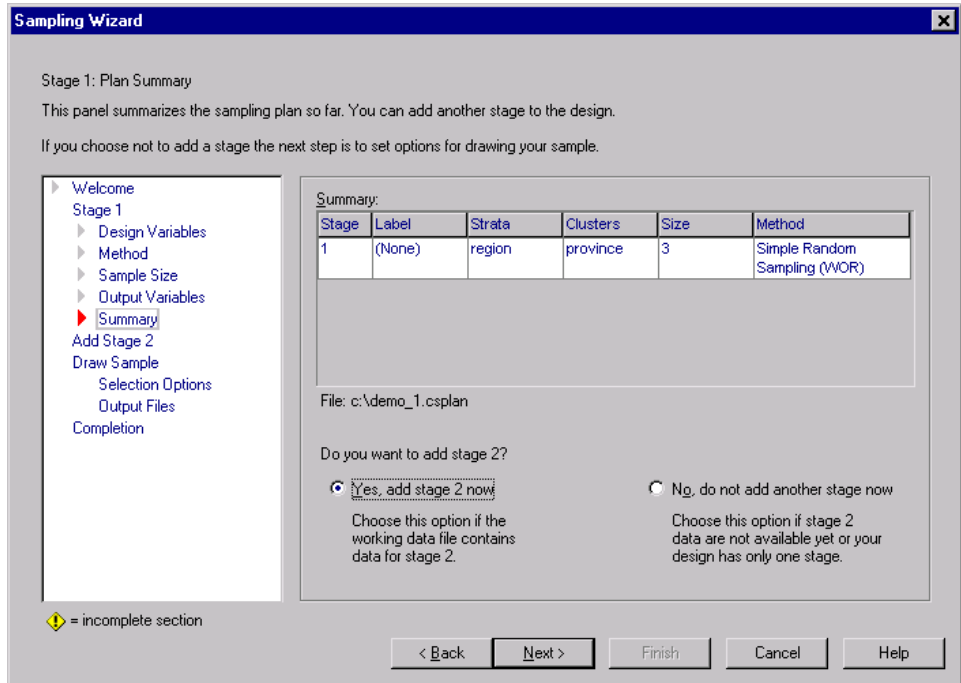
▶ Select Yes, add Stage 2 now.

▶ Click Next.

Figure 9-18
Design Variables step, stage 2
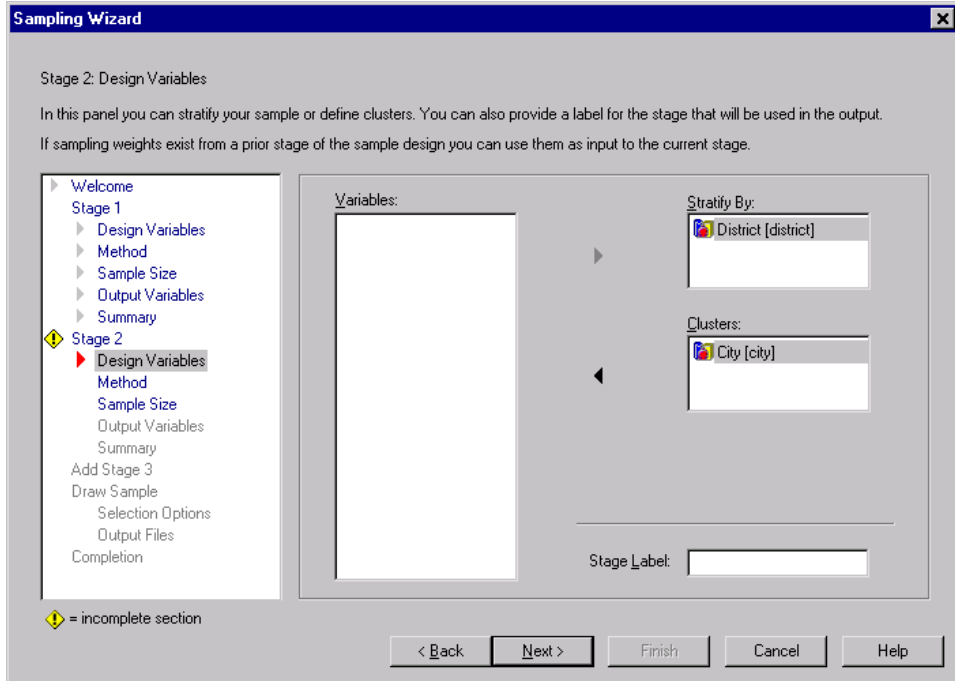
▶ Select *District* as a stratification variable.

▶ Select *City* as a cluster variable.

▶ Click Next, then click Next in the Method step.

This design structure means that independent samples are drawn for each district. In this stage, cities are drawn as the primary sampling unit using the default method, Simple Random Sampling.

Figure 9-19
Sample Size step
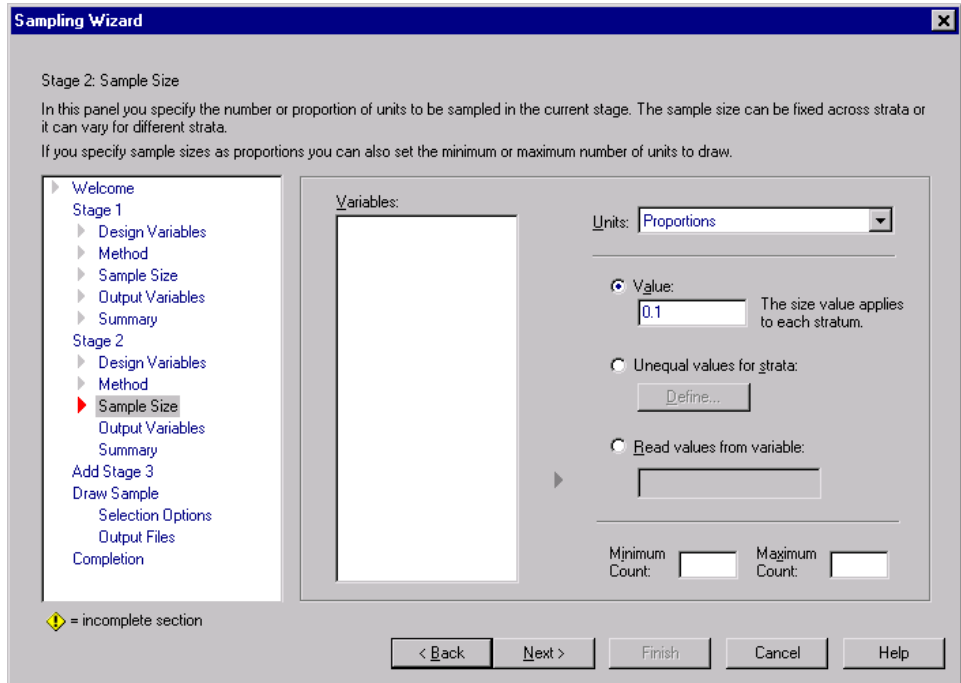
▶ Select Proportions from the Units drop-down list.

▶ Type 0.1 as the value of the proportion of units to sample from each strata.

▶ Click Next, then click Next in the Output Variables step.

Figure 9-20
Stage Summary step

▶ Look over the sampling design, then click Next.

Figure 9-21
Draw Sample:Selection Options step

▶ Select Custom value for the type of random seed to use and type 241972 as the value.

Using a custom value allows you to replicate the results of this example exactly.

▶ Click Next, then click Next in the Draw Sample:Output Files step.

Figure 9-22
Finish step
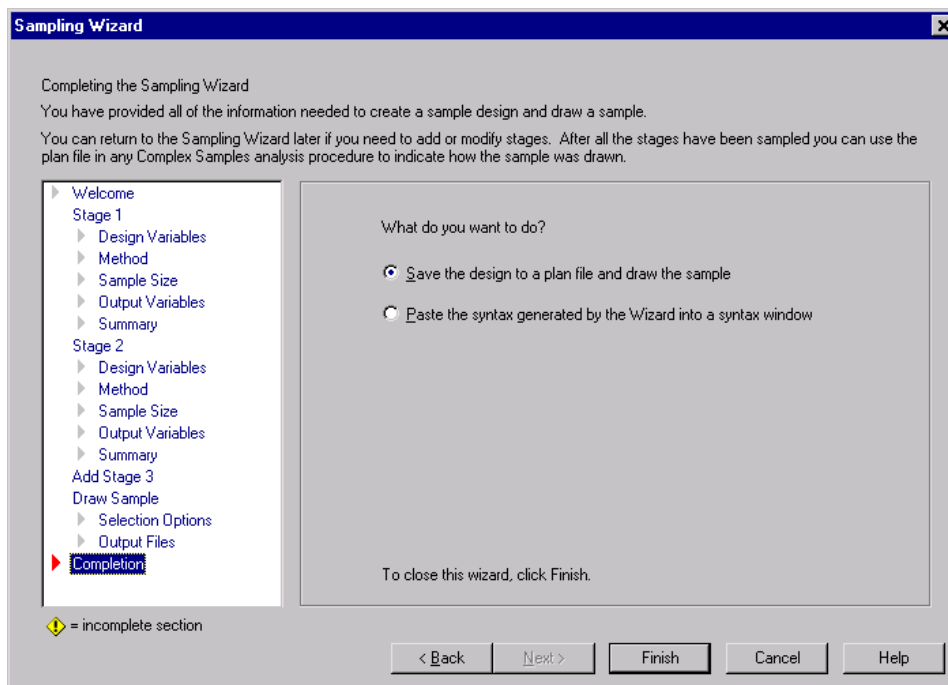
▶ Click Finish.

