

The background features a dark blue-grey field on the left, transitioning into a series of overlapping, semi-transparent green geometric shapes on the right. These shapes include triangles and polygons in various shades of green, from light lime to deep forest green, creating a layered, abstract effect.

National Park Biodiversity Project

by Lori Anne Madison

Data in the CSV

- ▶ The species_info.csv file contains data about all the organisms found in the national parks:
 - ▶ category
 - ▶ scientific name
 - ▶ common name(s) and
 - ▶ conservation status
- ▶ The file describes 5823 species
- ▶ Data includes plants as well as animals, further distinguishing them into vascular and nonvascular plants. However, no insects were included.

category	scientific_name
Mammal	Clethrionomys
Mammal	Bos bison
Mammal	Bos taurus
Mammal	Ovis aries
Mammal	Cervus elaphus

Protected Status Between Categories

GOAL: develop and run code to determine if different types of species are more likely to be endangered or threatened

Statistically significant

- ▶ Mammals are more likely to become endangered than reptiles
 - ▶ The Chi-square test shows a statistically significant difference
- ▶ All animals are more likely to become endangered than plants
 - ▶ The differences between either of the plant categories and any of the animal categories are statistically significant

Not significant

- ▶ The difference between the endangerment rates of mammals and birds is not significant
- ▶ Also, there are no significant differences between vascular and nonvascular plants, or between amphibians and mammals
- ▶ Although the Chi-square test shows the difference between fish and mammals is not significant, it just barely misses the .05 cutoff (p-val is 0.056), so it should be investigated further

Recommendation for Conservationists

According to the significance data, animals are a much higher conservation priority than plants. The differences between reptiles, fish, and amphibians are insignificant, but the difference between mammals and reptiles shows that mammals are more likely to be threatened than reptiles.

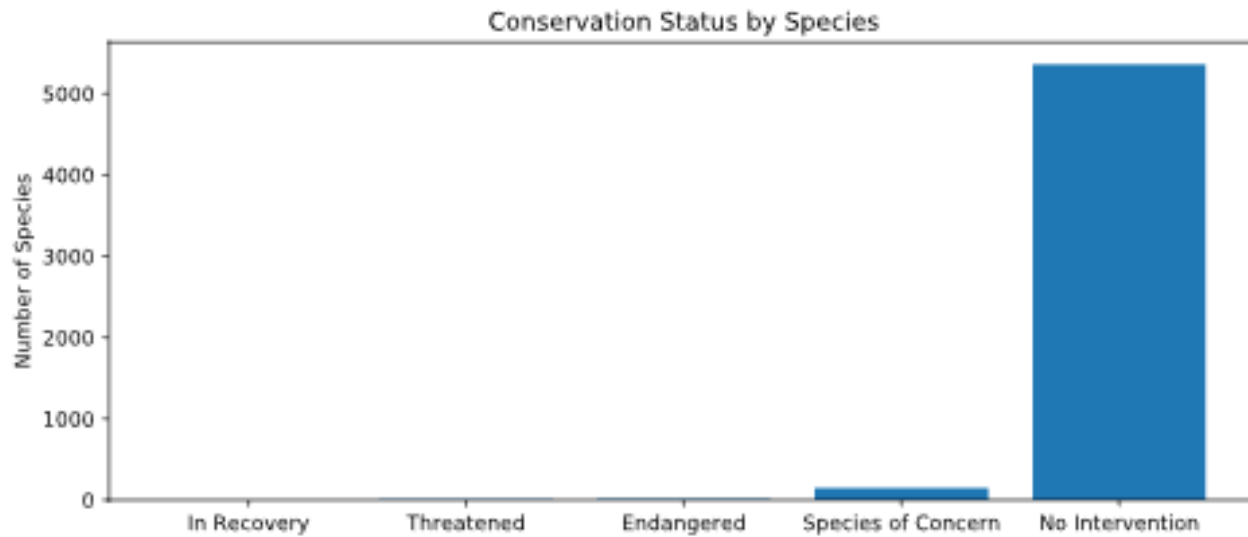
Therefore, mammals are a higher conservation priority than reptiles.

Preemptive action taken to prevent endangerment should first be taken on animals other than reptiles, then reptiles, then plants.

Sample Size Determination

- The sample size determination exercise for foot and mouth disease consisted of some relatively simple code, and the results of those calculations being plugged into a sample size calculator.
- The baseline percentage - the percentage of sheep at Bryce National Park with the disease - is 15%, so I wrote that into the baseline variable.
- The description says that the minimum percentage of change the scientists want to observe is 5%, so I used $(100 \times 5) / 15$ for that, as 15 was the baseline and the minimum detectable effect is a percentage of the baseline.
- After plugging these into the calculator, I found that the sample size per variant was 870. Then I simply divided this by the number of observations in each park to get the weeks.
- For Yellowstone Park, there were 507 observations so the scientists must spend 1.7 weeks (rounded, that would be two). For Bryce Park, there were only 250 observations, and so the scientists have to spend about 3.5 weeks there.

Graphs



These are the graphs I've created!
I don't think tables and DataFrames
count, so I didn't include them.

