National Endowment for the Arts

2012 SPPA PUBLIC-USE DATA FILE USER'S GUIDE

A Technical Research Manual

Prepared by Timothy Triplett
Statistical Methods Group
Urban Institute

September 2013

Table of Contents

Introduction

In 2012, the U.S. Census Bureau fielded the National Endowment for the Arts' seventh Survey of Public Participation in the Arts (SPPA). The SPPA is the nation's largest recurring cross-sectional survey of adult participation in arts and cultural activity. Besides informing NEA-commissioned reports and policy decisions, the raw data are shared with non-NEA researchers for their own use. This User's Guide presents basic information on the 2012 Survey of Public Participation in the Arts (SPPA) public-use data file.

The first section of the user's guide gives an overview of the survey, including descriptions of the sample design and data collection procedures. The following section discusses the major survey design changes implemented in 2012. Section 3 explains how missing data have been coded, and it includes recommendations on how to handle the missing data when doing your analysis. Section 4 and 5 provides information on how to use the survey weights and procedures for calculating variances and standard errors of survey estimates. Sections 6 and 7 describe some of the issues that need to be understood when doing multivariate analysis that may involve analyzing questions from different modules of the SPPA questionnaire or may involve doing analysis that uses CPS variables. Section 8, the final section, provides information on how to conduct trend analysis or how to combine SPPA data across years. The appendices to this guide include a data dictionary that displays both weighted and unweighted frequencies, a copy of the questionnaire, and U.S. Census Bureau source and accuracy statement.

The overall goal of this guide is to provide users with enough information about the data and the survey itself to be able to correctly use the public-use data. Some subsections of the guide will be in *italics. T*hese subsections will be preceded by one of the following icons:

 Indicating critical points that all users should understand

 Indicating useful tips, but not essential reading

 Indicating sections meant primarily for advanced users

Section 1: Overview of the Design of the Survey

The SPPA was used to collect arts participation data from American adults in 1982, 1985, 1992, 1997, 2002, 2008, and 2012. Until 1997, the SPPA was a supplement attached to the U.S. Census Bureau's National Crime Survey. In 1997, the SPPA was a stand-alone study conducted by WESTAT, a private contractor. Beginning with 2002, the SPPA is now being collected as a supplement attached to the Census Bureau's Current Population Survey.

The 2012 SPPA was administered as a supplement to the July 2012 Current Population Survey (CPS). The CPS is conducted by the U.S. Census Bureau for the Bureau of Labor Statistics (BLS). The BLS uses the data to provide a monthly report on the national employment situation. This report provides estimates of the number of employed and unemployed people in the United States. Approximately 60,000 households are eligible for the CPS each month. Sample households are selected by a multistage, stratified, statistical sampling scheme. A household is interviewed for 4 successive months, then not interviewed for 8 months, then returned to the sample for 4 additional months.

The SPPA supplement was administered to one-half of the eligible CPS households. These were the households that were in the exit round of the CPS sample rotation. After taking into account nonresponse, a total of 21,778 households provided at least one completed SPPA interview. The SPPA survey allows proxy responses for spouse or partners and in larger households a second supplemental interview was often conducted. The final distribution of 2012 SPPA interviews consisted of an initial 21,778 randomly selected adults, plus an additional 3,677 interviews were collected from adults in residing in larger households, and there were 11,811 spouse or partner proxy interviews. So, overall data on 37,266 adults was collected and the average number of completed adult interviews per household was 1.7.

The CPS is one of the nation's most respected surveys and still maintains a high response rate. Although the SPPA is conducted with only a portion of the overall CPS sample, it is a random portion, representative of the non-institutionalized U.S. adult population. Because the SPPA questions are asked after the Monthly CPS questions, the SPPA response rate is lower than the overall CPS response rate. However, the response rate for the CPS respondents selected to complete the 2012 SPPA was still 74.8% since 25,455 of the CPS respondents completed the SPPA supplement out of the 34,039 CPS respondents that were randomly selected to complete the supplement. For the sections of the survey that included spouse partner questions the response rate is actually 77.0% since when you include spouse partner responses, survey data was obtained from 37,266 people out of the 48,372 people who were sampled. The higher spouse partner

response rate is not surprising given that response rates are usually much lower for young adults and second interviews in large households which are also less likely to include spouse/partners. The 2012 response was less than the 81.6% spouse partner response rate achieved on the 2008 SPPA study. This decrease in SPPA response rate is likely in part a result of fielding the 2012 SPPA in July, a month in which many eligible respondents may be away from home during period in which the CPS data is collected.

The 2002 SPPA was the first SPPA survey to be conducted as a supplement to the CPS. Based on this experience, the 2008 and 2012 SPPA were redesigned to better handle a number of important design issues that arose from the 2002 version. The main goal was to develop a design that would be less burdensome to survey-takers. There were two major design changes that reduced the average burden:

> 1. Rather than attempt to interview all adults in the household (as in 2002), the 2008 and 2012 SPPA randomly sampled adults and, for many of the questions, accepted proxy responses for spouses or partners. In 2002, proxy responses were accepted only when repeated attempts to reach a designated respondent had failed. By contrast, for the 2008 and 2012 SPPA, the questions were worded so that they indicated which person the question specifically referred to. Proxy responses were clearly identified in the data file prepared by the Census Bureau.

> 2. Rather than administer the entire SPPA survey to all respondents, the questionnaire was separated into modules, so that any one respondent answered only a core set of arts attendance questions and 2 other modules.

Besides being a major design change, randomly sampling adults to complete the SPPA supplement was a particularly challenging task and something that departed from standard CPS sampling procedures. The CPS is administered to any person 16 or older who is able to report employment information on all persons 16 years or older in the household. The SPPA collects information only from adults 18 or older. In addition, the nature of many SPPA questions requires that the survey be based on direct self-reports rather than proxy reports. To make the SPPA a better fit for the CPS protocol, the 2008 and 2012 SPPA allowed some proxy responses from spouses or partners.

A small pilot study was conducted prior to the 2012 SPPA survey. Based on this study, NEA researchers were satisfied that respondents who had a spouse or partner would be able to readily answer questions on core arts participation activities and leisure activities on behalf of the spouse or partner. Therefore the questions asked in both the Core 1and Core 2 sections of the survey, plus the leisure activity questions asked in Module D of the questionnaire were asked of

the respondent and then repeated to get responses for their spouse or partners. In addition, the first four questions asked in Module A were also asked of both the respondent and their spouse or partner. More detail on what questions were asked in each module of the SPPA questionnaire will be covered in the next section of this report and Appendix B contains a complete copy of the final 2012 SPPA questionnaire. Overall, information on 11,811 SPPA respondents was obtained from spouse/partner proxy reports.

To ensure that all adults 18 and older had a chance of being included in the survey, sometime a second SPPA interview was conducted after the CPS and first SPPA supplement interviews had taken place. The second interview occurred in households that contained one or more adults in addition to the CPS respondent and his or her spouse or partner. Among this remainder group, an adult was randomly chosen for the second interview.

If the CPS respondent was 18 or older and had a spouse or partner, the initial survey was administered at random either to the CPS respondent or his or her partner, to avoid any potential bias that could have occurred from always choosing the CPS respondent. Similarly, in all households that had two interviews, the second respondent was chosen at random. The respondent selection strategy succeeded in achieving a higher response rate (77.0% vs. 69.6%) than the 2002 procedure of attempting to interview every adult in the household.

Researchers planning to compare the 2012 estimates with prior years should read Section 8 of this user's guide.

In some situations, a respondent was never available to complete the SPPA supplement—and, therefore, Census interviewers accepted a full proxy interview. This was done also for the 2002 and 2008 SPPA, though the number of full proxy interviews was much higher in 2002, given the need to attempt interviews with all adults 18 or older in the household. The Census Bureau did not flag the 2002 full proxy interviews, so it is difficult to study how these interviews may have impacted the 2002 estimates. For the 2012 survey, there were 5,196 full proxy interviews and 45 of these interviews included spouse/partner reports. This means that the 5,196 full proxy interviews collected information on 5,241 people. Full proxy interviews can be identified by using the full proxy flag variable (labeled "PUNXTPR3"in the data dictionary).