These selections produce the sampling plan file *demo_1.csplan* and draw a sample according to that plan.

Sample Results

Figure 9-23
Data Editor with sample results

You can see the sampling results in the Data Editor. Five new variables were saved to the working file, representing the inclusion probabilities and cumulative sampling weights for each stage, plus the "final" sampling weights for the first two stages.

■ Cities with values for these variables were selected to the sample.

■ Cities with system-missing values for the variables were not selected.

For each city selected, the company acquired subdivision and household unit information and placed it in *demo_cs_2.sav*. Use this file and the Sampling Wizard to sample the third stage of this design.

Using the Wizard to Sample from the Second Partial Frame
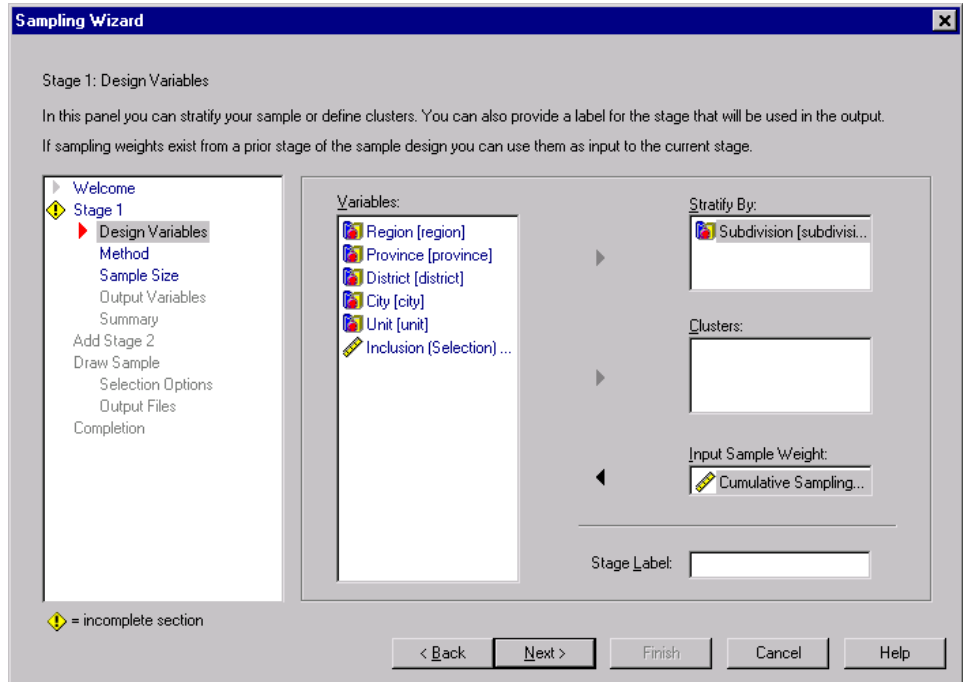
▶ To run the Complex Samples Sampling Wizard, from the menus choose:

Analyze
  Complex Samples
    Select a Sample...

Figure 9-24
Welcome step
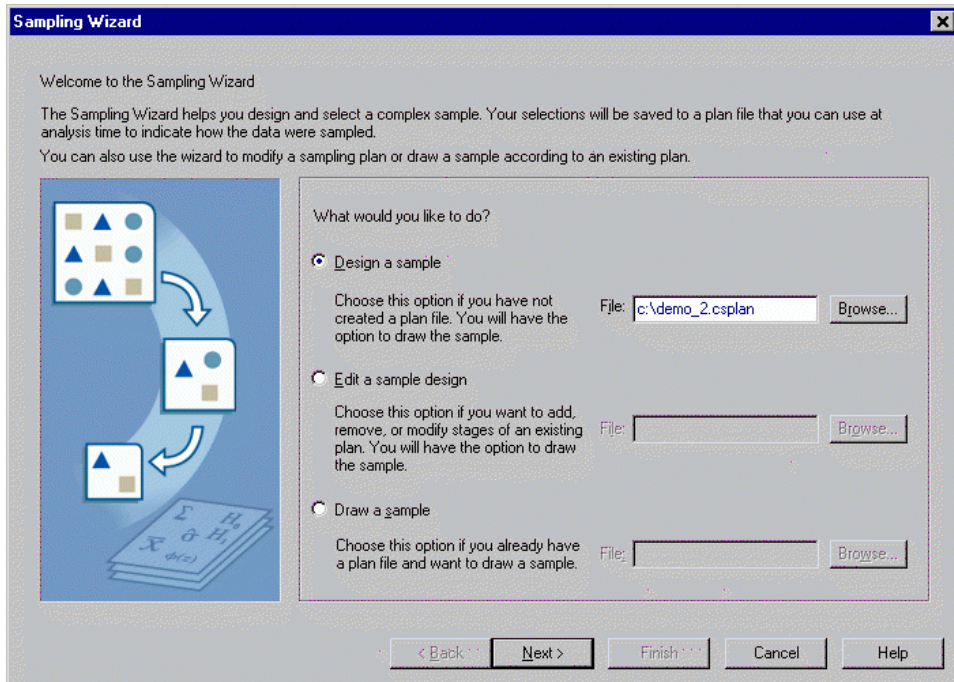
▶ Select Design a sample and type c:\demo_2.csplan as the name of the plan file.

▶ Click Next.

Figure 9-25
Design Variables step

▶ Select *Subdivision* as a stratification variable.

▶ Select *Cumulative Sampling Weight for Stage 2* as the input sample weight variable.

▶ Click Next, then click Next in the Method step.

This design structure means that independent samples are drawn for each subdivision. In this stage, household units are drawn as the primary sampling unit using the default method, Simple Random Sampling.

Figure 9-26
Sample Size step
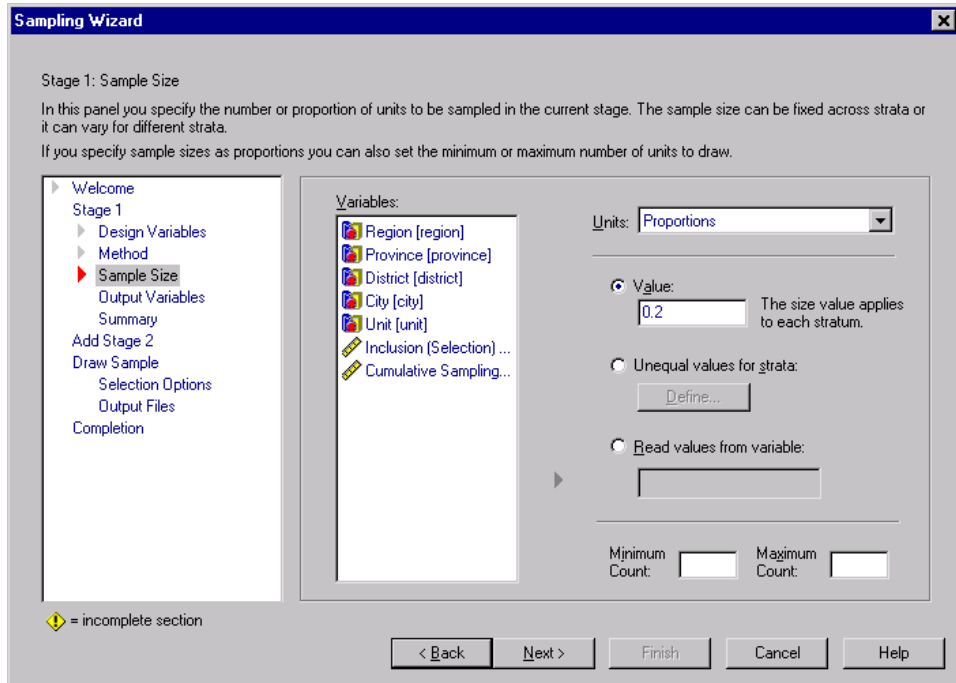
▶ Type 0.2 as the value for the proportion of units to select in this stage.

