

SENTINEL SAS MACRO TOOLKIT:

Age Stratification Tool

(%MS_AGESTRAT)

Documentation version: 1.3

Prepared by the Sentinel Operations Center

For use with Age Stratification Tool version 1.3

March 25, 2016

Sentinel is sponsored by the U.S. Food and Drug Administration (FDA) to monitor the safety of FDA-regulated medical products. Sentinel is one piece of the Sentinel Initiative, a multi-faceted effort by the FDA to develop a national electronic system that complements previously existing methods of safety surveillance. Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I.

Age Stratification Tool

1. Introduction

In many programs and queries, members (or claims) need to be categorized into time/age intervals. Often times, analysts approximate the number of years by dividing the number of days by 360, 365 or 365.25; or approximate the number of months by dividing the number of weeks by 4.3. Similarly, analysts may confine all computations to a common unit of time (e.g. all age groups defined in years or months, without mixing the two). Fractional approximations of these types can often yield discrepancies in their computed values and do not lend themselves to being flexible on the usage of more than one computational unit of time. The MS_AGESTRAT tool avoids these approximations and adds computational flexibility by using anniversary dates and allowing users to define age-groups using a variety of units of time (e.g. a mix of year, quarter, months, week, and/or days) within the same stratification scheme.

2. Program Objectives

This program creates an age group variable consisting of custom, user-specified groupings, with allowance for any mix of time units.

3. Parameter Specifications

Program Macro Variable Name	Short Description	Long Description
INFILE	File containing start date (required).	<p>Details The name of the source data set (.SAS7BDAT file format) containing the STARTDT variable (required, see STARTDT below). If a constant value for ENDDT will not be used, this file also must contain an ENDDT variable (see ENDDT below).</p> <p>Input type: Required, referenced with <i>LIBREF.SAS-data-set</i></p> <p>Format: text</p> <p>Example: <i>indata.demographic</i></p>
Mini	Output file	<p>Details The name of the resulting SAS data set containing the age group variable.</p> <p>Input type: Required</p> <p>Format: text</p> <p>Example: <i>outdata.results1</i></p>

Program Macro Variable Name	Short Description	Long Description
STARTDT	Start date to measure age	<p>Details: The name of the date variable included in the source data set used as the starting point to calculate the age. This is typically a birth date variable.</p> <p>Input type: Required; variable has to be in the source data set</p> <p>Format: sas date format (for variable name in the input data set)</p> <p>Example: the field <i>birth_date</i> in the <i>indata.demographic</i> data set</p>
ENDDT	End date to measure age	<p>Details: Value used as the end date at which the age will be calculated and the age group will be assigned. This parameter can either be a variable in the input data set or a constant (common value for all observations). This is the date “as of” the age will be calculated.</p> <p>Input type: Required;</p> <ul style="list-style-type: none"> • variable has to be in the input table; <p>OR</p> <ul style="list-style-type: none"> • constant; has to be supplied <p>Format:</p> <ul style="list-style-type: none"> • SAS date variable (for variable name in the input data set) or • mm/dd/yyyy (for date constant) <p>Examples:</p> <ul style="list-style-type: none"> • the variable <i>Adate</i> in the <i>indata.AgeStrat_demo_diagnosis</i> table (after the demographic and diagnosis datasets are merged together) • 01/01/2012, a given date constant
TIMESTRAT	Age groups for stratification	<p>Details: The user has the option to output results by custom age groupings or use the default age grouping. Each age group must be separated by a space and must contain a “-“(or a “+” if the last age category is open ended).</p> <p>For example, to have drug users stratified by 20 year increments in members 40+, the user will input AGESTRAT=40y-59y 60y-79y 80y-99y. Note that when a time increment is not specified, the program assumes it is years. For example, suppose that the user wants to use narrow age groups for early ages and broader for older, the user could want 00d-13d 02w-26w 27w-52w 01Y-10Y 11Y-25Y 25+.</p> <p>Note: Lower value is binding. If AGESTRAT=0-5 5-10, then all 5 year olds will be placed in the second age group. If AGESTRAT=0-5 6-10, then all 5 year olds will be placed in the first age group. Although the upper bound is not used in computation it is, however, still required.</p> <p>Input type: Optional (default value is 00-01 02-04 05-09 10-14 15-18 19-21 22-44 45-64 65-74 75+ years)</p> <p>Format: AAx-AAx BBx-BBx ZZx-ZZx with x = D, W, M, Q, Y. If the “x” parameter is blank, “Y” is assumed as the units.</p> <p>Example: AGESTRAT= 40-59 60-79 80-99</p>

4. Outputs

The output file consists of the input data set augmented by a text variable containing the Age Group for each observation. Each output data set is produced in the .SAS7BDAT file format.

5. Examples

- A.** The %MS_AGESTRAT macro is used to calculate the members' age on January 1st, 2011. The birth date is used as the start date and it is provided in the source dataset (AgeStrat_demographic). The end date parameter is initialized with the 01/01/2011 constant. The user-specified custom age groupings consist of a mix of days, weeks, months, quarters and years : 0D-6D 1W-3W 1M-2M 1Q-2Q 7M-11M 1Y-2Y 3Y-5Y 6Y+. The following parameters are used in the macro call:

```
%MS_AGESTRAT( INFILE=indata.ms_agestrat_demographic,
              OUTFILE=outdata.ms_agestrat_example1,
              STARTDT=birth_date,
              ENDDT=01/01/2011,
              TIMESTRAT=0D-6D 1W-3W 1M-2M 1Q-2Q 7M-11M 1Y-2Y 3Y-5Y 6Y+
              );
```

The Age Group field will be output in the outdata.example1 dataset.

- B.** The %MS_AGESTRAT macro is used to calculate the patients' ages as of a diagnosis date (ADATE) after merging the demographic and diagnosis tables. The resulting AgeStrat_demo_diagnosis data set is the source dataset, containing the start date – birth date – and the end date ADATE. No time vector is specified, therefore the default value will be used (00-01 02-04 05-09 10-14 15-18 19-21 22-44 45-64 65-74 75+ with all numbers representing YEARS). The following parameters are used in the macro call:

```
%MS_AGESTRAT( INFILE=indata.ms_agestrat_agestrat_demo_diagnosis,
              OUTFILE=outdata.ms_agestrat_example2,
              STARTDT=birth_date,
              ENDDT=ADate,
              TIMESTRAT=
```

```
);
```

The Age Group field will be output in the outdata.example2 dataset.