Biodiversity for the National Parks

Results and Observations

The Data - species.info.csv

Species.csv contains data about the species in U.S. National Parks. The data file tells us:

- The **scientific name** of each species
- The **common name** of each species
- The species **conservation status** (Endangered, In Recovery, No Intervention, Species of Concern, Threatened)

There are **5541 unique species** across the national parks.

There are **seven** different **species types**:

1. Mammal
2. Bird
3. Reptile
4. Amphibian
5. Fish
6. Nonvascular Plant
7. Vascular Plant

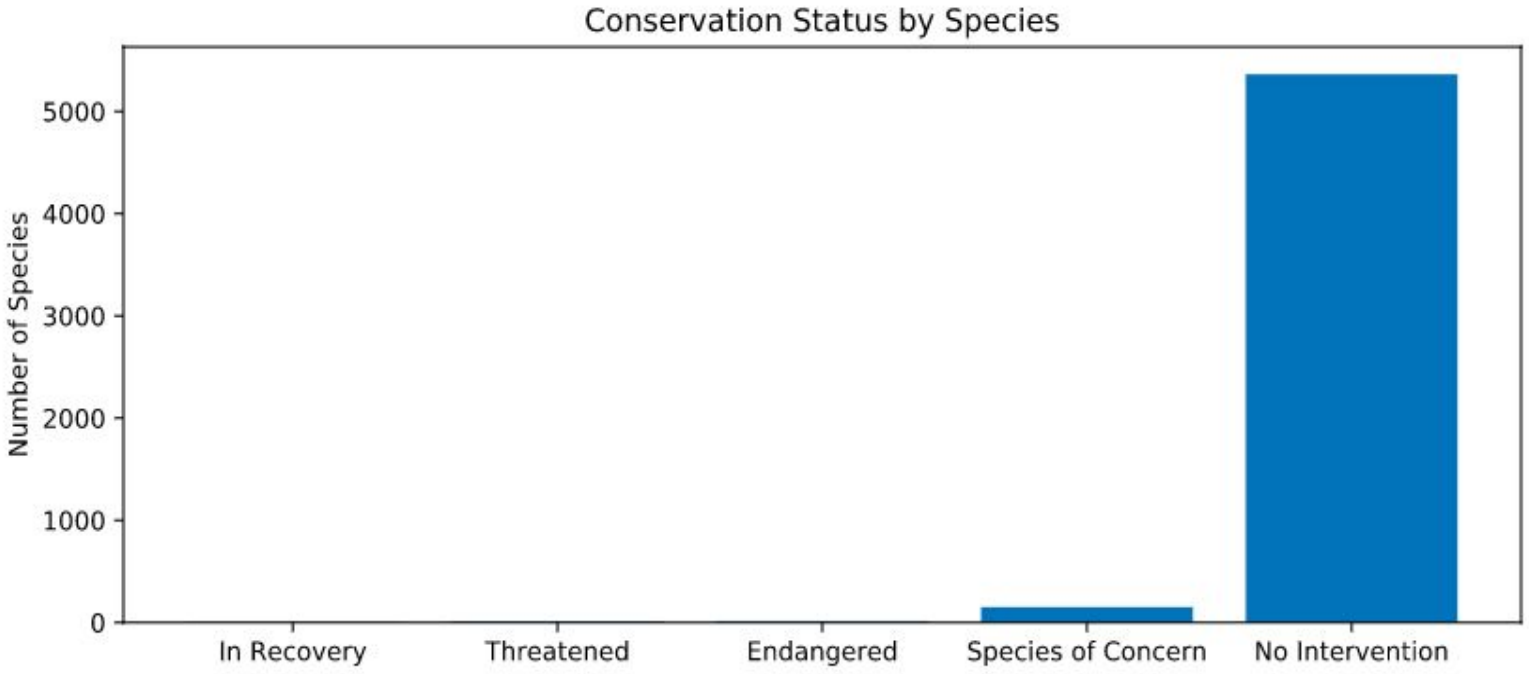
The Data - species.info.csv

Conservation Status

- **Species of Concern:** declining population or appears to be in need of conservation.
- **Threatened:** vulnerable to endangerment in the near future.
- **Endangered:** seriously at risk of extinction.
- **In Recovery:** formerly Endangered, but currently not in danger of extinction throughout all or a significant portion of its inhabitable range.
- **No Intervention:** Population is steady and has never been endangered or threatened

Out of the **5541** species in the National Parks: **15** are Endangered, **4** are In Recovery, **10** are Threatened, **151** are Species of Concern, and the vast majority, **5363**, require No Intervention. The following slide shows this information in bar graph form.

