Dictionary Tables and Views: Essential Tools for Serious Applications
Frank Dilorio, CodeCrafters, Inc., Chapel Hill NC
Jeff Abolafia, Rho, Inc., Chapel Hill NC

INTRODUCTION

Dictionary tables were introduced to the SAS System in the early 1990’s, in Version 6.07. Laden with information that is often difficult, and sometimes impossible, to get through other means, they still appear to be on the outside of many programmers’ Bag of Tricks. This is both perplexing and unfortunate for as we will see in this paper, once their content and organization is understood, they are readily adapted for a range of applications that, to use an old saw, “are only limited by your imagination.”

This paper describes dictionary tables and their associated SASHELP library views. It:

- presents scenarios that show how they can be used
- gives high-level descriptions of some of the most important (a relative term, to be sure) tables
- identifies features of SQL and the macro language that are commonly used when writing programs that effectively use the tables
- shows examples of the tables’ use, emphasizing the use of SQL and the macro language interface

The reader should come away from the discussion with an understanding of the tables as well as with a checklist of SQL skills that are required to use the tables most effectively.

CONSIDER THE ALTERNATIVES

SAS programmers, particularly those writing utilities, have always needed high-level information about the SAS environment (option settings, data set characteristics, etc.). Consider a few realistic scenarios and the coding strategy required, both not using and using the dictionary tables:

- You want to insert the name of the current (Windows OS) directory in a footnote.
  **Without the tables:** use a _NULL_ DATA step to write a temporary file, then parse the SAS Log to identify the file’s location. Ugh!
  **Using the tables:** allocate a file name of ‘.’, then store the EXTFILES table’s XPATH value of the filename in a macro variable. See Example 2 for details.

- As part of “resetting” the SAS environment in an interactive session, you want to delete all global macro variables.
  **Without the tables:** set the MVARSIZE system option to 0 (forcing all global macro variables to be written to a catalog, rather than be held in memory), then delete entries of type MACRO from the catalog. Depending on the number and length of the macro variables, the I/O required by this technique can affect performance. Ugh!
  **Using the tables:** using the MACROS table, save the names of the global macro variables, then use %SYMDEL. See Example 4 for details.

- A project uses numerous SAS datasets, some of them with like-named variables. You want to identify instances of like-named variables with conflicting data types or lengths. For example, the variable GENDER may be stored as a character, length 1 in one dataset and numeric, length 3 in another. Ugh!
  **Without the tables:** create an output data set from the CONTENTS procedure, then sort by variable name, and flag inconsistencies in a DATA step.
  **Using the tables:** use the COLUMNS dictionary table and SQL to create a data set containing the errant variables. See Example 1 for details.

- You want to print the first “n” observations from every data set in a library.
  **Without the tables:** write a macro parameterized for data set name and number of observations to print. Then, invoke the macro, manually specifying the data sets in the library.
  **Using the tables:** use the COLUMNS dictionary table and SQL to create a data set ready for a reporting procedure (PRINT, REPORT, et al.). See Example 8 for details.

ABOUT THE TABLES

Dictionary tables make these tasks and others feasible with a minimal amount of coding effort. Let’s take a look at some of their characteristics. They are:

- **Metadata,** or “data about data.” They are data that are one or more steps away from operational or summary data, which is what most people think of when they hear “data.” Rather than containing salary levels, event codes, dosage levels, and the like, metadata identify the dataset containing the data, its creation date, the data type of the variables, and a host of other high-level details.

- **Available only in Version 6.07 or later** (hopefully this isn’t an issue for anyone!)

- **Always and automatically created** during SAS System startup. There is no system option to suppress their creation or maintenance.

- **Automatically maintained** during the course of the interactive session or batch job.

- **Read-only.** You cannot change the table or view organization. You can affect their contents by making changes in the SAS environment. Whenever you change a system option, create a dataset, delete a member from a catalog, add a label to a variable, etc. one or more tables are updated.

- **Accessible from SQL** with the reserved LIBNAME of DICTIONARY (yes, that’s a 10-letter LIBNAME!)

- **Accessible outside SQL** by using views defined in the LIBNAME of SASHELP. SASHELP is allocated automatically during startup.

- **Usually more efficient when accessed from SQL.** That is, using SQL to access, say, DICTIONARY.TABLES will usually execute faster than SQL, a DATA step, or a procedure reading SASHELP.VTABLE.

So much for what they are. Let’s take a look at what’s inside them.
WHAT’S OUT THERE?

Before you can use the tables, you have to know what tables are available. In addition to this paper, there are several sources.

SAS Online Doc. Enter “dictionary tables” as the search term.

Viewtable. While in an interactive SAS session, use SAS Explorer to display the contents of the SASHELP library. The views of the dictionary tables begin with ‘V’. Double-click on any ‘V’ object to invoke Viewtable to display its contents.