▶ Click Next, then click Next in the Output Variables step.

Figure 9-27
Stage Summary step
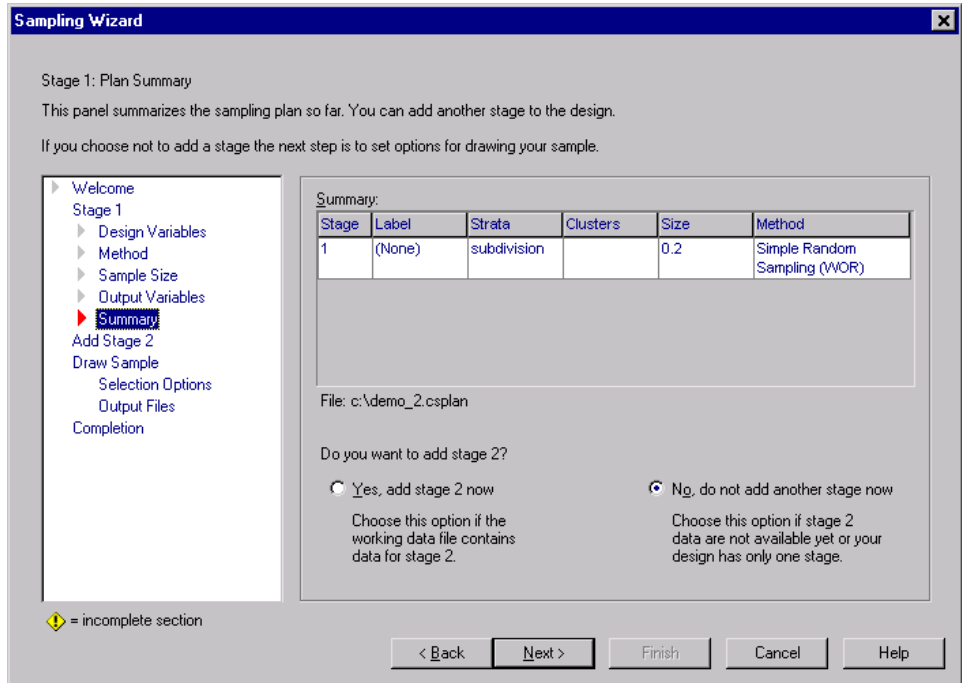
▶ Look over the sampling design, then click Next.

Figure 9-28
Draw Sample:Selection Options step
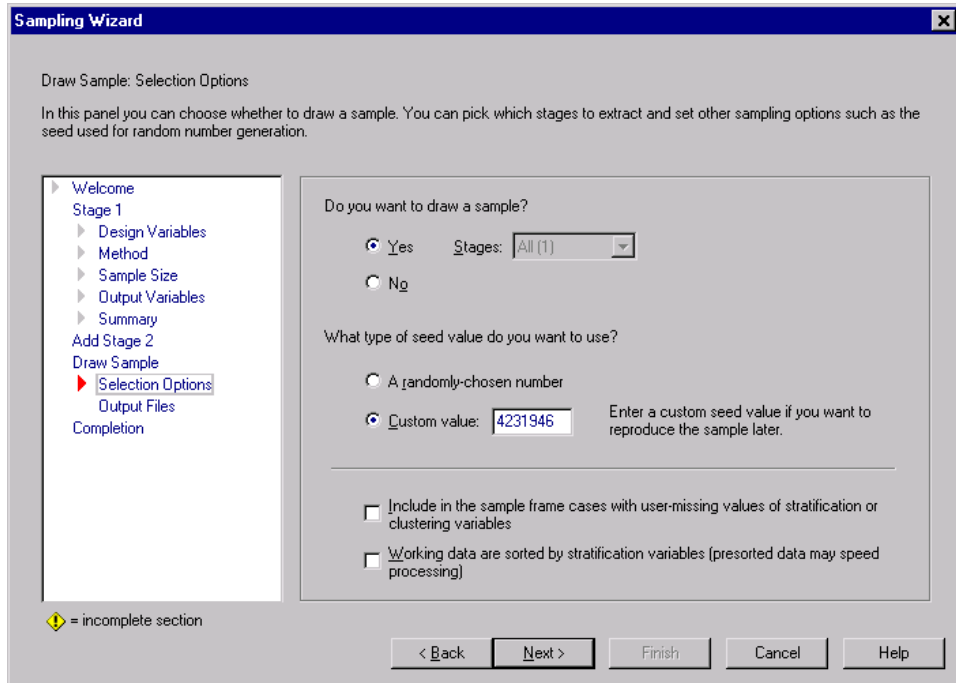
▶ Select Custom value for the type of random seed to use and type 4231946 as the value.

▶ Click Next, then click Next in the Draw Sample:Output Files step.

Figure 9-29
Finish step

▶ Click Finish.

These selections produce the sampling plan file *demo_2.csplan* and draw a sample according to that plan.

Sample Results

Figure 9-30
Data Editor with sample results

You can see the sampling results in the Data Editor. Three new variables were saved to the working file, representing the inclusion probabilities and cumulat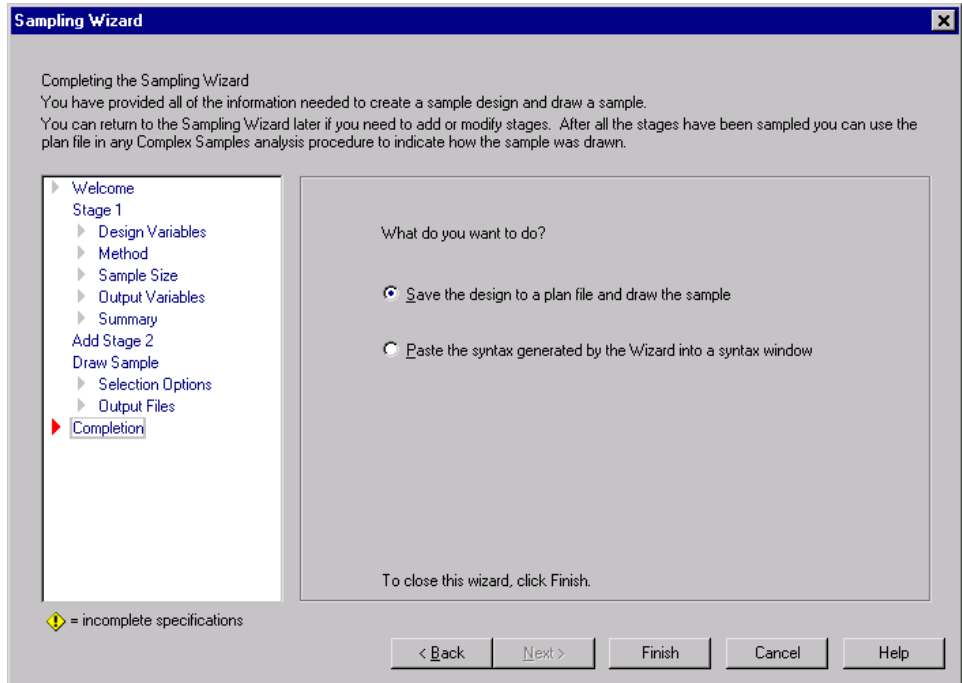ive sampling weights for the third stage, plus the final sampling weights. These new weights take into account the weights computed during the sampling of the first two stages.

- Units with values for these variables were selected to the sample.
- Units with system-missing values for the variables were not selected.

The company will now use its resources to obtain survey information for the housing units selected in the sample. Once the surveys are collected, you can process the sample with Complex Samples analysis procedures, using the sampling plan *demo_2.csplan* to provide the sampling specifications. As an alternative to creating a new sampling plan, you could also edit the existing plan, *demo_1.csplan*, add the third stage to the plan, and execute only the third stage.

Related Procedures

The Complex Samples Sampling Wizard procedure is a useful tool for creating a sampling plan file and drawing a sample.

■ To ready a sample for analysis when you do not have access to the sampling plan file, use the Analysis Preparation Wizard.

Complex Samples Analysis Preparation Wizard

The Analysis Preparation Wizard guides you through the steps for creating or modifying an analysis plan for use with the various Complex Samples analysis procedures. It is most useful when you do not have access to the sampling plan file used to draw the sample.

Using the Complex Samples Analysis Preparation Wizard to Ready NHIS Public Data

The National Health Interview Survey (NHIS) is a large, population-based survey of the U.S. civilian population. Interviews are carried out face-to-face in a nationally representative sample of households. Demographic information and observations about health behavior and status are obtained for members of each household.

A subset of the 2000 survey is collected in *nhis2000_subset.sav*. Use the Complex Samples Analysis Preparation Wizard to create an analysis plan for this data file so that it can be processed by Complex Samples analysis procedures.

