

# Biodiversity for the National Parks

Capstone Project

Introduction to Data Analysis Intensive Program

E. Reusi

# ENDANGERED SPECIES

# First things first...

# 5541

Categories of species

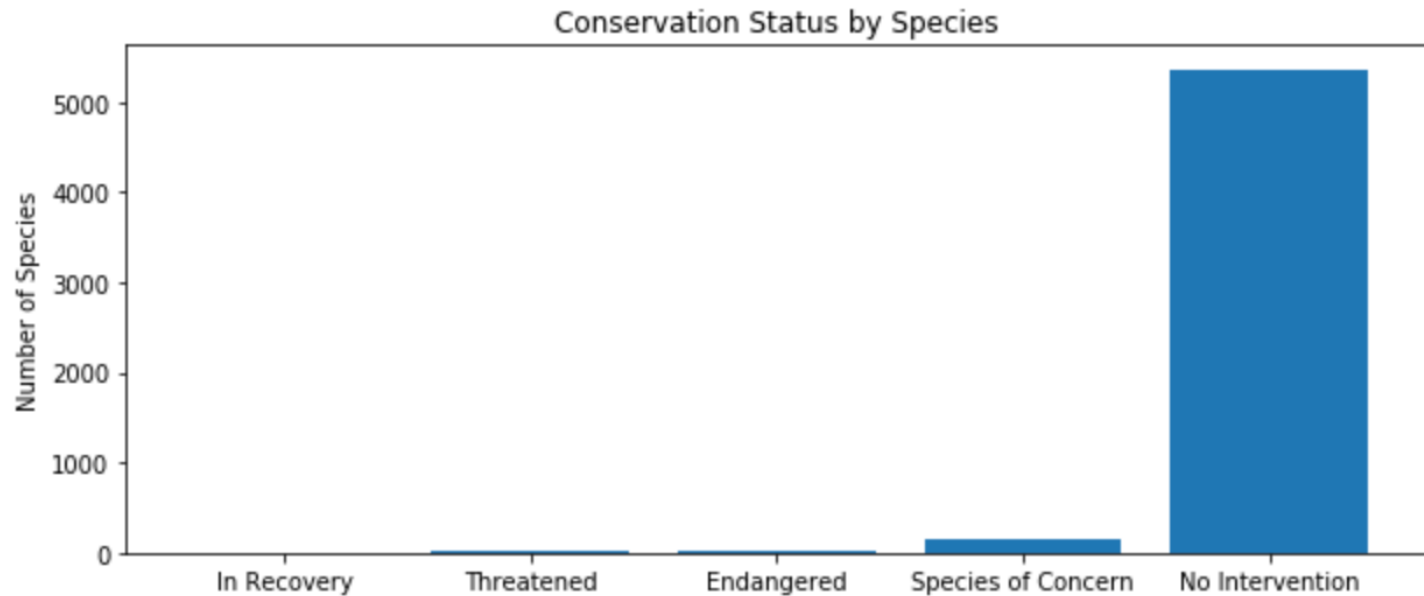
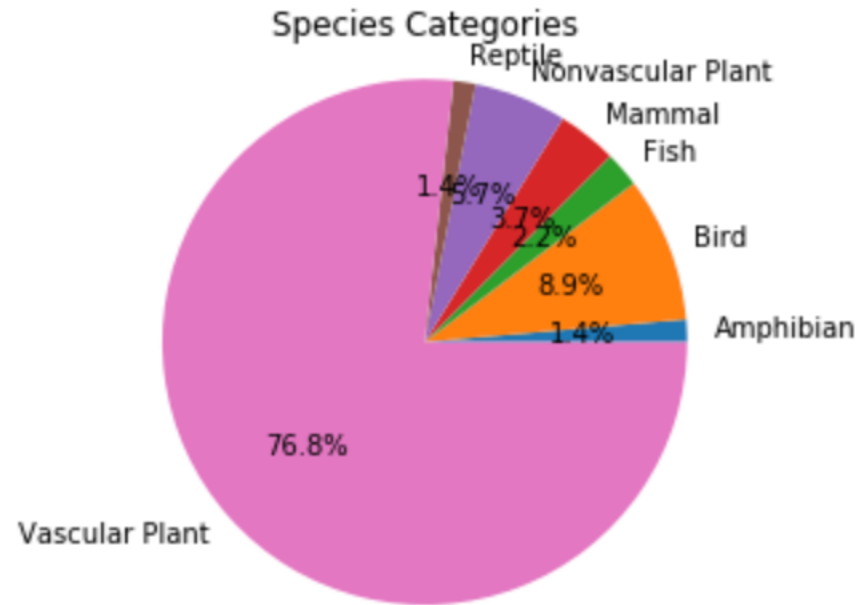
**ENDANGERED**  
**IN RECOVERY**  
**NO INTERVENTION**  
**SPECIES OF CONCERN**  
**THREATENED**

