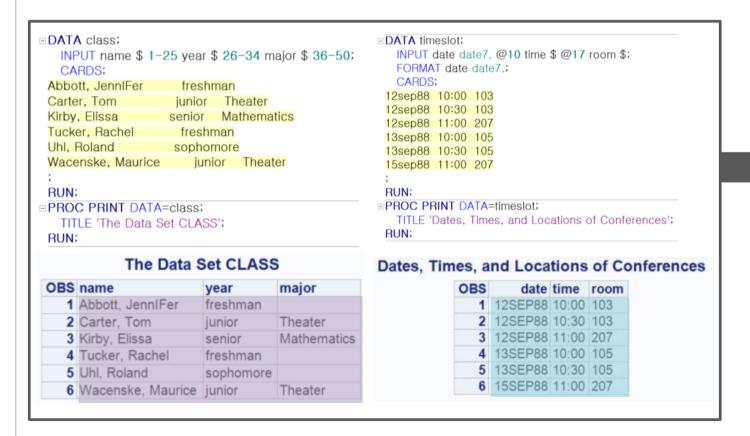


# Chapter 7. MERGE문 활용하기 Chapter 8. 새로운 변수의 생성

182STG27 임지연



1. 두 개의 DATASET 합치기



□ DATA schedule;

MERGE class timeslot;

RUN;

□ PROC PRINT DATA=schedule;

TITLE 'Merging CLASS and TIMESLOT to Assign';

TITLE2 'Conference Times to Students';

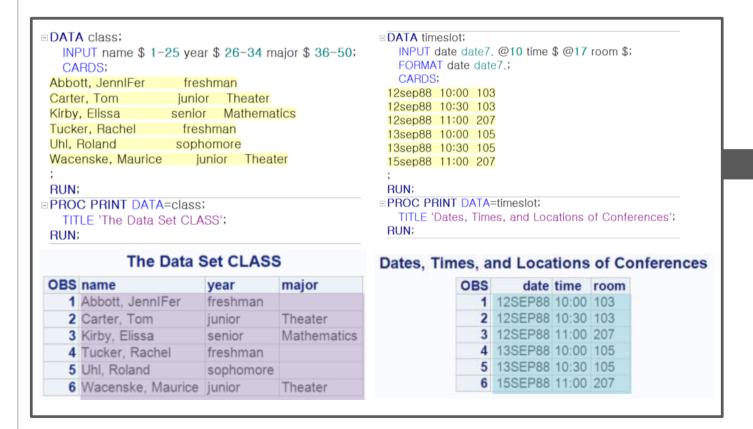
RUN;

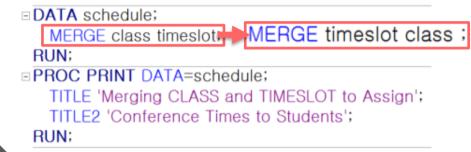
#### Merging CLASS and TIMESLOT to Assign Conference Times to Students

| OBS | name              | year      | major       | date    | time  | room |
|-----|-------------------|-----------|-------------|---------|-------|------|
| 1   | Abbott, JennIFer  | freshman  |             | 12SEP88 | 10:00 | 103  |
| 2   | Carter, Tom       | junior    | Theater     | 12SEP88 | 10:30 | 103  |
| 3   | Kirby, Elissa     | senior    | Mathematics | 12SEP88 | 11:00 | 207  |
| 4   | Tucker, Rachel    | freshman  |             | 13SEP88 | 10:00 | 105  |
| 5   | Uhl, Roland       | sophomore |             | 13SEP88 | 10:30 | 105  |
| 6   | Wacenske, Maurice | junior    | Theater     | 15SEP88 | 11:00 | 207  |



