

BIODIVERSITY IN THE NATURAL PARKS



Gerard Barbalich



Dataset analysis

- The species_info.csv dataset contains **5,824 observations** of four variables.
- **Four variables** for each observation:
 - category
 - scientific name
 - common names
 - conservation status
- **Five classes of conservation status**
 - Endangered
 - in recovery
 - no intervention
 - species of concern
 - threatened
- **Seven classes of animal:**
 - Mammal
 - Bird
 - Reptile
 - Amphibian
 - Fish
 - Vascular plant
 - Non vascular plant



Dataset analysis

`print(species.head())`

| | 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 |



Analysis of endangered status 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 |

This table seems to show that Mammals are more likely to be protected than other species. A chi-square test for independence was run to determine this.

The chi-square test for independence found:

- **Mammals** are more likely to be endangered than **Birds**
 - P-value 0.445 Non-significant difference
- **Mammals** are more likely to be endangered than **Amphibians**
 - P-value 0.084 Significant difference
- **Mammals** are more likely to be endangered than **Reptiles**
 - P-value 0.023 Significant difference



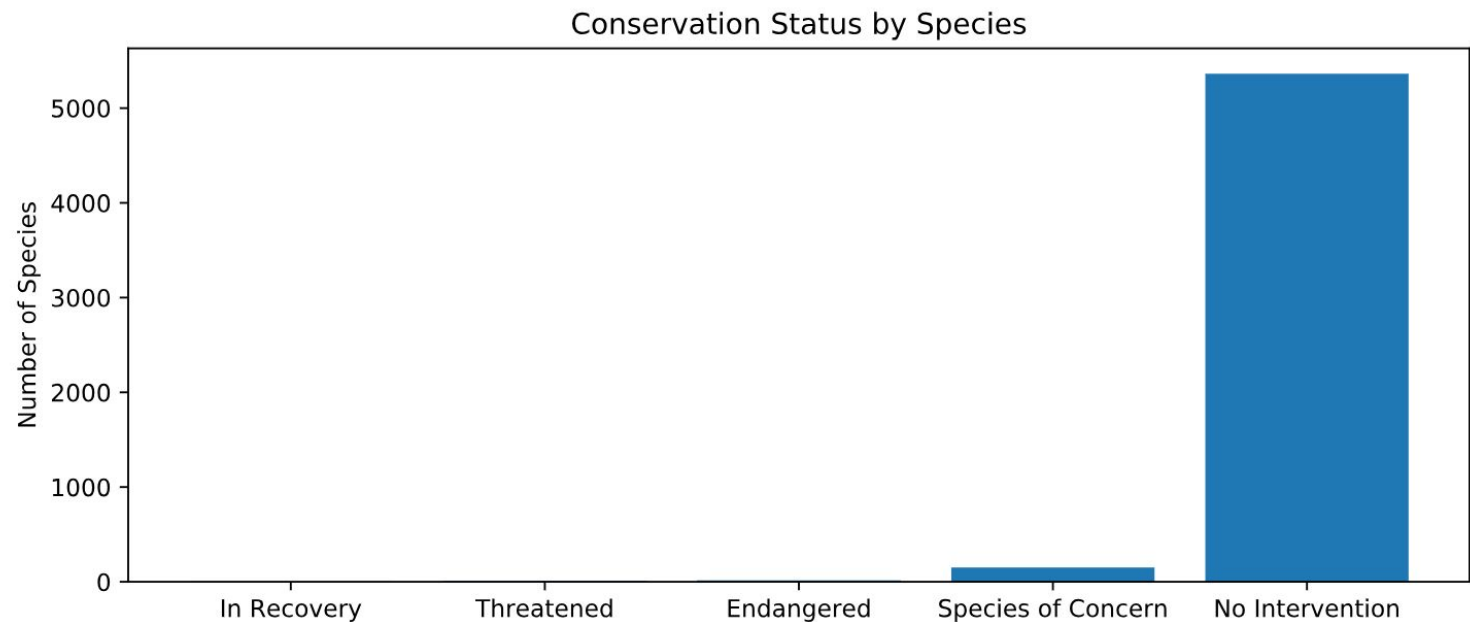
Recommendations based on analysis

Based on the significance hypothesis tests, the hypotheses that **mammals are significantly more likely to be endangered** than both Birds and Reptiles can be accepted.

Based on this analysis, more funding and priority should be given to the conservation of mammals.



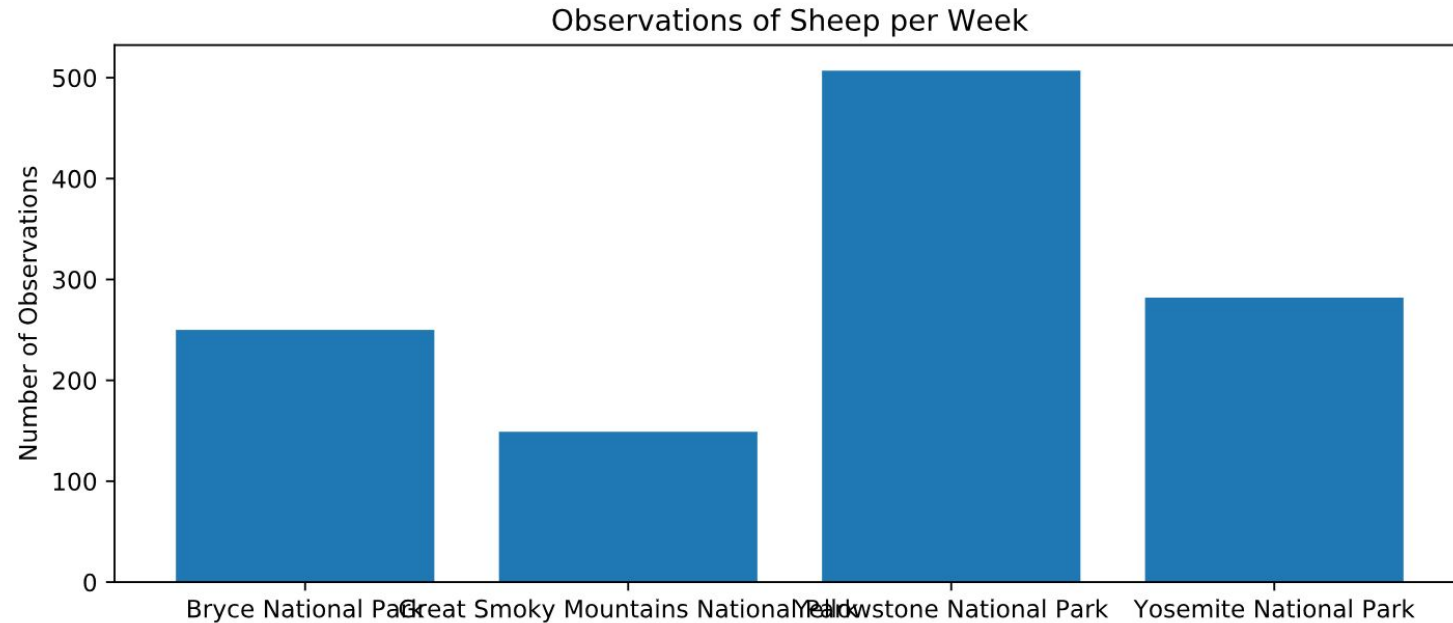
Conservation Status by Species





Foot and Mouth disease study

The sample size for the foot and mouth disease study was calculated based on sheep observations in four national parks





Foot and Mouth disease study

The input variables to calculate the sample size were:

- Baseline rate of 15%
- Minimum detectable effect of 33.33%
- Statistical significance of 90%

These variables produced a recommended sample size of 870 sheep. For each natural parks, this ment:

- 1.7 days sampling at Yellowstone National park
- 3.5 days sampling at Bryce National park
- 3.1 days sampling at Yosemite National park
- 5.8 days sampling at Great Smoky Mountains National park