DATA Step
How do I change the order of columns in a dataset?

Changing the order of columns in an output dataset requires the use of a compile-time-only statement appearing before the SET statement to order the columns in the Program Data Vector during compilation.

Unfortunately, not all compile-time-only statements will allow you to do this easily.

The dataset SASHELP.CLASS hass the variables: NAME. SEX, AGE, HEIGHT and WEIGHT in that order.

The best-known hack for ordering columns is the RETAIN statement: all variables read from a dataset are retained automatically anyway, so this is a no-harm no-foul approach.

If only some of the variables are listed on the RETAIN statement, their order will be as listed, the remaining variables will appear in their original order:

```
data ordersome ;
  retain age weight ;
  set sashelp.class ;
run ;
```

 gives columns: AGE, WEIGHT, NAME, SEX, HEIGHT

To re-order all columns they must all be listed. The FORMAT statement can be used in the same way as the RETAIN statement:

```
data orderall ;
  format age weight height sex name ;
  set sashelp.class ;
run ;
```

The FORMAT statement does not give a syntax error if a format is not specified. The ATTRIB and LENGTH statements however produce a syntax error if used in the same way:

```
data attributes ;
  attrib age weight height ;
  set sashelp.class ;
run ;

data lengths ;
  length age weight ;
  set sashelp.class ;
run ;
```

The KEEP statement (along with the DROP statement) is also a compile-time-only statement, but although it will control which columns appear in the output dataset it cannot be used to control their order. The code:

```
data keepsome ;
  keep weight age sex ;
  set sashelp.class ;
run ;
```

still leaves the columns in the order SEX, AGE, WEIGHT.
 Unique solution ID: #1004
 Author: Alan D Rudland
 Last update: 2017-02-23 11:49