Health Information National Trends Survey (HINTS) 5 Cycle 3 Data: A How-To Guide for Using the New Data in your Research

Behavioral Research Program (BRP)
Division of Cancer Control and Population Sciences
To ask a question

- Submit questions using the Q&A or Chat Panel and select *All Panelists*
- You may need to activate the appropriate box using the floating navigation panel found on the center of your screen

- The webinar recording will be posted in a few weeks at [cancercontrol.cancer.gov/brpwebinars](cancercontrol.cancer.gov/brpwebinars)
Webinar Overview

Richard Moser, Ph.D.
Training Director and Research Methods Coordinator
Behavioral Research Program

Ashley Murray, Ph.D.
Cancer Research Training Award Fellow
Behavioral Research Program
Health Information National Trends Survey
(HINTS) 5 Cycle 3 Tutorial

Richard P. Moser, Ph.D.
Chief Methodologist, HINTS Program
Research Methods Coordinator, Behavioral Research Program
National Cancer Institute

Ashley B. Murray, Ph.D.
Cancer Research Training Award Fellow
Behavioral Research Program
National Cancer Institute

March 26, 2020
Rapidly Changing Communication Environment
Rapidly Changing Communication Environment
HINTS History & Timeline

- **HINTS I**
  - List-assisted RDD
  - RDD-C.A.T.I. + Web
  - N = 6369

- **HINTS II**
  - Dual Frame:
    - RDD - C.A.T.I.
    - Postal Sample
    - N = 7674

- **HINTS III**
  - Postal Frame

- **HINTS IV**
  - Postal Frame
    - 4-1: n=3959
    - 4-2: n=3185
    - 4-3: n=3677
    - 4-4: n=3738
    - FDA-1: n=3630
    - FDA-2: n=3677
    - NCI-Designated Cancer Centers Letter RFA: n=5438

- **HINTS V**
  - Postal Frame
    - 5-1: n=3594
    - 5-2: n=3504
    - 5-3: n=5438
    - 5-4: tbd

Regions:
- Delaware
- Puerto Rico
- Hualapai
- Hong Kong
- China
- Guam
- Appalachia
- NCI-Designated Cancer Centers Letter RFA
HINTS History & Timeline

HINTS I
- List-assisted RDD
  - C.A.T.I. N = 6369
- RDD-C.A.T.I. + Web N = 5586

HINTS II
- Dual Frame:
  - RDD - C.A.T.I.
  - Postal Sample N = 7674

HINTS III
- Postal Frame

HINTS IV
- Postal Frame
  - Data available now!

HINTS V
- 5-1 + FDA-2 n=5438
  - 5-2 n=3504
  - 5-3 n=3285
  - 5-4 tbd

Locations:
- Delaware
- Puerto Rico
- Hualapai
- Hong Kong
- Guam
- Appalachia
- China
- NCI-Designated Cancer Centers Letter RFA

Years:
- 2003
- 2004
- 2005
- 2006
- 2007
- 2008
- 2009
- 2010
- 2011
- 2012
- 2013
- 2014
- 2015
- 2016
- 2017
- 2018
- 2019
- 2020
**HINTS History & Timeline**

**HINTS I**
- List-assisted RDD
- RDD-C.A.T.I. + Web
- N = 6369

**HINTS II**
- RDD-C.A.T.I. + Web
- N = 5586

**HINTS III**
- Dual Frame:
  - RDD - C.A.T.I.
  - Postal Sample
  - N = 7674

**HINTS IV**
- Postal Frame

**HINTS V**
- Postal Frame
  - Data available now!

- Currently in the field; available fall/winter, 2020

- 5-1 + FDA-2
  - n = 5438
  - n = 3285

- 5-3
  - n = 3504

- NCI-Designated Cancer Centers Letter RFA

**Locations**
- Puerto Rico
- Delaware
- Guam
- Hualapai
- Hong Kong
- China
- Appalachia
- NCI-Designated Cancer Centers Letter RFA

**Timeline**
- 2003
- 2004
- 2005
- 2006
- 2007
- 2008
- 2009
- 2010
- 2011
- 2012
- 2013
- 2014
- 2015
- 2016
- 2017
- 2018
- 2019
- 2020
Access to:

