```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
72
73
74         DATA population; /*Example of population dataset*/
75         input uniqueid gender $ age casecontrol; /*gender is defined as categorical variables by the following $*/
76         cards;

NOTE: The data set WORK.POPULATION has 24 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.02 seconds


101        ;
102        run;
103
104        %LET agerange = 5; /*For this example, we have chosen the age range to be 5. Controls can therefore be up to 5 years
104      ! younger or older than the case*/
105        %LET ratio = 3; /*We have chosen to match 3 controls for each case*/
106
107        DATA cases controls;
108           SET population;
109           IF casecontrol = 1 THEN OUTPUT cases;
110           ELSE OUTPUT controls;
111        RUN;

NOTE: There were 24 observations read from the data set WORK.POPULATION.
NOTE: The data set WORK.CASES has 7 observations and 4 variables.
NOTE: The data set WORK.CONTROLS has 17 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.02 seconds
      cpu time            0.01 seconds


112
113        PROC FREQ NOPRINT DATA=cases;
114           TABLES age*gender/OUT=caseout;
115        RUN;

NOTE: There were 7 observations read from the data set WORK.CASES.
NOTE: The data set WORK.CASEOUT has 5 observations and 4 variables.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


116
117        %MACRO sample(v_age, v_gender, v_count);
118
119           DATA qualify1;
120              SET controls;
121               WHERE (&v_age-&agerange <= age <= &v_age+&agerange)
122                 AND
123                 (gender = "&v_gender");
124
125
126           case_age = &v_age;
127           case_gender = "&v_gender";
128
129           SEED = RANUNI(0);
130           PROC SORT;
131              BY SEED;
132
133           DATA qualify2;
134              SET qualify1 NOBS=totobs;
135              IF _N_ <= &v_count*&ratio;
136              IF &v_count*&ratio <= totobs THEN tag = 'yes';
137                 ELSE tag = 'no';
138
139           PROC APPEND BASE=matches DATA=qualify2 force;
140
141           PROC SORT DATA=qualify2 OUT=temp1 (KEEP=uniqueid);
142              BY uniqueid;
143
144           PROC SORT DATA=controls OUT=temp2;
145              BY uniqueid;
146
147           DATA controls; /*the dataset controls is updated so that the controls already matched are removed and can not be
147      ! matched again*/
148                MERGE temp1(IN=in1) temp2(IN=in2);
149                BY uniqueid;
150                IF in2 AND NOT in1;
151
152        %MEND sample;
153
154        DATA _NULL_;
155           SET caseout;
156           CALL EXECUTE ('%sample('||age||','||gender||','||count||')');
157        RUN;

NOTE: Numeric values have been converted to character values at the places given by: (Line):(Column).
```

```
      156:30   156:53
NOTE: There were 5 observations read from the data set WORK.CASEOUT.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


NOTE: CALL EXECUTE generated line.
1         + DATA qualify1;       SET controls;     WHERE (25-5 <= age <= 25+5)       AND       (gender = "m");      case_age = 25;
  case_gender = "m";        SEED = RANUNI(0);

NOTE: There were 8 observations read from the data set WORK.CONTROLS.
      WHERE (age>=20 and age<=30) and (gender='m');
NOTE: The data set WORK.QUALIFY1 has 8 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


1         +
PROC SORT;       BY SEED;

NOTE: There were 8 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY1 has 8 observations and 7 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


1         +
DATA qualify2;       SET qualify1 NOBS=totobs;       IF
2         + _N_ <= 1*3;       IF 1*3 <= totobs THEN tag = 'yes';          ELSE tag = 'no';

NOTE: There were 8 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY2 has 3 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


2         +                                                                                PROC APPEND BASE=matches
DATA=qualify2 force;

NOTE: Appending WORK.QUALIFY2 to WORK.MATCHES.
NOTE: BASE data set does not exist. DATA file is being copied to BASE file.
NOTE: There were 3 observations read from the data set WORK.QUALIFY2.
NOTE: The data set WORK.MATCHES has 3 observations and 8 variables.
NOTE: PROCEDURE APPEND used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


2         +
PROC SORT DATA=qualify2 OUT=temp1 (KEEP=uniqueid);       BY uniqueid;     PROC SORT DATA=controls OUT=temp2;       BY
NOTE: There were 3 observations read from the data set WORK.QUALIFY2.
NOTE: The data set WORK.TEMP1 has 3 observations and 1 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


3         + uniqueid;

NOTE: There were 17 observations read from the data set WORK.CONTROLS.
NOTE: The data set WORK.TEMP2 has 17 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


3         +               DATA controls;        MERGE temp1(IN=in1) temp2(IN=in2);       BY uniqueid;        IF in2 AND NOT in1;

NOTE: There were 3 observations read from the data set WORK.TEMP1.
NOTE: There were 17 observations read from the data set WORK.TEMP2.
NOTE: The data set WORK.CONTROLS has 14 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds


NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

4         + DATA qualify1;       SET controls;     WHERE (26-5 <= age <= 26+5)       AND       (gender = "f");      case_age = 26;
  case_gender = "f";        SEED = RANUNI(0);

NOTE: There were 9 observations read from the data set WORK.CONTROLS.