1. 두 개의 DATASET 합치기





# Merging CLASS and TIMESLOT to Assign Conference Times to Students

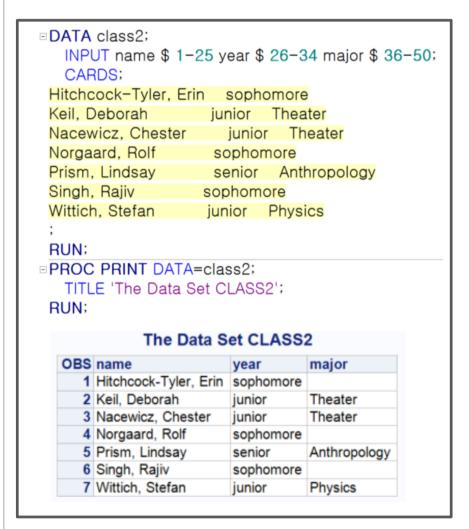
| OBS | name              | year      | major       | date    | time  | room |
|-----|-------------------|-----------|-------------|---------|-------|------|
| 1   | Abbott, JennIFer  | freshman  |             | 12SEP88 | 10:00 | 103  |
| 2   | Carter, Tom       | junior    | Theater     | 12SEP88 | 10:30 | 103  |
| 3   | Kirby, Elissa     | senior    | Mathematics | 12SEP88 | 11:00 | 207  |
| 4   | Tucker, Rachel    | freshman  |             | 13SEP88 | 10:00 | 105  |
| 5   | Uhl, Roland       | sophomore |             | 13SEP88 | 10:30 | 105  |
| 6   | Wacenske, Maurice | junior    | Theater     | 15SEP88 | 11:00 | 207  |

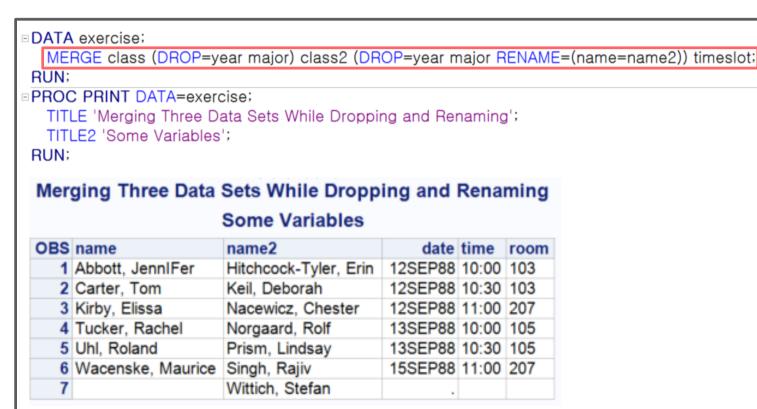
#### Merging TIMESLOT and CLASS to Assign Conference Times to Students

| OBS | date    | time  | room | name              | year      | major       |
|-----|---------|-------|------|-------------------|-----------|-------------|
| 1   | 12SEP88 | 10:00 | 103  | Abbott, JennIFer  | freshman  |             |
| 2   | 12SEP88 | 10:30 | 103  | Carter, Tom       | junior    | Theater     |
| 3   | 12SEP88 | 11:00 | 207  | Kirby, Elissa     | senior    | Mathematics |
| 4   | 13SEP88 | 10:00 | 105  | Tucker, Rachel    | freshman  |             |
| 5   | 13SEP88 | 10:30 | 105  | Uhl, Roland       | sophomore |             |
| 6   | 15SEP88 | 11:00 | 207  | Wacenske, Maurice | junior    | Theater     |



