



Biodiversity for the National Parks

Capstone Project



Species Information

The species_info.csv file is an array of data with 4 columns:

1. Category
 2. Scientific_name
 3. Common_name
 4. Conservation_status
- Sample Categories: mammal, bird, reptile, fish, amphibian, vascular plant
 - Species count: 5,541 species observed



Conservation Status: Meanings

- 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



Protected Species by Category

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

% of Protected Species, by Category:

- Mammal: 17%
- Bird: 15%
- Amphibian: 8.8%
- Fish: 8.7%
- Reptile: 6.4%

Are certain types of species more likely to be endangered? Significance Tests

Perform two chi-squared significance tests:

1. Mammal vs. Bird
2. Reptile vs. Mammal

Results:

- Mammal & bird p-value: 0.6875
- Reptile & mammal p-value: **0.0383**

Conclusion: Based on the outcome of the second test, some species are more likely to be endangered than others. Results of the second chi2 contingency test indicate significance in the reptile & mammal test (p-value < 0.5)

Recommendation

Conservationists concerned about protecting species should be aware that mammals are more at risk of being endangered than some of the other species observed.



Foot and Mouth Sample Size Determination Effort

Explanation of study: Park rangers at Yellowstone conducted a study on their program to reduce the rate of foot and mouth disease to determine if the program was working

Givens: prior year, 15% of sheep had the disease at Bryce Canyon park

Using the default significance, 90%

Calculations: Baseline = 15%

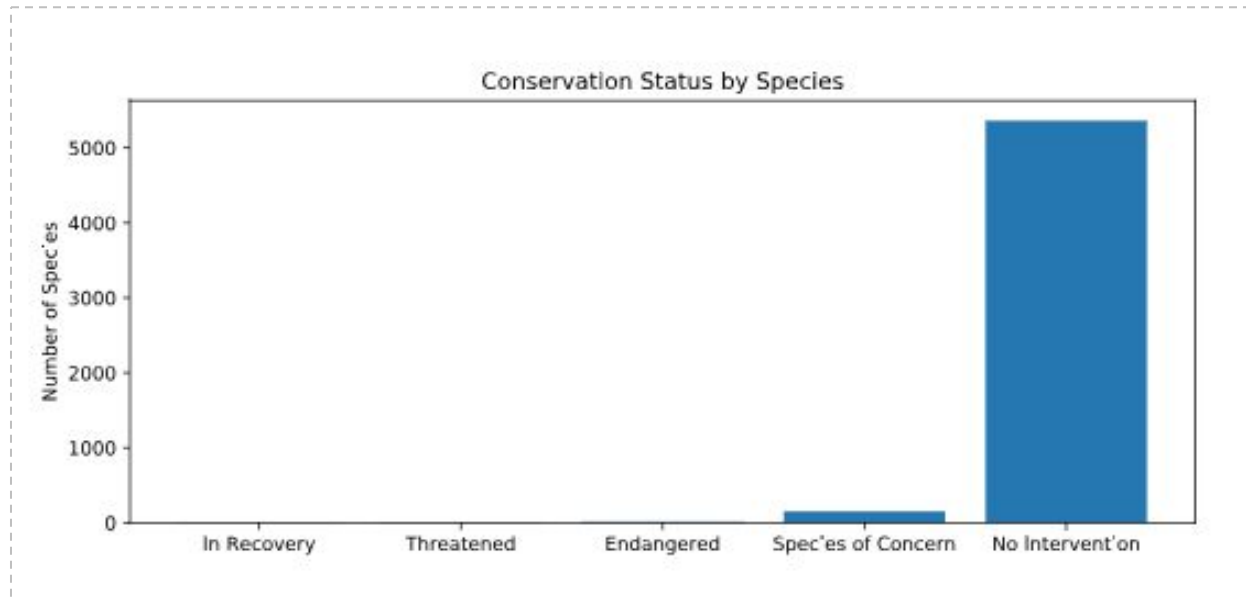
Minimum detectable effect = 33.33%

Sample size per variant = 870

Conclusion:

Scientists need to spend ~1 week in Yellowstone & 3 weeks in Bryce Canyon to observe enough sheep.

Graph 1: Conservation Status by Species





Graph 2: Observations of Sheep per Week

