Biodiversity for the National Parks

CAPSTONE PROJECT

BARBARA DUDAS

2018.03.27.

Raw Data: species_info.csv

- Our findings are based on a .csv file containing all the information about different species in our National Parks, including:
  - The categories of the animals: (Amphibian, Bird, Fish, Mammal, Nonvascular Plant, Reptile, Vascular Plant)
  - The scientific name of each species
  - The common names of each species
  - The species conservation status (species of concern, threatened, endangered, in recovery, no intervention)

Conservation statuses

What does each status mean?

- **Species of Concern**:
  declining population or appears to be in need of conservation.
- **Threatened**:
  vulnerable to endangerment in the near future.
- **Endangered**:
  seriously at risk of extinction.
- **In Recovery**:
  formerly Endangered, but currently not in danger of extinction throughout all or a significant portion of its inhabitable range.
- **No Intervention**:
  there is no need for any intervention in case of these species.

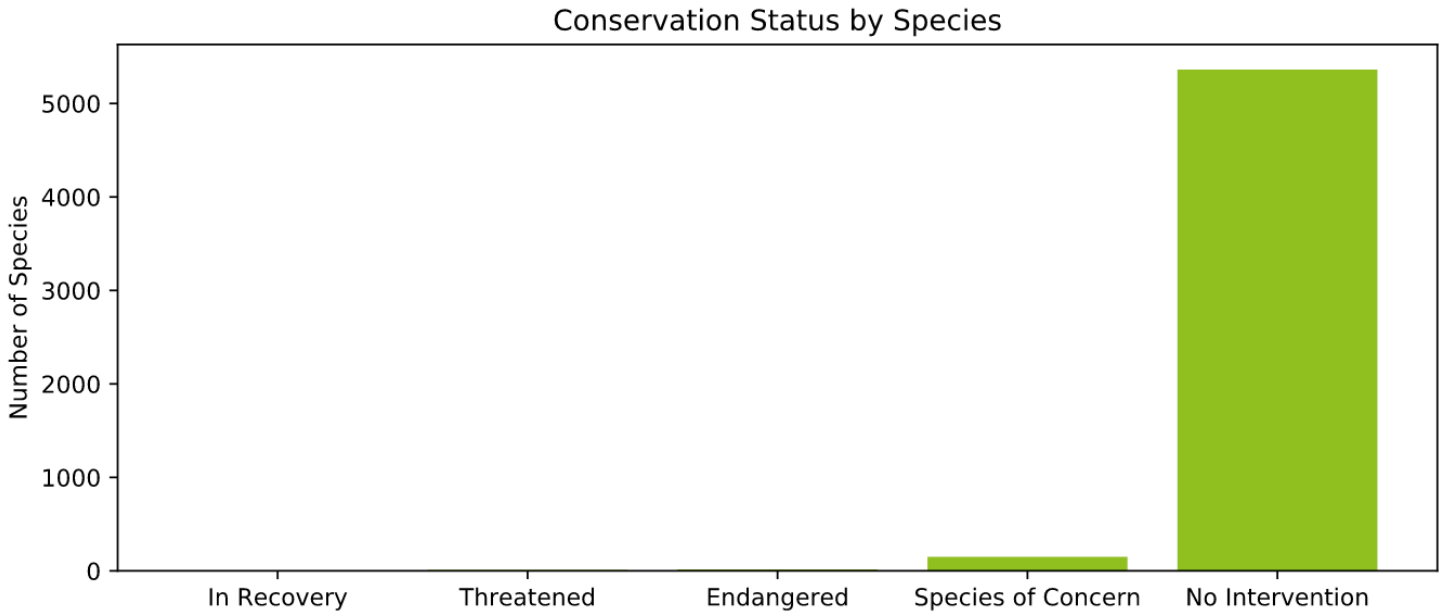
What we know based on the statuses:

- The following number of species fall into each categories:

STATUSES	NUMBER OF SPECIES
Endangered	15
In Recovery	4
Species of Concern	151
Threatened	10
No Intervention	5363

- Based on this, we can say altogether 180 species need some kind of help!