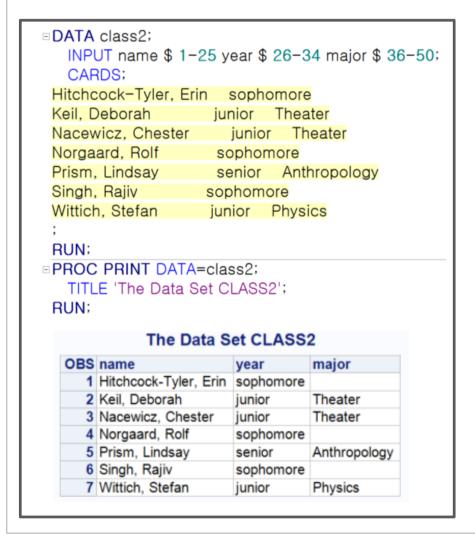
2. 세 개의 DATASET 합치기

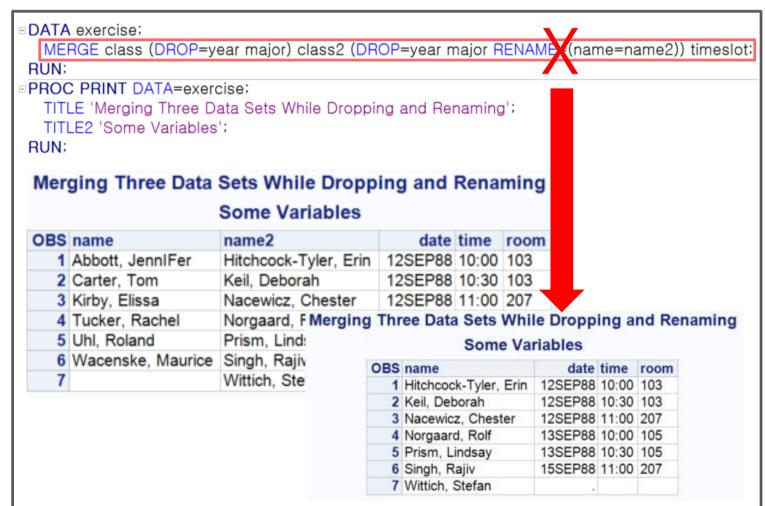






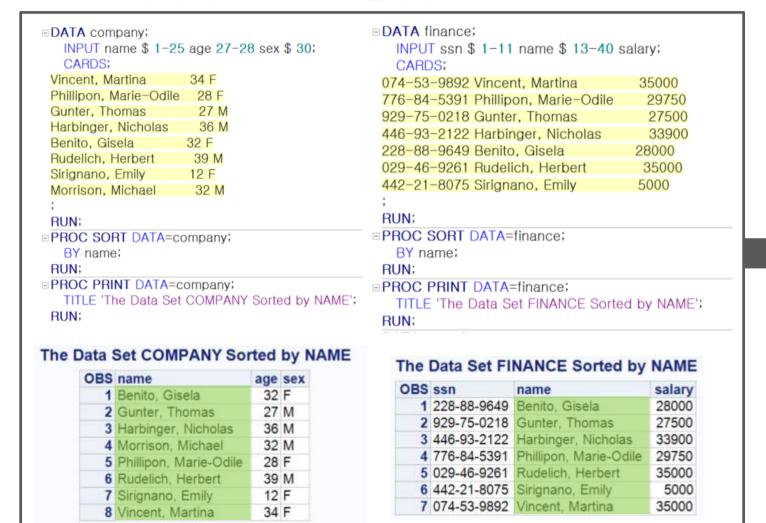
2. 세 개의 DATASET 합치기







3. 기준변수에 의한 DATASET 합치기



■ DATA compfin;

MERGE company finance;

BY name;

RUN;

PROC PRINT DATA=compfin;

TITLE 'Merging COMPANY and FINANCE';

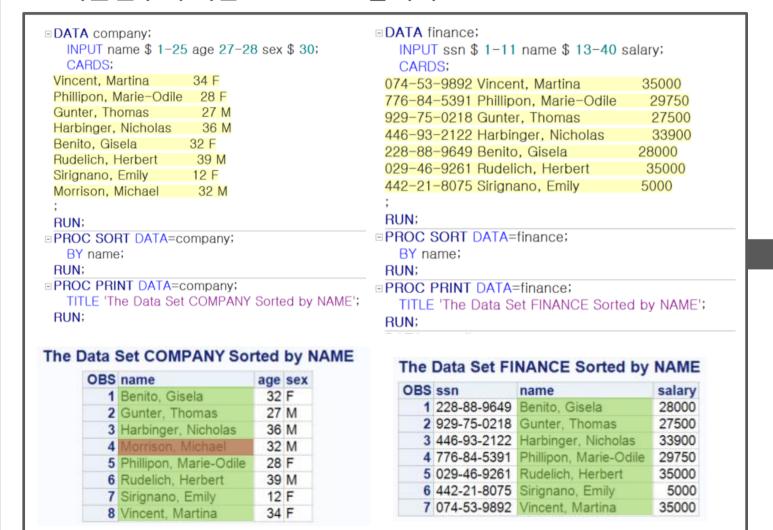
RUN;

