1. **Exploring Data with Procedures**

```
/*Parts A and B*/
/*list first 20 rows*/
proc print data=pg1.np_summary(obs=20);
    var Reg Type ParkName DayVisits TentCampers RVCampers;
run;

/*Part C*/
/*calculate summary statistics*/
proc means data=pg1.np_summary;
    var DayVisits TentCampers RVCampers;
run;

/*Part D*/
/*examine extreme values*/
proc univariate data=pg1.np_summary;
    var DayVisits TentCampers RVCampers;
run;

/*Part E*/
/*list unique values and frequency counts*/
proc freq data=pg1.np_summary;
    tables Reg Type;
run;
```

b. Do you observe any possible inconsistencies in the data?

Yes. The Type column has inconsistencies. Notice that national preserve locations have the code PRES and PRESERVE.

c. What is the minimum value for tent campers? Is that value unexpected?

The minimum value is zero. No, because it is possible that a park had zero tent campers.

d. Are there negative values for any of the columns?

No

e. Are there any lowercase codes? Are there any codes that occur only once in the table?

There are no lowercase codes. NC, NPRE, and RIVERWAYS occur once in the table.

2. **Using Procedures to Validate Data**

```
*Part A;
proc freq data=pg1.np_summary;
    tables Reg Type;
run;

*Part B;
proc univariate data=pg1.np_summary;
    var Acres;
run;
```

   a. What invalid values exist for **Reg**? **None**

   What invalid values exist for **Type**? *NPRE, PRESERVE, RIVERWAYS*

   c. What are the smallest and largest parks? **Observation 78 (African Burial Ground Monument, .35 acres) and observation 6 (Noatak National Preserve, 6,587,071.39 acres)**

3. **Generating Extreme Observations Output**

```
*Part A and B;
ods trace on;
proc univariate data=pg1.eu_occ;
    var camp;
run;
ods trace off;

*Part D and E;
ods select extremeobs;
proc univariate data=pg1.eu_occ nextrobs=10;
    var camp;
run;
```

End of Solutions