Alternately, if you know the object’s name, enter ‘viewtable object_name’ on the command line. For example, to see the view of the OPTIONS table, enter the following in the command line:

```sas
viewtable sashelp.voption
```

Viewtable is the ideal tool for gaining familiarity with the contents of the tables. You can browse the data, adjust the order of columns, and utilize the viewer to get an understanding of the data that would be difficult to gain from the documentation or PRINT procedure output alone.

CONTENTS. Provided you know the name of the view, you can use the CONTENTS procedure (or the CONTENTS statement in the DATASETS procedure) as follows:

```sas
proc contents data=sashelp.view;
```

In the above, view is the name of the view of interest (vcolumn, vcatalg, etc.).

SQL. The DESCRIBE statement in the SQL procedure provides some of the same information as CONTENTS. Results are written by default to the SAS Log. Examples for tables and views follow:

```sql
describe table dictionary.table;
describe view sashelp.view;
```

Programmatically, via a Macro. You can list the attributes (not the contents) of the tables and views programmatically. A macro to do this is shown below:

### EXHIBIT 1: PROGRAMMATIC LISTING OF TABLE ATTRIBUTES

```sas
%macro DictInfo;  
proc sql noprint;  
%if %scan(&sysver., 1) = 9 %then %do;  
select distinct memname, count(distinct memname) into :tbl separated by ' ', :ntbl from dictionary.dictionaries;  
%end;  
%else %do;  
%let tbl = CATALOGS COLUMNS INDEXES OPTIONS STYLES TABLES TITLES VIEWS;  
%let ntbl = 11;  
select memname, count(distinct memname) into :view separated by ' ', :nview from dictionary.views where memname like 'V%';  
%do i = 1 %to &ntbl.;  
%let item = %scan(&view., &i.);  
describe table dictionary.&item.;  
%end;  
%do i = 1 %to &nview.;  
describe view sashelp.&view.;  
%end;  
quit;  
%mend;
```

The macro uses SQL to write a description of each table to the SAS Log. The following, for example, is the description of the EXTFILES table.

```sql
NOTE: SQL table DICTIONARY.EXTFILES was created like:
create table DICTIONARY.EXTFILES

(  fileref char(8) label='Fileref',
  xpath char(1024) label='Path Name',
  xengine char(8) label='Engine Name'
);
```

The SQL output for Views is usually a bit more compact. For the EXTFILES table references above, we see:

```sql
NOTE: SQL view SASHELP.VEXTFL is defined as:
select * from DICTIONARY.EXTFILES;
```

It's important to emphasize that these tools simply identify what is available for use. Actually understanding their usefulness, nuances, and quirks is a bit of an art, most effectively developed while meeting real-world needs.

### TABLE AND VIEW ORGANIZATION

SAS Version 9 supports 22 dictionary tables (available in SQL) and 29 views of the tables (available with the reserved LIBNAME of SASHELP). Version 8 supports 11 tables and 17 views. The tables and their relationship to the views are presented in Exhibit 2.

### EXHIBIT 2: TABLE-VIEW CORRESPONDENCE

<table>
<thead>
<tr>
<th>DICTIONARY.table</th>
<th>SASHELP.view</th>
<th>Earliest Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>catalogs</td>
<td>vcatalg</td>
<td>6.07</td>
</tr>
<tr>
<td>columns</td>
<td>vcolumn</td>
<td>6.07</td>
</tr>
<tr>
<td>constraint_column_usage</td>
<td>vcnco</td>
<td>9.0</td>
</tr>
<tr>
<td>constraint_table_usage</td>
<td>vcntabu</td>
<td>9.0</td>
</tr>
<tr>
<td>dictionaries</td>
<td>vdictnry</td>
<td>9.0</td>
</tr>
<tr>
<td>engines</td>
<td>vengine</td>
<td>8.0</td>
</tr>
<tr>
<td>extfiles</td>
<td>vextfl</td>
<td>6.07</td>
</tr>
<tr>
<td>formats</td>
<td>vformat</td>
<td>9.0</td>
</tr>
<tr>
<td>goptions</td>
<td>vgopt, vallopt</td>
<td>9.0</td>
</tr>
<tr>
<td>indexes</td>
<td>vinix</td>
<td>6.07</td>
</tr>
<tr>
<td>libnames</td>
<td>vlibnam</td>
<td>9.0</td>
</tr>
<tr>
<td>macros</td>
<td>vmacro</td>
<td>6.11</td>
</tr>
<tr>
<td>members</td>
<td>vmember, vsacces, vscatlg, vslib, vstable, vsview, vstabvw</td>
<td>6.07</td>
</tr>
<tr>
<td>options</td>
<td>voption, vallopt</td>
<td>6.07</td>
</tr>
<tr>
<td>referential_constraints</td>
<td>vrefcon</td>
<td>9.0</td>
</tr>
<tr>
<td>remember</td>
<td>vrememb</td>
<td>9.0</td>
</tr>
<tr>
<td>styles</td>
<td>vstyle</td>
<td>8.0</td>
</tr>
<tr>
<td>tables</td>
<td>vtable</td>
<td>6.07</td>
</tr>
<tr>
<td>table_constraints</td>
<td>vlacbn</td>
<td>9.0</td>
</tr>
<tr>
<td>titles</td>
<td>vtitle</td>
<td>6.11</td>
</tr>
<tr>
<td>views</td>
<td>vview</td>
<td>6.07</td>
</tr>
</tbody>
</table>

Notes:

1. This is the first version that had the table and view. Readers familiar with the tables should note that some of the older,
familiar tables have new fields in Version 9. Notable among these are TABLES and COLUMNS.

