Biodiversity for the National Parks

Introduction to Data Analysis
Capstone Project

Relevant Findings

- The dataset provided by the National Parks Service contained information on 5541 different types of plants and animals, including their common name, scientific name, and conservation status.
- The 5541 species were broken down by category, as provided below:
  - Mammals
  - Birds
  - Reptiles
  - Amphibian
  - Fish
  - Vascular Plant
  - Nonvascular Plant

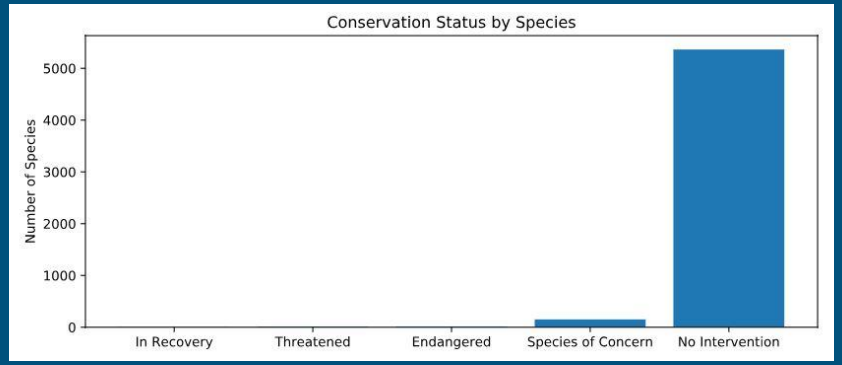
Relevant Findings

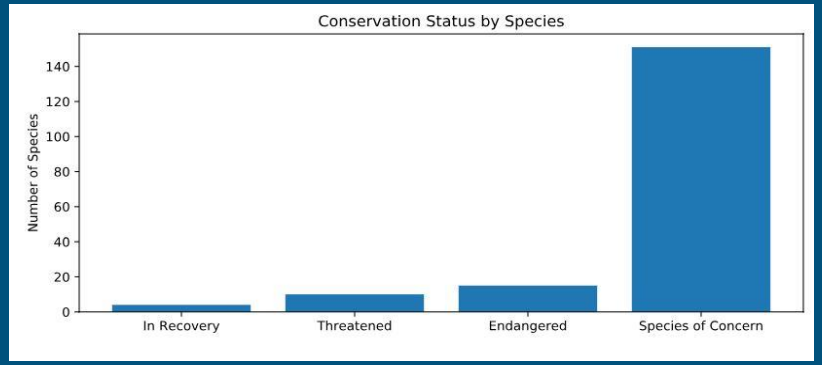
Of the species provided, 180 required some type of conservation effort, or had in the past. The complete breakdown of these statuses is shown in the chart and table below:

Conservation Status	Number of Species
Endangered	15
In Recovery	4
No Intervention	5353
Species of Concern	151
Threatened	10

Relevant Findings

Provided below are graphs representing the conservation status for the species in the dataset. The graph on the left includes species without intervention, while the graph on the right excludes this category in order to better display the other categories.





Significant Calculations and Analysis

Conducted calculations determining the percentage of each category of species from the dataset that had some level of protection:

- Amphibians - 8.86%
- Bird - 15.37%
- Fish - 8.73%
- Mammal - 17.04%
- Reptile - 6.41%
- Vascular Plant - 1.08%
- Nonvascular Plant - 1.50%

Significant Calculations and Analysis

Also sought to determine whether the difference between these percentages were significant, or merely the result of chance based on the available information.

Both mammals and birds had higher percentages of protected species than the rest, but a chi-squared test between these different categories showed that only the mammal percentage was significant compared to the others. Additionally, the difference in percentages for all species of animals compared to species of plants were determined to be significant.

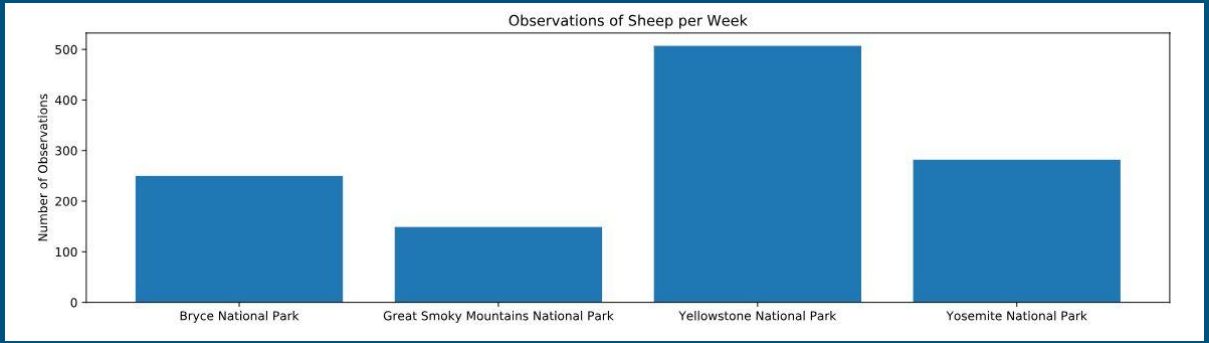
Recommendations

Based on the results of the significant calculations, it is clear that mammals are the most likely type of species to be endangered. It is recommended that these species be more closely monitored in the future and that conservationists take steps to further protect them.

Additionally, conservationists should attempt to gather more information on other animal species, in particular birds, to determine whether a larger dataset will affect the test for significance.

Foot and Mouth Study

In order to confidently determine whether the program to reduce Foot and Mouth Disease by 5% at Yellowstone National Park has been successful, it was first necessary to determine how many sheep were observed each week per park. This results of this calculation are shown in the bar graph below.



Foot and Mouth Study

Next, to determine the sample size necessary to show if the program has been successful, I used the 5% reduction goal and the current baseline of 15% of sheep at Bryce National Park that have the disease to calculate the ideal sample size.

Based on the figures provided, the Park Rangers will have to observe at least 870 sheep at each park in order to get an accurate representation of the effects. The below chart shows how long this will take at each park.

National Park	Observations Per Week	Time to Reach Sample Size
Bryce National Park	250	3.5 Weeks
Yellowstone National Park	507	1.7 Weeks
Great Smoky Mountains National Park	149	5.8 Weeks
Yosemite National Park	282	3.1 Weeks