Note: if you want to compare spouse/partner proxy responses with direct self-reports, you cannot simply compare the 11,811 spouse/partner responses with the 24,455 direct self-reported responses. Neither group is a representative sample. The 11,981 spouse/partner reports include responses only from people

who have a spouse or partner, while the 24,455 respondents who self-reported includes a disproportionate number of non-married respondents.

Section 8 of this manual provides a few examples and suggestions on how to properly compare proxy versus non-proxy reports and discusses in more detail the impact of proxy reporting when comparing 2008 estimates with those of prior years.

Section 2: Changes in the Survey Design

There were two major design changes implemented for the 2012 SPPA. First, the 2012 SPPA included two core components rather than one; the first set of core questions (Core 1) was nearly an exact replicate of the core questions used in previous years to ask about arts attendance[1]; and a second new experimental set of questions (Core 2) that also asked about arts attendance. Respondents were assigned randomly to answer either the Core 1 or Core 2 set of questions. The reason for developing two separate sets of core questions was to be able to maintain backward compatibility with earlier waves of the SPPA in order to look at historical trends while also trying to adapt the questionnaire to reflect some of the changes in the meaning of arts participation that has occurred over the 30-year period since the survey was first conducted. The 2012 survey in particular was marked by numerous changes, which benefited from consultations with arts and cultural researchers, practitioners, and policy-makers many of whom attended a full day SPPA planning workshop held on November 2, 2010. The 2012 sample size was roughly doubled from 2008, which afforded the opportunity for the SPPA to do some more experimenting with measuring arts participation while maintaining the ability to look at change over time.

The second major design change was that the 2012 included five separate questionnaire modules designed to capture other types of arts participation as well as participation in other leisure activities. So, in addition to completing a random set of set of core arts participation questions, respondents were assigned randomly to answer questions from two of the five SPPA modules. If a respondent was answering for themselves and on behalf of their spouse or partner, the respondent and spouse partner questions followed the same path through the instrument. However, as will be covered in section 4 (Using the Survey Weights) not all modules included spouse/partner versions of the questions.

While the Census Bureau again agreed to administer the SPPA as a supplement to the CPS, the Bureau expressed strong concerns about the length of time the interviewers had taken to conduct the prior two SPPA surveys, which were also supplements to the CPS. Hence, a big challenge in revising the 2012 SPPA was to ensure that it did not take too long to administer. Eliminating a few questions would not be enough and eliminating sections of the SPAA would have spoiled the richness of the survey. As a solution, the questionnaire was separated into modules, so that any one respondent would answer only one set of the core arts attendance questions and two of the five other modules (sections of the survey).

---

[1] The only change in the first set of core questions was that addition of the following question; "With the exception of elementary or high school performances did you go to any other music, theater, or dance performance during the last 12 months?" which was asked after the respondent answered all the performing arts questions.

Given that respondent did not answer each of the SPPA modules impacts users who plan to analyze data from any questions other than the core arts attendance questions. So, it is essential that data users who plan to analyze more than just the core questions understand the modular design of the 2012 questionnaire.

This modular approach allowed for NEA researchers to ask many new questions while being able to retain many of the SPPA questions that have been asked in the past. In particular, new questions were needed to deal with ever changing definition of art forms and the impact technology has had on how people engage in the arts. The following were the 2012 modules — Module A (Other arts participation questions and music preference questions), Module B (Accessing art through media), Module C (Creating arts through media), Module D (Creating, performing, and other activities), and Module E (Arts learning). The modules were randomly assigned so that that the questions within each module were asked of a representative national sample of adults. Also, the modules' respondents randomly overlapped with each other so that module question responses could be linked across modules with sample sizes that would be sufficient for most analyses. For example, this overlapping design allows a researcher to compare a person's music preference (from module A) with whether or not he/she took any music classes (from Module E).

The data user who plans to combine variables from different modules should review the examples shown in Section 6 of this manual.

The July 2012 timing of the interviews is another factor to consider when comparing results with prior SPPA studies. The SPPA studies in 1982 and 1992 were conducted throughout the year and the 2002 study was conducted in August and the 2008 in May. The SPPA always has asked about arts participation covering a 12-month period, but there are still potential recall differences depending on when the questions are asked. Survey research literature[2] has documented that respondents tend to report more accurately on recent events. In this case, one would expect more accurate reporting of outdoor arts festivals during the summer months while reports of theater attendance might be more accurate during the winter months. To the extent that an activity is seasonal, the reported participation rates could be impacted. However, in 1985 the funding for the SPPA was cut so that the SPPA survey was only conducted in six months rather than 12 months which lead to testing of seasonal effects using the 1982 SPPA and this research indicated no strong seasonality to the reporting of arts attendance.

---

[2] Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. The Psychology of Survey Responses. New York: Cambridge University Press, 2000.

The SPPA has a complex survey design. To produce unbiased estimates from the 2012 SPPA data, it is critical that researchers use the appropriate survey weight.

The survey weights are necessary to correct for the unequal probability of selection of adults within CPS households, to adjust for the fact that some questions did not ask for a spouse-or-partner proxy response, and to adjust to for that survey design of asking respondents to complete different forms of the questionnaire. Because of the complex design five final survey weights are included on the 2012 SPPA public-use data file. Which weight you choose to use will depend on which question you want to analyze and in what section of the questionnaire did the question get asked.

For analyzing questions in core 1 you should be using the weight variable labeled "PWOWGT" and for analyzing questions from the core 2 section of the survey you should use the weight variable "PWTWGT". Since only the first 4 questions in Module A (Other Art Activities and Music Preferences) included spouse/partner versions the questions (PEA1 through PEA4), only these 4 questions from Module A should use the weight that is labeled "PWSWGT". While the remaining questions in Module A (PEA5 through PEA8) did not include spouse partner versions of the questions and such when analyzing you should be using a unique weight variable labeled "PWAWGT". You should also use the "PWNWGT" weight variable for analyzing Module B (Accessing Art through Media), Module C (Creating Arts through Media), and Module E (Arts Learning) as each of these modules did not include questions about the respondent's spouse or partner. Finally, when analyzing questions from Module D (Creating, Performing, and Other Activities) where questions included spouse/partner version of the questions you should use the weight variable labeled "PWSWGT".

All five of the SPPA survey weights feature a Census non-response adjustment and have been post-stratified to match key demographics of the national non-institutionalized U.S. population. Consequently, the weight variable that you need to use for your analysis depends upon which module or core section of the survey the variable or variables you are interested reside in. Examples and more information on the survey weights are provided in Section 4 of this report. In section 4, we will also talk about how to use the weights when doing analysis that involves using variables from different modules. Examples of using the weights for multi-variable analysis when the variables come from different sections of the survey will covered in Section 6 of this report.

An easy way to figure out which weight variable to use would be to look the variable up in the data dictionary (Appendix A) located at the end of this report. Each SPPA variable is listed in the data dictionary which includes a description

that ends with the appropriate weight placed in parenthesis (i.e. (PWOWGT for an SPPA variable that was asked in the Core 1 section of the survey).

The 2012 SPPA public-use data file as well as earlier rounds of the SPPA will be archived at CPANDA (Cultural Policy & the Arts National Data Archive; www.CPANDA.org) and will also be available from the Census Bureau's CPS web site (www.census.gov/cps). The version of the SPPA public-use file at CPANDA has data records for only the 37,266 SPPA respondents while the Census Bureau's SPPA public-use file includes information on the entire 151,695 people who were in July 2012 CPS sample.

If you want to work with only SPPA respondents using the Census Bureau's version of the public-use file, then you need to select only those data records that either have a Core 1 weight (PWOWGT) value greater than zero or a Core 2 weight (PWTWGT) value greater than zero.

Although technically not a design change, there was an important change in how the Census Bureau structured and prepared the SPPA supplemental public use data file. In 2008, when a question had a spouse/partner version, the Census Bureau created a single variable for storing the respondent and spouse/partner responses. For 2012, the Census Bureau created separate variables for the spouse/partner responses and did not create a combined variable. So, for example the variable for going out to see a movie is "PE3A" for the respondent and "PE3B" for the spouse partner. To make the analysis easier for the researcher to produce nationally representative estimates of all adults 18 or older, the NEA version of the public use data file includes combined variables for each variable that has a spouse partner version of the question. For example the variable that can be used to estimate the number of people who went out to see a movie is "MOVIES". All the combined variables included in the NEA public-use data file are included in Appendix A, which is what we refer to as being the SPPA data dictionary. The data dictionary is arranged by section and so for those sections of the survey that had spouse/partner versions of the questions you will find combined variables at the end of the section after the listing of all the variables that are on the Census Bureau version of the public use file.

