Biodiversity Capstone Project

Investigating Protected Species

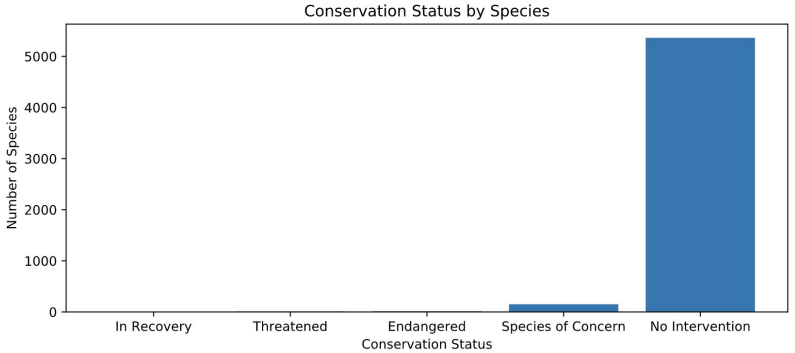
Analysis by Anna Campanelli
03/27/2018

Descriptions of species_info.csv

Upon review of the data provided by the National Park Service, the following observations can be made about the types of species in National Parks and conservation statuses:

- There are **7** different species_types in the DataFrame: *Mammal*, *Bird*, *Reptile*, *Amphibian*, *Fish*, *Vascular Plant,* and *Nonvascular Plant*. In total, the DataFrame captures over **5500** unique species.
- Upon first glance at conservation status data, the table was missing a large portion of the total species. Once updated to include species that do not need protection (*No Intervention*), we observed that most animals are thriving! Less than **5%** of species require protection efforts. The table to the right show the percentage of total species that fall into each conservation status, as does the figure below it.

	conservation_status	scientific_name	Percent of total
1	In Recovery	4	0.0721630885802
4	Threatened	10	0.18040772145
0	Endangered	15	0.270611582176
3	Species of Concern	151	2.7241565939
2	No Intervention	5363	96.7526610139



Significance calculations for endangered species

We observed that the *percent of species protected** varies by species_type, these values are shown in the table below. However, in order to prove that the differences in species protected are significant, so we ran *Chi-Squared significant tests* to compare across species. Finding below:

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

- **Bird to Mammal**: Found a p-value of ~0.688. The difference between the % of protected birds and mammals is not significant.
- **Amphibian to Mammal**: Found a p-value of ~0.128. The difference between the % of protected amphibians and mammals is not significant.
- **Fish to Mammal**: Found a p-value of ~0.056. The difference between the % of protected fish and mammals is not significant.
- **Vascular Plant to Mammal**: Found a p-value > 5%. The difference between the % of protected vascular plants and mammals is not significant.
- **Nonvascular Plant to Mammal**: Found a p-value of > 5%. The difference between the % of protected nonvascular plants and mammals is not significant.
- **Reptile to Mammal**: Found a p-value of **~0.038**. The difference between the % of protected reptiles and mammals is significant!

*Value is equal to the number of species that are protected divided by the total number of species.

Recommendations for conservationists

After comparing the rate at which two different species types are protected, we observed that most differences are a result of chance.

However, when testing the significance of the percentage of Reptiles that are protected (**6.4%**) to the percentage of Mammals (**1.7%**), we found that **Reptiles *<u>are</u>* more likely to be endangered than Mammals**.

Recommendation: Monitor Reptile populations closely in National Parks in order to prevent endangerment across the different Reptile species.



Determining sample size for Foot and Mouth Disease

Figure 1 to the right shows the total number of sheep observed in each National Park over a 7 day period. These observations were plugged into a sample size calculator (*Figure 2*) to determine the sample size required to achieve 90% confidence in results.

As shown the **Sample Size required is 890**.

In addition, using the 15% baseline provided and the goal to observe a 5% change, we determined a **Minimum Detectable Effect of 33%**.

Figure 3 to the right shows amount of time needed to observe sheep are shown in the table below

	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282

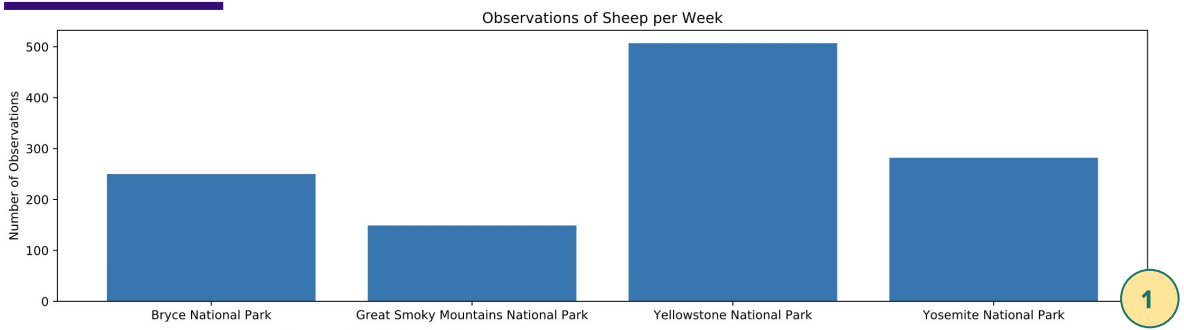
1

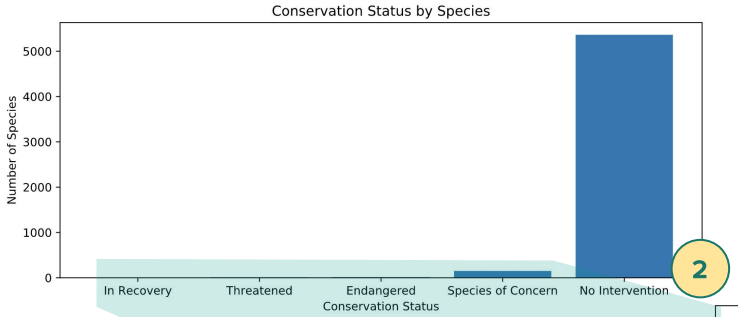
2

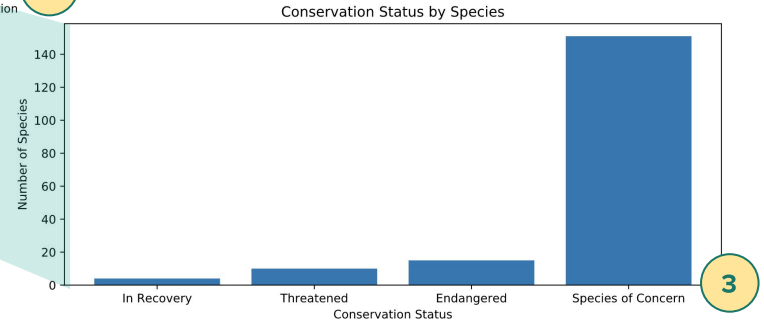
National Park	Observation Time Required (calculated in Weeks)
Bryce	3.56
Great Smoky Mnts	5.97
Yellowstone	1.75
Yosemites	3.16

3

Additional noteworthy findings







1. Shows sheep observations across National Parks. Pulled from a merge of the `observations` and `species` DataFrames

2. Shows the number of unique species that fall into each of the 5 conservation statuses. Majority fall into *No Intervention*. Pulled from `species` DataFrame. Also shown on Slide 2.

3. Shows the data as Figure 2, however *No Intervention* status is removed in order to show a sense of scale of active conversation effort for unique species