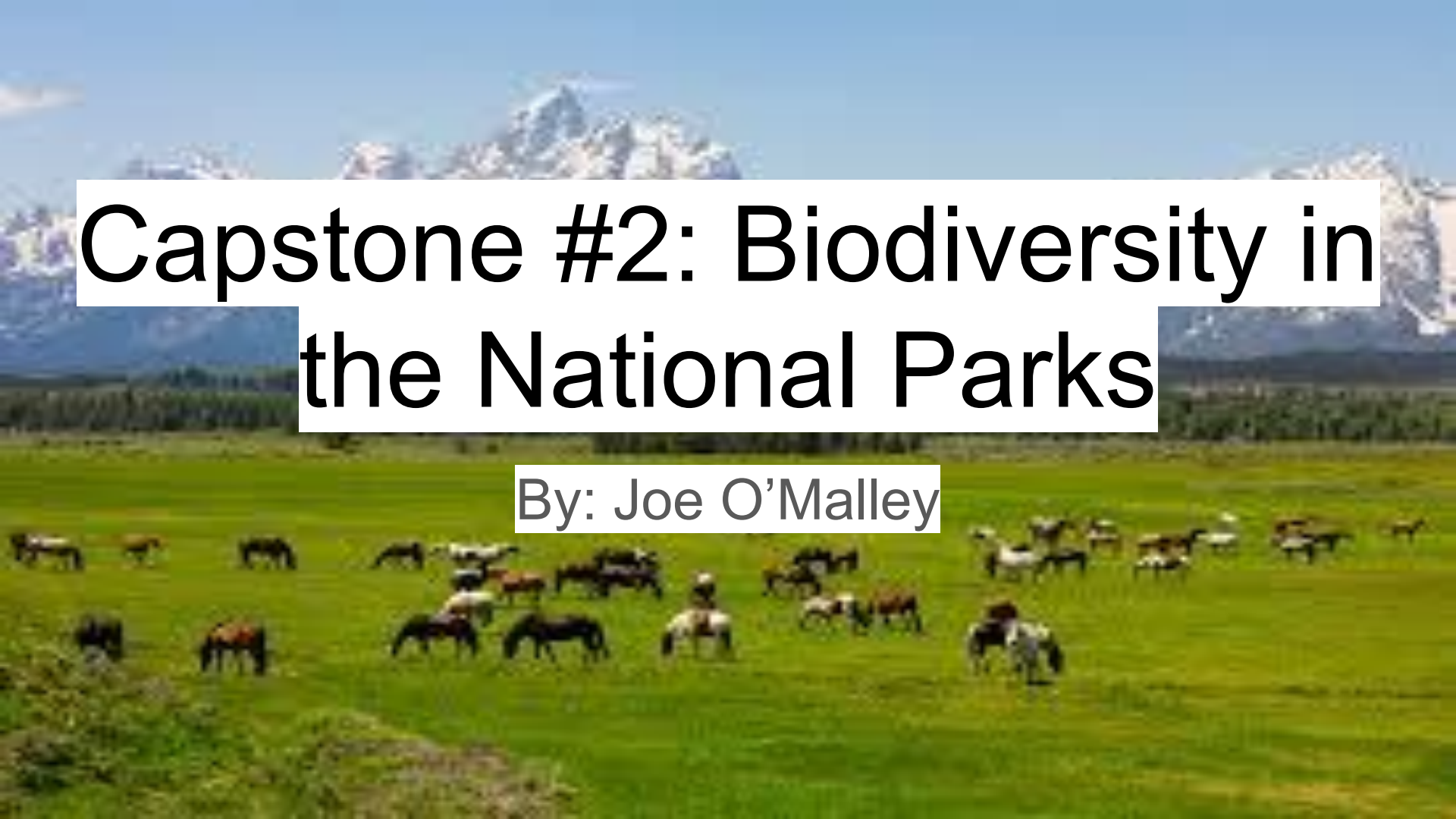


Capstone #2: Biodiversity in the National Parks

By: Joe O’Malley

Assignment

To investigate if there are any patterns or themes to the types of species that become endangered on behalf on the National Park Service of the United States

Data

National Park District provided information on “species_info.csv”, which included:

- Category: includes mammals, fish, vascular plants, etc
- Scientific_Name: 100s of scientific names, some shown below
- Common_Names: layman names
- Conservation_Status: includes endangered, in recovery, species of concern, threatened, and nan (not endangered)

This data needed to be broken into animals that are endangered and animals that are not.

