

**Collier County Jewish Community Study, 2017**  
**Cohen Center for Modern Jewish Studies, Brandeis University**  
**Documentation of Public Use Data Set**

*For more information contact Matthew Boxer, [mboxer@brandeis.edu](mailto:mboxer@brandeis.edu)  
Version July 20, 2017*

## **Introduction**

This document describes the Public Use Data Set for the 2017 Collier County Jewish Community Study. It explains the constructed variables in the dataset and the procedures for statistical weighting.

Variables that begin with the prefix “x\_” were constructed from original survey responses, some of which are not included in the public data set (this is noted in the descriptions where applicable). All other variables are documented in the codebook (Appendix B of the report) and are taken directly from the survey.

You can find the report and technical appendices at:

<http://www.brandeis.edu/ssri/pdfs/communitystudies/CollierCountyJewishCommStudy.pdf>

## **About the Public Use Data Set**

The Public Use Data Set contains all raw data provided by respondents except for data that might be used to identify individual respondents. Primarily, these removed variables were open-ended responses and household ZIP codes.

Descriptions of some variable naming conventions follow:

1. All variables with a “resp” in the name refer to the respondent (e.g., respage is the age of the respondent).
2. Variables with a “hhad” prefix refer to non-respondent adults in the household; these variables range from 2-5 because the respondent is considered as the first adult (e.g., hhadage3 is the age of the third adult in the household).
3. Variables with a “hhch” prefix refer to the children in the household; these range from 1-5 (e.g., hhchgrd1 is the grade-level of the first child in the household).
4. Variables with a “jed” prefix refer to Jewish education; those with a suffix of “ad2” “ad3” or “ad4” refer to adults still in high school; those without the “ad” suffix refer to children 17 and younger.

## *Potential Issues with Data Interpretation*

CMJS as a rule maintained the integrity of the data as collected. As such, one potential issue warrants caution. Due to programming conventions, both refusals and questions skipped through survey logic are coded the same way, as system missing. For example, those for whom charany=0 were not asked charanyamt.

It is up to the analyst's own interpretation as to how to account for these characteristics of the data set.

### **Definitions of Constructed Variables (x\_ prefix)**

Variables were constructed from raw data for two purposes:

1. Standardized coding of open-ended data
2. Analytical variables created through complex combinations of multiple raw variables. The syntax used to create complex variables will be included in the report appendix.

**x\_collier:** Respondents supplied the ZIP code of their primary residence in the Greater Naples area. To protect their identities, this variable groups ZIP codes into Collier County and not-Collier County.

**x\_hhchrelrsd1-5:** In the original survey, after identifying the religion in which the first child is being raised, respondents were asked if all children are being raised in the same religion (see variable **hhchrelsame**). These variables fill in the responses for children 2-5 who have the same religion as child 1.

**x\_localsyn:** The survey instrument asked respondents if their household belonged to each of the six synagogues based in Collier County. To avoid potentially revealing identifying information, this variable denotes if the household belongs to one of these six congregations.

**x\_hhjewish:** If the household is Jewish. In this dataset, all households include at least one adult Jew. This variable should only be used to estimate the number of households.

**x\_respjewish; x\_hhadjewish2-5:** These variables denote if the respondent, adult, or child is Jewish. Jews are defined as being Jews by Religion (JBR), Jews of No Religion (JNR), or Jewish of Multiple Religions (JMR). JBRs are those who say their religion is Jewish and have a Jewish background (i.e., a Jewish parent, was raised Jewish, or converted). JNRs are atheists and agnostics who consider themselves Jewish aside from religion and have a Jewish background, or say they are Jewish and atheist/agnostic, and have a Jewish background. JMRs either say they have two religions, one of which is Judaism, and have a Jewish background, or have another religion but consider themselves Jewish aside from religion, and have a Jewish background.

**x\_hhchjewish1-5:** Children are counted as Jews if they are being raised Jewish by religion, culturally Jewish, or Jewish and another religion (corresponding to the variables **x\_hhchrelrsd1-5**).

**x\_respmarjewtype:** This notes if the respondent's live-in spouse, fiancé/e, or significant other is Jewish (i.e., JBR, JNR, or JMR).

x\_hhmarjewtype: This notes if the household contains an inmarried or intermarried couple, or no couple (whether or not the respondent is part of the married couple).

x\_invinindex: This is an index created based on involvement in the local Jewish community, based on factors of organizational membership, programmatic participation, and prosocial behavior. Details on the construction of this variable can be found in the explanation of the Comparison Chart.

x\_primseas: This variable notes if a household is in the Naples area year-round, if it is seasonal and Naples is the primary residence, or if it is seasonal and Naples is not the primary residence. It combines data from the variables locsechome and locsecprim.

x\_naplesyrs10: This notes if the household has been living, year-round or part-time, in the Naples area for less or more than 10 years. It combines the cleaned variables locyears and locyearsseason (the originals of these variables are in the dataset).

## Weighting

Four weight variables are available for this dataset. Two are at the household level (x\_wtprimhh and x\_wtfullhh) and two are at the respondent level (x\_wtprimresp and x\_wtfullresp). Household-level weights should be used to calculate characteristics of the household, population counts, and anything involving children. Respondent-level weights should be used to calculate characteristics of respondents (e.g., behaviors and attitudes).

The weight variables are also segmented by whether they refer to the primary sample (x\_wtprimhh and x\_wtprimresp) or the full sample (x\_wtfullhh and x\_wtfullresp). Primary weights are the only ones appropriate for generating counts. The full-sample weights are appropriate only for generating characteristics.

Note: The public-use dataset includes all screener data, but households that screened out of the survey have all their weights set at 0. Weighting instructions are designed for use with Stata.

## Constructed variables for weighting

x\_strataemail: The strata identified is used for weighting.

x\_wtprimhh: The primary-sample household weight.

x\_wtfullhh: The full-sample household weight.

x\_wtprimresp: The primary-sample respondent weight.

x\_wtfullresp: The full-sample respondent weight.

## Primary Weights

*Use ONLY the primary sample for generating population counts*

For estimations on the number of households use wtprimhh. For estimations on the number of people—i.e., counts—use the household weights with totals of count variables—e.g., hhadnum, hhchnum.

```
svyset _n [pweight= x_wtprimhh], strata(x_strataemail) vce(linearized) singleunit(missing)
```

Primary weights should be used for generating counts of the population as a whole.

## *Example: Counts of Adults*

The number of adults in the community:

```
svyset _n [pweight= x_wtprimhh], strata(x_strataemail) vce(linearized) singleunit(missing)
svy: total hhadnum
```

## Full Weights

*DO NOT use the full weights to generate population counts.* Full weights should be used for characteristics of the population (e.g., percentage of synagogue-member households, or percentage of synagogue-member households that belong to other Jewish organizations)

```
svyset _n [pweight= x_wtfullhh], strata(x_strataemail) vce(linearized) singleunit(missing)
```

```
svyset _n [pweight= x_wtfullresp], strata(x_strataemail) vce(linearized) singleunit(missing)
```

### *Example: Household Subpopulations*

The proportion of synagogue-member households that also belong to Jewish organizations:

```
svyset _n [pweight= x_wtfullhh], strata(x_strataemail) vce(linearized) singleunit(missing)  
svy: tab orgmem rlsynany, col
```

### *Example: Respondent Attitudes*

The proportion of those who feel connected to the Greater Naples Jewish community:

```
svyset _n [pweight= x_wtfullresp], strata(x_strataemail) vce(linearized) singleunit(missing)  
svy, subpop(x_respjewish): tab jlconnnap
```