Conservation Status by Species



Are certain types of species more likely to be endangered?

category	not_protected	protected	percent_pr
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015

Diving deeper into the data, we created a table to answer the question, “Are certain types of species more likely to be endangered?”

When we broke each species down into the categories of “protected” and “not protected” and then calculated the “Percent of each species that is protected”, the numbers showed that two species, Bird and Mammal, were significantly higher in this number than the others, at 15 and 17 percent.

Are certain types of species more likely to be endangered?

We then did a Chi-Squared Significance Test to see if our numbers were showing what we thought they were, which was that Mammals were more likely to be endangered than Birds (17 percent vs. 15 percent). The Chi-Squared Significance Test would either prove or disprove our null hypothesis, which was that the difference between Birds and Mammals was due to chance and was not significant.

When we ran our Chi-Squared Test we found a p-value of .68, which since it is greater than .05, indicates that our null hypothesis was correct: The difference between Birds and Mammals was due to chance and that Mammals were **not more likely to be endangered** than Birds.

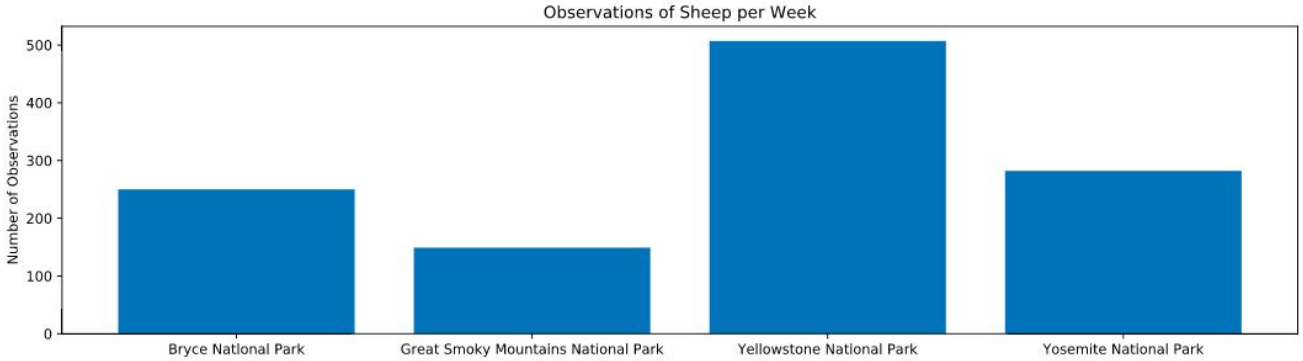
However, when we performed this same Chi-Squared test on **Reptiles** vs. Mammals, we found a p-value of .038, which indicates that we can not take the null hypothesis as correct, and Reptiles **ARE** more likely to be endangered than Mammals. Which means the answer to our question, “Are certain types of species more likely to be endangered?” is **YES**.

Recommendations

When looking at the “Percent Protected” numbers, Reptiles were actually quite low compared to Birds and Mammals with a value of around 6%. As it turns out, this number could be misleading, because as shown in the previous slide, when we did the significance test, Reptiles are actually more likely to be endangered than Mammals.

Conservationists working with the national parks should take advantage of the data available and the data analysis tools to get a better picture of which species need to be looked after more closely, and where the bulk of their efforts should be focused. We could do Significance Tests across all the categories and find which are most likely to be endangered and this could help the conservationists to keep even more species off the endangered list.

Foot and Mouth Disease Study



Using the Species.csv file we determined which species were sheep, and then sorting for which species of sheep were mammals, we came up with three total species of sheep. Merging this with our Observations file, we were able to graph the total observations of all three species at each National Park.

Foot and Mouth Disease Study

We determined the Sample Size of 870 (this is the number of sheep the scientists would need to observe at each park to ensure that their foot and mouth percentages are significant and to know whether their program to reduce the disease is working.) using a baseline conversion rate of 15%, a minimum detectable effect of 33.3%, and a statistical significance of 90%.

	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282

To observe enough sheep at Yellowstone, for example, the scientists would need to be there for (870/507) = 1.71 weeks.