**2014-2016 NSDUH: SAS files Containing Substate Region Definitions**

The SAS and Excel files described below consists of substate region definition variables along with Census geography variables (state, county, and tract). These files can be used to merge 2014-2016 substate region definitions onto any data file containing state, county and census tract codes (based on 2010 Census geography). Note, the SAS and Excel files are identical.

Four Files (two SAS and two Excel):

SUBSTATE\_TRACT141516.SAS7BDAT/ SUBSTATE\_TRACT141516.XLSX: Tract level files containing 2014-2016 substate region definitions for the following states: California, Connecticut, Delaware, District of Columbia, Illinois Massachusetts, and Rhode Island. Note, for California, Illinois, and Delaware, some regions are defined in terms of tracts and some are defined in terms of counties. However, all regions for those states are defined within the tract file.

SUBSTATE\_COUNTY141516.SAS7BDAT/ SUBSTATE\_COUNTY141516.XLSX: County level files containing 2014-2016 substate region definitions for the remaining states (i.e., states excluding the ones on the tract level file).

VARIABLES ON FILE:

1. SBST16N (2014-2016 SUBSTATE REGION NAME) is the substate region name used in the 2014-2016 substate report (<http://www.samhsa.gov/data>). The values of SBST16N are unique within a state, but two or more states may have substate regions with the same substate region name.
2. SBST16 (2014-2016 SUBSTATE REGION ID WITHIN STATE) is a numeric value that corresponds to the substate region name within a state. If a state has "n" substate regions, then SBST16 will take on values 1, 2. . . ,n, within that state. SBST16 was created by assigning a unique ID to the sorted order of SBST16N. For example if there are 10 substate regions in a state, Region 10 would come after Region 1, before Region 2 and hence Region 1 would get a SBST16 value of 1, Region 10 would get a SBST16 value of 2 and Region 2 would get a SBST16 value of 3. An actual example from the 2014-2016 substate regions is as follows: there are 10 substate regions (SBST16N) for Michigan, Region 1, Region 2 ... Region 10. For this example, Region 1 would get an SBST16 value of 1, Region 10 would get SBST16 value of 2, and Region 2 through Region 9 would get SBST16 values 3 through 10. Note, the order of SBST16 does not always match the order that the region names (SBST16N) are listed in published substate tables and other documents.

Several states also have aggregate substate areas. These aggregate areas are made up of two or more of the substate regions.

1. SBSTA16N (2014-2016 AGGREGATE SUBSTATE AREA NAME) is the name of the aggregate area variable; this variable is blank for the states with no aggregate areas. Within the states with aggregate areas, there are some substate regions that do not belong to an aggregate area. For these cases, SBSTA16N=SBST16N and AGGFLG16=0 (see below).
2. SBSTAG16 (2014-2016 AGGREGATE AREA ID WITHIN STATE) is a numeric variable within state assigned to each aggregate area (i.e., area assigned an aggregate name, SBSTA16N). If a state has "n" aggregate areas, then SBSTAG16 will take on values 1, 2,. . . ,n, within that state. SBSTAG16 was created by assigning a unique ID to the sorted order of SBSTA16N. For example, if a state has 3 substate areas (Region 1, Region 2, and Region 3) and two of these regions make up an aggregate region called Regions 2 and 3. Then SBSTAG16=1 for Region 1, and SBSTAG16=2 for the aggregate regions called Regions 2 and 3. An actual example from the 2014-2016 substate regions is as follows: there are 11 aggregate substate regions (SBSTA16N) for Texas, Region 1, Region 2. . . Region 11. For this example, Region 1 would get an SBSTAG16 value of 1, Region 10 and Region 11 would get SBSTAG16 values of 2 and 3, and Region 2 through Region 9 would get SBSTAG16 values 4 through 11. Note, the order of SBSTAG16 does not always match the order that the aggregate region names (SBSTA16N) are listed in published substate tables and other documents. Note, if a state does not have aggregate regions, this variable is blank.

1. [AGGFLG16:](#index_AGGFLG14) 2014-2016 AGGREGATE AREA FLAG 0 = 2014-2016 substate region not part of aggregate area

1 = 2014-2016 substate region part of aggregate area

1. [SBSTFG16:](#index_SBSTFG14) CHANGE INDICATOR BETWEEN 2012-2014 & 2014-2016 REGIONS 0 = Substate region changed between 2012-2014 & 2014-2016

1 = Substate region consistent between 2012-2014 & 2014-2016

1. STATE: State FIPS code
2. COUNTY: County FIPS code
3. TRACT: 2010 Geography census tract code (only on the tract files)

HOW TO USE:

Merge substate region definitions onto any data file using state and county for the states whose regions are defined in terms of counties; and using state, county, and tract identifier for the states whose regions are defined in terms of Census tracts.