codecademy

Capstone: Biodiversity for National Parks

Intro to Data Analysis

Gregory Caldwell

01/02/2019 Cohort

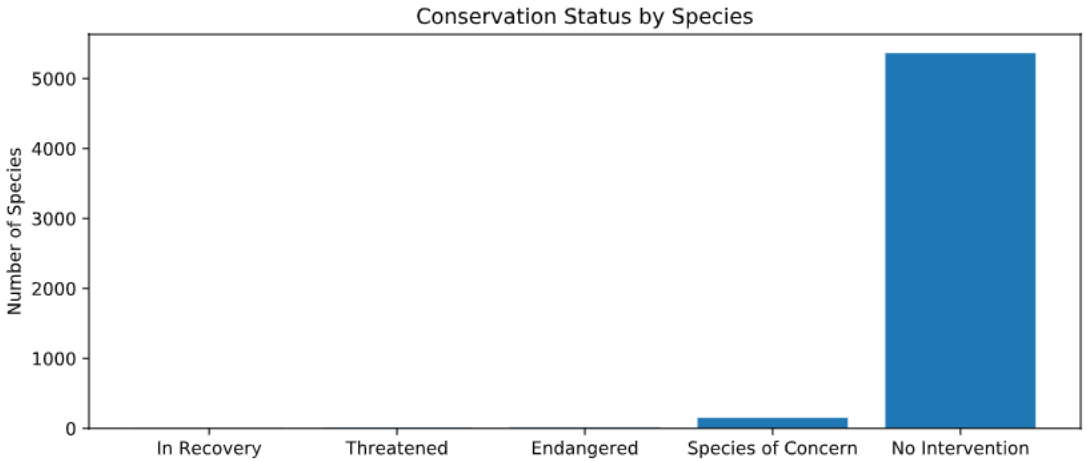
Table of Contents

Species Info Overview

Describing Species Info Data – Overview

- **Data Observed:** Category, Scientific Name, Common Names, Conservation Status
- **Species Count:** 5,541
- **Types of Animals:** Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, Nonvascular Plant
- **Conservation Statuses:** No Intervention, Species of Concern, Endangered, Threatened, In Recovery

Describing Species Info Data – Conservation Status of Parks



Conservation Status	Scientific Name
Endangered	15
In Recovery	4
No Intervention	5363
Species of Concern	151
Threatened	10

- 96.75% of species are listed as "No Intervention", with 0.27% Endangered

Endangered Status - Species Categories

(Significance Calculations)

Endangered Species Status – Objective

For the national parks we work with, are certain species more likely to be endangered than others?

Endangered Species – Overview

Category	Not Protected	Protected	Percent Protected
Amphibian	73	7	9%
Bird	442	79	15%
Fish	116	11	9%
Mammal	176	38	18%
Nonvascular Plant	328	5	2%
Reptile	74	5	6%
Vascular Plant	4424	46	1%

- **Most Protected Species by %:** Mammals, followed by Birds
- **Most Protected Species by #:** Birds, followed by Vascular Plants

Endangered Species – Category Comparison & Significance

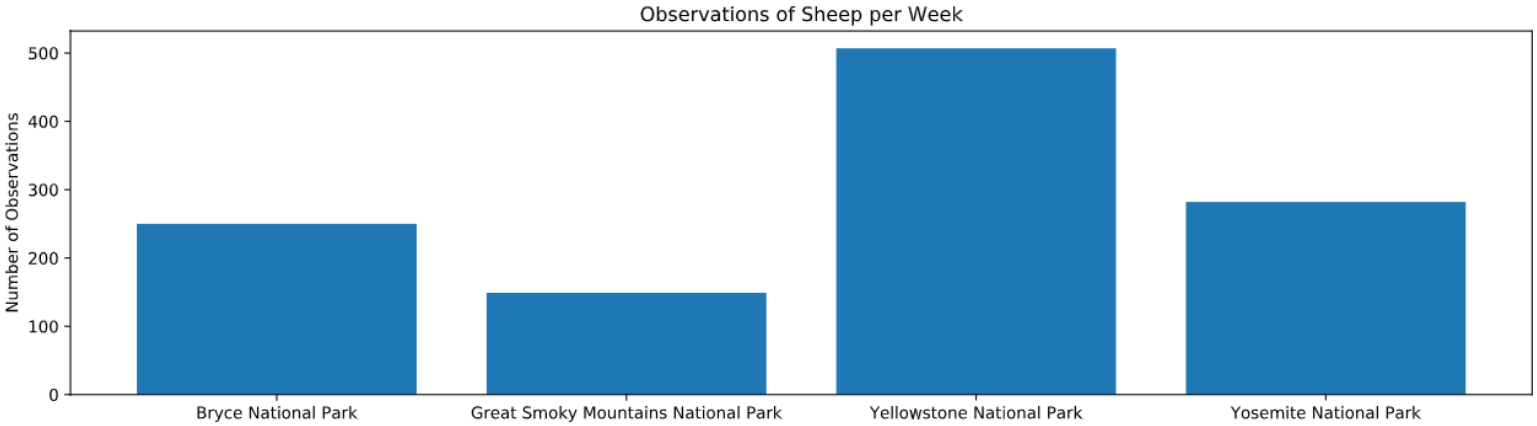
- Mammals **are not** more likely to be endangered than Birds
  - The difference in percent protected **is** due to chance
  - P-value = 0.0688, > 0.05 (null hypothesis is accepted)
- Mammals **are** more likely to be endangered than Reptiles
  - The difference in percent protected **is not** due to chance
  - P-value of 0.0384, < 0.05 (null hypothesis is rejected)

- Conclusion: Certain types of species **are** more likely to be endangered than others
  - Chi Squared test used to see if certain categories were more likely to be endangered than others
- Conservationist recommendations
  - Target conservation efforts towards those species that are more likely to need protection than others (e.g. mammals)

Foot & Mouth Disease – Sheep

(Sample Size Determination)

Sheep – Observations by Park



- **Data Observed:** Scientific Name, Park Name, Observations

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

Corresponds to 10 – 13 of 15

Foot and Mouth Disease - Sheep

- **Goal:** Establish the success of the program to reduce Foot and Mouth Disease in Sheep at various parks over the past year

- **Test outline:**
  - Baseline: 15% (last year, 15% had Foot and Mouth Disease)
  - Minimum detectable effect: 33% (i.e. detect reductions of at least 5 percentage points)
  - Level of significance: 90%

- **Survey Data Needed:**
  - Observations needed per park: 890
  - Weeks observing:
    - Yellowstone: 1.76
    - Bryce: 3.56

Thank You