Practice

If you restarted your SAS session, open and submit the **libname.sas** program in the course files.

Level 1

1. Creating One-Way Frequency Reports

The **pg1.np_species** table provides a detailed species list for selected national parks. Use this table to analyze categories of reported species.

a. Create a new program. Write a PROC FREQ step to analyze rows from **pg1.np_species**.

1) Use the TABLES statement to generate a frequency table for **Category**.
2) Use the NOCUM options to suppress the cumulative columns.
3) Use the ORDER=FREQ option in the PROC FREQ statement to order the results by descending frequency.
4) Use **Categories of Reported Species** as the report title.
5) Run the program and review the results.

Categories of Reported Species

Category	Frequency	Percent
Vascular Plant	65221	54.69
Bird	14612	12.25
Insect	14349	12.03
Fungi	6203	5.20
Nonvascular Plant	4278	3.59
Fish	3956	3.32
Mammal	3867	3.24
Invertebrate	1566	1.31
Reptile	1343	1.13
Algae	976	0.82
Slug/Snail	787	0.66
Spider/Scorpion	776	0.65
Amphibian	743	0.62
Crab/Lobster/Shrimp	582	0.49

b. Modify the PROC FREQ step to make the following changes:

1) Include only rows where **Species_ID** starts with *EVER* and **Category** is ***not*** *Vascular Plant*.

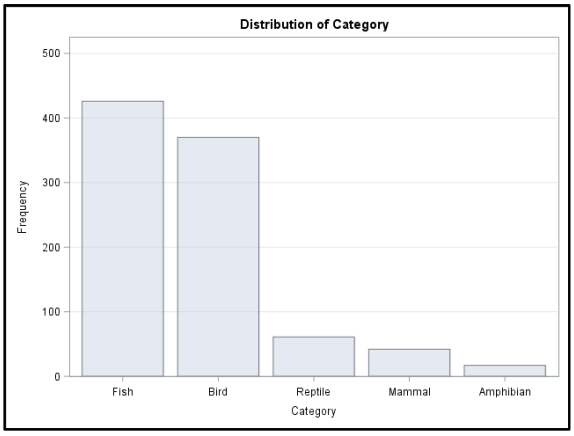
   Note: *EVER* represents Everglades National Park.

2) Turn on ODS Graphics before the PROC FREQ step and turn off the procedure title. Add the PLOTS=FREQPLOT option to display frequency plots.

3) Add **in the Everglades** as a second title. Run the program and review the results.

Categories of Reported Species in the Everglades

Category	Frequency	Percent
Fish	426	46.51
Bird	370	40.39
Reptile	61	6.66
Mammal	42	4.59
Amphibian	17	1.86



Level 2

2. Creating Two-Way Frequency Reports

The **pg1.np_codelookup** table is primarily used to look up a park name or park code. However, the table also includes columns for the park type and park region. Use this table to analyze the frequency of park types by the various regions.

a. Create a new program. Write a PROC FREQ step to analyze rows from **pg1.np_codelookup**. Generate a two-way frequency table for **Type** by **Region**. Exclude any park type that contains the word *Other*. The levels with the most rows should come first in the order. Suppress the display of column percentages. Use **Park Types by Region** as the report title.

b. Run the program and review the results. Identify the top three park types based on total frequency count.

Note: Statistics labels appear in the main table in Enterprise Guide if SAS Report is the output format.

Park Types by Region

Frequency
Percent
Row Pct

Table of Type by Region

Type	Region							
	Intermountain	Northeast	Southeast	Pacific West	Midwest	National Capital	Alaska	Total
National Historic Site	9 2.52 11.84	26 7.28 34.21	13 3.64 17.11	7 1.96 9.21	16 4.48 21.05	5 1.40 6.58	0 0.00 0.00	76 21.29
National Monument	34 9.52 49.28	8 2.24 11.59	8 2.24 11.59	9 2.52 13.04	8 2.24 11.59	0 0.00 0.00	2 0.56 2.90	69 19.33
National Park	18 5.04 30.51	2 0.56 3.39	7 1.96 11.86	17 4.76 28.81	7 1.96 11.86	0 0.00 0.00	8 2.24 13.56	59 16.53
National Historical Park	6	14	7	9	3	2	2	43

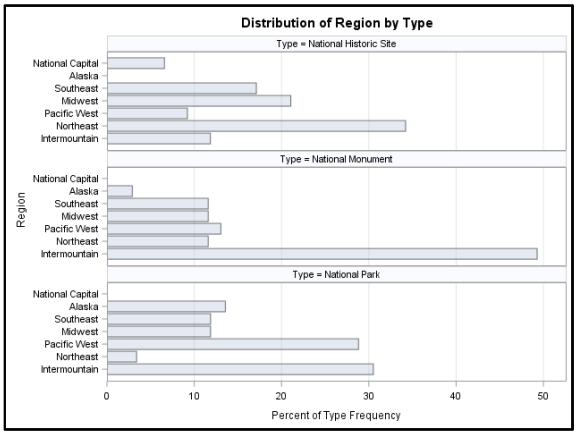
c. Modify the PROC FREQ step by limiting the park types to the three that were determined in the previous step. In addition to suppressing the display of column percentages, display the table using the CROSSLIST option. Add a frequency plot that groups the bars by the row variable, displays row percentages, and has a horizontal orientation. Use **Selected Park Types by Region** as the report title. Run the program and review the results.

Note: Use SAS documentation to learn how the GROUPBY=, SCALE=, and ORIENT= options can be used to control the appearance of the plot.

