Biodiversity for the National Parks

Dustin J. Mattingly

The Data

In the given data frame, species_info.csv, was a list of over 5000 species of plants and animals listed by both their scientific and common names.

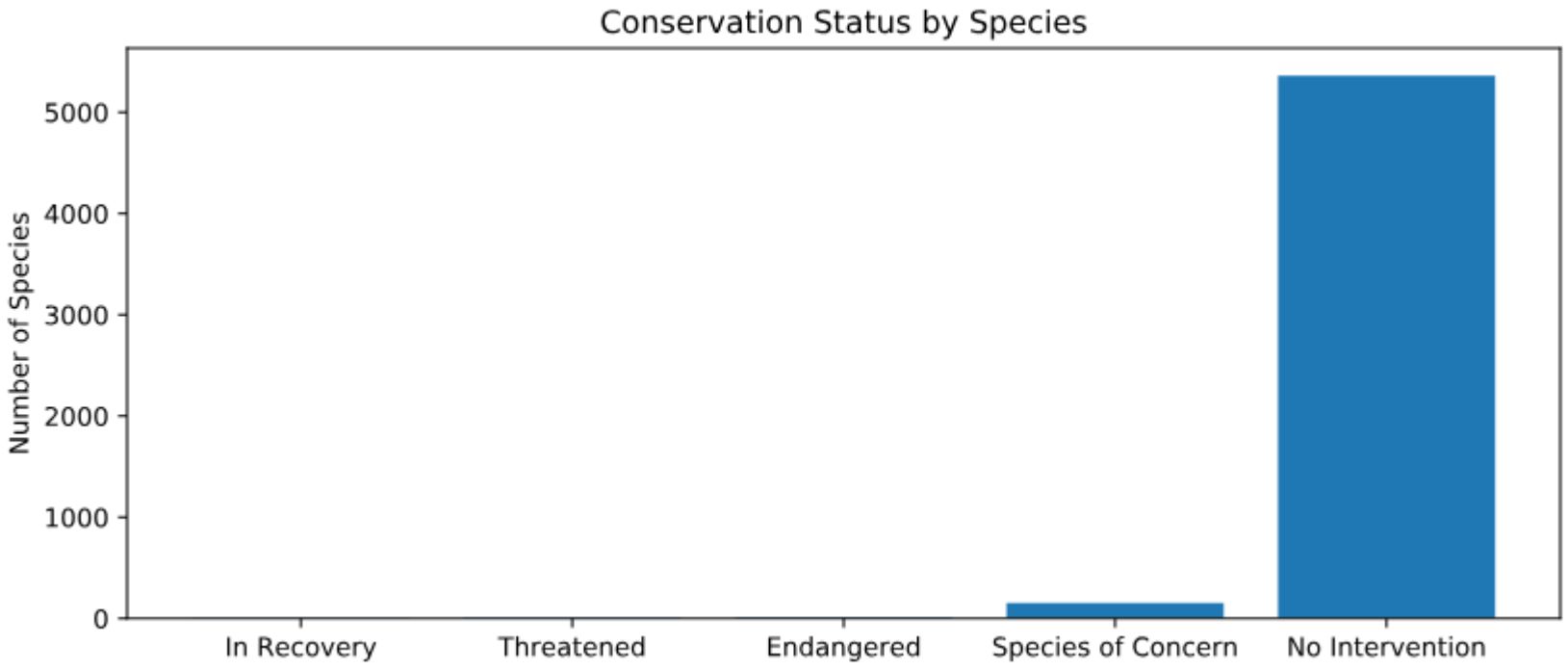
The Data

As well, each species was given a conservation status.

1. **Species of Concern**
2. **Threatened**
3. **Endangered**
4. **In Recovery**
5. **NaN (No Intervention)**

The Data

The figure below is a visual representation of the data as given.



Significance Calculations

Next, I created a table breakdown of the six categories of species to examine how many of each were protected.

is_protected	category	False	True
0	Amphibian	72	7
1	Bird	413	75
2	Fish	115	11
3	Mammal	146	30
4	Nonvascular Plant	328	5
5	Reptile	73	5
6	Vascular Plant	4216	46

Significance Calculations

To better present this data, I calculated the percentage of each category of species protected.

