

2장. SAS 데이터 셋 변경하기

2.1 기존 DS에서 변경하기

* DS을 다른 이름으로 저장하기

[prgm 1 데이터셋에서 변수 변경]

```
3 DATA EX1; SET TEST,SPORTS;
4
5
6
7 RUN;
```

` 해석하기.

SET DSN; 변경할 DS을 준비

★ 함수

` $X2=X2-2$; <- X2변수에서 2를 뺀 값을 X2로 저장

` $MEAN1=MEAN(X1,X2)$ <-X1 과 X2 의 평균

MEAN(OF X1-X10) : 연속된 변수명의 평균

` $MEAN2=(X1+X2)/2$;

※ 함수MEAN과 수식을 이용한 계산값 비교!!

```
3 DATA EX1; SET TEST,SPORTS;
4
5 X2=X2-2;
6 MEAN1=MEAN(X1,X2);
7 M=MEAN(OF X1-X2);
8 MEAN2=(X1+X2)/2;
9
10 RUN;
```

☆ 함수 변수명=함수(인수들)

절대값ABS / 정수INT / 제곱근SQRT / 지수EXP

자연지수e로그LOG / 밑이2인 로그LOG2 / LOG10

MAX MIN SUM MEAN VAR

※ 수치<->문자 형식바꾸기

수치->문자 변수명=PUT(수치변수명, 자리수.)

SEX=PUT(GENDER, 4.)

문자->수치 변수명=INPUT(문자변수명, 자리수.)

GG=INPUT(SEX, 4.)

※ 이름 바꾸기

RENAME 옛변수명=새변수명

RENAME MEAN1=MEANS_NEW

```
3 DATA EX1; SET TEST,SPORTS;
4 X2=X2-2; MEAN=(X1+X2)/2 ;
5 MEAN1=MEAN(X1,X2);
6
7 SEX=PUT(GENDER, 4.);
8 GG=INPUT(SEX, 4.);
9
10 RENAME MEAN1=MEAN_NEW;
11
12 RUN;
```

★★ 조건문

` IF 조건 THEN 결과;

IF X2="" THEN DELETE ;

IF GENDER=1 THEN G="남";

IF GENDER=2 THEN G="여";

☆ 원래 변수의 특성을 바꾸는 명령은 오류발생

예) IF GENDER=1 THEN GENDER='남' <- 오류!!

` IF THEN /ELSE IF THEN /ELSE

```
3 DATA EX1; SET TEST.SPORTS;
4 X2=X2-2; MEAN=(X1+X2)/2 ;
5
6 IF GENDER= 1 THEN G='남'; IF GENDER= 2 THEN G='여';
7
8 IF MEAN < 3 THEN GRADE='하';
9 ELSE IF MEAN < 4 THEN GRADE='중';
10 ELSE GRADE='상';
11 IF MEAN=" " THEN GRADE=" ";
12
13 RUN;
```

※ 변수 남기기(KEEP)/ 없애기(DROP)

남기기 KEEP 변수명1 변수명2 ...

KEEP MEAN_NEW

없애기 DROP 변수명1 변수명2 ...

DROP GENDER

[예제 2.1] SCORE 데이터셋을 이용하여 다음 데이터셋을 만드시오.

생성파일명 : SCORE COL1 COL2 DS1 DS2

<SCORE 데이터셋 ; SET TEST.SCORE>

- 1) Eng Math Sci MathPost
1~100 이외의 값을 갖으면 결측값 처리 하시오.
- 2) MJ MA의
1,2,3,4,5 이외의 값을 갖으면 결측값 처리 하시오.
NOT IN (1,2,3,4,5) <- 조건표현식 사용
- 3) MEAN 변수 생성
영어, 수학, 과학점수의 평균
※ 결측자료 있으면 결측값으로 변환하기
※ 소숫점4자리에서 반올림하기
~~>
- 4) DIFF 변수 생성 (특강 효과)
DIFF=MathPost - Math
- 5) MATH_JOY 변수 생성
수학불안 점수를 역채점으로 변환 6-MA
수학흥미와 수학불안 정방향값의 합
~~>
- 6) JOY 변수 만들기
MATH_JOY 4이하 JOY='낮음'
MATH_JOY 5이상7이하 JOY='낮음'
MATH_JOY 8이상 JOY='낮음'
- 7) DIFF 와 JOY 라벨 붙이기
DIFF "특강 효과"
JOY "수학흥미도 정도"

< COL1 COL2 데이터셋 ; SET SCORE>

- 1) COL1 데이터셋 만들기 : 아래 변수만 남기기
ID SEX_MF Eng Math Sci Math_Post MJ MA
- 2) COL2 데이터셋 만들기 : 아래 변수만 없애기
SEX_MF Eng Math Sci Math_Post MJ MA

< DS1 DS2 데이터셋 ; SET SCORE>

- 1) DS1 데이터셋 만들기
여자 케이스는 삭제
- 2) DS2 데이터셋 만들기
여자 케이스만 남기기

2.2 데이터셋 붙이기

* 케이스 추가하기 (SET 이용!!)

```
DATA 생성DSN ; SET 추가 DSN1 DSN1 ;
    DATA T_SET; SET DS1 DS2; RUN;
```

* 변수 추가하기 (MERGE 이용!!)

```
DATA 생성DSN ;
    MERGE 추가 DSN1 DSN1 ;    BY KEY;
```

☆ 변수추가 위해서 같은 케이스임을 확인하는 **KEY 필요!**

KEY 변수에 의하여 데이터셋은 정렬되어야 함.

```
58 DATA T_SET; SET DS1 DS2; RUN;
59
60 PROC SORT DATA=COL1 ; BY ID ;
61 PROC SORT DATA=COL2 ; BY ID ;
62
63 DATA T_MERGE ; MERGE COL1 COL2; BY ID; RUN;
64
```

※ 정렬하기~

```
PROC SORT DATA=T_SET OUT=T_SORT ;
    BY DESCENDING JOY DESCENDING SEX_MF;
    RUN;
```

` 해석하기.

SORT 정렬하는 과정입니다.

DATA=DSN 정렬할 데이터셋 제시

OUT= <- 정렬 후 결과 DSN 으로 저장

생략하면 원래 이름으로 다시 저장

BY <- 정렬의 기준이 되는 변수들 지정

DESCENDING Var ; Var 변수에 대하여 내림차순정렬