Selected Park Types by Region

Table of Type by Region

Type	Region	Frequency	Percent	Row Percent
National Historic Site	Intermountain	9	4.41	11.84
	Northeast	26	12.75	34.21
	Pacific West	7	3.43	9.21
	Midwest	16	7.84	21.05
	Southeast	13	6.37	17.11
	Alaska	0	0.00	0.00
	National Capital	5	2.45	6.58
	Total	76	37.25	100.00
National Monument	Intermountain	34	16.67	49.28
	Northeast	8	3.92	11.59

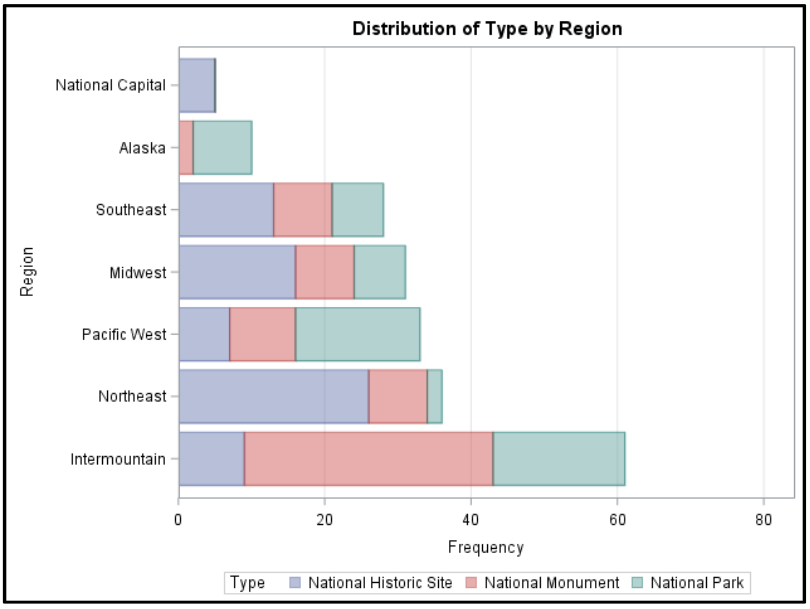


Challenge

3. Creating a Customized Graph of a Two-Way Frequency Table

The SGPLOT procedure can be used to create statistical graphics such as histograms and regression plots, in addition to simple graphics such as bar charts and line plots. Statements and options enable you to control the appearance of your graph and add additional features such as legends and reference lines.

a. Open **p105p03.sas** from the **practices** folder. Highlight the first TITLE statement and PROC FREQ step, run the selected code, and examine the generated plot. The program subsets the **pg1.np_codelookup** table for three park types: National Historic Site, National Monument, and National Park. The plot uses a stacked layout with a horizontal orientation.

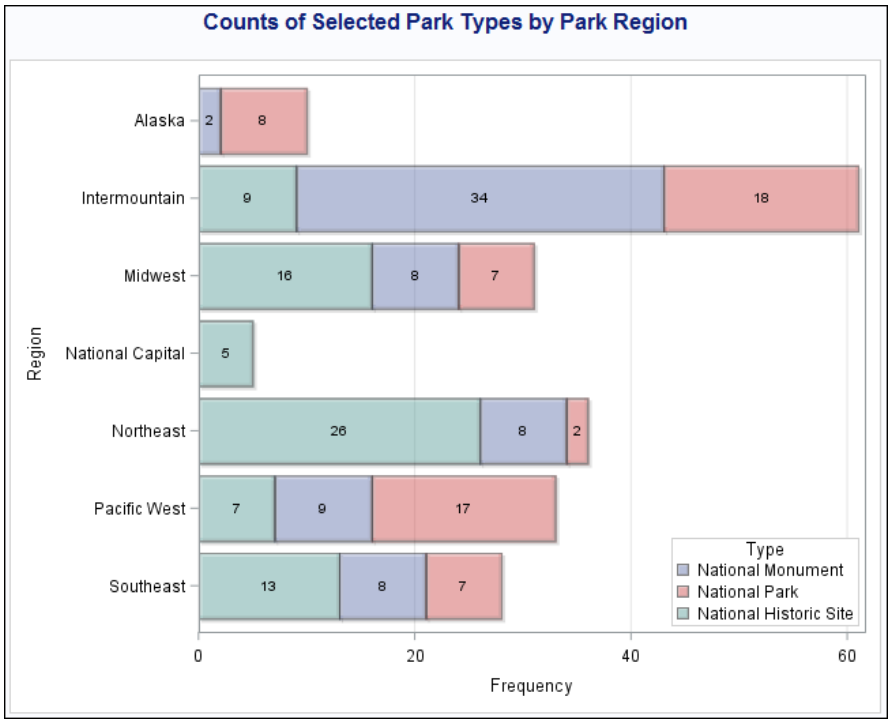


b. To create a more customized frequency bar chart, the SGPLOT procedure can be used with the **pg1.np_codelookup** table. Examine the PROC SGPLOT step in the demo program.

1) The HBAR statement creates a horizontal bar chart with separate bars for each **Region**. The GROUP= option segments each bar by the distinct values of **Type**.
2) The KEYLEGEND statement customizes the appearance and position of the legend.
3) The XAXIS statement adds reference lines on the horizontal axis.

c. Use SAS Help or autocomplete prompts to look for additional options in the HBAR statement to customize the appearance of the chart.

1) Display labels on each segment of the bars.
2) Change the fill attributes for each bar to make the color 50% transparent.
3) Apply different values for the DATASKIN option to change the color effect on the bars.



End of Practices