Marging COMPANY and FINANCE

| OBS n |                        |     |     |             |        |
|-------|------------------------|-----|-----|-------------|--------|
|       | iame                   | age | sex | ssn         | salary |
| 1 B   | Benito, Gisela         | 32  | F   | 228-88-9649 | 28000  |
| 2 0   | Gunter, Thomas         | 27  | M   | 929-75-0218 | 27500  |
| 3 ⊦   | Harbinger, Nicholas    | 36  | M   | 446-93-2122 | 33900  |
| 4 N   | Morrison, Michael      | 32  | M   |             |        |
| 5 P   | Phillipon, Marie-Odile | 28  | F   | 776-84-5391 | 29750  |
| 6 F   | Rudelich, Herbert      | 39  | M   | 029-46-9261 | 35000  |
| 7 S   | Sirignano, Emily       | 12  | F   | 442-21-8075 | 5000   |
| 8 V   | /incent, Martina       | 34  | F   | 074-53-9892 | 35000  |



3. 기준변수에 의한 DATASET 합치기



DATA compfin;

MERGE company finance;

BY name;

RUN;

PROC PRINT DATA=compfin;

TITLE 'Merging COMPANY and FINANCE';

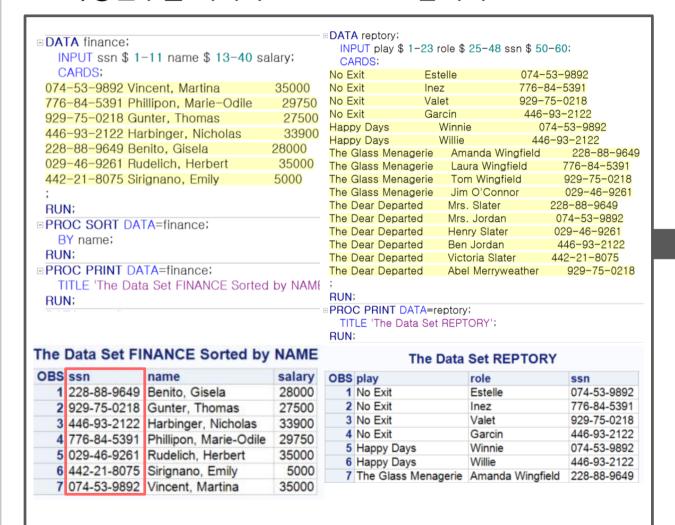
RUN;

Marging COMPANY and FINANCE

|     | Merging COMPA          | 414 I | anc | FINANCE     |        |
|-----|------------------------|-------|-----|-------------|--------|
| OBS | name                   | age   | sex | ssn         | salary |
| 1   | Benito, Gisela         | 32    | F   | 228-88-9649 | 28000  |
| 2   | Gunter, Thomas         | 27    | M   | 929-75-0218 | 27500  |
| 3   | Harbinger, Nicholas    | 36    | M   | 446-93-2122 | 33900  |
| 4   | Morrison, Michael      | 32    | M   |             |        |
| 5   | Phillipon, Marie-Odile | 28    | F   | 776-84-5391 | 29750  |
| 6   | Rudelich, Herbert      | 39    | M   | 029-46-9261 | 35000  |
| 7   | Sirignano, Emily       | 12    | F   | 442-21-8075 | 5000   |
| 8   | Vincent, Martina       | 34    | F   | 074-53-9892 | 35000  |



#### 4. 특정변수를 제거하고 DATASET 합치기



■ DATA newrep (DROP=ssn): MERGE finance (DROP=salary) reptory; BY ssn: RUN: □ PROC PRINT DATA=newrep; TITLE 'Replacing Players' Social Security Numbers'; TITLE2 'with Their Names': RUN: Replacing Players' Social Security Numbers with Their Names OBS name play role The Glass Menagerie 1 Rudelich, Herbert Jim O'Connor 2 Rudelich, Herbert The Dear Departed Henry Slater No Exit Estelle 3 Vincent, Martina 4 Vincent, Martina Happy Days Winnie The Dear Departed Mrs. Jordan 5 Vincent, Martina The Glass Menagerie Amanda Wingfield 6 Benito, Gisela 7 Benito, Gisela The Dear Departed Mrs. Slater 8 Sirignano, Emily The Dear Departed Victoria Slater 9 Harbinger, Nicholas No Exit Garcin 10 Harbinger, Nicholas Happy Days Willie 11 Harbinger, Nicholas The Dear Departed Ben Jordan 12 Phillipon, Marie-Odile No Exit Inez 13 Phillipon, Marie-Odile The Glass Menagerie Laura Wingfield 14 Gunter, Thomas No Exit Valet The Glass Menagerie Tom Wingfield 15 Gunter, Thomas 16 Gunter, Thomas The Dear Departed Abel Merryweather



