SAS进阶

第一部分 SAS格式

(1)SAS格式

• SAS变量包含数值型、字符型变量;

• 输入格式、输出格式;

• Format Categories;

(2)示例

$$\text{原始数据值} \xrightarrow[\text{输入格式}]{Informat} SAS\text{值} \xrightarrow[\text{输出格式}]{Format} \text{格式化}SAS\text{值}$$

$$\$548,231 \xrightarrow[dollar10.]{Informat} 548231 \xrightarrow[comma8.]{Format} 548,231$$

$$25jan2004 \xrightarrow[date9.]{Informat} 16095 \xrightarrow[MMDDYY10.]{Format} 01/25/2004$$

In [8]:
```
/* SAS INFORMAT */
DATA DEMOY;
INPUT NAME $11. BIRTH HEIGHT;
INFORMAT BIRTH YYMMDD10. HEIGHT 5.1;
CARDS;
LIXIAO1      1959/10/21 170.5
LIXIAO2      1959/10/22 176.5
WANGMING     1992/02/21 177.8
;
RUN;

PROC PRINT DATA = DEMOY;
RUN;
```

Out[8]:

The SAS System

Obs	NAME	BIRTH	HEIGHT
1	LIXIAO1	-72	170.5
2	LIXIAO2	-71	176.5
3	WANGMING	11739	177.8

In [9]:
```
DATA DEMON;
INPUT NAME $11. BIRTH $11. HEIGHT;
CARDS;
LIXIAO1      1959/10/21 170.5
LIXIAO2      1959/10/22 176.5
WANGMING     1992/02/21 177.8
;
RUN;

PROC PRINT DATA = DEMON;
RUN;
```

Out[9]:

The SAS System

Obs	NAME	BIRTH	HEIGHT
1	LIXIAO1	1959/10/2	1
2	LIXIAO2	1959/10/2	2
3	WANGMING	1992/02/2	1

In [10]:
```
DATA DDATE;
SDATE = "01JAN2018"D;
F1DATE = SDATE;
F2DATE = SDATE;
F3DATE = SDATE;
RUN;

PROC PRINT DATA = DDATE;
RUN;


DATA DDATE;
SDATE = "01JAN2018"D;
FORMAT F1DATE YYMMDD10. F2DATE YYMMDD8. F3DATE YYMMDD6.;
F1DATE = SDATE;
F2DATE = SDATE;
F3DATE = SDATE;
RUN;

PROC PRINT DATA = DDATE;
RUN;
```

Out[10]:

The SAS System

Obs	SDATE	F1DATE	F2DATE	F3DATE
1	21185	21185	21185	21185

The SAS System

Obs	SDATE	F1DATE	F2DATE	F3DATE
1	21185	2018-01-01	18-01-01	180101

第二部分 SAS日期

(1)日期转换

SAS 将日期和时间存储为一个唯一数字；

• SAS 日期值，如"21DEC2018"D；

介于 1960 年 1 月 1 日和指定日期之间的天数；

• SAS 时间值，如"09:39:00"T；

自当日午夜 12 时算起的秒数，SAS 时间值介于 0 和 86400 之间；

• SAS 日期时间值，如"21DEC2018 09:39:00"DT；

自 1960 年 1 月 1 日和指定日期内的小时/分钟/秒之间的秒数的值；

In [11]:
```
/* SAS DATE */
DATA SASDATE;
SDATE = "21DEC2018"D;
STIME = "09:39:00"T;
SDATETIME = "21DEC2018 09:39:00"DT;
RUN;

PROC PRINT DATA = SASDATE;
RUN;


DATA SASDATE_TR;
SDATE = "21DEC2018"D;
STIME = "09:39:00"T;
SDATETIME = "21DEC2018 09:39:00"DT;
FORMAT F1DATE YYMMDD10. F2DATE YYMMDD8. FTIME TIME10. FDATETIME DATETIME20.;
F1DATE = SDATE;
F2DATE = SDATE;
FTIME = STIME;
FDATETIME = SDATETIME;
RUN;

PROC PRINT DATA = SASDATE_TR;
RUN;
```

