National Park Biodiversity Project

C.G. Parker

Species Information Dataset

Analysis of Conservation Data

Dataset Coverage

The National Park service keeps information on 5541 unique species.

Dataset Coverage

The species in the dataset can be divided along the following taxonomic categories:

- Mammals (e.g., American bison, Gapper's red-backed vole)
- Birds (e.g., red-tailed hawk, Northern harrier)
- Reptiles (e.g., corn snake, Eastern slender glass lizard)
- Amphibians (e.g., American bullfrog, Tellico salamander)
- Fishes (e.g., bloodfin darter, mountain brook lamprey)
- Vascular plants (e.g., antelope bitterbrush, Kaweah River scorpion-weed)
- Nonvascular plants (e.g., forsstroemia moss, hyophila moss)

The slender glass lizard, one of many organisms in the dataset.

Conservation Categories

The species in the data set fall into several categories:

- Endangered species
- Species in recovery
- Species of concern
- Threatened species
- No intervention (no special conservation status)

Conservation Categories

Of the 5,543 species listed in the data set, 180 have a special conservation status.

Endangered	15
In Recovery	4
Species of Concern	151
Threatened	10
No Intervention	5363

Table 1: Species counts according to conservation status.

Conservation Categories

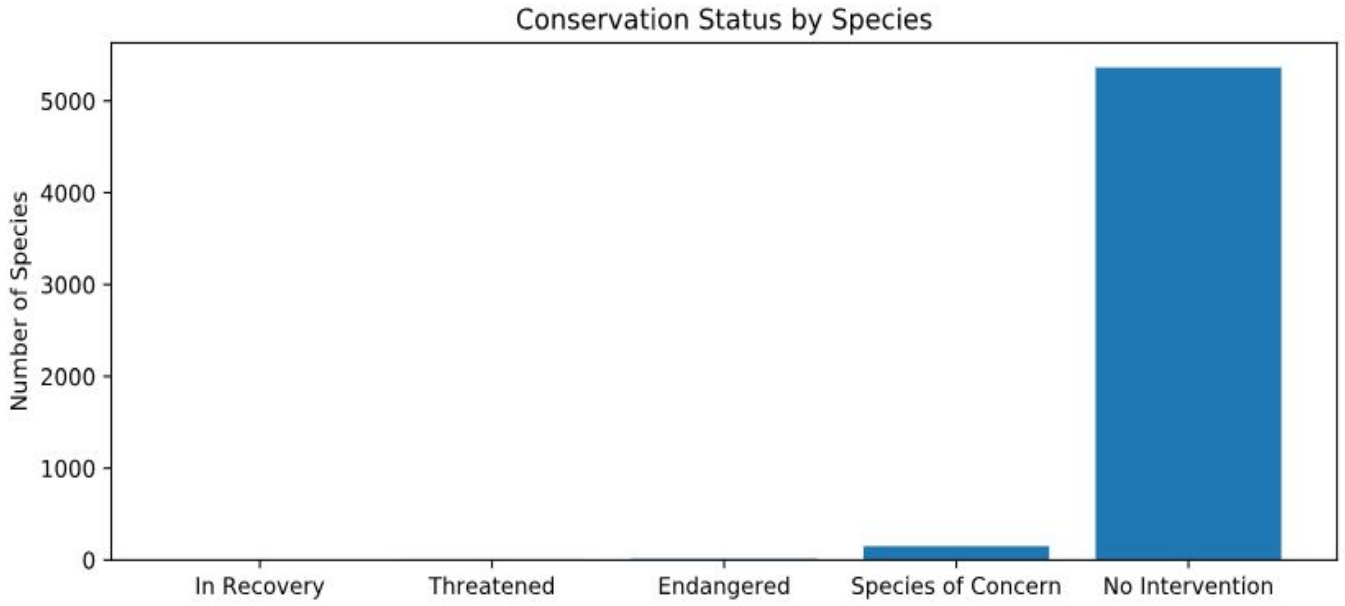


Figure 1: Species counts according to conservation status.

