

How Do You Use Look-up Tables?

Philip R Holland
Holland Numerics Ltd



How Do You Use Look-up Tables?

Agenda

- Introduction
- Data Step Merge
- PROC SQL Join
- Formats
- Generated IF .. THEN .. ELSE
- Generated SELECT .. WHEN .. OTHERWISE
- Comparison of Techniques
- Conclusions



How Do You Use Look-up Tables?

Introduction - 1

The MAIN SAS data set contains multiple copies of SASHELP.CARS (428 observations and 15 variables) saved in a single WORK data set to increase the size of this data set:

```
%LET mult = 1; /* 10, 100, 1000, 2000, 5000 */  
DATA main;  
    SET sashelp.cars;  
    DO i = 1 TO &mult.;  
        OUTPUT;  
    END;  
RUN;
```



How Do You Use Look-up Tables?

Introduction - 2

LOOKUP_ORIGIN (3 observations and 3 variables) is a summary of SASHELP.CARS:

LOOKUP_ORIGIN ▾

 |  Filter and Sort |  Query Builder | Data ▾ | Describe ▾ | Graph ▾ | Analyze ▾ | Export ▾ | Send To ▾ | 

	 Origin	 make_n	 type_n
1	Asia	14	6
2	Europe	10	4
3	USA	14	5



How Do You Use Look-up Tables?

Introduction - 3

LOOKUP_TYPE (15 observations and 6 variables) is a summary of SASHELP.CARS:

LOOKUP_TYPE ▾

Filter and Sort Query Builder Data ▾ Describe ▾ Graph ▾ Analyze ▾ Export ▾ Send To ▾

	Origin	Type	type_model_n	type_make_n	type_msrp_mean	type_horsepower_max
1	Asia	Hybrid	3	2	19920	110
2	Asia	SUV	25	11	29569	325
3	Asia	Sedan	93	13	22763.968085	340
4	Asia	Sports	17	9	32510.647059	300
5	Asia	Truck	8	4	20383.625	305
6	Asia	Wagon	11	9	23143.727273	315
7	Europe	SUV	10	6	48346	340
8	Europe	Sedan	76	8	42992.051282	493
9	Europe	Sports	23	5	71998.695652	493
10	Europe	Wagon	12	6	37851.25	340
11	USA	SUV	25	12	34589.2	325
12	USA	Sedan	90	12	25638.833333	302
13	USA	Sports	9	6	45257.222222	500
14	USA	Truck	16	5	27220.25	345
15	USA	Wagon	7	6	22345.714286	250



How Do You Use Look-up Tables?

Introduction - 4

LOOKUP_MAKE (38 observations and 6 variables) is a summary of SASHELP.CARS:

LOOKUP_MAKE ▾

Filter and Sort Query Builder Data ▾ Describe ▾ Graph ▾ Analyze ▾ Export ▾ Send To ▾

	Origin	Make	make_model_n	make_type_n	make_msrp_mean	make_horsepower_max
1	Asia	Acura	7	3	42938.571429	290
2	Asia	Honda	17	4	21434.705882	240
3	Asia	Hyundai	12	3	17476.5	194
4	Asia	Infiniti	7	2	36070	340
5	Asia	Isuzu	2	1	26149	275
6	Asia	Kia	11	3	15875.909091	195
7	Asia	Lexus	11	4	44215.454545	300
8	Asia	Mazda	11	4	21770.727273	238
9	Asia	Mitsubishi	13	4	23423.615385	271
10	Asia	Nissan	17	5	24730.941176	305
11	Asia	Scion	2	2	13565	108
12	Asia	Subaru	11	4	25501.818182	300
13	Asia	Suzuki	8	3	16230.25	185
14	Asia	Toyota	28	6	22524.464286	325
15	Europe	Audi	19	3	43307.894737	450



How Do You Use Look-up Tables?

Data Step Merge

```
PROC SORT DATA = main OUT = datastepmerge1; BY origin make; RUN;

DATA datastepmerge2;
  MERGE datastepmerge1 lookup_origin;
  BY origin;
RUN;

DATA datastepmerge3;
  MERGE datastepmerge2 lookup_make;
  BY origin make;
  IF msrp > make_msrp_mean THEN make_msrp_flag = 1;
                                ELSE make_msrp_flag = 0;
  make_horsepower_pct = 100 * horsepower / make_horsepower_max;
RUN;

PROC SORT DATA = datastepmerge3 OUT = datastepmerge4; BY origin type; RUN;

DATA datastepmerge5;
  MERGE datastepmerge4 lookup_type;
  BY origin type;
  IF msrp > type_msrp_mean THEN type_msrp_flag = 1;
                                ELSE type_msrp_flag = 0;
  type_horsepower_pct = 100 * horsepower / type_horsepower_max;
RUN;
```



How Do You Use Look-up Tables?

