Biodiversity in the National Parks

Using data analysis to guide and inform our stewardship of nature.

Our Diverse Ecosystem

- Our national parks are home to a wide variety of flora and fauna.
- The National Parks System maintains a publicly available database of the various plant and animal species that live in the parks.
- The database tells us the scientific and common names for organisms and categorizes them by type and conservation status.

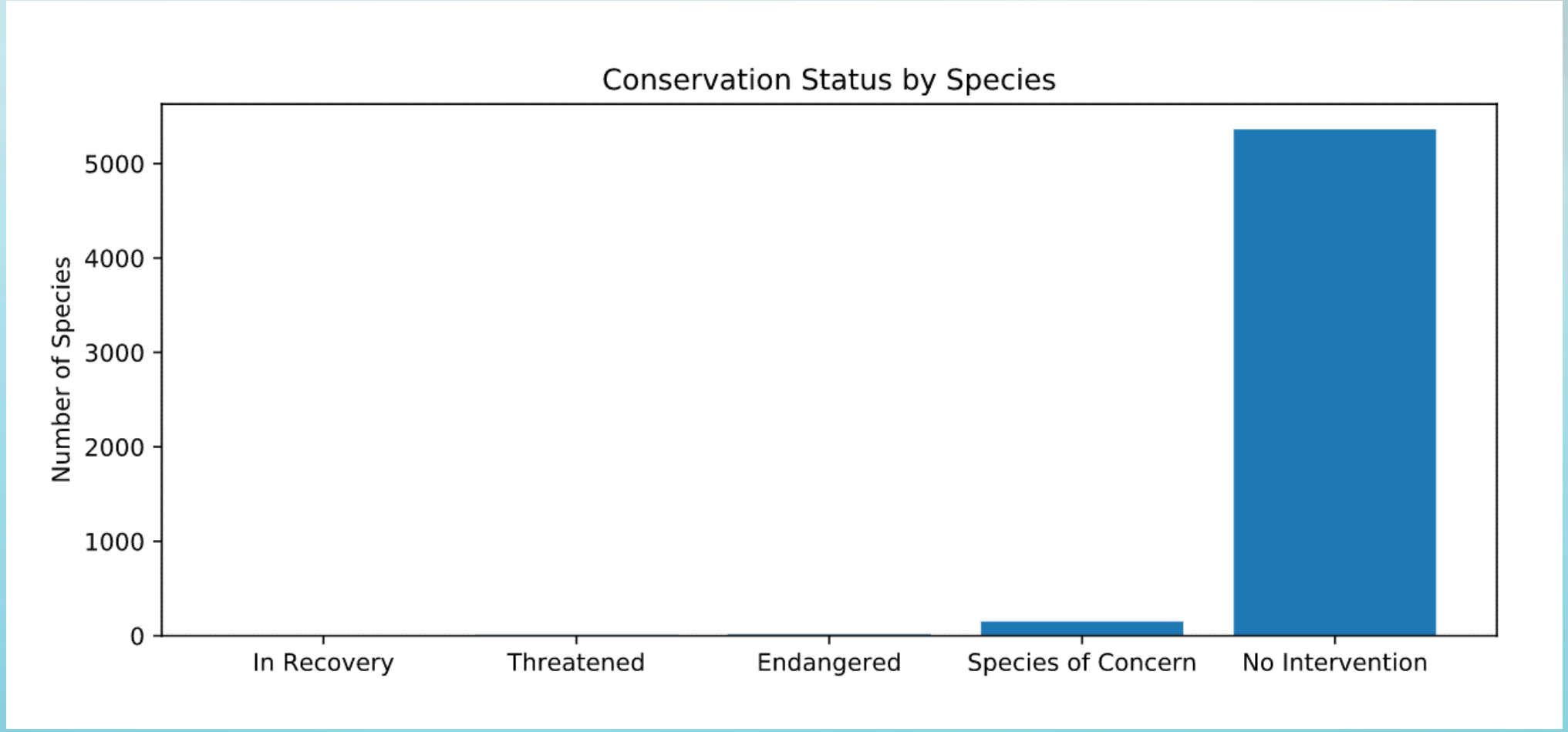
Organism Types

- Mammal
- Bird
- Reptile
- Amphibian
- Fish
- Vascular Plant
- Nonvascular Plant

