



National Park Services: Endangered Species



Are there species that the park should focus their conservation efforts?



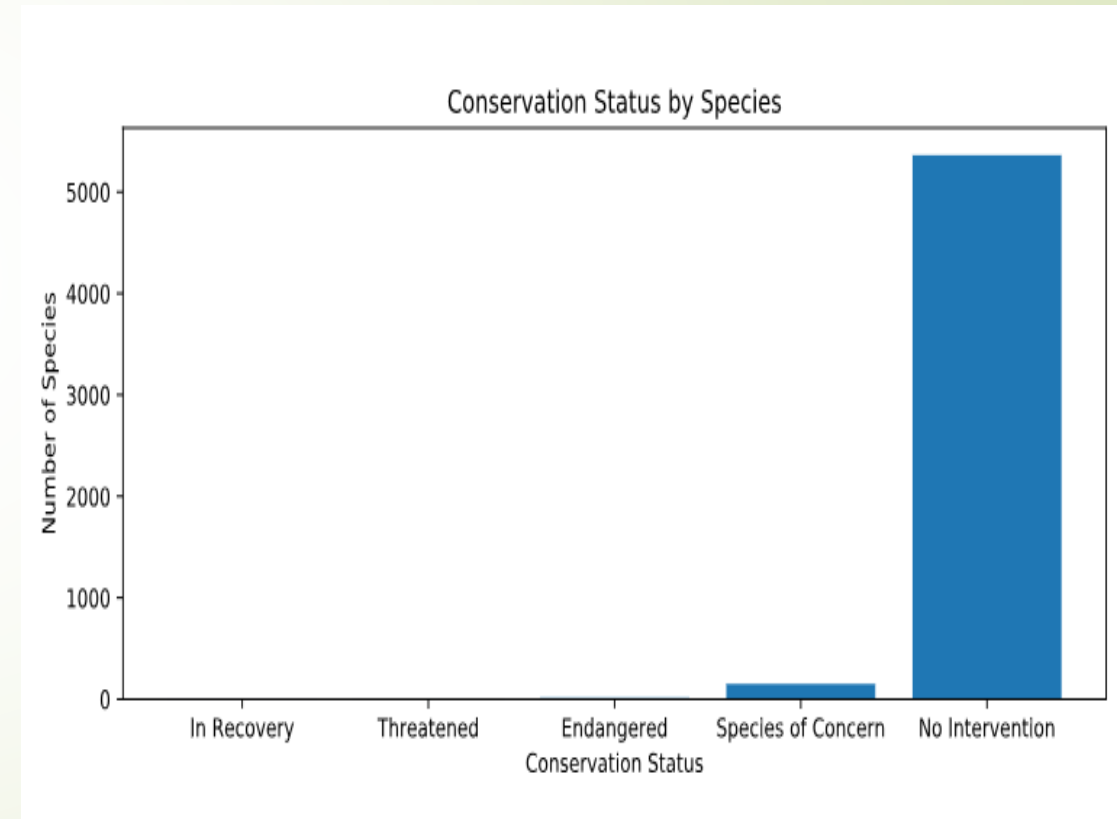
Data: Species Dataset Description



- Of the 5,824 species listed in the dataset, there are only 5,541 unique species (based on scientific name)
- There are 4 variables in the dataset: category, scientific name, common name, and conservation status
- The category variable is a general classification of the species, which includes: mammal, reptile, birds, fish, amphibian, vascular plant, and nonvascular plant
- The conservation status identify the level of threat to the species survival
 - The 4 tiers are: Species of Concern, Threaten, Endangered, In Recovery
 - The bulk (over 5600 species) are not threaten and therefore do not have a conservation status, which we recode as 'No Intervention'

Number of Species by Conservation Status



Conservation Status	Number of Species
Endangered	15
In Recovery	4
No Intervention	5,363
Species of Concern	151
Threatened	10





Proportion of species category that are endangered (Protected Conservation Status)

Category	Number Not Protected	Number Protected	Percent Protected
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015
Reptile	73	5	0.064103
Vascular Plant	4216	46	0.010793

- 
- 
- Mammal and Bird has the highest proportion of species that are endangered and, thus, considered protected under the conservation status
 - Other categories contain less than 10% of endangered species


Statistical Test: mammal vs. bird

	Protected	Not Protected
Mammal	30	146
Bird	75	413

Using a chi-squared test on the contingency table, the resulting p-value is 0.687595. This suggests that the probability of being protected for Mammal and Bird are from similar distribution. The proportion of mammal relative to the proportion of bird being protected is not significantly different from each other—the 2% difference is more likely due to chance.