Section 3: Dealing with Missing Respondent Data

It may seem as though most of the CPS variables have no missing data.   As with all surveys, however, some CPS respondents either cannot answer or they choose not to answer all of the questions. The reason for most of the CPS variables having no missing data is this: when they are not obtained from the respondent, the answers are imputed by the Census Bureau through elaborate imputation procedure. (See *"Current Population Survey Technical Paper 66";* www.census.gov/prod/2006pubs/tp-66.pdf.)

Imputation did not occur for any of the SPPA questions. Therefore, all SPPA questions have some missing data. For instance, if you add together the number of people who said yes or no for any participation question, that sum is always less than the total number of SPPA respondents who were asked the question because of missing data. It is sometimes important to differentiate between two types of missing data: 1) data missing because the respondent does not provide a useable answer; or, 2) data missing by design because the respondent was purposely not asked the question.

The data file uses three codes to describe missing data due to a respondent not providing a useable response. When the answer to the SPPA question is coded "-3", this indicates that the respondent refused to answer the item. The ".-2" value indicates that the respondent did not know the answer, and the "-9" value indicates that a useable answer was never ascertained. The "-1" value indicates that the respondent was not supposed to answer the question based on questionnaire skip patterns.

Usually researchers will exclude respondent missing data (codes -2, -3, and -9) when calculating percentage estimates.   This practice can be thought of as a form of pseudo-imputation—with the assumption that data missing from respondents would likely show a similar response pattern as the non-missing data.

 In order to match the arts activity and participation estimates that are released by the NEA, you should exclude respondents' missing data.

Information that is missing by design (code: -1) is also typically excluded by researchers when producing percentage estimates. It is always excluded for questions that were randomly skipped, for instance for each respondent three out of the five sets of module questions were randomly skipped. For some analyses, however, the -1 can be interpreted as having a null value.   For instance, question PEC2A can tell you how many people use a computer, a handheld device or mobile device, or the internet to create music. However, if you were to

simply exclude all the missing data, it would seem that the answer would be that 28.4% used a computer … to create music, during the last 12 months. However, the answer you probably want to report is that only 1.4% of adults used a computer … to create music during the last 12 months.   This is because anyone who said "NO" to the previous question about whether you created any music in the last 12 months did not get asked about whether they used a computer to create music. In this and many similar situations, you probably would want to interpret the -1 code as having a null or "NO" value for those respondents who were not asked the question because of their response to a previous question.

Whenever you are doing statistical analysis with variables that have missing data, make sure you either understand how the missing data is being treated or include in your program explicit instructions about how missing data will be handled.   Given the large sample size for many of the SPPA variables, you may not notice the unintentional impact that values of -1, -2, -3, or -9 may have on estimates such as means, medians, and regression coefficients.

The next section of this report will show that when you exclude missing data, your weighted population estimates will no longer sum to the total U.S. adult population and, therefore, will not produce accurate population estimates. The next section also explains, however, that there are acceptable procedures for producing reliable population estimates.

Section 4: Using the Survey Weights

Responses to SPPA questions should be weighted to provide approximately unbiased aggregate, national, or regional estimates. The weights should be applied to all survey items in an effort to

- Compensate for differential probabilities of selection for households and persons
- Reduce biases occurring where nonrespondents have different characteristics than respondents
- Adjust, to the extent possible, for under-coverage in the sampling frames and in the conduct of the survey

The procedures used by the Census Bureau to develop the CPS survey weights are quite complicated. The following bullets highlight the steps used to create the weights. More information on their construction can be found in Chapter 10 of the "Current Population Survey Technical Paper 66"; www.census.gov/prod/2006pubs/tp-66.pdf.)."

The steps are:

- *Preparation of simple unbiased estimates from base weights and special weights derived from each adult's probability of being sampled*
- *Adjustment for nonresponse*
- *First-stage ratio adjustment to reduce variances due to the sampling of primary sampling units (PSUs)*
- *National and state coverage adjustments to improve CPS coverage*
- *Second-stage ratio adjustment to reduce variances by controlling CPS estimates of the population to independent estimates of the current population*

Perhaps the most important task and certainly one of the first tasks facing the data user will be determining which of the survey weights you should be using to obtain accurate population estimates.

There are five survey weight variables attached to each sample adult's data record. Accordingly, if you specify which weight variable to use, your statistical analysis software will give each adult the appropriate weight when computing frequency distributions and statistics. The public-use data file has five weight variables because not all questions included a spouse partner proxy option, the sample was randomly split so that about ½ of the sample was assigned to each

of the two cores sections of the survey, and because respondents were randomly assign to only complete two of the five additional art questionnaire modules.

Table 1 below summarizes which weight to use for each module of the survey. Also, in appendix A of this document is the SPPA data dictionary, which lists each SPPA question by module and provides both the weighted and unweighted frequency counts for each response option. You will see in parentheses at the end of the variable description (PWOWGT, PWTWGT, PWSWGT, PWAWGT, or PWNWGT) indicating which weight was used to produce the weighted estimate. Later in this section we will explain which weight to use when doing analysis that uses variables from more than one module of the questionnaire.

TABLE 1	
The Weight that Should be Used for Each Questionnaire Module	
	Appropriate Weight
Core 1 Questions (arts participation trend questions)	PWOWGT
Core 2 Questions (experimental arts participation questions)	PWTWGT
Module A1 (Other arts activities: PEA1 through PEA4)	PWSWGT
Module A2 (Music preference questions: PEA5 through PEA8)	PWAWGT
Module B (Accessing Art through Media)	PWNWGT
Module C (Creating Arts through Media)	PWNWGT
Module D (Creating, Performing, and Other Activities)	PWSWGT
Module E (arts learning)	PWNWGT

The following table (Table 2) shows the weighted and unweighted estimate for going to a film festival during the past 12 months. The shaded portion of the table shows unweighted estimates, while the non-shaded region of the table shows the weighted estimate.

TABLE 2			
Unweighted Frequency For Going to a Film Festival within the last 12 months			
Response Option	Frequency	Percent	Percent excluding missing data
1= YES	439	1.2	2.5
2= NO	17,020	45.7	97.5
Total	17,459	46.8	100.0
Weighted Frequency For Going to a Film Festival within the last 12 months (using PWTWGT)			
Response Option	Frequency	Percent	Percent excluding missing data
1= YES	5,399,608	2.3	2.4
2= NO	216,609,898	92.2	97.6
Total	222,009,506	94.5	100.0

 By checking the frequency count, you usually can tell that you are looking at a weighted estimate. A weighted frequency count will be a very large number,

since the survey weight is used to multiply each sampled response. In Table 2 (above), for example, we produce an unbiased estimate of the number (5.4 million) of adults who went to a film festival within the last 12 months.

Although a large frequency count tends to indicate a weighted estimate, the statistical output is usually not helpful in determining whether the correct survey weight was applied. Table 3 demonstrates how similar the two sets of film festival attendance estimates are when the wrong weight is applied. The shaded portion of the table shows the estimate that would have resulted by using the PWNWGT instead of the PWTWGT. The percentage estimates are so similar that even an experienced SPPA researcher may be unable to tell just from the statistical output whether or not the appropriate weight was used.

TABLE 3			
Incorrectly Weighted Frequency For Going to a Film Festival (using PWNWGT)			
Response Option	Frequency	Percent	Percent excluding missing data
1= YES	2,662,064	1.1	2.4
2= NO	108,367,384	46.1	97.6
Total	111,029,448	47.2	100.0
Correctly Weighted Frequency For Going to a Film Festival (using PWTWGT)			
Response Option	Frequency	Percent	Percent excluding missing data
1= YES	5,399,608	2.3	2.4
2= NO	216,609,898	92.2	97.6
Total	222,009,506	94.5	100.0

In May 2012, the U.S. Census Bureau estimated that the number of non-institutionalized adults 18 or older living in the United States was 234,993,563. The survey weights can be used to estimate how many of these approximately 235 million adults 18 or older attended or did a particular activity. When doing population estimates, however, you must be careful of how you handle missing data. Missing data typically occur when a person refuses to answer a particular question or does not know the answer to the question being asked. These situations are usually lumped together and classified as "missing data". Refer to the previous section of the guide for more information on how missing data for the SPPA variables have been coded.