- HINTS data and supporting documents
- Electronic codebook
- Tutorial
- Reports
- List of publications
- Briefs

https://hints.cancer.gov/
Overview of HINTS 5 Cycle 3

• Population: US Non-Institutionalized Adults (18+)

• Conducted January to April 2019

• Total N = 5,438
Web Pilot

- Test multi-mode survey
- Goal: Improve data quality
  - Increase response rates
  - Increase coverage
  - Reduce undesirable respondent behavior
    - Non-completion
    - Speeding
    - Straight lining
- Assess cost effectiveness

- Random assignment into 3 groups
  - Paper only
  - Web option
  - Web bonus
  - All get $2 incentive in mail
- Web groups could answer on internet or return by mail
- Prompting vs. no prompting (web)
- Considered as independent samples from same population
  - Weights available for each
## Web Pilot: Results

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample Size</th>
<th>Response Rate *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Only</td>
<td>3,372</td>
<td>30.2%</td>
</tr>
<tr>
<td>Web Option</td>
<td>986</td>
<td>29.6%</td>
</tr>
<tr>
<td>Web Bonus</td>
<td>1,080</td>
<td>31.5%</td>
</tr>
<tr>
<td>Total Sample</td>
<td>5,438</td>
<td>30.3%</td>
</tr>
</tbody>
</table>

*Not statistically different*
Health Information National Trends Survey 5 (HINTS 5)

Web Pilot Results Report
March 2020

Prepared for
National Cancer Institute
9609 Medical Center Drive
Bethesda, MD 20892-9760

Prepared by
Westat
1600 Research Boulevard
Rockville, MD 20850

See For More Information
Conducting Analyses
Statistics Overview

• HINTS can be analyzed using a variety of programs
• When you are interested in conducting inferential statistics (i.e. anything that involves calculating a p-value or confidence interval), it is important to consider which program you use
• A program should be able to:
  – Compute the correct variance estimates when analyzing survey data that employ a complex sampling method (e.g. HINTS)
  – Analyze data using jackknife replicate weights OR Taylor Linearization
Statistics Overview

- Code and results for SAS, SPSS, and STATA are provided in the Overview of the HINTS 5 Cycle 3 Survey and Data Analysis Recommendations Document
  - Descriptive analyses, Chi Square, Logistic Regression, & Linear Regression using both Jackknife replicates and Taylor Linearization in SAS and STATA (SPSS can only use Taylor Linearization; code & results provided)
  - SPSS analyses use Complex Samples module (add-on to Base SPSS)
  - We suggest using a program that can integrate Jackknife replicate weights when possible
- Tutorial will highlight both SAS and SPSS
Determining and Using Weights for Analyses

SAS and SPSS
Determining Weights to Use for HINTS 5 Cycle 3

Are variables of interest measured in the Web Pilot Report?

- No
  - Conduct Analyses to Assess for Group Differences (NWGT weights)
    - Pg 8

- Yes
  - Based on results from Web Pilot Report or your analyses, determine whether to utilize the total sample or only one group*

  - Total Sample
    - Yes
      - Were Group Differences Found?
        - Yes
          - Use the combined sample and control for group differences using the Treatment variable
            (NWGT weights & Treatment_H5C3)
            - Pg 10
        - No
          - Use the combined sample without controlling for group differences
            (TG_ALL_FINWT weights)
            - Pg 10
    - No
      - Using Only One Group
        - Without accounting for potential group differences (e.g., using the Paper Only sample)
          (TG1_FINWT weights)
          - Pg 10

* Only One Group
Are variables of interest measured in the Web Pilot Report?

No

Conduct Analyses to Assess for Group Differences (NWGT weights)

Pg 8

Yes

Based on analyses, do the total sample for only one gender...
Conducting Analyses to Assess for Group Differences in SAS

- It is strongly recommended that analysts first assess for possible group differences within their target variables.

