

# **BIODIVERISTY CAPSTONE PROJECT - INVESTIGATING PROTECTED SPECIES**

By Chris Clarke

# OBJECTIVE OF THE PROJECT

The objective of this project was to help the National parks analyze data for endangered species and observations of species across a number of different parks.

There were two parts to the project;

The first was to identify if certain types of species are more likely to be endangered than others.

The second was to provide the parks a sample size for observations so that they could confidently determine if a 5% drop in foot and mouth cases was significant.

# SPECIES DATA — WHAT I NOTICED

- There were 7 different types of Species ('Mammal' 'Bird' 'Reptile' 'Amphibian' 'Fish' 'Vascular Plant' & 'Nonvascular Plant').
- Vascular Plants made up the majority of the not protected species with 4216.
- There were only four conservation statuses originally 'Endangered', 'In Recovery', 'Species of Concern' and 'Threatened' until we added 'No Intervention' when we amended the data from 'None' or NaN to capture all the species.
- The majority (96%) of the Species had a status of No Intervention.
- Only 0.2% of the Species were actually Endangered.

# SIGNIFICANT CALCULATIONS BETWEEN SPECIES

The significant calculations done between species were the following;

- Pivot Table

The pivot table didn't show us a great deal except there was a slight difference between mammals but didn't know prove if it was significant or not.

- Chi Squared

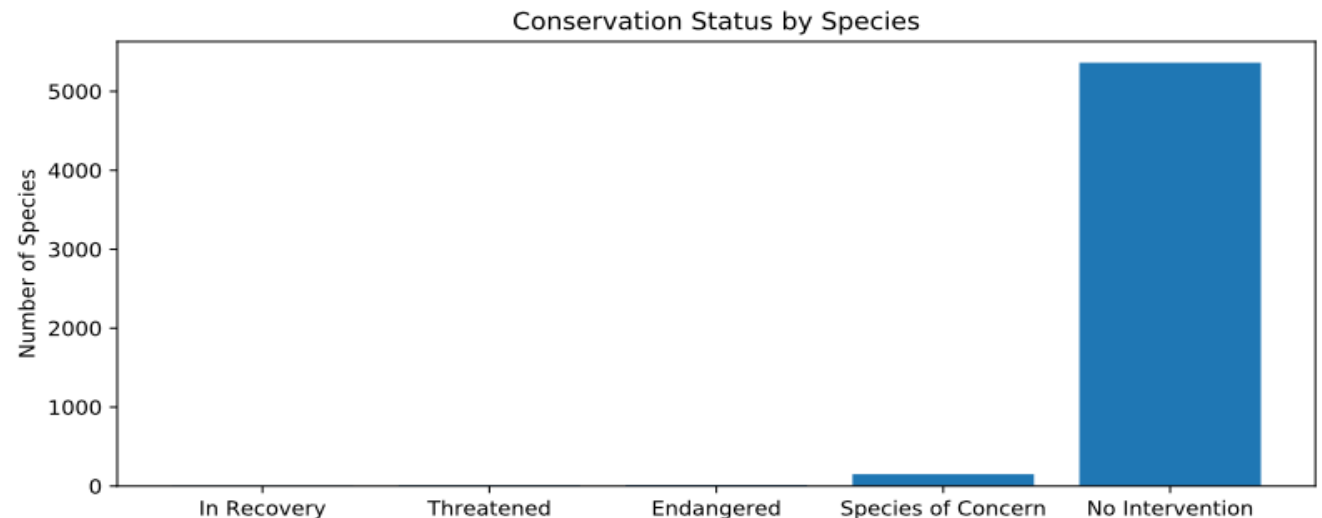
The chi squared test which is great tool to use when you have two or more categorical data sets that you wish to compare. This helped to show that there wasn't a significance when we compared birds and mammals but there was a significance reptiles and mammals helping us to conclude that certain types of species are more likely to be endangered than others.

