

NOTE #4: ARRAYS

Many cases we have to deal with a group of many variables for data manipulation, summary, and analysis. SAS array can be used wisely to have your tasks more efficient and simple. It can be better explained by examples.

```
/* Example 1-1 */
/* replacing missing score with observation min */

DATA score1;
  input ID $ test1-test5 ;
  datalines;
001 52 63 98 54 25
002 54 . 89 74 45
003 63 65 87 89 74
004 36 25 45 . 39
005 96 98 97 89 74
006 85 . 74 65 .
;
DATA array1; set score1;
  Array array_test{5} test1-test5;
  Do i = 1 to 5;
    if missing(array_test{i}) then array_test{i}=smallest(1,of test1-test5);
  end;
  Drop i;
  average=mean(of test1-test5);

proc print; run;

/* Example 1-2 */
/* replacing missing score with variable min */

DATA score1;
  input ID $ test1-test5 ;
  datalines;
001 52 63 98 54 25
002 54 . 89 74 45
003 63 65 87 89 74
004 36 25 45 . 39
005 96 98 97 89 74
006 85 . 74 65 .
;

PROC TRANSPOSE data=score1 out=score2; run;
DATA score_tr; set score2;
Array array_tr{6} coll1-col6;
Do i = 1 to 6;
  if missing(array_tr{i}) then array_tr{i}=min(of coll1-col6);
end;
Drop i;

PROC TRANSPOSE data=score_tr out=score1 (drop=_NAME_);
run;

DATA score3; set score1;
  average=mean(of test1-test5);
```

```

PROC PRINT; RUN;

/* Example 2 */
DATA score2;
  input name $ (grade1-grade3) ($) unit1-unit3 ;
  datalines;
Mike A B C 4 5 4
Ruth C D B 3 3 4
Kate B B A 5 3 4
Scott B A A 4 4 5
;
DATA array2; set score2;
  Array array_grade{3} grade1-grade3;
  Array array_unit{3} unit1-unit3;
  Array array_GP{3};

Do i = 1 to 3;
  if array_grade{i}='A' then array_GP{i}=4*array_unit{i};
  if array_grade{i}='B' then array_GP{i}=3*array_unit{i};
  if array_grade{i}='C' then array_GP{i}=2*array_unit{i};
  if array_grade{i}='D' then array_GP{i}=1*array_unit{i};
end;
Drop i;
total_GP=sum(of array_GP1-array_GP3); total_unit=sum(of unit1-unit3);
GPA=total_GP/total_unit;
format GPA f3.1;

proc print; run;

/* Example 3 */
data array3;
  input code $ dept $ Q1 Q2 status $;
DATALINES;
AAA QA 2.3 5.9 Green
AAB SD 5.2 4.7 NA
XXX RT 6.2 4.1 Blue
ACC NA 1.2 2.3 NA
AAD XXX 1.0 5.3 Green
;
proc print; run;
data array3_1;
  set array3;
  array char_array{*} $ _CHARACTER_; * SAS will calculate the size of the;
  * array. _character_ (_numeric_)is
  * used to use all character (numeric);
  * variables in data;

  array num_array{*} _NUMERIC_;

  do i = 1 to dim(char_array); * dim( ) will return the number of ;
  * column in the array;
    if char_array{i} in ('NA' 'XXX') then
      call missing(char_array{i});
    char_array{i} = lowercase(char_array{i});
  end;

  do i = 1 to dim(num_array);
    if num_array{i} LE 3.0 then
      call missing(num_array{i});
  end;
drop i;
run;
proc print; run;

```

```
/* CLASS PRACTICE 1*/

/* Using SAS array convert the following Fahrenheit degrees to Celsius.*/

/* 45 85 47 96 102 58 78 65 87 58    */

/* Example 4 */
/* Example of using multidimensional array to read 2-dim table vaule */
/* This example will read percent points of t distribution at          */
/* df=1 to 10 and prob points .90, .95, .975, .99                      */

DATA tval;
  input df p;
Datalines;
1 .90
3 .99
9 .95
10 .975
;
Data t_table;

array ttable{10,4} _temporary_;
if _N_= 1 then
  Do df=1 to 10;
    Do prob=1 to 4;
      input ttable{df,prob} @;
    END;
  END;

SET tval;
IF p=.90 then prob=1;
ELSE IF p=.95 then prob=2;
ELSE IF p=.975 then prob=3;
ELSE IF p=.99 then prob=4;

t_value=ttable{df,prob};

DATALINES;
3.078 6.314 12.706 31.821
1.886 2.920 4.303 6.965
1.638 2.353 3.182 4.541
1.533 2.132 2.776 3.747
1.476 2.015 2.571 3.365
1.440 1.943 2.447 3.143
1.415 1.895 2.365 2.998
1.397 1.860 2.306 2.896
1.383 1.833 2.262 2.821
1.372 1.812 2.228 2.764
;
proc print data=t_table NOOBS;
var df p t_value;
run;
```

```
/* CLASS PRACTICE 2 */
```

To enter a college your SAT scores must be at least 670, 650, 610 for Reading, writing, and math, respectively, or total score of 2050. Consider the following score list.

```
001 580 590 680
002 780 710 600
003 680 650 600
004 690 680 710
005 780 800 760
006 520 780 780
007 630 680 720
008 780 710 640
009 580 780 710
010 670 650 620
```

Use SAS array to count the number of tests passed and totals scores. Then, determine whether each student is accepted or rejected.