Assessing for Group Differences with *Binary Outcomes* (with SEEKCANCERINFO as example):

```sas
data DATAFILENAME;
set DATAFILENAME;
*Set negative values to missing;
if SeekCancerInfo < 0 then SeekCancerInfo=-.;
run;
proc surveylogistic data=DATAFILENAME varmethod=jackknife;
weight nwgt0;
repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
class TREATMENT_H5C3;
model SeekCancerInfo = TREATMENT_H5C3;
run;
```

Assessing for Group Differences with *Continuous Data* (with GENERALHEALTH as example):

```sas
data DATAFILE;
set DATAFILE;
*Set negative values to missing;
if GeneralHealth < 0 then GeneralHealth=.;
run;
proc surveyreg data=DATAFILENAME varmethod=jackknife;
weight nwgt0;
repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
class TREATMENT_H5C3;
model GeneralHealth = TREATMENT_H5C3 /solution;
run;
```
Conducting Analyses to Assess for Group Differences in SAS

- It is strongly recommended that analysts first assess for possible group differences within their target variables.

Assessing for Group Differences with *Binary Outcomes* (with SEEKCANCERINFO as example):

```sas
data DATAFILENAME;
    set DATAFILENAME;
    *Set negative values to missing;
    if SeekCancerInfo < 0 then SeekCancerInfo=.;
run;
proc surveylogistic data=DATAFILENAME varestmethod=jackknife;
    weight nwgt0;
    repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
    class TREATMENT_H5C3;
    model SeekCancerInfo = TREATMENT_H5C3;
run;
```

Assessing for Group Differences with *Continuous Data* (with GENERALHEALTH as example):

```sas
data DATAFILE;
    set DATAFILE;
    *Set negative values to missing;
    if GeneralHealth < 0 then GeneralHealth=.;
run;
proc surveyreg data=DATAFILENAME varestmethod=jackknife;
    weight nwgt0;
    repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
    class TREATMENT_H5C3;
    model GeneralHealth = TREATMENT_H5C3 /solution;
run;
```
Conducting Analyses to Assess for Group Differences in SAS

- It is strongly recommended that analysts first assess for possible group differences within their target variables.

Assessing for Group Differences with *Binary Outcomes* (with SEEKCANCERINFO as example):
```sas
data DATAFILENAME;
    set DATAFILENAME;
    *Set negative values to missing;
    if SeekCancerInfo < 0 then SeekCancerInfo=.;
run;
proc surveylogistic data=DATAFILENAME varmethod=jackknife;
    weight nwgt0;
    repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
    class TREATMENT_H5C3;
    model SeekCancerInfo = TREATMENT_H5C3;
run;
```

Assessing for Group Differences with *Continuous Data* (with GENERALHEALTH as example):
```sas
data DATAFILE;
    set DATAFILE;
    *Set negative values to missing;
    if GeneralHealth < 0 then GeneralHealth=.;
run;
proc surveyreg data=DATAFILENAME varmethod=jackknife;
    weight nwgt0;
    repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
    class TREATMENT_H5C3;
    model GeneralHealth = TREATMENT_H5C3 /solution;
run;
```
Conducting Analyses to Assess for Group Differences in SAS

• It is strongly recommended that analysts first assess for possible group differences within their target variables.

Assessing for Group Differences with *Binary Outcomes* (with SEEKCANCERINFO as example):

```sas
data DATAFILENAME;
  set DATAFILENAME;
  *Set negative values to missing;
  if SeekCancerInfo < 0 then SeekCancerInfo=.;
run;
proc surveylogistic data=DATAFILENAME varmethod=jackknife;
  weight nwgt0;
  repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
  class TREATMENT_H5C3;
  model SeekCancerInfo = TREATMENT_H5C3;
run;
```

Assessing for Group Differences with *Continuous Data* (with GENERALHEALTH as example):

```sas
data DATAFILE;
  set DATAFILE;
  *Set negative values to missing;
  if GeneralHealth < 0 then GeneralHealth=.;
run;
proc surveyreg data=DATAFILENAME varmethod=jackknife;
  weight nwgt0;
  repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
  class TREATMENT_H5C3;
  model GeneralHealth = TREATMENT_H5C3 /solution;
run;
```
Conducting Analyses to Assess for Group Differences in SAS

- It is strongly recommended that analysts first assess for possible group differences within their target variables.

