Capstone 2:Biodiversity Project

By Ewa Jodlowska

Species data findings

Below is a list of all the data findings from species_info.csv

- The data tells us the category of species, their respective scientific name, common name, and conservation status. The status can be one of 5: Species of Concern, Endangered, Threatened, In Recovery, and No Intervention.
- There are 5,541 unique species that belong to 7 different categories: Mammal, Bird, Reptile, Amphibian, Fish, and Vascular Plant.
- There are 161 species that have the status "species of concern", which is the highest status that needs some sort of intervention. There is a big difference between the highest category and the second highest of species that need intervention. The second highest category is endangered with 16 species and the lowest status belongs to species that are in recovery with 4 species.
- We also saw that mammals are the highest category of species that need some sort of intervention because 18% of mammals are protected. The second highest category of species that need intervention is birds. The two lowest are nonvascular plants and vascular plants with 1.5% and 1% respectively.
- When doing a chi squared test, we found that the difference between mammals and birds is not significant. Additionally, we found that the difference between reptiles and mammals is significant.
- When looking at sheep specifically, we found that there are 3 sheep species and 2 of them are protected.

Endangered species

- We saw that mammals are the highest category of species that need some sort of intervention because 18% of mammals are protected. The second highest category of species that need intervention is birds. The two lowest are nonvascular plants and vascular plants with 1.5% and 1% respectively.
- When doing a chi squared test, we found that the difference between mammals and birds is not significant. Additionally, we found that the difference between reptiles and mammals is significant.

Recommendation for conservationists

My recommendation to conservationists is to pay close attention to mammals and birds since they are the highest categories of protected species. Scientists need to work on creating safer environments for those species and work on increasing their populations. The categories of species that need less intervention are the two categories that belong to plants.

Sample size determination

In order for scientists to determine if they are able to detect reductions of at least 5% from the current 15% of known sheep that have foot to mouth disease, they need to have the correct sample size so they can know their results with confidence.

We first needed to determine the minimum detectable effect by taking the minimum difference we want to detect and dividing it by the baseline conversion rate and turning it into a percentage...which gets us 33.33%.

After determining the minimum detectable effect, we use Optimizely's sample size calculator, we see that we need a sample size of 510.

In order to determine the amount of weeks scientists would need to observe sheep at Bryce National Park, we divided the sample size needed by the amount of observations they made per week. For Bryce NP that was 510/250 and for Yellowstone NP that was 510/507. Those calculations told us scientists would have to gather observations for about 2 weeks at Bryce NP and 1 week at Yellowstone.

Status of Species Graph
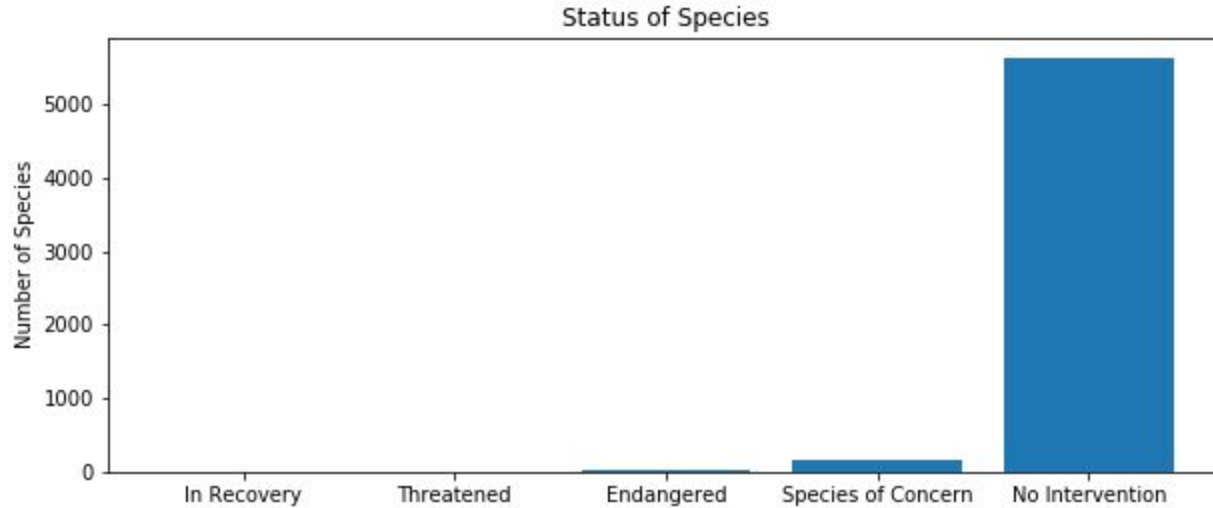
Status of Species

From the data collected, we see that the majority of species do not need protection. The second highest status is those species that are of concern, which is the category right before becoming endangered.

Observations of Sheep per Week

Observations of Sheep per Week
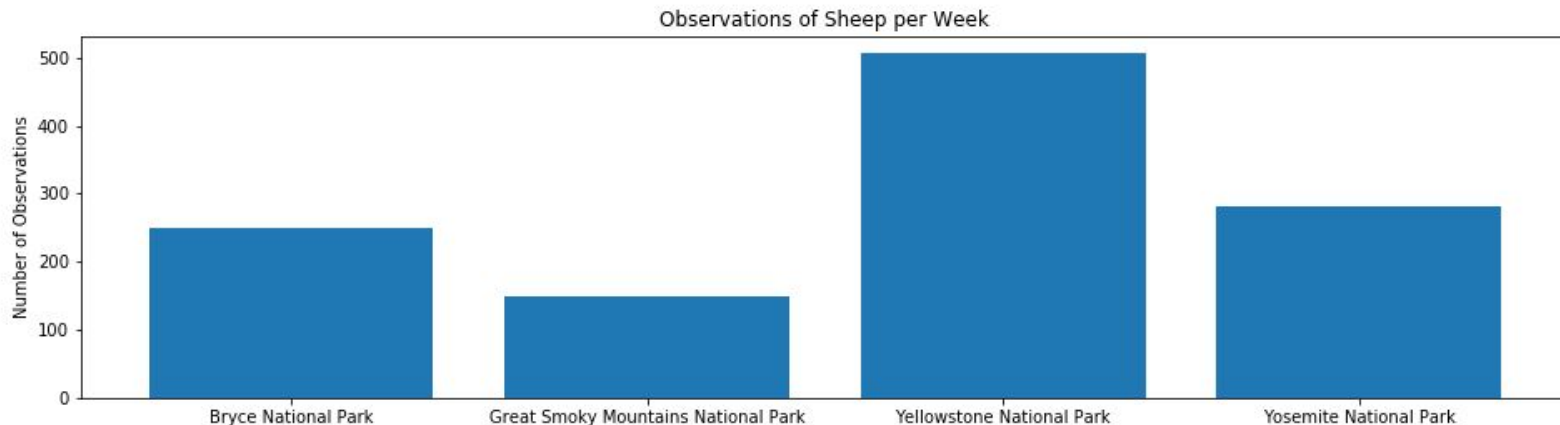
From the data collected, we see that Yellowstone National Park had the highest amount of sheep observations while Great Smoky Mountains National Park had the lowest amount.