Biodiversity Capstone Project: Investigating Protected Species

By: Eric Mendez

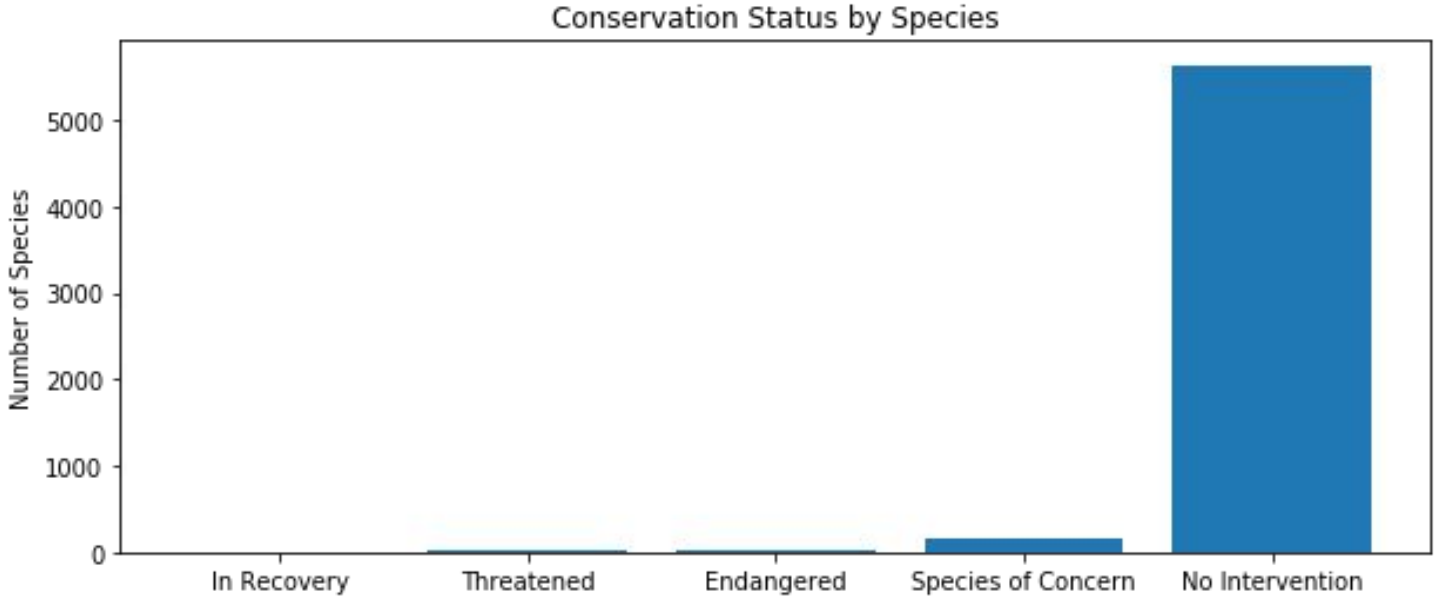
Inspecting the Data

What the data is composed of:

- 5541 species
- 7 types of species:
  - Mammal
  - Bird
  - Reptile
  - Amphibian
  - Fish
  - Vascular Plant
  - Nonvascular Plant
- Scientific name & common name
- Conservation status
  - Endangered
  - In Recovery
  - No Intervention
  - Species of Concern
  - Threatened

Inconsistencies in the data set:

- Over 200 duplicated species (entries)
- A majority of species unclassified with any conservation status



More than 95% of species are not classified with a conservation status

Analyzing the Data

Category	Not Protected	Protected	Percent Protected
Amphibian	72	7	9%
Bird	413	75	15%
Fish	115	11	9%
Mammal	146	30	17%
Nonvascular Plant	328	5	2%
Reptile	73	5	6%
Vascular Plant	4216	46	1%

Observations:

- All species are less than 20% protected
- Birds and mammals are the most protected
- Many vascular plants are not protected

Are certain types of species more likely to be endangered?

After running a chi-squared test to compare any two species to see if they are related, in regards to being protected, we found that:

- The slight difference in birds (15% protected) and mammals (17%) **is due to chance**; a p-value of ~0.688
- The difference in reptiles (6%) and mammals (17%) **is significant**; a p-value of ~0.038

Therefore, certain types of species **are** more likely to be endangered than others.

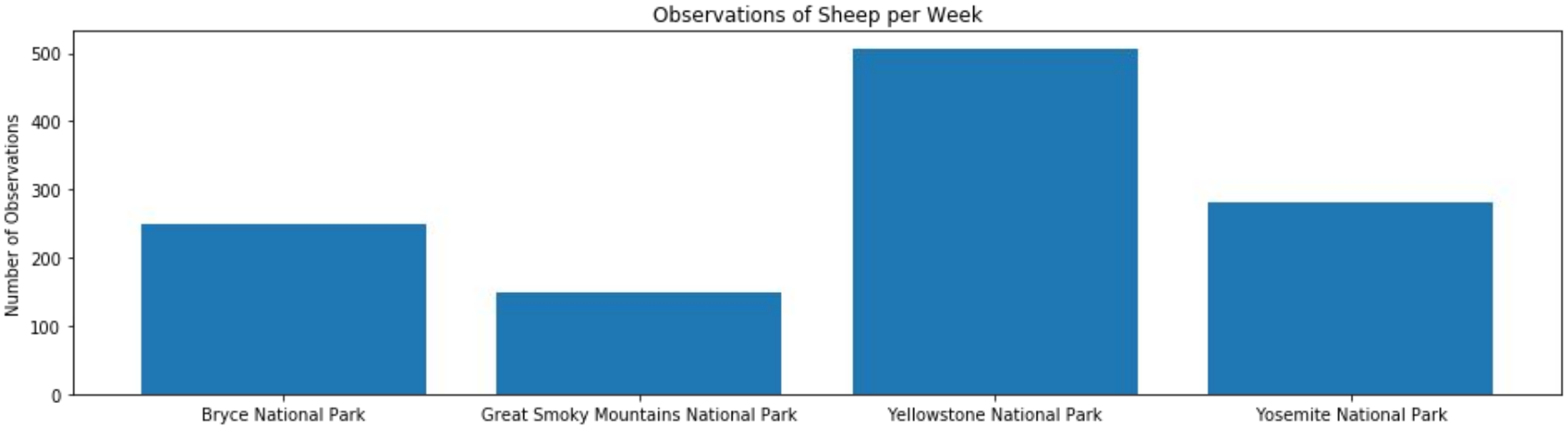
Based on these significance calculations,

amphibians, reptiles, and fish displayed significant difference when running the chi-squared test.

I advise that conservationists look into these 3 types of species, as they are more likely to be endangered than the other types.

Vascular and non-vascular plants are the least to worry about.

Foot & Mouth Disease Reduction Program: Findings



Last year, 15% of sheep at Bryce National Park had foot and mouth disease.

So, using:

- this 15% **baseline of occurence**
- any 5% **drop in observed cases of the disease**
- the observation data
- A/B test sample size calculator (by Optimizely)

The results were...

A sample size per variation of 870.

This means that...

- You would need to observe at least 510 sheep
  - That’s approximately 1 week of observing in Yellowstone, or
  - Approximately 2 weeks of observing in Bryce