Conservation of Endangered Species: Recommendation

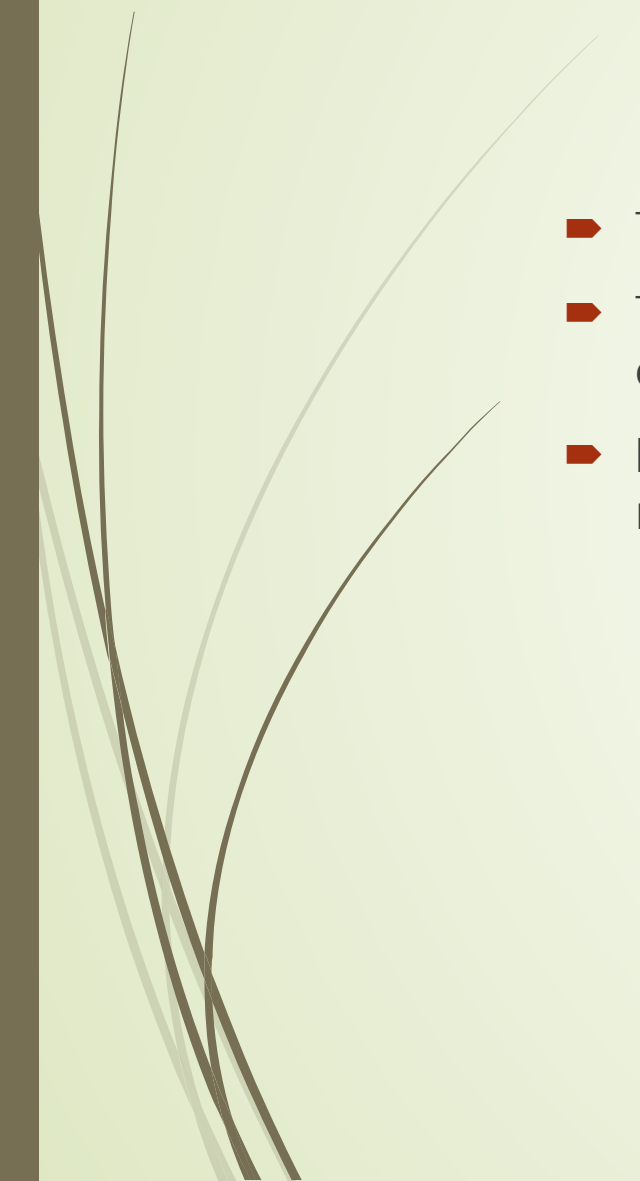
- While species of mammal and bird are more likely to be endangered, it is not statistically significant that mammal is more endangered than birds.
 - However, these two species are statistically more significantly more likely to be endangered than species from the other categories such as reptiles.
 - Therefore, conservation effort should be focused on mammals and birds.
- 



Foot and mouth disease in sheep
population: sample size



Data: Observation Dataset Description

- There are 23,296 observations in the dataset
 - There are three variables in the data: scientific name, park name, and observation
 - In order to consider only sheep observation in the dataset, it needs to be merged with the filtered species dataset.
- 

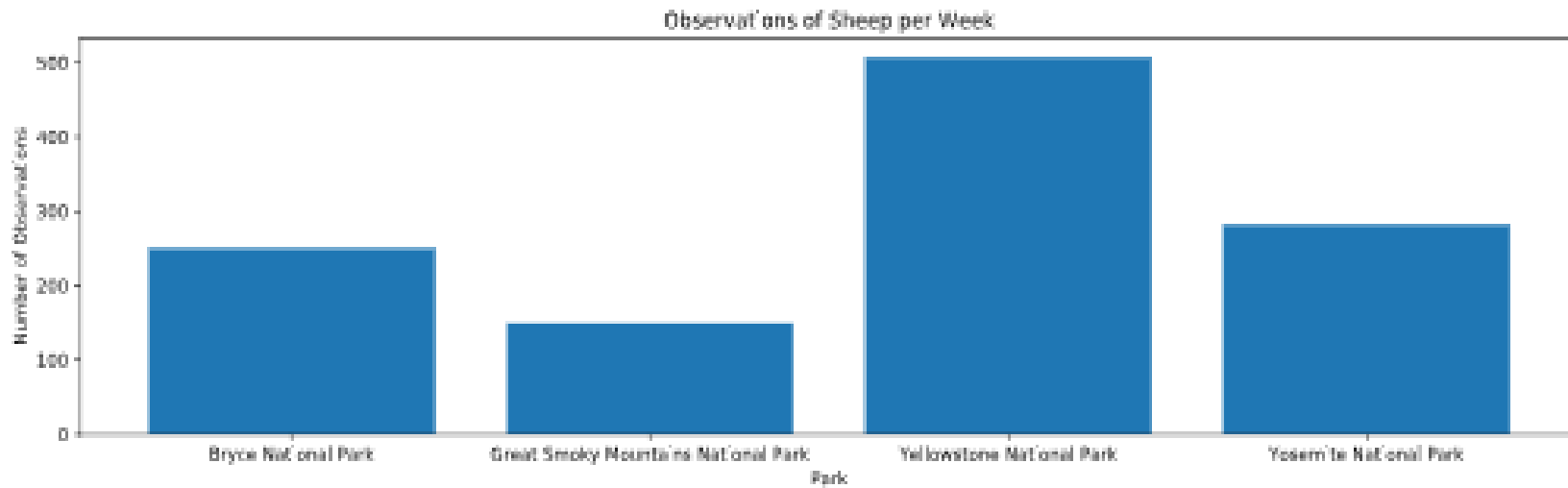


Data: Merged & Cleaned Dataset



Park	Number of Sheep Observed
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite National Park	282

Observations of Sheep by Park Location





Sample Size Calculation

- Requirements for sample size calculation:
 - Baseline conversion rate: the observed rate of foot and mouth disease in the population
 - Minimum detectable effect: the percentage of the reduction the park wants to see relative to the baseline
 - Statistical significance: the level at which the difference should be datable
- Sample size: 890 under the conditions
 - Baseline :15%
 - Minimum detectable effect: 33% ($100 * 5\% \text{ reduction} / 15\% \text{ baseline}$)
 - Statistical significance: 90%

Time to Collect Sample Size

Park	Week
Bryce National Park	3.56
Great Smoky Mountains National Park	5.17
Yellowstone National Park	1.75
Yosemite National Park	3.16

The table above shows the amount of time for each park to collect enough sample size to perform a statistical test at the 90% confidence level to determine if there has been a reduction of 5% in foot and mouth disease in the sheep population in the respective parks relative to the baseline population.