Biodiversity in the National Parks

Codecademy Pro: Introduction to Data Analysis

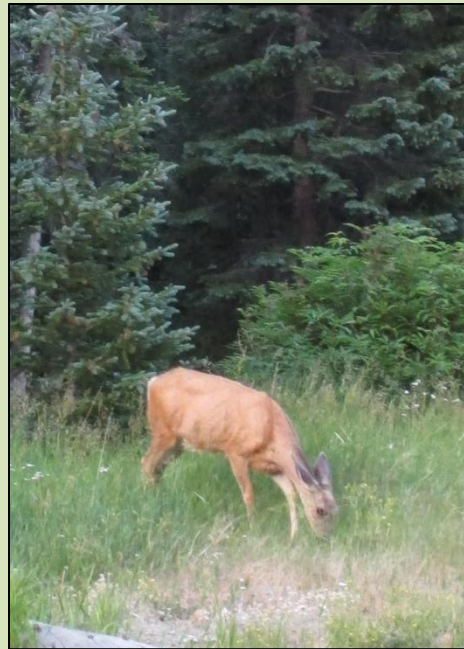Ashley Gingeleski

Species Data

- Category
  - Most species:

    Vascular plants

- Scientific Name

- Common Name

- Conservation Status
  - Most common protected status: Species of Concern

  - Highest percent protected: Mammals

Significance Calculations

- Chi-squared Test
  - 2 categories with 2 values (Protected & Not Protected)
- Chi-squared statistical significance
  - Based on p-values < 0.05
- Mammal vs. Bird
  - p-value: ~0.69
  - No significance
- Reptile vs. Mammal
  - p-value: ~0.03
  - Significance

Recommended Steps for Conservationists

- Resources for protecting mammals should take priority over those for reptiles. Mammals are more likely to be endangered.

- Resources should be evenly allocated for mammals and birds. They have comparable likelihoods for being endangered.
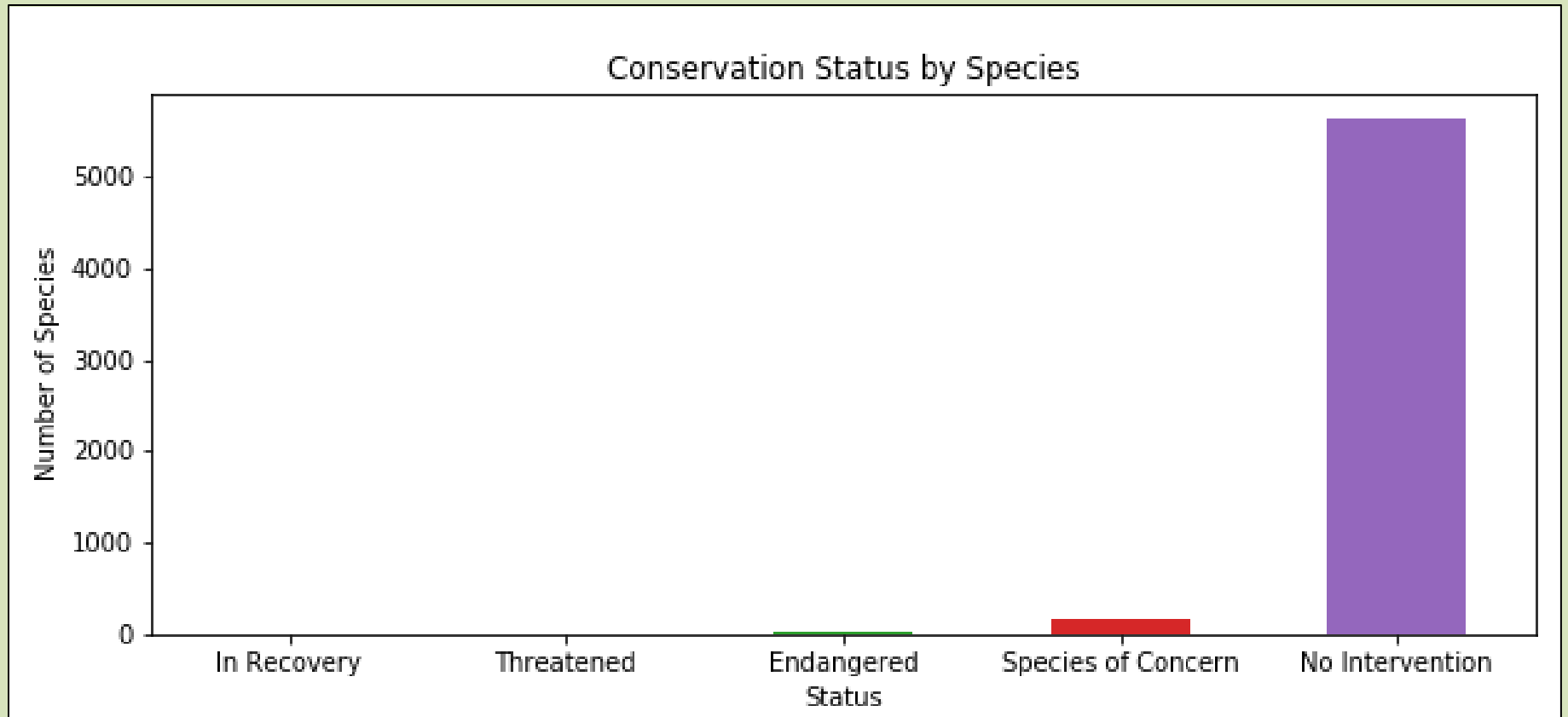
Sample Size Determination

- Baseline Conversion Rate: **15%**
  - Percent of sheep in Bryce Canyon with the disease
- Minimum Detectable Effect: **5%**
  - Conservationists want to be able to detect at least 5% with confidence

- Statistical Significance: **90%**
  - Default level of significance
- Sample Size: **35,000**

Weeks in Bryce to meet sample size: **140**

Weeks in Yellowstone to meet sample size: **70**

Conservation Status of Species

Conservation Status by Species

Weekly Sheep Observations

Observations of Sheep per Week