Conservation Statuses

- **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
- **Species of Concern**: declining population or appears to be in need of conservation
- **Threatened**: vulnerable to endangerment in the near future

Distribution of Conservation Status



Conservation Status Discussion

- The good news is that most of the known species are not in need of special conservation attention.
- It is also good that the number of threatened and endangered species is relatively low compared to the total number of species.
- However, considering that a species is “in recovery” after having been endangered, this data seems to indicate that not many species have recovered from an endangered status.

Status	Species Count
In Recovery	4
Threatened	10
Endangered	15
Species of Concern	151
No Status	5363

Drilling Into Protected Species

- Let's investigate what kinds of plants and animals are most commonly given a protected status.
- Animals are more commonly protected than plants.
- Amphibians, fish and reptile species have moderate rates of protection.
- Birds and mammals are the species most likely to be given protected status.

Category	Not protected	Protected	%
Amphibian	72	7	8.86
Bird	413	75	15.37
Fish	115	11	8.73
Mammal	146	30	17.05
Nonvascular Plant	328	5	1.5
Reptile	73	5	6.41
Vascular Plant	4216	46	1.08

Protected Species, Statistically Speaking

- When looking at data like we see in the table on the previous slide, it is tempting to compare and contrast the data and draw conclusions just by eyeballing it.
- However, we're far from unbiased when it comes to making fair judgments about the data. For example, based on the available data, are birds more likely to be protected than reptiles? Are mammals more likely than birds to be protected?
- We can use a chi-squared test to rid us of our prejudices and tell us whether or not any given category of organism is significantly more likely to be protected than another, given the data we have.
- If a chi-squared test between two categories of organism yields a p-value greater than 0.05, then **no significant difference can be said to exist between the populations, given the available data.**

Chi-Squared Test Outcomes (p-values)

	% protected	Bird	Fish	Mammal	Nonvascular Plant	Reptile	Vascular Plant
Amphibian	8.86%	0.1759361	0.8247942	0.1275766	**0.0017785**	0.7814504	**1.044e-08**
Bird	15.37%		0.0766819	0.6875948	**1.054e-10**	0.0531354	**4.612e-79**
Fish	8.73%			0.0561483	**0.0004960**	0.7406524	**1.486e-12**
Mammal	17.05%				**1.481e-10**	**0.0383555**	**1.440e-55**
Nonvascular Plant	1.5%					**0.0336269**	0.6623419
Reptile	6.41%						**0.0001450**
Vascular Plant	1.08%						

Chi-Squared Results Discussion

- This table shows the p-value obtained by performing a chi-squared comparison between each pair of organism categories. It also shows the percent of protected organisms in each category for reference.
- Some of the results are surprising. For example, even though 15.37% of bird species are protected and only 6.41% of reptile species are protected, the p-value of our chi-squared test is above 0.05, indicating that we don’t quite have enough data to conclusively say that birds are actually more likely to be protected than reptiles. In other words, the difference in our data between reptiles and birds may be the result of random chance rather than a true indication of a difference between populations.
- What other data in this table surprises you?

