Biodiversity in National Parks: Capstone Project

By: Javier Chacon

Inspection of the Dataframe

The data-frame provided for us 5 different columns of data

- Category of species (mammal, reptile, bird, etc.(
- The scientific name of the species
- The multiple common names of the species
- Their conservation status (endangered, species of concern, etc.)

Comparisons between species endangerment across species categories

We performed a chi square test on the proportion of endangered mammals to see if there was a significant difference between the species categories

- Our null hypothesis In this case was that there is no significant difference in between the proportions of endangered mammals and birds/reptiles. Our alternative hypothesis says that there is a significant difference in the proportions. Our p value will test whether the proportions that exist are probable given that the null hypothesis is true.

The categories we compared were

- Mammals and Birds: We found that we have a p-value of .6875, meaning that it is not significant at the 95% confidence level. There is no significant difference b/w the proportion of endangered mammals and birds
- Mammals and Reptiles:We found that we have a p-value of .038, meaning that it is significant at the 95% confidence level. There is a significant difference b/w the proportion of endangered mammals and reptiles.

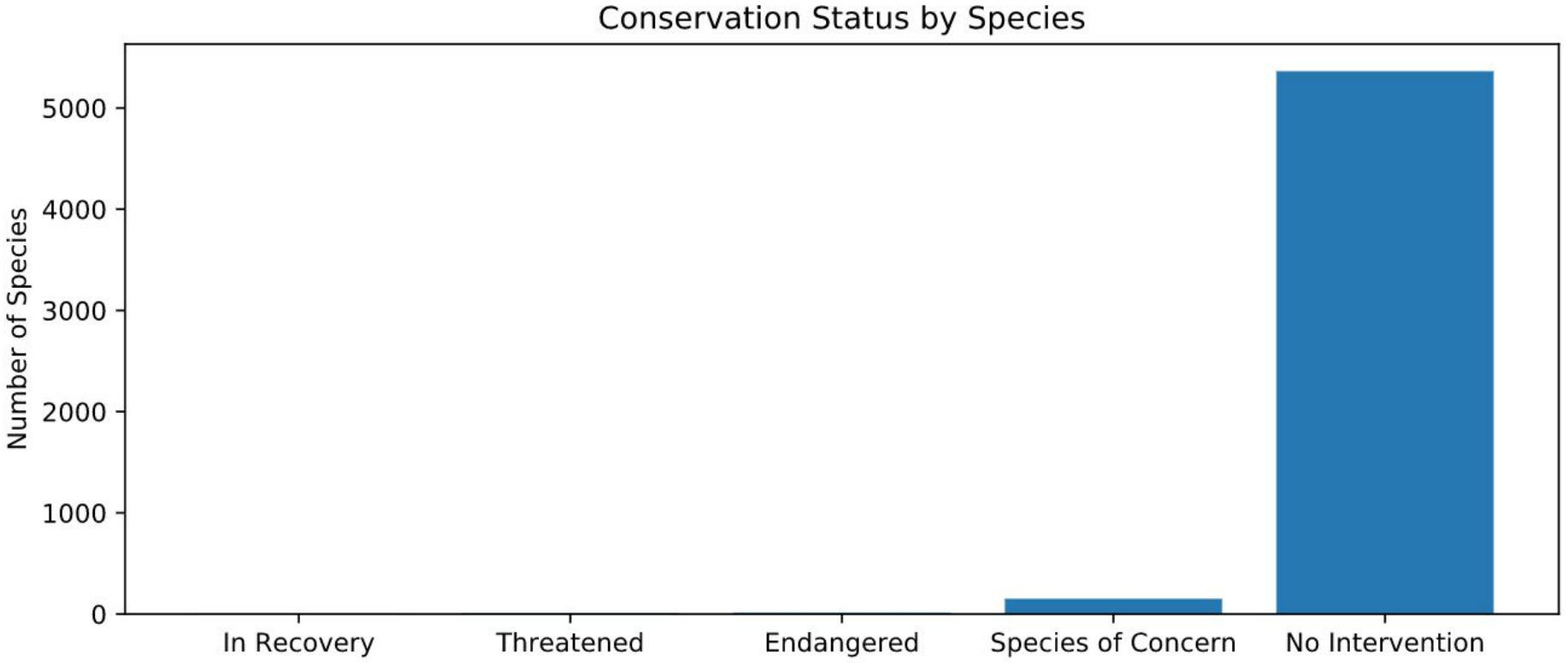
