SAS拼表

第一部分:SAS拼表

一、SET

使用SET语句串接多个数据集;

DATA t3;

SET t1 t2;

RUN;

In [11]:
```
OPTIONS COMPRESS = YES;
```

Out[11]:

```
152  ods listing close;ods html5 (id=saspy_internal) file=stdout options(bitmap_mode='inline') devi
ce=svg style=HTMLBlue; ods
152! graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: STDOUT
153
154  OPTIONS COMPRESS = YES;
155
156  ods html5 (id=saspy_internal) close;ods listing;

157
```

In [12]:
```
DATA TEST_X1;
INPUT NAME $ PRODUCT $ TYPE $;
CARDS;
A CAR 40
B CAR 42
C BUS 44
D MOTO 9
E BUS 10
;
RUN;

PROC PRINT DATA = TEST_X1;
RUN;
```

Out[12]:

The SAS System

Obs	NAME	PRODUCT	TYPE
1	A	CAR	40
2	B	CAR	42
3	C	BUS	44
4	D	MOTO	9
5	E	BUS	10

In [13]:
```
DATA TEST_X2;
INPUT NAME $ PRODUCT $ TYPE $;
CARDS;
A CAR 40
C CAR 42
B BUS 44
H MOTO 9
J BUS 10
;
RUN;


PROC PRINT DATA = TEST_X2;
RUN;
```

Out[13]:

The SAS System

Obs	NAME	PRODUCT	TYPE
1	A	CAR	40
2	C	CAR	42
3	B	BUS	44
4	H	MOTO	9
5	J	BUS	10

In [14]:
```
DATA TEST_Y;
INPUT NAME $ NPRODUCT $ NTYPE $;
CARDS;
A APPLE 38
B BANANA 42
C CAT 44
D DOG 9
E EGG 10
;
RUN;

PROC PRINT DATA = TEST_Y;
RUN;
```

Out[14]:

The SAS System

Obs	NAME	NPRODUCT	NTYPE
1	A	APPLE	38
2	B	BANANA	42
3	C	CAT	44
4	D	DOG	9
5	E	EGG	10

In [15]:
```
DATA TEST_Z;
INPUT NAME $ NPRODUCT $ NTYPE $;
CARDS;
A APPLE 38
A BANANA 42
B CAT 44
B DOG 9
B EGG 10
;
RUN;

PROC PRINT DATA = TEST_Z;
RUN;
```

Out[15]:

The SAS System

Obs	NAME	NPRODUCT	NTYPE
1	A	APPLE	38
2	A	BANANA	42
3	B	CAT	44
4	B	DOG	9
5	B	EGG	10

In [16]:
```
DATA A1_1;
SET TEST_X1 TEST_X2;
RUN;

PROC PRINT DATA = A1_1;
RUN;
```

Out[16]:

The SAS System

Obs	NAME	PRODUCT	TYPE
1	A	CAR	40
2	B	CAR	42
3	C	BUS	44
4	D	MOTO	9
5	E	BUS	10
6	A	CAR	40
7	C	CAR	42
8	B	BUS	44
9	H	MOTO	9
10	J	BUS	10

In [17]:
```
DATA A1_2;
SET TEST_X1 TEST_Y;
RUN;

PROC PRINT DATA = A1_2;
RUN;
```

Out[17]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40		
2	B	CAR	42		
3	C	BUS	44		
4	D	MOTO	9		
5	E	BUS	10		
6	A			APPLE	38
7	B			BANANA	42
8	C			CAT	44
9	D			DOG	9
10	E			EGG	10

In [22]:
```
DATA A1_3;
SET
    TEST_X1
    TEST_Y(
        RENAME = (
            NPRODUCT = PRODUCT
            NTYPE = TYPE
        )
    )
;
RUN;

PROC PRINT DATA = A1_3;
RUN;
```

Out[22]:

The SAS System

Obs	NAME	PRODUCT	TYPE
1	A	CAR	40
2	B	CAR	42
3	C	BUS	44
4	D	MOTO	9
5	E	BUS	10
6	A	APPLE	38
7	B	BANANA	42
8	C	CAT	44
9	D	DOG	9
10	E	EGG	10

二、*MERGE*

使用MERGE语句并接多个数据集，数据集需按并接变量排序(PROC SORT);

DATA t3;

MERGE t1(IN=A) t2(IN=B);

BY n1;

IF A; / *左连接，右连接，内连接，全连接* /

RUN;

In [21]:
```
PROC SORT DATA = TEST_X2;
BY NAME;
RUN;

PROC SORT DATA = TEST_Y;
BY NAME;
RUN;

PROC SORT DATA = TEST_Z;
BY NAME;
RUN;
```

