



CHILDREN'S HEALTHY LIVING (CHL) CENTER OF EXCELLENCE

Complex Survey Weights for Data Analysis

Developed by the CHL Data Work Group
for use in the CHL Pacific Region

For further information, please contact:

CHL Data Center

Children's Healthy Living Center of Excellence

Ag Sciences 302, 1955 East West Road, Honolulu, HI 96822

Email: chldata@hawaii.edu

Website: www.chl-pacific.org

This project is supported by the Agriculture and Food Research Initiative Grant no. 2021-68012-35899 and Grant no. 2018-69001-27551 from the USDA National Institute of Food and Agricultural Science Enhancement Coordinated Agricultural Program and Grant no. 1P20GM139753-01A1 from the Centers of Biomedical Research Excellence (COBRE).



United States Department of Agriculture
National Institute of Food and Agriculture

BACKGROUND

A community-randomized intervention was conducted in the 5 jurisdictions of Alaska, Hawai'i, American Samoa, Guam and CNMI (called intervention jurisdictions or INT). Child and community level data were collected at baseline (time 1) and post-intervention at 24 months (time 2) and 72 months (time 3). A prevalence survey was conducted in the Freely Associated States (FAS) of Federated States of Micronesia (FSM), Republic of Palau, and the Republic of the Marshall Islands (RMI). In all jurisdictions, CHL data were collected within community clusters. For more details on the CHL's study design, see paper by [Wilkins et al.](#)¹

Due to this complex survey design, specific adjustments must be made for analyses with CHL individual-level data. For studies including 2 or more CHL jurisdictions in analyses (e.g., studies including the whole CHL Pacific region), there must be adjustments on the clustering of participants in communities (variable: `community_cluster`) within jurisdiction strata (variable: `jurisnum`) and with survey weights (variable: `wt_anthro_adj`). For studies including only 1 jurisdiction in analyses, there must be adjustments on the clustering of participants in communities only (variable: `community_cluster`).

Sample program codes for SAS, R, and STATA and instructions for SPSS to include this complex sample design are shown in the next sections.

For any questions, please email the CHL Data Center at chldata@hawaii.edu.

¹Wilken LR, Novotny R, Fialkowski MK, et al. Children's Healthy Living (CHL) Program for remote underserved minority populations in the Pacific region: rationale and design of a community randomized trial to prevent early childhood obesity. *BMC Public Health*. 2013;13:944. Published 2013 Oct 9. doi:10.1186/1471-2458-13-944

SAMPLE PROGRAM CODE FOR SAS, R, & STATA

Highlighted are codes for adding complex survey weights.

1. SAS

```
/*Survey design for analyses that include 2 or more jurisdictions*/
proc surveyreg/surveylogistic data=dataset;
class covariates;
model outcome = exposures covariates;
weight wt_anthro_adj;
strata jurisnum;
cluster community_cluster;
run;

/*Survey design for analyses that include only 1 jurisdiction*/
proc surveyreg/surveylogistic data=dataset;
class covariates;
model outcome = exposures covariates;
cluster community_cluster;
run;
```

2. R

```
# By using the svyglm function from the survey package and setting up your survey design,
# you can adjust for the survey weights, strata, and clusters. This provides more accurate
# estimates and standard errors given the survey design.
```

```
# Survey design for analyses that include 2 or more jurisdictions
design <- svydesign(ids =~ id, cluster = ~community_cluster, strata = ~jurisnum, weights
= ~wt_anthro_adj, data = dataset)
```

```
# Survey design for analyses that include only 1 jurisdiction
design <- svydesign(ids =~ id, cluster = ~community_cluster, data = dataset)
```

```
outcome <- svyglm(exposure ~ covariate + covariate, design = design, family =
"binomial")
```

NOTE:

To use *svydesign*, you must install the survey package and load the library with the same name.

If the package is already installed in your environment then you can load it:

```
install.packages("survey")
library(survey)
```