PROC SQL Join - 1

```
PROC SQL;
  CREATE TABLE sqljoin1 AS
    SELECT a.*
      ,b.make_n
      ,b.type_n
      ,c.make_model_n
      ,c.make_type_n
      ,c.make_msrp_mean
      ,(CASE
        WHEN a.msrp > c.make_msrp_mean THEN 1
        ELSE 0
        END) AS make_msrp_flag
      ,c.make_horsepower_max
      ,(100 * a.horsepower / c.make_horsepower_max) AS make_horsepower_pct
      ,d.type_model_n
      ,d.type_make_n
      ,d.type_msrp_mean
      ,(CASE
        WHEN a.msrp > d.type_msrp_mean THEN 1
        ELSE 0
        END) AS type_msrp_flag
      ,d.type_horsepower_max
      ,(100 * a.horsepower / d.type_horsepower_max) AS type_horsepower_pct
```



How Do You Use Look-up Tables?

PROC SQL Join - 2

```
FROM    main a
LEFT JOIN
        lookup_origin b
ON      a.origin = b.origin
LEFT JOIN
        lookup_make c
ON      a.origin = c.origin AND a.make = c.make
LEFT JOIN
        lookup_type d
ON      a.origin = d.origin AND a.type = d.type
;
QUIT;
```



How Do You Use Look-up Tables?

Formats - 1

```
DATA format_origin;  
  LENGTH fmtname $7 start $80 label 8 type hlo $1;  
  SET lookup_origin;  
  type = 'I';  
  hlo = ' ';  
  start = origin;  
  fmtname = 'originm';  
  label = make_n;  
  output;  
  fmtname = 'origint';  
  label = type_n;  
  output;  
RUN;  
  
PROC SORT DATA = format_origin NODUPKEY; BY fmtname start; RUN;  
  
PROC FORMAT CNTLIN = format_origin; RUN;
```



How Do You Use Look-up Tables?

Formats - 2

```
%MACRO generate_format(level1=, level2=);
  DATA format_&level1.;
    LENGTH fmtname $7 start $80 label 8 type hlo $1;
    SET lookup_&level1.;
    type = 'I';
    hlo = ' ';
    start = CATX('|', origin, &level1.);
    fmtname = "&level1.c";
    label = &level1._model_n;
    output;
    fmtname = "&level1.x";
    label = &level1._&level2._n;
    output;
    fmtname = "&level1.p";
    label = &level1._msrp_mean;
    output;
    fmtname = "&level1.h";
    label = &level1._horsepower_max;
    output;
  RUN;

  PROC SORT DATA = format_&level1. NODUPKEY; BY fmtname start; RUN;

  PROC FORMAT CNTLIN = format_&level1.; RUN;
%MEND generate_format;
```

How Do You Use Look-up Tables?

Formats - 3

```
%generate_format(level1=make, level2=type);
%generate_format(level1=type, level2=make);

DATA format1;
  SET main;
  make_n = INPUT(origin, originm.);
  type_n = INPUT(origin, origint.);
  make_model_n = INPUT(CATX('|', origin, make), makec.);
  make_type_n = INPUT(CATX('|', origin, make), makex.);
  make_msrp_mean = INPUT(CATX('|', origin, make), makep.);
  IF msrp > make_msrp_mean THEN make_msrp_flag = 1;
  ELSE make_msrp_flag = 0;
  make_horsepower_max = INPUT(CATX('|', origin, make), makeh.);
  make_horsepower_pct = 100 * horsepower / make_horsepower_max;
  type_model_n = INPUT(CATX('|', origin, type), typec.);
  type_make_n = INPUT(CATX('|', origin, type), typex.);
  type_msrp_mean = INPUT(CATX('|', origin, type), typep.);
  IF msrp > type_msrp_mean THEN type_msrp_flag = 1;
  ELSE type_msrp_flag = 0;
  type_horsepower_max = INPUT(CATX('|', origin, type), typeh.);
  type_horsepower_pct = 100 * horsepower / type_horsepower_max;
RUN;
```



How Do You Use Look-up Tables?

Generated IF .. THEN .. ELSE - 1

```
FILENAME srcif CATALOG "work.generateif";

DATA _NULL_;
  SET lookup_origin END = eof;
  FILE srcif(origin.source);
  IF _N_ = 1 THEN PUT "IF origin = '" origin +(-1) "'" THEN DO;";
  ELSE PUT "ELSE IF origin = '" origin +(-1) "'" THEN DO;";
  PUT "make_n = " make_n ";";
  PUT "type_n = " type_n ";";
  PUT "END;";
RUN;
```



How Do You Use Look-up Tables?

