Complex Sampling Weights
and
Preparing 2024 BRFSS Module Data for Analysis

August 2025

Overview

The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based telephone survey that collects data on a number of health outcomes, health-related risk behaviors, use of preventive services, and chronic conditions from noninstitutionalized adults who reside in each of the states and participating US territories. In 2011, BRFSS changed its data collection procedures, structure, and weighting methodology so the traditional landline telephone-based dataset could include data from participants using cell phones. The BRFSS uses a core set of questions and gives states the option to include multiple modules that focus on specific health issues. Not all optional modules are collected by all states, and states may opt to collect module data from subsets of their survey participants. Core questions are used during all interviews. Researchers using BRFSS data should conduct analyses with complex sampling procedures; they also should appropriately stratify and weight the data in their work. Potential bias resulting from selection probabilities and noncoverage among segments of the population can be reduced through weighting.

Researchers analyzing variables from the core-only section should use the variable _LLCPWT for weighting.

BRFSS has created this document to guide users analyzing variables from 2024 BRFSS optional modules and/or combinations of module and core variables. The 2024 BRFSS datasets include data that respondents provided by landline telephone or cell phone. Data users should note that newer weighting procedures are likely to affect trend lines when comparing BRFSS data collected up to 2010 with data from 2011 and afterward; because of these changes, users are advised NOT to make direct comparisons with pre-2011 data, and instead, begin new trend lines with that year.

Data users should become familiar with the information presented in this document prior to performing analyses. Information about the changes to 2024 BRFSS data collection is available in the 2024 Data Comparability Report.

Weighting BRFSS Core Data

BRFSS has grouped the 2024 data into four datasets:

- 2024 BRFSS Questionnaire data (Combined Landline Telephone and Cellular Telephone)

  And three versions

- 2024 BRFSS Combined Landline Telephone and Cellular Telephone Version 1
- 2024 BRFSS Combined Landline Telephone and Cellular Telephone Version 2
- 2024 BRFSS Combined Landline Telephone and Cellular Telephone Version 3

Researchers should understand that BRFSS data were collected by asking core and/or module questions on the annual questionnaire. Users, therefore, may need to use data from up to 4 different datasets, described below.

- Use the Combined Landline Telephone and Cell phone data if questions are exclusively from the core section ***or*** if questions are in common modules that were asked on both the landline survey and cellphone survey *(Note: See the section below,* ***Using BRFSS Data from Multiple Datasets****, for more about common modules.)*

- Use Version 1, Version 2, and/or Version 3 Combined Landline Telephone and Cell Phone datasets if states collected multiple version questionnaires.

In all cases, researchers conducting complex sampling analyses from the *core-only* section should use the variable **_LLCPWT** for weighting, **_STSTR** for stratification, and the variable **_PSU** for primary sampling unit, or clustering for short. The following table displays the description of the data, names of the datasets, and the variable names of the final weight.

Data description	Dataset name	Final weight variable name
Combined Landline Telephone and Cellular Telephone	LLCP2024	_LLCPWT
Combined Landline Telephone and Cellular Telephone Version 1	LLCP24V1	_LCPWTV1
Combined Landline Telephone and Cellular Telephone Version 2	LLCP24V2	_LCPWTV2
Combined Landline Telephone and Cellular Telephone Version 3	LLCP24V3	_LCPWTV3

The examples below demonstrate how to use _LLCPWT, _STSTR & _PSU when analyzing variable **HAVARTH4** (*Ever told you had some form of arthritis, rheumatoid arthritis, gout, lupus, or fibromyalgia*) and **PHYSHLTH** (*how many days during the past 30 days was your physical health not good?*) in the core section using the 2024 Louisiana dataset. Examples are given in both SAS and R.