3. STATA

```
// svyset [psu] [weight] [, strata(varname) fpc(varname)]
```

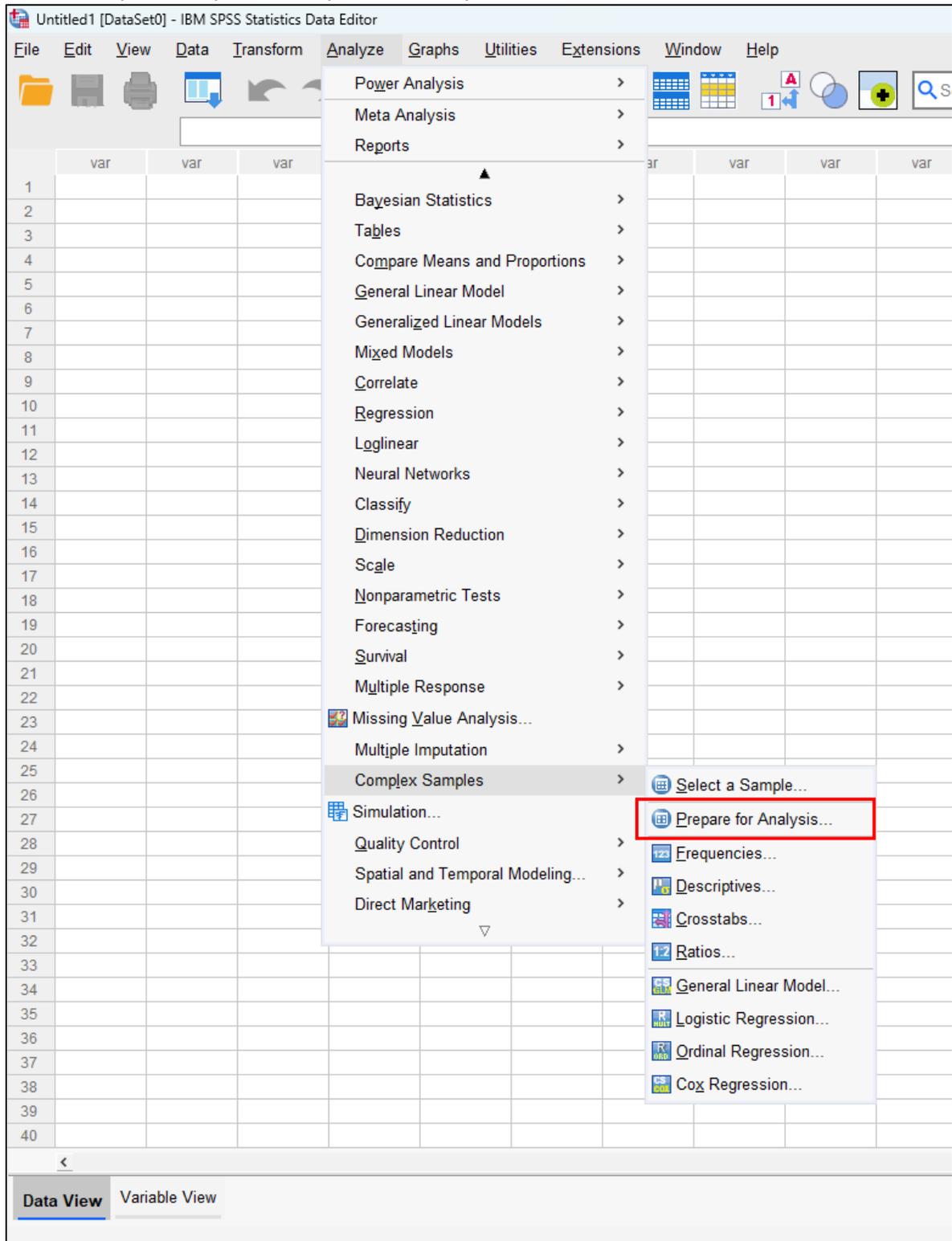
```
// Survey design for analyses that include 2 or more jurisdictions  
svyset community_cluster [pweight = wt_anthro_adj], strata(jurisnum)
```

```
// Survey design for analyses that include only 1 jurisdiction  
svyset community_cluster
```