Generated IF .. THEN .. ELSE - 2

```
%MACRO generate_if(level1=, level2=);
  DATA _NULL_;
    SET lookup_&level1. END = eof;
    FILE srcif(&level1..source);
    IF _N_ = 1 THEN PUT "IF origin = '" origin +(-1) "' AND &level1. = '"
                        &level1. +(-1) "' THEN DO;";
    ELSE PUT "ELSE IF origin = '" origin +(-1) "' AND &level1. = '"
            &level1. +(-1) "' THEN DO;";
    PUT "&level1._model_n = " &level1._model_n ";";
    PUT "&level1._&level2._n = " &level1._&level2._n ";";
    PUT "&level1._msrp_mean = " &level1._msrp_mean ";";
    PUT "IF msrp > &level1._msrp_mean THEN &level1._msrp_flag = 1;";
    PUT "                                ELSE &level1._msrp_flag = 0;";
    PUT "&level1._horsepower_max = " &level1._horsepower_max ";";
    PUT "&level1._horsepower_pct = 100 * horsepower / &level1._horsepower_max;";
    PUT "END;";
  RUN;
%MEND generate_if;
```



How Do You Use Look-up Tables?

Generated IF .. THEN .. ELSE - 3

```
%generate_if(level1=make, level2=type);  
%generate_if(level1=type, level2=make);  
  
DATA generateif1;  
    SET main;  
    %INCLUDE srcif(origin.source);  
    %INCLUDE srcif(make.source);  
    %INCLUDE srcif(type.source);  
RUN;
```



How Do You Use Look-up Tables?

Generated SELECT .. WHEN .. OTHERWISE - 1

```
FILENAME srcsel CATALOG "work.generateselect";

DATA _NULL_;
  SET lookup_origin END = eof;
  FILE srcsel(origin.source);
  IF _N_ = 1 THEN PUT "SELECT;";
  PUT "WHEN (origin = '" origin + (-1) "') DO;";
  PUT "make_n = " make_n ";";
  PUT "type_n = " type_n ";";
  PUT "END;";
  IF eof THEN DO;
    PUT "OTHERWISE;";
    PUT "END;";
  END;
RUN;
```



How Do You Use Look-up Tables?

Generated SELECT .. WHEN .. OTHERWISE - 2

```
%MACRO generate_select(level1=, level2=);
  DATA _NULL_;
    SET lookup_&level1. END = eof;
    FILE srcsel(make.source);
    IF _N_ = 1 THEN PUT "SELECT;";
    PUT "WHEN (origin = '" origin +(-1) "' AND &level1. = '" &level1. +(-1) "') DO;";
    PUT "&level1._model_n = " &level1._model_n ";";
    PUT "&level1._&level2._n = " &level1._&level2._n ";";
    PUT "&level1._msrp_mean = " &level1._msrp_mean ";";
    PUT "IF msrp > &level1._msrp_mean THEN &level1._msrp_flag = 1;";
    PUT "                               ELSE &level1._msrp_flag = 0;";
    PUT "&level1._horsepower_max = " &level1._horsepower_max ";";
    PUT "&level1._horsepower_pct = 100 * horsepower / &level1._horsepower_max;";
    PUT "END;";
    IF eof THEN DO;
      PUT "OTHERWISE;";
      PUT "END;";
    END;
  RUN;
%MEND compare_select;
```



How Do You Use Look-up Tables?