The Importance of Sample Size

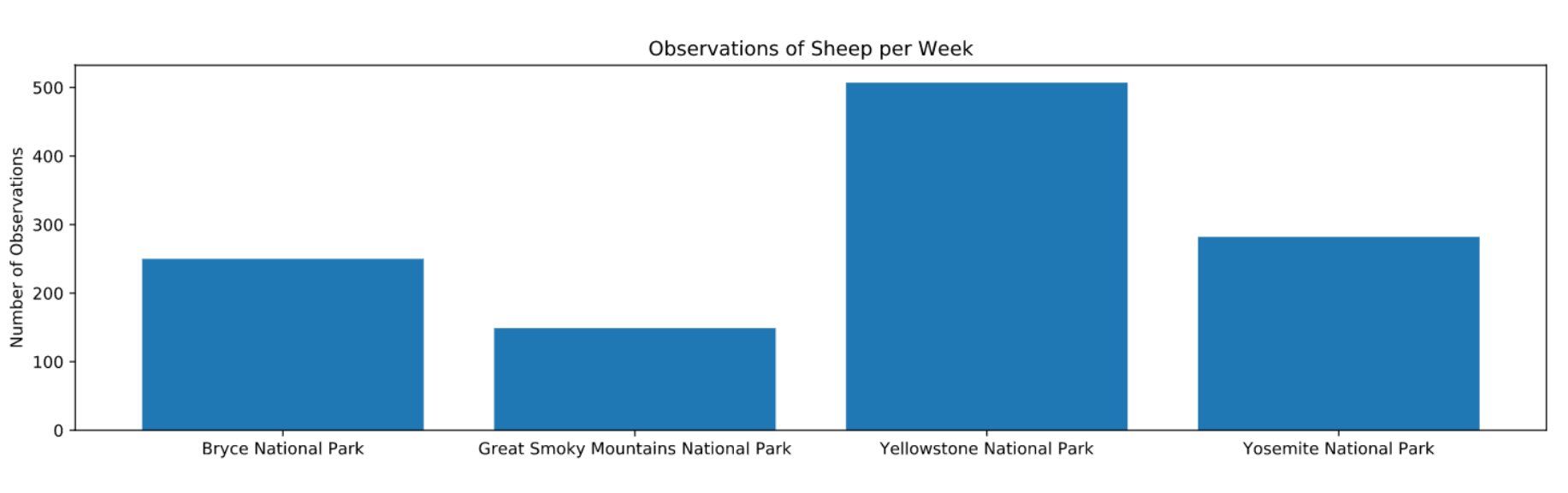
- Sometimes when we end up with surprising results like the ones we got from our chi-squared significance test, it may simply be the case that we don’t have enough data to *confidently* make a conclusion.
- It turns out that statistics also provides formal mechanisms for determining how large our sample size must be in order to achieve an adequate confidence level.
- There are many calculators available to help us apply statistics theory to determine the sample size required to answer a question with a desired level of confidence.

Foot and Mouth Reduction: An Exercise In Sample Size Determination

- Over the past year, rangers at Yellowstone National Park have been running a program to reduce foot and mouth disease in the park.
- We know that 15% of sheep at Bryce National Park had foot and mouth disease last year.
- We would like to be able to detect a 5% difference in rate of disease (33.33% of our baseline) at a 90% confidence level.
- This is all we need to know in order to use a sample size calculator to tell us that we need to sample 870 sheep to determine whether or not efforts to reduce disease occurrence have been successful!

Sheep Observations

The length of time it will take us to survey 870 sheep depends on how frequently sheep are observed in the various parks. Fortunately, we have this data for several national parks. Over the past 7 days, 250 sheep have been sighted at Bryce National Park, 149 at Great Smoky Mountains National Park, 507 at Yellowstone National Park, and 282 at Yosemite National Park:



We can divide our required sample size by observation frequency to determine how many weeks a survey will take at each park. At Bryce it will take 3-4 weeks, at Great Smoky Mountains 5-6 weeks, only 2 weeks at Yellowstone, and maybe slightly more than 3 weeks at Yosemite. It's potentially a long time to wait, but we'll be patient if we care about getting accurate results!

The Conservationist Take-Away

Conservationists who are interested in deciding where their efforts are needed the most should think about what this means. Statistical analyses like these are important because they free us from making judgments based on what we see through the narrow lens of the data we happen to have. In many cases, the unbiased answers that statistical calculations yield may be surprising, and may simply mean that more data points are required before we can be confident about saying one type of organism is more likely to require protection than another, but without them we're subject to biases that we cannot see.