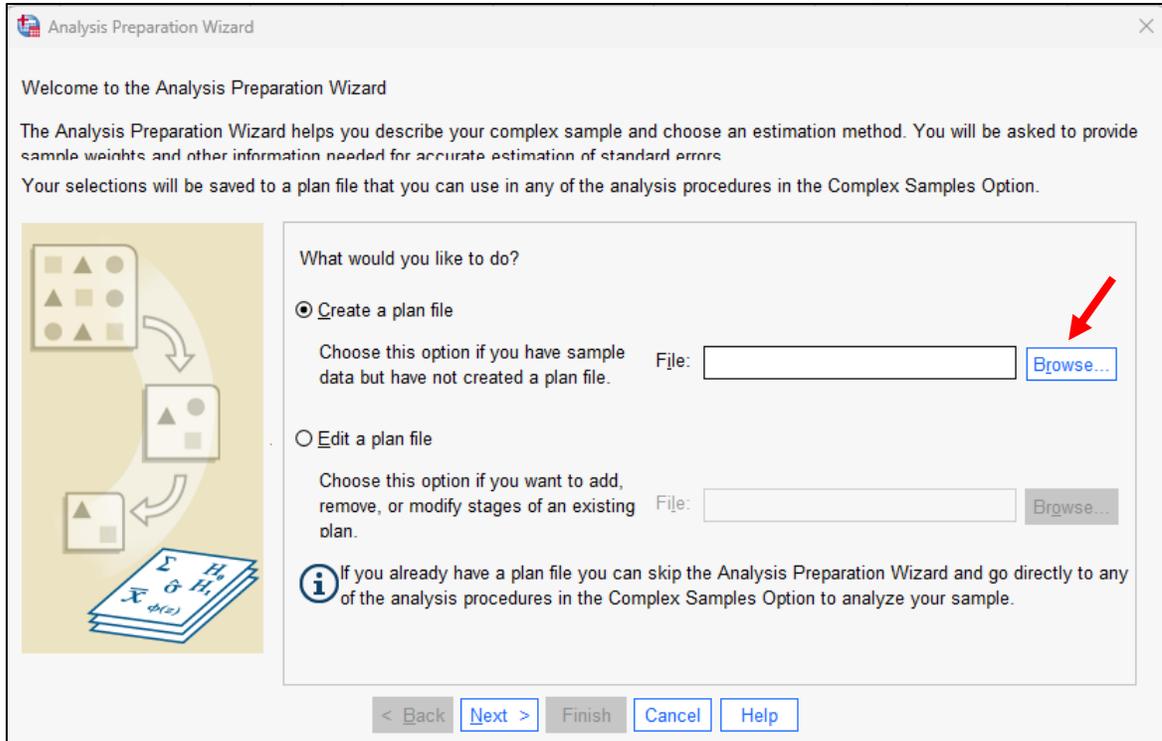
```
svy: regress outcome exposures covariates
```

INSTRUCTIONS FOR PREPARING COMPLEX SAMPLES IN SPSS

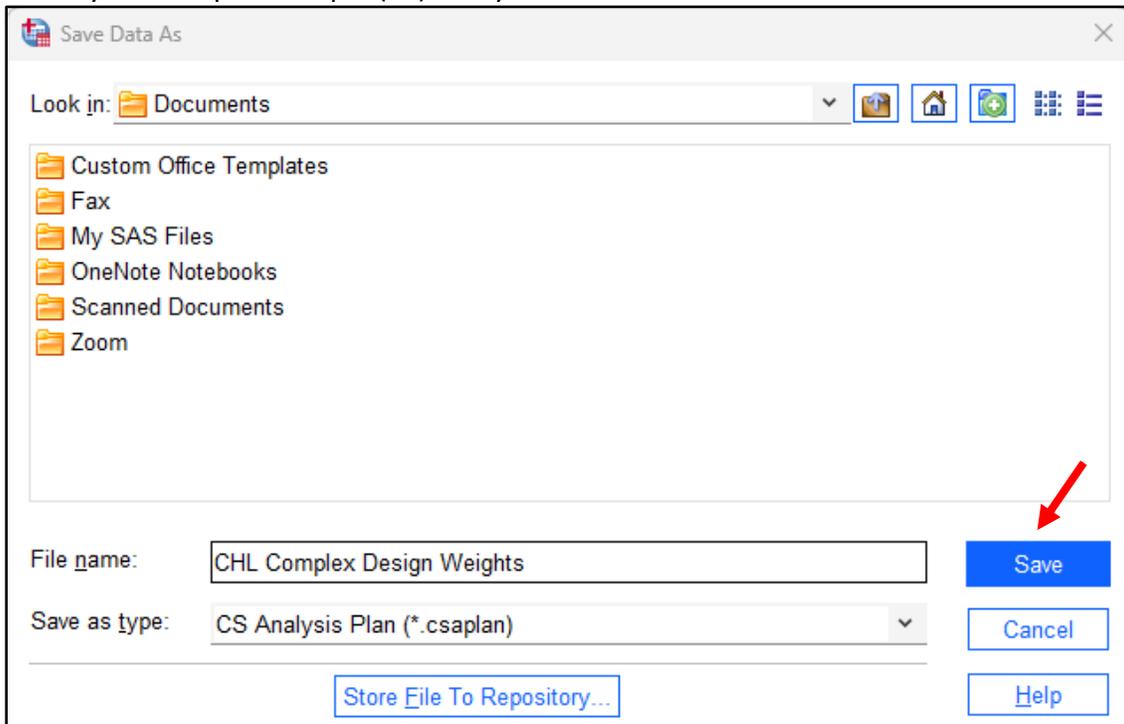
Select Complex Samples → Prepare for Analysis.