Recommendation

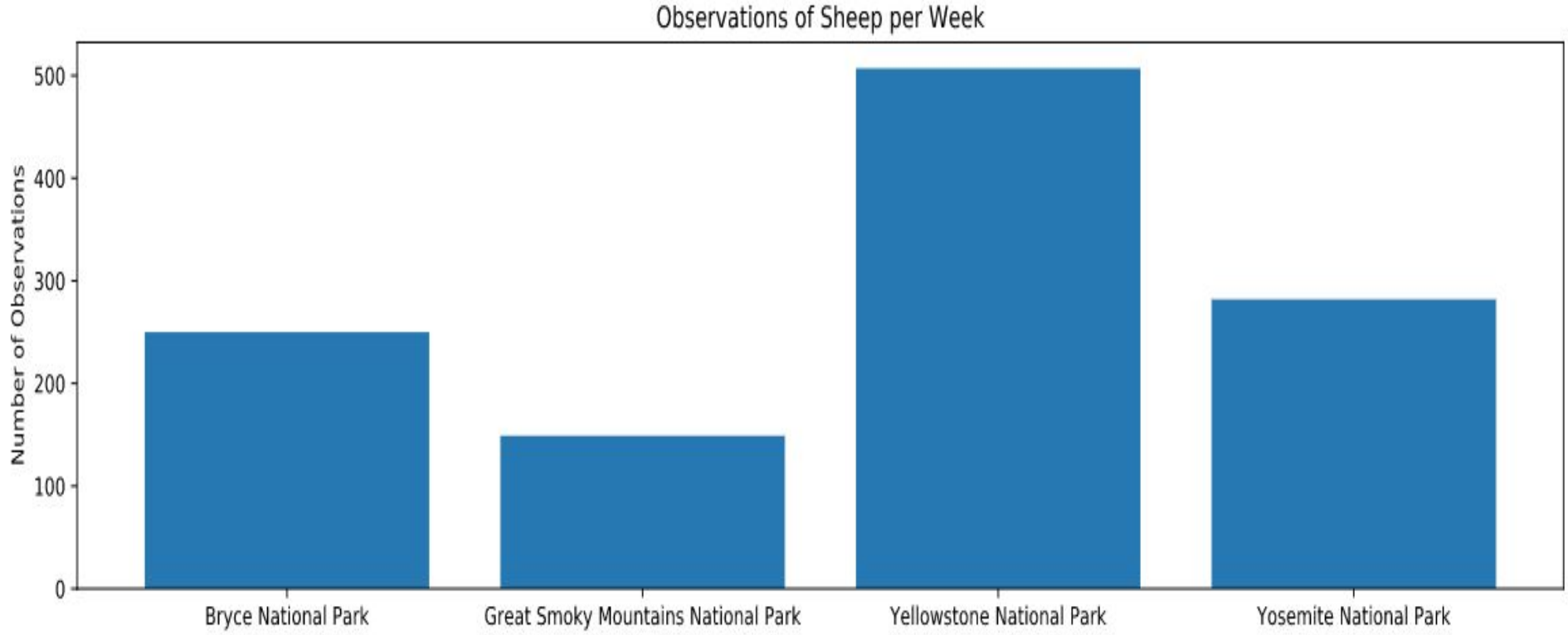
I would suggest that the national parks perform a significance tests across all species and see where they should focus their efforts in protecting and making sure to keep track of endangered species. In our calculations, they should pay specific attention to how they are protecting their mammals as a significant proportion more of mammals are endangered. This may mean allocating more resources and imposing stricter park rules to protect these mammals, as well as tracking them.

Sample Size calculations

In order to know what sample size would deliver us significant proportion results without succumbing to inaccurate proportion sizes due to variation in the sample population, we had to take into account the current proportion (15%) the minimum detectable difference in the number of those infected (33%), and a 90% confidence interval. A sample size of 890 is what we need to give us significant enough observations and samples to come to a significant conclusion on the real proportion of infected sheep in the parks. The number of weeks needed to obtain enough of observations in

- Yellowstone- is 1.755 weeks
- Brice- is -3.56 weeks


Conservation Status by Species
Number of Species
0
1000
2000
3000
4000
5000
In Recovery
Threatened
Endangered
Species of Concern
No Intervention


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