

Introduction to the National Longitudinal Study of Adolescent to Adult Health (Add Health)

Hsueh-Sheng Wu

Center for Family and Demographic Research

January 23, 2023

BGSU

 Center for Family and Demographic Research

Outline

- Introduction
- What is special about Add Health?
- Survey design
- Subject areas
- Add Health Codebook Explorer
- Data files
- File location
- Unit of analysis
- Analytic tips
- Studies using Add Health data
- Help with Add Health analyses
- Conclusions

Introduction

- National Longitudinal Study of Adolescent Health (Add Health) is a study that the Carolina Population Center at the University of North Carolina-Chapel Hill has conducted to follow a nationally representative sample of adolescents in grades 7-12 since 1994.
- These adolescents were first interviewed in 1994-1995 (Wave I) and followed up in 1996 (Wave II), 2001-2002 (Wave III), 2007-2008 (Wave IV) and finally 2016-2019 (Wave V).
- Add Health also has supplemental education data. In fall 2001, the Population Research Center at the University of Texas-Austin collected data that supplement Add Health. These supplementary data focus on (1) educational achievement, (2) course taking patterns, (3) curricular exposure, and (4) educational contexts of Add Health respondents at Wave III.
- In addition to original Add Health study, three satellite projects have been initiated, including (1) the Add Health Parent Study (Parents 2015-2017), (2) Sexual Orientation/Gender Identity, Socioeconomic Status, and Health across the Life Course (SOGI-SES)", and (3) the Omics project that examines the relations between biological factors (i.e., gene, proteins, molecules) and various outcomes.

What Is Special about Add Health?

- Add Health is an ongoing study that has the largest representative sample of adolescents and collects the most comprehensive data of these adolescents for over 20 years.
- Add Health started with 20,745 adolescents at wave 1 (1994-1995) and more than 20 years later, it still have 12,300 adolescents at Wave V (2016-2019).
- Add Health collected comprehensive information on biological and psychological developments of adolescents and the social contexts, such as home, friends, intimate relationships, schools, and neighborhoods.
- Because of its sample size, scopes of data, and panel design, Add Health is invaluable for examining how social, psychological, and biological factors influence an individual's life through adolescence, young adulthood, and later adulthood.

Survey Design

- Add Health used a stratified two-stage sampling method:
 - The sampling frame is stratified by region, urbanization, school size, school type, and race composition
 - 80 high schools and 52 middle schools were selected with an unequal probability at the first stage
 - 90,000 students were selected to fill out in-school Add Health questionnaire, and 20,745 of them fill out in-home questionnaire
- Add Health oversampled twins and siblings of twins; non-related adolescents residing together; disabled minority students; Blacks from well-educated families; and minority students who are Chinese, Cubans, and Puerto Ricans.
- Data were collected with Computer-Assisted Personal Interviewing (CAPI) and questionnaire.

Survey Design (Cont.)

Supplemental Education Data:

- The sample consists of all respondents at Wave III of Add Health.
- The study collected high school transcripts and other data from high schools that Add Health respondents last attended.
- The data were collected from 130 Add Health high schools and 1,400 additional high schools.
- Education data were collected for approximately 12,000 respondents, which is about 80% of Add Health respondents at Wave III.

Survey Design (Cont.)

Tab 1. The Life Segment covered by Wave V of Add Health (N=12,300)

Age at W1	Adolescence							Young Adulthood							Later Adulthood										Midlife	Total												
	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34			35	36	37	38	39	40	41	42	43	44	45-64	
11	1	2	2				3	3					4	4								5	5															6
12		1,2	2	2			3	3					4	4	4							5	5	5	5												321	
13			1,2	2	2		3	3	3				4	4	4	4						5	5	5	5	5											1,421	
14				1,2	2	2		3	3	3			4	4	4	4						5	5	5	5	5	5										1,735	
15					1,2	2	2		3	3	3			4	4	4	4					5	5	5	5	5	5										2,152	
16						1,2	2	2		3	3	3			4	4	4	4				5	5	5	5	5	5										2,362	
17							1,2	2		2		3	3	3			4	4	4	4			5	5	5	5	5	5									2,300	
18								1,2		2	2		3	3	3			4	4	4	4			5	5	5	5	5									1,693	
19									1,2	2	2		3	3		3		4	4	4	4			5	5	5	5	5									262	
20										1,2					3	3		4	4	4							5	5	5								41	
21											1,2					3	3		4	4							5	5	5	5							7	

Subject Areas

- Add Health covers many interesting subject areas.
- Some of the areas have been covered at each wave, whereas others are covered at only certain wave or waves.
- The list of sections in the codebooks at each wave shows what subject areas are covered at each wave.
- When a subject area is covered across waves, it does not necessarily mean all questions about the subject area are asked across waves. So, it is very important to check at which waves a certain question is asked.

Add Health Codebook Explorer

Home » Documentation » Codebook Explorer

Add Health Codebook Explorer (ACE)

Welcome to the Add Health Codebook Explorer (ACE). This tool allows you to view questions asked of respondents, parents, and interviewers during the In-School and In-Home interviews in Waves I through IV.

You can browse by topic or search for questions by variable name, keyword, or phrase. The questions are organized by topic, subtopic, and variable. Collections of variables were constructed specifically for this site to show similar questions asked across several waves of data collection and do not represent grouping for research purposes. The questions included in ACE are from the In-School and In-Home interviews. There are many other data sets available; please visit the [Codebooks page](#) for a complete list of available data, including contextual and biological data.

Please refer to the [ACE Instructions](#) before searching and browsing to maximize your results.

Search

Field matches **Search Text** **Find**

Question or Resp contains

Browse

[Topics](#)

[Instruments](#)

Add Health Codebook Explorer (Cont.)

The screenshot shows the Add Health Codebook Explorer website. The browser address bar displays the URL: <https://addhealth.cpc.unc.edu/documentation/codebook-explorer/>. The website header includes the Add Health logo and navigation links for About, Documentation, Data, and Publications. The main content area features a search interface with the following components:

- Search** section containing:
 - Field**: A dropdown menu currently set to "Question or Resp".
 - matches**: A dropdown menu currently set to "contains".
 - Search Text**: An empty text input field.
 - Find**: A teal button to execute the search.
- Browse** section with links for [Topics](#) and [Instruments](#).
- Topics** section listing various categories, each with a plus sign icon:
 - Contextual Data
 - Crime/Delinquency & Victimization
 - Demographic Characteristics
 - Education
 - Family
 - Friends & Social Network
 - Medication & Substance Use and Abuse
 - Meta Data
 - Physical Health
 - Psychological Well-being & Cognition
 - Reproductive Health
 - Risk Behavior
 - Romantic Relationships
 - SES, Labor Market & Occupation

