



BIODIVERSITY CAPSTONE PROJECT

Fanny Ummelen

19-11-2019

Introduction

The (fictional) situation in this capstone project is the following:

I am a biodiversity analyst working for the National Parks Service. I am going to help them analyze some data about species at various national parks.

The fictional dataset contains entries of 5541 species. Each of them is assigned a category (mammal, bird, reptile, amphibian, fish, vascular plant or nonvascular plant) and a conservation status (species of concern, endangered, threatened, in recovery or no intervention).



Table of Contents

- Exploring the complete data set
 - *How many species are endangered?*
 - *Are certain types of species more likely to be endangered?*
- Feasibility of disease prevention study
 - *How many total sheep observations (across all three species) were made at each national park?*
 - *How many weeks would you need to observe sheep at each Park in order to observe enough sheep to determine the effectiveness of disease prevention?*

EXPLORING THE COMPLETE DATA SET



How many species are endangered?

- Only a small percentage (3.45%) of all species are categorized as needing some sort of protection.
 - *In recovery:* 0.07%
 - *Threatened:* 0.18%
 - *Endangered:* 0.29%
 - *Species of Concern:* 2.91%



Are certain types of species more likely to be endangered?

- For each species category the number of protected species and how many percent this is of the whole category are listed
- Using a χ^2 test we determined that:
 - *Mammals are not significantly more often protected than birds (**p value = 0.69**)*
 - *Mammals are significantly more often protected than reptiles (**p value = 0.04**)*

Category	# protected	% protected
Amphibian	7	8.9
Bird	75	15.4
Fish	11	8.7
Mammal	30	17.0
Nonvascular plant	5	1.5
Reptile	5	6.4
Vascular plant	46	1.1

FEASIBILITY OF DISEASE PREVENTION STUDY



Background

Our scientists know that 15% of sheep at Bryce National Park have foot and mouth disease. Park rangers at Yellowstone National Park have been running a program to reduce the rate of foot and mouth disease at that park. The scientists want to test whether or not this program is working. They want to be able to detect reductions of at least 5 percentage points.

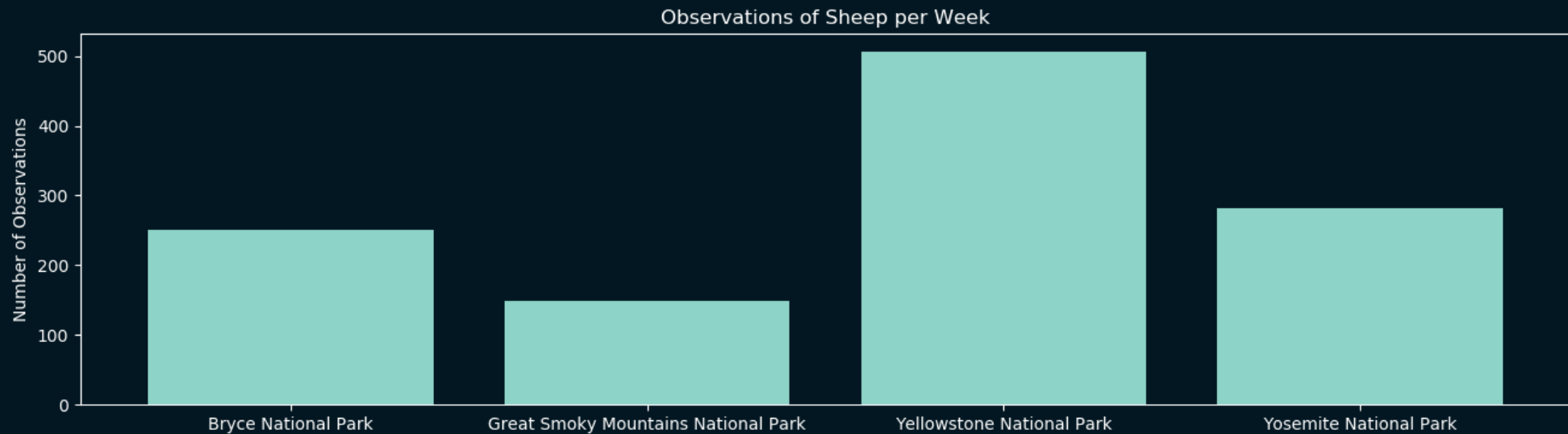
There are 3 species of sheep present in the national parks, each with a different conservation status.



Scientific Name	Conservation Status
Ovis aries	No Intervention
Ovis canadensis	Species of Concern
Ovis canadensis sierrae	Endangered

How many total sheep observations were made at each national park?

- The total number of sheep observations (all three species included) in one week is shown for each national park.



Feasibility of disease prevention study

Based on the requirements listed in the table, the minimum sample size is **610** observations.

The number of weeks of observations required to reach this sample size can be calculated for each national park, based on the number of sheep observations per week listed on the previous slide.

Base conversion rate	15%
Statistical significance	90%
Minimum detectable effect	50%

National Park	# weeks of observation
Bryce	3
Yellowstone	2

Conclusions:

- Only **3.45%** of all species in the national parks are categorized as needing some sort of protection.
- Mammals are significantly more often protected than reptiles (**p value = 0.04**)
- The difference in how often mammals or birds are protected is not significant (**p value = 0.69**)
- A minimum sample size of **610** sheep observations is required to conduct the disease prevention study
 - *This would amount to **3 weeks** of observations at Bryce National Park*
 - *This would amount to **2 weeks** of observations at Yellowstone National Park*