Assessing for Group Differences with **Binary Outcomes** (with SEEKCANCERINFO as example):
```sas
data DATAFILENAME;
set DATAFILENAME;
*Set negative values to missing;
if SeekCancerInfo < 0 then SeekCancerInfo=.
run;
proc surveylogistic data=DATAFILENAME varmethod=jackknife;
weight nwgt0;
repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
class TREATMENT_H5C3;
model SeekCancerInfo = TREATMENT_H5C3;
run;
```

- Assessing for Group Differences with **Continuous Data** (with GENERALHEALTH as example):
```sas
data DATAFILE;
set DATAFILE;
*Set negative values to missing;
if GeneralHealth < 0 then GeneralHealth=.
run;
proc surveyreg data=DATAFILENAME varmethod=jackknife;
weight nwgt0;
repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
class TREATMENT_H5C3;
model GeneralHealth = TREATMENT_H5C3 /solution;
run;
```
Conducting Analyses to Assess for Group Differences in SAS

- It is strongly recommended that analysts first assess for possible group differences within their target variables.

Assessing for Group Differences with *Binary Outcomes* (with SEEKCANCERINFO as example):
```sas
data DATAFILENAME;
  set DATAFILENAME;
  *Set negative values to missing;
  if SeekCancerInfo < 0 then SeekCancerInfo=.;
run;
proc surveylogistic data=DATAFILENAME varmethod=jackknife;
  weight nwgt0;
  repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
  class TREATMENT_H5C3;
  model SeekCancerInfo = TREATMENT_H5C3;
run;
```

Assessing for Group Differences with *Continuous Data* (with GENERALHEALTH as example):
```sas
data DATAFILE;
  set DATAFILE;
  *Set negative values to missing;
  if GeneralHealth < 0 then GeneralHealth=.;
run;
proc surveyreg data=DATAFILENAME varmethod=jackknife;
  weight nwgt0;
  repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
  class TREATMENT_H5C3;
  model GeneralHealth = TREATMENT_H5C3 /solution;
run;
```
SPSS Complex Samples
SPSS Complex Samples

Welcome to the Analysis Preparation Wizard

The Analysis Preparation Wizard helps you describe your complex sample and choose an estimation method. You will be asked to provide sample weights and other information needed for accurate estimation of standard errors.

Your selections will be saved to a plan file that you can use in any of the analysis procedures in the Complex Samples Option.

What would you like to do?

- Create a plan file
  - Choose this option if you have sample data but have not created a plan file.

- Edit a plan file
  - Choose this option if you want to add, remove, or modify stages of an existing plan.

If you already have a plan file you can skip the Analysis Preparation Wizard and go directly to any of the analysis procedures in the Complex Samples Option to analyze your sample.
SPSS Complex Samples
Controlling or Assessing for Group Differences

SPSS Complex Samples
Controlling or Assessing for Group Differences

Full Sample (No Group Differences)
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.
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.
Conducting Analyses to Assess for Group Differences in SPSS
Conducting Analyses to Assess for Group Differences in SPSS
Conducting Analyses to Assess for Group Differences in SPSS
Conducting Analyses to Assess for Group Differences in SPSS
Conducting Analyses to Assess for Group Differences in SPSS
Determining Weights to Use for HINTS 5 Cycle 3

Conduct Analyses to Assess for Group Differences (NWGT weights)

Based on results from Web Pilot Report or your analyses, determine whether to utilize the total sample or only one group*

- Only One Group
  - Total Sample
    - Were Group Differences Found?
      - Yes
        - Using Only One Group
          - Without accounting for potential group differences (e.g. using the Paper Only sample) (TC1, FINWT weights)
      - No
Determining Weights to Use for HINTS 5 Cycle 3

Based on results from Web Pilot Report or your analyses, determine whether to utilize the total sample or only one group*