1. 새로운 변수 생성





#### Increasing the Air Fare by \$10 in All Observations

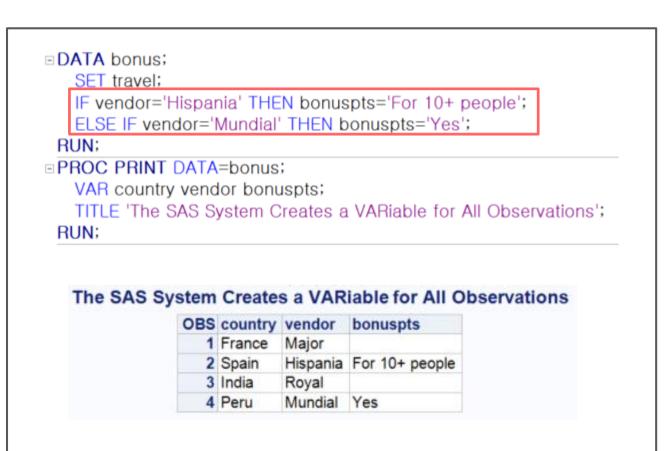
| OBS | country | aircost | newacost |
|-----|---------|---------|----------|
| 1   | France  | 793     | 803      |
| 2   | Spain   | 805     | 815      |
| 3   | India   |         |          |
| 4   | Peru    | 722     | 732      |



# Chapter8. 새로운 변수의 생성 2. IF문의 형식

#### IF 문 형식

IF 조건식1 THEN 실행문1; ELSE IF 조건식2 THEN 실행문2; ELSE 조건식3 THEN 실행문3;





3. 변수의 값 바꾸기

■ DATA newair;

SET travel;

newacost=aircost+10;

RUN;

■ PROC PRINT DATA=newair;

VAR country aircost newacost;

TITLE 'Increasing the Air Fare by \$10 in All Observations';

RUN:

#### Increasing the Air Fare by \$10 in All Observations

| OBS | country | aircost | newacost |
|-----|---------|---------|----------|
| 1   | France  | 793     | 803      |
| 2   | Spain   | 805     | 815      |
| 3   | India   |         |          |
| 4   | Peru    | 722     | 732      |

□ DATA newair2;
 SET travel;
 aircost=aircost+10;
 RUN;
 □ PROC PRINT DATA=newair2;
 VAR country aircost;
 TITLE 'Changing the InFORMATion in a VARiable';

#### Changing the InFORMATion in a VARiable

RUN;

| country | aircost                  |
|---------|--------------------------|
| France  | 803                      |
| Spain   | 815                      |
| India   |                          |
| Peru    | 732                      |
|         | France<br>Spain<br>India |