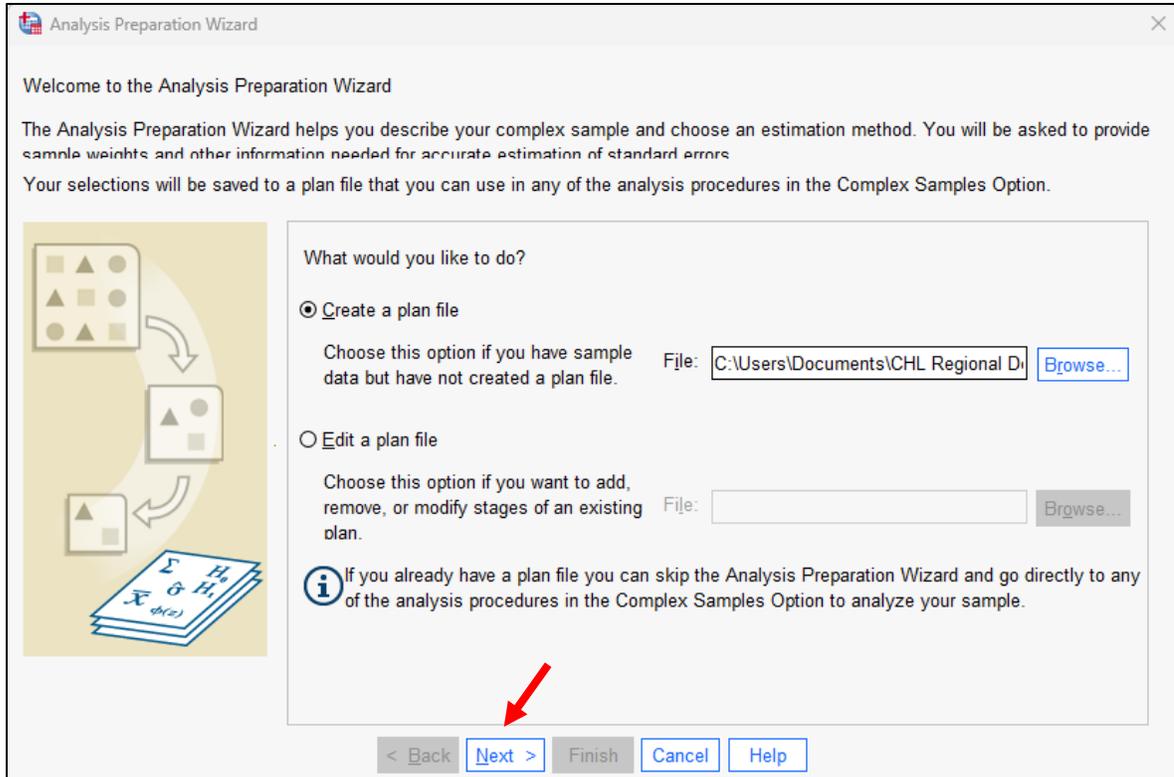
Use the Analysis Preparation Wizard to set you complex survey weights.



Name your Complex Sample (CS) Analysis Plan and Save.

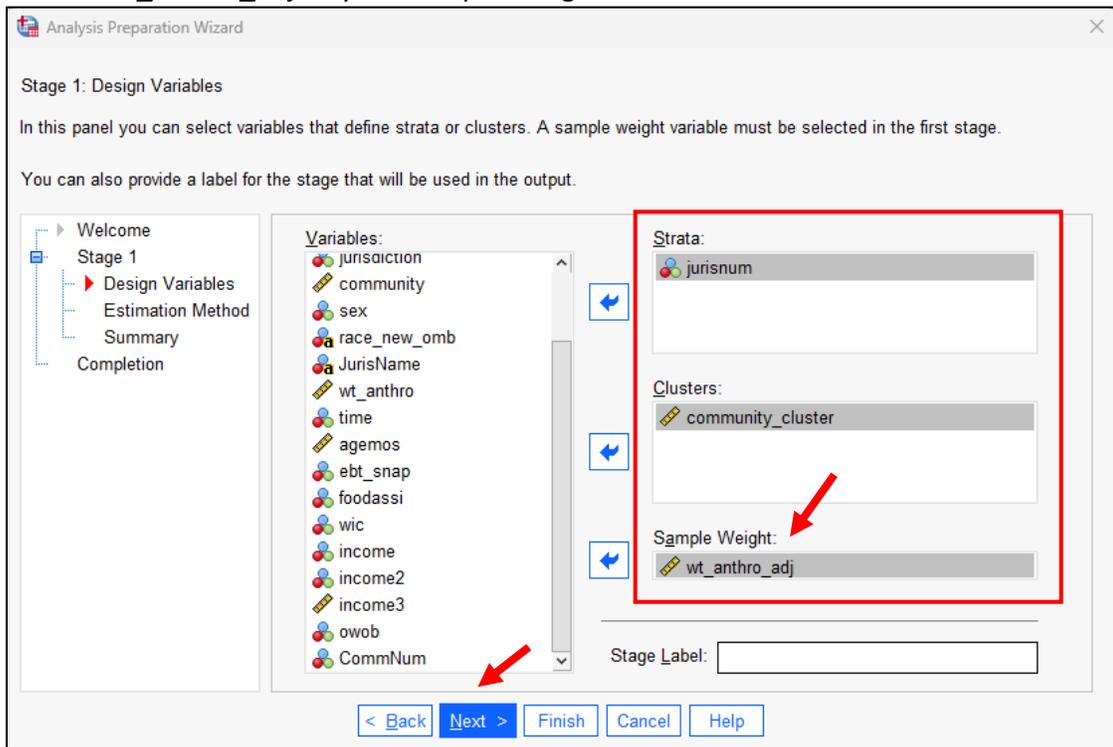


Click Next.



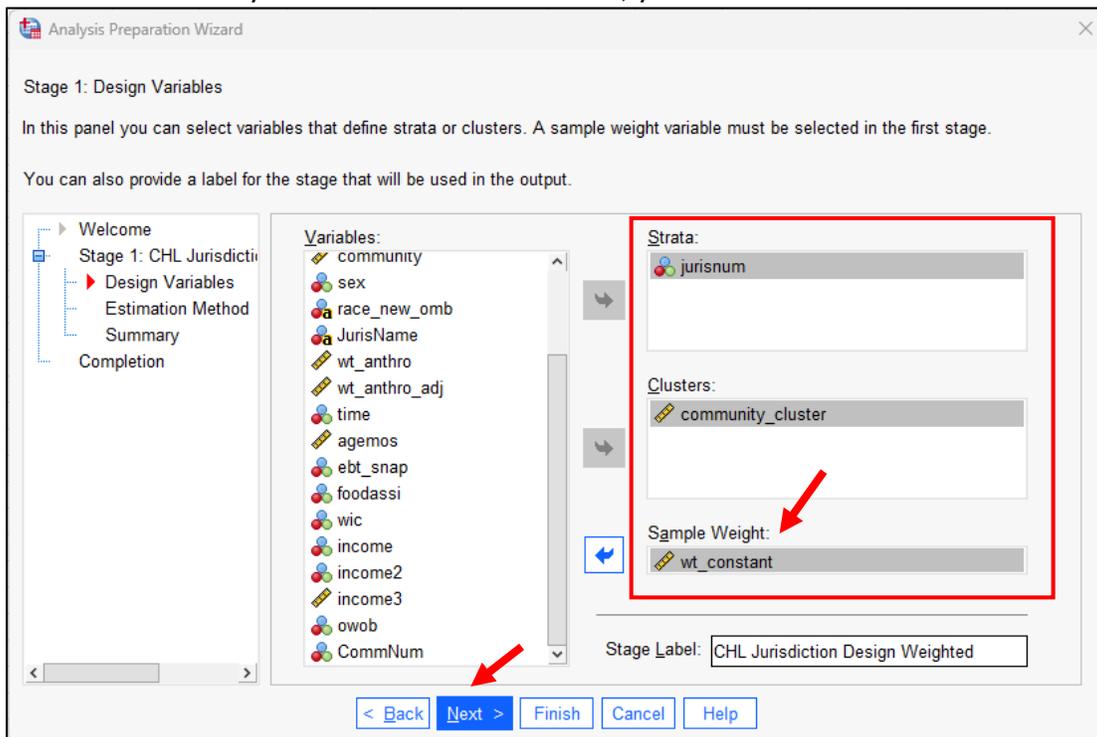
****The next window shows what Sample Weight to include depending on if you are including 2 or more jurisdictions OR only 1 jurisdiction in your analyses. Carefully follow the next page to decide what Sample Weight to include.**

The following step is for analyses including 2 or more jurisdictions:
Include *wt_anthro_adj* as your Sample Weight



The following step is for analyses including only 1 jurisdiction:

Include *wt_constant* as your Sample Weight, *wt_constant* is a default variable set to 1 for all observations. If you do not have this variable, you can create it.



