Biodiversity for the National Parks

By Tia Hunt

Hi, my name is Tia, and welcome to my presentation for my Capstone project – Biodiversity for the National Parks. In this presentation, I shall investigate endangered species, determine if there is a significant difference in the number of endangered species between two sets of categories, and determine a recommendation for conservationists concerned about endangered species. I will also compare the number of sheep in each park, and then determine a sample size needed to be confident that there is at least a 5% decrease in the number of Foot and Mouth cases.

Investigating Endangered Species

First, I will be talking about the information stored in the species database.

Observations

There are 5824 species in the species list.

Category	No. of species	Percent
Amphibian	80	1.4%
Bird	521	8.9%
Fish	127	2.2%
Mammal	214	3.7%
Nonvascular Plant	333	5.7%
Reptile	79	1.4%
Vascular Plant	4470	76.8%

Conservation Status	No. of species	Percent
Endangered	15	0.3%
In Recovery	4	0.1%
No Intervention	5363	96.8%
Species of Concern	151	2.7%
Threatened	10	0.2%

Data has been recorded for 5,824 species. This data includes, the category, the species scientific and common names, as well as the Conservation Status of the individual species.

In the first table, we see that there are 7 different animal classes, with 76.8% of species recorded being a Vascular Plant, and 1.4% being an Amphibian or a Reptile.

The second table shows that there are 5 different conservation status' that a species could fall into,
- Endangered - the species are at serious risk of extinction.
- In Recovery - the species was formerly Endangered, but currently they are no longer in danger of extinction throughout all or a significant portion of its inhabitable range.
- No intervention - the species are in no danger
- Species of Concern - the species have a declining population or appears to need conservation.
- Threatened - the species are vulnerable to endangerment soon.

We can see that 96.8% of species fall into the conservation status of No Intervention,

whereas 0.1% of species fall into the conservation status of recovery.
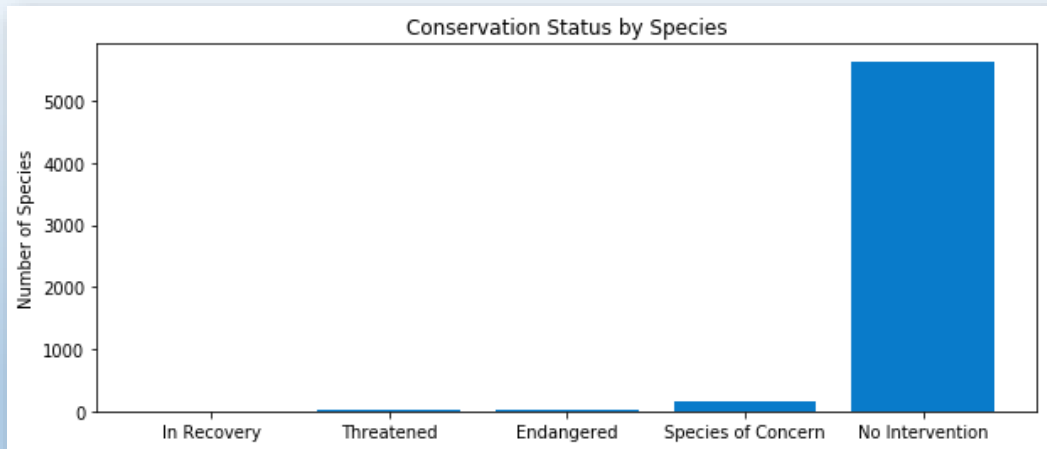
Here is a visual representation of each conservation status, and how many species are in each status. It's visually obvious that there are more species in the No Intervention status then the other status'.

Protected or not protected, is there is difference?

Category	Not Protected	Protected	Percent Not Protected	Percent Protected
Amphibian	72	7	91.1%	8.9%
Bird	413	75	84.6%	15.4%
Fish	115	11	91.3%	8.7%
Mammal	146	30	83.0%	17.0%
Nonvascular Plant	328	5	98.5%	1.5%
Reptile	73	5	93.6%	6.4%
Vascular Plant	4216	46	98.9%	1.1%

