

The background of the slide is a dense, close-up photograph of green leaves, likely from a tree or shrub, filling the entire frame. The leaves are various shades of green, from bright to dark, and are arranged in a way that creates a textured, natural background.

Biodiversity Analysis for the National Parks

Biodiversity Project of Codecademy

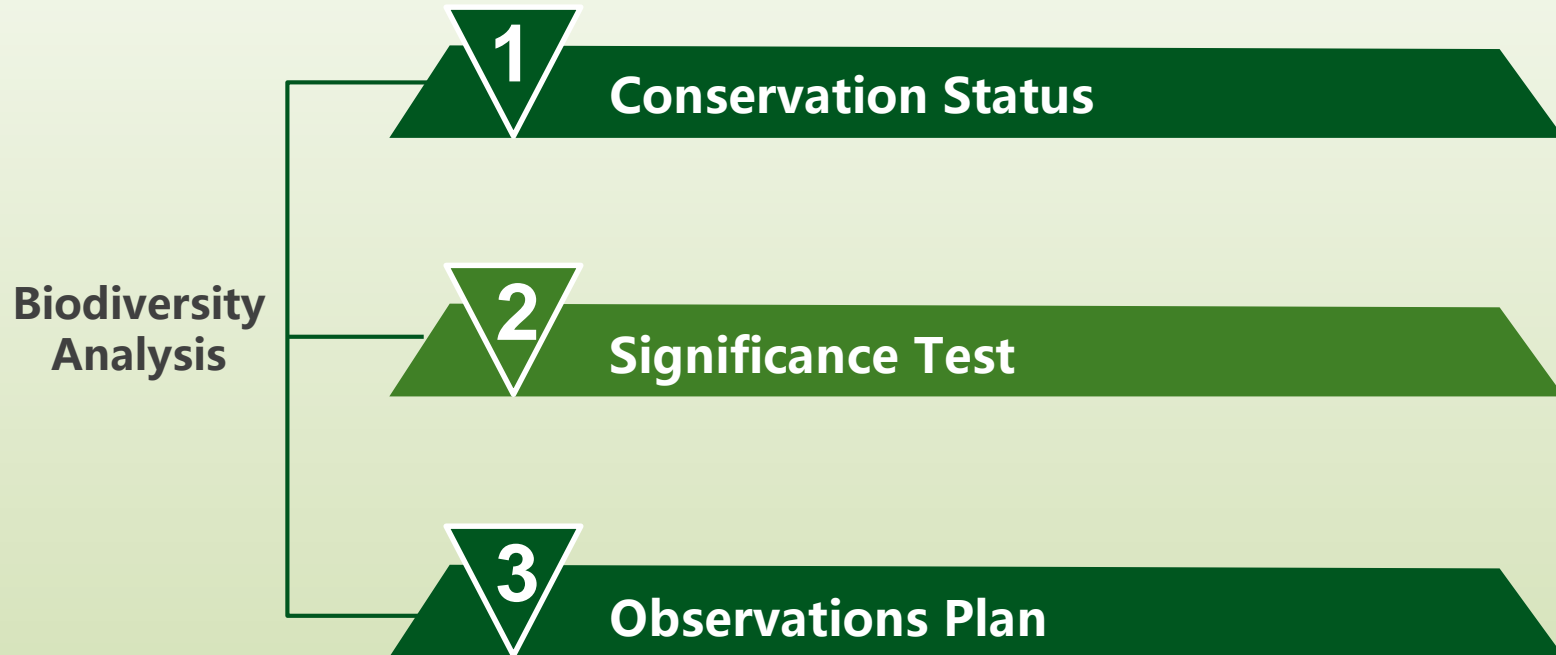
Yezhao Huai
yhuai001@fiu.edu



This a biodiversity analysis from the National Parks Service.

We focus on analyzing data about species at various national parks.