      WHERE (age>=21 and age<=31) and (gender='f');
NOTE: The data set WORK.QUALIFY1 has 9 observations and 7 variables.

4         +
PROC SORT;       BY SEED;
```

```
NOTE: There were 9 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY1 has 9 observations and 7 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


4         +
DATA qualify2;       SET qualify1 NOBS=totobs;       IF
5         + _N_ <= 3*3;       IF 3*3 <= totobs THEN tag = 'yes';          ELSE tag = 'no';

NOTE: There were 9 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY2 has 9 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


5         +                                                                                PROC APPEND BASE=matches
DATA=qualify2 force;

NOTE: Appending WORK.QUALIFY2 to WORK.MATCHES.
NOTE: There were 9 observations read from the data set WORK.QUALIFY2.
NOTE: 9 observations added.
NOTE: The data set WORK.MATCHES has 12 observations and 8 variables.
NOTE: PROCEDURE APPEND used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


5         +
PROC SORT DATA=qualify2 OUT=temp1 (KEEP=uniqueid);       BY uniqueid;     PROC SORT DATA=controls OUT=temp2;       BY
NOTE: There were 9 observations read from the data set WORK.QUALIFY2.
NOTE: The data set WORK.TEMP1 has 9 observations and 1 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.02 seconds


6         + uniqueid;

NOTE: There were 14 observations read from the data set WORK.CONTROLS.
NOTE: The data set WORK.TEMP2 has 14 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


6         +               DATA controls;        MERGE temp1(IN=in1) temp2(IN=in2);       BY uniqueid;        IF in2 AND NOT in1;

NOTE: There were 9 observations read from the data set WORK.TEMP1.
NOTE: There were 14 observations read from the data set WORK.TEMP2.
NOTE: The data set WORK.CONTROLS has 5 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


7         + DATA qualify1;       SET controls;     WHERE (26-5 <= age <= 26+5)       AND       (gender = "m");      case_age = 26;
  case_gender = "m";        SEED = RANUNI(0);

NOTE: There were 5 observations read from the data set WORK.CONTROLS.
      WHERE (age>=21 and age<=31) and (gender='m');
NOTE: The data set WORK.QUALIFY1 has 5 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


7         +
PROC SORT;       BY SEED;

NOTE: There were 5 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY1 has 5 observations and 7 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


7         +
DATA qualify2;       SET qualify1 NOBS=totobs;       IF
8         + _N_ <= 1*3;       IF 1*3 <= totobs THEN tag = 'yes';          ELSE tag = 'no';

NOTE: There were 5 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY2 has 3 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


8         +                                                                                PROC APPEND BASE=matches
DATA=qualify2 force;
```

```
NOTE: Appending WORK.QUALIFY2 to WORK.MATCHES.
NOTE: There were 3 observations read from the data set WORK.QUALIFY2.
NOTE: 3 observations added.
NOTE: The data set WORK.MATCHES has 15 observations and 8 variables.
NOTE: PROCEDURE APPEND used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


8         +
PROC SORT DATA=qualify2 OUT=temp1 (KEEP=uniqueid);       BY uniqueid;     PROC SORT DATA=controls OUT=temp2;       BY
NOTE: There were 3 observations read from the data set WORK.QUALIFY2.
NOTE: The data set WORK.TEMP1 has 3 observations and 1 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


9         + uniqueid;

NOTE: There were 5 observations read from the data set WORK.CONTROLS.
NOTE: The data set WORK.TEMP2 has 5 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


9         +               DATA controls;        MERGE temp1(IN=in1) temp2(IN=in2);       BY uniqueid;        IF in2 AND NOT in1;

NOTE: There were 3 observations read from the data set WORK.TEMP1.
NOTE: There were 5 observations read from the data set WORK.TEMP2.
NOTE: The data set WORK.CONTROLS has 2 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


10        + DATA qualify1;       SET controls;     WHERE (27-5 <= age <= 27+5)       AND       (gender = "m");      case_age = 27;
  case_gender = "m";        SEED = RANUNI(0);

NOTE: There were 2 observations read from the data set WORK.CONTROLS.
      WHERE (age>=22 and age<=32) and (gender='m');
NOTE: The data set WORK.QUALIFY1 has 2 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.02 seconds


10        +
PROC SORT;       BY SEED;

NOTE: There were 2 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY1 has 2 observations and 7 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


10        +
DATA qualify2;       SET qualify1 NOBS=totobs;       IF
11        + _N_ <= 1*3;       IF 1*3 <= totobs THEN tag = 'yes';          ELSE tag = 'no';

NOTE: There were 2 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY2 has 2 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


11        +                                                                                PROC APPEND BASE=matches
DATA=qualify2 force;

NOTE: Appending WORK.QUALIFY2 to WORK.MATCHES.
NOTE: There were 2 observations read from the data set WORK.QUALIFY2.