Note, too, that some fields have subtly different content in Version 9 – CRDATE and MODATE in TABLES, for example, is now a date-time value, rather than a simple date value.

VALLOPT has all rows and columns from the GOPTIONS and OPTIONS tables.

The remainder of this section is devoted to descriptions of some of the more popular tables. “Popular” is, of course, a loaded term. In this context, it simply means tables that the authors have used most frequently in their daily work. Descriptions of most of the other tables are found in Appendix A. Exhibit 3, below, explains the format of the table descriptions.

**EXHIBIT 3: TABLE DESCRIPTION FORMAT**

<table>
<thead>
<tr>
<th>Dictionary</th>
<th>Table/View Name</th>
<th>Column</th>
<th>Type/Length</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name</strong></td>
<td><strong>Libname</strong></td>
<td><strong>Memname</strong></td>
<td><strong>Mentype</strong></td>
</tr>
<tr>
<td><strong>Column</strong></td>
<td><strong>Type</strong></td>
<td><strong>Length</strong></td>
<td><strong>Npos</strong></td>
</tr>
</tbody>
</table>

**COLUMNS**

Content: Information about variables in currently allocated data sets and views. See Exhibit 4 for details.

Granularity: Variable in a data set or view.

Comments:

1. Users familiar with this table should pay particular attention to new fields in Version 9.
2. Variable NAME, containing the name of a variable, is stored with the upper-lower case mix used at its creation. Subsetting on NAME should, therefore, either be explicit about case or use a function to fold NAME to consistently upper or lower-case.
3. Users of CONTENTS output data sets will be relieved to see the inscrutable-but-predictable 1-2 TYPE values replaced by the dictionary table’s more straightforward ‘num’ and ‘char’.
4. There are some situations where using a CONTENTS procedure output dataset is more effective than the dictionary tables and views. This is because these data sets are not normalized, unlike the dictionary tables.

A CONTENTS data set will have information at both the table and variable levels in the same observation (e.g., number of observations, number of variables, individual variable name, type, and length). Getting the same information from the dictionary tables requires a join of the MEMBERS, COLUMNS, and INDEXES tables.

This is not a huge problem, but is generally not as convenient as the “old” way. Indeed, there are items in the CONTENTS data sets that are simply not available anywhere in the dictionary data. Among these are whether the NODUPREC and NODUPKEY options were used during creation of the data set.

Used in Examples: 1, 5, 6.

**MACROS**

Content: Values of all current macro variables. See Exhibit 5 for details.

Granularity: Macro scope, name, and offset.

Comments: Long (greater than 200 character) values are broken into multiple observations, then distinguished by their offset value (first observation is offset 0, next is 200, and so on). To work with unique macro variable names, use a WHERE clause like “offset = 0”.

Used in Examples: 3, 4.

**Exhibit 4: COLUMNS Table**

<table>
<thead>
<tr>
<th>Dictionary</th>
<th>Column</th>
<th>Type/Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>dictionary.columns</td>
<td>$8 libname</td>
<td>Library name [upper case]</td>
</tr>
<tr>
<td>sashelp.vcolumn</td>
<td>$32 memname</td>
<td>Member name [upper case]</td>
</tr>
<tr>
<td></td>
<td>$8 mentype</td>
<td>Member type [DATA</td>
</tr>
<tr>
<td></td>
<td>$32 name</td>
<td>Column name [case as-is from data set creation]</td>
</tr>
<tr>
<td></td>
<td>$4 type</td>
<td>Column type [char</td>
</tr>
<tr>
<td></td>
<td>num length</td>
<td>Column length</td>
</tr>
<tr>
<td></td>
<td>num npos</td>
<td>Column position [offset within observation, e.g., 0, 1, 20]</td>
</tr>
<tr>
<td></td>
<td>num varnum</td>
<td>Column number in table [1, 2, 3, ...]</td>
</tr>
<tr>
<td></td>
<td>$256 label</td>
<td>Column label</td>
</tr>
<tr>
<td></td>
<td>$16 format</td>
<td>Column format [DATE9, $HEX22]</td>
</tr>
<tr>
<td></td>
<td>$16 informat</td>
<td>Column informat [MMDDYY10, 8.2]</td>
</tr>
<tr>
<td></td>
<td>$9 idxusage</td>
<td>Column index type [SIMPLE</td>
</tr>
<tr>
<td></td>
<td>num sortedby</td>
<td>Order in key sequence [0, 1, 2, ...]</td>
</tr>
<tr>
<td></td>
<td>$12 xtype</td>
<td>Extended type</td>
</tr>
<tr>
<td></td>
<td>$3 notnull</td>
<td>Not NULL? [no</td>
</tr>
<tr>
<td></td>
<td>num precision</td>
<td>Precision</td>
</tr>
<tr>
<td></td>
<td>num scale</td>
<td>Scale</td>
</tr>
</tbody>
</table>

**MEMBERS**

Content: Information about SAS data stores – data sets, views, template item stores, catalogs, MDDB, etc. See Exhibit 6 for details.