Contents



Conservation Status

Key Indicators

- Conservation Status
- Number of species

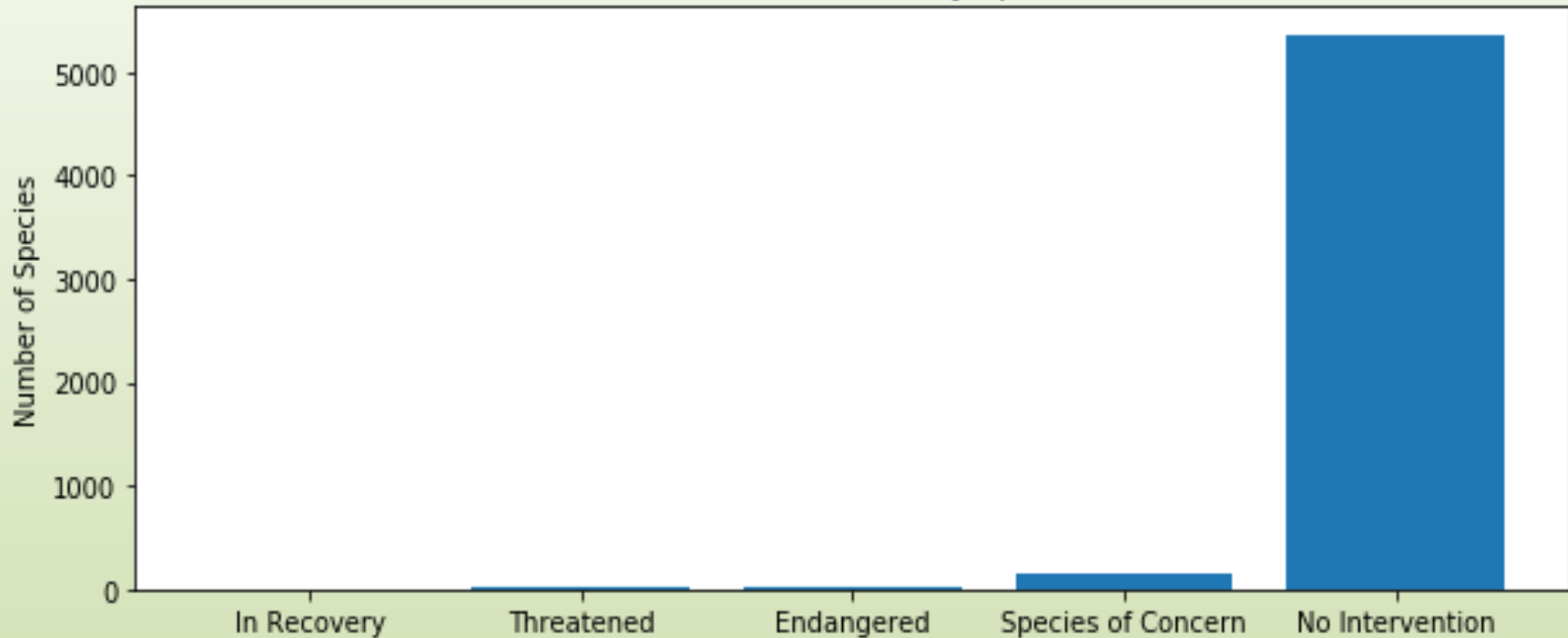
Conservation status	Number of species
In Recovery	4
Threatened	10
Endangered	15
Species of Concern	151
No Intervention	5363

The initial data include the scientific name, common names, and conservation status of each species. There are 5541 different species, 7 categories, and 4 conservation status.

We add “No Intervention” conservation status, and then get the number of species for each status by counting unique scientific name.

Conservation Status

Conservation Status by Species



Species of Concern: Declining or appear to be in need of conservation

Threatened: Vulnerable to endangerment in the near future

Endangered: Seriously at risk of extinction

In Recovery: Formerly Endangered, but currently neither in danger of extinction throughout all or a significant portion of its range

No Intervention: Have no relevant evaluation

Significance Test

We already know that most of the species in the National Parks have no intervention. Now we are going to test the significant difference between the species from different categories.

First of all, we define the protected state based on the conservation status and calculate the protected percent for each category.