Data Files

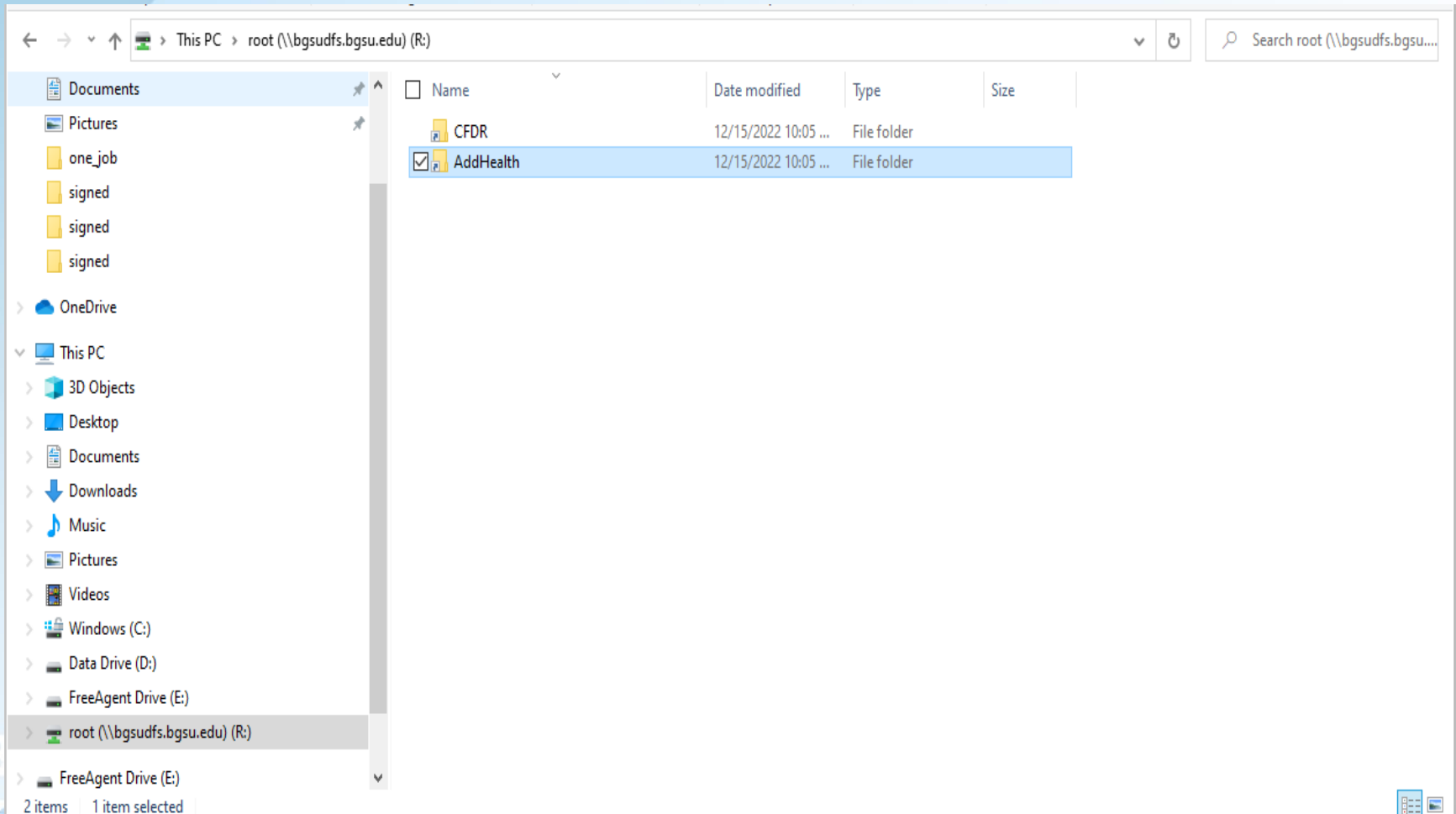
- CFDR stores a copy of public data in the public folder (R:\CFDR\Public\Data\AddHealth). In addition, public data can be downloaded from ICPSR website (<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/21600>).
- CFDR stores a copy of restricted data on the secured server (<R:\AddHealth>). Only people who have obtained the permission from the Carolina Population Center at the University of North Carolina-Chapel Hill can access the data.
- The difference between the public data and the restricted data is that public data contain about only one-third of respondents, whereas the restricted data have all respondents.

Data Files (Cont.)

- Supplemental Education Data
 - All education data are restricted data
 - CFDR stores a copy of restricted data on the secured server (<R:\AddHealth>). Only people who have obtained the permission from the Carolina Population Center at the University of North Carolina-Chapel Hill can access the data
- CFDR has constructed some SAS data sets from the restricted Add Health data, including the in-home interview data, weighted data, and family structure measures from Wave I through IV. These constructed data are stored in the folder “<R:\AddHealth\ADD Health\Add Health study\CFDR SAS data.>”
- If you need to use restricted Add Health data, please contact Dr. Jenjira Yahirun(yahirun@bgsu.edu).