Example in SAS:

```
LIBNAME IN '\BRFSS\';          /* Use LIBNAME to set the location of the BRFSS dataset */

DATA LA24;                     /* Create a temporary dataset for Louisiana */
 SET IN.LLCP2024 (WHERE=(_STATE = 22));
 RUN;

PROC SURVEYFREQ                /* HAVARTH4, Calculate percentage of population with arthritis */
  DATA = LA24;
  WEIGHT _LLCPWT;                    /* Use _LLCPWT for weighted analysis*/
  STRATA _STSTR;
  CLUSTER _PSU;
  TABLE HAVARTH4;
RUN;

PROC SURVEYMEANS               /* PHYSHLTH, Calculate average number of physical healthy days */
DATA =   LA24;
WEIGHT _LLCPWT;                /* Use _LLCPWT for weighted analysis */
STRATA _STSTR;
CLUSTER _PSU;
VAR PHYSHLTH;
RUN;
```

Example in R:

```
# Install the survey package if it is not already installed
install.packages("survey")

# Call the library for the current R session
library(survey)

# Read in BRFSS data
load("\\BRFSS\\BRFSS.rdata")

# Subset the data for Louisiana(22)
BRFSS <- BRFSS[BRFSS$state == 22, ]

# Set options for allowing a single observation per stratum
options(survey.lonely.psu = "adjust")

# Create survey design
brfssdsgn <- svydesign(
     id=~1,
    strata = ~ststr,
    weights = ~llcpwt,
    data = BRFSS)

# calculate average number of physical healthy days
svymean(~physhlth,
      brfssdsgn,
      na.rm = TRUE)

# calculate percent in each arthritis category
svymean(~factor(havarth4),
      brfssdsgn,
      na.rm = TRUE)
```

Using BRFSS Data from Multiple Datasets

Due to the complex nature of state-based data collection processes, users may have to create a dataset that fits their research needs. The following two examples illustrate how to prepare module data for analysis when states have collected module data in a variety of ways.

Prior to conducting analyses using optional module data, users should always review all relevant documents for the given data year, such as the Overview, Codebook, Comparability of Data, and Module Weights in particular to identify states that collected data of interest and determine which questionnaire version states used. Remember to check FIPS state codes in each dataset to avoid duplication.

Individual states may have chosen to use a number of optional modules, depending on each state's needs. Individual states also may have chosen to divide their samples and use different modules in the subsamples that were distinguished by version of the surveys. Modules that appeared in each version of a state's questionnaire are called common modules. Each state that collected common modules did so on all its versions of combined landline telephone and cell phone.

Example 1: Cognitive Decline

We will use the module **Cognitive Decline** to demonstrate how to combine and reweight data from multiple datasets for analysis. This module was selected because states collected the data either as a common module in the combined landline telephone and cell phone or by splitting the sample and offering the module on one or more versions of the data collection. For ease in using the 2024 optional modules, please refer to the two documents published with the annual dataset: BRFSS Modules Used by State and BRFSS Modules Used by Category. In addition to showing which states used which modules, these tables list names and descriptions of the datasets, questionnaire versions, and weight variables from the corresponding datasets. If users want to analyze variables from certain modules, these two documents are a good place to start. The following information has been copied from the two module documents to show segments of the tables:

Module weights for 2024

States by Module

Module	Description	Dataset	Data Weight	State(s)
Cognitive Decline	Combined Land Line and Cell Phone data	LLCP2024	_LLCPWT	Alabama, Alaska, Delaware, Florida, Idaho, Indiana, Maine, Missouri, New Hampshire, North Carolina, Oregon, Pennsylvania, Puerto Rico, Texas, Utah, Vermont, West Virginia, Wyoming
	Combined Land Line and Cell Phone data, version 1	LLCP24V1	_LCPWTV1	Arizona, **Michigan**
	Combined Land Line and Cell Phone data, version 2	LLCP24V2	_LCPWTV2	Iowa, Kansas, New York, Ohio
	Combined Land Line and Cell Phone data, version 3	LLCP24V3	_LCPWTV3	**Michigan**