NOTE: 2 observations added.
NOTE: The data set WORK.MATCHES has 17 observations and 8 variables.
NOTE: PROCEDURE APPEND used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


11        +
PROC SORT DATA=qualify2 OUT=temp1 (KEEP=uniqueid);       BY uniqueid;     PROC SORT DATA=controls OUT=temp2;       BY
NOTE: There were 2 observations read from the data set WORK.QUALIFY2.
NOTE: The data set WORK.TEMP1 has 2 observations and 1 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


12        + uniqueid;
```

```
NOTE: There were 2 observations read from the data set WORK.CONTROLS.
NOTE: The data set WORK.TEMP2 has 2 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


12        +               DATA controls;        MERGE temp1(IN=in1) temp2(IN=in2);       BY uniqueid;        IF in2 AND NOT in1;

NOTE: There were 2 observations read from the data set WORK.TEMP1.
NOTE: There were 2 observations read from the data set WORK.TEMP2.
NOTE: The data set WORK.CONTROLS has 0 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


13        + DATA qualify1;       SET controls;     WHERE (28-5 <= age <= 28+5)       AND       (gender = "m");      case_age = 28;
  case_gender = "m";        SEED = RANUNI(0);

NOTE: There were 0 observations read from the data set WORK.CONTROLS.
      WHERE (age>=23 and age<=33) and (gender='m');
NOTE: The data set WORK.QUALIFY1 has 0 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds


13        +
PROC SORT;       BY SEED;

NOTE: Input data set is empty.
NOTE: The data set WORK.QUALIFY1 has 0 observations and 7 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


13        +
DATA qualify2;       SET qualify1 NOBS=totobs;       IF
14        + _N_ <= 1*3;       IF 1*3 <= totobs THEN tag = 'yes';          ELSE tag = 'no';

NOTE: There were 0 observations read from the data set WORK.QUALIFY1.
NOTE: The data set WORK.QUALIFY2 has 0 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


14        +                                                                                PROC APPEND BASE=matches
DATA=qualify2 force;

NOTE: Appending WORK.QUALIFY2 to WORK.MATCHES.
NOTE: There were 0 observations read from the data set WORK.QUALIFY2.
NOTE: 0 observations added.
NOTE: The data set WORK.MATCHES has 17 observations and 8 variables.
NOTE: PROCEDURE APPEND used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


14        +
PROC SORT DATA=qualify2 OUT=temp1 (KEEP=uniqueid);       BY uniqueid;     PROC SORT DATA=controls OUT=temp2;       BY
NOTE: Input data set is empty.
NOTE: The data set WORK.TEMP1 has 0 observations and 1 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


15        + uniqueid;

NOTE: Input data set is empty.
NOTE: The data set WORK.TEMP2 has 0 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


15        +               DATA controls;        MERGE temp1(IN=in1) temp2(IN=in2);       BY uniqueid;        IF in2 AND NOT in1;
158
159        /*The next part is for testing if any of the cases have not received the wanted amount of controls*/

NOTE: There were 0 observations read from the data set WORK.TEMP1.
NOTE: There were 0 observations read from the data set WORK.TEMP2.
NOTE: The data set WORK.CONTROLS has 0 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds


160        PROC FREQ NOPRINT DATA=matches;
161           TABLES case_age*case_gender/OUT=con_out;
```

```
162

NOTE: There were 17 observations read from the data set WORK.MATCHES.
NOTE: The data set WORK.CON_OUT has 4 observations and 4 variables.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


163        PROC SORT DATA = caseout(RENAME=
164                         (age=case_age gender=case_gender count=case_cnt));
165           BY case_age case_gender;
166

NOTE: There were 5 observations read from the data set WORK.CASEOUT.
NOTE: The data set WORK.CASEOUT has 5 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


167        PROC SORT DATA = con_out (RENAME= (count=con_cnt));
168           BY case_age case_gender;
169

NOTE: There were 4 observations read from the data set WORK.CON_OUT.
NOTE: The data set WORK.CON_OUT has 4 observations and 4 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds


170        DATA final (DROP=percent);
171           MERGE caseout con_out;
172           BY case_age case_gender;
173
174        con_need = case_cnt*&ratio;
175        IF con_cnt = . THEN con_cnt = 0;
176        diff = con_cnt-con_need;
177

NOTE: There were 5 observations read from the data set WORK.CASEOUT.
NOTE: There were 4 observations read from the data set WORK.CON_OUT.
NOTE: The data set WORK.FINAL has 5 observations and 6 variables.
NOTE: DATA statement used (Total process time):
      real time           0.02 seconds
      cpu time            0.00 seconds


178        PROC PRINT DATA = final; /*creates a table showing what characterizes the cases who have not received enough matches and
178      ! how many matches they are missing*/
179           WHERE diff < 0;
180           TITLE 'Insufficient Matches';
181        RUN;

NOTE: There were 2 observations read from the data set WORK.FINAL.
      WHERE diff<0;
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.08 seconds
      cpu time            0.08 seconds


182
183
184        OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
196
```