Out[11]:

The SAS System

Obs	SDATE	STIME	SDATETIME
1	21539	34740	1861004340

The SAS System

Obs	SDATE	STIME	SDATETIME	F1DATE	F2DATE	FTIME	FDATETIME
1	21539	34740	1861004340	2018-12-21	18-12-21	9:39:00	21DEC2018:09:39:

(2)日期计算

• 直接进行数值运算；

• 内置函数；

INTNX

计算从特定日期开始相隔指定间隔的日期；

INTNX(interval, start-from, increment, <**'alignment'**>)；

其中，**'alignment'**：

“B”：Beginning，返回的日期时间对齐间隔时间的开始；

“M”：Middle，返回的日期时间对齐间隔时间的中间；

“E”：End，返回的日期时间对齐间隔时间的结束；

“S”：Same，返回的日期时间和开始时间具有相同的对齐；

INTCK

计算两个日期间的间隔长度；

INTNX(interval, start-date, end-date, <**'method'**>)；

其中，**'method'**：

“C”：Continuous，测量连续时间，间隔是根据开始日期而开始计算；

“D”：Discrete，测量离散时间，离散方法计算间隔边界；

• 其它函数：

TODAY() ，YEAR()，MONTH()，DAY() ，WEEKEND()等等

In [12]:
```
/* INTNX */
DATA DINTNX;
SDATE = "01JAN2018"D;
FORMAT FDATE SDATED SDATEM SDATEY YYMMDD10.;
FDATE = SDATE;
SDATED = INTNX("DAY", SDATE, 1);
SDATEM = INTNX("MONTH", SDATE, 1);
SDATEY = INTNX("YEAR", SDATE, 1);
RUN;

PROC PRINT DATA = DINTNX;
RUN;


DATA DINTNXM;
SDATE = "02JAN2018"D;
FORMAT FDATE SDATEMB SDATEMM SDATEME SDATEMS YYMMDD10.;
FDATE = SDATE;
SDATEMB = INTNX("MONTH", SDATE, 1, "B");
SDATEMM = INTNX("MONTH", SDATE, 1, "M");
SDATEME = INTNX("MONTH", SDATE, 1, "E");
SDATEMS = INTNX("MONTH", SDATE, 1, "S");
RUN;

PROC PRINT DATA = DINTNXM;
RUN;


DATA DINTNXY;
SDATE = "02JAN2018"D;
FORMAT FDATE SDATEMB SDATEMM SDATEME SDATEMS YYMMDD10.;
FDATE = SDATE;
SDATEMB = INTNX("YEAR", SDATE, 1, "B");
SDATEMM = INTNX("YEAR", SDATE, 1, "M");
SDATEME = INTNX("YEAR", SDATE, 1, "E");
SDATEMS = INTNX("YEAR", SDATE, 1, "S");
RUN;

PROC PRINT DATA = DINTNXY;
RUN;
```

Out[12]:

The SAS System

Obs	SDATE	FDATE	SDATED	SDATEM	SDATEY
1	21185	2018-01-01	2018-01-02	2018-02-01	2019-01-01

The SAS System

Obs	SDATE	FDATE	SDATEMB	SDATEMM	SDATEME	SDATEMS
1	21186	2018-01-02	2018-02-01	2018-02-14	2018-02-28	2018-02-02

The SAS System

Obs	SDATE	FDATE	SDATEMB	SDATEMM	SDATEME	SDATEMS
1	21186	2018-01-02	2019-01-01	2019-07-02	2019-12-31	2019-01-02

In [13]:
```
/* INTCK */
DATA DINTCK;
SDATE_M11 = "01DEC2018"D;
SDATE_M12 = "05NOV2018"D;
GAP_DAYS = INTCK("DAY", SDATE_M12, SDATE_M11);
GAP_MONS = INTCK("MONTH", SDATE_M12, SDATE_M11);
RUN;

PROC PRINT DATA = DINTCK;
RUN;
```

Out[13]:

The SAS System

Obs	SDATE_M11	SDATE_M12	GAP_DAYS	GAP_MONS
1	21519	21493	26	1

第三部分 SAS类型转换

SAS类型转换 PUT/INPUT

基本形式：

数值转字符: VAR_A = PUT(VAR_B, $w.)