****The remaining steps are applied to analyses including EITHER 2 or more jurisdictions OR only 1 jurisdiction.**

Select WR (sampling with replacement).

Analysis Preparation Wizard

Stage 1: Estimation Method

In this panel you select a method for estimating standard errors.

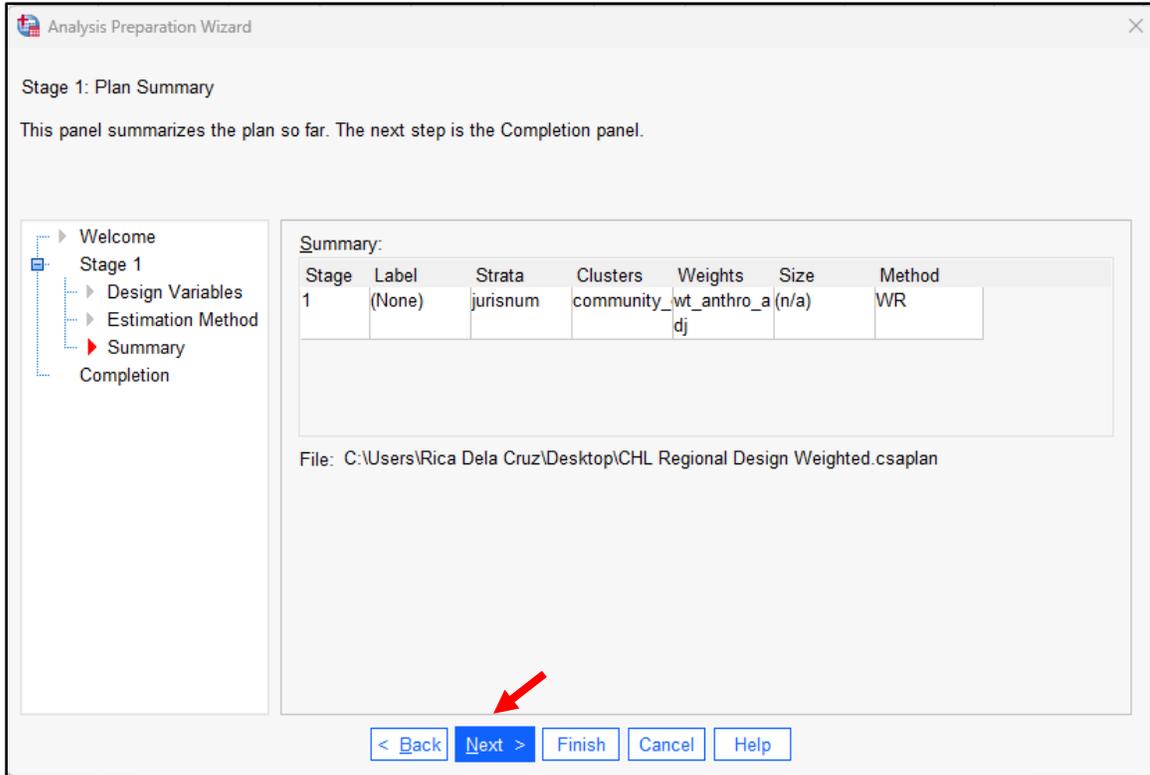
The estimation method depends on assumptions about how the sample was drawn.

Which of the following sample designs should be assumed for estimation?