Granularity: Unique combination of library name, entity name, and entity type.

Comments:

1. The views of this table facilitate subsetting by entity type (VSTABLE selects only data sets) and library (VSLIB is useful when identifying a physical path for a LIBNAME).
2. See Note 4 in the description of the COLUMNS table, above.
**Exhibit 6: MEMBERS Table**

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
</tr>
<tr>
<td>memtype</td>
<td>$8</td>
</tr>
<tr>
<td>engine</td>
<td>$8</td>
</tr>
<tr>
<td>index</td>
<td>$32</td>
</tr>
<tr>
<td>path</td>
<td>$1024</td>
</tr>
</tbody>
</table>

**dictionary.members**

- **libname**: Library name [upper case]
- **memname**: Member name [upper case]
- **memtype**: Member type [DATA|VIEW|ITEMSTOR| ...]
- **engine**: Engine name [upper case]
- **index**: Indexes [yes|no] [not index names!]
- **path**: Path name [mixed case] [if concatenated, enclosed in parentheses, each path name quoted; otherwise, no parentheses or quotes]

**sashelp.vmember**

- All fields, all rows from dictionary table

**sashelp.vstable**

- libname

**sashelp.vscatlg**

- memname

**sashelp.vsview**

- memname

**sashelp.vstabvw**

- memname

**sashelp.vmember**

- memtype

**OPTIONS / GOPTIONS**

**Content**: Information about Base SAS and SAS/GRAPH options. See Exhibit 7 for details.

**Granularity**: Option.

**Comments**:

1. The OPTDESC field contains brief, useful descriptions of the option.
2. The GETOPTION function, used with the KEYWORD parameter, provides a useful and somewhat more powerful way to retrieve option settings.

**Used in Example**: 12.

**Exhibit 7: OPTIONS Table**

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>optname</td>
<td>$32</td>
</tr>
<tr>
<td>setting</td>
<td>$1024</td>
</tr>
<tr>
<td>optdesc</td>
<td>$160</td>
</tr>
<tr>
<td>level</td>
<td>$8</td>
</tr>
<tr>
<td>group</td>
<td>$32</td>
</tr>
</tbody>
</table>

**TABLES**

**Content**: Information about data sets and views. See Exhibit 8 for details.

**Granularity**: Unique combinations of library name, data set/view name.

**Comments**: Users familiar with this table should pay particular attention to new fields in Version 9. Also note the DATETIME formats and informats associated with CRDATE and MODATE. These used to be date values. The increased date-time precision is a welcome change, but is also one that will require modification of Version 8 programs that use these values.

**Used in Examples**: 8, 9, 10, 11.

**Exhibit 8: TABLES Table**

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
</tr>
</thead>
</table>

**TITLES**

**Content**: Current titles and footnotes. See Exhibit 9 for details.

**Granularity**: Title / footnote

**Comments**: Only the first 256 characters of a title or footnote are stored.
TOOLS TO PUT THE TABLES TO WORK

Ask “n” SAS programmers for a solution to a problem and you’ll get at least “n” distinct answers! Almost all SAS users would acknowledge that it is a robust tool. At the same time, though, one would be hard-pressed to argue that effective use of the metadata described in this paper would not require familiarity with Structured Query Language (SQL) and the macro language.

It’s no accident that each of the examples below employs SQL and most use the macro language. Its table-joining ability, ease of identifying group characteristics, and macro language interface make it the ideal tool for building applications with the SAS metadata.

The macro language’s ability to conditionally execute some of, all of, or multiple statements, along with the %SYSFUNC function’s on-the-fly use of common functions, provides the programmer with a powerful toolset for building powerful, generalized utilities.

Exhibit 10 identifies some of the features of SQL and briefly describes why they are useful for building applications using the metadata. A complete description of each feature is well beyond the scope of this paper. The Exhibit is simply intended to be a starting point, based on the authors’ experience, of key SQL features used in building metadata-based applications.

Exhibit 11 identifies key macro language features used in metadata-based programs. As with the SQL discussion, above, we do not enter the syntax and usage thickets. The table simply identifies parts of the language the authors most frequently use.

EXAMPLES

There are many applications for the information contained in the dictionary tables. In this section we present a series of examples taken from “real world” applications. These are summarized in Exhibit 12. Application background, code and, when appropriate, output are shown. Notice that most of the examples show the tables and views used as part of macros. This is in keeping with the high-level, meta-data nature of the information. Abstracted information and generalized tools such as macros are a logical, complementary fit.

