Biodiversity in our National Parks

Codecademy Capstone Project, Option 2

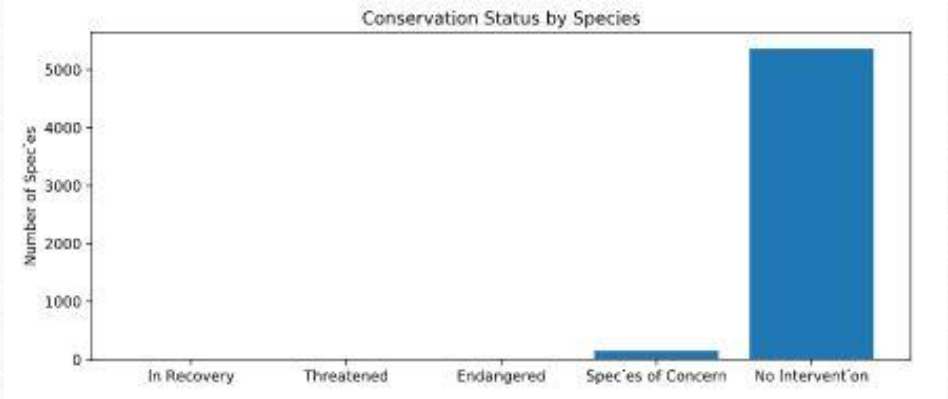
Karl Jarvis, pyRunner75544, “Kbet”

Objectives

- Analyze data on various species in our National Parks including conservation status, collected by field researchers.
- Analyze a second data set on species of wildlife spotted in the parks that were collected over just one week.
- Use the data to draw useful conclusions to aide our researchers in their work protecting our parks.

Observations on Species_info.csv

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



- Our initial look at the data reveals one pleasant bit of information; the vast majority of the species are not in fact in need of protection.
- Thankfully, of the 5541 species recorded only 3.4% are labeled under 'Species of Concern', 'Endangered', 'Threatened' or 'In Recovery.'

Portion of Each Category Protected

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	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	Vascular Plant	4216	46	0.010793

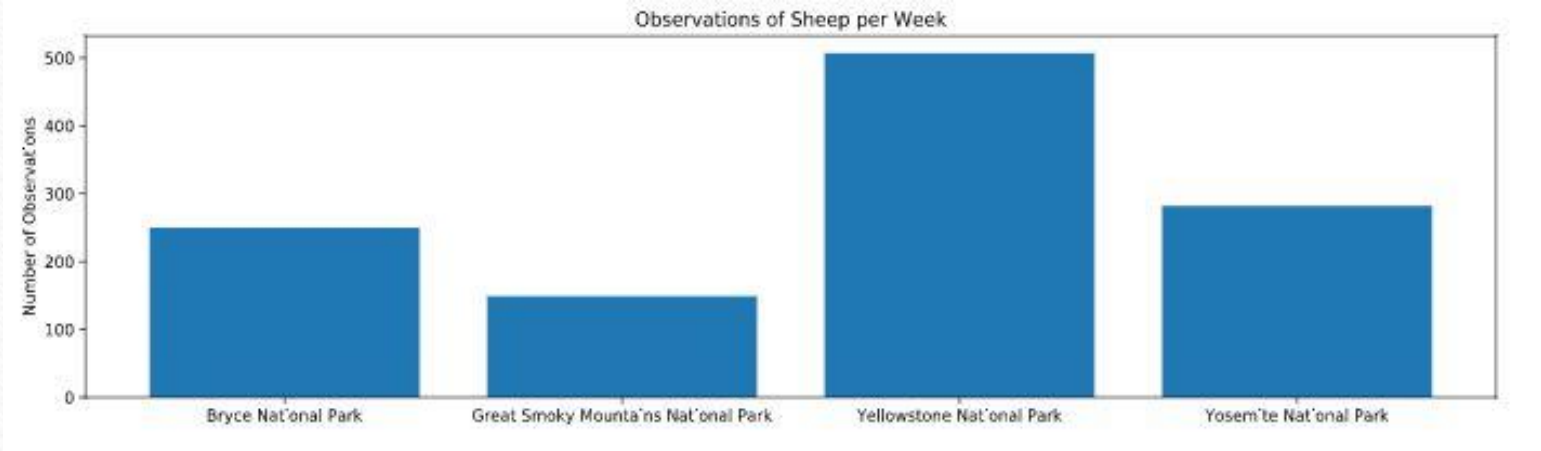
- Most notably, the highest percentage of protected species protected seems to be birds and mammals and it seems that mammals are more likely to be endangered than Birds.
- To verify this assertion we chose to run a chi squared significance test, as our data involves multiple categories.
- The test revealed a p-value of .6876, indicating that the difference is not significant and therefore our assertion has a high probability of being valid.
- A later test revealed that reptile and mammal data comparison was significantly different, with a p-value of 0.0383.
- Of these two groups we recommend that researchers focus on birds and mammals given the significance test outcomes.

Work with Observations.csv Data

	scientific_name	park_name	observations	category	common_names	conservation_status	is_protected	is_sheep
0	Ovis canadensis	Yellowstone National Park	219	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
1	Ovis canadensis	Bryce National Park	109	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
2	Ovis canadensis	Yosemite National Park	117	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
3	Ovis canadensis	Great Smoky Mountains National Park	48	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
4	Ovis canadensis sierrae	Yellowstone National Park	67	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
5	Ovis canadensis sierrae	Yosemite National Park	39	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
6	Ovis canadensis sierrae	Bryce National Park	22	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
7	Ovis canadensis sierrae	Great Smoky Mountains National Park	25	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
8	Ovis aries	Yosemite National Park	126	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
9	Ovis aries	Great Smoky Mountains National Park	76	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
10	Ovis aries	Bryce National Park	119	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
11	Ovis aries	Yellowstone National Park	221	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True

Ruminations on 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



- One group of researchers working with sheep asked us to focus on their interests in dealing with foot and mouth disease. We began with organizing the sheep by park.

Foot and Mouth Reduction

- We know that 15% of Bryce National Park sheep have the disease from earlier work.
- Researchers consider a minimum 5% reduction significant, and that level would mean that their efforts are working.
- Using our sample size calculator with our minimum detectable effect of 33.33% and the baseline of 15% we can determine that our sample size per variant should be 870.
- We suggest some guidance for researchers as to the time they'll need to spend observing sheep in Yellowstone and Bryce National Parks given the observations that were done during the one week time frame of the data we were given.
- Yellowstone will take about 1.7 weeks and Bryce will take about 3.5 weeks to complete a survey.

Thanks for your interest in our ongoing work with the parks!