Capstone Option 2: Biodiversity for the National Parks

Rachel Colyer
rachel.colyer@gmail.com

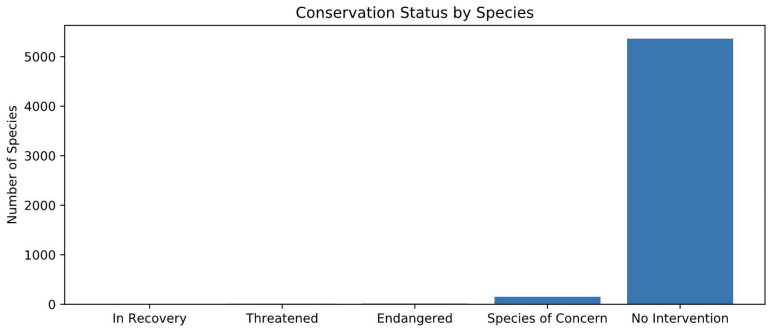
Species_info

- Contained the category,scientific_name,common_names,conservation_status.
- Some of the data points collected were
  a. Species count (5541), species types, conservation status, and counts of conservation status within the file.
- Then, we implemented a few fixes to the data to fill in missing information and transform it into a more useful form

```
5541
['Mammal' 'Bird' 'Reptile' 'Amphibian' 'Fish' 'Vascular Plant'
 'Nonvascular Plant']
[nan 'Species of Concern' 'Endangered' 'Threatened' 'In Recovery']
  conservation_status  scientific_name
0          Endangered               15
1         In Recovery                4
2     No Intervention             5363
3  Species of Concern              151
4          Threatened               10
```

Conservation Status Bar Chart

I thought the scaling on the bar chart should have been different. I think adding in the actual numbers for each item on the x-axis would make this information more useable in a real-world scenario.



Are certain types of species more likely to be endangered?

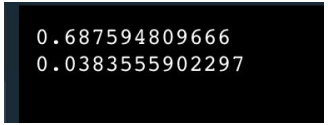
Using pivots we were able to restructure the data in a useful way.

	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

Chi-Squared tests to determine if some species are more likely to be endangered.

This test found no statistical significance between mammals and birds, which have the highest numbers in that category.

There was a statistical significance between reptiles and mammals. This shows the possibility that some species could be more likely to become endangered, voiding the original null hypothesis.



Recommendations:

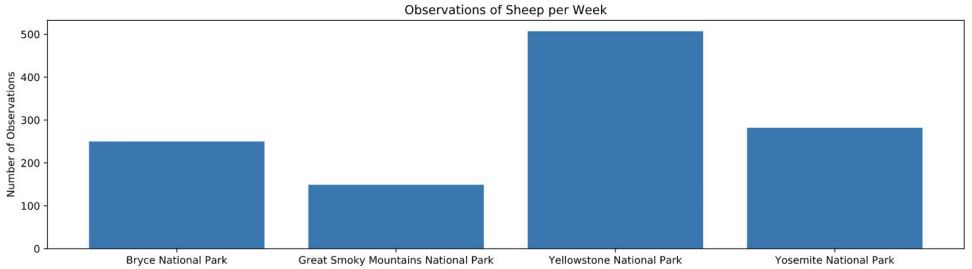
Conservationists should watch species populations that are more vulnerable and work towards protecting them before the reach endangered status.

SHEEP!

Sheep in yellowstone are on the move and visible.

During this exercise we learned how to find more specific data within fields that likely could have been categorized in additional columns, but were not.

	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



Foot and Mouth Disease

Scientist would need to observe 510 sheep, and that would likely be in yellowstone, as it was the leading park for sheep observation. This number is also slightly higher than recorded observations so it might be difficult to reach these conclusions, especially in the smaller parks though they might be able to meet these numbers over a period of time. If that works for the program, then they would have a reliable way to track herds for foot and mouth disease.