

## Practice: Generating an Accumulating Column within Groups The

**pg2.np\_yearlytraffic** table contains annual traffic counts at locations in national parks. Park names are grouped into park types.

1. Open **p202p04.sas** from the **practices** folder. Complete the PROC SORT step to sort the **pg2.np\_yearlytraffic** table by **ParkType** and **ParkName**.

```
proc sort data=pg2.np_yearlyTraffic
          out=sortedTraffic(keep=ParkType ParkName Location Count);
  by ParkType ParkName;
run;
```

2. Modify the DATA step as follows:

- Read the sorted table created in PROC SORT.
- Add a BY statement to group the data by **ParkType**.
- Create a column, **TypeCount**, that is the running total of **Count** within each value of **ParkType**.
- Format **TypeCount** so that values are displayed with commas.
- Keep only the **ParkType** and **TypeCount** columns.
- Submit the program and confirm that **TypeCount** is reset at the beginning of each **ParkType** group.

```
data TypeTraffic;
  set work.sortedTraffic;
  by ParkType;
  if first.ParkType=1 then TypeCount=0;
  TypeCount+Count;
  format typeCount comma12.;
  keep ParkType TypeCount;
run;
```

3. Examine the **TypeTraffic** table and answer the following questions:

- How many rows and columns are in the table?
  - What are the values of **ParkType** and **TypeCount** in row 85?
- 
- 478 rows, 2 columns
  - In row 85, **ParkType** is *National Park* and **TypeCount** is 377,759.

4. Modify the program to write only the last row for each value of **ParkType** to the output table. Submit the program and examine the output data.

```
data TypeTraffic;
  set work.sortedTraffic;
  by ParkType;
  if first.ParkType=1 then TypeCount=0;
```

```
TypeCount+Count;  
if last.ParkType=1;  
format typeCount comma12.;  
keep ParkType TypeCount;  
run;
```

5. Examine the **TypeTraffic** table and answer the following questions:

- How many rows and columns are in the table?
  - What are the values of **ParkType** and **TypeCount** in row 5?
- 
- Five rows, two columns
  - In row 5, **ParkType** is *National Seashore* and **TypeCount** is 6,622,359.