Appendix A shows the number of endangered species per endangered status

	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

Calculations & Results

- In appendix B, we see that 17% of Mammals and 15% of Birds are protected.
- Conversely, we see that 1% of Vascular plants and ~2% of non-vascular plants are protected
- We then performed a Chi-Squared test to see if the percentage difference in protected statuses between species was due to chance
  - Null-Hypthesis: There is no difference between species and the percentage breakdown we're seeing in protected status is due to chance
  - Result:
    - P-value between mammal and birds is 0.68, Do Not Reject Null Hypothesis
    - P-value between mammals and reptiles is 0.04, Reject Null Hypothesis, there is a significant difference between protected statuses for Mammals and Reptiles
- Conclusion: there are differences in endangered statuses between species not due to chance

Recommendation for conservationists

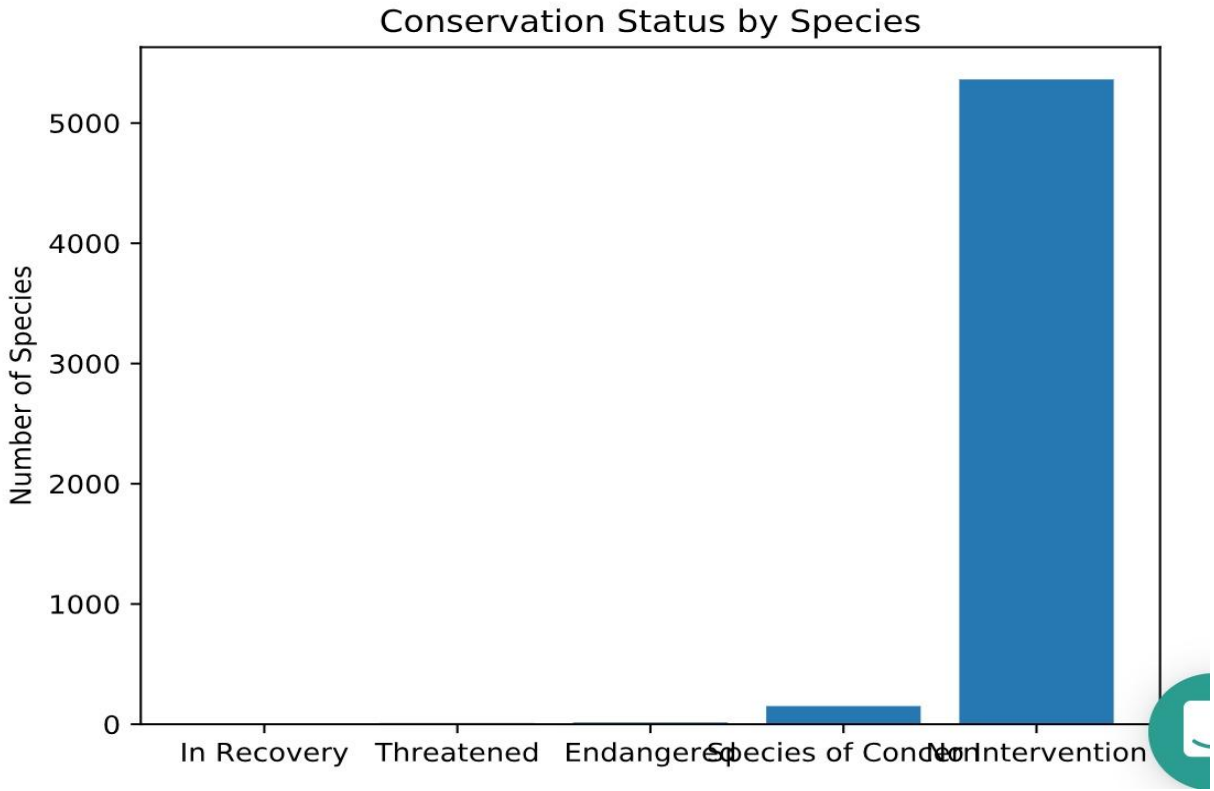
The Chi-Square test proved that certain species are more likely to be endangered, meaning conservation efforts for one species should be prioritized based relative to other species (i.e. mammals over reptiles).