Conservation Categories

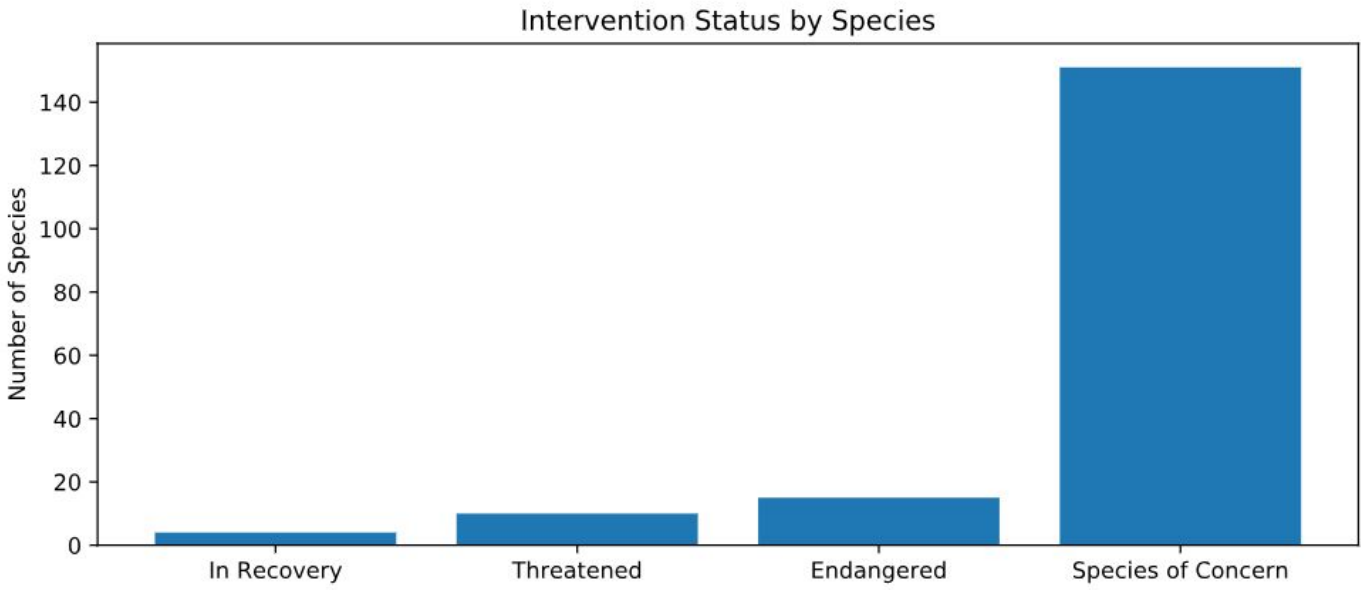


Figure 2: Species counts according to intervention status.

Conservation Categories

Important question: are certain species categories more likely to require intervention of some kind?

Conservation Categories

Category	Not Protected	Protected	% Protected
Amphibian	72	7	8.86%
Bird	413	75	15.37%
Fish	115	11	8.73%
Mammal	146	30	17.05%
Reptile	73	5	6.41%
Nonvascular Plant	328	5	1.50%
Vascular Plant	4216	46	1.08%

Table 2: Proportion of protected versus unprotected species, by category. Percentage values rounded.

- Mammals and birds have the highest proportion of protected species.
- Vascular and nonvascular plants have the lowest proportion of of protected species.

Conservation Statistics

To investigate whether the differences in proportion between these groups suggested an underlying difference, versus chance effects, we performed a number of chi-squared tests:

Test 1: birds vs mammals
Test 2: reptiles vs mammals

Conservation Statistics

To investigate whether the differences in proportion between these groups suggested an underlying difference or merely chance effects, we performed a number of chi-squared tests:

Test 1: birds vs mammals
Test 2: reptiles vs mammals

Birds vs mammals: $p = 0.6876$ (non-significant)
Reptiles vs mammals: $p = 0.0384$ (significant at $p < .05$ level) *

The reptiles vs mammals test rejected the null hypothesis of no difference between the categories.

That result provides evidence that mammals are more likely to be endangered than reptiles.

Conservation Statistics

In light of these results, we looked at three additional comparisons:

Test 3: vascular vs nonvascular plants
Test 4: amphibians vs non-amphibian animals
Test 5: reptiles vs birds

Vascular vs nonvascular plants: $p = 0.9092$ (non-significant)
Amphibian vs non-amphibian animals: $p = 0.2747$ (non-significant)
Reptiles vs birds: $p = 0.0531$ (non-significant)

None of these were significant, so no operations were performed to correct for multiple comparisons.

Conservation Statistics

- Given that mammals appear to be significantly more likely to require protection than other species, based on this data set we suggest focusing conservation efforts on mammalian species
- However, it is worth investigating whether mammals are in greater need of protection than other species, or if their increased intervention status is a byproduct of other factors
  - Easy to locate/track?
  - Attractive choices for securing research funding?

