Practice

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

Level 1

4. Filtering Rows in a Listing Report Using Character Data

The **pg1.np_summary** table contains public use statistics from the National Park Service. The park type codes are inconsistent for national preserves. Examine these inconsistencies by producing a report that lists any national preserve.

a. Open **p103p04.sas** from the **practices** folder. Add a WHERE statement to print only the rows where **ParkName** includes *Preserve*.

Note: **ParkName** contains character values. These values are case sensitive.

b. Submit the program and view the results. Which codes are used for preserves?

Note: If you use double quotation marks in the WHERE statement, you receive a warning in the log. To eliminate the warning, use single quotation marks.

Obs	Type	ParkName
4	PRE	Yukon-Charley Rivers National Preserve
5	PRE	Bering Land Bridge National Preserve
6	PRESERVE	Noatak National Preserve
58	PRESERVE	Big Thicket National Preserve
74	PRE	Tallgrass Prairie National Preserve
113	PRESERVE	Mojave National Preserve
127	NPRE	Little River Canyon National Preserve
135	PRESERVE	Big Cypress National Preserve

5. Creating a Listing Report for Missing Data

Use PROC PRINT and the WHERE statement to examine the **pg1.eu_occ** table.

a. Create a new program. Write a PROC PRINT step to read the **pg1.eu_occ** table. Use a WHERE statement to list rows where **Hotel**, **ShortStay**, and **Camp** are missing. Run the program. How many rows are included?

b. Modify the WHERE statement to list rows with **Hotel** values greater than 40,000,000. Run the program. Which months are included in the report?

Obs	Geo	Country	YearMon	Hotel	ShortStay	Camp
1322	ES	Spain	2017M08	46720017	14976087	10627605
1323	ES	Spain	2017M07	43651610	12905564	8040503
1334	ES	Spain	2016M08	46502956	14449380	10425182
1335	ES	Spain	2016M07	42948946	12349138	7525715
1346	ES	Spain	2015M08	44813404	13742674	9708860
1358	ES	Spain	2014M08	43038904	12906393	9651606

Level 2

6. Using Macro Variables to Subset Data in Procedures

a. Create a new program. Write a PROC FREQ step to analyze rows from **pg1.np_species**. Include only rows where **Species_ID** starts with *YOSE* (Yosemite National Park) and **Category** equals *Mammal*. Generate frequency tables for **Abundance** and **Conservation_Status**.

b. Write a PROC PRINT step to list the same subset of rows from **pg1.np_species**. Include **Species_ID**, **Category**, **Scientific_Name**, and **Common_Names** in the report. Run the program.

Abundance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Abundant	1	1.19	1	1.19
Common	30	35.71	31	36.90
Occasional	2	2.38	33	39.29
Rare	16	19.05	49	58.33
Uncommon	34	40.48	83	98.81
Unknown	1	1.19	84	100.00
Frequency Missing = 4				

Conservation_Status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Endangered	2	10.53	2	10.53
Species of Concern	17	89.47	19	100.00
Frequency Missing = 69				

Obs	Species_ID	Category	Scientific_Name	Common_Names
115376	YOSE-1000	Mammal	Ovis canadensis sierrae	Sierra Nevada Bighorn Sheep
115377	YOSE-1001	Mammal	Cervus nannodes	None
115378	YOSE-1002	Mammal	Odocoileus hemionus	Mule Deer, Mule Deer
115379	YOSE-1003	Mammal	Sus scrofa	Pig, Pig (Feral), Wild Boar, Wild Boar
115380	YOSE-1004	Mammal	Canis latrans	Coyote

c. Create a macro variable named **ParkCode** to store *YOSE*, and another macro variable named **SpeciesCat** to store *Mammal*. Modify the code to reference the macro variables. Run the program and confirm that the same results are generated.

Note: The macro variable values are case sensitive when they are used in a WHERE statement.

d. Change the values of the macro variables to *ZION* (Zion National Park) and *Bird*. Run the program.

Abundance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Abundant	9	4.33	9	4.33
Common	83	39.90	92	44.23
Occasional	40	19.23	132	63.46
Rare	26	12.50	158	75.96
Uncommon	47	22.60	205	98.56
Unknown	3	1.44	208	100.00
Frequency Missing = 93				

Conservation Status				
Conservation_Status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Endangered	2	3.51	2	3.51
In Recovery	1	1.75	3	5.26
Species of Concern	52	91.23	55	96.49
Threatened	2	3.51	57	100.00
Frequency Missing = 244				

Obs	Species_ID	Category	Scientific_Name	Common_Names
117544	ZION-1080	Bird	Setophaga occidentalis	Hermit Warbler
117545	ZION-1081	Bird	Accipiter cooperii	Cooper's Hawk
117546	ZION-1082	Bird	Accipiter gentilis	Northern Goshawk
117547	ZION-1083	Bird	Accipiter striatus	Sharp-Shinned Hawk
117548	ZION-1084	Bird	Aquila chrysaetos	Golden Eagle

Challenge

7. Eliminating Case Sensitivity in WHERE Conditions

Character comparisons in a WHERE statement are case sensitive. Use SAS functions to make comparisons case insensitive.

a. Open **pg1.np_traffic**. Notice that the case of **Location** values is inconsistent.

b. Create a new program. Write a PROC PRINT step that lists **ParkName**, **Location**, and **Count**. Print rows where **Count** is not equal to 0 and **Location** includes *MAIN ENTRANCE*. Submit the program. Use the log to confirm that 38 rows are listed.

Note: If you use double quotation marks in the WHERE statement, you receive a warning in the log. To eliminate the warning, use single quotation marks.

c. The UPCASE function can be used to eliminate case sensitivity in character WHERE expressions. Use the UPCASE function on the **Location** column to include any case of *MAIN ENTRANCE*. Run the program and verify that 40 rows are listed.

UPCASE(*column*)

Note: The UPCASE function in a WHERE statement does not permanently convert the values of the column to uppercase.

Obs	ParkName	Location	Count
2	Abraham Lincoln Birthplace National Historical Park	TRAFFIC COUNT AT MAIN ENTRANCE	1,302
8	Allegheny Portage Railroad National Historic Site	Traffic Count at Main Entrance	784
22	Andersonville National Historic Site	TRAFFIC COUNT AT MAIN ENTRANCE	2,146
101	Booker T. Washington National Monument	TRAFFIC COUNT AT MAIN ENTRANCE	364
125	Cabrillo National Monument	TRAFFIC COUNT AT MAIN ENTRANCE	30,994
130	Canyon De Chelly National Monument	TRAFFIC COUNT AT MAIN ENTRANCE	151,500

End of Practices