Out[21]:

```
311  ods listing close;ods html5 (id=saspy_internal) file=stdout options(bitmap_mode='inline') devi
ce=svg style=HTMLBlue; ods
311! graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: STDOUT
312
313  PROC SORT DATA = TEST_X2;
314  BY NAME;
315  RUN;
NOTE: There were 5 observations read from the data set WORK.TEST_X2.
NOTE: The data set WORK.TEST_X2 has 5 observations and 3 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

316
317  PROC SORT DATA = TEST_Y;
318  BY NAME;
319  RUN;
NOTE: There were 5 observations read from the data set WORK.TEST_Y.
NOTE: The data set WORK.TEST_Y has 5 observations and 3 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

320
321  PROC SORT DATA = TEST_Z;
322  BY NAME;
323  RUN;
NOTE: There were 5 observations read from the data set WORK.TEST_Z.
NOTE: The data set WORK.TEST_Z has 5 observations and 3 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

324
325  ods html5 (id=saspy_internal) close;ods listing;

326
```

In [33]:
```
DATA A2;
MERGE TEST_X2(IN=A) TEST_Y(IN=B);
BY NAME;
IF A;
RUN;

PROC PRINT DATA = A2;
RUN;
```

Out[33]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40	APPLE	38
2	B	BUS	44	BANANA	42
3	C	CAR	42	CAT	44
4	H	MOTO	9		
5	J	BUS	10		

In [24]:
```
DATA A2_1;
MERGE TEST_X2(IN=A) TEST_Y(IN=B);
BY NAME;
IF NOT A;
RUN;

PROC PRINT DATA = A2_1;
RUN;
```

Out[24]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	D			DOG	9
2	E			EGG	10

In [25]:
```
DATA A3;
MERGE TEST_X2(IN=A) TEST_Y(IN=B);
BY NAME;
IF B;
RUN;

PROC PRINT DATA = A3;
RUN;
```

Out[25]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40	APPLE	38
2	B	BUS	44	BANANA	42
3	C	CAR	42	CAT	44
4	D			DOG	9
5	E			EGG	10

In [26]:
```
DATA A4;
MERGE TEST_X2(IN=A) TEST_Y(IN=B);
BY NAME;
IF A AND B;
RUN;

PROC PRINT DATA = A4;
RUN;
```

Out[26]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40	APPLE	38
2	B	BUS	44	BANANA	42
3	C	CAR	42	CAT	44

In [27]:
```
DATA A5;
MERGE TEST_X2(IN=A) TEST_Y(IN=B);
BY NAME;
RUN;

PROC PRINT DATA = A5;
RUN;
```

Out[27]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40	APPLE	38
2	B	BUS	44	BANANA	42
3	C	CAR	42	CAT	44
4	D			DOG	9
5	E			EGG	10
6	H	MOTO	9		
7	J	BUS	10		

In [28]:
```
DATA A51;
MERGE TEST_X2(IN=A) TEST_Y(IN=B);
BY NAME;
IF A OR B;
RUN;

PROC PRINT DATA = A51;
RUN;
```

Out[28]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40	APPLE	38
2	B	BUS	44	BANANA	42
3	C	CAR	42	CAT	44
4	D			DOG	9
5	E			EGG	10
6	H	MOTO	9		
7	J	BUS	10		

In [34]:
```
DATA A6;
MERGE TEST_X2(IN=A) TEST_Z(IN=B);
BY NAME;
IF A;
RUN;

PROC PRINT DATA = A6;
RUN;
```

Out[34]:

The SAS System

Obs	NAME	PRODUCT	TYPE	NPRODUCT	NTYPE
1	A	CAR	40	APPLE	38
2	A	CAR	40	BANANA	42
3	B	BUS	44	CAT	44
4	B	BUS	44	DOG	9
5	B	BUS	44	EGG	10
6	C	CAR	42		
7	H	MOTO	9		
8	J	BUS	10		

并接匹配

1.一对一匹配

2.一对多匹配

3.多对多匹配(PROC SQL)

MERGE不能实现多对多，非则通过SQL实现；

In [40]:
```
DATA A7;
MERGE TEST_Z(IN=A) TEST_Z(IN=B);
BY NAME;
IF A AND B;
RUN;

PROC PRINT DATA = A7;
RUN;
```

Out[40]:

The SAS System

Obs	NAME	NPRODUCT	NTYPE
1	A	APPLE	38
2	A	BANANA	42
3	B	CAT	44
4	B	DOG	9
5	B	EGG	10

In [41]:
```
PROC SQL;
SELECT
    A.NAME AS ANAME,
    A.NPRODUCT AS ANPRODUCT,
    A.NTYPE AS ANTYPE,
    B.NPRODUCT AS BNPRODUCT,
    B.NTYPE AS BNTYPE
FROM TEST_Z AS A LEFT JOIN TEST_Z AS B
ON A.NAME = B.NAME
;
QUIT;
```

Out[41]:

The SAS System

ANAME	ANPRODUCT	ANTYPE	BNPRODUCT	BNTYPE
A	APPLE	38	APPLE	38
A	BANANA	42	APPLE	38
A	APPLE	38	BANANA	42
A	BANANA	42	BANANA	42
B	CAT	44	CAT	44
B	DOG	9	CAT	44
B	EGG	10	CAT	44
B	CAT	44	DOG	9
B	DOG	9	DOG	9
B	EGG	10	DOG	9
B	CAT	44	EGG	10
B	DOG	9	EGG	10
B	EGG	10	EGG	10