Modules by State

State	Description	Dataset	Data weight	Module(s)
Michigan	Combined Land Line and Cell Phone data	LLCP2024	_LLCPWT	Arthritis, Cancer Survivorship: Type of Cancer, Childhood Asthma Prevalence, Family Planning, Firearm Safety, Pre-Diabetes, Prostate Cancer Screening, Random Child Selection, Sexual Orientation, Sugar Sweetened Beverages
	Combined Land Line and Cell Phone data, version 1	LLCP24V1	_LCPWTV1	**Cognitive Decline**, Other Tobacco Use
	Combined Land Line and Cell Phone data, version 2	LLCP24V2	_LCPWTV2	Social Determinants
	Combined Land Line and Cell Phone data, version 3	LLCP24V3	_LCPWTV3	Caregiver, **Cognitive Decline**, Other Tobacco Use, Social Determinants

1) The **Cognitive decline** module listed in the 2024 document **Modules by Category** shows:

Michigan (26) used version 1 and 3 of the Cognitive Decline module, Arizona (4) used version 1, Iowa (19), Kansas (20), New York (36) and Ohio (39) used version 2 of the module. The following 18 states used the common version of the Cognitive Decline module:

Alabama (1) Alaska (2) Delaware (10) Florida (12) Idaho (16) Indiana (18) Maine (23) Missouri (29) New Hampshire (33) North Carolina (37) Oregon (41) Pennsylvania (42) Texas (48) Utah (49) Vermont (50) West Virginia (54) Wyoming (56) and Puerto Rico (72)

Note: numbers in parentheses are the FIPS codes for states.

We will use 2024 data to explain how to reweight data for the Cognitive Decline module analysis.

a. Extract data from LLCP2024 for the 18 states that collected the common version of the module.

b. Extract data from LLCP24V1 for Arizona(4) and Michigan (26) that collected version 1 of the module

c. Extract data from LLCP24V2 for Iowa(19), Kansas(20), New York(36) and Ohio(39) that used version 2 of the module

d. Extract data from LLCP24V3 for *Michigan (26) that used version 3 of the module in addition to version 1.*

2) Obtain datasets:

- *To download the Combined Landline Telephone and Cellular Telephone dataset (LLCP2024):*

  - Go to the 2024 data webpage, choose the 2024 data. From there, read all of the 2024 documents including the overview, codebook, and the modules by category list. The dataset LLCP2024 is available in ASCII and SAS transport formats. Record layout, formats, and SAS code that converts an ASCII or SAS transportable file to SAS data-sets also are available in the **SAS Resources** section of the page.

- *To download the Combined Landline Telephone and Cellular Telephone Questionnaire Version 1 dataset:*

  - Go to the 2024 data webpage, then find The Combined Landline and Cellular Telephone Survey Multiple Questionnaire Version Data. There are 3 separate SAS datasets corresponding to Questionnaire Versions 1, 2, and 3. Choose Version 1: LLCP24V1.

- *To download the Combined Landline Telephone and Cellular Telephone Questionnaire Version 2 dataset:*

  - Follow the same steps above for Version 1, but choose Version 2: LLCP24V2.

- *To download the Combined Landline Telephone and Cellular Telephone Questionnaire Version 3 dataset:*

  - Follow the same steps above for Version 1, but choose Version 3: LLCP24V3

3) Generate a working dataset:

Create a new and uniform final weight variable from each version of the datasets downloaded in the steps above and combine data into one working dataset. This will require users to follow the following steps:

a. Keep all states that collected the module data in each version of the datasets,
b. Rename the corresponding weight variable to a new and consistent weight variable,
c. Combine all the datasets into one that contains the renamed weight variable.

Since Michigan used versions 1 and 3 of the Cognitive Decline module, we need to manually adjust the Weight variable for each version of its data following these steps:

Michigan:

1) Sample size for version 1, n1 = 3618
2) Sample size for version 3, n3 = 3711
3) Combined sample size for version 1 and version 3, (n1 + n3) = (3618 + 3711) = 7,329
4) the proportion for version 1 data is **P1**= n1/(n1+n3) = 3618/7329 ≈ **0.49**
   the proportion for version 3 data is **P3**= n3/(n1+n3) = 3711/7329 ≈ **0.51**

As we can see, the sample sizes for version 1 and version 3 of the 2024 Cognitive Decline module for Michigan are 3618 and 3711. We can use the proportions we calculated above (0.49 and 0.51) or simply divide the FINALWT by 2 since the two sample sizes are similar. If the sample sizes are quite different, however, we then should use the calculated proportions instead.

