National Parks Biodiversity

By Kerry Driscoll

The Data

Working with two data sets

- species_info.csv
    - detailed classification information for each species
    - Category, Scientific Name, Common Names, Conservation Status

- observations.csv
    - the number of times each species was observed within a national park in a week
    - Scientific Name, National Park, Number of Observations

species_info.csv

- ## Category
  - 7 classification types:
  - Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, Nonvascular Plant

- ## Scientific Name
  - i.e. *Cervus elaphus*
  - 5,541 unique species

- ## Common Names
  - i.e. Wapiti or Elk

- ## Conservation Status
  - 5 classification types:
    - NOT PROTECTED: No Intervention
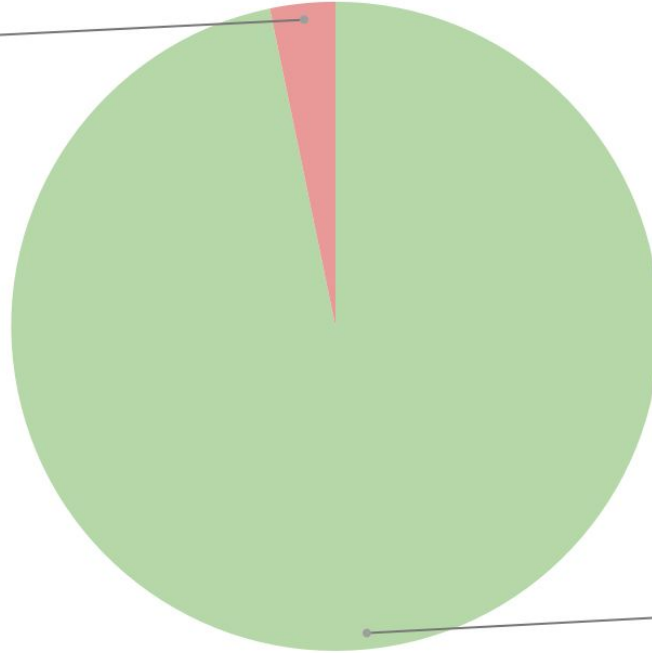   - PROTECTED: Species of Concern, Endangered, Threatened, In Recovery

Count by Conservation Status

Protected Classification	Category	# of Species	% of Species
Not Protected	No Intervention	5,363	96.79 %
Protected	Species of Concern	151	2.73 %
	Endangered	15	0.27 %
	Threatened	10	0.18 %
	In Recovery	4	0.07 %