<table>
<thead>
<tr>
<th>Exhibit 12: Index of Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>#</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
</tbody>
</table>
Example 1: Conflicting Variable Attributes

A harsh and dreary reality of working with "real world" data is the potential for inconsistent attributes. Data sets that should have identical lengths, type, formats, etc. can have mismatches. A demographic data set, for example, might store GENDER as character, length 1 ("M" and "F"). Outsource, laboratory data can represent the same information as numeric (1 and 2), or character, length 6 ("Male" and "Female"). Havoc ensues when the data sets are combined.

Example 1 uses the COLUMNS table to identify inconsistencies between like-named variables. Regardless of the number of times a variable appears in WORK library data sets, it should have the same attribute (we look at length and type here, but could extend it to format, informat, and label). Thus if we have more than one value, we have an inconsistency. The Example sets up two data sets, A and B, then uses SQL to ferret out the problems.

Example 1 (Part 1)

```sas
data a;
x = 'a';
y = 'ppp';
run;
data b;
length y 3;
x = 'xx';
y = 1;
z = '2';
run;
proc sql noprint;
    create table temp as
    select libname, name, length, type
    from dictionary.columns
    where libname = 'WORK'
    group by name
    having count(distinct length) > 1
    | count(distinct type) > 1
    ;
quit;
```

Example 1 (Part 2)

```sas
Obs memname name length type
1 A x 1 char
2 B x 2 char
3 B y 3 num
4 A y 3 char
```

Example 2: Identify Current Directory

Programs often need to know the name of the directory from which they are executing. A program could, for instance, need to identify whether it is running in a test or production environment. This could be determined from the directory name.

In the following Example, run in Windows, we use the EXTFILES table to write the directory name (variable XPATH) to global macro variable CURRENT. We assign a FILENAME before the SQL code, then clean up and deassign it after SQL terminates. Once macro %CurrDir completes, the macro variable CURRENT will be available to the calling program.

Example 2

```sas
$macro CurrDir;
$global CURRENT;
    filename __temp '_.';
proc sql noprint;
    select xpath into :Current
        from dictionary.extfiles
        where fileref = '__TEMP';
quit;
    filename __temp clear;
%end;
```

Example 3: Identify All Global Macro Variables

If you have read to this point, you’re probably sold on the idea of using metadata as often as possible. And if you read this Example’s title, you’ll say “sure, I can write %put_global; to list macro variables, but I’d rather exercise the MACROS table instead.” With that in mind …

Example 3

```sas
proc sql noprint;
    create table _macvars_ as
        select *
            from dictionary.macros
        order by name;
quit;
```

```
Data set _MACVARS_ can be used by any reporting procedure to display the first 200 characters of all global macro variables (we use only OFFSET=0). The advantage of this approach compared to %put __global_is that it is cleaner. The %PUT statement spews variable names and values to the SAS Log in a non-obvious manner. This may be fine for a few macro variables, but when “lots” are involved, the more precise and controlled presentation shown in the Example is preferable.

Example 4: Delete Global Macro Variables

Repeated execution of programs in an interactive SAS session can create a mass of unwanted and/or incorrectly assigned macro variables. It would be helpful to have a way clean up the session and delete all global macro variables. Prior to Version 8 and the introduction of the %SYMDEL statement, this was difficult (but not impossible). The simplest way was to end the SAS session and begin a new one. Crude, but effective.

The Example below uses the MACROS table and the macro variable interface to create a list, stored in macro variable __GLOB, of all Global macro variables (the list will include __GLOB). We then pass the list to %SYMDEL.

Example 4

```sas
%mend DelMacVar;
proc sql noprint;
    select name into __glob separated by ' ' 
        from dictionary.macros
    order by name 
    scope = 'GLOBAL'
    ;
quit;
%global __glob;
%symdel __glob;
%mend;
```

A simple extension to DelMacVar is the addition of KEEP and/or DROP parameters, which would control the selection of macro variables in __GLOB.

Example 5: Use Variables Meeting Criteria

This Example demonstrates how the dictionary tables can be used in the first small step toward generalized code. Rather than hard-code a list of variable names – AE1, AE2, and so on – we use the COLUMNS table to identify these variables. Thus if one
data set has AE1 through AE5 and another has AE1 through AE3, no changes to the program will be needed. In the Example below, we create a list with the names and use it in an array definition in a DATA step.

Example 5 (Original)

```sas
proc sql noprint;
   select trim(name) into :vars separated by ' ' from dictionary.columns
       where libname='INDATA' and memname='AE' and
       substr(NAME,1,2)='AE' and type = 'char';
quit;
data test;
   set indata.ae;
   array ae(*) &vars.;
   do i = 1 to dim(ae);
      if ae(i) ^= ' ' then do;
         NonMissingAE = 1;
         leave;
      end;
   end;
   if NonMissingAE then output;
run;
```

A few caveats are warranted. First, this type of coding requires that you know the data – the program assumes that you want any character variable starting with AE. Be sure that the program performs the selection correctly. You want to avoid the inclusion of AESEV1, AERSLT1, and the like. Also consider another data-driven possibility, that there are no character variables beginning with AE. The ARRAY statement will be handed a null value of macro variable VARS, and will fail. A good practice when making lists is to create a count of elements in the list, then use the count to control the statements downstream of SQL. This is demonstrated below.

Example 5 (Revised)