Only One Group

Using Only One Group
Without accounting for potential group differences (e.g. using the Paper Only sample)
(TG1_FINWT weights)
Pg 10
Conducting Analyses Using Only One Group

```sas
proc surveylogistic data=DATAFILENAME varmethod=jackknife;
  weight tgl_finwt0;
  repweights tgl_finwt1-tgl_finwt50 /df=49 jkcoefs=.98;
  *Predictor# variables in model statement are placeholders to substitute with your desired predictors;
  model SeekCancerInfo = predictor1 predictor2 predictor3 predictor4...;
run;
```

Note: example code above is for the paper-only sample (“TG1”). Weight and repweight statements may be replaced with the “TG2” weights for the web option sample or “TG3” weights for the web bonus sample.
Determining Weights to Use for HINTS 5 Cycle 3

Conduct Analyses to Assess for Group Differences (NWGT weights) Pg 8

Based on results from Web Pilot Report or your analyses, determine whether to utilize the total sample or only one group*

Only One Group

Total Sample

Using Only One Group
Without accounting for potential group differences (e.g. using the Paper Only sample) (TC1, FINWT weights)

Were Group Differences Found?

Yes

No
Were Group Differences Found?

- Yes: Use the combined sample and control for group differences using the Treatment variable (NWGT weights & Treatment_H5C3) Pg 10
- No: Use the combined sample without controlling for group differences (TG_ALL_FINWT weights) Pg 10

Total Sample

*only one group*
Conducting Analyses Combined Sample, Controlling For Group Differences

• **Logistic Regression**
  
  ```
  proc surveylogistic data=DATAFILENAME varmethod=jackknife;
      weight nwgt0;
      repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
   *Predictor# variables in model statement are placeholders to substitute with your desired predictors;
      model SeekCancerInfo = TREATMENT_H5C3 predictor1 predictor2 predictor3 predictor4...;
  
  run;
  ```

• **Linear Regression**
  
  ```
  proc surveyreg data=DATAFILENAME varmethod=jackknife;
      weight nwgt0;
      repweights nwgt1-nwgt150 /df=147 jkcoefs=.98;
   *Predictor# variables in model statement are placeholders to substitute with your desired predictors;
      model GeneralHealth = TREATMENT_H5C3 predictor1 predictor2 predictor3 predictor4... /solution;
  
  run;
  ```
Conducting Analyses
Combined Sample, Controlling For Group Differences
Conducting Analyses Combined Sample, Controlling For Group Differences
Were Group Differences Found?

Yes:
Use the combined sample and control for group differences using the Treatment variable (NWGT weights & Treatment_H5C3)
Pg 10

No:
Use the combined sample without controlling for group differences (TG_ALL_FINWT weights)
Pg 10
Conducting Analyses on the Combined Sample, Without Controlling for Group Differences

If no group differences are found, it is suggested that analysts use the combined sample to increase statistical power

- **Frequency Table and Chi-Square Test**
  ```
  proc surveyfreq data = hints5cycle3 varmethod = jackknife;
  weight TG_all_FINWT0;
  repweights TG_all_FINWT1-TG_all_FINWT50 / df = 49 jkcoefs = 0.98;
  tables edu*gender / row col wchisq;
  run;
  ```

- **Logistic Regression**
  ```
  proc surveylogistic data= hints5cycle3 varmethod=jackknife;
  weight TG_all_FINWT0;
  repweights TG_all_FINWT1-TG_all_FINWT50 / df=49 jkcoefs=0.98;
  class edu (ref="Less than high school")
  gender (ref="Male")/param=REF;
  model seekcancerinfo (descending) = gender edu /tech=newton xconv=1e-8 CLPARM EXPB;
  run;
  ```

- **Linear Regression**
  ```
  proc surveyreg data= hints5cycle3 varmethod=jackknife;
  weight TG_all_FINWT0;
  repweights TG_all_FINWT1-TG_all_FINWT50 / df=49 jkcoefs=0.98;
  class edu (ref="Less than high school") gender (ref="Male");
  model generalhealth = edu gender /solution;
  run;
  ```
Conducting Analyses on the Combined Sample, Without Controlling for Group Differences
Combining Cycles
Merging HINTS Survey Iterations