Using the Wizard

▶ To prepare a sample using the Complex Samples Analysis Preparation Wizard, from the menus choose:

Analyze
  Complex Samples
    Prepare for Analysis...

Figure 10-1
Analysis Preparation Wizard Welcome step
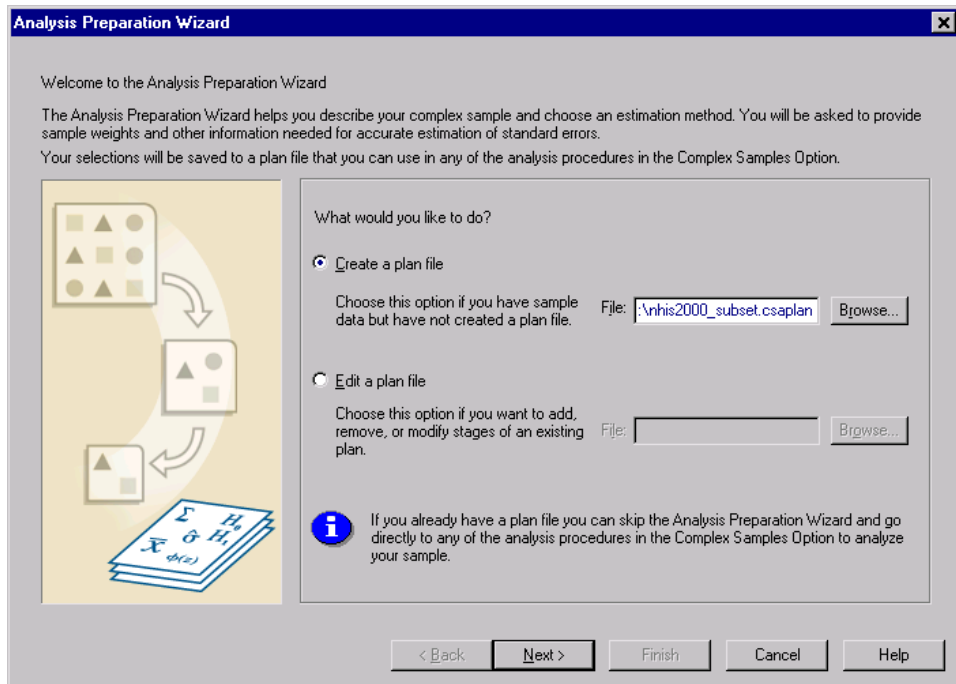
▶ Enter c:\nhis2000_subset.csaplan as the name for the analysis plan file.

▶ Click Next.

Figure 10-2
Analysis Preparation Wizard Design Variables step
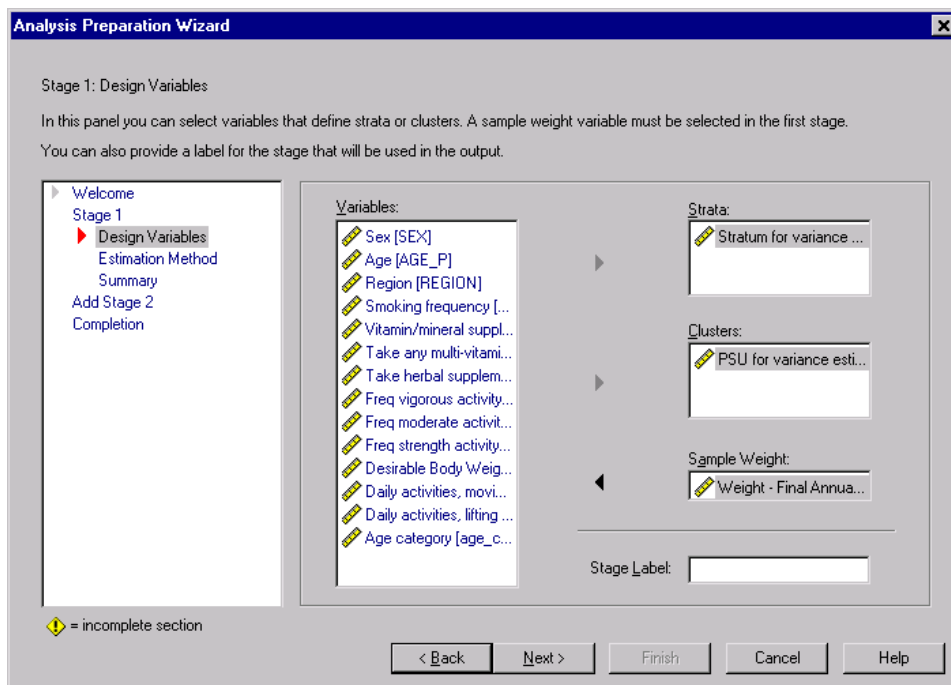
The data are obtained using a complex multistage sample; however, for end users, the original NHIS design variables were transformed to a simplified set of design and weight variables whose results approximate those of the original design structures.

▶ Select *Stratum for variance estimation* as a strata variable.

▶ Select *PSU for variance estimation* as a cluster variable.

▶ Select *Weight - Final Annual* as the sample weight variable.

▶ Click Finish.

Summary

Figure 10-3
Summary

			Stage 1
Design Variables	Stratification	1	Stratum for variance estimation
	Cluster	1	PSU for variance estimation
Analysis Information	Estimator Assumption		
			Sampling with replacement[a]

Plan File: c:\nhis2000_subset.csaplan
Weight Variable: Weight - Final Annual

a.
  The requested estimator assumes sampling with replacement at
  stage 1. Any subsequent stages will not be used for estimation.

The summary table reviews your analysis plan. The plan consists of one stage with a design of one stratification variable and one cluster variable. With-replacement (WR) estimation is used, and the plan is saved to *c:/nhis2000_subset.csaplan*. You can now use this plan file to process *nhis2000_subset.sav* with Complex Samples analysis procedures.

Related Procedures

The Complex Samples Analysis Preparation Wizard procedure is a useful tool for readying a sample for analysis when you do not have access to the sampling plan file.

■ To create a sampling plan file and draw a sample, use the Sampling Wizard.

Complex Samples Frequencies

The Complex Samples Frequencies procedure produces frequency tables for selected variables and displays univariate statistics. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

Frequencies Analysis of Nutritional Supplements Usage

A researcher wants to study the use of nutritional supplements among U.S. citizens, using the results of the National Health Interview Survey (NHIS) and a previously created analysis plan. For more information, see "Using the Complex Samples Analysis Preparation Wizard to Ready NHIS Public Data" in Chapter 10 on page 95.

A subset of the 2000 survey is collected in *nhis2000_subset.sav*. The analysis plan is stored in *nhis2000_subset.csaplan*. Use Complex Samples Frequencies to produce statistics for nutritional supplement usage.

Running the Analysis

▶ To run a Complex Samples Frequencies analysis, from the menus choose:

Analyze
  Complex Samples
    Frequencies...

Figure 11-1
Plan dialog box
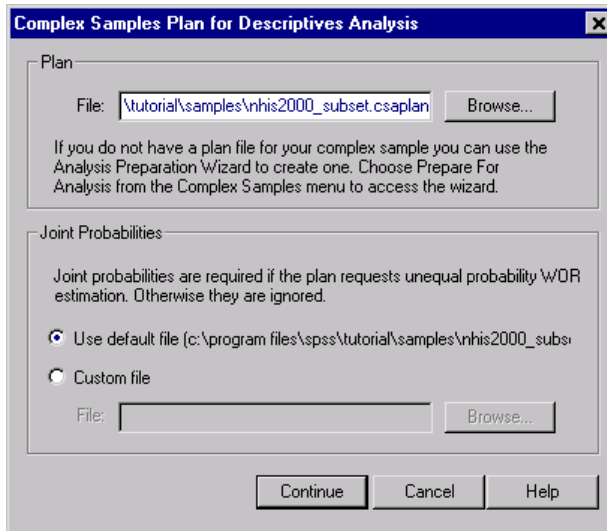
▶ Browse to the *\tutorial\sample_files* subdirectory of the directory in which you installed SPSS and select *nhis2000_subset.csaplan*.

▶ Click Continue.