Example in SAS:

```
LIBNAME IN  '\BRFSS\';

DATA LLCP;     /* Extract data for states that used common version of the module*/
SET IN.LLCP2024 (WHERE=(_STATE IN (1 2 10 12 16 18 23 29 33 37 41 42 48 49 50 54 56 72)));
RENAME _LLCPWT=FINALWT;
RUN;

/* Extract data for states that used V1 of the module,
/* Adjust weights for MI  that used V1 & V3 of the module */
DATA LLCPV1;
   SET IN.LLCP24V1 (WHERE=(_STATE IN (4, 26)));
   IF _STATE = 26 THEN FINALWT= (_LCPWTV1*.49);   * _LCPWT1 times the calculated proportion;
   ELSE FINALWT= _LCPWTV1;
   DROP _LCPWTV1;
RUN;

/* Extract V2 data for IA, KS, NY & OH, Rename LCPWTV2 to FINALWT, no weight adjustment needed */
DATA LLCPV2;
   SET IN.LLCP24V2(WHERE=(_STATE IN (19 20 36 39)));
```

```
   FINALWT=_LCPWTV2;
   DROP _LCPWTV2;
RUN;

/* Extract V3 data for MI, and adjust weights because MI used both V1 & V3 of the module */
DATA LLCPV3;
   SET IN.LLCP24V3(WHERE=(_STATE IN (26)));
   FINALWT= (_LCPWTV3*.51);    * _LCPWTV3 times the calculated proportion;
   DROP _LCPWTV3;
RUN;

DATA COGNTV;     *combine all versions of the dataset into one;
   SET   LLCP   LLCPV1   LLCPV2   LLCPV3;
RUN;
```

The combined and reweighted dataset COGNTV is ready for analysis.

Example in R:

```
# Extract data from LLCP2024 for states that used the common version of the module# Note - R does not allow
variable names to begin with underscores, Such as _LLCPWT, _LCPWTV1, _LCPWTV2
# and _LCPWTV3, remove underscores before analysis

llcp <- llcp2024[llcp2024$state %in% c(1, 2, 10, 12, 16, 18, 23, 29, 33, 37, 41, 42, 48, 49, 50, 54, 56, 72), ]

llcp$finalwt <- llcp$llcpwt      # Rename weight variable LLCPWT to FINALWT for the main dataset

# Extract V1 data for AZ, rename weight var
azllcpv1 <- llcp24v1[llcp24v1$state %in% c(4), ]
azllcpv1$finalwt <- azllcpv1$lcpwtv1

# Extract V1 data for MI & Adjust weight
millcpv1 <- llCP24v1[llCP24v1$state %in% c(26), ]
millcpv1$finalwt <- millcpv1$lcpwtv1* (0.49)

#V1 - combine V1 datasets for AZ & MI
 llcpv1<- rbind(azllcpv1, millcpv1)

# Extract V2  data for IA, KS, NY & OH that used only version 2 of the module, rename lcpwtv2 to FINALWT
llcpv2 <- llCP24v2[llCP24v2$state %in% c(19,20,36,39), ]
llcpv2$finalwt <- llcpv2$lcpwtv2

# V3 – Extract Version3 data for MI - Note that MI used both V1 & V3 version of the module,
# weight variable LCPWTV3 need to be adjusted and renamed to FINALWT
llcpv3<- llCP24v3[llCP24v3$state %in% c(26), ]
llcpv3$finalwt <- llcpv3$lcpwtv3* (0.51)

# Combine all datasets for analysis
cogntv <- rbind(llcp, llcpv1, llcpv2, llcpv3)
```

The combined and reweighted dataset **cogntv** is ready for analysis.

Example 2: Family planning

The example below uses the **Family planning** Module to demonstrate how to combine and reweight data from multiple datasets.

