



Capstone Option 2:

Biodiversity for the National Parks

Data included in *species_info.csv*



- Category;
- Scientific names;
- Common names;
- Conservation status.



Data included in *species_info.csv*

Number of species in database:

5541



Data included in *species_info.csv*

Categories:

- Mammals;
- Birds;
- Reptiles;
- Amphibians;
- Fish;
- Vascular Plants;
- Nonvascular Plants.

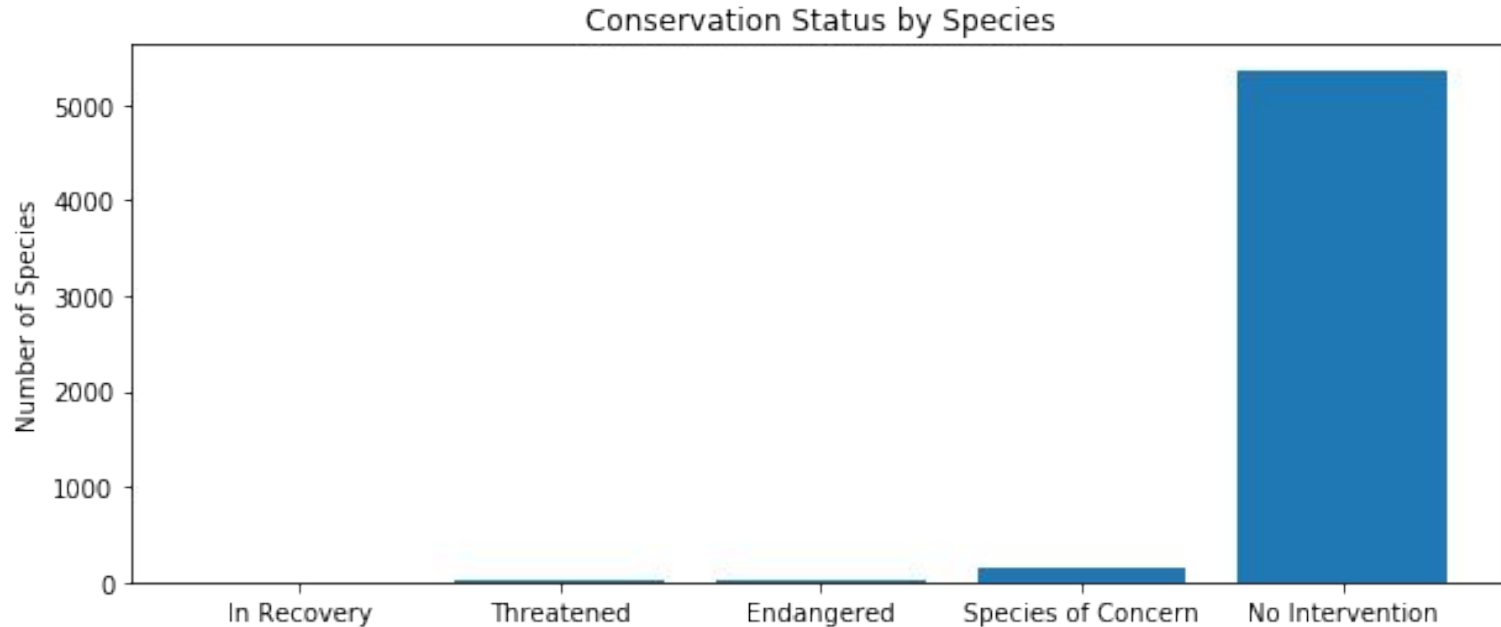


Data included in *species_info.csv*

Conservation Statuses:

- No intervention;
- Species of concern;
- Endangered;
- Threatened;
- In recovery.

Data included in *species_info.csv*





Endangered status of different categories

Are certain types of species more likely to be endangered?



Endangered status of different categories

- Amphibian 8.8%;
- Bird 15.3%;
- Fish 8.7%;
- Mammal 17%;
- Reptile 6.4%;
- Vascular Plant 1%;
- Nonvascular Plant 1.5%.



Results of Chi-Squared Test for Significance

p-value of Bird and Mammal Chi-Squared test - 0.6875

p-value of Reptile and Mammal Chi-Squared test -
0.0383



Results of Chi-Squared Test for Significance

There is a significant difference between Reptile and Mammal

There is no significant difference between Bird and Mammal



Recommendations based on Chi-Squared Test

It turns out Mammals and birds are the most endangered categories. The chi-squared test we ran showed a p-value of 0.6875, so our conclusion is that the difference between percentage of protected mammals and birds is not significant.



Sheep Observations By Park

Bryce National Park - 250

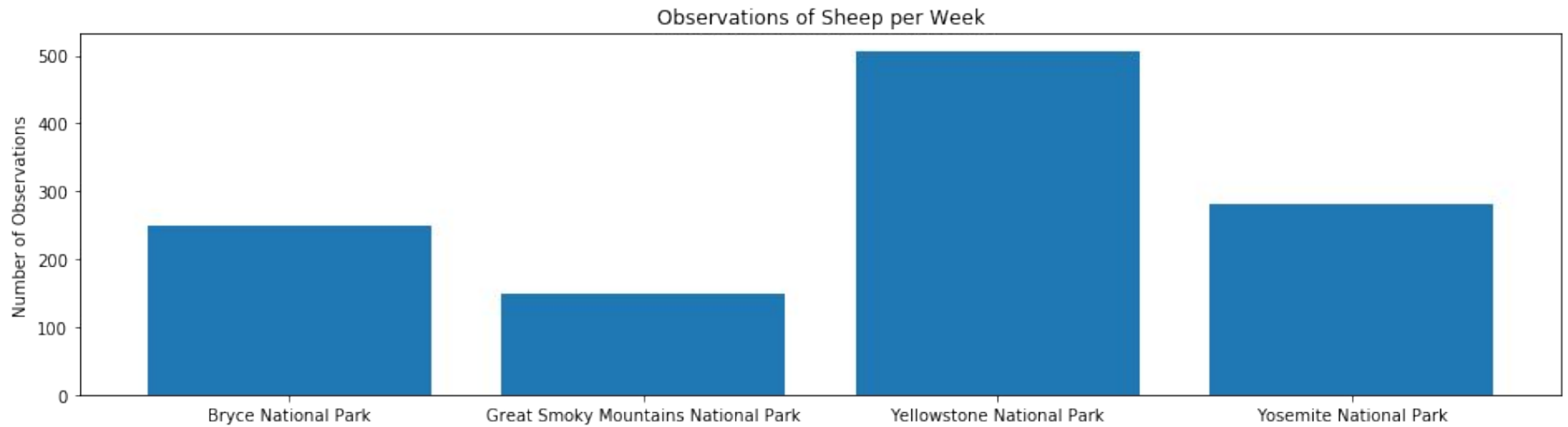
Great Smoky Mountains Park - 149

Yellowstone National Park - 507

Yosemite National Park - 282



Sheep Observations By Park





Foot and mouth disease baseline

15 %



Minimum detectable effect

33 %



Sample size

Sheep to observe from each park at level of significance
90%

Sample size - 510



Weeks of Sheep Observations By Park

Bryce National Park - 2 weeks

Yellowstone National Park - 1 week