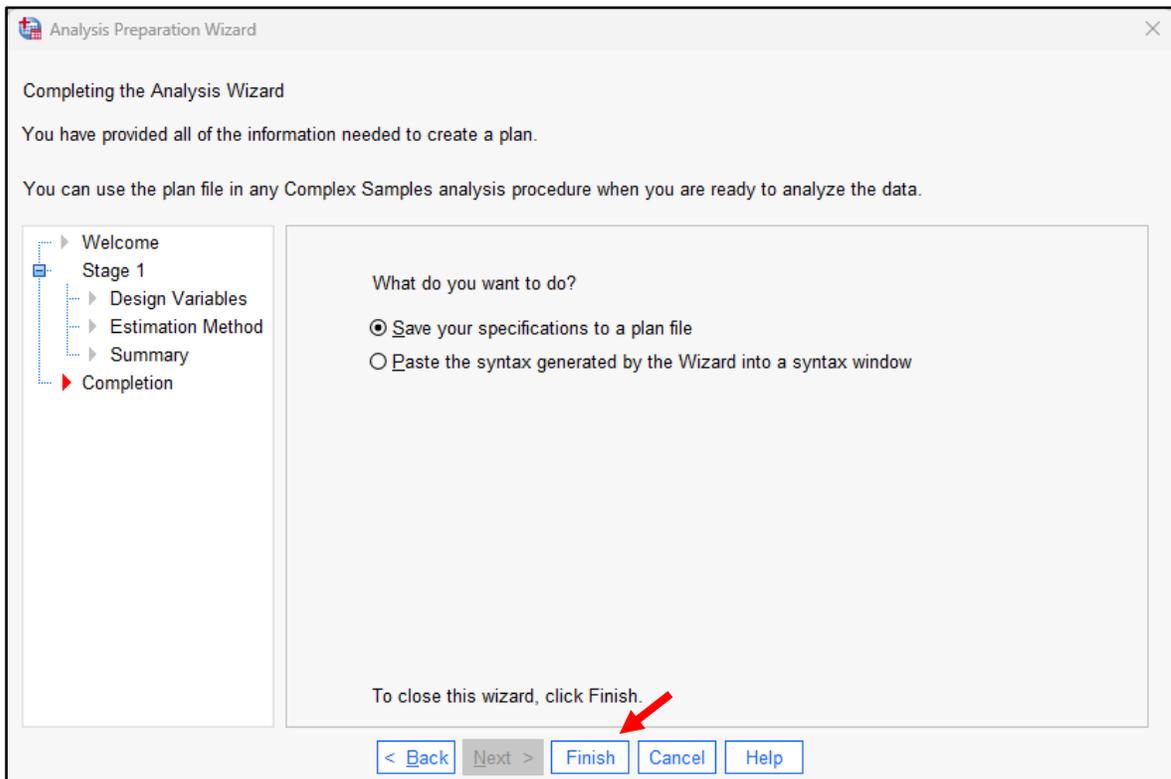
- WR (sampling with replacement)**
If you choose this option you will not be able to add additional stages. Any sample stages after the current stage will be ignored when the data are analyzed.
 Use finite population correction (FPC) when estimating variance under simple random sampling assumption
- Equal WOR (equal probability sampling without replacement)**
The next panel will ask you to specify inclusion probabilities or population sizes.
- Unequal WOR (unequal probability sampling without replacement)**
Joint probabilities will be required to analyze sample data. This option is available in stage 1 only.

< Back Next > Finish Cancel Help

Click Next.

