

Protecting BioDiversity in Our National Parks

CodeAcademy Capstone Project

Introduction to Data Analysis

Jefferson Underwood

August 10, 2018

Overview

The National Park System has a species inventory of 5541 species falling into 7 different categories: Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, Nonvascular Plant.

These species fall into five different conservation status designations:

- No Intervention needed
- Species of Concern: declining or appear 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 neither in danger of extinction throughout all or a significant portion of its range

Bottom Line Up Front

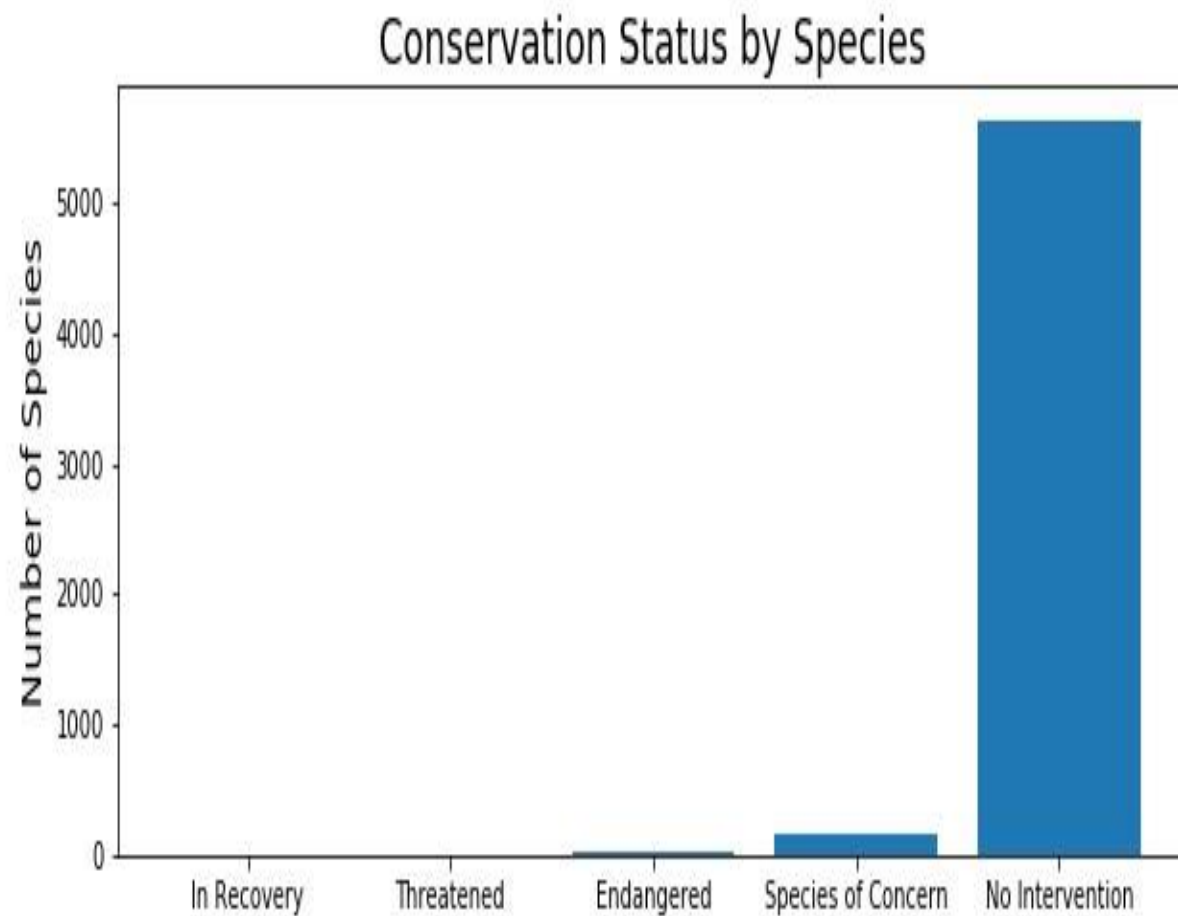
Descriptive analysis indicates three distinct groups of species of comparable endangerment. Further inferential and predictive analyses of additional data is needed to identify the shared factors defining these three groups.

Protection strategies across species and parks may be useful.