Generated SELECT .. WHEN .. OTHERWISE - 3

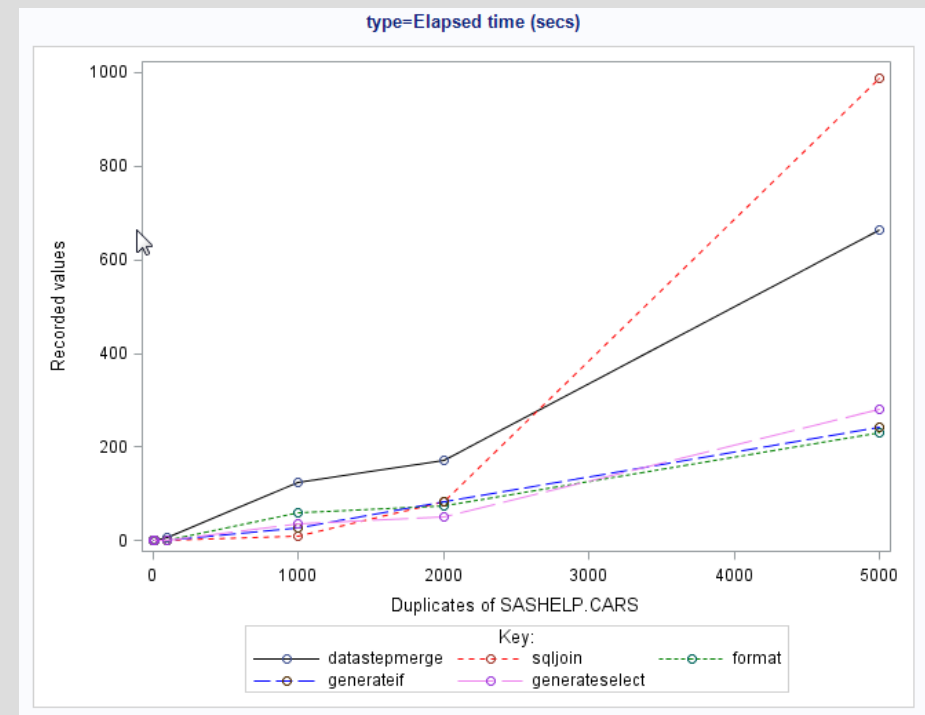
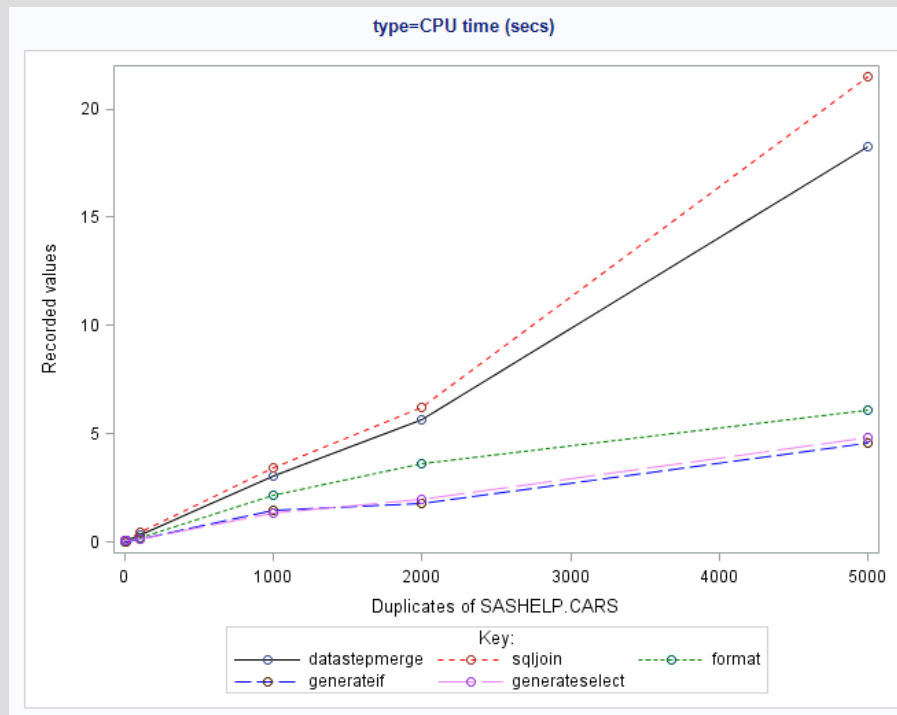
```
%generate_select(level1=make, level2=type);  
%generate_select(level1=type, level2=make);
```

```
DATA generateselect1;  
  SET main;  
  %INCLUDE srcsel(origin.source);  
  %INCLUDE srcsel(make.source);  
  %INCLUDE srcsel(type.source);  
RUN;
```



How Do You Use Look-up Tables?

Comparison of Techniques



How Do You Use Look-up Tables?

Conclusions

- All techniques for using look-up tables generate the same data, but which technique you use is often a personal choice.
- Only having 1 technique to choose from is limiting, so learning a new technique may help you:
 - speed-up your processing,
 - write your validation program in a completely different way to the source program,
 - improve your SAS programming.
- It is your choice!!



How Do You Use Look-up Tables?

Contact details

Philip R Holland, SAS Consultant

Holland Numerics Ltd

94 Green Drift, Royston, Herts SG8 5BT, UK

tel: +44-7714-279085

fax: +44-1763-242486

email: phil@hollandnumerics.com

web: www.hollandnumerics.com/SASPAPER.HTM