- Analysts should first assess for group differences in the HINTS 5, Cycle 3 data on variables of interest
- Sample code is created assuming there are no differences between groups in HINTS 5 Cycle 3
- If group differences are found:
  - Create a new variable in both data files that would allow the analyst to differentiate between the 4 groups (H5C2, H5C3 Paper Only, H5C3 Web Option, and H5C3 Web Bonus groups)
  - Use the Rizzo, et al., (2008) method to create 200 replicate weights
Merging HINTS Survey Iterations Using SAS

/*FIRST CREATE THE FORMAT FOR THE SURVEY VARIABLE*/
proc format;
  value survey
    1="HINTS 5 CYCLE 2"
    2="HINTS 5 CYCLE 3"
  ;
run;

/******************************************/
/*CREATE TWO SEPARATE TEMPORARY DATA FILES THAT CONTAIN THE NEW 'SURVEY' VARIABLE.*/
/*PUT NAME OF LIBRARY WHERE HINTS 5 CYCLE 2 FORMATS ARE STORED*/;
options fmtsearch=(LibH5C2);

data tempHINTS5CYCLE2;
  /*PUT NAME OF LIBRARY AND NAME OF EXISTING HINTS 5 CYCLE 3 DATA FILE*/
  set LibH5C2.DataH5C2;
    survey=1;
    format survey survey.;
run;

/* PUT NAME OF LIBRARY WHERE HINTS 5 CYCLE 3 FORMATS ARE STORED*/
options fmtsearch=(hints5c3);

data tempHINTS5CYCLE3;
  /*PUT NAME OF LIBRARY AND NAME OF EXISTING HINTS 5 CYCLE 3 DATA FILE*/
  set hints5c3.hints5cycle3_formatted;
    survey=2;
    format survey survey.;
run;
Merging HINTS Survey Iterations Using SAS

/*FIRST CREATE THE FORMAT FOR THE SURVEY VARIABLE*/
proc format;
  value survey
    1 = "HINTS 5 CYCLE 2"
    2 = "HINTS 5 CYCLE 3"
  ;
run;

/**************************************************************************/
/*CREATE TWO SEPARATE TEMPORARY DATA FILES THAT CONTAIN THE NEW ‘SURVEY’*/
/*VARIABLE.*/

/**************************************************************************/
/*PUT NAME OF LIBRARY WHERE HINTS 5 CYCLE 2 FORMATS ARE STORED*/;
options fmtsearch=(LibH5C2);

data tempHINTS5CYCLE2;
  /*PUT NAME OF LIBRARY AND NAME OF EXISTING HINTS 5 CYCLE 3 DATA FILE*/
  set LibH5C2.DataH5C2;
    survey=1;
    format survey survey.;
run;

/* PUT NAME OF LIBRARY WHERE HINTS 5 CYCLE 3 FORMATS ARE STORED*/
options fmtsearch=(hints5c3);

data tempHINTS5CYCLE3;
  /*PUT NAME OF LIBRARY AND NAME OF EXISTING HINTS 5 CYCLE 3 DATA FILE*/
  set hints5c3.hints5cycle3formatted;
    survey=2;
    format survey survey.;
run;
Merging HINTS Survey Iterations Using SAS

/*FIRST CREATE THE FORMAT FOR THE SURVEY VARIABLE*/

proc format;
   value survey
      1 = "HINTS 5 CYCLE 2"
      2 = "HINTS 5 CYCLE 3"
   ;
run;

/******************************************************************************/

/*CREATE TWO SEPARATE TEMPORARY DATA FILES THAT CONTAIN THE NEW ‘SURVEY’ VARIABLE.*/

/*PUT NAME OF LIBRARY WHERE HINTS 5 CYCLE 2 FORMATS ARE STORED*/;
options fmtsearch=(LibH5C2);

data tempHINTS5CYCLE2;
   /*PUT NAME OF LIBRARY AND NAME OF EXISTING HINTS 5 CYCLE 3 DATA FILE*/
   set LibH5C2.DataH5C2;

      survey=1;
      format survey survey. ;
run;