1) The module **Family planning** listed in the document Modules by Category shows that a total of 25 states used it as an optional module in 2024.

- 20 states used the common version of this module; therefore, we need to extract data for these states from the dataset LLCP2024:

  Alaska (2) California (6) Connecticut (9) Delaware (10) Illinois (17) Indiana (18) Iowa (19) Louisiana (22) Michigan (26) Minnesota (27) Montana (30) New Jersey (34) New Mexico (35) North Carolina (37) Oregon (41) Rhode Island (44) Utah (49) Vermont (50) Virginia (51) Wyoming (56)

- ***Maryland** (24) and Massachusetts (25)* used **version 1** of the module, therefore, we need to extract data from LLCP24V1 for Maryland and Massachusetts.

- *Arizona (4) Kansas(20) and **Maryland** (24)* used **version 2** of the module and therefore, we need to extract data from LLCP24V2 for them.

- *New York (36)* used version 3 of the module, and we need to extract data from LLCP24V3 for New York.

  Note: Maryland used both version 1 & version 2 of the Family Planning module so the weight variable will need to be adjusted for that.

2) Obtain datasets:

- *To download the Combined Landline Telephone and Cellular Telephone dataset (LLCP2024):*

  - Go to the 2024 data webpage, and choose the 2024 dataset. From there, read all of the 2024 documents including the Overview, Codebook, and Module Weights. The dataset is available in ASCII and SAS transport formats. The record layout, formats, and SAS code that converts ASCII (or the SAS transportable file) to SAS datasets are also available in the **SAS Resources** section of the page.

- *To download the Combined Landline Telephone and Cellular Telephone Questionnaire Version 1 dataset:*

  - Go to the 2024 data webpage, select BRFSS the Combined Landline and Cellular Telephone Survey Multiple Questionnaire Version Data. There are 3 separate SAS datasets corresponding to Questionnaire Versions 1, 2, and 3. Choose Version 1: LLCP24V1.

- *To download the Combined Landline Telephone and Cellular Telephone Questionnaire Version 2 dataset:*

- Follow the same step above for Version 1, but instead, choose Version 2: LLCP24V2.

- To download the Combined Landline Telephone and Cellular Telephone Questionnaire Version 3 dataset:

  - Follow the same step above for Version 1, but, choose Version 3: LLCP24V3.

3) **Generate a working dataset**:
Create a new and uniform final weight variable from each of the datasets, then combine data into one working dataset. This step will require users to do the following:

a. Keep all states that collected the module data in each version of the datasets.
b. Rename the corresponding weight variable to a consistent weight variable and adjust the weight variable if any state uses more than 1 version of the module.
c. Combine all the datasets into one that contains the renamed weight variable. Details on how to adjust weight are shown below.

4) **Weight adjustment**:
Since Maryland used both V1 & V2 of the **Family Planning** module, the weight variable should be adjusted for the module data analysis.

Maryland:

1) Sample size for version 1, n1 = 4672
2) Sample size for version 2, n2 = 4744
3) Combined sample size for version 1 and version 2, (n1 + n2) = (4672 + 4744) = 9,416
4) the proportion for version 1 data is **P1**= n1/(n1+n2) = 4672/9416 ≈ **0.50**
   the proportion for version 2 data is **P2**= n2/(n1+n2) = 4744/9416 ≈ **0.50**

As we can see, the sample sizes for version 1 and version 2 of 2024 Maryland data are 4672 and 4744. The proportions we calculated above are roughly the same so we can simply divide the weight variable for each version by 2 to get the new FINALWT. If the sample sizes are quite different, however, we should use the calculated proportions instead.