字符转数值: VAR_B = INPUT(VAR_A, w.)

VAR_A为字符类型，VAR_B为数值类型，w为字段长度；

In [14]:
```
/* PUT INPUT */
DATA DPUT1;
DATE18 = 20181221;
DATEP18 = PUT(DATE18, $8.);
RUN;

PROC PRINT DATA = DPUT1;
RUN;


DATA DPUT2;
DATE18 = 20181221;
DATEP18 = PUT(DATE18, $8.);
DATEI18 = INPUT(DATEP18, YYMMDD8.);
/*      DATEI18 = INPUT(PUT(DATE18, $8.), YYMMDD8.);*/
DATEPN18 = PUT(DATEI18, YYMMDDN8.);
DATEIN18 = INPUT(DATEPN18, 8.);
RUN;

PROC PRINT DATA = DPUT2;
RUN;
```

Out[14]:

The SAS System

Obs	DATE18	DATEP18
1	20181221	20181221

The SAS System

Obs	DATE18	DATEP18	DATEI18	DATEPN18	DATEIN18
1	20181221	20181221	21539	20181221	20181221

补充 @time 20191004

In [15]:
```
DATA TA;
A1 = PUT(20190101, $8.);
A2 = PUT(20190101, 8.);

B1 = INPUT("20190101", $8.);
B2 = INPUT("20190101", 8.);
B3 = INPUT("20190101", YYMMDD8.);
RUN;

PROC PRINT DATA = TA;
RUN;
```

Out[15]:

The SAS System

Obs	A1	A2	B1	B2	B3
1	20190101	20190101	20190101	20190101	21550

In [16]:
```
PROC SQL;
CREATE TABLE TB1 AS
SELECT

/* -ERROR- */
/*  PUT(20190101, $8.), */
/* -ERROR- */
    PUT(20190101, 8.),

    INPUT("20190101", $8.),
    INPUT("20190101", 8.),
    INPUT("20190101", YYMMDD8.),
    1
FROM TA;
QUIT;

PROC PRINT DATA = TB1;
RUN;
```

Out[16]:

The SAS System

Obs	_TEMA001	_TEMA002	_TEMA003	_TEMA004	_TEMA005
1	20190101	20190101	20190101	21550	1

In [17]:
```
PROC SQL;
CREATE TABLE TB2 AS
SELECT

/* -ERROR- */
    PUT(20190101, $8.),
/* -ERROR- */
    PUT(20190101, 8.),

    INPUT("20190101", $8.),
    INPUT("20190101", 8.),
    INPUT("20190101", YYMMDD8.),
    1
FROM TA;
QUIT;

PROC PRINT DATA = TB2;
RUN;
```

Out[17]:

```
405  ods listing close;ods html5 (id=saspy_internal) file=stdout options(bitmap_mode='inline') devi
ce=svg style=HTMLBlue; ods
405! graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: STDOUT
406
407  PROC SQL;
408  CREATE TABLE TB2 AS
409  SELECT
410
411  /* -ERROR- */
412      PUT(20190101, $8.),
413  /* -ERROR- */
414      PUT(20190101, 8.),
415
416      INPUT("20190101", $8.),
417      INPUT("20190101", 8.),
418      INPUT("20190101", YYMMDD8.),
419      1
420  FROM TA;
ERROR: Character format $ in PUT function requires a character argument.
NOTE: PROC SQL set option NOEXEC and will continue to check the syntax of statements.
421  QUIT;
NOTE: The SAS System stopped processing this step because of errors.
NOTE: PROCEDURE SQL used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

422
423  PROC PRINT DATA = TB2;
ERROR: File WORK.TB2.DATA does not exist.
424  RUN;
NOTE: The SAS System stopped processing this step because of errors.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

425
426  ods html5 (id=saspy_internal) close;ods listing;

427
```

In [ ]: