Biodiversity for the National Parks

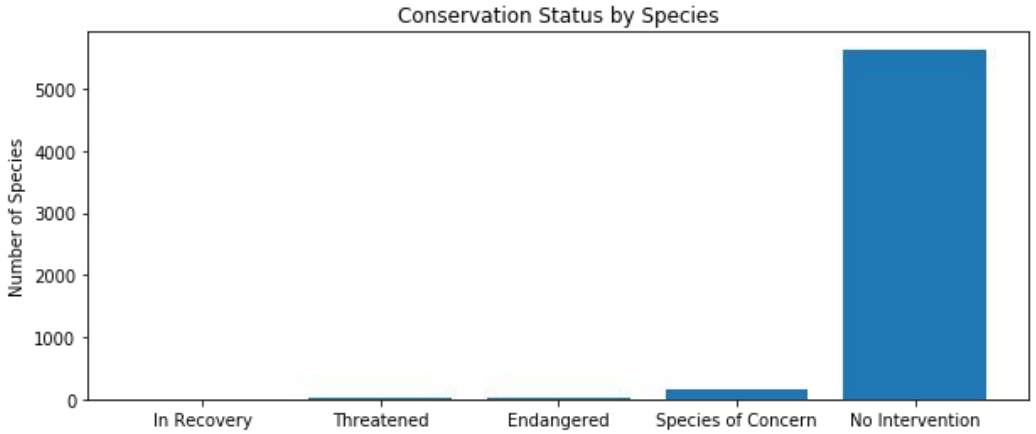
Project Capstone Option 2

Species Analyzed

Category	Not Protected	Protected	Percent Protected
Amphibian	72	7	8%
Bird	413	75	15%
Fish	115	11	8%
Mammal	146	30	17%
Nonvascular Plant	328	5	1%
Reptile	73	5	6%
Vascular Plant	4216	46	1%

More than 5K species observed across various National Parks:

- 151 Species of Concern, 15 Endangered, 10 Threatened, 4 In Recovery
- Mammals and Birds are more likely to be endangered



Significance Testing Across Categories

Because we wanted to compare categorical data sets, we needed to use the Chi Square Test to determine if there is a statistical difference between the data sets.

The null hypothesis is that there is no statistical difference between the data sets.

The results of the test show that we must reject the null hypothesis given p-values less than 0.05

Results:

Chi Square Test	P-value	Reject Null
Mammals & Birds	0.68	No
Mammals & Reptiles	0.03	Yes
Mammals & Vascular Plants	1.4 e-55	Yes

Recommendation

The significance testing revealed that there is no difference between the data sets for Mammals in comparison to Birds, Amphibians, and Fish. There is significant difference between the data sets for Mammals in comparison to Reptiles, Vascular Plants, and Nonvascular Plants.

Because Mammals have the highest percent of species protected at 17%, the National Parks should focus resources on saving Mammals. The National Parks can prioritize Mammals, Birds, Amphibians, and Fish based on the percent of species protected within the respective categories.

Foot and Mouth Disease in Sheep

Additionally, scientists have been observing sheep activity in Bryce National Park, Great Smoky Mountains National Park, Yellowstone National Park, and Yosemite National Park, in order to reduce Foot and Mouth Disease.

We understand that scientists know that 15% of sheep at Bryce National Park have foot and mouth disease. Park rangers at Yellowstone National Park have been running a program to reduce the rate of foot and mouth disease at that park. The scientists want to test whether or not this program is working. They want to be able to detect reductions of at least 5 percentage point.

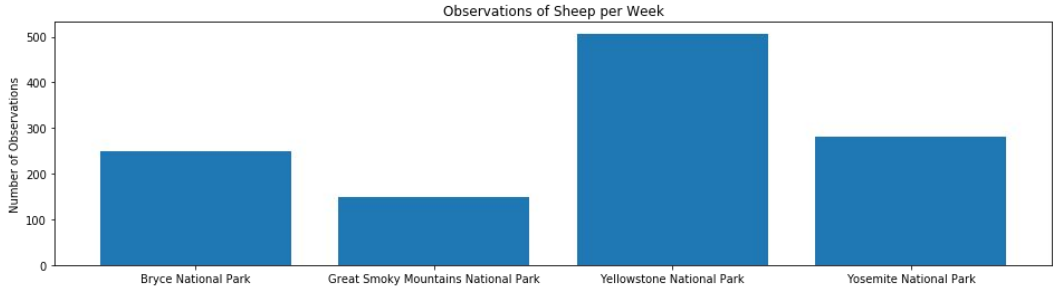
We've calculated the sample size the scientists would need to understand whether they have been reducing the rate of the disease.

To calculate sample size we used 3 inputs:

- Baseline Conversion rate = 15% of sheep at Bryce National Park
- Minimum Detectable Effect = 33.3% (calculated as 100 x 0.05 / 0.15 or 100 times the detectable reduction divided by control baseline conversion rate)
- Statistical significance = 90%

The Sample Size should be: 510 observations

Time to reach sample size



Based on the observations of Sheep per Week across the 4 National Parks:

National Park	# of weeks to reach sample
Great Smoky Mountain National Park	3.42
Bryce National Park	2.04
Yosemite National Park	1.80
Yellowstone National Park	1.00