Sample size determination

- Determined sample size for foot and mouth disease study. Here are the relevant figures:
  - baseline = 15
  - minimum_detectable_effect = (float(15-10)/15)* 100
  - sample_size_per_variant = 870
  - yellowstone_weeks_observing = float(sample_size_per_variant) / 507
  - bryce_weeks_observing = float(sample_size_per_variant)/250

See Appendix D for more details regarding sample_size calculation of 870

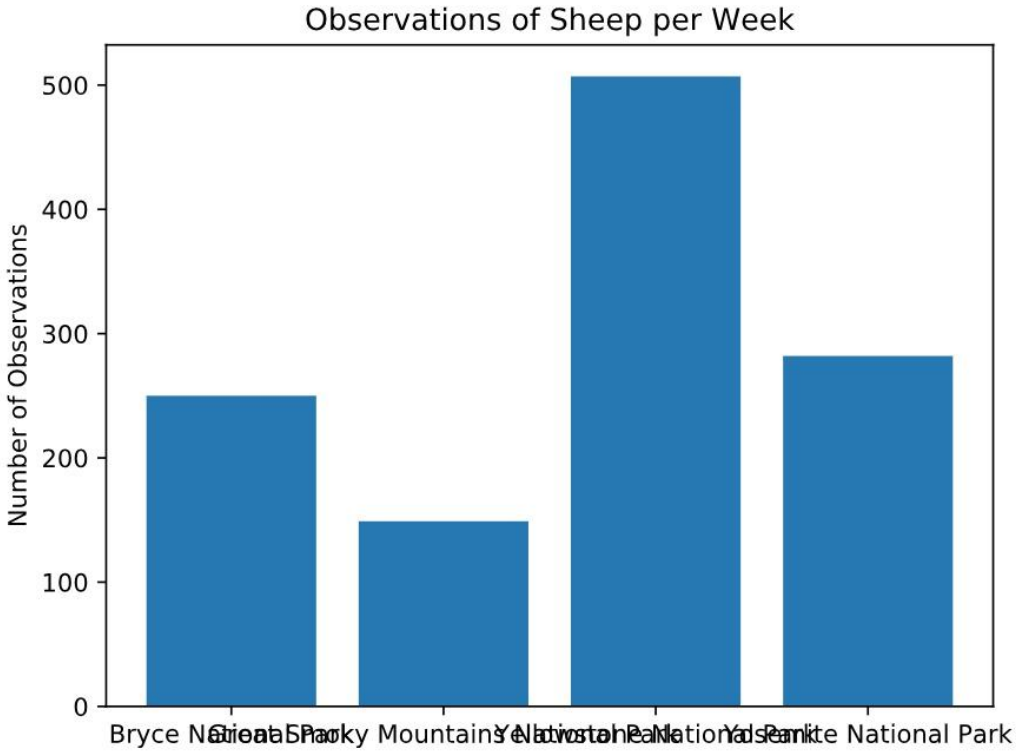
Appendix A


Conservation Status by Species
Number of Species
5000
4000
3000
2000
1000
0
In Recovery
Threatened
Endangered
Species of Concern
Non Intervention

Appendix B

	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

Appendix C


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

Sample Size Calculator