Most Species Don't Need Special Attention

Conservation Status	Scientific Name
No Intervention	5363
Species of Concern	151
Threatened	10
Endangered	15
In Recovery	4

Conservation Status By Species



Conservation Status Groups

Category	Not protected	Protected	Percent Protected
Mammal	146	30	17
Bird	413	75	15.4
Amphibian	72	7	8.9
Fish	115	11	8.7
Reptile	73	5	6.4
Nonvascular Plant	328	5	1.5
Vascular Plant	4216	46	1.1

The seven species appear to fall within three distinct groupings –
Mammal/Bird, Amphibian/Fish/Reptile, Nonvascular Plant/Vascular Plant

Confirmation of three main groupings

Chi-2 Analysis confirms three distinct groups of comparable endangerment:

- birds and mammals (Pval = 0.7)
- fish, reptiles, and amphibians (Pval = 0.8)
- non-vascular and vascular plants (Pval = 1)

Recommendation

NPS analysts should conduct inferential and predictive analyses of additional data to explore the shared factors defining these three groups of comparable endangerment and to develop candidate strategies benefiting multiple species.

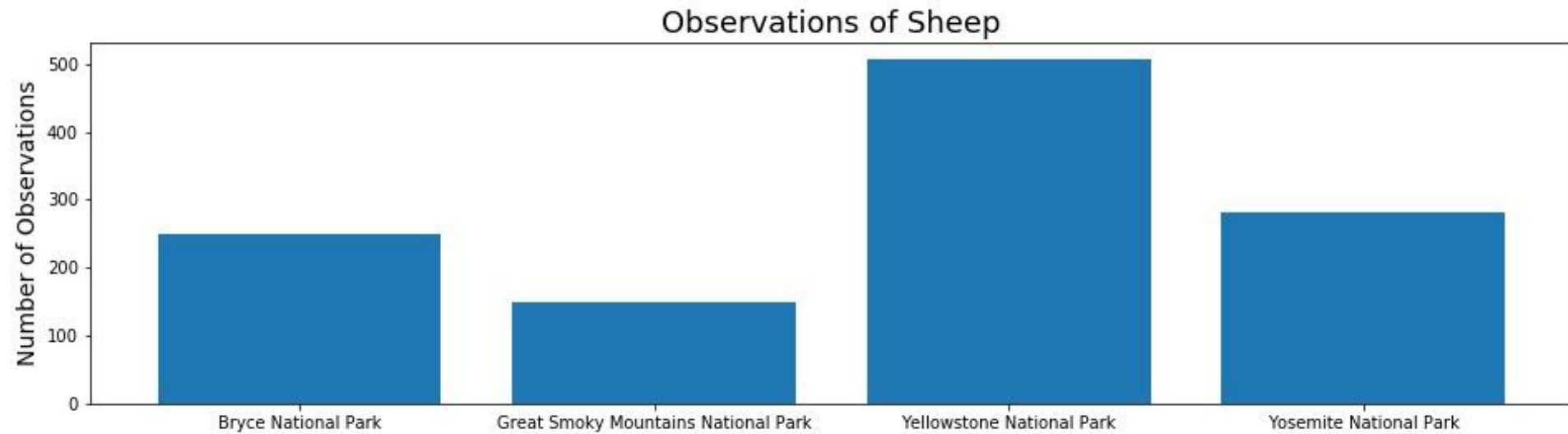
Protection Strategies May Span Multiple Parks

Reduction in foot and mouth (FAM) disease will remove one threat factor to sheep.

The Bryce Canyon Park has a 15% occurrence rate for FAM.

Yellowstone Park has been conducting a program to reduce FAM in that park to a 10% occurrence rate (a 30% reduction in the rate of infection).

Sheep Populations of Selected Parks



Verification Of Effectiveness

The CodeAcademy sample size calculator requires 870 samples to validate a 10% reduction to a 15% baseline at a 90% confidence rate

At the current observation rates this will require 2 weeks ($870/507 = 1.72$) at Yellowstone and 4 weeks ($870/250 = 3.48$) at the control site (Bryce)

Sample size calculators differ in their predictions. Using the same inputs (15% ,30%, and 90%) Optimizely advises 510 samples and Abtasty advises 808 samples. Both are less than the CodeAcademy advised sample size of 870.

Data collections have a one week process cycle, hence 2 weeks are needed at Yellowstone to collect data on the treated population and 4 weeks at Bryce to collect a sufficiently large number of samples, with additional weeks if the initial sample period doesn't include at 870 samples at each park (the largest number of samples indicated by the three sample size calculators.)

Conclusion

Descriptive analysis indicates three distinct groups of species of comparable endangerment. Further inferential and predictive analyses of additional data is needed to identify the shared factors defining these three groups.

Protection strategies across species and parks may be useful.