

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light mint green. They are positioned diagonally, with the blue one partially covering the green one.

Codecademy Capstone Project - Biodiversity in National Parks

by Riley Adams

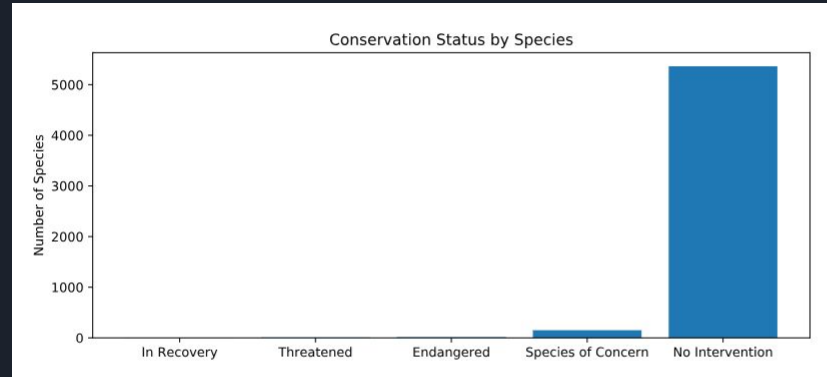


Conservation Data

Conservation Status	Species Count	Description
Endangered	15	Seriously at-risk of extinction
In Recovery	4	Formerly endangered but no longer in danger of extinction
No Intervention	5,363	No threat faced at this time
Species of Concern	151	Declining population or in need of conservation
Threatened	10	Vulnerable to endangerment in near future

- Inspected dataframe to become familiar with data
 - Performed count of the species falling into each category
- To include all species in dataframe, needed to assign value to those not needing intervention
 - Shown above as “No Intervention”
- Vast majority of species no in danger, however several are on the verge of slipping into “Endangered” category

Conservation Status by Species



- Created chart providing visual of conservation status by species
 - Most fall do not need intervention for continued survival, but several crossing into concerning categories
- But this masks the story behind these buckets - some species are more likely to be endangered than others
- Using a pivot table, we were able to organize the data to have a better view for analysis

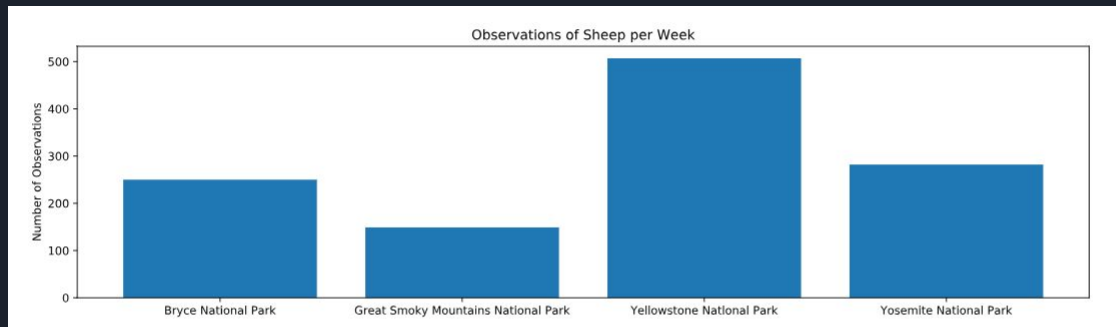


Testing for Significance

- At first blush, it would seem that some species categories are more protected than others
- To test this finding, we conducted some significance tests
 - Chi-Square Test
 - Running the Chi-Square Test, we saw no significant difference in percent protection between mammals and birds
 - p-value of 0.688 - indicates the difference between the percentages of protected birds and mammals is not significant and is a result of chance
 - We did find a significant difference between reptiles and mammals, however
 - P-value of 0.038 indicates there is a significant difference between reptiles and mammals
 - We can conclude that certain types of species are more likely to be endangered than others
 - We recommend that conservations use this information to inform their efforts regarding protection of certain species categories

Counting Sheep, but not to Fall Asleep

- After sorting some additional data based on sheep counts, we decided to identify opportunities to pinpoint necessary data thresholds for examining disease propagation
- Data was grouped by park name and common name
- Below shows observations of sheep per week by National Park





Sheep Sampling

- Identified Baseline of 15%
- Calculated Minimum Detectable Effect of $100 * 5/15$
- Using data from Bryce National Park, to be sure of a >5% drop in observed disease cases at Yellowstone National Park we would need a sample of 870 sheep, which would take ~1.7 weeks to collect a sufficient amount