Example code in SAS:

```
LIBNAME IN '\BRFSS\'; *specify the libname where BRFSS datasets are downloaded and stored;

DATA FAM;      * Extract data for states that used common version of the module;
SET IN.LLCP2024 (WHERE=(_STATE IN (2 6 9 10 17 18 19 22 26 27 30 34 35 37 41 44 49 50 5156)));
FINALWT=_LLCPWT;      *FINALWT, new weight variable for new datasets;
RUN;

DATA FAMV1;      *Extract V1 data MD & MA;
SET IN.LLCP24V1 (WHERE=(_STATE IN (24 25)));
IF _STATE = 24 THEN FINALWT = _LCPWTV1/2;
ELSE FINALWT=_LCPWTV1;
RUN;

DATA FAMV2;      *Extract V2 data for AZ, KS & MD;;
```

```
SET IN.LLCP24V2 (WHERE=(_STATE IN (4,20,24)));
IF _STATE = 24 THEN FINALWT = _LCPWTV2/2;
ELSE FINALWT = _LCPWTV2;
RUN;

DATA FAMV3;   *Extract V3 data FOR NY;
SET IN.LLCP24V3 (WHERE=(_STATE IN (36)));
FINALWT = _LCPWTV3;
RUN;

DATA FAMPLAN;    *Create combined dataset for analysis;
SET   FAM   FAMV1   FAMV2   FAMV3;
RUN;
```

The combined SAS dataset, FAMPLAN, is ready to be analyzed.

Example code in R:

```
# Begin with the main dataset LLCP2024 for states that used the common version of the Family Planning module
# Since R does not allow variable names to begin with underscore, such as _LLCPWT, _LCPWTV1, _LCPWTV2 and _LCPWTV3
# Remove underscores from the dataset before analysis

llcp <- llcp2024[llcp2024$state %in% c(2, 6, 9, 10, 17, 18, 19, 22, 26, 27, 30, 34, 35, 37, 41, 44, 49, 50, 51, 56 ),]
llcp$finalwt <- llcp$llcpwt

# Extract V1 data for MD
mdllcpv1[llcp24v1$state %in%  c(24),]
# Adjust weight variable for MD V1
mdllcpv1$finalwt <- llcpv1$lcpwtv1*0.5

# Extract V1 data for MA
mallcpv1[llcp24v1$state %in%  c(25),]
# Rename LCPWTV1 to FINALWT
mallcpv1$finalwt <-mallcpv1$lcpwtv1

#Combine V1 datasets for MD & MA
llcpv1 <- rbind(mdllcpv1, mallcpv1)

# Extract V2 data for MD
mdllcpv2[llcp24v2$state %in%  c(24),]
# Adjust weight variable for MD V2
mdllcpv2$finalwt <- mdllcpv2$lcpwtv1*0.5

# Extract V2 data for AZ, KS
azksllcpv2[llcp24v2$state %in%  c(4,20), ]
# Rename LCPWTV2 to FINALWT
azksllcpv2$finalwt <- azksllcpv2$lcpwtv2

#Combine V2 datasets for MD,AZ & KS
llcpv2 <- rbind(mdllcpv2, azksllcpv2)

# Extract V3 data for NY & rename LCPWTV3 to FINALWT
llcpv3[llcp24v3$state %in%  c(36), ])
llcpv3$finalwt <- llcpv3$lcpwtv3

# Combine all datasets and it's ready for analysis
famplan <- rbind(llcp, llcpv1, llcpv2, llcpv3)
```

Combining Multiple Years of BRFSS Data

There are times we need to analyze data from multiple years, such as when calculating the average prevalence of a variable in core sections for 2022 and 2024. In this case, we need to combine data from 2022 and 2024. The weight variable for 2022 and 2024 needs to be adjusted proportionally for the combined dataset. There are two ways to adjust the weight variable. The sample size from each year helps to determine how we should adjust them. Again, if the sample sizes are very similar, we can simply divide the original weight variable by 2. If they are different, we need to adjust the weight variable proportionally by following these steps:

1) Find out the sample sizes for 2022 and 2024
2) Sum up the total sample size for 2022 and 2024
3) Find the proportion for 2022 (Use the sample size for 2022 divided by the combined sample size for 2022 & 2024)
4) Find the proportion for 2024 (Use the sample size for 2024 divided by the combined sample size for 2022 & 2024)
5) The adjusted weight for 2022 is (_LLCPWT(2022) * proportion for 2022)
6) The adjusted weight for 2024 is (_LLCPWT(2024) * proportion for 2024)

The example below uses the 2022 and 2024 data from New York to demonstrate how to combine and reweight data from multiple years.

