

The background features abstract, overlapping green geometric shapes in various shades, creating a modern and dynamic feel. The shapes are primarily triangles and polygons, some solid and some semi-transparent, arranged in a way that suggests movement and depth. The colors range from light lime green to deep forest green.

# Biodiversity for the National Parks

Codecademy Capstone Project Option #2

Calla Glavin

# Overview

- ▶ Description of National Parks' Biodiversity Data
- ▶ Exploration of Species' endangered status
- ▶ Recommendations for conservationists
- ▶ Sample size determination for Foot & Mouth Disease study

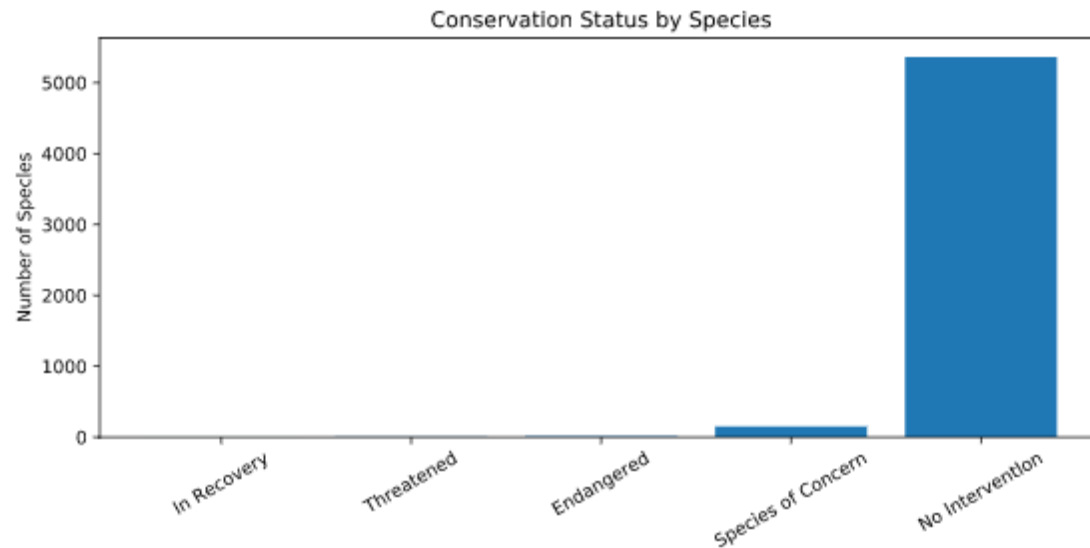
# Basic Description: National Parks' Biodiversity Data

- ▶ Data set includes:
  - ▶ 5541 unique species
  - ▶ 7 categories of species: mammal, bird, reptile, amphibian, fish, vascular plant, nonvascular plant
  - ▶ 5 categories of endangerment: No intervention (no endangered status), species of concern, endangered, threatened, in recovery
- ▶ **Takeaway: The majority of species have a “no intervention” conservation status; that is, they are not endangered, nearing endangerment or in recovery.**

| Conservation Status | Count |
|---------------------|-------|
| Endangered          | 15    |
| In Recovery         | 4     |
| Species of Concern  | 151   |
| Threatened          | 10    |
| No Intervention     | 5363  |

# Key Takeaway #1 from Data Overview

- Takeaway #1: The majority of species have a “no intervention” conservation status; that is, they are not endangered, nearing endangerment or in recovery.



| Conservation Status | Count |
|---------------------|-------|
| Endangered          | 15    |
| In Recovery         | 4     |
| Species of Concern  | 151   |
| Threatened          | 10    |
| No Intervention     | 5363  |

# Key Takeaway #2 from Data Overview

- **Takeaway #2:** From a descriptive statistic standpoint, the majority of species within different categories are not protected; however, without significance testing, we do not know if there are significant differences between the percentage of protected species between different categories.

| Category          | Not protected | Protected | Percentage Protected |
|-------------------|---------------|-----------|----------------------|
| Amphibian         | 72            | 7         | 8.86%                |
| Bird              | 413           | 75        | 15.36%               |
| 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%                |