```sas
%macro temp;
   proc sql noprint;
      select count(*), trim(name)
          into :nVars,   :vars separated by ' ' from dictionary.columns
             where libname = 'INDATA' and
             memname = 'AE' and
             substr(NAME,1,2) = 'AE' and
             type = 'char';
   quit;
   %if &nVars. > 0 %then %do;
      data test;
         set indata.ae;
         array ae(*) &vars.;
         do i = 1 to dim(ae);
            if ae(i) ^= ' ' then do;
               NonMissingAE = 1;
               leave;
            end;
         end;
         if NonMissingAE then output;
   %end;
%mend;
```

Example 6: Identify Features Incompatible with Transport Format

Remember the days of limits such as eight-character variable names, 200 byte character variables, and the like? They live on in reality as well as our memory, namely in SAS transport files. Rather than let SAS issue warnings and errors when it encounters attributes incompatible with the transport standard, we can use the COLUMNS table to identify problems. In the example, data set V8 has a number of features that are problematic: variable HAS GAP has an embedded blank; TOOLOONG, not surprisingly, exceeds the length limit; and LONGLBL has a label that exceeds the transport file maximum.

```sas
filename xDir pipe 'dir /b /on "c:\tmp\*.xpt" ';
data XPT_files;
   infile xDir;
   input;
   if index(_infile_, '.') then do;
      memname = scan(_infile_, 1, '.');
   end;
   output;
   end;''
run;
proc sql noprint;
   select "libname TRANS sasv5xpt 'c:\temp'"
     trim(memname)
     .".xpt";
quit;
```

PRINT procedure output of data set WHOOPS is shown below. Note that this use of the dictionary data is very basic. It’s not hard to think of extensions to the program. We could, for example, make it into a macro, passing it a one or two-level data set name.

Example 6 (Part 2)

```sas
   name label
has gap has gap
LongVarName clean $1 longlbl 3
   TooLong TooLong $1000
longlbl TooLong

   name         length
has gap           1
LongVarName       1
longlbl           3
TooLong        1000
```

Example 7: Generate Code from SQL

And then there are times when even the metadata isn’t enough. Such is the case with transport files, where data sets have to be referenced individually, rather than collectively (a LIBNAME must point to a specific dataset, rather than a directory). You could manually create the requisite LIBNAME and CONTENTS procedure statements to get what you need, maybe even making these into a macro that you could invoke for each data set. A non-obvious but effective approach is shown in Example 7. We identify the transport files programmatically, via a DOS command (DIR). We then process the command’s output, creating data set XPT_FILES, with variable MEMNAME. Once this is done, we can create the LIBNAME and other statements to process each data set using SQL. The SELECT statement concatenates literals containing SAS statements and variables containing data set names. The result is stored in macro variable CONTENTS.

Example 7 (Part 1)

```sas
filename xDir pipe 'dir /b /on "c:\tmp\*.xpt" ';
data XPT_files;
   infile xDir;
   input;
   if index(_infile_, '.') then do;
      memname = scan(_infile_, 1, '.');
   end;
   output;
   end;''
run;
proc sql noprint;
   select "libname TRANS sasv5xpt 'c:\temp'"
     trim(memname)
     .".xpt";
quit;
```
If data sets AE and DEMOG were in the directory, the following would be the value of macro variable CONTENTS.

**Example 7 (Part 2)**

```sql
libname TRANS sasv5xpt 'c:\temp\ae.xpt'; proc contents data=trans.ae;run; libname TRANS sasv5xpt 'c:\temp\dmg.xpt'; proc contents data=trans.dmg;run;
```

The Example doesn’t use dictionary tables, but is useful for demonstrating the power of the SELECT statement.

**Example 8: Print from Every Data Set in a Library**

Just as we went to some length in Example 7 to avoid manually writing statements for each data set in a library, here we also automatically generate code. This time, we use the TABLES table, identifying data sets in a library and printing up to 10 observations from each data set. *A priori* knowledge of the data set names is not needed. The macro gets this information from the metadata.

**Example 8**

```sql
libname inmeta 'J:\Client\Project\DATA\DERIVE';
%Macro Printit ;
  proc sql noprint;
  select memname into :dsns separated by ' ' from dictionary.tables where libname="INMETA";
  quit;
  %put &sqlobs. data sets in INMETA are: &dsns ;
  %if &sqlobs. = 0 %then %do;
  %put Nothing to do!;
  %goto bottom;
  %end;
  %let i =1 ;
  %do %until (%scan(&dsns,&i)= );
  %let dsn = %scan(&dsns,&i);
  proc print data= inmeta.&dsn(obs=10);
  title2 "First 10 records from &DSN";
  run;
  %let i = %eval(&i + 1);
  %end;
%Mend;
```

As in previous Examples, it’s interesting to consider improvements that we could make. An obvious tweak is controlling the number of observations to print. This could be done via a parameter defaulting to, say, 10 observations. A more significant improvement would be changes to the handling of empty (0-observation) data sets. Currently, if a data set is present but empty, nothing will be printed. We could count the observations in each data set (macro variable DSN) and print the data or write a message to the output file saying the data set was empty.

**Example 9: Enhance Example 8**

