A common task is to combine variables from two or more SAS datasets into one master dataset. In SAS, this operation is referred to as *merging datasets*.

In the example below:

- the SAS dataset *census* has information on the Household ID (variable: houseID) and Census District where the house is located (variable: censusDistrict).
- The SAS dataset *households* has information on the income (variable: income) and number of occupants (variable: occupants) for the households.

This program illustrates how to create a new SAS dataset named: *combined*, which has all the information for the household (i.e. houseID, censusDistrict, income, occupants) in one observation. In SAS terminology, the program merges the two SAS datasets *census* and *households* to form a new dataset that will be named: *combined*.

```sas
/* Merging datasets in SAS */
data census;
  input houseID censusDistrict ;
datalines;
56543 380
23914 265
76103 455
;
run;

data households ;
  input houseID income occupants ;
datalines;
23914 265 4
56543 380 1
76103 455 7
;
rung;
```

**Step 1:** Create the SAS dataset: *census*

**Step 2:** Create the SAS dataset: *households*
**SAS – MERGING SAS DATASETS**

```sas
proc sort data = census ;
   by houseID ;
run ;
proc sort data = households ;
   by houseID ;
run ;
data combined ;
   merge census households ;
   by houseID ;
run ;
proc print data = combined ;
title "Combined dataset" ;
run ;
```

**Step 3:**
Sort both SAS datasets by the variable that uniquely identifies the household (in this case, the variable is: `houseID`)

**Step 4:**
Merge the two datasets to create the new SAS dataset named: `combined`.

**Step 5:**
Print the combined dataset to check that the merge worked as expected.

**Tips:**
- Each SAS dataset to be joined must have a common variable that uniquely identifies the observation. (in this example, the variable is `houseID` – this variable exists in both the `census` and `households` SAS datasets – see Steps 1 and 2 above).
- All datasets must be sorted by the common variable with the PROC SORT procedure before they can be joined by the MERGE statement (see Step 3 above).
- It is helpful to print the combined dataset to verify that the merge worked as expected (see Step 5 above).
- Always examine the SAS log for any error messages or warnings.