

Biodiversity in National Parks



Sarah Hand
shand@Cambridge.org

The Conservation Status of Different Species

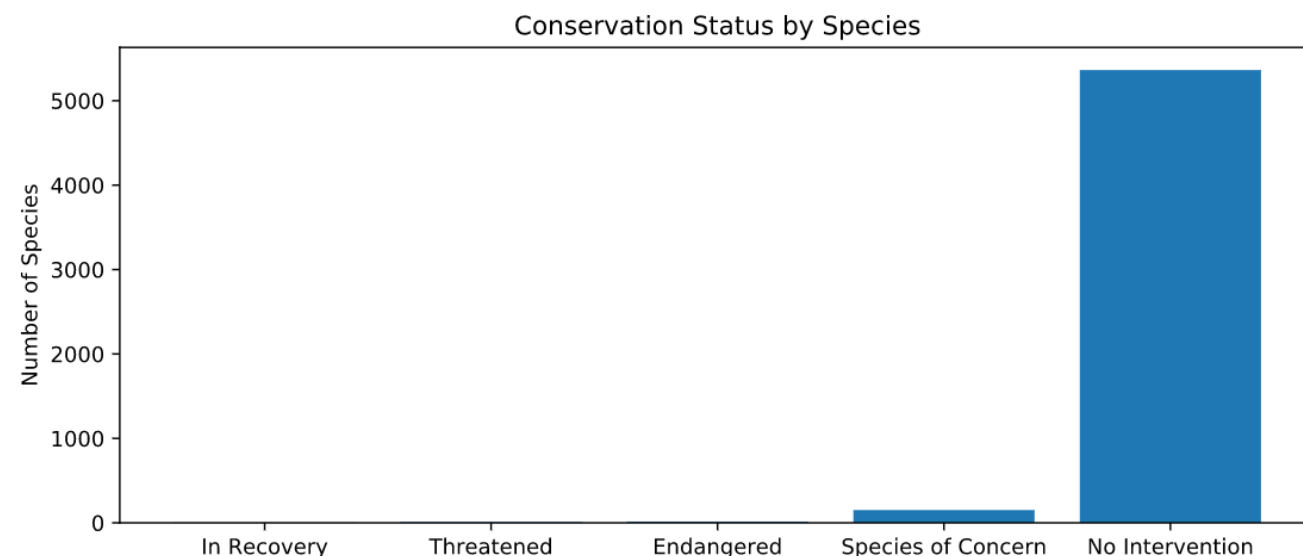
The study looked a total of **5543** different species across national parks.

These were filtered into conservation statuses:

- **Species of Concern**: declining population or appears to be in need of conservation.
- **Threatened**: vulnerable to endangerment in the near future.
- **Endangered**: seriously at risk of extinction.
- **In Recovery**: formerly Endangered, but currently not in danger of extinction throughout all or a significant portion of its inhabitable range.
- **No Intervention**: the population is steady and there is no current need for intervention.

When viewed as a table or bar chart we can clearly see that although No intervention is by far the largest category there are 180 species that currently require some level of 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	Percentage Protected
Amphibian	72	7	9%
Bird	413	75	15%
Fish	115	11	9%
Mammal	146	30	17%
Nonvascular Plant	328	5	2%
Reptile	73	5	6%
Vascular Plant	4216	46	1%

In order to answer the question the first thing I did was categorize by species and assign everything with the status no intervention as 'not protected', and everything else as 'protected'. Although no species had a need for protection higher than 20% it is interesting to see the difference by species.

Is this a significant difference?

To check that our results weren't just a result of chance I did a couple of Chi Square tests:

Percentage comparison of Mammals and Birds:

pvalue = ~ 0.688 = not significant / is a result of chance

Percentage comparison of Mammals and Reptiles:

pvalue = ~ 0.038 = significant/ is not a result of chance

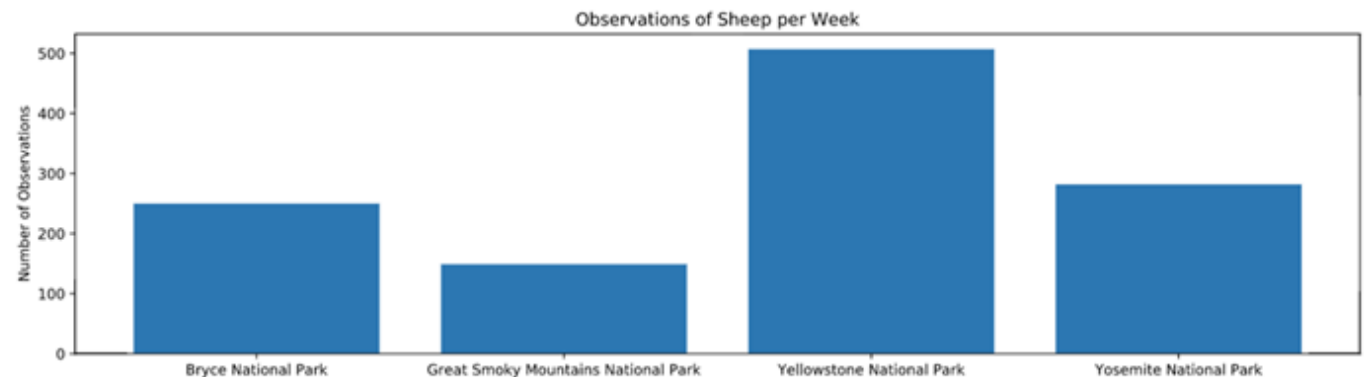
This confirms that certain types of species are more likely to be endangered than others!



In search of Sheep

Taking data from several National Parks across 7 days I filtered to only show the data for species 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 Reduction Effort - Sample Size Determination

I then calculated the sample sizes scientists would need to observe from each park to make sure their foot and mouth percentages are significant.*

Baseline = **15%** (*based on past investigations at Bryce*)

Minimum detectable effect = **33%** (*to detect reductions of at least 5 percentage points*)

Statistical significance = **90%** (*default level of significance*)

Sample size = **520***

520 sheep from each park will need observing!

*I used an [optimizely](#) sample size calculator.

How long would scientists have to observe sheep at each park?

Using the sample size of 520 and based on the amount of sheep seen at each park across 7 days we can determine that scientists would have to observe sheep across the following number of weeks:*

- Yellowstone National Park: **1 week**
- Bryce National Park: **2 weeks**
- Great Smoky Mountains National Park: **4 weeks**
- Yosemite national Park: **2 weeks**

*rounded to complete weeks