

\*\*\*\*\*  
\*\*\*\*\*  
NEW YORK CITY YOUTH RISK BEHAVIOR SURVEY (YRBS)  
SAMPLE CODE FOR GENERATING NYC CITYWIDE AND BOROUGH PREVALENCE ESTIMATES FROM  
THE CDC'S 2019 DISTRICT COMBINED DATASET

FOR MORE INFORMATION, PLEASE CONTACT:  
NYC DEPARTMENT OF HEALTH & MENTAL HYGIENE  
BUREAU OF EPIDEMIOLOGY SERVICES  
EPIDATAREQUEST@HEALTH.NYC.GOV

\*\*\*\*\*  
\*\*\*\*\*

\*\*\*\*\*  
\*\*\*\*\*

USERS WILL NEED TO RUN THE CDC'S 'SAS FORMAT PROGRAM' AND 'SAS INPUT PROGRAM'  
TO CALL-IN THE DATA  
BEFORE RUNNING THIS PROGRAM.

THE PROGRAMS ARE AVAILABLE HERE:  
[HTTPS://WWW.CDC.GOV/HEALTHYYOUTH/DATA/YRBS/DATA.HTM](https://www.cdc.gov/healthyyouth/data/yrbs/data.htm)

ENTER IN THE FOLDER PATH WHERE YOU HAVE SAVED THE DATASET AS INDICATED BY <  
>.

REMOVE THE < > BEFORE RUNNING THE PROGRAM.

\*\*\*\*\*  
\*\*\*\*\*;

\*\*\*\*\*  
\*\*\*\*\*

YRBS DATA ARE AVAILABLE FOR NEW YORK CITY OVERALL STARTING IN 1997 AND FOR  
EACH OF THE FIVE BOROUGHS (BOROUGH OF SCHOOL)  
STARTING IN 2003.

SURVEY SITE NAME (SITECODE)	AVAILABLE YEARS (YEAR)
NEW YORK CITY, NY (NYC)	1997-2019
BOROUGH OF BRONX (NYG)	2003-2019
BOROUGH OF BROOKLYN (NYH)	2003-2019
BOROUGH OF MANHATTAN (NYI)	2003-2019
BOROUGH OF QUEENS (NYJ)	2003-2019
BOROUGH OF STATEN ISLAND (NYK)	2003-2019

\*\*\*\*\*  
\*\*\*\*\*;

\*\*\*\*\*  
NEST: Stratum psu  
SURVEY WEIGHT: Weight  
DESIGN: With replacement  
\*\*\*\*\*;

LIBNAME YRBS '< >';

**DATA** YRBS\_NYI19 ; **set** YRBS.SADC\_2019\_DISTRICT;  
**WHERE** SITECODE='NYI' AND YEAR=2019; /\*SPECIFY SITE (MANHATTAN) & YEAR OF  
INTEREST (SEE LIST ABOVE)\*/  
**RUN;**

```

PROC CONTENTS DATA=YRBS_NYI19;
RUN;

/*SAMPLE SUDAAN CODE*/

/*PREVALENCE OF EVER TOLD ASTHMA - IN MANHATTAN (BOROUGH OF SCHOOL), 2019 */
PROC DESCRIPT DATA= YRBS_NYI19 FILETYPE = SAS DESIGN = WR NOMARG NOTSORTED;
  NEST STRATUM PSU / MISSUNIT;
  WEIGHT WEIGHT;
  VAR QN87 QN87;
  CATLEVEL 1 2;
  TABLES _ONE_;
  CLASS _ONE_ / NOFREQ;
  SETENV DECWIDTH = 1;
  PRINT PERCENT SEPERCENT LOWPCT UPPCT / STYLE = NCHS;
  OUTPUT / FILENAME = CHECK FILETYPE = SAS TABLECELL = DEFAULT REPLACE;
run;

/*USE OUTPUT DATASET TO CALCULATE RSE AND RELIABILITY*/
data check1;
  set check;
  if percent in (0.00, 100.00) then do;
    if nsum >= 50 then flag = '**';
    if nsum >= 50 then flag_detail = '**';
    if nsum < 50 then flag = '^';
    if nsum < 50 then flag_detail = '^1';
  end;
  else if percent not in (0.00, 100.00) then do;
    rse = sepercent/percent;
    ciband = uppct-lowpct;
    halfw = ciband/2;
    if sepercent = 0.0 and ciband = 0.0 then flag='^';
    if sepercent = 0.0 and ciband = 0.0 then flag_detail='^2';
    else if rse =>0.5 then do;
      if ciband >=6 then flag='^';
      else if ciband < 6 then flag = '*';
      if ciband >=6 then flag_detail='^3';
      else if ciband < 6 then flag_detail = '*1';
    end;
    else if rse < 0.3 then do;
      if nsum <50 then flag='*';
      else if nsum >= 50 then do;
        if halfw > 10 then flag = '*';
      end;
      if nsum <50 then flag_detail='*2';
      else if nsum >= 50 then do;
        if halfw > 10 then flag_detail = '*3';
      end;
    end;
  end;
  else if 0.5 > rse >=0.3 then do;
    flag='*';
    flag_detail='*1';
  end;
end;
run;

```

```

/*print out all observations that meet flag/suppression criteria*/

options ls = 150;
proc print data = check1 noobs;
  var flag_detail flag rse nsum ciband halfw percent lowpct uppct;
  where flag in ('*', '^', '**');
run;

/*remember what each of the different flags mean!

Flag
* = reliability note ("Estimate should be interpreted with caution.
Estimate's Relative Standard Error (a measure of estimate precision) is
greater than 30%, or the 95% Confidence Interval's half width is greater than
10, or the sample size is too small making the estimate potentially
unreliable")

^ = suppression note ("Data are suppressed due to imprecise and unreliable
estimates")

** = confidence interval note ("Estimate should be interpreted with caution.
95% Confidence Interval and Relative Standard Error are not calculated")

Flag_detail
*1 = reliability note ("Estimate should be interpreted with caution due to
large Relative Standard Error")

*2 = reliability note ("Estimate should be interpreted with caution due to
small sample size")

*3 = reliability note ("Estimate should be interpreted with caution due to
wide 95% Confidence Interval")

^1 = suppression note ("Data are suppressed due to imprecise and unreliable
estimates caused by small sample size")

^2 = suppression note ("Data are suppressed due to imprecise and unreliable
estimates caused by 0 standard error")

^3 = suppression note ("Data are suppressed due to imprecise and unreliable
estimates caused by wide 95% Confidence Interval and large Relative Standard
Error")

** = confidence interval note ("Estimate should be interpreted with caution.
95% Confidence Interval and Relative Standard Error are not calculated")

/*Be sure to include other relevant notes as necessary*/

/* GET THE ESTIMATED NUMBER OF STUDENTS WHO REPORTED EVER TOLD ASTHMA - IN
MANHATTAN (BOROUGH OF SCHOOL), 2019 */

PROC FREQ DATA=YRBS_NYI19;
TABLE QN87/list missing;
WEIGHT WEIGHT;
run;

```