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 green. They are positioned diagonally, with the blue one partially covering the green one.

# Biodiversity for the National Parks Capstone Project

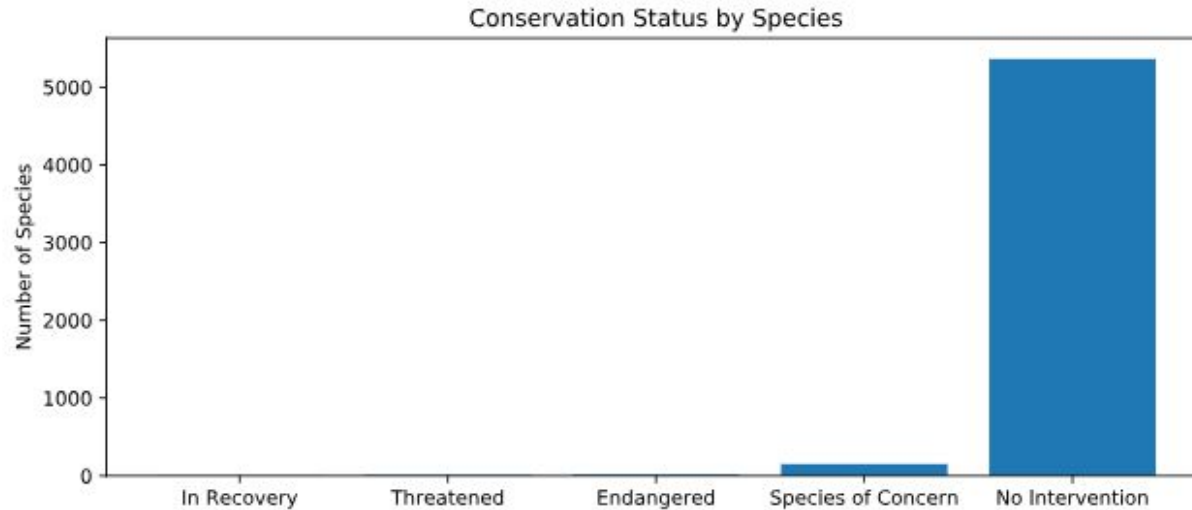
Kelsey Siegel

# species\_info.csv

- What did our original data display?
  - Initially, our data displayed a table with columns titled: category, scientific name, conservation status, and common name.
  - There were a total of 5,823 species listed.
  - The categories given were: amphibian, bird, fish, mammal, nonvascular plant, reptile, and vascular plant.
  - Conservation status was given as: Endangered, In Recovery, Species of Concern, or Threatened.

	index	category	scientific_name	common_names	conservation_status
Amphibian	80	1	79	80	3
Bird	521	1	488	481	3
Fish	127	1	125	122	3
Mammal	214	1	176	205	4
Nonvascular Plant	333	1	333	246	1
Reptile	79	1	78	79	1
Vascular Plant	4470	1	4262	4291	3

# Species Sorted by Conservation Status





# Investigating Endangered Species

	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

Provoking Questions:

1. Which species appear to be endangered?
2. Which species are protected more than others?
3. Does this mean that certain species are more likely to be endangered than others?



# Chi-Squared Test for Significance

- Are certain types of species more likely to be endangered?

## MAMMALS vs. BIRDS

- On the surface, it may look as though mammals are more likely to be endangered than birds. However, using a chi-squared test for significance, we can see that the resulting p-value of 0.6875948 indicates no significant difference between mammals and birds.

## MAMMALS vs. REPTILES

- On the other hand, when we compared mammals to reptiles and ran a chi-squared test for significance, we notice that the resulting p-value was 0.03835559. This indicates that there is a significant difference between mammals and reptiles.

In conclusion, our chi-squared test for significance illustrates that certain types of species are more likely to be endangered than others.