If you do not take into account the missing data, then your population counts will total less than the overall population of 235 million adults. For instance, in our previous example, the estimated total number of adults who went to a film festival plus the number who said they did not go to a film festival equaled just over 222 million (222,009,506) because it does not account for the missing data. Also, if you do not omit the missing data, you would estimate that 2.3% of the population went to a film festival. Yet most researchers exclude missing data when estimating the percentage of the population that attends or participates in a particular activity. Hence, they would report that 2.4% of the adult population

went to a film festival. This percentage can be obtained from the final column in tables 2 and 3.

If you do decide to exclude the missing data, then a more accurate population estimate will be obtained by multiplying the percentage that excludes missing data by the total population. For instance, when we exclude the missing data from the previous example, we see that 2.4% of the adult population went to a film festival. This proportion translates to approximately 5.6 million (.024 x 235 million adults) adults went to a film festival as opposed to the estimate of 5.4 million (5,399,608) shown in tables 2 and 3. Again, this difference occurs because the population estimates in tables 2 and 3 do not adjust for the missing data.

Because of the random assignment of questionnaire modules you can do analysis that combines variables from different modules. In addition, you can also combine variables from different modules with questions asked in either the core 1 or core 2 section of the survey. However you cannot do analysis that combines variables from the core1 with variables from core2 because respondents were assigned to either complete core 1 or core 2, but never both core 1 and core2.

Table 4 summarizes which weight to use for the various possible combinations of variables in your analysis. Keep in mind, you cannot use variables from both core1 and core 2 at the same time. Also, you cannot use variables from more than two modules in the same runs, since no respondents answered more than two modules and using variables from two different modules will sometimes raise sample size concerns. Module sample sizes and combined module sample sizes are discussed in more detail in section 6 of this user's guide.

TABLE 4	
Which Weight to Use When Combining Modules	
If your analysis includes:	Appropriate Weight
Any variable from Module B, C or E	PWNWGT
Any variable from Module A2 (*PEA5 through PEA8*) and no variables from Module B, C, or E	PWAWGT
Any variable from Core 1 and no variables from Module A2, B, C, E	PWOWGT
Any variable from Core 2 and no variables from Module A2, B, C, or E	PWTWGT
Any variables from Module A1 (*PEA1 through PEA4*) or Module D and no variables from Core1, Core2, Module A2, B, C, or E	PWSWGT

Office of Research & Analysis
National Endowment for the Arts

💡 *Always remember that if your statistical run includes any variable from a module that doesn't include spouse partner versions of the questions (Module B, C, E, and Module A2), than you need to weight your runs using the PWNMWGT variable or PWAWGT if you are only using questions PEA5 to PEA8 (Module A2). This is very important because the other survey weights on the data file will produce biased estimate because they do not adjust the estimate to take into account that spouse partner are omitted from your analysis.*

🧑‍🔬 *If you are unsure whether a particular question permitted a spouse/partner response, you can always use the flag variable (labeled "PRINTSFLG" in the data dictionary) included in the public-use data file. Simply run a cross-tabulation of the variable of interest by "PRINTSFLG." If you find there are no cases, then you have a question without a spouse/partner option. Hence, you will need to weight your analysis using PWNWGT. The only exception would be If you are using a variable from Module A2 (PEA5 to PEA8) and not using a variable from Module B, C, or E, in which case you would weight your analysis using the PWAWGT weight variable.*

When your analysis includes only questions that were asked of the respondent's spouse or partner then you will most likely use the weight variable associated with the core questions (core 1 PWOWGT; core 2 PWTWGT) that you are including in your analysis. The only exception would be if you are using a variable from Module A1 (PEA1 to PEA4) combined with a Module D and not using a variable from either of the core sections of the survey, in which case you would weight your analysis using the PWSWGT weight variable.

The Table 5 cross-tabulation example shows that the likelihood of a person playing a musical instrument is strongly influenced by the level of his or her mother's education. Because the question about playing a musical instrument is asked in Module D of the SPPA, one would use the PWSWGT variable to get a frequency of playing an instrument. Yet the mother's level of education is an additional demographic question asked at the end of the arts learning (Module E) section of the survey. Thus, we would use PWNWGT to get a frequency distribution of mother's education. In this example, the correct estimates are shown in the non-shaded portion of the Table 5, following the rule of always using PWNWGT when doing analyses that include any variable from Module B, C or E. The shaded portion of Table 5 shows the incorrect estimates that would be obtained had you used the PWSWGT. More examples and guidance on doing analyses combining variables from different modules of the SPPA are included section six of this guide.

TABLE 5			
Incorrectly Weighted Cross-tabulation: Playing a Musical Instrument with Mother's Education Level (using PWSWGT)			
	Played a Musical Instrument		
Mother's Education	**YES**	**NO**	**TOTAL**
Less than 9th Grade	197,403 (5.8%)	3,186,524 (94.2%)	3,383,927 (100%)
Some High School	249,563 (10.2%)	2,197,783 (89.8%)	2,447,346 (100%)
High School Graduate	884,242 (11.1%)	7,054,680 (88.9%)	7,938,922 (100%)
Some College	539,734 (24.1%)	1,695,515 (75.9%)	2,235,249 (100%)
College Graduate	770,245 (29.1%)	1,873,038 (70.9%)	2,643,283 (100%)
Advanced Degree	176,368 (23.0%)	590,210 (77.0%)	766,578 (100%)
Total	2,817,555 (14.5%)	16,597,750 (85.5%)	19,415,305 (100%)
Correctly Weighted Cross-tabulation: Playing a Musical Instrument with Mother's Education Level (using PWNWGT)			
	Played a Musical Instrument		
Mother's Education	**YES**	**NO**	**TOTAL**
Less than 9th Grade	255,801 (6.6%)	3,602,292 (93.4%)	3,858,093 (100%)
Some High School	259,349 (10.2%)	2,273,900 (89.8%)	2,533,249 (100%)
High School Graduate	982,260 (11.0%)	7,940,610 (89.0%)	8,922,870 (100%)
Some College	562,190 (23.4%)	1,843,131 (76.6%)	2,405,321 (100%)
College Graduate	765,296 (27.4%)	2,025,588 (72.6%)	2,790,884 (100%)
Advanced Degree	211,197 (24.9%)	637,821 (75.1%)	849,018 (100%)
Total	3,036,093 (14.2%)	18,323,342 (85.6%)	21,359,435 (100%)

In the Table 5 example, both survey weights produce similar percentage estimates, but the population counts are smaller for estimates that use PWSWGT weight variable.   This occurs because PWNWGT correctly increases each sampled person's relative weight to account for the fact that the respondents were not asked about their spouse/partner's mother's education. While the percentage estimates using PWNWGT are correct, note that the total population counts sums to only 21 million adults far less than the estimated 235 million adults living in the United States. A small fraction of this difference is due to people not knowing or unwilling to answer one of these two questions. The main reason that the total population counts add to fewer than 235 million adults is that the question about playing a musical instrument is from Module D (Creating, Performing, and Other Activities) and Mother's Education comes from Module E (Arts Learning). Questions from module A, B, C, D, or E are asked only of a random 2/5's or 40 percent of all respondents and any analysis that uses questions from two different modules (such as the current example) will have had both questions asked of about 10% of all respondents. Thus the survey weights will produce population counts that sum to less than the total adult population.

Each respondent was asked questions from only 2 of the 5 modules (A, B, C, D, or E) therefore, the weighted population counts for the module questions when combined with the core questions or another module will not sum to the

total adult population and should not be used to project the total number of adults participating or doing a particular activity.

You can obtain population estimates for module question by using the weighting adjustments that are published in the Census Bureau's source and accuracy statement. These module factor adjustments are simply a single number that you multiply your population count by to get a full population estimate. These adjustments correct for the fact only a random portion of the overall sample get asked each module question, However, these module factor adjustment do not account for respondent missing data. Since all SPPA question have some respondent missing data (don't know, refused to answer, valid answer not ascertained) the population counts using these adjustment factors will still not fully sum to the total adult population. A simple strategy for projecting total number of adults would be to multiply the percentage who participate or do a particular activity by the estimated total adult population. For example, 8.1% of adults reported attending a live jazz performance, so to get the estimated number of adults who attended you multiply the 8.1 percent by 234,993,544 (estimated overall number of adults) which tells you that 19,034,477 (roughly 19 million) adults attend a live jazz performance in past 12 months.