Category	Not protected	Protected	Percent of protected
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015
Reptile	73	5	0.064103
Vascular Plant	4216	46	0.010793

Significance Test

Then, we create contingency table prepared for Chi Square Significant Test

Category	Protected	Not protected
Category1	num1	num2
Category2	num3	num4

If we get P value less than 0.05, it is significant. For example, the P value of test of the difference between *Reptile* and *Mammal* significant is $0.0383556 < 0.05$, so it looks like there is a significant difference between *Reptile* and *Mammal*!

Observations Plan

Background

Conservationists have been recording sightings of different species at several national parks for the past 7 days. Some scientists are studying the number of sheep sightings at different national parks. There are several different scientific names for different types of sheep.

Condition

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.

Goal

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 point. We also want to know how many weeks would you need to observe sheep.

Observations Plan

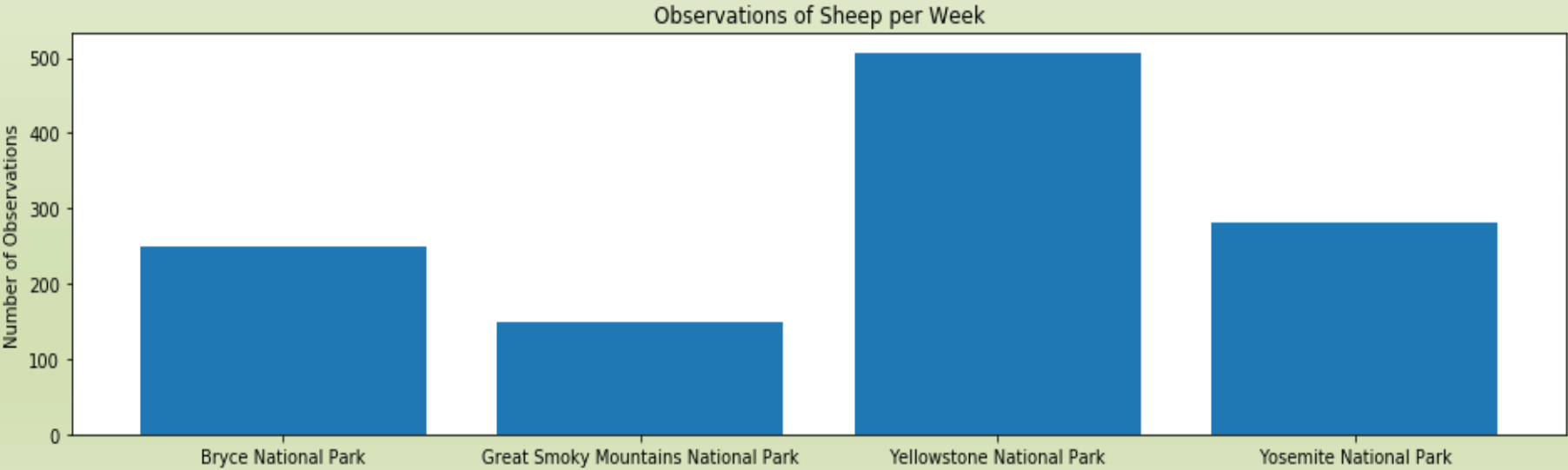
At first, we need to get all the “Sheep” occur in common names and belongs to Mammal category merge with observations data. After that, we can group observations of sheep per week by different parks.

Scientific Name	Park Name	Observations	Category	Common Names	Conservation Status	Protected
Ovis canadensis	Yellowstone National Park	219	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True
Ovis canadensis	Bryce National Park	109	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True
Ovis canadensis	Yosemite National Park	117	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True
Ovis canadensis	Great Smoky Mountains National Park	48	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True

⋮

Observations Plan

Park Name	Observations
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite National Park	282



Observations Plan Conclusion

Finally, we calculate the sample based on baseline, minimum detectable effect, and level of significance, and we get sample size per variant is 510.

The result means we need 2 weeks to observe sheep at Bryce National Park and 1 week at Yellowstone National Park in order to observe enough sheep.

Thank You!