The conservationists would like to know if certain types of species are more likely to be endangered? Frist, we need to count how many species in each category are protected, and how many are not protected. To fall under the not protected status, the specie would have a conservation status of no intervention, if a specie falls under any other conservation status, then they fall under the protected status. This table shows the different counts in each category, along with the percent in either not protected or protected. From the percent protected column, it looks as though the species in the Mammal category are more likely endangered than the species in the Bird category. To test for a difference, a chi square test shall be used since there are two or more categorical datasets that we are going to compare.
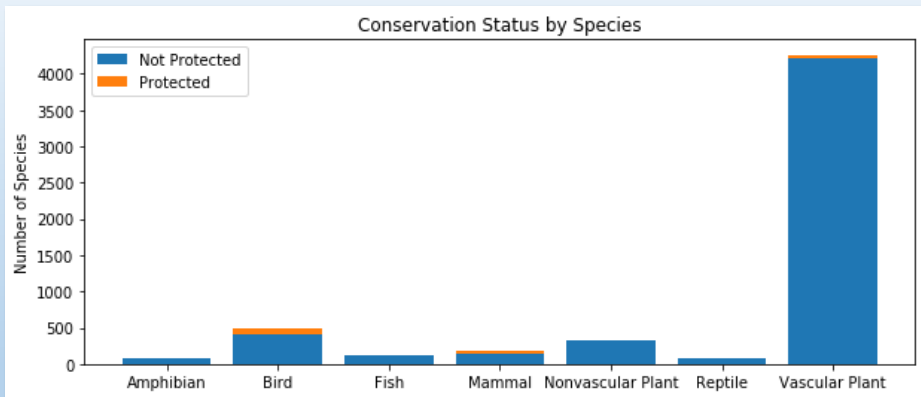
First, here is a stacked bar graph, to visually show the difference between the number of not protected species, compared to the number of protected species in each category. In each category there are more species not protected than there are protected.

- Null Hypothesis: Any difference between species was a result of chance.
- Alternative Hypothesis: There is a significance difference between species.

There is a difference between the following sets of animal classes:

	Mammal	Bird	Reptile	Amphibian	Fish	Vascular Plant	Nonvascular Plant
Mammal	N/A	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE
Bird	FALSE	N/A	FALSE	FALSE	FALSE	TRUE	TRUE
Reptile	TRUE	FALSE	N/A	FALSE	FALSE	TRUE	TRUE
Amphibian	FALSE	FALSE	FALSE	N/A	FALSE	TRUE	TRUE
Fish	FALSE	FALSE	FALSE	FALSE	N/A	TRUE	TRUE
Vascular Plant	TRUE	TRUE	TRUE	TRUE	TRUE	N/A	FALSE
Nonvascular Plant	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	N/A

To test if certain species are more likely to be endangered, we want to state our null and alternative hypothesis. The null hypothesis states that any difference between species was a result of chance. Whereas, the alternative hypothesis states that there is a significance difference between species. When running a chi square test on the Mammal and Bird category, we achieve a p-value of 0.69, which is greater than 0.05, therefore the null hypothesis states that any difference between the Mammal and Bird categories was a result of chance, in this case there is no evidence to accept the alternative hypothesis. Running a chi square test on the Mammal and Reptile categories produces a p-value of 0.039, which is less than 0.05. Therefore, the alternative hypothesis states that these is a significance difference between the Mammal and Reptile categories. Hence, at the 5% level, there is evidence that we can reject the null hypothesis in favour of the alternative hypothesis, and can therefore conclude that some species are more likely to be endangered than others.

Here I have produced a table showing all possible category combinations, and the table consist of the value False when there is not a significant difference between the two categories, True when there is a significant difference between the categories, or N/A when the two categories are the same, so there won't be a difference between them. We can see from a quick glance, that the Vascular and Nonvascular plant are

significantly different from other categories, apart from with each other.

Which category is more likely to be at risk? And Recommendations

Category	More likely
Mammal, Reptile	Mammal
Mammal, Vascular Plant	Mammal
Mammal, Nonvascular Plant	Mammal
Bird, Vascular Plant	Bird
Bird, Nonvascular Plant	Bird
Reptile, Vascular Plant	Reptile
Reptile, Nonvascular Plant	Reptile
Amphibian, Vascular Plant	Amphibian
Amphibian, Nonvascular Plant	Amphibian
Fish, Vascular Plant	Fish
Fish, Nonvascular Plant	Fish

If funding is not already in place, I suggest the following as a rough approximation:
- 20 % funding to mammal
- 18% into bird amphibian and fish
- 16% into reptile
- 5% into Nonvascular and Vascular Plant.

The table we just saw tells us if there is a difference between pairs, but not which category is more likely to be endangered. Using the table before last, which gave the percent protected by each category, we can determine which category is more likely to be endangered. This table shows the pairs of categories that showed a significant difference, and which one is more likely to be endangered.