In addition to the five survey weights on the public-use data file, the Census Bureau developed for each of these weights a set of replicate weights. Replicate weights allow the computation of replicate estimates and, more importantly, they provide a more reliable approximation of the variance of an estimate. The idea underlying replication is to draw subsamples from the sample, compute the estimate from each subsample, and estimate the variance from the variability of the subsample estimates. Subsamples from the original full sample are used to calculate subsample estimates of a parameter for which a full-sample estimate of interest has been generated. The variability of these subsample estimates compared to the estimate for the full sample can then be computed. The subsamples are called replicates or replicate subsamples, and the estimates from the subsamples are called replicate estimates.

Each replicate weight is derived using the same estimation steps as the full sample weight (i.e. PWOWGT, PWTWGT, PWAWGT, PWNWGT, or PWSWGT), but including a replicate weighting factor to each base weight. Each base replicate weight then goes through the weighting process. Once the replicate weights are developed, computing estimates of variance for sample estimates of interest is computer-intensive but straightforward. Researchers interested in using the replicate weights should contact the National Endowment for the Arts research office to obtain a copy of the replicate weights data file. Procedures for using these replicate weights are discussed in the next section ("Procedures for estimating standard errors") of this manual.

Some of the same questions were asked in both Core1 and Core2. Estimates from core 2 are generally not recommended for conducting trend comparisons. Also producing estimates based on combined responses from Core 1 and Core 2 is not recommended because of differences that could be a result of the questions being asked in a different order or subject to different skip patterns, and some questions that appear to be the same use slightly different wording. However, it may still be useful to use both the responses from Core 1 and Core 2 for analysis where you would like to have a larger sample size of respondents, keeping in mind the limitations of using the combined variable to produce population estimates.

Every respondent was asked either Core 1 or Core 2, but never both, thus they have only been assigned a value for the PWOWGT weight or the PWTWGT variable. To use responses from both core1 and core 2 you will need to create a composite weight variable that has as its value either the PWOWGT value for those who were asked core 1 or the PWTWGT value for those asked core2. For example, the questions did you read any books and if yes, how many was asked in both core 1 and core2. To get a larger sample size of avid readers (i.e. people who read more than ten books) you could create a new variable that combines the responses on books read from both core1 and core2. However, if you were to do this, then your core weight variable would be neither PWOWGT nor PWTWGT, but a composite weight variable instead. A simple way to compute a composite core weight variable would be to set the value of a new weight variable "CORE12WT" equal to PWOWGT if the value of PWOWGT is greater than zero or equal to PWTWGT if the value of PWTWGT is greater than zero.

Section 5: Procedures for Estimating Standard Errors

The sample of households and persons surveyed for the SPPA is just one of many possible samples that could have been obtained. Sampling error refers to error in survey estimates that arise from the fact that estimates are based on a sample of observations rather than the population of observations. This form of error is usually expressed in terms of the standard error of an estimate, or the square of the standard error, the sampling variance. Standard errors are also required to conduct hypothesis tests or tests of statistical significance. A clear presentation of estimates from a survey or hypothesis test should include measures of uncertainty associated with using a sample for inference, as opposed to using the entire population.

This section explains the process of obtaining standard errors for SPPA estimates. The SPPA sample and respondents are subsets of the CPS sample and respondents, which is a sample design that includes stratification and clustering. Although survey estimates obtained from the default options in most statistical packages will be correct, the standard error estimates will often understate the true standard errors because they do not account for the survey design.

Stratification generally leads to a gain in efficiency over simple random sampling. On the other hand, clustering usually leads to deterioration in efficiency. This latter effect arises because of the positive intra-cluster correlation among the subunits in the clusters. For example, the cluster effect is larger for larger households because we sometimes sampled more than one adult from the same household. This clustering effect increases the variance over what would pertain in a simple random sampling of adults.

To determine the total effect of any complex survey design on the sampling variance, you first calculate the variance associated with an estimate assuming a complex sample design. Then you calculate the variance you would expect from a simple random sample design. The ratio of the complex variance estimate over the variance associated with a simple design is what we call the design effect, often referred to as the DEFF, and it measures the overall efficiency of the sampling design.

In a wide range of situations, the adjusted standard error of a statistic should be calculated by multiplying the usual formula by the square root of the DEFF. Thus, the formula for computing the 95% confidence interval around a percentage is:

$$\hat{p} \pm \left(deft \times 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

where $\hat{p}$ is the sample estimate, n is the unweighted number of sample cases in the group being considered and, deft is the square root of DEFF.

The remainder of this section discusses how to obtain and use design effects. The following are five recommended ways in which you can get an estimate of the 2012 SPPA design effects. This list starts with the easier but less precise methods and ends with the more difficult but more precise methods.

1. Depending upon which section of the questionnaire your variable is from, choose the appropriate design effect from table 6.
2. Use the tables and procedures for obtaining design effects described in the Census Bureau's Source and Accuracy Statement for the July 2012 Arts Supplement.
3. Use a Taylor series linearization approach to approximate the design effect separately for each estimate.
4. Use the actual replicate weights that the Census Bureau created to estimate the overall average design effect to estimate design effect separately for each estimate.

The U.S. Census Bureau does provide replicate weights that can be used to obtain standard errors reflecting the complexity of the sample design. However, for researchers who may not have access to the necessary computer hardware and software or technical ability to use these replicate weights to calculate standard errors appropriately, we provide instructions on how to use average design effects to obtain approximate standard errors for survey estimates.

Technically, each variable has its own design effect. One way to represent the approximate impact of the design effect is to compute design effects for a number of similar variables and then average those numbers to produce an overall estimated average design effect. Table 5 shows the average design effects by questionnaire module.

TABLE 6	
Overall Average Design Effects by SPPA Questionnaire Module	
	Average Design Effect
Core 1 Questions (arts participation trend questions)	2.1
Core 2 Questions (experimental arts participation questions)	2.1
Module A1 (other arts participation PEA1 – PEA4)	2.2
Module A2 (musical preferences PEA5 – PEA8)	2.3
Module B (accessing art through media)	2.3
Module C (creating arts through media)	2.3
Module D (creating, performing, and other activities)	2.2
Module E (arts learning)	2.3

Table 7 uses the estimated design effects shown in Table 6 and generates standard error estimates and 95% confidence intervals for select 2012 SPPA variables.

TABLE 7						
Variance, Standard Error and Confidence Intervals for Select 2012 Variables						
	Percent	Variance	Design Effect	Adjusted Variance	Adjusted Standard Error	95% Confidence Interval
Jazz	8.1	.07	2.1	.15	.39	7.3 – 8.9
FILMFESTIVAL	2.4	.02	2.1	.04	.2	2.2 – 2.6
PEA74	26.6	.20	2.3	.46	.68	25.9 – 27.3
PEC3A	1.3	.01	2.3	.02	.14	1.2 – 1.4
KNIT	13.2	.11	2.2	.24	.49	12.7 – 13.7
PEE7A	17.6	.15	2.3	.35	.59	17.1 – 18.2

JAZZ = Attended a live jazz performance during the last 12 months
FIlmfestival = Go to a film festival during the last 12 months
PEA74 = Like to listen to Jazz
PEC3 = Create or perform any dance during the last 12 months
KNIT = Do any weaving, crocheting, quilting, needlepoint, knitting, or sewing during the last 12 months
PEE7A = Ever taken any lessons or classes in art appreciation or art history

Using the estimated overall design effects will greatly improve upon the standard error estimates associated with your SPPA estimates. Still, it is important to keep in mind that each estimate has its own design effect. Therefore, the design effect for attending jazz may be higher or lower for, say, men versus women or for any other subgroup of the population. If getting more precise standard estimates is a concern, then you should read the remainder of this section, which describes how to use replicate or linearization methods to estimate standard errors. Another option is to use the Census Bureau's Source and Accuracy Statement for the July 2012 CPS Arts Supplement, which includes tables and procedures that can be used to improve the design effect estimates for different subgroups. The Census Bureau's Source and Accuracy Statement is included as Appendix C.