Figure 11-2
Complex Samples Frequencies dialog box
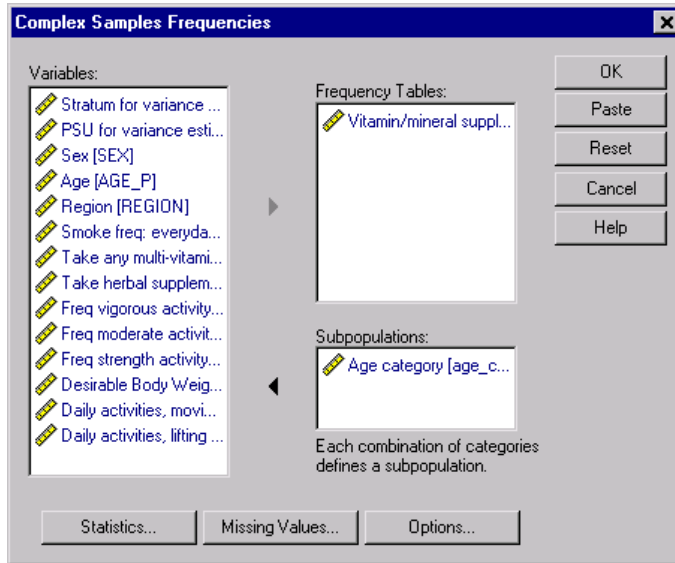
► Select *Vitamin/mineral supplmnts-past 12 m* as a frequency variable.

► Select *Age category* as a subpopulation variable.

► Click Statistics.

Figure 11-3
Statistics dialog box

▶ Select Table percent in the Cells group.

▶ Select Confidence interval in the Statistics group.

▶ Click Continue.

▶ Click OK in the Complex Samples Frequencies dialog box.

Frequency Table

Figure 11-4
Frequency table for variable/situation

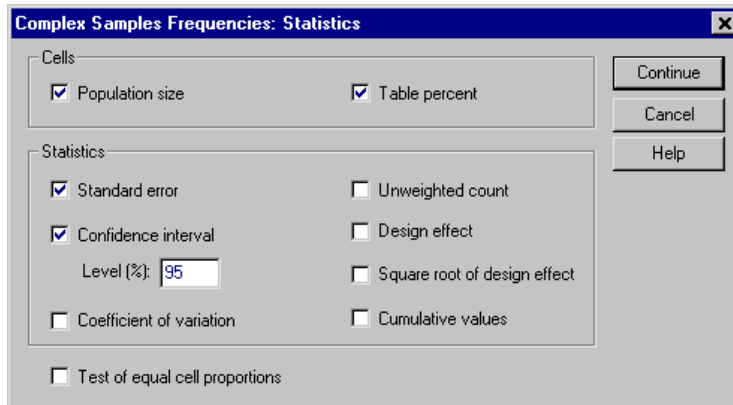|		Estimate	Standard Error	95% Confidence Interval	
				Lower	Upper
Population Size	Yes	1.0E+08	1185126.7	1.0E+08	1.1E+08
	No	9.1E+07	1094401.9	8.9E+07	9.3E+07
	Total	1.9E+08	1789098.7	1.9E+08	2.0E+08
Percent ot total	Yes	53.1%	.4%	52.4%	53.8%
	No	46.9%	.4%	46.2%	47.6%
	Total	100.0%	.0%	100.0%	100.0%

Each selected statistic is computed for each selected cell measure. The first column contains estimates of the number and percentage of the population that do or do not take vitamin/mineral supplements. The confidence intervals are non-overlapping; thus, you can conclude that, overall, more Americans take vitamn/mineral supplements than not.

Frequency by Subpopulation

Figure 11-5
Frequency table by subpopulation

Age category			Estimate	Standard Error	95% Confidence Interval Lower	95% Confidence Interval Upper
18-24	Population Size	Yes	1.0E+07	350602.35	9328882	1.1E+07
		No	1.5E+07	499182.39	1.4E+07	1.6E+07
		Total	2.5E+07	680732.81	2.4E+07	2.7E+07
	Percent of total	Yes	39.3%	1.0%	37.4%	41.2%
		No	60.7%	1.0%	58.8%	62.6%
		Total	100.0%	.0%	100.0%	100.0%
25-44	Population Size	Yes	3.9E+07	660855.72	3.8E+07	4.0E+07
		No	4.0E+07	645934.19	3.8E+07	4.1E+07
		Total	7.9E+07	961114.33	7.7E+07	8.1E+07
	Percent of total	Yes	49.8%	.6%	48.7%	50.9%
		No	50.2%	.6%	49.1%	51.3%
		Total	100.0%	.0%	100.0%	100.0%
45-64	Population Size	Yes	3.4E+07	598603.73	3.3E+07	3.5E+07
		No	2.4E+07	497723.83	2.3E+07	2.5E+07
		Total	5.8E+07	814680.41	5.7E+07	6.0E+07
	Percent of total	Yes	58.7%	.6%	57.5%	60.0%
		No	41.3%	.6%	40.0%	42.5%
		Total	100.0%	.0%	100.0%	100.0%
65+	Population Size	Yes	1.9E+07	439459.79	1.9E+07	2.0E+07
		No	1.2E+07	314238.08	1.1E+07	1.2E+07
		Total	3.1E+07	587623.44	3.0E+07	3.2E+07
	Percent of total	Yes	62.2%	.7%	60.7%	63.6%
		No	37.8%	.7%	36.4%	39.3%
		Total	100.0%	.0%	100.0%	100.0%

When computing statistics by subpopulation, each selected statistic is computed for each selected cell measure by value of *Age category*. The first column contains estimates of the number and percentage of the population of each category that do or do not take vitamin/mineral supplements. The confidence intervals for the table

percentages are all non-overlapping; thus, you can conclude that with increasing age category, greater proportions of Americans take vitamn/mineral supplements.

Summary

Using the Complex Samples Frequencies procedure, you have obtained statistics for the use of nutritional supplements among U.S. citizens.

- Overall, More Americans Take Vitamn/Mineral Supplements Than Not.
- When broken down by age category, greater proportions of Americans take vitamin/mineral supplements with increasing age.

Related Procedures

The Complex Samples Frequencies procedure is a useful tool for obtaining univariate descriptive statistics of categorical variables for observations obtained via a complex sampling design.

- The Complex Samples Sampling Wizard is used to specify complex sampling design specifications and obtain a sample. The sampling plan file created by the Sampling Wizard contains a default analysis plan and can be specified in the Plan dialog box when analyzing the sample obtained according to that plan.
- The Complex Samples Analysis Preparation Wizard is used to set analysis specifications for an existing complex sample. The analysis plan file created by the Sampling Wizard can be specified in the Plan dialog box when analyzing the sample corresponding to that plan.
- The Complex Samples Crosstabs procedure provides descriptive statistics for the crosstabulation of categorical variables.
- The Complex Samples Descriptives procedure provides univariate descriptive statistics for scale variables.

Complex Samples Descriptives

The Complex Samples Descriptives procedure displays univariate summary statistics for several variables. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

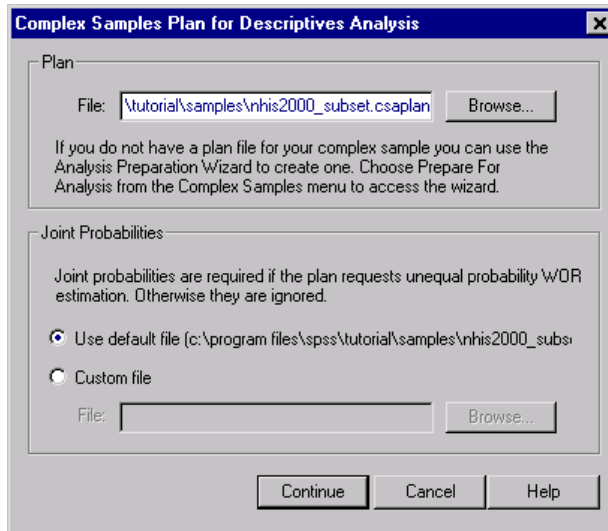## *Using Complex Samples Descriptives to Analyze Activity Levels*

A researcher wants to study the activity levels of U.S. citizens, using the results of the National Health Interview Survey (NHIS) and a previously created analysis plan. For more information, see "Using the Complex Samples Analysis Preparation Wizard to Ready NHIS Public Data" in Chapter 10 on page 95.

A subset of the 2000 survey is collected in *nhis2000_subset.sav*. The analysis plan is stored in *nhis2000_subset.csaplan*. Use Complex Samples Descriptives to produce univariate descriptive statistics for activity levels.
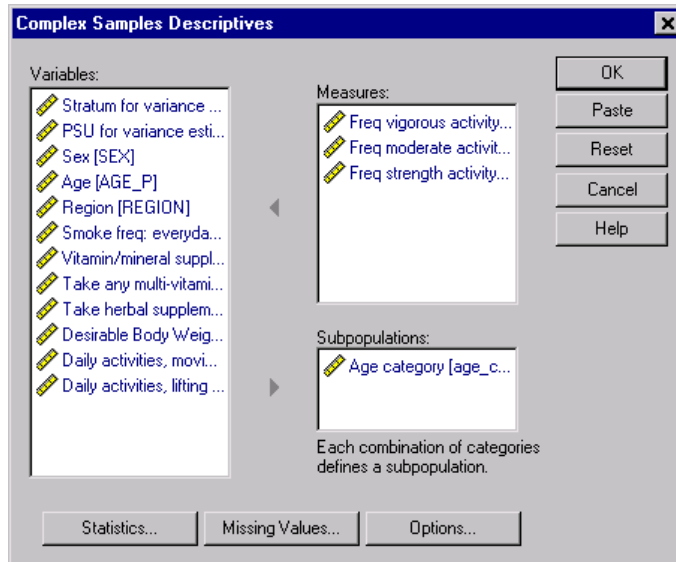
Running the Analysis

▶ To run a Complex Samples Descriptives analysis, from the menus choose:

Analyze
  Complex Samples
    Descriptives...