Species by Conservation Status

Conservation Status

Protected
3.2%

Not Protected
96.8%

observations.csv

- **Scientific Name**
  - 5,541 unique species

- **National Park**
  - 4 National Parks:
  - Bryce, Great Smoky Mountains, Yellowstone, Yosemite
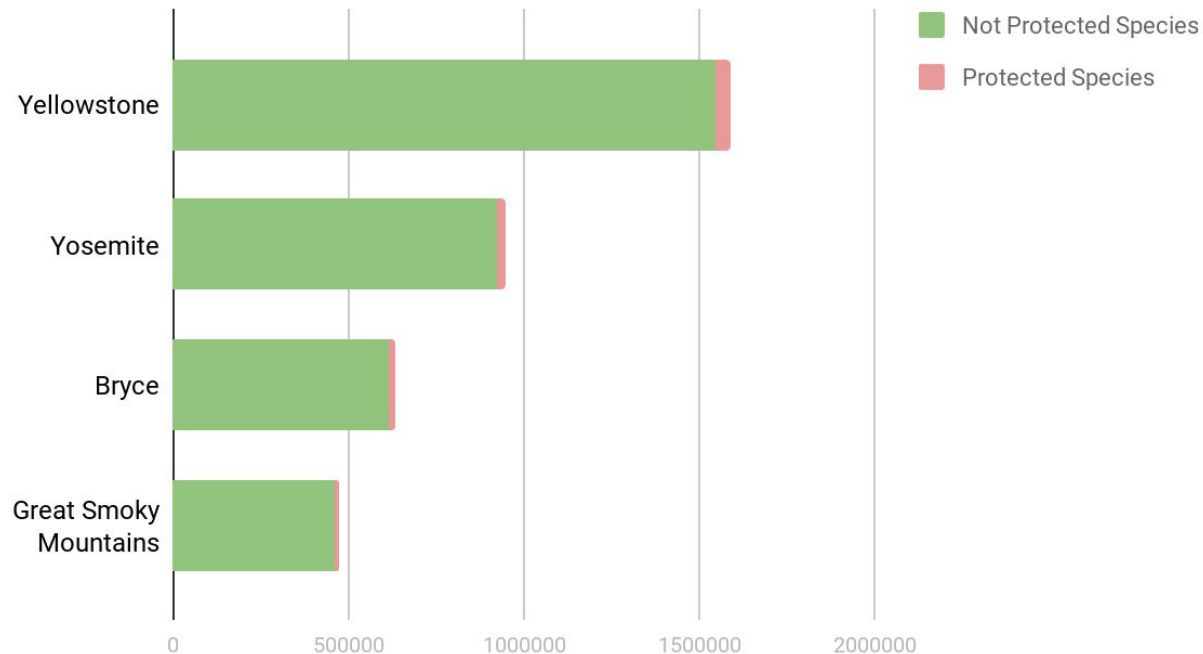
- **Number of Observations**
  - Min: 9
  - Median: 124
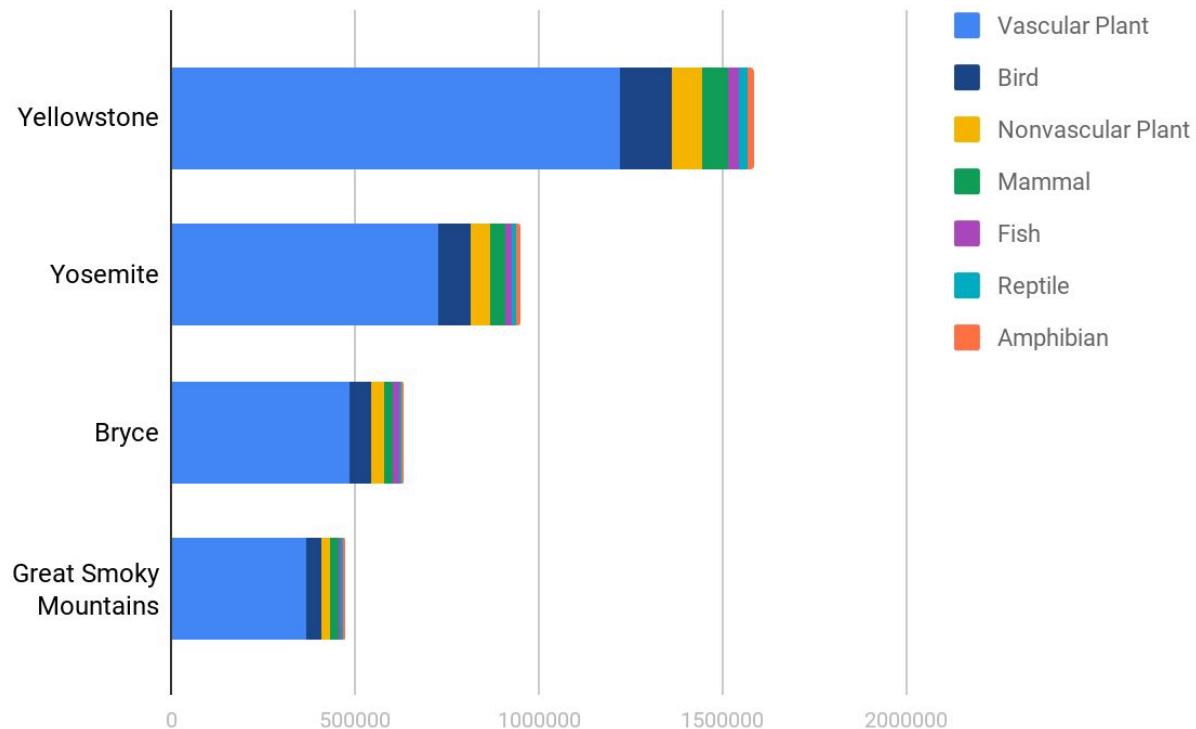 - Max: 321

Number of Observations by Park

National Park	# of Observations	% of Observations
Yellowstone	1,443,562	43.55%
Yosemite	863,332	26.05%
Bryce	576,025	17.38%
Great Smoky Mountains	431,820	13.03%
Total	3,314,739	100.0%

Number of Observations by Park

Observations of Species Category by Park

SIGNIFICANCE CALCULATIONS

- Trying to determine if a category of species is more likely to be "protected" than another

- Using Chi-Squared Test
  - Tests if the distribution of categories are statistically different
  - Determined by p-value
    - If p-value >.10: we **cannot** say the incidence of "protected" status is statistically different between the two categories
    - If p-value ≤.10: we can say with at least 90% confidence that the incidence of "protected" status is statistically different between the two categories
    - If p-value ≤.05: we can say with at least 95% confidence that the incidence of "protected" status is statistically different between the two categories

Protected Status by Species Category

Category	# of Species Not Protected	# of Species Protected	% of Species Protected
Mammal	146	30	17.05%
Bird	413	75	15.37%
Amphibian	72	7	8.86%
Fish	115	11	8.73%
Reptile	73	5	6.41%
Nonvascular Plant	328	5	1.50%
Vascular Plant	4216	46	1.09%

Statistically Different "Protected" Status Rates

P-Value	Mammal	Bird	Amphibian	Fish	Reptile	Nonvasc.
Bird	0.687					
Amphibian	0.128	0.176				
Fish	0.056	0.077	0.825			
Reptile	0.038	0.053	0.781	0.741		
Nonvascular Plant	0.000	0.000	0.002	0.000	0.034	
Vascular Plant	0.000	0.000	0.000	0.000	0.000	0.662

not statistically different	statistically different at < 10%	statistically different at < 5%

RECOMMENDATION

- Conservationists should place a **high priority on Mammals and Birds**. They are more likely to be in need of protection relative to Fish, Reptiles and Plants at a 10% significance level.

- Conservationists should place a **low priority on Plant Species** (both nonvascular and vascular). They are statistically significantly less likely to be in need of protection relative to all other categories.
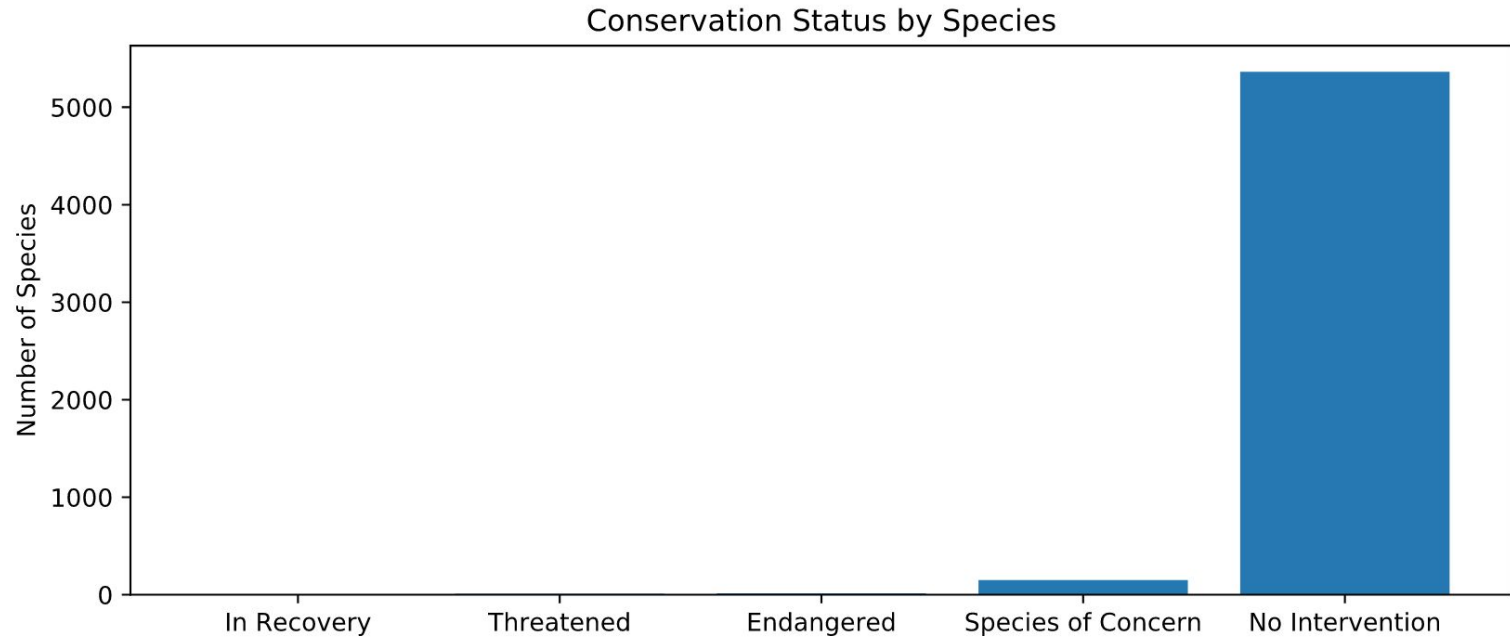
SHEEP FOOT AND MOUTH DISEASE STUDY

- Need to determine appropriate sample size to determine the effectiveness of the Park Rangers' program to reduce foot and mouth disease among sheep in Yellowstone National Park

- Assumptions:
  - Baseline: 15% of sheep in Bryce NP have foot and disease
  - Minimum detectable effect: 33.33% = (desired reduction ÷ baseline) = (5% ÷ 15%)
  - Statistical Significance: 90%

Result:
510 sheep

- Given that Yellowstone observes 507 sheep per week, it will take rangers **1 week** to capture this sample size

Figures

Conservation Status by Species

Figures