A more precise way of calculating standard errors for SPPA estimate is to use the 160 replicate weights that the Census Bureau has created separately for each of the fives SPPA survey weights. The basic idea behind replication is to draw subsamples from the sample, compute the estimate from each of the subsamples, and estimate the variance from the variability of the subsample estimates. Specifically, subsamples of the original full sample are selected to calculate subsample estimates of a parameter for which a full-sample estimate of interest has been generated. The variability of these subsample estimates around the estimate for the full sample provides an estimate of the standard error

of the estimate. The subsamples are called replicates and the estimates from the subsamples are called replicate estimates.

Replicate weights are created to derive the corresponding set of replicate estimates. Each replicate weight is derived using the same estimation steps as the full sample weight but applying a replicate factor to each case before the weighting process begins.

Once the replicate weights are developed, it is a straightforward matter to compute estimates of variance for sample estimates of interest. Imagine using each of the 160 replicate weights, one replicate weight at a time, to obtain 160 separate weighted estimates of the same statistic, such as a mean. Take these 160 estimated means and calculate the sum of the squared deviations from the mean that was estimated using the full sample weight. This sum of the squared deviations from these 160 means divided by 160 (the number of replicates) times 4 will give you the estimated variance. In turn, the standard error is the square root of the estimated variance.

Although the logic behind using replicate weights is not unduly complicated, it is fairly computer-intensive to produce standard errors using the replicate weights. To use the replicate weights, you need to either use specialized software designed to make use of replicate weights when generating standard errors— examples include SUDAAN and WesVar— or you should use specialized advanced sampling modules in software such as STATA, SAS or SPPS.   The replicate weights are not included in the public-use data file, but they can be obtained upon request from the NEA's Office of Research & Analysis. More information about using the replicates may also be included in the Census Bureau's forthcoming July 2012 CPS Arts Supplement technical documentation on the Bureau's CPS website; www.census.gov/cps.

Another way of obtaining standard errors for SPPA estimates is to use a Taylor series linearization approach. The Taylor series approximation (or any other related linearization method) uses the full sample weight along with variables describing the structure of the replicate weights (PSU and Stratum) to obtain corrected standard errors. Users interested in a linearization method can opt to use SUDAAN, the "SVY" commands in STATA, the "PROC SURVEYMEANS" and "PROC SURVEYREG" commands in SAS, or the "CSELECT" complex samples procedures in the SPSS complex samples module. In addition, there are many online statistical tools such as the Princeton University's Cultural Policy & the Arts National Data Archive (CPANDA) that can produce adjusted standard error estimates based on a Taylor series approximation.

Section 6: Multi-variable Analyses that Combines Questions from Different Modules

Analysis of more than one question or variable at a time is often referred to as performing multivariate statistical analysis. There are two key issues that you should account for in doing multivariate analysis with the 2012 SPPA public-use data. First, there is the issue covered in the previous section – deciding which survey weight to use – as the data file contains two survey weights. Again, it is important that you weight your analysis with the correct weight variable when combining variables even those variables which, by themselves, would use a different weight.

The other important issue to consider when running multivariate analysis is whether you have a large enough sample size, given that respondents received different versions of the survey based on which set of questionnaire modules they were randomly assigned to answer. Table 8 shows the various sample sizes that your analysis will be drawn from, depending on which combination of questionnaire modules you decide to use in your analysis. For example, if you wanted to know what percentage of bluegrass music-listening adults (from Module A) also attended a sporting event last year (from module D); your sample size for that analysis would be 10,153 respondents.

TABLE 8 SPPA 2012 SAMPLE SIZES							
	Core 1	Core 2	Module A	Module B	Module C	Module D	Module E
Core1 (arts participation trend questions)	18,803	0	7,512 5,094	5,114	5,218	7,512	5,159
Core 2 (experimental arts participation questions)	0	18,463	7,413 5,059	4,983	5,043	7,413	5,102
Module A (other arts participation and music preference) [3]	7,512 5,094	7,413 5,059	14,925 10,153	2,513	2,573 2,573	14,925 10,153	2,585 2,585
Module B (accessing art through media)	5,114	4,983	2,513	10,097	2,502	2,513	2,549
Module C (creating arts through media)	5,218	5,043	2,573 2,573	2,502	10,261	2,573	2,595
Module D (creating, performing, and other activities)	7,512	7,413	14,925 10,153	2,513	2,573	14,925	2,585
Module E (arts learning)	5,159	5,102	2,585 2,585	2,549	2,595	2,585	10,261

---

[3] The top number in Module A indicates the number of individual responses for questions PEA1 to PEA4 which included spouse/partner responses. The bottom number represents the number of responses to the questions PEA5 to PEA8.

All the cells in Table 8 are greater than 2,500 — which demonstrates the strength of the survey design[4]. The samples sizes are such that you can easily compare responses to a question from one module with the responses to a question from another module. However, you do need to be careful when you further break down the data by small demographic groups. For instance, your combined module sample sizes often will not be large enough to allow you to draw conclusions about different response patterns observed in small geographical areas nationwide.

When working with survey weights, it is always a good practice to run your analysis as both weighted and unweighted. While you should not report unweighted estimates, the unweighted runs will show you exactly how many respondents your weighted estimate is based on.

Table 9 shows that people who play a musical instrument are much more likely to attend a film festival (5.9%) than the overall adult population (2.4%). This is an example of your most basic type of multi-variable analysis, often referred to as two variable or bi-variate analyses. This is also an example of an analysis that uses one variable from a SPPA module (Module D) combined with a core variable (Core 2); hence, the sample size for this run is 7,413. Because the survey weights do not produce full population estimates (see the prior section for details on obtaining population counts), this table reports only the percentage estimates.

Table 9			
Weighted Cross-tabulation: Plays a Musical Instrument with Attending a Film Festival (n=7,413) (weight=PWTWGT)			
Plays a Musical Instrument			
Attended a Film Festival	**YES**	**NO**	**TOTAL**
YES	5.9%	1.9%	2.4%
NO	94.1%	98.1%	97.6%
Total	100%	100%	100%

Table 10 shows that people who play a music instrument are almost three times more likely to take photographs as an artistic activity (32.2%) than the overall adult population (12.3%). This is an example of an analysis that uses variables from two different modules (Module C and D). Thus, the sample size for this run is 2,573 which is much smaller than you would get from an analysis that uses variables from the same module only.

---

[4] The margin of error due to sampling at the 95% confidence interval for any estimated proportion based on the sample size of 2,500 would be at most ±2%.

Table 10			
Weighted Cross-tabulation: Plays a Musical Instrument with Taking Photographs as an Artistic Activity (n=2,573) (weight=PWNWGT)			
Plays a Musical Instrument			
Take Photographs as an Artistic Activity	**YES**	**NO**	**TOTAL**
YES	32.2%	9.9%	12.3%
NO	67.8%	90.1%	87.7%
Total	100%	100%	100%

💡 *Since each respondent was asked to complete only two of the questionnaire modules, it is impossible to do analyses that combine variables from more than two questionnaire modules.*

It is possible to do analysis that combines variables from two different modules with questions from the core section of the survey, since all respondents were asked either the Core 1 or Core 2 questions. Table 11 is an example of a run that combines two modules with a question from the Core 1 section of the survey. In this example, you can see that the combination of playing a musical instrument and taking photographs as an artistic activity is even greater for those who report visiting an art museum (39.9%) versus those who did not report visiting an art museum (16.2%). Notice, as expected, the sample size for the three way cross-tabulation is much smaller (n=1,241) than the sample sizes in tables 9 and 10.