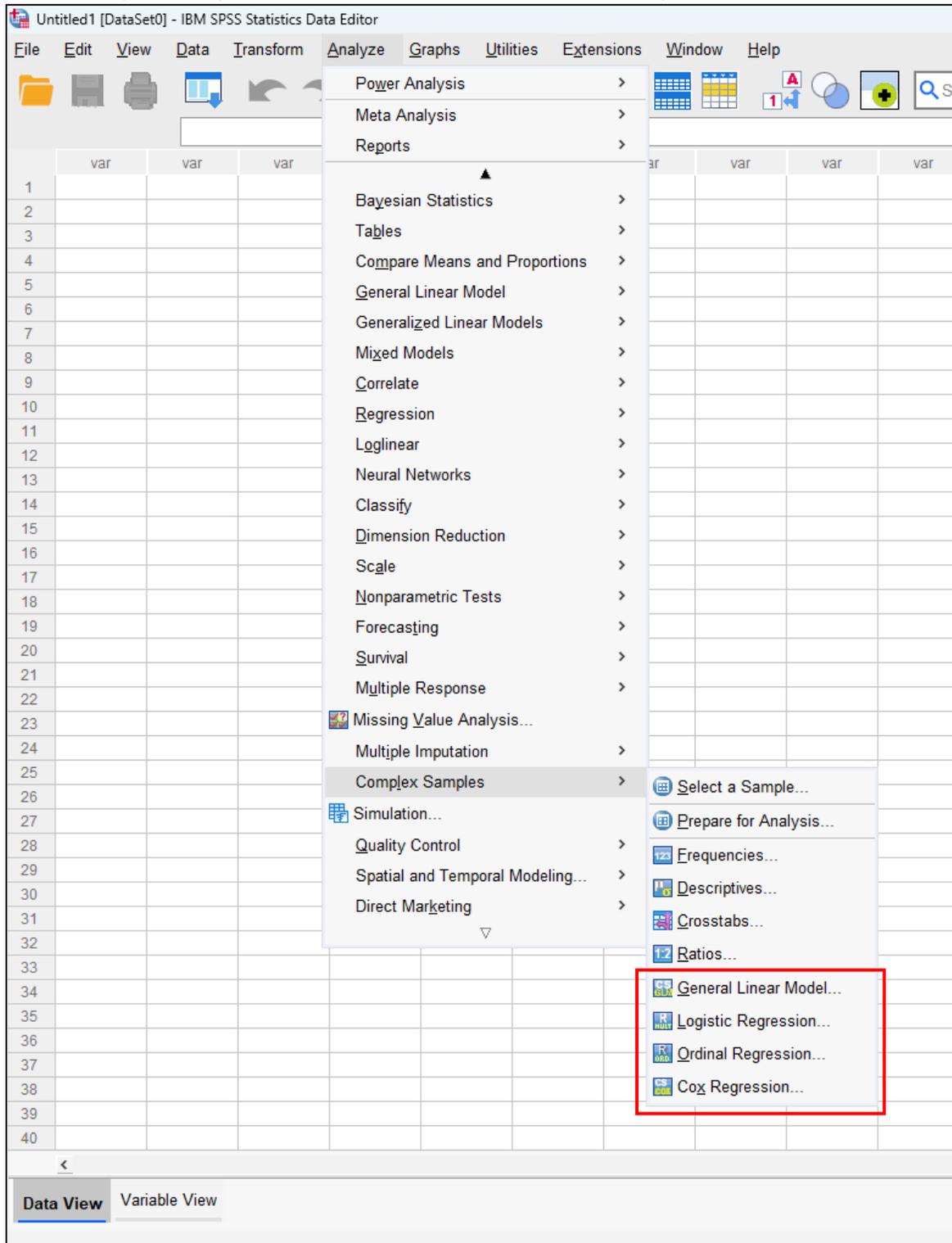


Click Finish.

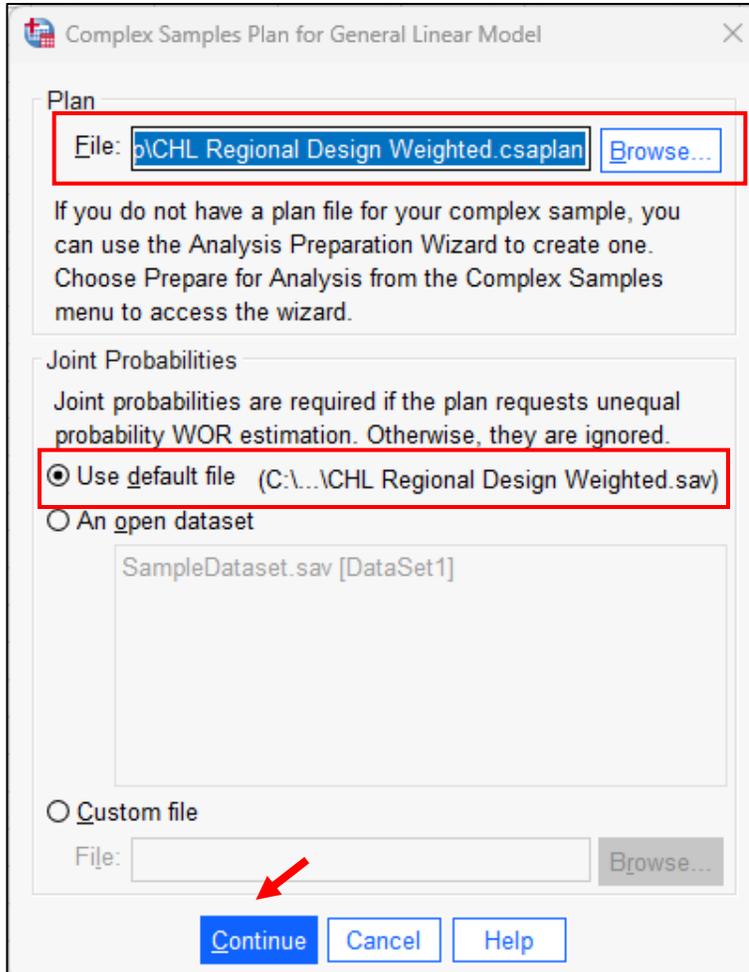


Once complex sampling design has been prepared, you can follow the steps below to continue to run your desired analyses:

Select Complex Samples → Choose statistical method for analysis.



Find your saved complex sample analysis plan (.csaplan) & Select Use default file.



Choose your variables of interest and run your model.