# A RECOMMENDATION ABOUT ENDANGERED SPECIES

Based on the significance tests there are 2 species with a high protected percentage indicating these should be the priority for any conservationists;

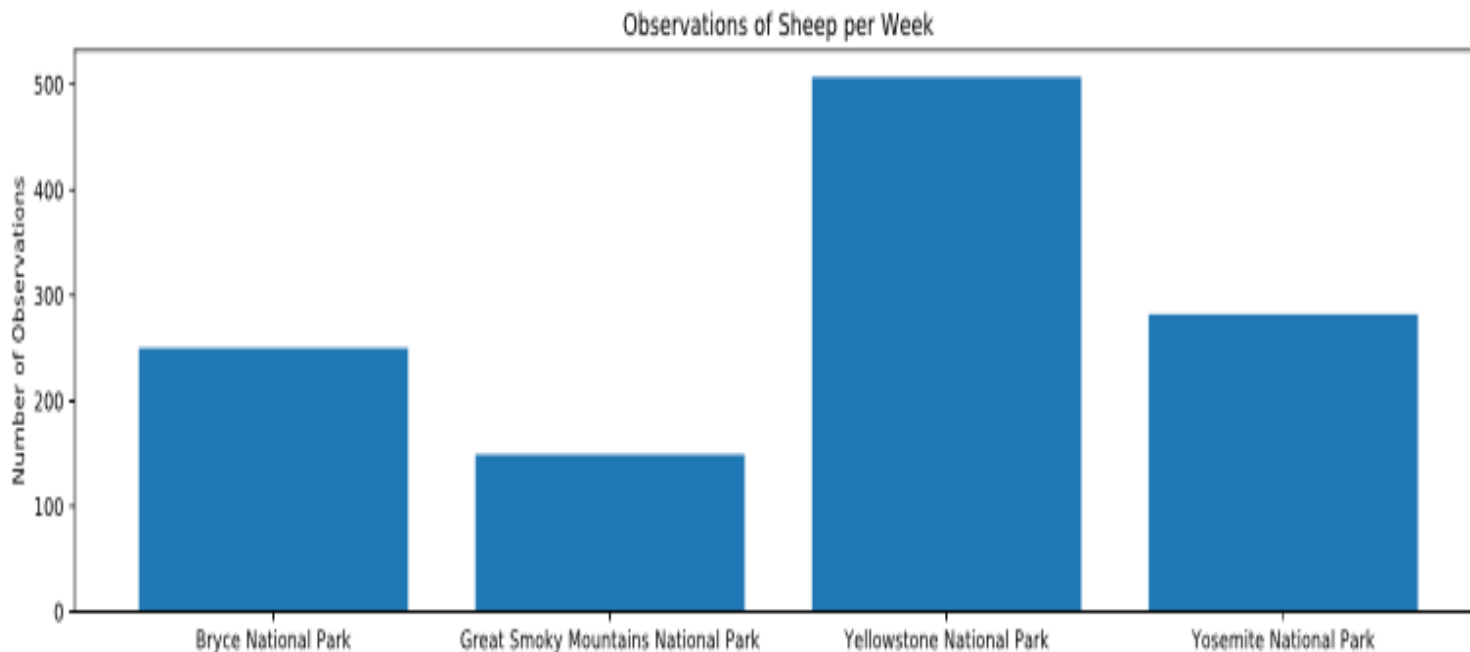
- Birds with 75 protected out of 488 (15%)
- Mammals with 30 protected out of 176 (17%)

Vascular Plants had the biggest disparity between protected (46) and not protected (4216) indicating only 1% of the species is endangered which should be the least concern for conservationists



# SAMPLE SIZE DETERMINATION

This started by narrowing down the data into rows where the species was a sheep and was a mammal and creating a data frame 'sheep\_species'. From here with then merged with the 'observations' data frame created a new data frame called 'sheep\_observations'. We then used this data frame to create a bar chart so we could see how many sheep were viewed at each park across the week.



Now we worked on the sample size determination in order for the Yellowstone National Park Rangers to be able to detect reductions of at least 5 percentage points so that they could know with confidence the % of sheep at their park with foot and mouth. We were able to calculate the sample size using data from the previous year at Bryce National Park, which worked at 870 and with a baseline of 15% it would take just under 2 weeks(1.7) for Yellowstone National Park and approximately 3 and a half weeks for Bryce National Park to see this many sheep.