

Kirk's Korner



Quick & Simple Tips

Kirk Paul Lafler, Software Intelligence Corporation

Tip: *Performing Table Lookup to Convert Coded Data*

Have you ever needed to substitute a coded value in a SAS dataset with more descriptive information? Here's a tip that will help transform coded data into meaningful information quickly and easily.

The SAS[®] System provides an effective way of converting coded data into more descriptive information using a table lookup process. When data is stored in a dataset as coded information you may be faced with having to convert the coded data into descriptive information before further processing can continue. The reasons for coding data are many, but it is often to reduce data storage requirements, safeguard sensitive information, or validate the authenticity or existence of a coded value. Whatever the reason, SAS software provides a simple way of handling this type of situation.

To illustrate how data can be converted from a coded form to a more descriptive and expanded view, we'll explore an example often found while completing online surveys. A survey may ask you to select one value from a list of choices such as 1 - 5 where "1" represents "Strongly Agree", "2" represents "Agree", "3" represents "Unsure", "4" represents "Disagree", and "5" represents "Strongly Disagree". After making the selection from the list of choices, the selection is often stored as coded data as either a numeric or character value in a dataset. Since this code may be useful as a means of indicating a particular response to a given question, it may be void of meaning or unintelligible by itself. In these situations, you can use the FORMAT procedure to convert the coded data (1 - 5) so it can display more descriptive information.

Say, for example, that we have a SAS dataset called SURVEY that contains the coded data we want to convert as well as other survey information. A common method of approaching this situation is to use the FORMAT procedure with a VALUES statement to assign individual entries consisting of coded data with more descriptive information. The example illustrates that any value outside the valid range of values is automatically assigned a description of "Miscoded".

The PRINT procedure is then specified in our example (although you could use any output-producing procedure or DATA step) along with a FORMAT statement to generate the desired output containing the description for each coded data value. The FORMAT statement associates the variable CODE in the SURVEY dataset with the user-defined format \$CODES. By using PROC FORMAT in these situations, you can take advantage of storage efficiencies provided by coding data while being able to display more descriptive information quickly and easily.

```
PROC FORMAT;
  VALUE $CODES `1' = `Strongly Agrees'
              `2' = `Agrees'
              `3' = `Unsure'
              `4' = `Disagrees'
              `5' = `Strongly Disagrees'
              OTHER = `Miscoded';
RUN;

PROC PRINT DATA=SURVEY NOOBS N;
  FORMAT CODE $CODES.;
RUN;
```

Contact Information

If you would like more information or have any questions about this tip, please contact: Kirk Paul Lafler, Software Intelligence Corporation at KirkLafler@cs.com. Kirk has been working with the SAS System since 1979 and is a SAS Certified Professional®. His company provides custom SAS programming, application design and development, consulting services, and hands-on SAS training to clients around the world. Kirk is the author of four books including PROC SQL: Beyond the Basics Using SAS by SAS Institute, and more than two hundred peer-reviewed articles and papers that have appeared in professional journals and SAS User Group proceedings. Kirk can be reached at:

Kirk Paul Lafler
Software Intelligence Corporation
World Headquarters
P.O. Box 1390
Spring Valley, California 91979-1390
E-mail: KirkLafler@cs.com



SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.