Different types of species

**AMPHIBIAN**  
**BIRD**  
**FISH**  
**MAMMAL**  
**NONVASCULAR PLANT**  
**REPTILE**  
**VASCULAR PLANT**

Conservation statuses

# Categorization



> 76% of species in the park are vascular plants

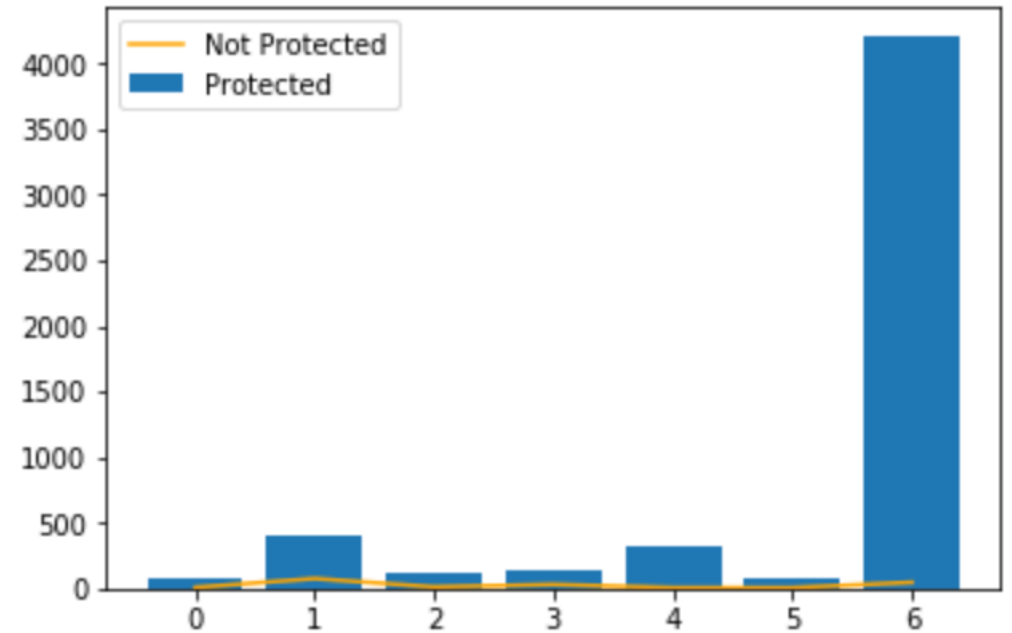
> 96% of species in the park requires no intervention

# Protection by species

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

	category	not_protected	protected	percent_protected
3	Mammal	146	30	0.170455
1	Bird	413	75	0.153689
0	Amphibian	72	7	0.088608
2	Fish	115	11	0.087302
5	Reptile	73	5	0.064103
4	Nonvascular Plant	328	5	0.015015
6	Vascular Plant	4216	46	0.010793

\*Data sorted by percent\_protected



< 20% of each species are in need of protection

Relative to the number of species in each category, mammals have the highest protection percentage (most endangered), while vascular plants have the lowest.

# Are certain types of species more likely to be endangered?

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

## Null hypothesis:

The difference in percentages is a result of chance.

If p-value > 0.05: accept null hypothesis, not significant, chance

If p-value < 0.05: reject null hypothesis, significant, not chance

## Tested P-values:

Mammal vs Bird = 0.68759480966613362

Mammal vs Reptile = 0.038355590229698977

Bird vs Nonvascular Plant = 1.0546306904975004e-10

Amphibian vs Fish = 0.8247942981524834

All = 5.5108280473137505e-89

Majority of the tested p-values generated by performing the chi-squared test are greater than 0.05.

The difference between the protection percentage of most species is not significant and is a result of chance.

Certain types of species are more likely to be endangered than others.