From the findings, Nonvascular and Vascular Plant both show significant difference against the other categories, but not against one another. When looking back at the percentages of protected species, this indicates that these are less likely to be endangered. When comparing, Bird, Amphibian and Fish categories, they are identical in their relationship to all categories. Therefore, concluding, they have the same risk of being endangered as one another. Mammal showed a significant difference to Reptile, when looking back at the percentages, suggests Mammal is more likely to be endangered. When these were compared to Bird, Amphibian and Fish, there was no significant difference. Concluding from this, my recommendation, would be to put most of the funding into protecting the Mammal, very closely followed by, Bird, Amphibian and Fish, which would receive the same level of funding as each other, then closely followed by Reptile with marginally less funding, then followed by Nonvascular and Vascular Plant, receiving significantly less level of funding than the

rest, but the same as one another. My recommendation, as a quick approximation, would be to put 20 % funding towards Mammal, 18% into Bird, Amphibian and Fish, 16% into Reptile, and 5% into Nonvascular and Vascular Plant.

Sheep Observation

Across 4 National Parks, for seven days, conservationists have been recording sightings of different species. Next, we're going to be investigating the observations made for sheep species.

National Park	Percent of Sheep
Bryce	21.0%
Great Smoky Mountains	12.5%
Yellowstone	42.7%
Yosemite	23.7%

Observations of Sheep per Week
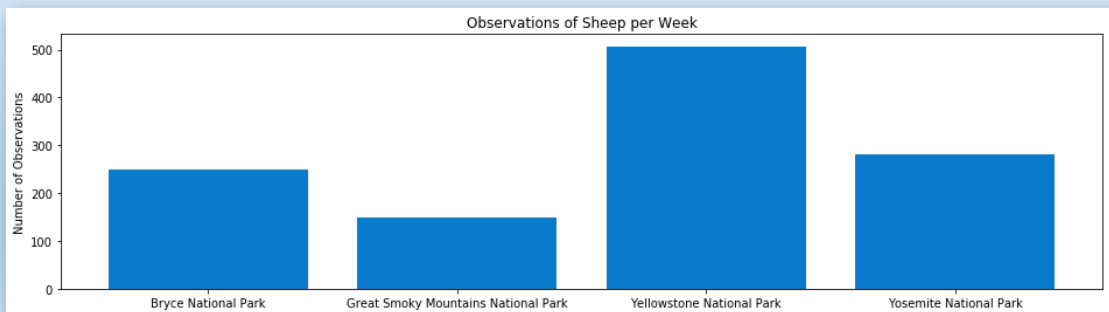
The table shows the percentage of sheep observations across all four parks. The graph shows that, 250 sheep were observed at Bryce National Park, 149 at Great Smoky Mountains National Park, 507 at Yellowstone National Park and 282 at Yosemite National park. There have been more sheep sightings at the Yellowstone National Park.

Foot and Mouth Reduction Effort

There have been cases of the foot and mouth disease within the sheep species. The next couple slides will be how we decide the sample size needed to test if there has been a reduction in the cases.

How many sheep need to be observed to be confident of a 5% decrease?

Last year, it was recorded that 15% of sheep at Bryce National Park had foot and mouth disease.

Hence, the values needed to calculate the sample size are:

• Baseline – 15

• Minimum detectable effect – 33

• Statistical Significance – 90%

Using Optimizely - https://www.optimizely.com/sample-size-calculator/?conversion=15&effect=33&significance=90

Gives a sample size of 520.

We know that last year, it was recorded that 15% of sheep at Bryce National Park had the foot and mouth disease. Park rangers at Yellowstone would like to know if their program to reduce the disease has been successful in reducing the number of sheep with the disease by at least 5%. They would like to know the sample size that they need to check, to be confident that there has been a reduction. For this, we can use the Optimizely sample size calculator. The baseline is 15, the minimum detectable effect is 33 (rounded down), and the statistical significance used is 90%. When these values are put into the Optimizely calculator, we find that 520 sheep would need to be observed, to confidently check if there has been a drop in the foot and mouth disease by at least 5%.

How long will it take to observe enough sheep at each National Park?

National Park	Number of Weeks
Bryce	2.1
Great Smoky Mountains	3.5
Yellowstone	1.0
Yosemite	1.8

To find how long it would take for them to observe the number of sheep needed, divide the sample size by the number of observed sheep at the park. Therefore, it would take the rangers at Yellowstone approximately 1 week to observe enough sheep to be confident in their findings. If each of the parks ran the same program as Yellowstone in an effort to reduce the disease, and that 15% of the sheep previously had this disease, this table shows approximately how long it would take each park to observe enough sheep in order to be confident of their findings.

Thank you for your time!