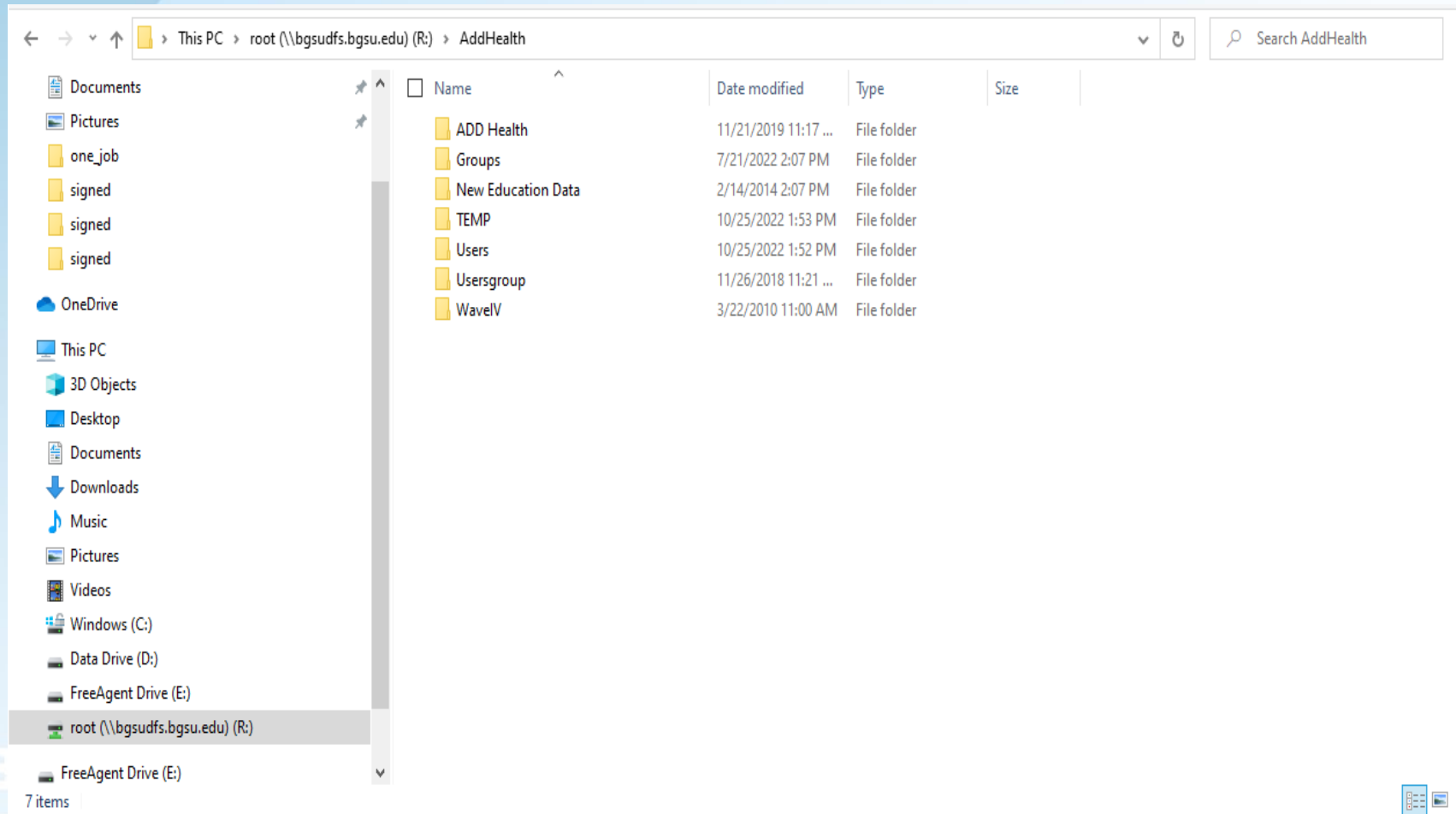
File Location

Folder name: R:\



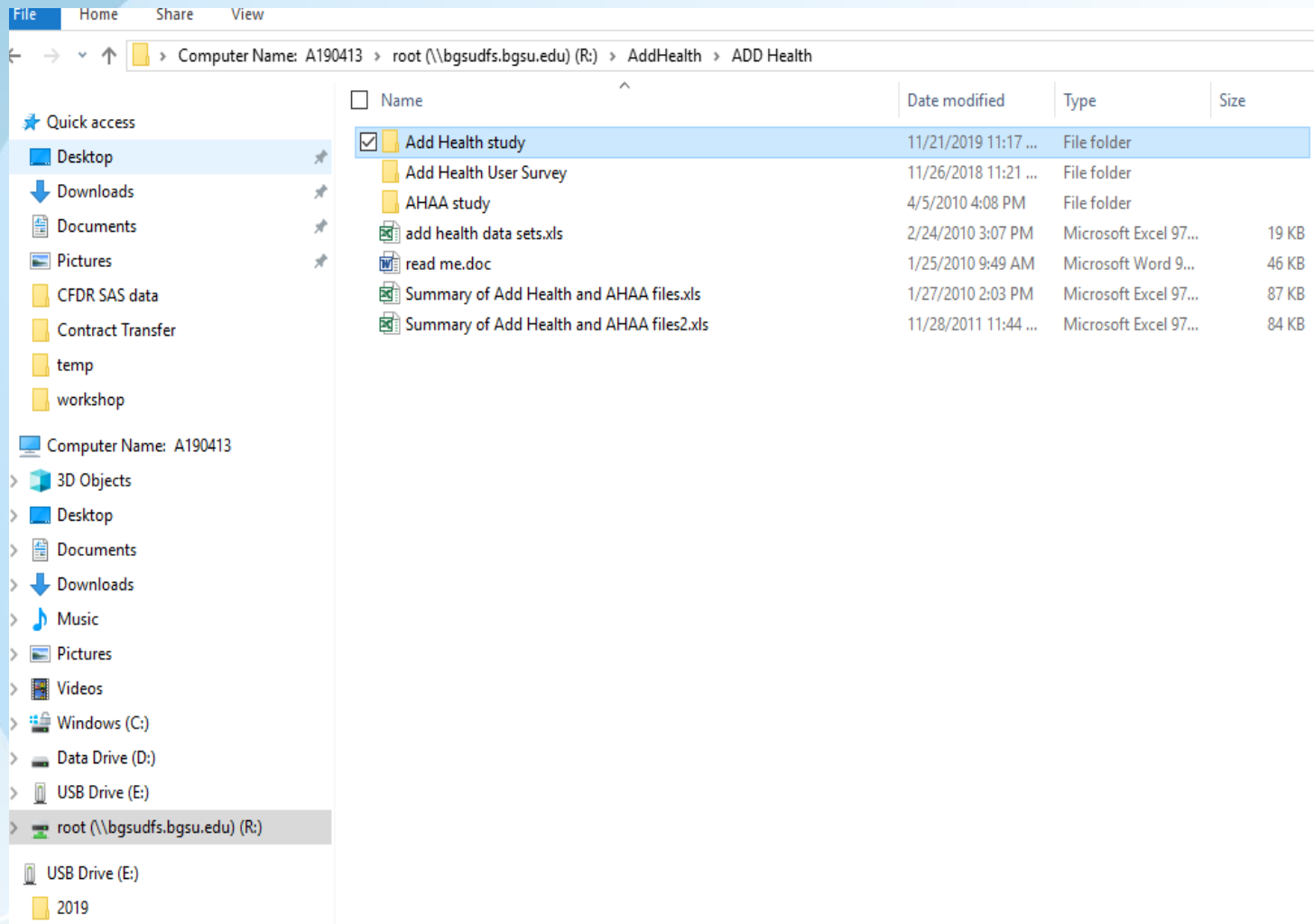
File Location (Cont.)

Former name: R:\AddHealth



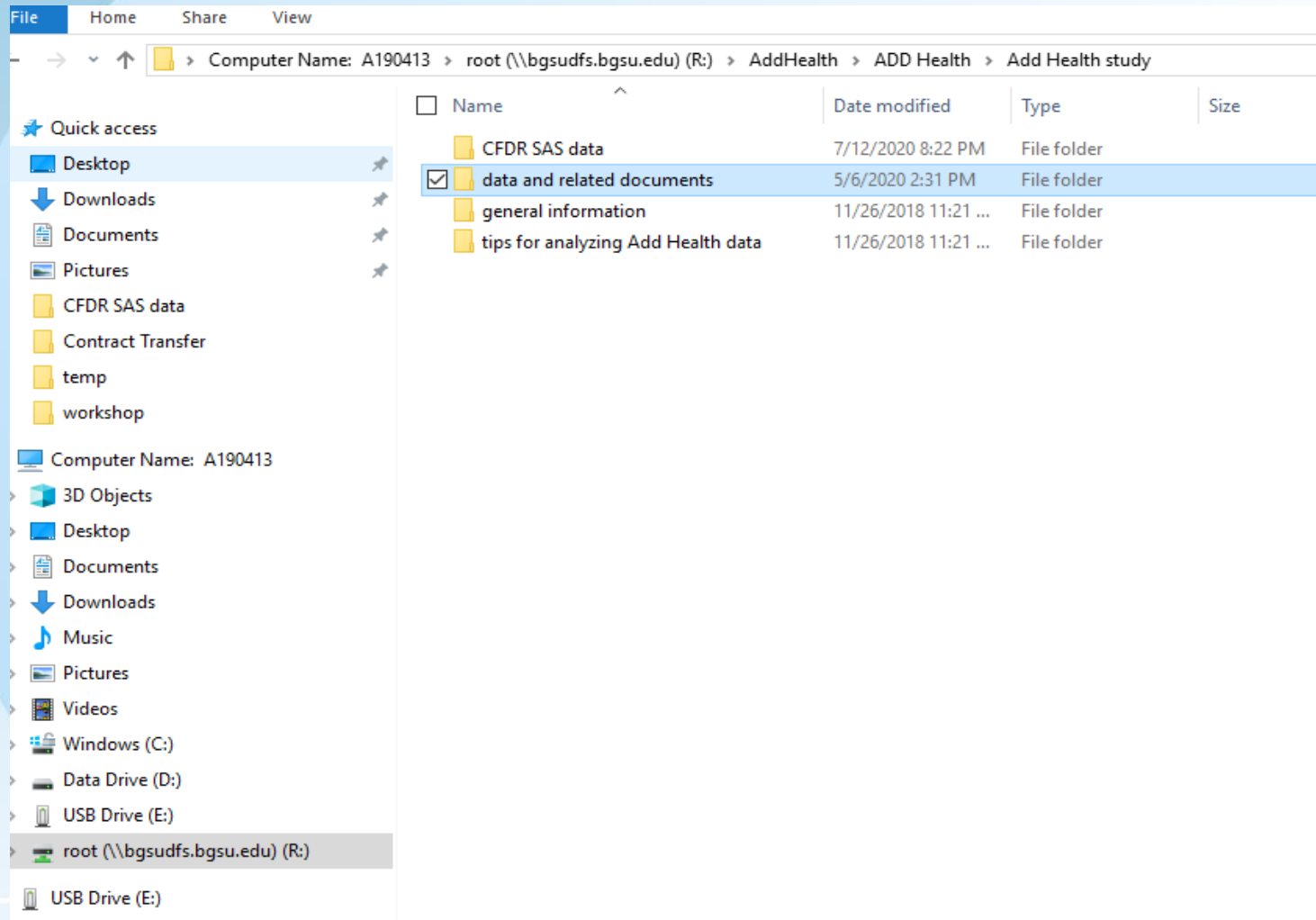
File Location (Cont.)

Folder name: R:\AddHealth\ADD Health



File Location (Cont.)

Folder name: R:\AddHealth\ADD Health\Add Health study



File Location (Cont.)

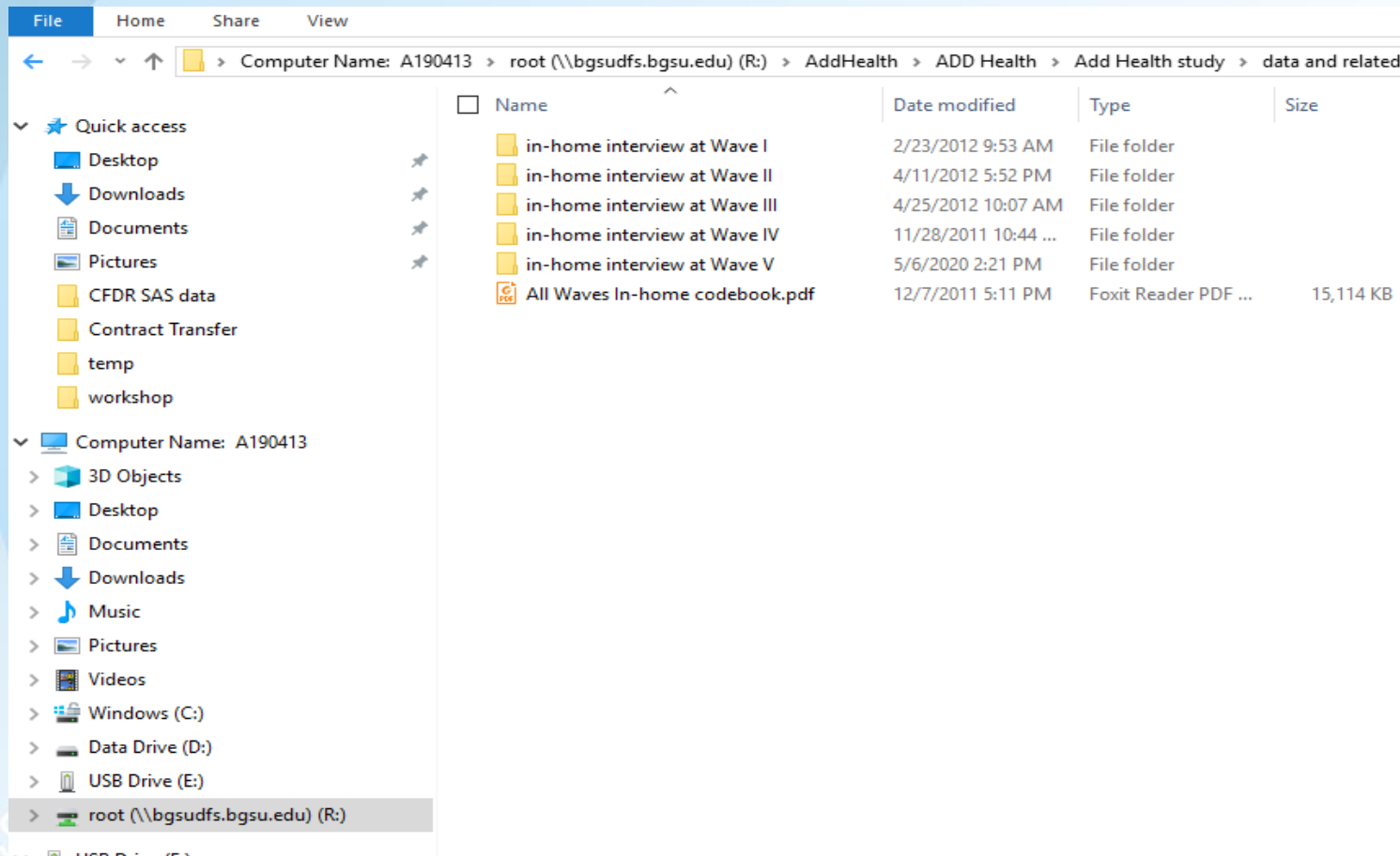
Main Add Health Data Folders: “R:\AddHealth\ADD Health\Add Health study\data and related documents”

