Objectives

- Determine if certain categories of animals are more at risk than others

- Make recommendations based on this analysis for further understanding and protection of biodiversity in the national park system

- Analyze a dataset of the number of animal observations in each national park
  - Help determine sample size requirements for a future study into the efficacy of a foot-and-mouth disease intervention carried out by Yellowstone National Park

Understand our data on national park biodiversity
What categories of organisms are at risk at what rates?

- We looked at 5541 unique species of plants and animals

- Most (4595) are plants

- ~6.5% of species are vulnerable (under any protection category other than "no intervention"

Organisms by Category	
Amphibian	79
Bird	488
Fish	125
Mammal	176
Nonvascular Plant	333
Reptile	78
Vascular Plant	4262

Protection Categories	
In Recovery	4
Threatened	10
Endangered	15
Species of Concern	151
No Intervention	5363

Conservation Status by Species
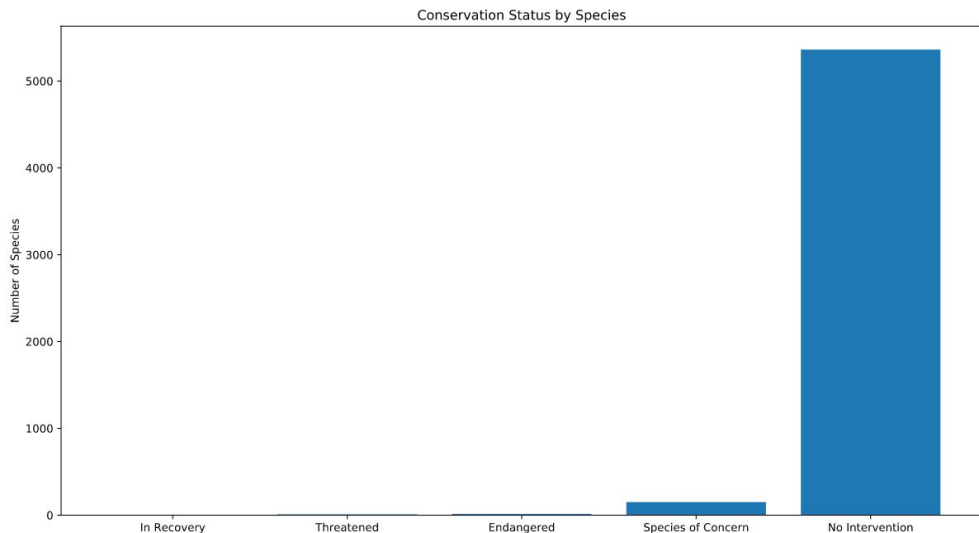
Are certain categories more vulnerable than others?

- Categorize into "protected" vs "not protected"

- Calculate percent protected `(protected / (protected + not protected)`

- Hypothesis: Category of species is significantly correlated to rate of endangerment (not protected)

- How do we test this hypothesis?
  - Data is categorical
  - More than two categories
  - Therefore we should use a chi-squared test

- Chi-squared tests:
  - Mammals vs birds: p-value = ~0.69 (not significant)
  - Mammals vs reptiles: p-value = ~0.04 (significant)
  - All animals (plants excluded): p-value = ~0.03 (significant)

```
contingency = [[7, 72],
               [75, 413],
               [11, 115],
               [30, 146],
               [5, 73]]
chi2, pval_all, dof, expected = chi2_contingency(contingency)
```

Category	Not protected	Protected	% protected
Amphibian	72	7	8.86%
Bird	413	75	15.37%
Fish	115	11	8.73%
Mammal	146	30	17.05%
Reptile	73	5	6.41%
Nonvascular Plant	328	5	1.50%
Vascular Plant	4216	46	1.08%

Conclusions & recommendations

Conclusion - some species of animals are more likely to be endangered than others
- E.g Mammals are not statistically significantly more likely to be endangered than birds
- E.g Mammals are statistically significantly more likely to be endangered than reptiles

Recommendation - further study required
- Investigate how environmental factors affect categories of animals differentially
- Investigate why certain categories of animals like mammals struggle to thrive more than others like reptiles
- This will help us determine what future interventions can be done to help the most vulnerable species

Foot-and-mouth disease intervention efficacy

Context: Yellowstone National Park has been running a foot-and-mouth disease intervention

Objective: Determine if the intervention has been effective

- We will need to sample animals from Yellowstone and check their post-intervention rate of disease
- We must determine the required sample size to determine the rate with confidence
- We must estimate how many weeks of observations it will take to collect the necessary samples

Existing data: The rate of foot-and-mouth disease last year at **Bryce National Par**k was **15%**

- This is the only prior data we have to base our analysis on
- We will use this data to determine our required sample size

Foot-and-mouth disease intervention efficacy

- We analyzed data of number of observations of each species at each park

- We filtered the data to focus on sheep species
  - This is because we only have data on rates of sheep disease from Bryce National Park

- We must be able to detect reductions of at least 5 percentage points

- We will calculate within a 90% significance level

Park name	# sheep observations
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite National Park	282

Observations of Sheep per Week
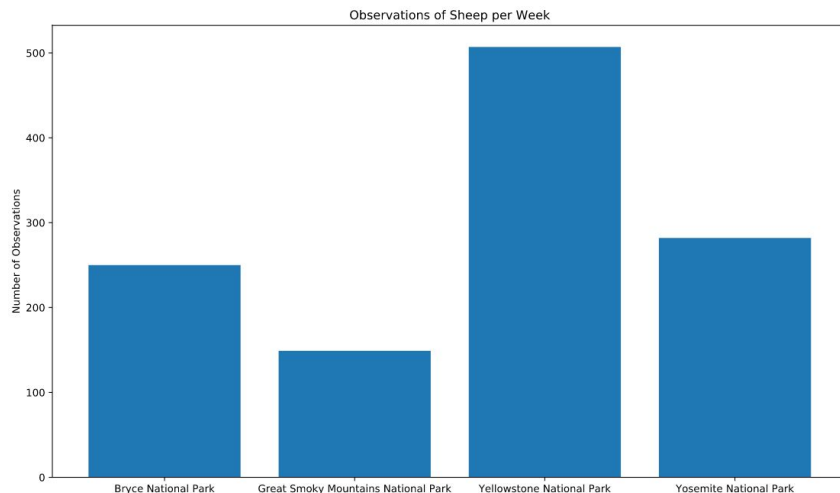
Foot-and-mouth disease intervention efficacy

Sample size calculations (calculated with tool at optimizely.com)**:**

```
minimum_detectable_effect = 100 * 0.05 / 0.15 = 33.33
baseline = 15
sample_size_per_variant = 510
```

Estimated observation weeks calculations:

```
Bryce = 510 / 250 = 2.04
Yellowstone = 510 / 507 = 1.00
```

Park name	# sheep observations / week
Bryce National Park	250
Yellowstone National Park	507

Based on these calculations, it will take about two weeks of observations at Bryce National Park and one week at Yellowstone to determine if the intervention was effective

Conclusion

Biodiversity and species vulnerability

By analyzing the biodiversity data provided by the National Park Service, we have determined that certain categories of animals are more likely to be at risk than others.

Foot-and-mouth disease intervention

Using the known rate of foot-and-mouth disease from Bryce National Park last year we determined that it will require 510 sheep observations to establish the efficacy of the foot-and-mouth intervention. This will require one week of observation at Yellowstone National Park and two at Bryce.