Table 11			
Weighted Three Variable Cross-Tabulation: Plays a Musical Instrument with Taking Photographs as an Artistic Activity (Controlling for Whether or Not the Person Visited an Art Museum During the Past 12 Months) (n=1,241) (weight=PWNWGT)			
Plays a Musical Instrument **(Visited an Art Museum During the Past 12 Months)**			
Take Photographs as an Artistic Activity	**YES**	**NO**	**TOTAL**
YES	39.9%	27.6%	30.2%
NO	60.1%	72.4%	69.8%
Total	100%	100%	100%
Plays a Musical Instrument **(Did not Visit an Art Museum During the Past 12 Months)**			
Take Photographs as an Artistic Activity	**YES**	**NO**	**TOTAL**
YES	16.2%	6.7%	7.5%
NO	83.8%	93.3%	92.5%
Total	100%	100%	100%

Many researchers will no doubt want to use the SPPA data to explore relationships between variables. Regression analysis is perhaps the most

common statistical tool for doing multivariate analysis that explores relationships between groups of variables. It is beyond the scope of this users-guide to give advice or examples of various regression techniques that could be done using the SPPA data. Given the modular design of the SPPA questionnaire, however, a researcher should be careful in choosing which and how many independent variables to put into their models in order to ensure adequate sample size.

While the weights are needed for producing unbiased population estimates, the decision to use them for regression analysis is more complicated. When doing regression analysis, researchers often prefer to try and control for the main variables (gender, age, race, and region) used to create the survey weights rather than use the weights in order to improve the precision of their estimates. However, because of the complexity of the SPPA survey weights, it is strongly recommended that you at least run your regression both with and without using the survey weights to see if the results differ.   And, in general, it is probably preferable to use the weights in regression analyses when your dependent variable is strongly correlated with the weighting factors. The following is a good article on this subject: "Sampling Weights and Regression Analysis" by Christopher Winship (Harvard University) and Larry Radbill (Joint Economic Committee of Congress) in the Sociological Methods and Research Journal *23:2 (November 1994): 230-257.*

Section 7: Useful CPS variables

A major advantage of surveys that are conducted as part of the CPS is their access to a rich assortment of personal, household, and geographic variables. These variables are already included on the 2012 SPPA data file and are listed in the SPPA data dictionary (Appendix A). Most users will be more than satisfied with simply using the CPS variables that are already included on the data file. Others may want additional CPS variables from the March 2012 Annual Social and Economic (ASEC) Supplement. The samples of SPPA households were in the exit rounds of their CPS commitment; all therefor only one-quarter of the sample was in the CPS sample rotation in March or April 2012. Therefore, a user can merge variables from the 2012 SPPA survey with the 2012 ASEC only for those SPPA households that were in their final month of their CPS commitment. This manual explains how the procedure is done.

The 2012 SPPA public-use file includes all of the most commonly used personal and household demographic questions. Since the CPS survey is used to measure the U.S. employment situation, there are also an abundance of variables about people's employment. There are more CPS variables on the 2012 SPPA public-use data file than there are SPPA supplement variables. Given the large number of CPS variables on the data file, it would be wise to search the 2012 SPPA data dictionary before investing time searching for variables that you would like to merge on to the data file.

Due to federal data disclosure rules and confidentiality concerns, you cannot obtain geographic variables other than those already released as part of the 2012 SPPA public-use data file. The public-use versions of the ASEC, and all other CPS supplements, omit provide geographic variables that let you analyze smaller regions of the county. Furthermore, the sample design and methods of weighting CPS data are geared toward producing estimates for the entire Nation. Consequently, estimates for states and smaller areas of the country are not as reliable as national estimates.

While there are 385 CPS variables on the 2012 SPPA public-use data file, there is still a wealth of variables that you can merge on to the SPPA file from the March 2012 ASEC Supplement. The additional household variables that are available from the ASEC include more variables on the economic characteristics of the household, dwelling characteristics, information about appliances, and many other variables that describe the physical characteristics of a person's household. The additional personal variables available from the ASEC include family interrelationship variables, and more detailed ethnicity, education and income variables. Also, the ASEC includes tax, disability, migration, health

insurance, welfare, and poverty variables not on the 2012 SPPA public-use data file.

So, how do you merge the ASEC variables on to the 2012 SPPA public-use data file? To merge household-level variables, you need to match the 2012 SPPA and ASEC file by *household*. To merge family- or person-level variables, you need to match by *household respondent*. To match households, you must ensure that the HRHHID (household identifier 1 variable) variables on the 2012 SPPA data file have the same value as the H-IDNUM1 variable on the 2012 ASEC data file. To correctly match the respondents, you must make sure that the HURESPLI (respondent line number) variable on the 2012 SPPA data file has the same value as the H-RESPNM on the 2012 ASEC data file—but you also need to ensure that the household id variables match.

When merging two data files containing many of the same respondents, it is always important that the two data files be sorted by the variables that you will use to find the matches.

Anyone who has routinely merged data files can attest that no matter what software program you use to carry out the match and merger, incorrect mergers do occur and sometimes can go undetected. Therefore, a good strategy for verifying that your merger was done properly would be to include in your merger a variable that exists on both the 2012 SPPA data file and the file you are merging from. For instance, you could merge the respondent's gender from the ASEC on to the 2012 SPPA data file. Then all you need to do to ensure that the merger was done correctly is verify that values for the SPPA gender variable are the same as the values for the ASEC gender variable. Be careful not to over-write the original variable when merging variables that exist on both files.

Office of Research & Analysis
National Endowment for the Arts

Section 8: Comparing 2012 to Earlier SPPA Estimates

The comparison of SPPA estimates over time is something that many researchers often do. The 2012 SPPA was designed to maintain this important capability. As with any survey, however, there are limits to how much a change can be considered "real" and not reflect larger differences in the sampled population or in the methodologies used for each round of data collection. This section of the user's guide will describe key factors that should be considered when comparing the 2012 SPPA with earlier SPPA studies. At the end of this section are a couple of examples of how to use the SPPA combined data file that the NEA provides for making it easier to do trend analysis on the benchmark Core 1 SPPA variables.

Due to the considerable differences in survey methodologies, the 1997 SPPA telephone survey produced results that are not comparable to the 2012 SPPA or any of the other SPPA surveys. The 1997 survey should be analyzed only as a stand-alone, snapshot survey. It should not be used in research that examines change over time. Hence, the discussion and considerations covered in this section are applicable only to comparisons done between the 2012 SPPA and earlier SPPA studies other than the 1997 SPPA.

To compare 2012 estimates with prior estimates, you need to obtain a copy of earlier questionnaires and compare the wording of those questions you plan to analyze with the 2012 wording found in Appendix B of this user's guide. Differences in question wording do not necessarily mean that you cannot compare changes in estimates over time. Such changes may have been necessary to make the questions appropriate for the current population. For instance, if you want to analyze arts participation through the media you will notice that the 2012 questions have been modified to reflect the growing use of the Internet and other electronic devices used to watch or listen to arts performances. In this situation, the wording changes likely improved a researcher's ability to make comparisons that would have suffered had the wording of the media questions been exactly the same as in prior years. Admittedly, the impact of wording changes is a matter for subjective judgment. Even if the perceived impact is minor, it is generally good practice to acknowledge in an endnote or footnote when there are wording differences.

Again, we are not recommending that you make any comparisons with the 1997 SPPA. Comparisons with all other SPPA studies benefit from the fact that the survey sampling was done by the Census Bureau.   More importantly, all of the SPPA samples are weighted so that they all provide a representative snapshot of the non-institutionalized U.S. adult population 18 years of age or older. As discussed earlier in this user's guide, however, there are some differences in the

sampling procedures. In the 2008 and 2012 SPPA, there was a random selection of adults rather than an attempt to interview all adults in the household – as was done in the earlier Census SPPA studies. Also, the first three SPPA surveys were conducted as a supplement to the National Crime Victimization Survey (NCVS). Although both the NCVS and CPS are based on a multistage, stratified, random sample of households, the two sampling designs do differ.

Now, the good news: the impact of these design differences—regarding how households or people within households are sampled—has already been measured by the Census Bureau. This measurement is called the survey design effect and was discussed in more detail in section 4 of this user's guide. Once you have adjusted the variance of your estimates by the appropriate design effect (this is done by simply multiplying the variance by the design effect), it is straight forward to estimate change over time. This benefit accrues in part because each SPPA study is an independent sample of the U.S. adult population.