The screenshot shows a Windows File Explorer window with the following path: Computer Name: A190413 > root (\\bgsudfs.bgsu.edu) (R:) > AddHealth > ADD Health > Add Health study > data and related documents. The left sidebar shows the navigation pane with 'Computer Name: A190413' expanded to 'root (\\bgsudfs.bgsu.edu) (R:)'. The main pane displays a list of folders with columns for Name, Date modified, Type, and Size. The folder 'in-home interview at Waves I, II, III, and IV' is selected and highlighted in blue.

Name	Date modified	Type	Size
parents at Wave I	5/6/2020 2:38 PM	File folder	
Genetics Wave III and IV	5/6/2020 2:31 PM	File folder	
Meds data	5/6/2020 2:29 PM	File folder	
weight variables at Wave I, II, and III	5/6/2020 2:27 PM	File folder	
<input checked="" type="checkbox"/> in-home interview at Waves I, II, III, and IV	5/6/2020 2:21 PM	File folder	
Sexual Minority Policies Waves III and IV	6/17/2019 1:29 PM	File folder	
Wave V	1/17/2018 2:15 PM	File folder	
Wave IV	1/17/2018 2:12 PM	File folder	
Newly requested data	11/13/2017 11:49 ...	File folder	
Interviewers I, II, III, and IV	11/16/2016 10:56 ...	File folder	
Grouping at Wave 1	8/21/2014 2:20 PM	File folder	
Political Context at Waves I, II, and III	8/21/2014 2:18 PM	File folder	
Education at Wave III	8/21/2014 2:17 PM	File folder	
school at Waves I and II	2/24/2014 2:04 PM	File folder	
Residential Location at Waves at III and IV	1/31/2014 1:03 PM	File folder	
contextual information at Waves I, II, and III	6/13/2013 3:29 PM	File folder	
Supplemental Contextual data at Wave III and IV	5/16/2013 10:16 AM	File folder	
friends at Waves I, II, and III	9/28/2011 11:43 PM	File folder	
ONE Data	10/7/2010 2:54 PM	File folder	
sibling at Waves I and III	12/10/2009 10:45 ...	File folder	
geographic information at Wave III	12/10/2009 10:11 ...	File folder	
neighborhood at Wave I and II	12/10/2009 10:07 ...	File folder	
Peabody Picture Vocabulary Test at Wave III	12/10/2009 10:05 ...	File folder	

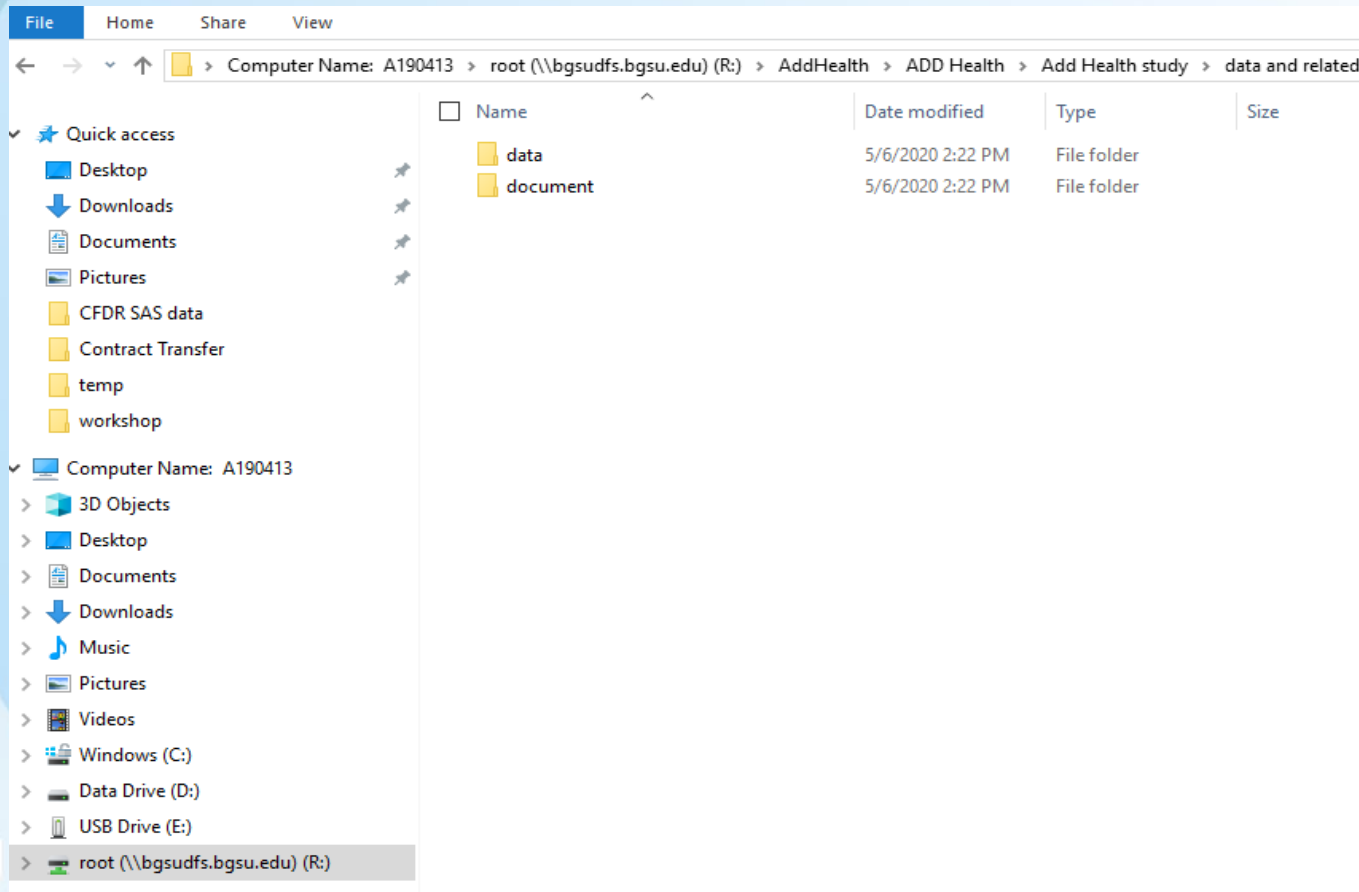
File Location (Cont.)