3. 변수의 값 바꾸기



805

722

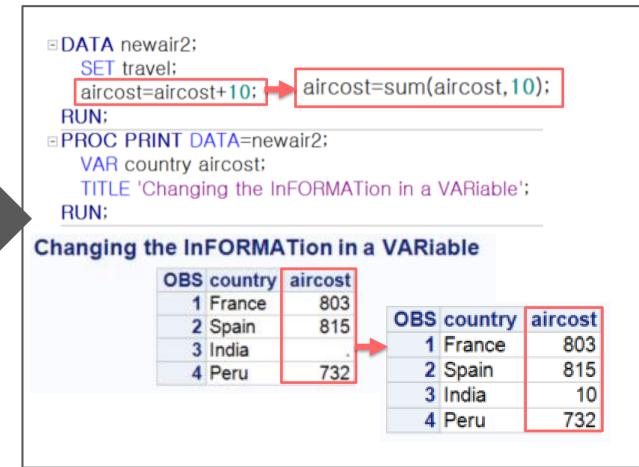
815

732

2 Spain

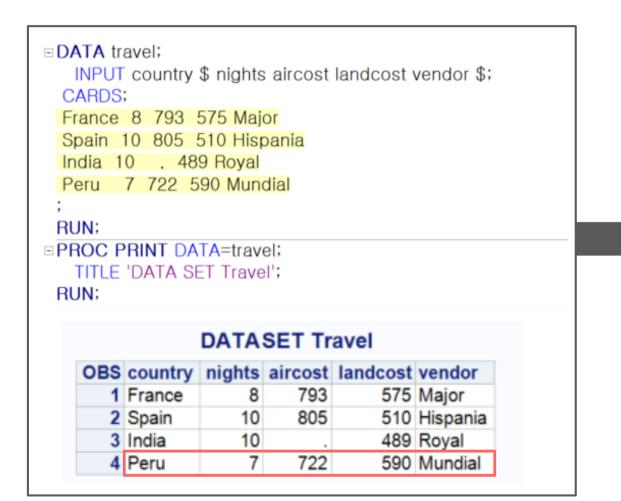
3 India

4 Peru





4. 특정 관찰값들 삭제하기 (IF, DELETE문)



■ DATA subset;
 SET travel;
 IF country='Peru' THEN delete;
 RUN;
 ■ PROC PRINT DATA=subset;
 TITLE 'Omitting a Discontinued Tour';
 RUN;

# Omitting a Discontinued Tour OBS country nights aircost landcost vendor 1 France 8 793 575 Major 2 Spain 10 805 510 Hispania 3 India 10 . 489 Royal



# Chapter8. 새로운 변수의 생성 5. ROUND, SUM 함수 이용해 새로운 변수생성

#### **ROUND, SUM**

ROUND(변수, 반올림숫자);

SUM(변수, 변수);

#### 참고)

X < 반올림숫자 - 반올림숫자/2: 버림

반올림숫자 - 반올림숫자/2 <X < 반올림숫자 + 반올림숫자/2: 반올림숫자

X > 반올림숫자 + 반올림숫자/2: 반올림

# ■ DATA moretour; SET tours; roundair=ROUND(aircost,50); totcostr=ROUND(aircost+landcost,100); sumcost=SUM(aircost,landcost); roundsum=ROUND(SUM(aircost,landcost),100); RUN; ■ PROC PRINT DATA=moretour; VAR country aircost landcost roundair totcostr sumcost roundsum; TITLE 'Rounding and Summing Values'; RUN:

|     |             | Rouna   | ing and  | Summir   | ig value | es      |          |
|-----|-------------|---------|----------|----------|----------|---------|----------|
| OBS | country     | aircost | landcost | roundair | totcostr | sumcost | roundsum |
| 1   | Japan       | 982     | 1020     | 1000     | 2000     | 2002    | 2000     |
| 2   | Greece      |         | 748      |          |          | 748     | 700      |
| 3   | New Zealand | 1368    | 1539     | 1350     | 2900     | 2907    | 2900     |
| 4   | Ireland     | 787     | 628      | 800      | 1400     | 1415    | 1400     |
| 5   | Venezuela   | 426     | 505      | 450      | 900      | 931     | 900      |
| 6   | Italy       | 852     | 598      | 850      | 1500     | 1450    | 1500     |
| 7   | USSR        | 1106    | 1024     | 1100     | 2100     | 2130    | 2100     |
| 8   | Switzerland | 816     | 834      | 800      | 1700     | 1650    | 1700     |
| 9   | Australia   | 1299    | 1169     | 1300     | 2500     | 2468    | 2500     |
| 10  | Brazil      | 682     | 610      | 700      | 1300     | 1292    | 1300     |