Figure 12-1
Plan dialog box


▶ Browse to the *\tutorial\sample_files* subdirectory of the directory in which you installed SPSS and select *nhis2000_subset.csaplan*.

▶ Click Continue.

Figure 12-2
Complex Samples Descriptives dialog box

▶ Select *Freq vigorous activity (times per wk)* through *Freq strength activity (times per wk)* as measure variables.

▶ Select *Age category* as a subpopulation variable.

▶ Click Statistics.

Figure 12-3
Statistics dialog box
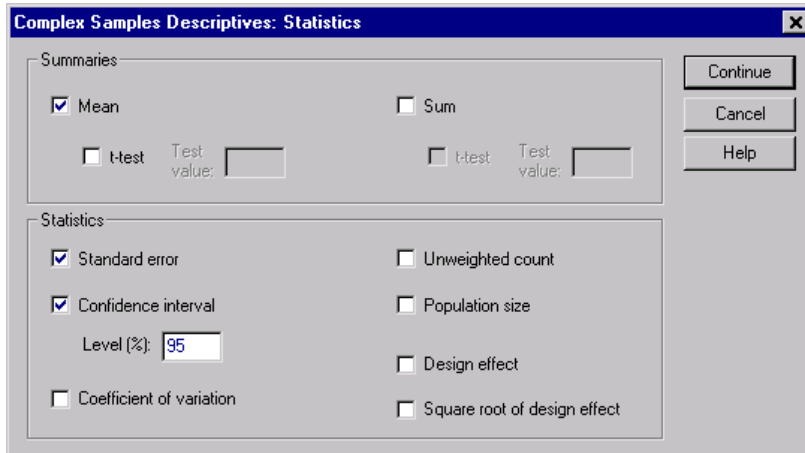
▶ Select Confidence interval in the Statistics group.

▶ Click Continue.

▶ Click OK in the Complex Samples Descriptives dialog box.

Univariate Statistics

Figure 12-4
Univariate statistics

		Estimate	Standard Error	95% Confidence Interval	
				Lower	Upper
Mean	Freq vigorous activity (times per wk)	3.73	.033	3.66	3.79
	Freq moderate activity (times per wk)	4.90	.041	4.82	4.98
	Freq strength activity (times per wk)	3.52	.042	3.43	3.60

Each selected statistic is computed for each measure variable. The first column contains estimates of the average number of times per week a person engages in a particular type of activity. The confidence intervals for the means are non-overlapping; thus, you can conclude that, overall, Americans engage in a strength activity less often than vigorous activity, and they engage in vigorous activity less often than moderate activity.

Univariate Statistics by Subpopulation

Figure 12-5
Univariate statistics by subpopulation

Age category			Estimate	Standard Error	95% Confidence Interval	
					Lower	Upper
18-24	Mean	Freq vigorous activity (times per wk)	3.92	.087	3.75	4.09
		Freq moderate activity (times per wk)	5.18	.137	4.91	5.45
		Freq strength activity (times per wk)	3.45	.085	3.28	3.62
25-44	Mean	Freq vigorous activity (times per wk)	3.55	.048	3.46	3.65
		Freq moderate activity (times per wk)	4.73	.056	4.62	4.84
		Freq strength activity (times per wk)	3.28	.052	3.18	3.38
45-64	Mean	Freq vigorous activity (times per wk)	3.79	.063	3.66	3.91
		Freq moderate activity (times per wk)	4.88	.070	4.74	5.02
		Freq strength activity (times per wk)	3.65	.092	3.47	3.84
65+	Mean	Freq vigorous activity (times per wk)	4.18	.111	3.96	4.39
		Freq moderate activity (times per wk)	5.22	.084	5.06	5.39
		Freq strength activity (times per wk)	4.66	.155	4.36	4.97

Each selected statistic is computed for each measure variable by values of *Age category*. The first column contains estimates of the average number of times per week that people of each category engage in a particular type of activity. The confidence intervals for the means allow you to make some interesting conclusions.

- In terms of vigorous and moderate activities, 25–44-year-olds are less active than those 18–24 and 45–64, and 45–64-year-olds are less active than those 65 or older.

- In terms of strength activity, 25–44-year-olds are less active than those 45–64, and 18–24 and 45–64-year-olds are less active than those 65 or older.

Summary

Using the Complex Samples Descriptives procedure, you have obtained statistics for the activity levels of U.S. citizens.

- Overall, Americans spend varying amounts of time at different types of activities.

- When broken down by age, it roughly appears that post-collegiate Americans are initially less active than they were while in school but become more conscientious about exercising as they age.

Related Procedures

The Complex Samples Descriptives procedure is a useful tool for obtaining univariate descriptive statistics of scale measures for observations obtained via a complex sampling design.

- The Complex Samples Sampling Wizard is used to specify complex sampling design specifications and obtain a sample. The sampling plan file created by the Sampling Wizard contains a default analysis plan and can be specified in the Plan dialog box when analyzing the sample obtained according to that plan.

- The Complex Samples Analysis Preparation Wizard is used to specify analysis specifications for an existing complex sample. The analysis plan file created by the Sampling Wizard can be specified in the Plan dialog box when analyzing the sample corresponding to that plan.

- The Complex Samples Ratios procedure provides descriptive statistics for ratios of scale measures.

- The Complex Samples Frequencies procedure provides univariate descriptive statistics of categorical variables.

Complex Samples Crosstabs

The Complex Samples Crosstabs procedure produces crosstabulation tables for pairs of selected variables and displays two-way statistics. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

## *Using Crosstabs to Measure the Relative Risk of an Event*

A company that sells magazine subscriptions traditionally sends monthly mailings to a purchased database of names. The response rate is typically low, so you need to find a way to better target prospective customers. One suggestion was to focus mailings on people with newspaper subscriptions, on the assumption that people who read newspapers are more likely to subscribe to magazines.

Use the Complex Samples Crosstabs procedure to test this theory by constructing a two-by-two table of *Newspaper subscription* by *Response* and compute the relative risk that a person with a newspaper subscription will respond to the mailing. This information is collected in *demo_cs.sav*, and should be analyzed using the sampling plan file *demo_2.csplan*.

Running the Analysis

► To run a Complex Samples Crosstabs analysis, from the menus choose:

Analyze
  Complex Samples
    Crosstabs...

Figure 13-1
Plan dialog box
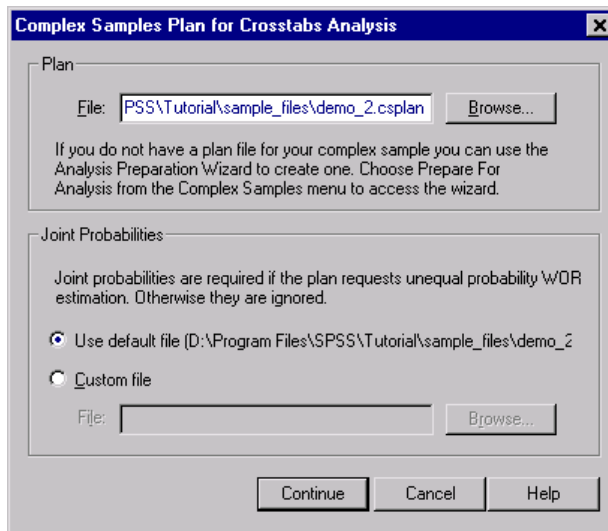

▶ Browse to the *\tutorial\sample_files* subdirectory of the directory in which you installed SPSS and select *demo_2.csplan*.

▶ Click Continue.