The location of first four waves of In-home interview data:
“R:\AddHealth\ADD Health\Add Health study\data and related documents\in-home interview at Waves I, II, III, and IV”



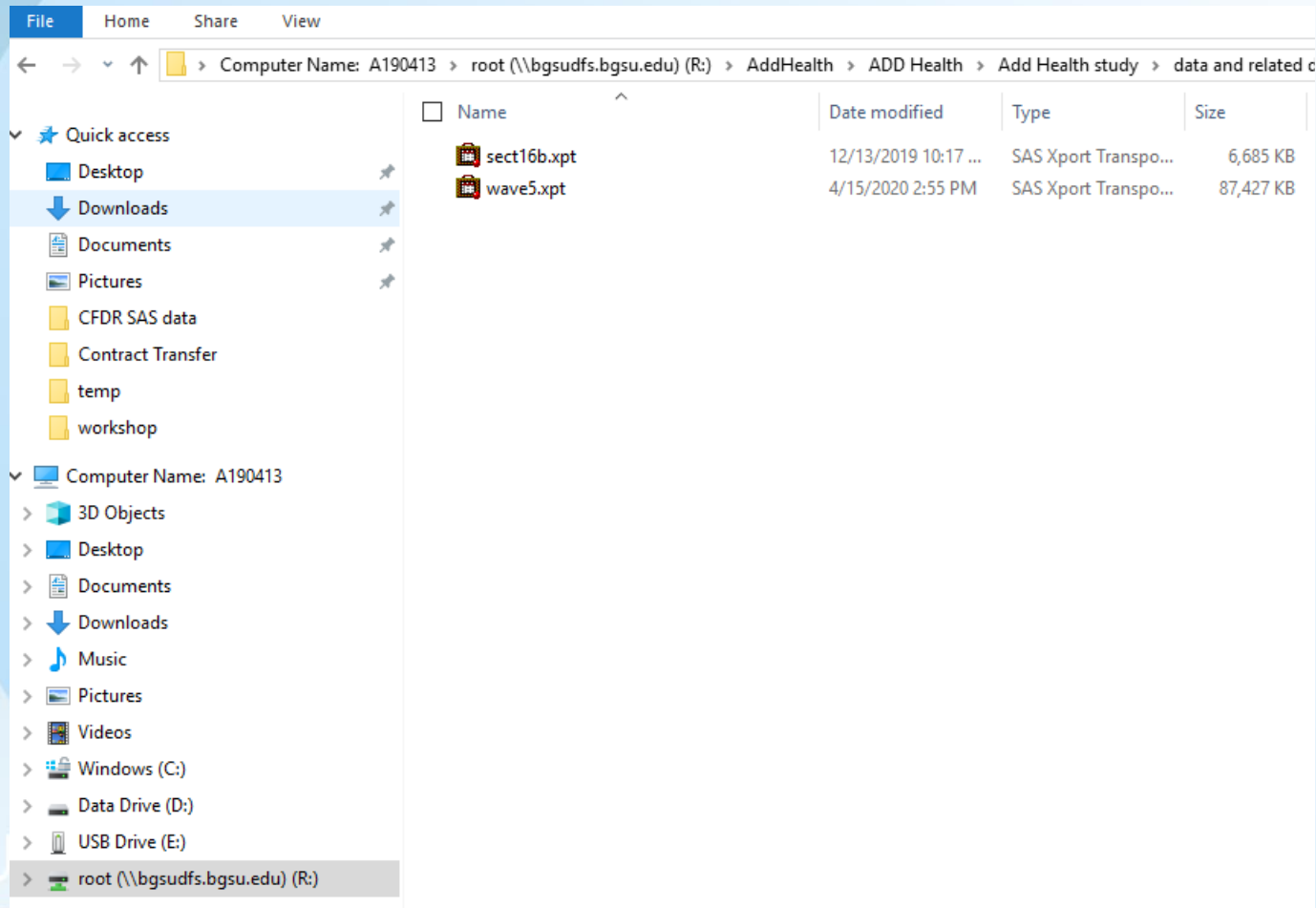
File Location (Cont.)

The location of In-home interview data at Wave V: “R:\AddHealth\ADD Health\Add Health study\data and related documents\in-home interview at Waves I, II, III, and IV\in-home interview at Wave V”



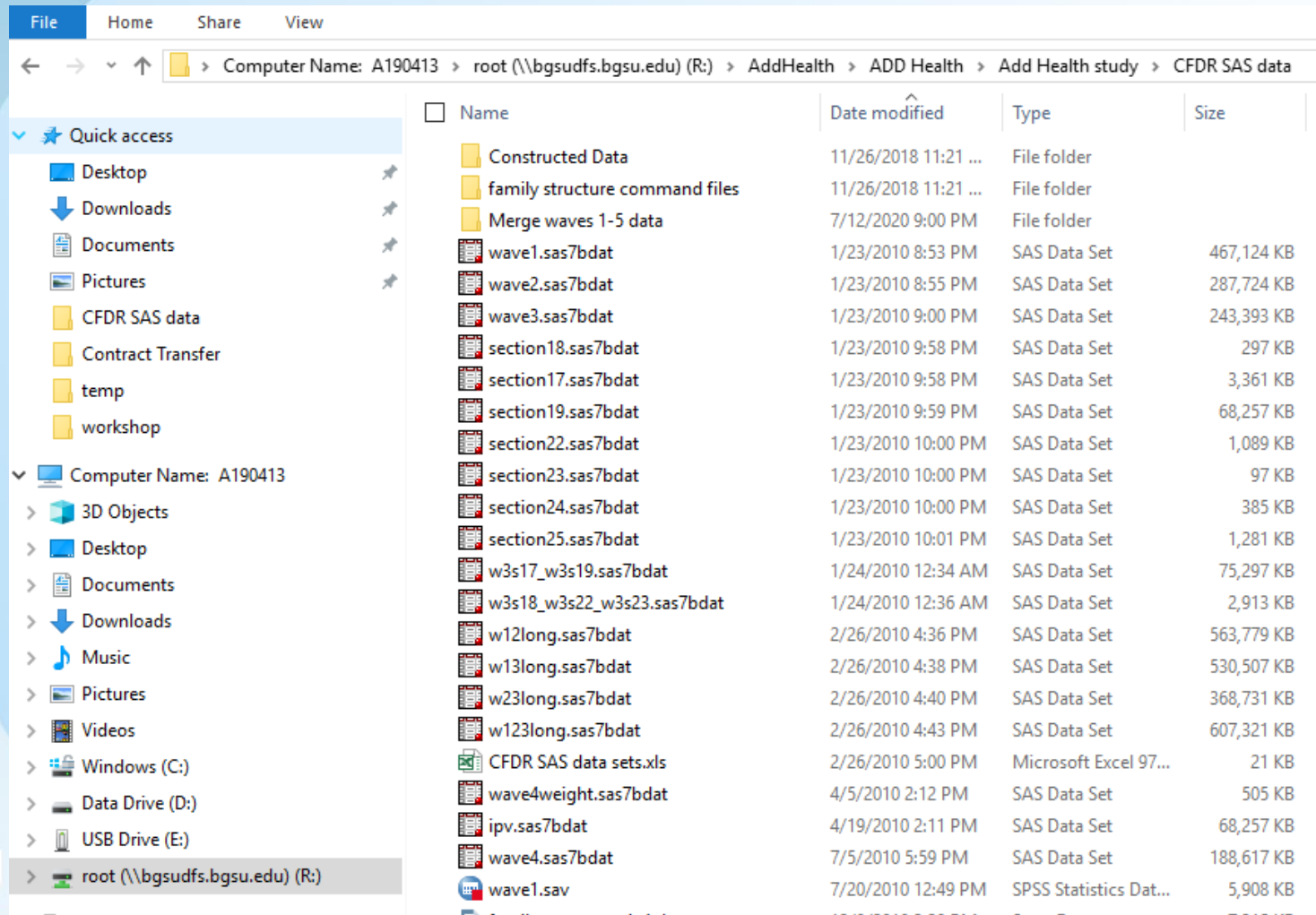
File Location (Cont.)

The location of In-home interview data at Wave V: “R:\AddHealth\ADD Health\Add Health study\data and related documents\in-home interview at Waves I, II, III, and IV\data”



File Location (Cont.)

CFDR puts our constructed data and variables in a specific folder:
R:\AddHealth\ADD Health\Add Health study\CFDR SAS data



Name	Date modified	Type	Size
Constructed Data	11/26/2018 11:21 ...	File folder	
family structure command files	11/26/2018 11:21 ...	File folder	
Merge waves 1-5 data	7/12/2020 9:00 PM	File folder	
wave1.sas7bdat	1/23/2010 8:53 PM	SAS Data Set	467,124 KB
wave2.sas7bdat	1/23/2010 8:55 PM	SAS Data Set	287,724 KB
wave3.sas7bdat	1/23/2010 9:00 PM	SAS Data Set	243,393 KB
section18.sas7bdat	1/23/2010 9:58 PM	SAS Data Set	297 KB
section17.sas7bdat	1/23/2010 9:58 PM	SAS Data Set	3,361 KB
section19.sas7bdat	1/23/2010 9:59 PM	SAS Data Set	68,257 KB
section22.sas7bdat	1/23/2010 10:00 PM	SAS Data Set	1,089 KB
section23.sas7bdat	1/23/2010 10:00 PM	SAS Data Set	97 KB
section24.sas7bdat	1/23/2010 10:00 PM	SAS Data Set	385 KB
section25.sas7bdat	1/23/2010 10:01 PM	SAS Data Set	1,281 KB
w3s17_w3s19.sas7bdat	1/24/2010 12:34 AM	SAS Data Set	75,297 KB
w3s18_w3s22_w3s23.sas7bdat	1/24/2010 12:36 AM	SAS Data Set	2,913 KB
w12long.sas7bdat	2/26/2010 4:36 PM	SAS Data Set	563,779 KB
w13long.sas7bdat	2/26/2010 4:38 PM	SAS Data Set	530,507 KB
w23long.sas7bdat	2/26/2010 4:40 PM	SAS Data Set	368,731 KB
w123long.sas7bdat	2/26/2010 4:43 PM	SAS Data Set	607,321 KB
CFDR SAS data sets.xls	2/26/2010 5:00 PM	Microsoft Excel 97...	21 KB
wave4weight.sas7bdat	4/5/2010 2:12 PM	SAS Data Set	505 KB
ipv.sas7bdat	4/19/2010 2:11 PM	SAS Data Set	68,257 KB
wave4.sas7bdat	7/5/2010 5:59 PM	SAS Data Set	188,617 KB
wave1.sav	7/20/2010 12:49 PM	SPSS Statistics Dat...	5,908 KB

File Location (Cont.)

