

```

options pagesize=50 linesize=80;
options formchar='|----|+|---+=|-\<>';
title1 "SYC Data Preprocessing";
footnote "SYC Data Preprocessing";

/* creating dataset */
data SYC;
  infile "/home/u63744989/dataset/syc.txt" missover firstobs = 2 delimiter = ',';
  /* before read in the data, we manually delete the instructions on the top of the txt file */
  input stratum psu psusize initwt finalwt randgrp age race ethnicty educ
  sex livewith famtime crimtype everviol numarr probtn corrinst evertime
  prviol prprop prdrug prpub prjuv agefirst usewepn alcuse everdrug;

  /* creating labels */
  label age = "age";
  label race = "race";
  label sex = 'gender';
  label stratum = "stratum";
  label numarr = "number of prior arrest";
  label prviol = "previously arrested for violent crime";
  label everdrug = "ever used illegal drugs";
  label finalwt = "final sampling weight";

  /* Change the missing values to . for those variables whose missing values is 9*/
  array missing_ls_1 race sex numarr prviol everdrug;
  do over missing_ls_1;
    if (missing_ls_1 = 9) then missing_ls_1 = .;
  end;

  /* Change the missing values to . for those variables whose missing values is 99*/
  array missing_ls_2 age;
  do over missing_ls_2;
    if (missing_ls_2 = 99) then missing_ls_2 = .;
  end;

  /* Keep only the necessary columns, i.e. the 8 variables mentioned */
  keep stratum numarr age race sex prviol everdrug finalwt;
run;

/* sort the dataset by stratum in an ascending order*/
proc sort data=SYC;
  by stratum;
run;

/* output dataset to a csv file saved as the syc for sas processing*/
proc export data=SYC
  outfile='/home/u63744989/dataset/syc_sas.csv'
  dbms=csv
  replace;
run;

data SYC;
  set SYC;
  keep stratum finalwt;
run;

proc export data=SYC
  outfile = '/home/u63744989/dataset/syc_r.csv'
  dbms = csv
  replace;
run;

/* View overview of the stratum variable */
proc freq data=SYC;
  tables stratum / nocol;
  title "Overview of Stratum Variable";
run;

ods pdf file='/home/u63744989/figures/output_1.pdf';

ods _all_ close;

```

