

CAPSTONE OPTION 2: BIODIVERSITY FOR THE NATIONAL PARKS

CHRIS GREEN

SPECIES_INFO.CSV

Species.csv is a csv file which contains a table of data which has records of different species and details the category, scientific name, common names and conservation status of each.

Amongst other things we discovered:

- The scientific name of each species
- The common names of each species
- The species conservation status

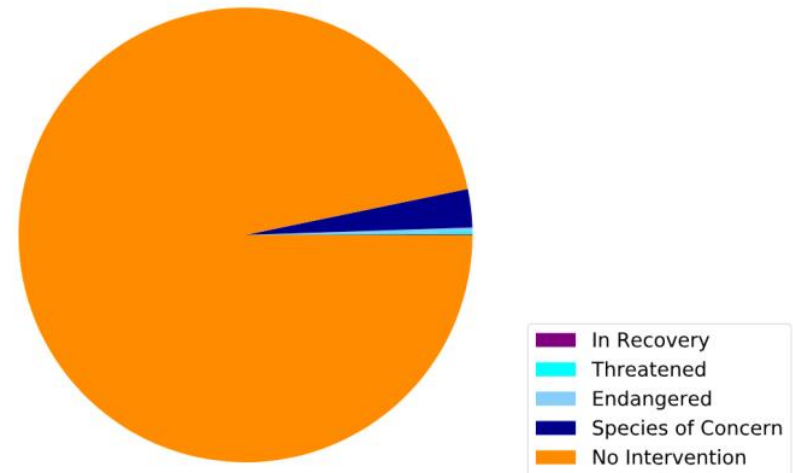
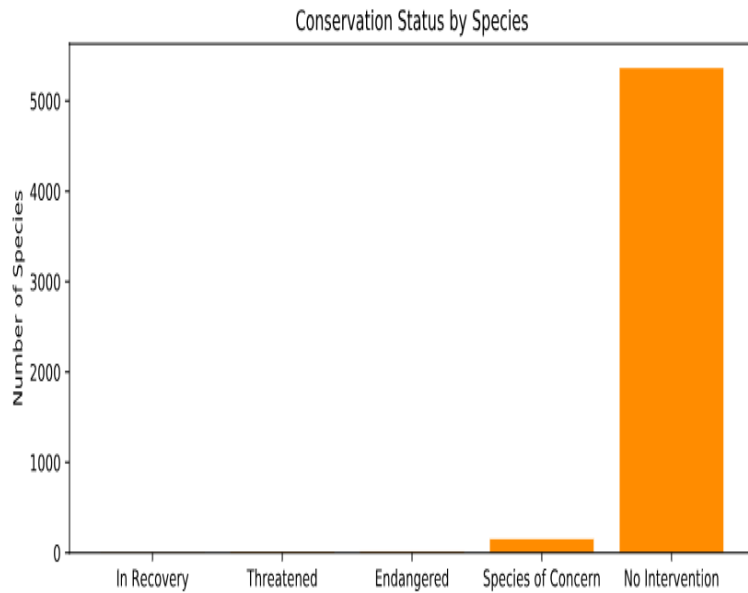
We learnt from the data that there were 5,541 unique species recorded and that there were 6 unique categories of species: Mammal, Bird, Reptile, Amphibian, Fish and Vascular Plant.

Each of these records fell into a different category of conservation status which helps to determine the endangerment of the species. These categories were 'No intervention', 'Species of concern', 'Endangered', 'Threatened' and 'In recovery'.



SPECIES_INFO.CSV

From the data we learned that the vast majority of species had 'no intervention' listed as the conservation status.



SIGNIFICANCE

At face value it was suggested that some categories of species are more likely than others to require intervention.

	category	not_protected	protected	percent_protected
3	Mammal	146	30	17%
1	Bird	413	75	15%
0	Amphibian	72	7	9%
2	Fish	115	11	9%
5	Reptile	73	5	6%
4	Nonvascular Plant	328	5	2%
6	Vascular Plant	4216	46	1%

However, further analysis was require to reach a conclusion

RECOMMENDATION

Having looked deeper into the data our recommendation to scientists concerned about endangered species would be that we should focus more on some categories of species than others. We reached this conclusion through a Chi-squared test for significance.

Chi squared was the most reliable test to use as we have more than one categorical dataset

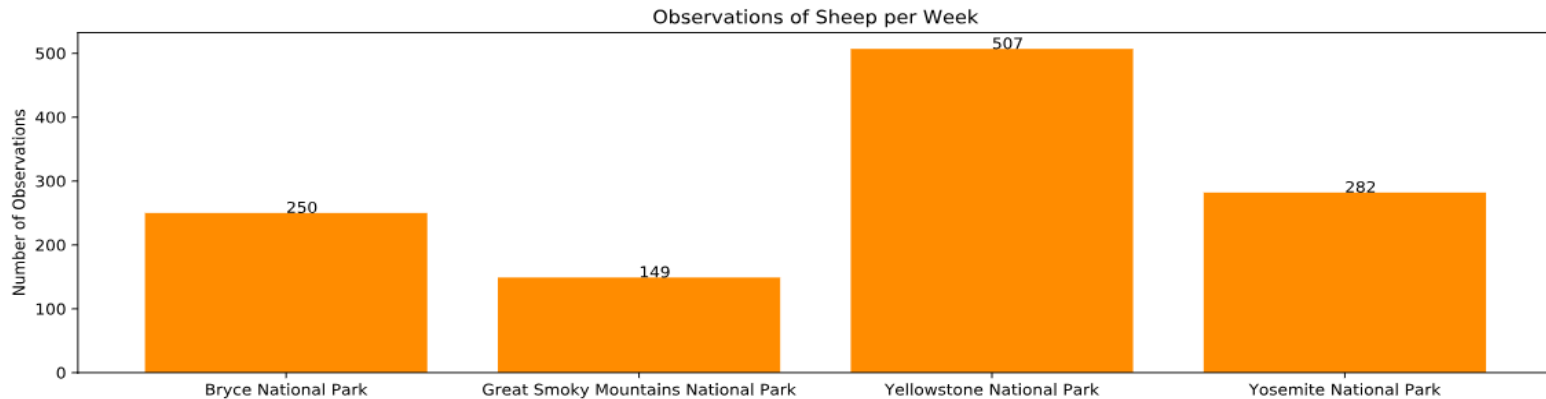
When running the test we found that there was no significant difference between the likelihood mammals being endangered over birds (p-value of 0.68), however, there was a significantly greater likelihood that Mammals would be endangered over reptiles (p-value of 0.04).

As such we can conclude that some categories of animals are more likely to be endangered than others.



SAMPLE SIZE DETERMINATION

With a baseline 15% occurrence of foot and mouth disease in sheep at Bryce National Park, we found that if the scientists wanted to be sure that a >5% drop in observed cases of foot and mouth disease in the sheep at Yellowstone was significant they would have to observe at least 870 sheep.



This would take the scientists approximately 1 week to complete their observation. If the scientists wanted to repeat their measurements at Bryce National Park, it would take 3.48 weeks.