

Biodiversity in National Parks

- Analysis

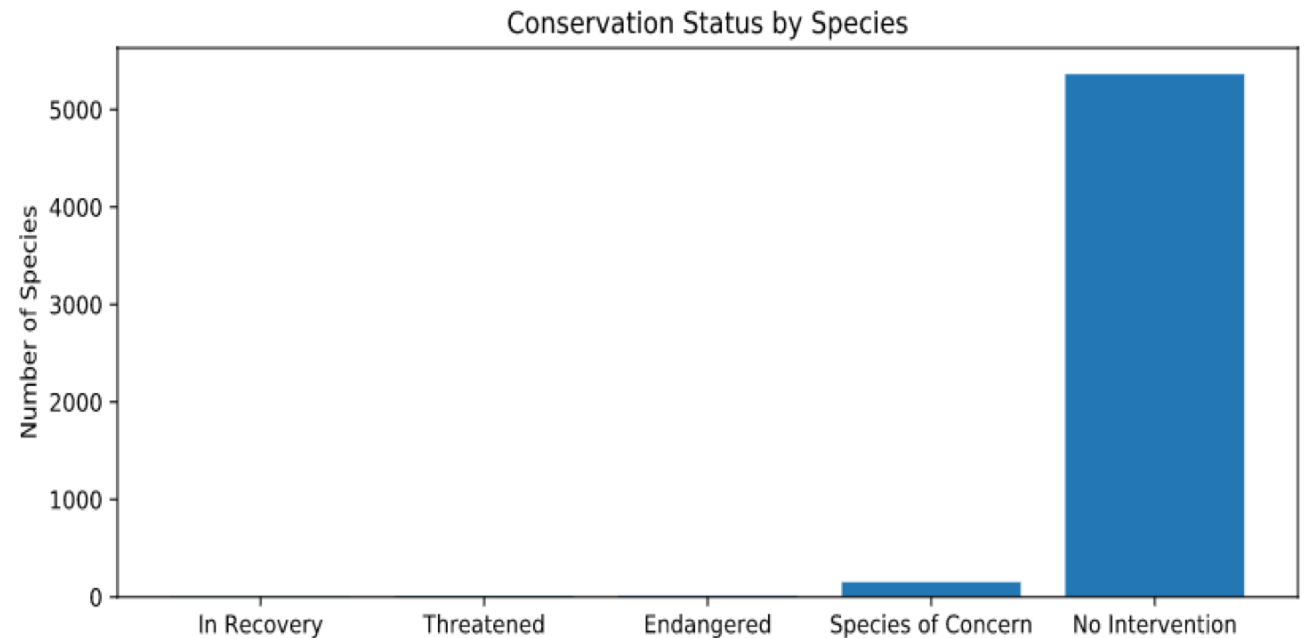
Data Summary

- This dataset includes information about bio-diversity in National parks
- Original dataset includes 5825 species and 4 specifics of the species
 - Category: Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, Non-vascular Plant
 - Scientific Name of the species
 - Common Name of the species
 - Conservation Status: Species of Concern, Endangered, Threatened, In recovery, No information (n/a)

Analysis-I

- Initial analysis on the dataset shows that there are around 5543 unique species
- 96.75% (5363 species) of the species require no intervention and are considered being protected
- 3.25% (180 species) of the species need some sort of protection
- 14% of the species (25 species) that need protection are at the verge of extinction.

conservation_status	scientific_name
0	Endangered 15
1	In Recovery 4
2	No Intervention 5363
3	Species of Concern 151
4	Threatened 10



Analysis-II

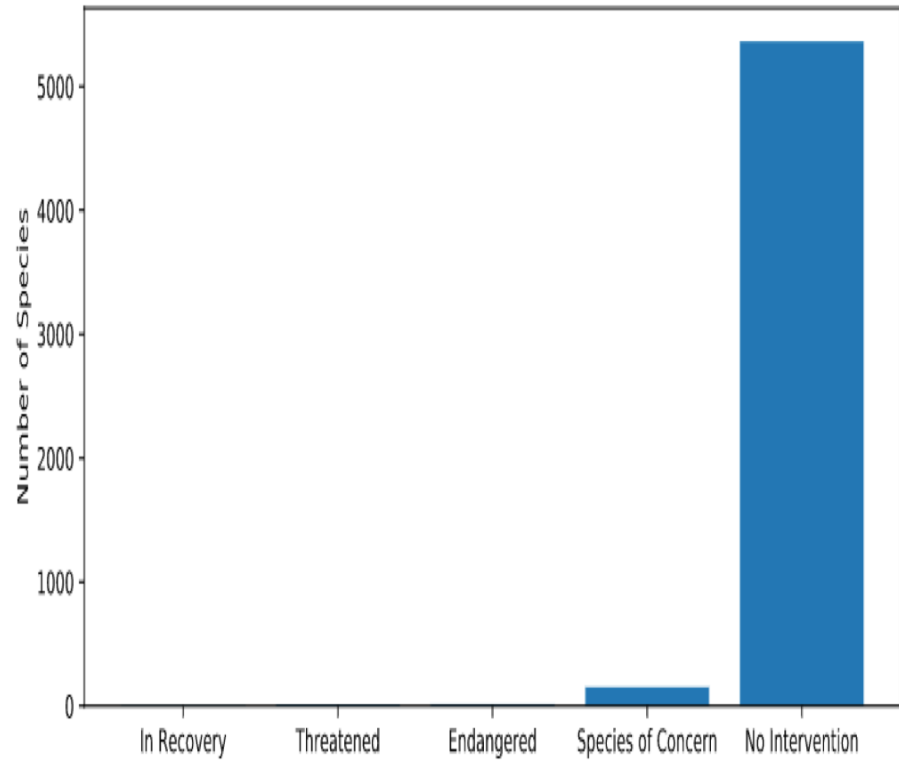
- The following observations are made from Chi Square tests,
 - Species in Mammal and Bird categories are more likely to be endangered than species in Amphibian, Reptile , Fish, Vascular and Non-vascular plant
 - Between species in Mammal and Bird, there is no significant difference in datasets and hence they both have equal probability to be endangered
 - There is a significant difference in Mammal and Reptile datasets – inferred from calculated p-value which is <0.05 . This observation is also the same between Bird and Reptile datasets.
- Based on significance calculations, it would be ideal to focus on protecting species in categories **Mammal** and **Bird**.

Study on Foot and Mouth disease

- Baseline conversion rate: Calculated from historic data. Last year's data from Bryce National Park which is **15%** of the sheep had foot and mouth disease is used as baseline percentage
- Minimum Detectable Effect: at least 5% drop in observed cases of foot and mouth disease in sheep is the requirement and hence minimum detectable effect is **33.33%** $((100-5)/15)$
- Significance level: **90%**
- Sample Size: **870**
- With sheep observations data in multiple n needed to observe enough sheep are,
 - Yellowstone – 1 week and 5 days
 - Bryce – 3 weeks and 3 days
 - Great Smoky Mountains – 5 weeks and 5 days
 - Yosemite – 3 weeks

	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282

Conservation Status by Species



Observations of Sheep per Week

