Introduction to Data Analysis

Capstone Option 2:

Biodiversity for the National Parks

Data Description in species_info.csv

	category	scientific_name	common_names	conservation_status
0	Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole	nan
1	Mammal	Bos bison	American Bison, Bison	nan
2	Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Domesticated Cattle	nan
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	nan
4	Mammal	Cervus elaphus	Wapiti Or Elk	nan

- 5 different species in category Mammal
- Values of conservation_status with species: nan (replaced by No Intervention =5363), Species of Concern (151), Endangered (15), Threatened (10), In Recovery (4)
- See Visual of this Data on slide 6

Significance calculations for endangered status between different categories of 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

- Mammals are more likely to be endangered than Birds
- Based on categorical data, a chi-squared test is appropriate
- But the difference between the birds and mammals is not significant (p-value = 0.68) and is therefore a result of chance
- On the other side the significance between reptiles and mammals is significant, because p-value is 0.038

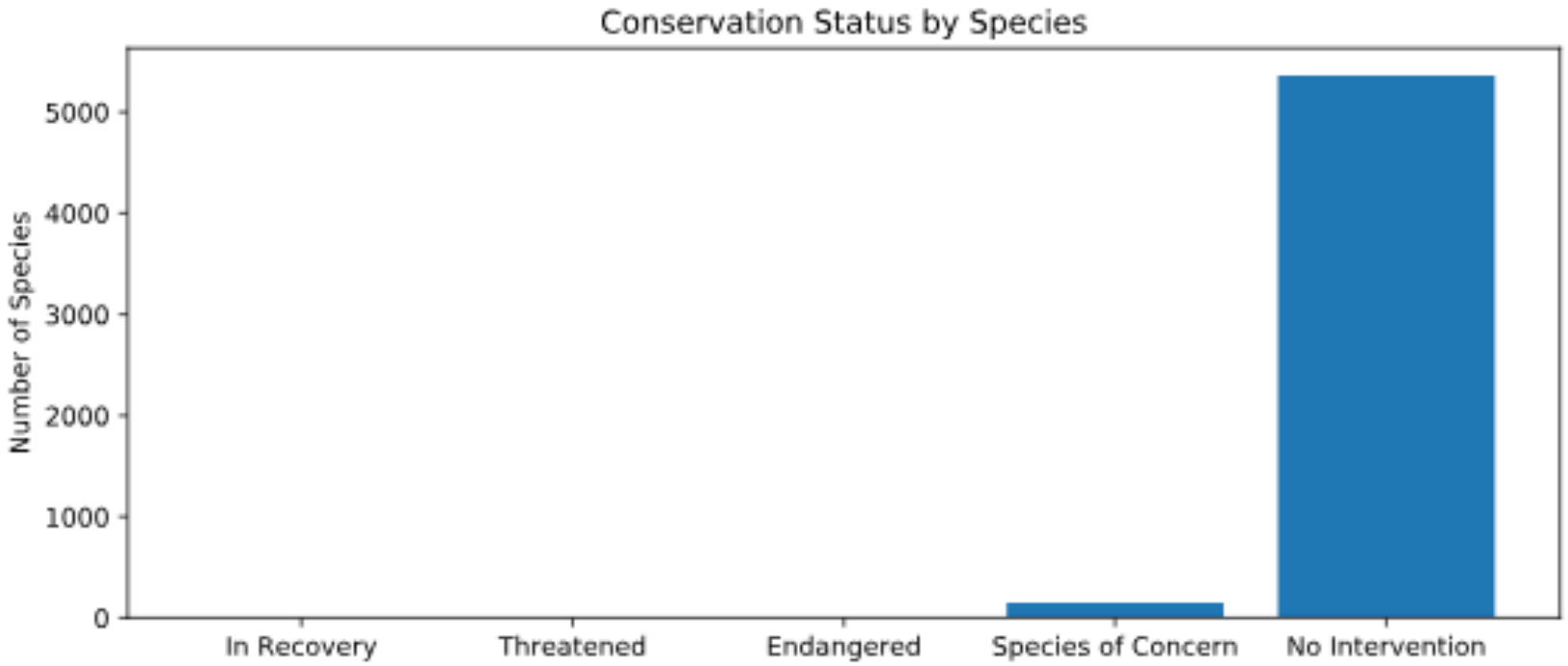
Recommendation for conservationists based on significance calculations

- Certain types of species are more likely to be endangered than others, e.g. Mammals

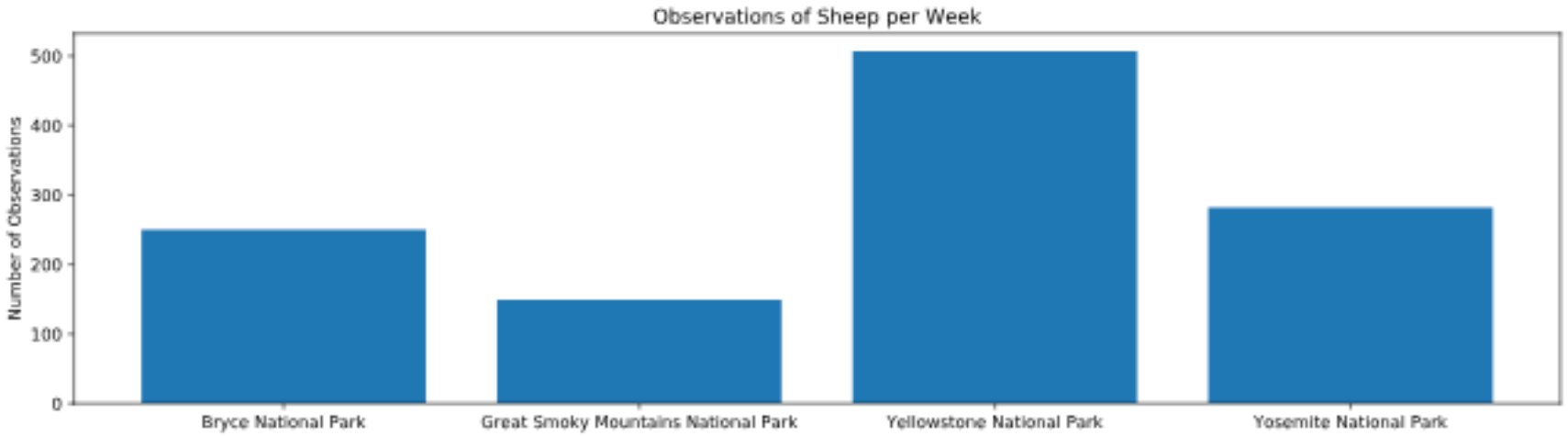
Description of sample size determination for the foot and mouth disease study

- Baseline conversion rate is 15% with statistical significance of 90% and minimum detectable effect of 33%
- So Sample size would be 890

All graphs:


Conservation Status by Species
Number of Species
5000
4000
3000
2000
1000
0
In Recovery
Threatened
Endangered
Species of Concern
No Intervention

All graphs:


Observations of Sheep per Week
Number of Observations
0
100
200
300
400
500
Bryce National Park
Great Smoky Mountains National Park
Yellowstone National Park
Yosemite National Park