# Recommendation

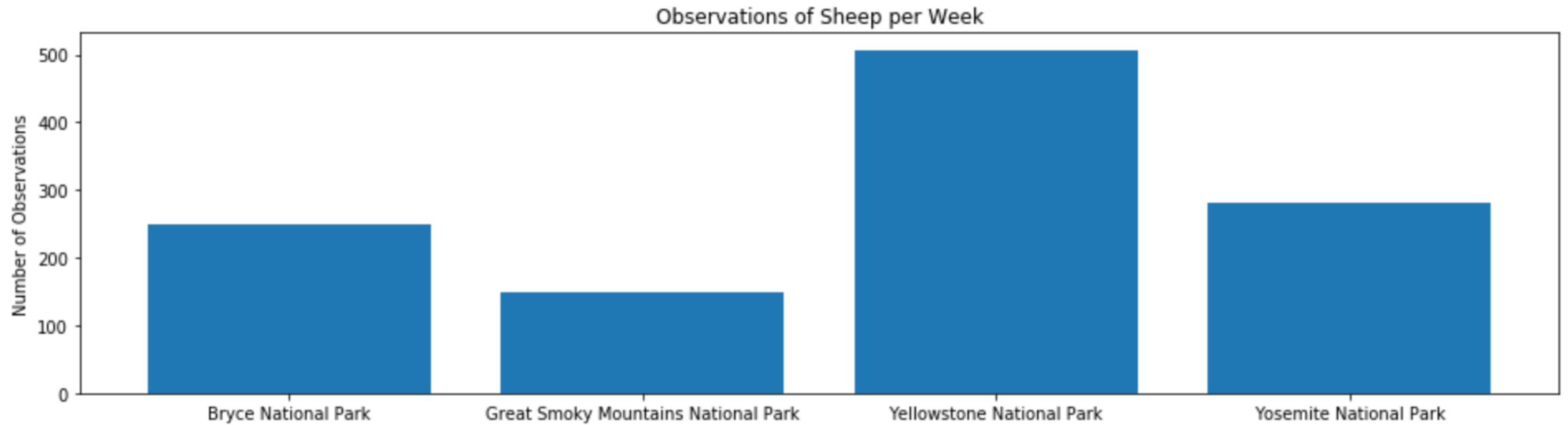
Although we found that the difference between the protection percentage of most species is not significant and is a result of chance, there are still certain types of species that are more likely to be endangered than others.

It is therefore recommended to give the protection of these species higher priority and monitor their conservation status more closely.

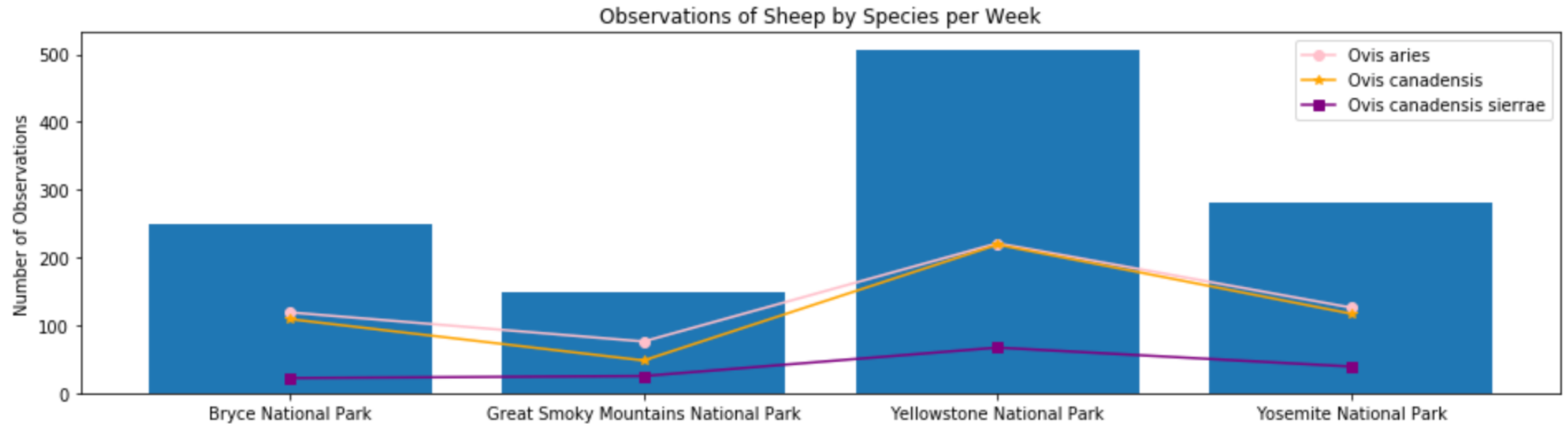
# FOOT & MOUTH DISEASE STUDY



# Observations of Sheep per Week



# Observations of Sheep by Species per Week



# Sample Size Determination

Given a 15% occurrence of foot and mouth disease at Bryce National Park, park rangers at Yellowstone National Park will need to observe 510 sheep for approximately 1 week to make sure that there will be at least a 5% drop in observed cases.

Park rangers running the same program at Bryce National Park will have to observe the same number of sheep for about 2 weeks.

Baseline Conversion Rate

15 %

Your control group's expected conversion rate. [\[?\]](#)

Minimum Detectable Effect

33.3 %

The minimum relative change in conversion rate you would like to be able to detect. [\[?\]](#)

Statistical Significance

90%

95% is an accepted standard for statistical significance, although Optimizely allows you to set your own threshold for significance based on your risk tolerance. [\[?\]](#)

[EDIT](#)

Sample Size per Variation

510