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

Significance Calculations

Looking at this data we can see that the categories which need the most protecting are mammals and birds.

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

Significance Calculations

The data also implies that mammals are more endangered than birds.

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

Significance Calculations

A chi-squared test was used to determine if the difference between mammals and birds was significant.

When the test was ran, the p-value was ~0.688. Therefore, the difference between the percent protected of these two category of species is not significant, but just the result of chance.

Significance Calculations

However, when the same chi-squared test was done to compare the percent protected of mammals with every other category of species, I found three of the p-values to be less than 0.05 marking a significant difference.

This demonstrates that there are certain types of species more likely to be endangered than others

Recommendation

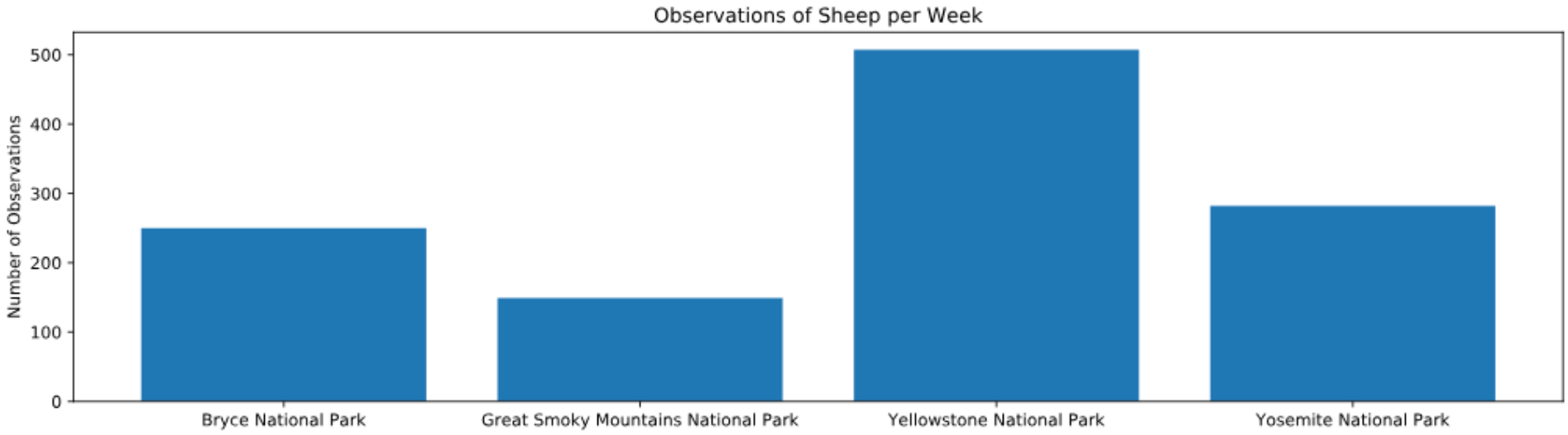
It is recommended the conservationists concerned about endangered species should continue to protected mammal species in our national parks as they are more likely to be endangered than reptiles, nonvascular plant, and vascular plants.

The Data

In the given data frame, observations.csv, was a list of sightings recorded by conservationists at several national parks over a span of seven days. This list included over 23000 species listed by scientific name.

The Data

This data was combined with the previous data set as to narrow the list to those that are both mammals and sheep. With this data the figure below was constructed to display the number of sheep sightings in the given parks within seven days



Significance Calculations

Using a sample size calculator, I plugged in the default level of significance of 90%. As well, the recorded 15% baseline of foot and mouth disease among sheep at Bryce National Park the previous year.

To calculate the minimum detectable effect I calculated

$$100 \times 5/15$$

to get 33.33%.

Significance Calculations

Once all the data was placed into the sample size calculator, it was determined that a sample size of 870 sheep would need to be observed to ensure the foot and mouth percentages be significant.

This paired with the data of how many observations of sheep were made in each park over seven days can be used to determine approximately how long it would take to make 870 observations in each park.