Figure 13-2
Crosstabs dialog box
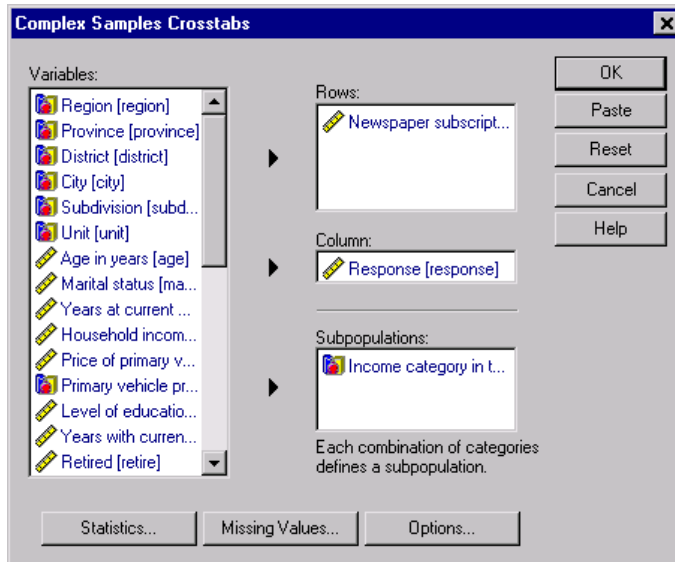
▶ Select *Newspaper subscription* as a row variable.

▶ Select *Response* as the column variable.

▶ There is also some interest in seeing the results broken down by income categories, so select *Income category in thousands* as a subpopulation variable.

▶ Click Statistics.

Figure 13-3
Statistics dialog box
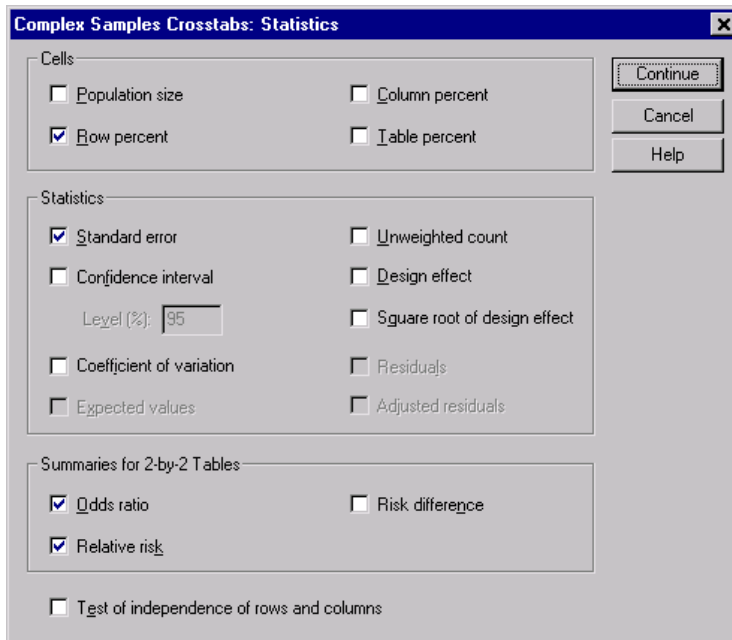
▶ Deselect Population size and select Row percent in the Cells group.

▶ Select Odds ratio and Relative risk in the Summaries for 2-by-2 Tables group.

▶ Click Continue.

▶ Click OK in the Complex Samples Crosstabs dialog box.

These selections produce a crosstabulation table and risk estimate for *Newspaper subscription* by *Response*. Separate tables with results split by *Income category in thousands* are also created.

Crosstabulation

Figure 13-4
Crosstabulation for Newspaper subscription by Response

Newspaper subscription			Response		
			Yes	No	Total
Yes	% within Newspaper subscription	Estimate	18.6%	81.4%	100.0%
		Standard Error	1.2%	1.2%	.0%
No	% within Newspaper subscription	Estimate	10.6%	89.4%	100.0%
		Standard Error	.7%	.7%	.0%
Total	% within Newspaper subscription	Estimate	13.5%	86.5%	100.0%
		Standard Error	.7%	.7%	.0%

The crosstabulation shows that, overall, not very many people responded to the mailing. However, a greater proportion of newspaper subscribers responded.

Risk Estimate

Figure 13-5
Risk estimate for Newspaper subscription by Response

			Estimate
Newspaper subscription * Response	Odds Ratio		1.912
	Relative Risk	For cohort Response = Yes	1.743
		For cohort Response = No	.911

Statistics are computed only for 2-by-2 tables with all cells observed.

The relative risk is a ratio of event probabilities. The relative risk of a response to the mailing is the ratio of the probability that a newspaper subscriber responds, to the probability that a nonsubscriber responds. Thus, the estimate of the relative risk is simply 18.6%/10.6% = 1.743. Likewise, the relative risk of nonresponse is the ratio of the probability that a subscriber does not respond, to the probability that a nonsubscriber does not respond. Your estimate of this relative risk is 0.911. Given these results, you can estimate that a newspaper subscriber is 1.743 times as likely to respond to the mailing as a nonsubscriber, or 0.911 times as likely as a nonsubscriber not to respond.

The odds ratio is a ratio of event odds. The odds of an event is the ratio of the probability that the event occurs, to the probability that the event does not occur. Thus, the estimate of the odds that a newspaper subscriber responds to the mailing

is 18.6%/81.4% = 0.229. Likewise, the estimate of the odds that a nonsubscriber responds is 10.6%/89.4% = 0.118. The estimate of the odds ratio is therefore 0.229/0.118 = 1.912. Note that the odds ratio is the ratio of the relative risk of responding, to the relative risk of not responding, or 1.743/0.911 = 1.912.

Odds Ratio versus Relative Risk

Since it is a ratio of ratios, the odds ratio is very difficult to interpret. The relative risk is easier to interpret, so the odds ratio alone is not very helpful. However, there are certain commonly occurring situations in which the estimate of the relative risk is not very good, and the odds ratio can be used to approximate the relative risk of the event of interest. The odds ratio should be used as an approximation to the relative risk of the event of interest when both of the following conditions are met:

- The probability of the event of interest is small (<0.1). This condition guarantees that the odds ratio will make a good approximation to the relative risk. In this example, the event of interest is a response to the mailing.

- The design of the study is case control. This condition signals that the usual estimate of the relative risk will likely not be good. A case-control study is retrospective, most often used when the event of interest is unlikely or when the design of a prospective experiment is impractical or unethical.

Neither condition is met in this example, since the overall proportion of respondents was 13.5% and the design of the study was not case control, so it's safer to report 1.743 as the relative risk, rather than the value of the odds ratio.

Risk Estimate by Subpopulation

Figure 13-6
Risk estimate for Newspaper subscription by Response, controlling for Income category

Income category				Estimate
Under $25	Newspaper subscription * Response	Odds Ratio		2.608
		Relative Risk	For cohort Response = Yes	2.149
			For cohort Response = No	.824
$25 - $49	Newspaper subscription * Response	Odds Ratio		1.923
		Relative Risk	For cohort Response = Yes	1.737
			For cohort Response = No	.903
$50 - $74	Newspaper subscription * Response	Odds Ratio		1.538
		Relative Risk	For cohort Response = Yes	1.493
			For cohort Response = No	.971
$75+	Newspaper subscription * Response	Odds Ratio		1.200
		Relative Risk	For cohort Response = Yes	1.182
			For cohort Response = No	.985

Statistics are computed only for 2-by-2 tables with all cells observed.

Relative risk estimates are computed separately for each income category. Note that the relative risk of a positive response for newspaper subscribers appears to gradually decrease with increasing income, which indicates that you may be able to further target the mailings.

Summary

Using Complex Samples Crosstabs' risk estimates, you found that you can increase your response rate to direct mailings by targeting newspaper subscribers. Further, you found some evidence that the risk estimates may not be constant across *Income category*, so you may be able to increase your response rate even more by targeting lower-income newspaper subscribers.

Related Procedures

The Complex Samples Crosstabs procedure is a useful tool for obtaining descriptive statistics of the crosstabulation of categorical variables for observations obtained via a complex sampling design.

■ The Complex Samples Sampling Wizard is used to specify complex sampling design specifications and obtain a sample. The sampling plan file created by the Sampling Wizard contains a default analysis plan and can be specified in the Plan dialog box when analyzing the sample obtained according to that plan.

■ The Complex Samples Analysis Preparation Wizard is used to specify analysis specifications for an existing complex sample. The analysis plan file created by the Sampling Wizard can be specified in the Plan dialog box when analyzing the sample corresponding to that plan.

■ The Complex Samples Frequencies procedure provides univariate descriptive statistics for categorical variables.

Complex Samples Ratios

The Complex Samples Ratios procedure displays univariate summary statistics for ratios of variables. Optionally, you can request statistics by subgroups, defined by one or more categorical variables.

Using Complex Samples Ratios to Aid Property Value Assessment

A state agency is charged with ensuring that property taxes are fairly assessed from county to county. Taxes are based on the appraised value of the property, so the agency wants to track property values across counties to be sure that each county's records are equally up-to-date. Since resources for obtaining current appraisals are limited, the agency chose to employ complex sampling methodology to select properties.

