Codecademy – Intro to data Analysis Capstone Project

Biodiversity of the National Parks

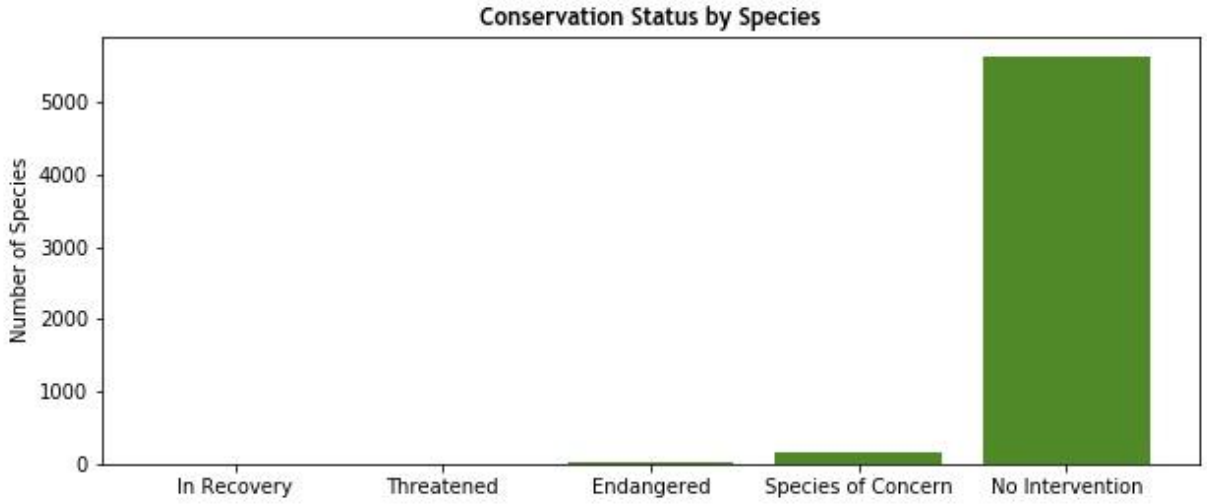
by

Alex Teixeira

The Data

- Table comprising all species identified in our National Parks
- 5824 entries to date with each entry containing the category, scientific name, common names and conservation status (if any) of all species identified
- Categories – Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant
- Conservation status – Species of concern, Endangered, Threatened, In recovery

Conservation Status



- As seen in the bar plot above, most species are not categorized as needing protection (5363) and were allocated a No Intervention status
- The data was analysed in order to understand differences in endangered status between different categories of species. These results are presented in the next slide

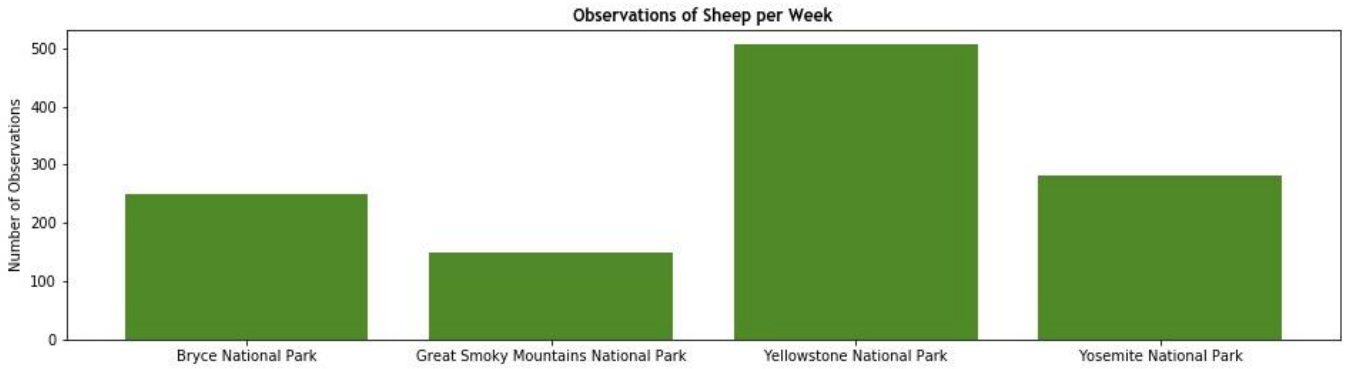
Conservation status

Difference between categories of species

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	8.86
1	Bird	413	75	15.37
2	Fish	115	11	8.73
3	Mammal	146	30	17.05
4	Nonvascular Plant	328	5	1.5
5	Reptile	73	5	6.41
6	Vascular Plant	4216	46	1.08

- The table above shows the protection status of each category of species
- A *chi squared test* was performed to asses if there's a significant difference between endangered status of some the categories of species (not protected vs. protected):
  - *There is NO significant difference between Mammals and Birds (pvalue = 0.69)*
  - *There is a significant difference between Mammals and Reptiles (pvalue = 0.04)*

Foot and mouth disease in Sheep – observations



- The number of observations of the different species across the different national parks was recorded in a different dataset
- This dataset was merged with the Species dataset in order to analyse the occurrence of Foot and Mouth disease in all Sheep species.
- In the graph above we plotted the number of observations of all sheep species across the different national parks

Foot and mouth disease in Sheep – recommendations

- In order to asses if the Program you are conducting in Yellowstone Park is significantly reducing the occurrence of foot and mouth disease in sheep you will need to make a number of observations from both your park and Bryce National Park where we obtained a baseline occurrence of 15%
- You want to detect a reduction of 5% (approx. 33% minimum detectable effect) with a 90% statistical significance, so you will need:

520 observations

- In each park. At your current observation pace this would take approx. 2 weeks in Bryce National Park and 1 week in Yellowstone National Park