The content of the folder: R:\AddHealth\ADD Health\Add Health study\CFDR SAS data\Constructed Data

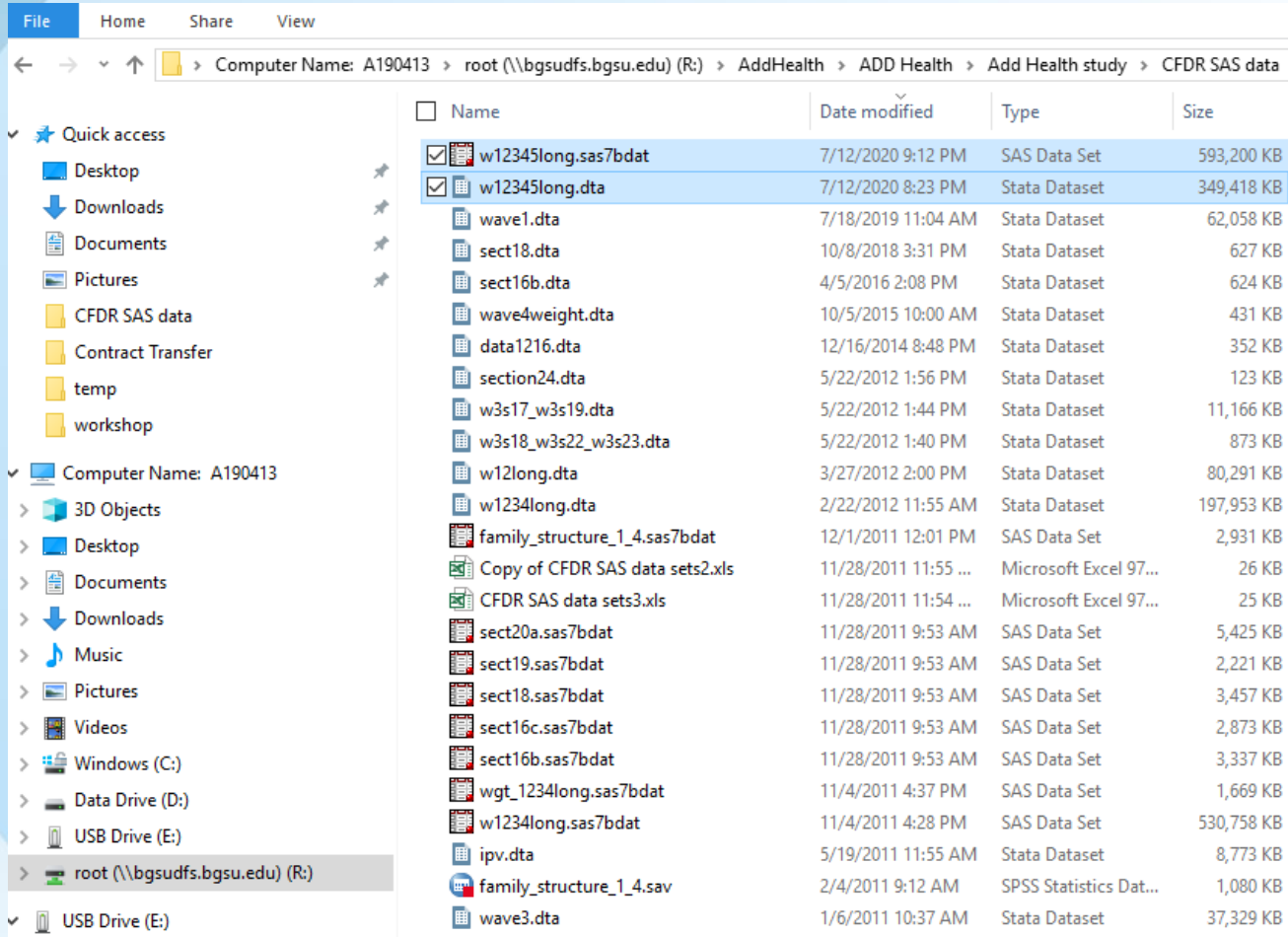
The screenshot shows a Windows File Explorer window with the following details:

- Address Bar:** Computer Name: A190413 > root (\\bgsudfs.bgsu.edu) (R:) > AddHealth > ADD Health > Add Health study > CFDR SAS data
- Left Pane (Navigation):** Quick access (Desktop, Downloads, Documents, Pictures, CFDR SAS data, Contract Transfer, temp, workshop) and Computer Name: A190413 (3D Objects, Desktop, Documents, Downloads, Music, Pictures, Videos, Windows (C:), Data Drive (D:), USB Drive (E:), root (\\bgsudfs.bgsu.edu) (R:), USB Drive (E:)).
- Right Pane (File List):**

Name	Date modified	Type	Size
focal partner	11/26/2018 11:21 ...	File folder	
SES	11/26/2018 11:21 ...	File folder	
UNC race and family structure variables	6/24/2020 3:45 PM	File folder	
first_cohabitation.dta	1/9/2013 5:00 PM	Stata Dataset	371 KB
first_cohabitation.sas7bdat	1/9/2013 5:00 PM	SAS Data Set	333 KB
first_marriage.dta	10/1/2012 9:18 AM	Stata Dataset	371 KB
first_marriage.sas7bdat	1/9/2013 4:36 PM	SAS Data Set	333 KB
first_union.dta	8/3/2012 5:41 PM	Stata Dataset	371 KB
first_union.sas7bdat	1/9/2013 4:36 PM	SAS Data Set	333 KB

File Location (Cont.)

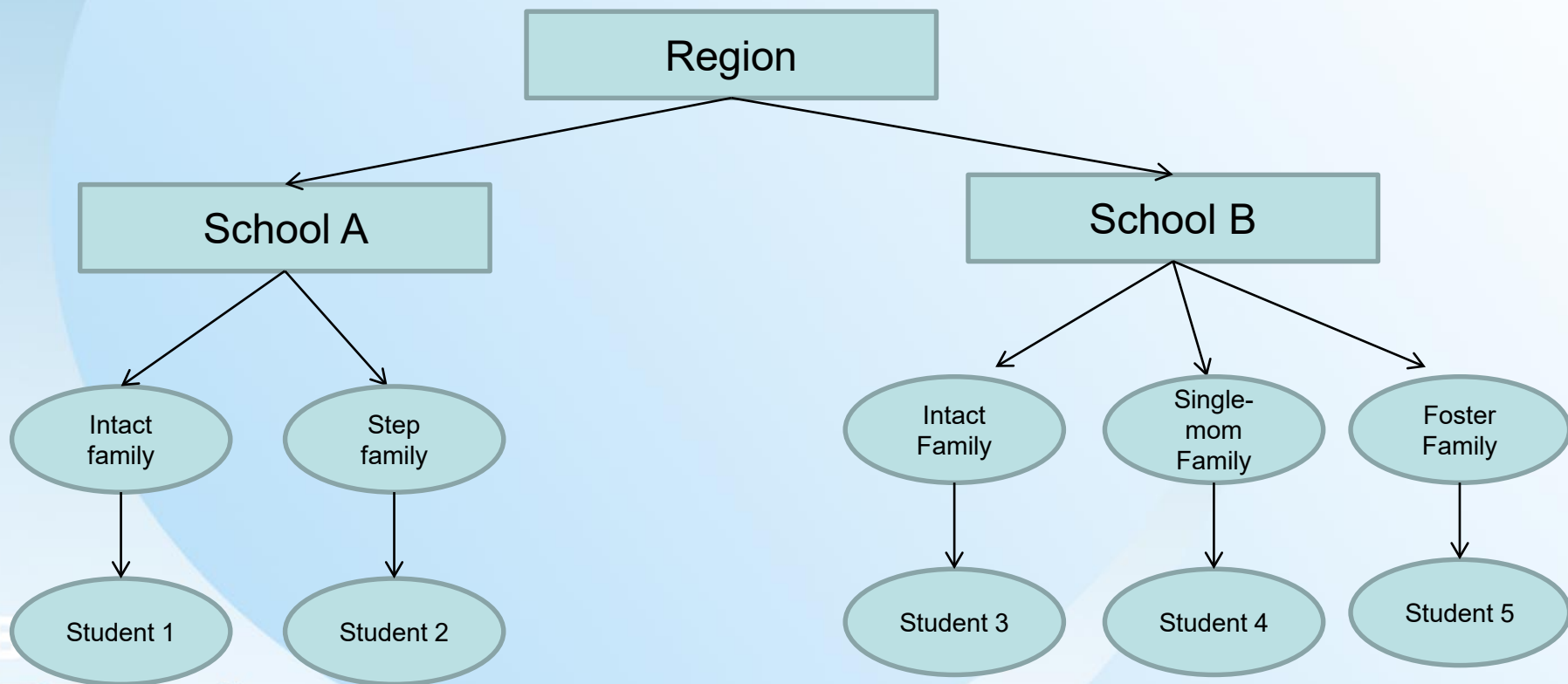
All Five Waves of In-Home interview individual data are in the folder “R:\AddHealth\ADD Health\Add Health study\CFDR SAS data”