Species Observations Dataset

Analysis & Sample Size Determination

Sheep Sightings by Park

The second data set under analysis contained information about recent sightings for various species, as well as the locations where they were sighted.

We used this data to analyze sightings of several sheep species.

Sheep Sightings by Park

The dataset contained information on three species of sheep:

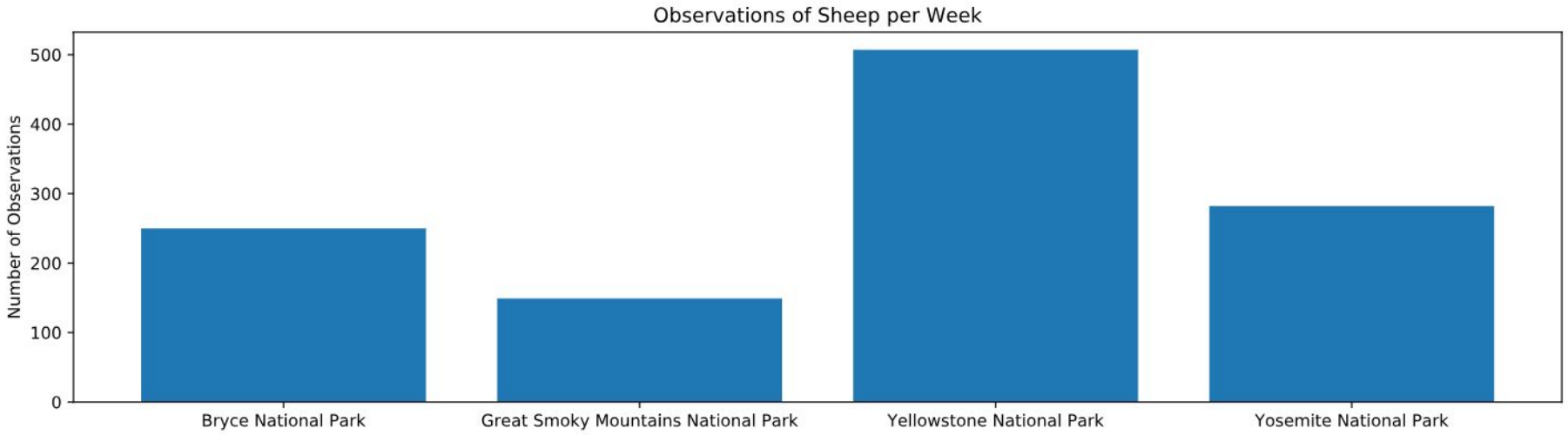
- *Ovis aries* (domestic sheep)
- *Ovis canadensis* (bighorn sheep)
- *Ovis canadensis sierrae* (Sierra Nevada bighorn sheep)

The sightings were recorded across four parks:

- Yellowstone National Park
- Yosemite National Park
- Bryce National Park
- Great Smoky Mountains National Park

Sheep Sightings by Park

Here is a graph showing the amount of sightings by park per week:



- The park with the most sightings was Yellowstone National Park.
- The park with the least sightings was Great Smoky Mountains National Park.

Study Plan: Reducing Foot and Mouth Disease

We want to know whether the program to reduce the rate of foot and mouth disease is working - specifically, if it has been reduced by $\geq 5\%$.

An analysis was performed to determine the necessary sample size to observe such an effect.

Study Plan: Reducing Foot and Mouth Disease

Based on information about last year's rate of infection from Bryce National Park, the baseline conversion rate was assumed to be 15%.

Since the scientists are interested in a reduction of at least 5%, the minimum detectable effect is equivalent to 100 x (5 / baseline) = 33.3%.

Study Plan: Reducing Foot and Mouth Disease

Given this baseline conversion rate the and minimum detectable effect of 33.3%, the desired statistical significance level (90%) means that the necessary sample size per variant would be $n = 870$.

Study Plan: Reducing Foot and Mouth Disease

With this sample size and the number of observed sheep per site, these figures can be used to determine how many weeks of observation would be necessary to get the required sample size.

- Yellowstone weeks of observation = 870 / 507 = 1.72 weeks
- Bryce weeks of observation = 870 / 250 = 3.48 weeks

Study Plan: Reducing Foot and Mouth Disease

Fortunately, this means the necessary data can be collected in a very short time period!