Davingling and Committee Values

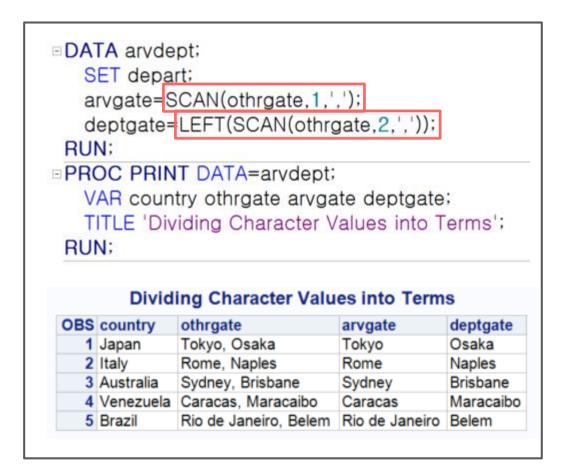


# Chapter8. 새로운 변수의 생성 6. SCAN, LEFT 함수 이용해 새로운 문자변수 생성

#### **SCAN, LEFT**

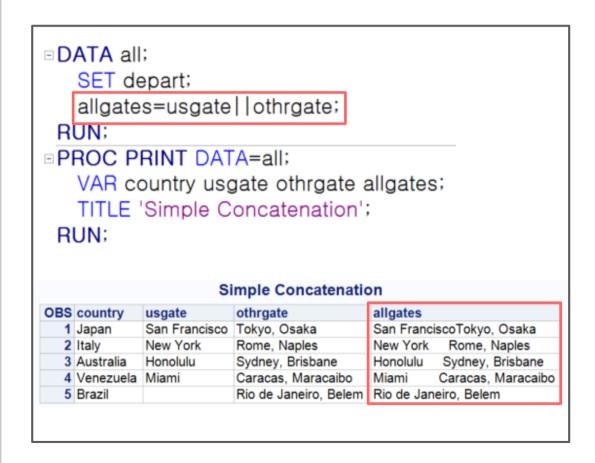
새변수 = SCAN(변수, 숫자, 구분자);

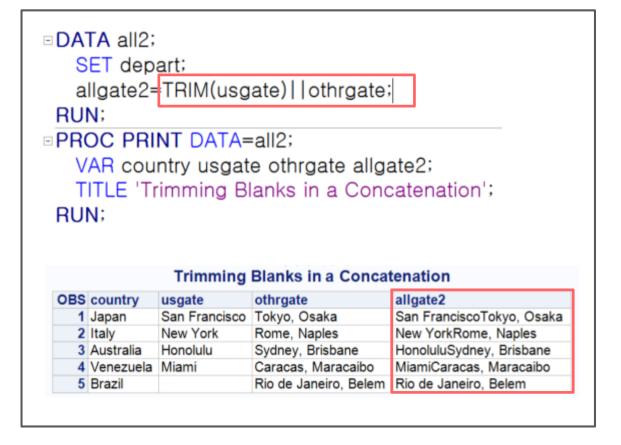
새변수 = LEFT(변수);





7. 연결연산자(||,TRIM함수)를 이용한 새로운 문자변수의 생성(1)







7. 연결연산자(II,TRIM함수)를 이용한 새로운 문자변수의 생성(2)

```
■DATA all3;

SET depart;

allgate3=TRIM(usgate)||', '||othrgate;

IF country='Brazil' THEN allgate3=othrgate;

RUN;

PROC PRINT DATA=all3;

VAR country usgate othrgate allgate3;

TITLE 'Readable Concatenated Values';

RUN;
```

|     |           | /alues        |                       |                             |
|-----|-----------|---------------|-----------------------|-----------------------------|
| OBS | country   | usgate        | othrgate              | allgate3                    |
| 1   | Japan     | San Francisco | Tokyo, Osaka          | San Francisco, Tokyo, Osaka |
| 2   | Italy     | New York      | Rome, Naples          | New York, Rome, Naples      |
| 3   | Australia | Honolulu      | Sydney, Brisbane      | Honolulu, Sydney, Brisbane  |
| 4   | Venezuela | Miami         | Caracas, Maracaibo    | Miami, Caracas, Maracaibo   |
| 5   | Brazil    |               | Rio de Janeiro, Belem | Rio de Janeiro, Belem       |
|     |           |               |                       |                             |

■DATA rebuild2;
 LENGTH arvgate \$ 14 deptgate \$ 9;
 SET arvdept;
 othrgat2=TRIM(arvgate)||', '||deptgate;
 RUN;
 ■PROC PRINT DATA=rebuild2;
 VAR country othrgate othrgat2;
 TITLE 'All Concatenated Pieces Appear';
 RUN;

| OBS | country   | othrgate              | othrgat2              |
|-----|-----------|-----------------------|-----------------------|
| 1   | Japan     | Tokyo, Osaka          | Tokyo, Osaka          |
| 2   | Italy     | Rome, Naples          | Rome, Naples          |
| 3   | Australia | Sydney, Brisbane      | Sydney, Brisbane      |
| 4   | Venezuela | Caracas, Maracaibo    | Caracas, Maracaibo    |
| 5   | Brazil    | Rio de Janeiro, Belem | Rio de Janeiro, Belem |