# Recommendation for Conservationists



Based on our significance calculations, it is clear to see that some species are at a higher risk of becoming endangered than others. After calculating p-value significance, we can see that reptiles are more likely to become endangered species than mammals. Conservationists should make an effort to identify and locate the specific species that are more likely to become extinct. They can help protect them in our national parks by sustaining a healthy and suitable habitat for these animals.





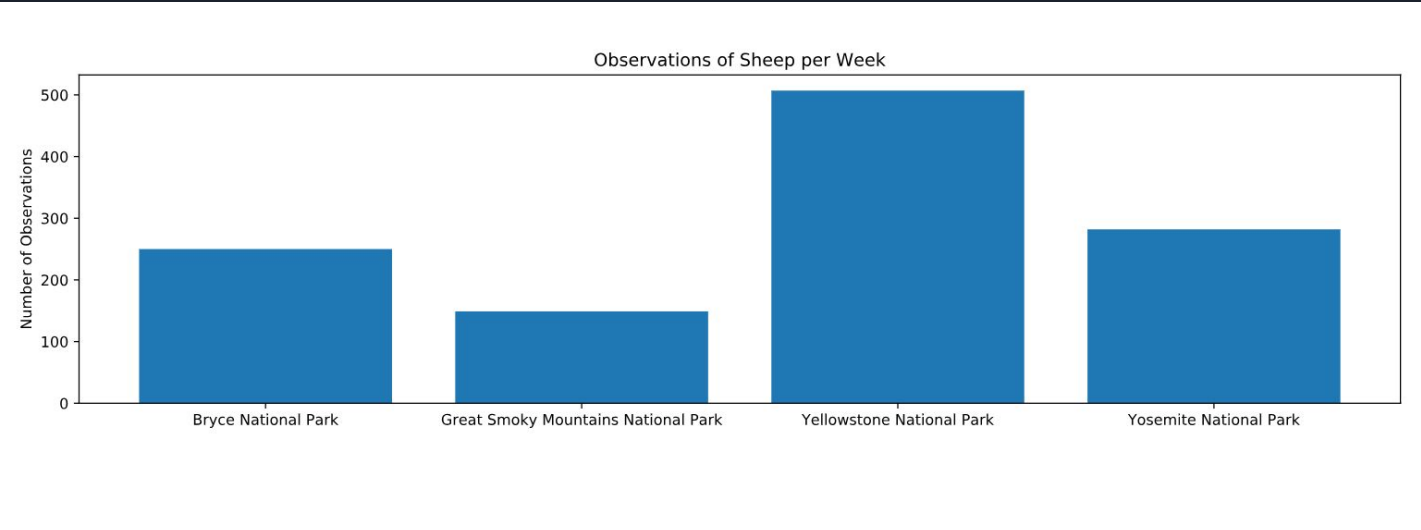
# Total Sheep Sightings per National Park

The table below indicates how many sightings of sheep (across all three species) were found in each of the national parks listed.

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

# Total Sheep Sightings per National Park

The graph below indicates the number of sightings of sheep (across all three species) that were found in each of the national parks listed.





# Foot and Mouth Disease Reduction Effort

## Determining the Sample Size

script.py

```
1 baseline = 15
2
3 minimum_detectable_effect = 100*5./15
4
5 sample_size_per_variant = 870
6
7 yellowstone_weeks_observing = sample_size_per_variant/507.
8
9 bryce_weeks_observing = sample_size_per_variant/250.
```

<https://s3.amazonaws.com/codecademy-content/courses/learn-hypothesis-t>

Baseline conversion rate: 15 %

Statistical significance: 85% 90% 95%

Minimum detectable effect: 33 %

Sample size: 890

Given a baseline of 15% occurrence of foot and mouth disease in sheep at Bryce National Park, we discovered that at least 890 sheep should be observed if the scientists wanted to be sure that a >5% drop in observed cases in the sheep at Yellowstone was significant. Using the data we found earlier, we calculated that this would take approximately one week of observing in Yellowstone and approximately two weeks in Bryce National Park.