To understand the simplicity of estimating the significance of changes over time, consider estimating a proportion or count at time t — say, t θ_t. Let $v(\theta_t)$ be its estimated variance (the square of the standard error). The estimated change between times t1 and t2 for this proportion or count is $\Delta = \theta_{t1} - \theta_{t2}$. The variance of the difference is the sum of the variances for the two time periods, which is $v\Delta = (v\theta_{t1\_} + v\theta_{t2})$. The two variances on the right side of the equation should be computed separately. To get the standard error of the differences between the two estimates you would then take the square root of $v\Delta$. If the difference between the two estimates is greater than 1.96 times the standard error of the differences, then you can say with 95% certainty that the differences between the different SPPA estimates are significant.

Table 12 provides a real example of how you would go about determining whether a change in an estimate is significant. In this example, we see that the percentage of adults who reported attending a live classical music performance dropped from 2002 to 2012 by 2.8 percentage points. Is that a statistically significant change?

The first step we need to take to answer this question is to sum the variances of the two estimates. The sum would be equal to 0.18 (third column) — except that we also need to adjust the variance estimate to account for the design effect. So, to answer our question, we really need to sum the adjusted variances (variance * design effect), to arrive at 0.46. The next step is to estimate the standard error of the difference, which is the square root of the adjusted variance, or 0.67. Finally we can build a confidence interval by multiplying the adjusted standard error by 1.96 (this step gives you a 95% confidence interval), and then adding and subtracting that number to and from the 2.8 percentage point change that occurred between 2002 and 2012.

The answer to our hypothetical question is yes. We are at least 95% confident that there was a decline in live classical music attendance between 2002 and 2012, as zero (no change) is not within the confidence interval (1.0 to 3.6).

	Percent	Variance	Design Effect	Adjusted Variance	Adjusted Standard Error	95% Confidence Interval
TABLE 12						
Testing to see if the change in attending a live classical music between the 2002 and 2012 was statistically significant at the 95% confidence interval						
2002	11.6	.10	2.8	.28	.53	10.6 - 12.6
2012	8.8	.08	2.3	.18	.42	8.0 - 9.6
(2002-2012)	2.8	.18	NA	.46	.67	1.5 - 4.1

Table 13 provides the estimated overall design effect for all SPPA studies. In addition to the overall average design effect, the Census Bureau does provide in their Source and Accuracy Statements, tables and procedures that produce more precise design effect estimates for different questions or population subgroups. However, these tables and procedures for calculating more precise design effects are available only for the 2002, 2008, and 2012 SPPA studies.

TABLE 13	
Overall Average SPPA Design Effects by Year	
	Average Design Effect
1982	1.8
1985	2.0
1992	4.8
1997	1.6
2002	2.8
2008	2.9
2012	2.3

If the 2012 variable that you are interested in has a higher or lower design effect than the overall 2012 average design effect, then it is somewhat likely that the earlier estimates for this variable had a higher or lower design effect. This is because certain variables are likely to have higher or lower response correlations within a household which leads to higher or lower design effects and this is likely to hold true over time. There is less concern if the variable has a lower design effect than the overall design effect since this would likely result in your test of differences over time being more conservative.

When comparing responses over time, there are a few other things worth thinking about besides question-wording and sampling. For instance, even though the Census Bureau performs a nonresponse adjustment, differences in SPPA response rates may explain small changes in the estimates. There is not any strong evidence as to how differences in response rate may impact estimates, but one can surmise that it may be harder to interview people who are less involved in the arts and hence a higher response rate may in fact slightly lower arts participation estimates.

Another factor is that more of the earlier SPPA studies were done in person rather than over the phone. Again, this probably has had minimal or no impact on the estimates, but some people would argue that social desirability (pleasing the interviewer) is higher for in-person interviews and that in-person respondents may over-report participation relative to phone respondents.

There have been some researchers concerned with the impact of earlier SPPA studies being a supplement to the crime victimization survey versus more recently becoming a supplement to the CPS. However, it would seem to be a big stretch to imagine how either of these two survey topics (crime or labor force activity) would likely have any influence on how people report participating in the arts. Another potential concern is that the SPPA CPS supplemental interviews are all completed within a few weeks while the earlier SPPA crime-victimization-survey supplement was done over the course of a full year. To the extent that an activity is seasonal, the reported participation rates could be impacted. However, examination of the monthly reports from the 1982 SPPA indicated no strong seasonality to the reporting of arts attendance[5]

The decision in 2008 and 2012 to collect participation information from the spouse or partner is a change that may have led to slightly lower estimates for a few activities. For most variables the spouse/partner proxy reports yielded slightly lower participation rates, but usually too small to have any impact on the overall participation estimates. In fact, for the most part estimates from spouse/partner proxy reports were remarkably similar to self-reported estimates. However, there were a few SPPA estimates that may have been slightly higher had there been only self-reported data (reading, visiting art museums, exercise). Some research into the impact of spouse/partner did find a few gender differences. But no differences were found when comparing spouse/partner proxy reporting versus self-report by age, race, and region.

Finally, since researchers often compare SPPA estimates over time the National Endowment for the Arts provides a SPPA data file that combines data from 1982, 1985, 1992, 2002, 2008, and 2012. This combined data file can be obtained by

---

[5] *Public Participation in the Arts: Final Report on the 1985 Survey*, Survey Research Center, University of Maryland, 1985, ERIC # ED 264168.

contacting the NEA research office or can be accessed <mark>online at the CPANDA arts data file archive.</mark>

This file only includes Core 1 2012 questions, since these are the only questions that have been asked consistently since 1982. In addition to the Core SPPA variables the file includes demographic and geographic variables. Most importantly the data file includes two weight variables: "WEIGHT", which is the weight variable to use when producing estimates for a particular year or comparing years; and "WEIGHT_NORMALIZED", which is the weight variable to use when pooling or combing data across years. There are also two variables— "VARSTRAT" and "VARUNIT"—that are included to help people estimate the variance or standard error associated with the estimates. If you decide use this combined data file it is highly recommended that you also obtain a copy of the data dictionary that describes each variable and provides more detail on the weighting and variance estimation procedures.

Table 14 uses the combined data file, and shows us that the number of people who reported that they attended a live ballet performance with the last 12 months has declined from 4.2% in 1982 to 2.7% in 2012. This is an example of your most basic type of trend analysis, which uses the weight variable "WEIGHT" to produce the year by year estimates. Notice that the sample size for this run is 98,047 which is the total number of completed SPPA interviews since 1982, excluding the 1997 SPPA, but including the 2008 and 2012 spouse/partner proxy interviews.

Table 14			
Weighted Trend analysis: Attended a Live Ballet Performance (n=98,047) (weight=WEIGHT)			
Attended a Live Ballet Performance during the past 12 Months			
Year	YES	NO	TOTAL
1982	4.2%	95.8%	100%
1985	4.3%	95.7%	100%
1992	4.7%	95.3%	100%
2002	3.9%	96.1%	100%
2008	2.9%	97.1%	100%
2012	2.7%	97.3%	100%

Table 15 provides an interesting example of why a researcher might want to pool data across years. The sample sizes for the SPPA studies are generally sufficient for producing national and regional estimates but not larger enough in any particular year to produce statewide estimates. In fact, it becomes even more difficult to look at state or county estimates when participation rates are quite small such as attending the opera or a ballet. However, if we were to pool data across all rounds of the SPPA, then we would have at least 100 observations in every state including Washington D.C.   Therefore, we could look to see (as shown in Table 15) which states over the past 30 years has had the

highest percentage of people reporting that they attended a ballet in the past 12 months (i.e. Washington D.C with 8.3%) and which state has had lowest (i.e. West Virginia with 0.9%). Analysis that pools data across year requires that the researcher use the weight variable labeled "WEIGHT_NORMALIZED" to produce pooled estimates. Notice that the sample sizes for each state shown in the final column are all greater than 100.

Table 15			
Weighted Pooled data analysis: Attended a Live Ballet Performance by State (weight=WEIGHT_NORMALIZED)			
States with the highest percentage of people attending a Live Ballet Performance			
State	YES	NO	Sample Size
Washington DC	8.3%	95.8%	109
Washington	7.0%	93.0%	1,165
Colorado	5.4%	94.6%	831
Montana	4.6%	95.4%	175
Alaska	4.4%	95.6%	113
States with the lowest percentage of people attending a Live Ballet Performance			
West Virginia	0.9%	99.1%	338
Indiana	1.1%	98.9%	1,120
Arkansas	1.2%	98.8%	506
Missouri	1.2%	98.8%	1,025
Arizona	1.5%	98.5%	1,044