```sql
libname VACC 'S:\RHO\VASOCOR\VVS_Acc\data\crf' ;
%Macro Printit ;
  proc sql noprint;
  select quote(trim(id)) into :ids separated by ',' from vacc.riskmstr where CVDYN='Y';
  select memname into :dsns separated by ' ' from dictionary.tables where libname='VACC' ;
  quit;
  %put ids with CVD are: &ids ;
  %put data set in VACC are: &dsns ;
  %Macro printit ;
  %let i =1 ;
  %do %until (%scan(&dsns,&i)= ) ;
  %let dsn = %scan(&dsns,&i);
  proc print data= VACC.&dsn where id in(&ids) ;
  title2 "Records with CVD from &DSN.";
  run;
  %let i = %eval(&i + 1);
  %end;
%Mend;
```

The previous Example was helpful, since it quickly listed observations from a data set, giving a feel for its contents. A more practical, operational example follows. It creates two lists, one from the TABLES table, as before, and one from a data set with cardiovascular disease (CVD) indicators. Macro variable IDS is a comma-separated list of patients with a CVD history. This list is used as a filter for the PRINT procedure executed within the macro %DO loop. Notice that we code multiple SELECT statements in SQL; it is not necessary to invoke SQL separately.

**Example 10: Calculation As Part of an Expression**

We have already seen a use of the SELECT statement for creating groups of executable statements (Example 7). To emphasize the statement’s power, and its importance in reducing the amount of code that’s written and maintained, we present another example. Using the TABLES table and a knowledge of our system’s disk sector size, we compute the size of each member in a library. Variable SIZEMB is created within SQL, ready to use from data set SIZES. Since we create the variable straight from the table, no post-processing DATA step is required. You could, of course, process SASHELP.VTABLE in a DATA step to achieve the same result, but recall earlier comments about the typically less-efficient view processing.

**Example 10**

```sql
proc sql noprint;
create table sizes as
  select memname,
      (npage * 16384) / 1048576 as sizeMB /* divide by 1,024 for KB */
  from dictionary.tables
  where libname = 'WORK';
quit;
```

**Example 11: Count Observations in a Data Set**

A data set’s *n* is one of its most fundamental characteristics, and one that is often needed for decision-making in utility programs. This Example presents one of many possible approaches to identifying the count. The user passes the one or two-level name of the data set and, optionally, the name of the global macro variable that will hold the count. The variable is populated using the TABLES table, and is reset to -1 if the data set was not found. The COUNT variable, therefore, communicates whether the data set exists (value >= 0) and if it is populated (> 0).

**Example 11 (Part 1)**