# Testing Significance of Differences in Protected Status Between Species

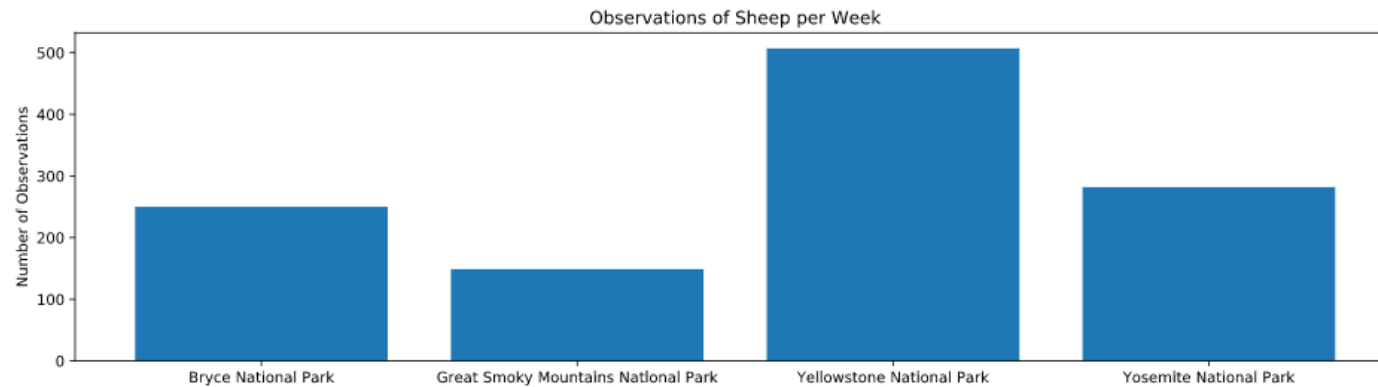
- ▶ There are significant differences between the protected status of mammals and reptiles (p-value < 0.05, chi-square test); but not between mammals and birds (p-value > 0.05; chi-square test).
- ▶ Thus, certain species have a significantly higher protected percentage than do others.
- ▶ This could suggest one of two possibilities:
  - ▶ Some species are in greater endangerment due to their habitats, or food sources (birds and mammals share similar habitats, and have similar needs, for example).
  - ▶ Some species may receive greater attention due to their size, visibility, or similarity to mankind (e.g. more mammals may be on the endangered list because they are easier to identify due to size, and easier to sympathize with than non-vascular plants, for example).

# Recommendations to Conservationists

- ▶ Due to the significant differences between the protected status of species, and potential explanations of these differences it is important to address the following lines of inquiry:
  - ▶ Explore the reasons **WHY** species are endangered
  - ▶ Investigate whether smaller, less visible species are endangered
  - ▶ While significance tests can describe the degree to which percentage protected status are different between categories of species, **these do not describe how species are related: thus, investigate how endangered species impact the vitality of other species**

# Sheep Observations at National Parks

- Scientists wanted to investigate how many sheep were observed per week at each national park: Yellowstone National Park had the most observations of sheep per week





# Foot & Mouth Disease in Sheep Populations

- ▶ Scientists running a program to reduce Foot & Mouth Disease in sheep populations want to be able to deduce reductions of **at least 5 percentage points**
  - ▶ Baseline percentage: 15% (of sheep at Bryce National Park had Foot & Mouth Disease)
  - ▶ Minimum detectable effect: 33.33% (5% reduction / 15% baseline percentage)
  - ▶ Using a sample size calculator, at a 90% significance rate: 870 observations are needed to discern this minimum detectable effect
  - ▶ Based on the observations to date:
    - ▶ Scientists will need to observe sheep at Yellowstone Park for 3.5 weeks to observe this minimum detectable effect (if it exists)
    - ▶ Scientists will need to observe sheep at Bryce Park for at least 1.7 weeks to observe this minimum detectable effect (if it exists)

# Conclusions

- ▶ Simple data can be interpreted in a variety of ways
- ▶ Simple significance calculations can lead to thought-provoking questions but cannot answer these questions alone
- ▶ In order to investigate certain questions, sample size determinations provide tangible recommendations for policymakers to ascertain significant differences