/* PUT NAME OF LIBRARY WHERE HINTS 5 CYCLE 3 FORMATS ARE STORED*/;
options fmtsearch=(hints5c3);

data tempHINTS5CYCLE3;
   /*PUT NAME OF LIBRARY AND NAME OF EXISTING HINTS 5 CYCLE 3 DATA FILE*/
   set hints5c3.hints5cycle3_formatted;
      survey=2;
      format survey survey. ;
run;
**SAS Code to Set Up Final and Replicate Weights for the Replicate Variance Estimation Method**

/*THIS CODE MERGES THE TWO TEMPORARY DATA SETS CREATED ABOVE. IT ALSO CREATES ONE FINAL SAMPLE WEIGHT (Merged_NWGT0) AND 100 REPLICATE WEIGHTS (Merged_NWGT1 THRU Merged_NWGT100)*/

```sas
data mergeHINTS5C2_HINTS5C3;
  set tempHINTS5CYCLE2 tempHINTS5CYCLE3;
  /*Create Replicate Weights for trend tests*/
  **Replicate Weights;**
  array hints52wgts [50] person_finwt1-person_finwt50;
  array hints53wgts [50] TG_all_finwt1-TG_all_finwt50;
  array Merged_NWgt [100] Merged_NWGT1-Merged_NWGT100;

  **Adjust Final And Replicate Weights;**
  if survey eq 1 then do i=1 to 50; *HINTS 5 CYCLE 2;
    Merged_NWGT0=person_finwt0;
    Merged_NWgt[i]=hints52wgts[i];
    Merged_NWgt[50+i]=person_finwt0;
  end;

  else if survey eq 2 then do i=1 to 50; *HINTS 5 CYCLE 3;
    Merged_NWGT0=TG_all_finwt0;
    Merged_NWgt[i]=TG_all_finwt0;
    Merged_NWgt[50+i]=hints53wgts[i];
  end;
run;
```
Merging HINTS Survey Iterations Using SAS

*SAS Code to Set Up Final and Replicate Weights for the Replicate Variance Estimation Method*

/*THIS CODE MERGES THE TWO TEMPORARY DATA SETS CREATED ABOVE. IT ALSO CREATES ONE FINAL SAMPLE WEIGHT (Merged_NWGT0) AND 100 REPLICATE WEIGHTS (Merged_NWGT1 THRU Merged_NWGT100)*/

data mergeHINTS5C2_HINTS5C3;
  set tempHINTS5CYCLE2 tempHINTS5CYCLE3;
  /*Create Replicate Weights for trend tests*/
  **Replicate Weights;**
  array hints52wgts [50] person_finwt1-person_finwt50;
  array hints53wgts [50] TG_all_finwt1-TG_all_finwt50;
  array Merged_NWgt [100] Merged_NWGT1-Merged_NWGT100;

  **Adjust Final And Replicate Weights;**
  if survey eq 1 then do i=1 to 50; *HINTS 5 CYCLE 2;
    Merged_NWGT0=person_finwt0;
    Merged_NWgt[i]=hints52wgts[i];
    Merged_NWgt[50+i]=person_finwt0;
  end;

  else if survey eq 2 then do i=1 to 50; *HINTS 5 CYCLE 3;
    Merged_NWGT0=TG_all_finwt0;
    Merged_NWgt[i]=TG_all_finwt0;
    Merged_NWgt[50+i]=hints53wgts[i];
  end;
run;
Merging HINTS Survey Iterations Using SAS

**SAS Code to Set Up Final and Replicate Weights for the Replicate Variance Estimation Method**

/* THIS CODE MERGES THE TWO TEMPORARY DATA SETS CREATED ABOVE. IT ALSO CREATES ONE FINAL SAMPLE WEIGHT (Merged_NWGT0) AND 100 REPLICATE WEIGHTS (Merged_NWGT1 THRU Merged_NWGT100) */

data mergeHINTS5C2_HINTS5C3;
  set tempHINTS5CYCLE2 tempHINTS5CYCLE3;

/* Create Replicate Weights for trend tests */
**Replicate Weights;**
array hints52wgts [50] person_finwt1-person_finwt50;
array hints53wgts [50] TG_all_finwt1-TG_all_finwt50;
array Merged_NWgt [100] Merged_NWGT1-Merged_NWGT100;

**Adjust Final And Replicate Weights;**
if survey eq 1 then do i=1 to 50; *HINTS 5 CYCLE 2;
  Merged_NWGT0=person_finwt0;
  Merged_NWgt[i]=hints52wgts[i];
  Merged_NWgt[50+i]=person_finwt0;
end;

else if survey eq 2 then do i=1 to 50; *HINTS 5 CYCLE 3;
  Merged_NWGT0=TG_all_finwt0;
  Merged_NWgt[i]=TG_all_finwt0;
  Merged_NWgt[50+i]=hints53wgts[i];
end;
run;
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav'
/COMPRESSED.
DATASET NAME MERGED_DATA.

DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=TG_all_FINWT0.
COMPUTE Survey=2.
EXECUTE.
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav' /COMPRESSED.
DATASET NAME MERGED_DATA.

DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=FINWT0.
COMPUTE Survey=2.
EXECUTE.
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav' /COMPRESSED.
DATASET NAME MERGED_DATA.
DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=TG_all_FINWT0.
COMPUTE Survey=2.
EXECUTE.
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav' /COMPRESSED.
DATASET NAME MERGED_DATA.

DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=TG_all_FINWT0.
COMPUTE Survey=2.
EXECUTE.
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav' /COMPRESSED.
DATASET NAME MERGED_DATA.
DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=TG_all_FINWT0.
COMPUTE Survey=2.
EXECUTE.

Save as new file
Rename Dataset
Rename Final Sample Weight
Create a variable that allows us to distinguish between participants from each cycle
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav'
/COMPRESSED.
DATASET NAME MERGED_DATA.

DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=TG_all_FINWT0.
COMPUTE Survey=2.
EXECUTE.

**below, you should insert the filepath for your HINTS 5 Cycle 2 data**.

GET
  FILE='H:\HINTS\5 Cycle 2\HINTS-5_Cycle2_SPSS\hints5_cycle2_public.sav'.
ALTER TYPE ALL(A=AMIN).
DATASET NAME H5C2 WINDOW=FRONT.
COMPUTE MERGED_FINWT0=Person_FINWT0.
COMPUTE Survey=1.
EXECUTE.
Merging HINTS Survey Iterations Using SPSS

SAVE OUTFILE='H:\HINTS\5 Cycle 3\SPSS\MERGED_H5C3_H5C2.sav'
/COMPRESSED.
DATASET NAME MERGED_DATA.

DATASET ACTIVATE MERGED_DATA.
COMPUTE MERGED_FINWT0=TG_all_FINWT0.
COMPUTE Survey=2.
EXECUTE.

**below, you should insert the filepath for your HINTS 5 Cycle 2 data**.
GET
FILE='H:\HINTS\5 Cycle 2\HINTS-5_Cycle2_SPSS\hints5_cycle2_public.sav'.
ALTER TYPE ALL(A=AMIN).
DATASET NAME H5C2 WINDOW=FRONT.
COMPUTE MERGED_FINWT0=Person_FINWT0.
COMPUTE Survey=1.
EXECUTE.

Open other cycle data
Rename Dataset
Rename Final Sample Weight
Create a variable that allows us to distinguish between participants from each cycle
Merging HINTS Survey Iterations Using SPSS
Merging HINTS Survey Iterations Using SPSS

Add Cases to hints5_cycle3_public.sav[DataSet1]

Select a dataset from the list of open datasets or from a file to merge with the active dataset

- An open dataset
  - hints5_cycle2_public.sav[DataSet3]

- An external SPSS Statistics data file
  - Browse...

Non-SPSS Statistics data files must be opened in SPSS Statistics before they can be used as part of a merge.

Continue  Cancel  Help
OVERVIEW OF THE HINTS 5 CYCLE 3 SURVEY AND DATA ANALYSIS RECOMMENDATIONS

January 2020
Questions?
Back Pocket Slides
SAS Taylor Linearization

- Frequencies and Chi Square, Combined Sample (No Differences Found)

```
proc surveyfreq data = hints5cycle3
  varmethod = TAYLOR;
  strata VAR_STRATUM;
  cluster VAR_CLUSTER;
  weight TG_all_finwt0;
  tables edu*gender / row col wchisq;
run;
```
Questions?