Capstone Project

Biodiversity for the National Parks

Michael Lu

species_info.csv Data Frame

	category	scientific_name	common_names	conservation_status
0	Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole	nan
1	Mammal	Bos bison	American Bison, Bison	nan
2	Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Domesticated Cattle	nan
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	nan
4	Mammal	Cervus elaphus	Wapiti Or Elk	nan

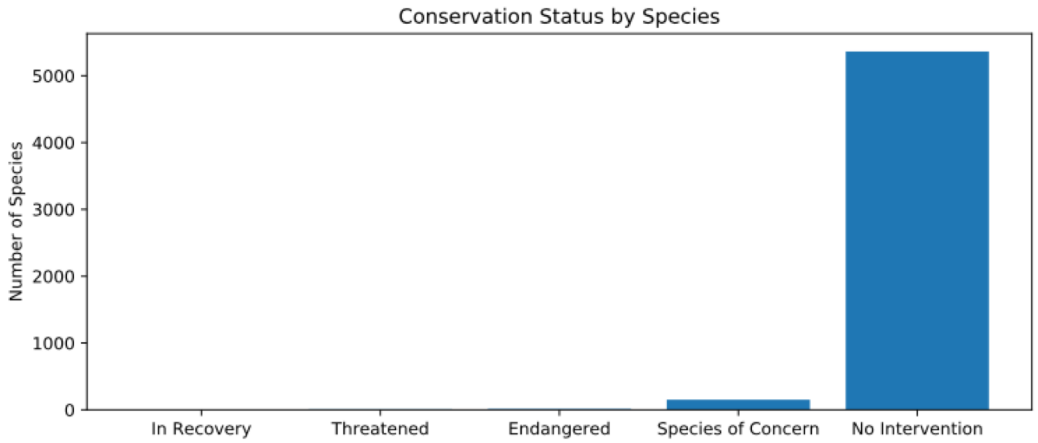
< Inspect the first five row of the data frame>

- 5824 rows(data) in the data frame
- 5541 unique species in the data frame
- 6 categories in the data frame
- 4 types of conservation status in the data frame

Conservation Status

- 4 types of conservation status in the original data frame
- Replace null data into “No Intervention”
- Most of the species’ conservation status are “No Intervention”

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	No Intervention	5363
3	Species of Concern	151
4	Threatened	10



Are certain types of species more likely to be endangered?

	Category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Dird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vescular Plant	4216	46	0.010793

Chi-Squared Test for Significance

Amphibian	Bird	Fish	Mamal	Nonvascular Plant	Reptile	Vascular Plant	
X	0.176	0.825	0.128	0.0018	0.781	1.04E-08	Amphibian
	X	0.077	0.688	1.05E-10	0.053	4.61E-79	Bird
		X	0.056	0.0005	0.741	1.49E-12	Fish
			X	1.48E-10	0.038	1.44E-55	Mamal
				X	0.034	0.662	Nonvascular Plant
					X	1.45E-04	Reptile
						X	Vascular Plant

<P-value Table>

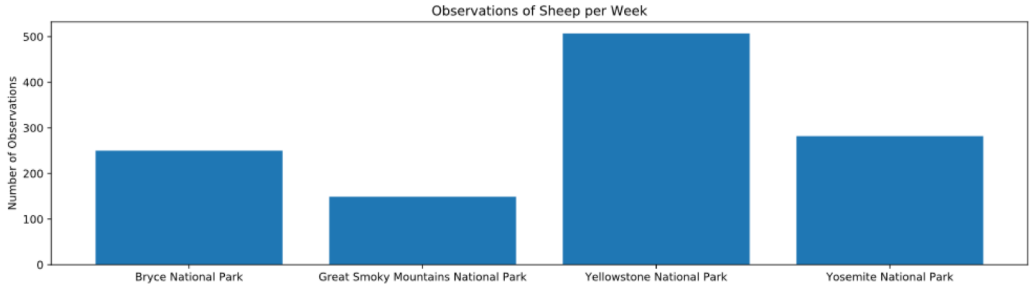
- If p-value > 0.05 No significant difference
- If p-value < 0.05 Significant difference

- Nonvascular Plant and Vascular Plant seem to be the species that need more concerned and protection.
- Although, there are several tests that results in a significant difference, it doesn’t seem to be much concerned compare to Nonvascular Plant and Vascular Plant.

Sheep Observation in each Park

- Number of sheep observed in each park over the past 7 days
- Yellowstone National Park observed the most numbers of sheep

	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282



Foot and Mouth Disease

- Baseline Conversion Rate: 15%
- Minimum Detectable Effect: 33.33%
- Statistical Significance: 90%
- Desired Sample size: 870
- Desired Observation Times: Bryce – Approx. 3.5 weeks
  Great Smoky – Approx. 6 weeks
  Yellowstone – Approx. 2 weeks
  Yosemite – Approx. 3 weeks