Name	Date modified	Type	Size
<input checked="" type="checkbox"/> w12345long.sas7bdat	7/12/2020 9:12 PM	SAS Data Set	593,200 KB
<input checked="" type="checkbox"/> w12345long.dta	7/12/2020 8:23 PM	Stata Dataset	349,418 KB
wave1.dta	7/18/2019 11:04 AM	Stata Dataset	62,058 KB
sect18.dta	10/8/2018 3:31 PM	Stata Dataset	627 KB
sect16b.dta	4/5/2016 2:08 PM	Stata Dataset	624 KB
wave4weight.dta	10/5/2015 10:00 AM	Stata Dataset	431 KB
data1216.dta	12/16/2014 8:48 PM	Stata Dataset	352 KB
section24.dta	5/22/2012 1:56 PM	Stata Dataset	123 KB
w3s17_w3s19.dta	5/22/2012 1:44 PM	Stata Dataset	11,166 KB
w3s18_w3s22_w3s23.dta	5/22/2012 1:40 PM	Stata Dataset	873 KB
w12long.dta	3/27/2012 2:00 PM	Stata Dataset	80,291 KB
w1234long.dta	2/22/2012 11:55 AM	Stata Dataset	197,953 KB
family_structure_1_4.sas7bdat	12/1/2011 12:01 PM	SAS Data Set	2,931 KB
Copy of CFDR SAS data sets2.xls	11/28/2011 11:55 ...	Microsoft Excel 97...	26 KB
CFDR SAS data sets3.xls	11/28/2011 11:54 ...	Microsoft Excel 97...	25 KB
sect20a.sas7bdat	11/28/2011 9:53 AM	SAS Data Set	5,425 KB
sect19.sas7bdat	11/28/2011 9:53 AM	SAS Data Set	2,221 KB
sect18.sas7bdat	11/28/2011 9:53 AM	SAS Data Set	3,457 KB
sect16c.sas7bdat	11/28/2011 9:53 AM	SAS Data Set	2,873 KB
sect16b.sas7bdat	11/28/2011 9:53 AM	SAS Data Set	3,337 KB
wgt_1234long.sas7bdat	11/4/2011 4:37 PM	SAS Data Set	1,669 KB
w1234long.sas7bdat	11/4/2011 4:28 PM	SAS Data Set	530,758 KB
ipv.dta	5/19/2011 11:55 AM	Stata Dataset	8,773 KB
family_structure_1_4.sav	2/4/2011 9:12 AM	SPSS Statistics Dat...	1,080 KB
wave3.dta	1/6/2011 10:37 AM	Stata Dataset	37,329 KB

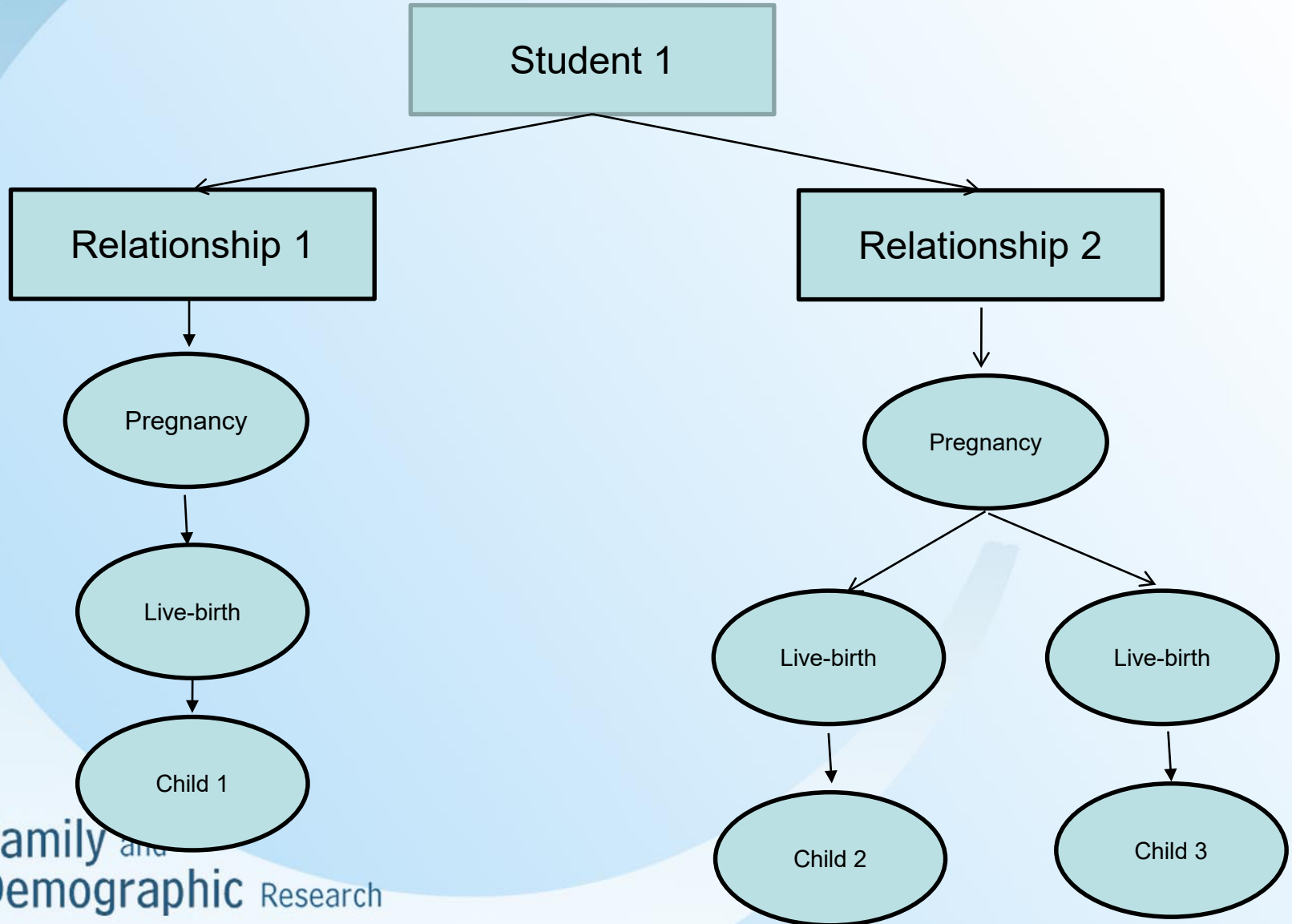
Unit of Analysis

- Add Health collects information on individual adolescents, their social environment (e.g., neighborhood, school, family) and various aspects of social relations and experiences (e.g., intimate relationship, pregnancy, live births, and parent-children relationship).
- An example of the nested structure of neighborhood, school, family, and individual adolescents is as follows:



Unit of Analysis (Cont.)

An example of the nested structure of Individual, relationship, pregnancy, live births, and parent-children relationship is as follows:



Analytic Tips

- How to find the variables you need?
- How to read Add Health data?
- How to merge data?
- How to weight Add Health data?
- How to change the unit of analysis?

How to Find the Variables You Need?

- Use codebooks or Add Health Codebook Explorer to locate the variables of interest
- Add Health provides codebooks that list all of the names and wordings of the variables at each wave. Thus, if you are interested in the in-home interview data, you should start finding your variables by reading through the following codebooks:
 - WAVE1NDX.PDF
 - WAVE2NDX.PDF
 - wave3ndx.pdf
 - wave4ndx.pdf
 - WaveVMixed_ModeSurveyCodebook.pdf and WaveVSection16BCodebook
 - Each subject area usually has its own codebook, and you can only find value labels in each codebook

How to Read Add Health Data?