The sample of properties selected and their current appraisal information is collected in *property_assess_cs_sample.sav*. Use Complex Samples Ratios to assess the change in property values since the last appraisal across the five counties.

Running the Analysis

▶ To run a Complex Samples Ratios analysis, from the menus choose:

Analyze
  Complex Samples
    Ratios...

Figure 14-1
Plan dialog box
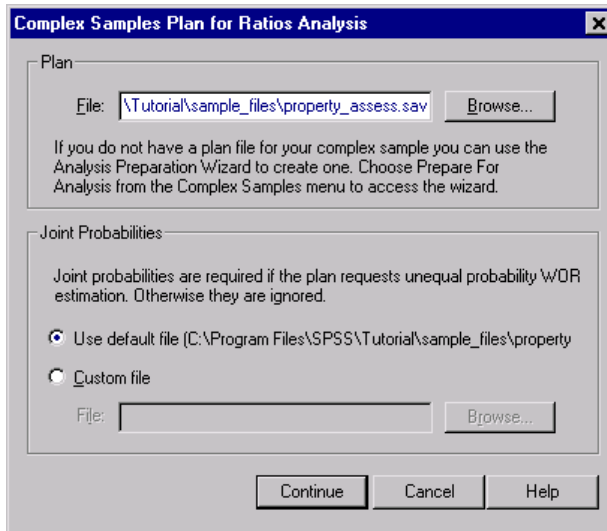
▶ Browse to the *\tutorial\sample_files* subdirectory of the directory in which you installed SPSS and select *property_assess.csplan*.

▶ Click Continue.

Figure 14-2
Complex Samples Ratios dialog box
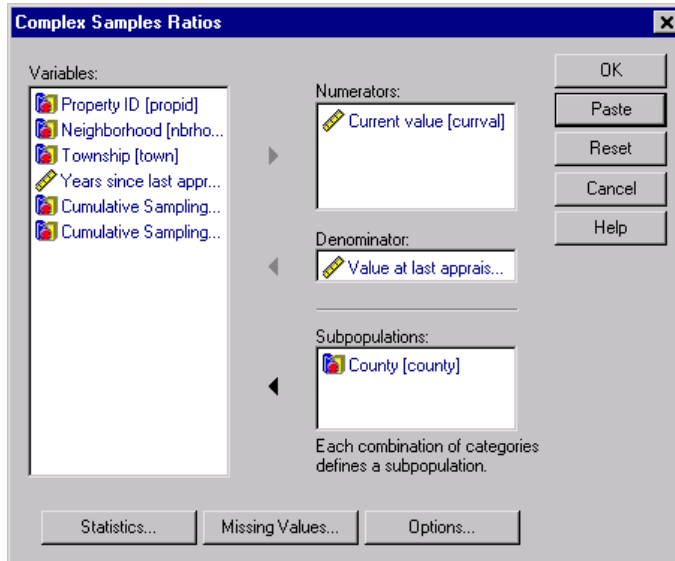
▶ Select *Current value* as a numerator variable.

▶ Select *Value at last appraisal* as the denominator variable.

▶ Select *County* as a subpopulation variable.

▶ Click Statistics.

Figure 14-3
Statistics dialog box
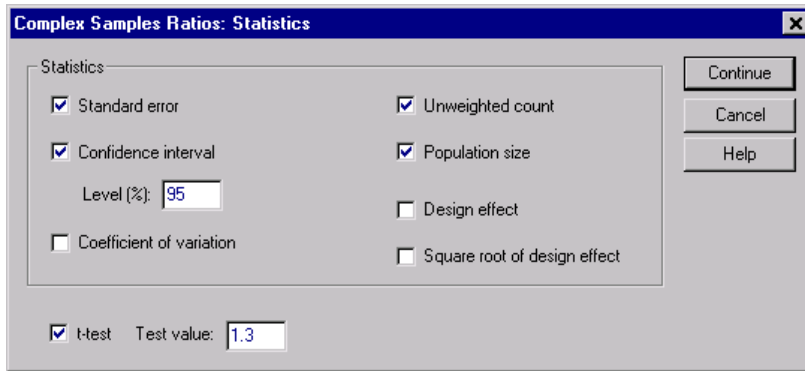
▶ Select Confidence interval, Unweighted count, and Population size in the Statistics group.

▶ Select t-test and enter 1.3 as the test value.

▶ Click Continue.

▶ Click OK in the Complex Samples Ratios dialog box.

Ratios

Figure 14-4
Ratios table

County	Numerator	Denominator	Ratio Estimate	Standard Error	95% Confidence Interval	
					Lower	Upper
Eastern	Current value	Value at last appraisal	1.381	.068	1.236	1.52
Central	Current value	Value at last appraisal	1.364	.064	1.227	1.50
Western	Current value	Value at last appraisal	1.524	.053	1.410	1.63
Northern	Current value	Value at last appraisal	1.277	.032	1.208	1.34
Southern	Current value	Value at last appraisal	1.195	.029	1.134	1.25

The default display of the table is very wide, so you will need to pivot it for a better view.

Pivoting the Ratios Table

▶ Double-click the table to activate it.

▶ From the Viewer menus choose:
  Pivot
    Pivoting Trays

▶ Drag *Numerator* and then *Denominator* from the row to the layer.

▶ Drag *County* from the row to the column.

▶ Drag *Statistics* from the column to the row.

▶ Close the pivoting trays window.

Pivoted Ratios Table

Figure 14-5
Pivoted Ratios table

Numerator: Current value
Denominator: Value at last appraisal

		County				
		Eastern	Central	Western	Northern	Southern
Ratio Estimate		1.381	1.364	1.524	1.277	1.195
Standard Error		.068	.064	.053	.032	.029
95% Confidence Interval	Lower	1.236	1.227	1.410	1.208	1.134
	Upper	1.525	1.502	1.638	1.346	1.256
Hypothesis Test	Test Value	1.3	1.3	1.3	1.3	1.3
	t	1.191	.997	4.201	-.702	-3.646
	df	15	15	15	15	15
	Sig.	.252	.334	.001	.493	.002
Population Size		1883.250	1557.500	4044.000	2306.250	2204.000
Unweighted Count		168	179	202	205	220

The ratios table is now pivoted so that statistics are easier to compare across counties.

■ The ratio estimates range from a low of 1.195 in the Southern county to a high of 1.524 in the Western county.

■ There is also quite a bit of variability in the standard errors, which range from a low of 0.029 in the Southern county to 0.068 in the Eastern county.

■ Some of the confidence intervals do not overlap; thus, you can conclude that the ratios for the Western county are higher than the ratios for the Northern and Southern counties.

■ Finally, as a more objective measure, note that the significance values for the *t*-tests for the Western and Southern counties are less than 0.05. Thus, you can conclude that the ratio for the Western county is greater than 1.3 and the ratio for the Southern county is less than 1.3.

Summary

Using the Complex Samples Ratios procedure, you have obtained various statistics for the ratios of *Current value* to *Value at last appraisal*. The results suggest that there may be certain iniquities in the assessment of property taxes from county to county, namely:

■ The ratios for the Western county are high, indicating that their records are not as up-to-date as other counties with respect to the appreciation of property values. Property taxes are probably too low in this county.

■ The ratios for the Southern county are low, indicating that their records are more up-to-date than the other counties with respect to the appreciation of property values. Property taxes are probably too high in this county.

■ The ratios for the Southern county are lower than those of Western county, but still within the objective goal of 1.3.

Resources used to track property values in Southern county will be reassigned to Western county to bring these counties' ratios in line with the others and the goal of 1.3.

Related Procedures

The Complex Samples Ratios procedure is a useful tool for obtaining univariate descriptive statistics of the ratio of scale measures for observations obtained via a complex sampling design.

- The Complex Samples Sampling Wizard is used to specify complex sampling design specifications and obtain a sample. The sampling plan file created by the Sampling Wizard contains a default analysis plan and can be specified in the Plan dialog box when analyzing the sample obtained according to that plan.

- The Complex Samples Analysis Preparation Wizard is used to specify analysis specifications for an existing complex sample. The analysis plan file created by the Sampling Wizard can be specified in the Plan dialog box when analyzing the sample corresponding to that plan.

- The Complex Samples Descriptives procedure provides descriptive statistics for individual scale measures.

Bibliography

Cochran, W. G. 1977. *Sampling Techniques.* New York: John Wiley and Sons.

Hays, W. L. 1981. *Statistics.* New York: Holt, Rinehart, and Winston.

Kish, L. 1965. *Survey Sampling.* New York: John Wiley and Sons.

Kish, L. 1987. *Statistical Design for Research.* New York: John Wiley and Sons.

Murthy, M. N. 1967. *Sampling Theory and Methods.* Calcutta, India: Statistical Publishing Society.

Index