Biodiversity for the National Parks

Marshall McDougall

Species Data

The data included in the ‘species_info’ csv that was loaded into a pandas dataframe included four columns of information: category, scientific name, common names, and conservation status.

The 5541 species are placed into one of the 6 following categories: mammal, bird, reptile, amphibian, fish, vascular plant, and non-vascular plant.

The species were then given a status based on their population dynamics to see where they lie in terms of endangerment. The status’ used were: endangered, in recovery, species of concern, and threatened. If a species did not fit into any of these categories the status column was left blank. For these values I modified the data frame to indicate that ‘no intervention’ was necessary. The majority of the species required no intervention.

Calculations

Modifications were performed on the data frame to separate the species status into ‘protected’ or ‘not protected’ (protected including all species that did not have ‘no intervention’ as a status). After pivoting the table it could be easily calculated to see which category of species had the most percentage requiring protection (by dividing the amount of protected species by the total species). This showed that birds and mammals had the highest percentage of protected species.

Two Chi-Squared tests were performed to test to see if a species was more or less likely to be protected based on their species category. The first Chi-Squared test yielded a non-statistically significant result comparing mammals and birds, however, the second Chi-Squared test indicated that mammals were statistically more likely to be protected compared to reptiles.

Recommendations

Based on the results of the Chi-Squared tests and what is visible in the bar chart produced while analyzing the ‘species_info’ csv file, we can deduce that mammals have the highest percentage of its species protected, which is statistically significant when compared to reptiles.

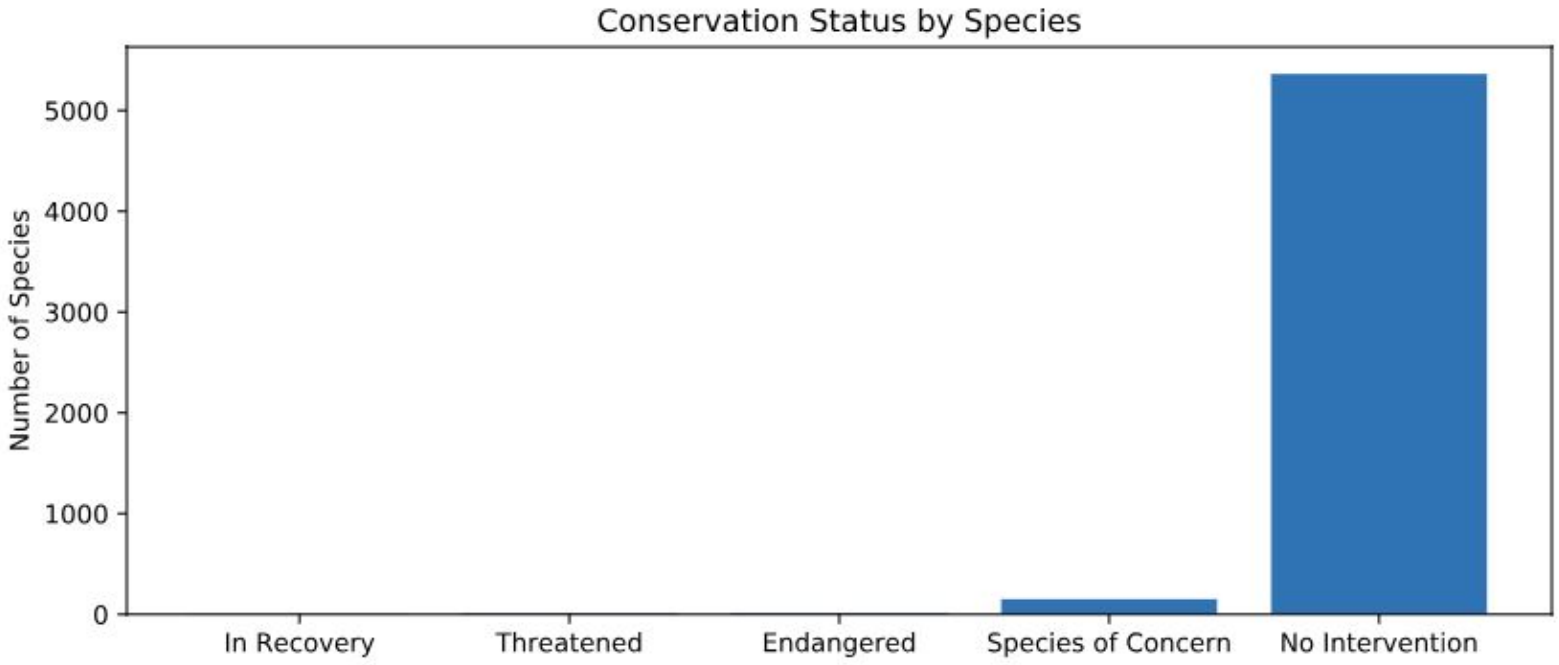
Additional effort in helping species recover from endangerment and preventative measures aimed at keeping species from becoming endangered should be focused on mammals.

Foot and Mouth Disease Sample Sizes

With the goal of detecting a 5% difference in foot and mouth disease among sheep in four national parks from a base of 15%, it was determined that to gain a statistical significance of 90% 890 observations would need to take place. With the observation data from the four national parks the following table shows how many weeks the observation study would take to complete at each of the four locations.

Park	# of Weeks to Complete Foot and Mouth Observation
Bryce National Park	3.56
Great Smoky Mountains National Park	5.97
Yellowstone National Park	1.75
Yosemite National Park	3.16

Graphs Produced



Graphs Produced