The sample size for 2022 New York data is 17,800 and 43,913 for 2024. Since these sample sizes are quite different, the weights need to be adjusted proportionally. Follow the steps listed above to adjust weights for New York for 2022 & 2024.

1) The sample size is 17,800 for 2022, and 43,913 for 2024
2) The combined sample size is 61,713
3) The proportion for 2022 is 17,800 / 61,713 ≈ .29
4) The proportion for 2024 is 43,913 / 61,713 ≈ .71;
5) The adjusted weight for 2022 is (_LLCPWT * **.29**)
6) The adjusted weight for 2024 is (_LLCPWT * **.71**)

Obtain datasets:

- *To download the 2022 Combined Landline Telephone and Cellular Telephone Questionnaire dataset (LLCP2022):*
  - Go to the 2022 data Webpage, and choose 2022 dataset. From there, read all of the 2022 documents including the Overview, Codebook, and the Modules by Category. The dataset is available in ASCII and SAS transport formats. A record layout, formats, and SAS code that converts an ASCII or SAS transportable file to SAS datasets also are available in **SAS Resources** section of the page.

- *To download the 2024 Combined Landline Telephone and Cellular Telephone Questionnaire dataset (LLCP2024):*
  - Go to the 2024 data Webpage, and choose the 2024 dataset and read all of the 2024 documents including the Overview, Codebook, and the Modules by Category. The

dataset is available in ASCII and SAS transport formats. Record layout, formats, and SAS code are also available in the **SAS Resources** section of the page.

- *Generate a new and uniform final weight variable (the adjusted weight variables) from each of the datasets and combine the data into one working dataset. This step will require users to follow these steps:*

  - Rename the corresponding weight variable to a consistent weight variable; and
  - Combine all datasets into one with the consistent weight variable.

  Again, users should note that new weighting procedures are likely to affect trend lines when comparing BRFSS data collected before and after 2011; because of the changes, users are advised NOT to make direct comparisons with pre-2011 data, and instead, to begin new trend lines with that year.

Example code in SAS:

```
Libname D22 "\BRFSS2022";
Libname D24 "\BRFSS2024";

DATA NY22;    /* Check sample size for each YEAR, 2022 & 2024 */
  SET D22.LLCP2022(WHERE=(_STATE=36));
  RUN;   /* There were 17800 observations read from the dataset D22.LLCP2022 for NY (36) */

DATA NY24;
  SET D24.LLCP2024(WHERE=(_STATE=36));
  RUN;   /*There were 43913 observations read from the dataset D24.LLCP2024 for NY (36) */

DATA NY2YR;      /* Create a final dataset with the adjusted weight variable (FINALWT) */
  SET NY22 (IN=A)
      NY24 (IN=B);
  IF A THEN FINALWT = _LLCPWT * .29;      * Use the calculated proportion to adjust _LLCPWT;
  IF B THEN FINALWT = _LLCPWT * .71;
  RUN;
```

The combined and reweighted dataset NY2YR is ready for analysis.

Example code in R:

```
# Subset 2021 and 2024 data for New York
# Since R does not allow variable names to begin with underscore, such as _LLCPWT, _LCPWTV1, _LCPWTV2 and
#_LCPWTV3, remove underscores from the dataset before analysis

NY22 <- llcp2022[llcp2022$state == 36,]
NY24 <- llcp2024[llcp2024$state == 36,]

# Count number of observations in each year
nrow(NY22)
nrow(NY24)

# Create a final weight based on proportion of observations in each year
NY22$finalwt <- NY22$llcpwt * (.29)
NY24$finalwt <- NY24$llcpwt * (.71)

# Combine the datasets and reweight
NY2YR <- rbind(NY22, NY24)
```

The combined and reweighted dataset NY2YR is ready for analysis.

~~~ END~~~