

Writing Matches and Nonmatches to Separate Tables

The **pg2.np_2016** table contains monthly public use statistics from the National Park Service for parks by **ParkCode**. The **pg2.np_codelookup** table contains the full name for each **ParkCode** value. Create a table, **parkStats**, that contains all park codes found in the **np_2016** table. Create a second table, **parkOther**, that contains **ParkCode** values in the **np_codelookup** table, but **not** in the **np_2016** table.

1. Determine the name of the common column in the **pg2.np_codelookup** and **pg2.np_2016** tables.

The common column is **ParkCode**.

2. Create a new program.

- Ensure that the data in both tables is sorted by the matching column.
- Using a DATA step, merge the **pg2.np_codelookup** and **pg2.np_2016** tables to create two new tables:
 - The **work.parkStats** table should contain only **ParkCode** values that are in the **np_2016** table.
 - Keep only the **ParkCode**, **ParkName**, **Year**, **Month**, and **DayVisits** columns.
 - The **work.parkOther** table should contain all other rows.
 - Keep only the **ParkCode** and **ParkName** columns.
- Submit the program and examine the output data.

```
proc sort data=pg2.np_CodeLookup
          out=work.sortedCodes;
  by ParkCode;
run;

proc sort data=pg2.np_2016
          out=work.sorted_code_2016;
  by ParkCode;
run;

data work.parkStats(keep=ParkCode ParkName Year Month DayVisits)
  work.parkOther(keep=ParkCode ParkName);
  merge work.sorted_code_2016(in=inStats) work.sortedCodes;
  by ParkCode;
  if inStats=1 then output work.parkStats;
  else output work.parkOther;
run;
```

3. How many rows are in the **parkStats** table?

The **parkStats** table has 4416 rows.

4. How many rows are in the **parkOther** table?

The **parkOther** table has 347 rows.