- The public data of Add Health may be in SAS, Stata, or SPSS format. You can use Stat/Transfer to change the data from one format to another
- The restricted data of Add Health are initially in SAS export format. The following codes provide instructions on how to use SAS and Stata to read in the SAS export file

How to Read Add Health Data? (Cont.)

SAS code:

```
LIBNAME wave1 xport "R:\ADD Health\Add Health study\data and related documents\in-home interview at Waves I, II, and III\in-home interview at Wave I\data\allwave1.exp";
```

```
LIBNAME out "r:\Add Health\temp";
```

```
Data out.wave1;
```

```
Set wave1.allwave1;
```

```
Run;
```

Stata code:

```
fdause "R:\ADD Health\Add Health study\data and related documents\in-home interview at Waves I, II, and III\in-home interview at Wave I\data\allwave1.exp"
```


How to Merge Data?

- When do data need to be merged?
 - If you want to combine data from different waves of Add Health
 - If you want to combine data with different unit of measurements
 - If you want to use both Add Health data and Education data
- SAS and Stata sample commands to merge Waves I and II data are shown in the following slides:

How to Merge Data? (Cont.)

- SAS code:

```
Libname in "R:\AddHealth";  
*****;  
PROC SORT DATA=in.wave1;  
BY aid;  
RUN;  
*****;  
PROC SORT DATA=in.wave2;  
BY aid;  
RUN;  
*****;  
DATA in.wave12;  
MERGE in.wave1 (IN=in_wave1) in.wave2 (IN=in_wave2);  
BY aid;  
RUN;
```

How to Merge Data? (Cont.)

- Stata code:

```
use "R:\AddHealth\wave1.dta"  
sort aid  
save "R:\AddHealth\wave1_2.dta", replace  
*****  
use "R:\AddHealth\wave2.dta"  
sort aid  
save "R:\AddHealth\wave2_2.dta", replace  
*****  
use "R:\AddHealth\wave1_2.dta", clear  
sort aid  
merge aid using "R:\AddHealth\wave2_2.dta"  
tab1 _merge  
rename _merge wave12  
label variable wave12 "indicator for merging waves 1 and 2"  
sort aid  
save "R:\AddHealth\wave12.dta", replace
```

How to Weight Add Health Data?

- Add Health data are collected with a complex survey design. Therefore, each respondent does not have the same probability of being selected into the sample and thus needs to be reweighted
- Clustering of students from the same regions and schools
- The analysis of Add Health data always needs to be weighted in order to adjust for the effects of its complex survey design
- SAS and Stata differ in their abilities of performing statistical analyses, while controlling for the effects of the complex survey design

How to Weight Add Health Data? (Cont.)

Table 3. Select Stata and SAS procedures for Analyzing Survey Data

Analysis	Stata command	SAS command
Estimate means for survey data	svy: mean	Proc Surveymeans
Estimate proportions for survey data	svy: tab	Proc Surveyfreq
Linear regression for survey data	svy: regress	Proc Surveyreg
Logistic regression for survey data, reporting odds ratios	svy: logistic	Proc Surveylogistic
Cox proportional hazards model for survey data	svy: stcox	Proc Surveyphreg
Ordered logistic regression for survey data	svy: ologit	
Ordered probit regression for survey data	svy: oprobit	
Multinomial (polytomous) logistic regression for survey data	svy: mlogit	
Multinomial probit regression for survey data	svy: mprobit	
Parametric survival models for survey data	svy: streg	
Generalized linear models for survey data	svy: glm	
Generalized negative binomial regression for survey data	svy: gnbreg	
Poisson regression for survey data	svy: poisson	
Zero-inflated negative binomial regression for survey data	svy: zinb	
Zero-inflated Poisson regression for survey data	svy: zip	

How to Weight Add Health Data? (Cont.)

If you use the full sample in the analysis, you can use either SAS or Stata for the analysis.

SAS code:

```
Libname in "R:\AddHealth\TEMP";  
  
proc surveylogistic data= in.logit3;  
cluster psuscid3;  
weight gswgt3;  
strata region3;  
model h3ed3 = bio_sex3 calcage3;  
run;
```

Stata code:

```
use "R:\AddHealth\TEMP\logit3.dta", clear  
svyset psuscid3 [pweight =gswgt3], strata(region3)  
svy: logit h3ed3 bio_sex3 calcage3
```

How to Weight Add Health Data? (Cont.)

If you use only part of the sample in the analysis, you should use Stata for the analysis because SAS has the sub-population options for the Proc Surveymeans command only.

Stata code:

```
use "R:\AddHealth\TEMP\logit3.dta", clear
svyset psuscid3 [pweight =gswgt3], strata(region3)
svy, subpop(marker): logit h3ed3 bio_sex3 calcage3
```


How to Change the Unit of Analysis?

- Changing the unit of analysis means changing the unit of observations in the data set. Because of the nested structure of Add Health data, you can change the unit of observations from one level to another
- When the unit of analysis changes, the number of valid observation changes, too

Table 2. The number of Units at Different Levels of Analysis for Section 25 of the Wave III of Add Health

Unit of Analysis	Number of Analysis Units
CHILD	4,181
BIRTH	4,181
PREGNANCY	4,055
RELATION	3,293
RESPONDENT	2,960

- SAS and Stata examples of changing unit from birth to pregnancy

How to Change the Unit of Analysis? (Cont.)

- SAS code:

```
PROC SORT DATA = out.sect25 OUT=out.preg;  
BY aid rrelno rpregno;  
RUN;
```

```
PROC FREQ DATA = out.preg;  
TABLES birthno;  
RUN;
```

```
PROC TRANSPOSE DATA =out.preg  
OUT=out.f_preg PREFIX =birth;  
BY aid rrelno rpregno;  
ID birthno;  
VAR c_age;  
RUN;
```

```
PROC CONTENTS DATA = out.sect25;  
RUN;
```

```
PROC CONTENTS DATA = out.f_preg;
```

```
RUN;
```

How to Change the Unit of Analysis? (Cont.)

- Stata code:

```
use t:\temp\sect25.dta, clear
```

```
des aid rrelno rpregno
```

```
sum rrelno rpregno
```

```
tostring rrelno, generate(srrelno)
```

```
tostring rpregno, generate(srpregno)
```

```
gen said_rp3 = aid + srrelno + srpregno
```

```
replace said_rp3 = aid + "0" + srrelno + srpregno if rrelno >=1 & rrelno <=9
```

```
label variable said_rp3 "string id for pregnancy record"
```

```
sort said_rp3
```

```
des
```

```
tab1 birthno
```

```
reshape wide c_age, i(said_rp3) j(birthno)
```

```
des
```

```
rename c_age1 birth1
```

```
rename c_age2 birth2
```

```
save "t:\temp\pregnancy.dta", replace
```

Studies Using Add Health Data

- There have been more than 4,000 publications using Add Health. You can locate them through Add Health or ICPSR web site:
 - Add Health Web site
 - <https://addhealth.cpc.unc.edu/publications/>
 - ICPSR website:
 - After you find Add Health data (<https://www.icpsr.umich.edu/web/ICPSR/series/1006/studies>), click on the tab “Data-related Publications”

Help with Add Health Analyses

- **Official Add Health listserv**

Listserve is a place where Add Health users ask and answer questions about analyzing Add Health data. To subscribe the Add Health listserv, send e-mail to: listserv@unc.edu and in the body of the message put: subscribe addhealth2 <firstname lastname>

- **Add Health Users Conference**

Carolina Population Center at University of North Carolina has hosted 14 Add Health Users Conferences on how to construct and analyze Add Health data. Add Health website (<http://www.cpc.unc.edu/projects/addhealth/news>) provides information on the upcoming Add Health User Conference.

- **CFDR Programming Help**

If you have programming problems, contact Hsueh-Sheng Wu at whu@bgsu.edu

BGSU



Center for Family and Demographic Research

Conclusions

- Add Health is an excellent data set for studying how adolescents make transitions into adulthood and how such transitions are influenced by their families, peers, schools, and neighborhoods.
- When using Add Health data, researchers should allow themselves sufficient time to construct variables because they may need to correct for possible inconsistencies in data from different measurement units and/or across different waves before they can construct variables for their analyses.
- The analysis of Add Health data always needs to be weighted in order to adjust for its complex survey design. This workshop covers how to weight the results if the single-level analysis is conducted. If multilevel analysis is to be conducted, researchers need to rescale the weights, so units at different levels have their respective weights. For details, consult the document, Guidelines for Analyzing Add Health data (https://addhealth.cpc.unc.edu/wp-content/uploads/docs/user_guides/GuidelinesforAnalysisofAddHealthData_202004.pdf)
- If you have any questions about using, coding, or analyzing Add Health data, please feel free to contact me (wuh@bgsu.edu).