```sql
%Macro countobs(datastor=, count=_count_);
```
The Example below uses COUNTOBS, and takes full advantage of the information contained in the macro variable. If it has a value greater than 0, we print the data set. If the value is 0, we write a message to the print file. Finally, if the value is -1, we write to the print file, saying it could not be located. Rather than just not print empty or missing data sets, we take the time to say why we didn’t print anything. The cost of extra macro language coding is far outweighed by the benefits of improved user understanding of the data.

Example 12 (Part 2)

%macro PrettyPrint(data=);%reset(action=start);options leftmargin=1in rightmargin=1in nodate nonumber dkricond=nowarn; ... other statements ...%reset(action=end);%mend;
APPENDIX A: OTHER TABLES AND VIEWS

**dictionary.catalogs**
sashelp.vcatalog

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
<td>Library name [upper case]</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name [upper case]</td>
</tr>
<tr>
<td>memtype</td>
<td>$8</td>
<td>Member type [CATALOG]</td>
</tr>
<tr>
<td>objname</td>
<td>$32</td>
<td>Object name [upper case]</td>
</tr>
<tr>
<td>objtype</td>
<td>$8</td>
<td>Object type [SCL</td>
</tr>
<tr>
<td>objdesc</td>
<td></td>
<td>Object description</td>
</tr>
<tr>
<td>created</td>
<td>num</td>
<td>Date created [DATETIME informat, format]</td>
</tr>
<tr>
<td>modified</td>
<td>num</td>
<td>Date modified [DATETIME informat, format]</td>
</tr>
<tr>
<td>alias</td>
<td>$8</td>
<td>Object alias [upper case]</td>
</tr>
<tr>
<td>level</td>
<td>num</td>
<td>Library concatenation level [0, 1, 2, …]</td>
</tr>
</tbody>
</table>

**dictionary.dictionaries**
sashelp.vdctnry

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name</td>
</tr>
<tr>
<td>memlabel</td>
<td>$256</td>
<td>Member label</td>
</tr>
<tr>
<td>name</td>
<td>$32</td>
<td>Column name</td>
</tr>
<tr>
<td>type</td>
<td>$4</td>
<td>Column type</td>
</tr>
<tr>
<td>length</td>
<td>num</td>
<td>Column length</td>
</tr>
<tr>
<td>npos</td>
<td>num</td>
<td>Column position</td>
</tr>
<tr>
<td>varnum</td>
<td>num</td>
<td>Column number in table</td>
</tr>
<tr>
<td>format</td>
<td>$49</td>
<td>Column format</td>
</tr>
<tr>
<td>informat</td>
<td>$49</td>
<td>Column informat</td>
</tr>
</tbody>
</table>

**dictionary.engines**
sashelp.vengine

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>engine</td>
<td>$8</td>
<td>Engine name [upper case]</td>
</tr>
<tr>
<td>alias</td>
<td>$8</td>
<td>Alias [upper case]</td>
</tr>
<tr>
<td>description</td>
<td>$40</td>
<td>Description</td>
</tr>
<tr>
<td>preferred</td>
<td>$3</td>
<td>Preferred? [yes</td>
</tr>
<tr>
<td>properties</td>
<td>$1024</td>
<td>Engine dialog properties</td>
</tr>
</tbody>
</table>

**dictionary.indexes**
sashelp.vindex

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
<td>Library name [upper case]</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name [upper case]</td>
</tr>
<tr>
<td>memtype</td>
<td>$8</td>
<td>Member type [DATA]</td>
</tr>
<tr>
<td>name</td>
<td>$32</td>
<td>Column name [case varies]</td>
</tr>
<tr>
<td>idxusage</td>
<td>$32</td>
<td>Column index type [COMPOSITE</td>
</tr>
<tr>
<td>indxname</td>
<td>$32</td>
<td>Index name [upper case]</td>
</tr>
<tr>
<td>indxpos</td>
<td>num</td>
<td>[Position of column in concatenated key</td>
</tr>
<tr>
<td>nomiss</td>
<td>$3</td>
<td>NOMISS option [yes</td>
</tr>
<tr>
<td>unique</td>
<td>$3</td>
<td>UNIQUE option [yes</td>
</tr>
</tbody>
</table>

**dictionary.formats**
sashelp.vformat

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
<td>Library name [upper case] [source='C']</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name [upper case] [source='C']</td>
</tr>
<tr>
<td>path</td>
<td>$1024</td>
<td>Path name [case preserved] [source='U']</td>
</tr>
<tr>
<td>objname</td>
<td>$32</td>
<td>Object name [upper case] [e.g. GROUP, TYPE] [source='U']</td>
</tr>
<tr>
<td>fmtname</td>
<td>$32</td>
<td>Format name [upper case] [e.g., GROUP, TYPE] [no decimal point]</td>
</tr>
<tr>
<td>fmttype</td>
<td>$1</td>
<td>Format type [F (format) I (informat)]</td>
</tr>
<tr>
<td>source</td>
<td>$1</td>
<td>Format source [U (user) B (built-in) C (from a catalog, i.e., created by PROC FORMAT)]</td>
</tr>
<tr>
<td>minw</td>
<td>num</td>
<td>Minimum width</td>
</tr>
<tr>
<td>mind</td>
<td>num</td>
<td>Minimum decimal width</td>
</tr>
<tr>
<td>maxw</td>
<td>num</td>
<td>Maximum width</td>
</tr>
<tr>
<td>maxd</td>
<td>num</td>
<td>Maximum decimal width</td>
</tr>
<tr>
<td>defw</td>
<td>num</td>
<td>Default width</td>
</tr>
<tr>
<td>defd</td>
<td>num</td>
<td>Default decimal width</td>
</tr>
</tbody>
</table>

**dictionary.libnames**
sashelp.vlibnam

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
<td>Library name [upper case]</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name [upper case]</td>
</tr>
<tr>
<td>memtype</td>
<td>$8</td>
<td>Member type [VIEW]</td>
</tr>
<tr>
<td>engine</td>
<td>$8</td>
<td>Engine name [SASESQL</td>
</tr>
</tbody>
</table>

**dictionary.styles**
sashelp.vstyle

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
<td>Library name</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name</td>
</tr>
<tr>
<td>style</td>
<td>$32</td>
<td>Style name</td>
</tr>
<tr>
<td>crdate</td>
<td>num</td>
<td>Date created [DATETIME informat, format]</td>
</tr>
</tbody>
</table>

**dictionary.views**
sashelp.vview

<table>
<thead>
<tr>
<th>Field</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libname</td>
<td>$8</td>
<td>Library name [upper case]</td>
</tr>
<tr>
<td>memname</td>
<td>$32</td>
<td>Member name [upper case]</td>
</tr>
<tr>
<td>memtype</td>
<td>$8</td>
<td>Member type [VIEW]</td>
</tr>
<tr>
<td>engine</td>
<td>$8</td>
<td>Engine name [SASESQL</td>
</tr>
</tbody>
</table>