So the second largest conservation status is the species of concern



How many species need help from each animal category?

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	8.86%
1	Bird	413	75	15.36%
2	Fish	115	11	8.73%
3	Mammal	146	30	17.04%
4	Nonvascular Plant	328	5	1.50%
5	Reptile	73	5	6.41%
6	Vascular Plant	4216	46	1.07%

Are certain types of species more likely to be endangered?

- We initially saw that there was a slight difference in the percentages of birds and mammals that fall into a protected category. Our **null hypothesis** here was that this difference was a result of chance.
- When we ran our chi-squared test, we found a p-value of **~0.688**, so we can conclude that the difference between the percentages of protected birds and mammals is not significant and is a **result of chance**.
- We also ran our chi-squared test for checking the difference between the percentages of protected vascular and non-vascular plants, and we found a p-value of **~0.662**, so we can conclude that the difference between them is not significant and also a **result of chance**.
- But, when we compared the percentages of protected reptiles and mammals and ran the same chi-squared test, we calculated a p-value of **~0.038**, which **is significant**.
- Therefore, we can conclude that **certain types of species *are* more likely to be endangered** than others.

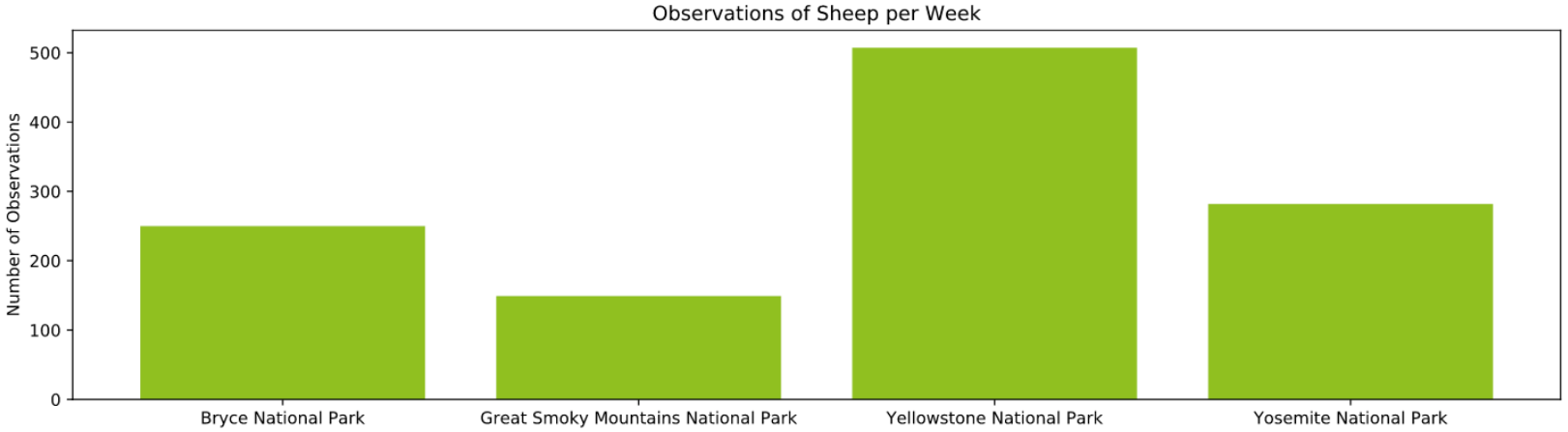
Recommendations

- I would recommend you to keep an eye on each species. We can confidently say that **mammals are more likely to need help than e.g. reptiles**.
- We also can say that **there is no significant difference between vascular or non-vascular plants** and their endangerment.
- Also we need to highlight here that there is a significant difference based on the p-value for reptiles and vascular plants, so we can safely say that **reptiles are more likely to need help than vascular plants**.
- Based on the protected percent of each species we can say more species need help from the birds or mammal category than e.g. any plant category.

Observation of diseased animals from last year for each parks

	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

Check out the same data on a graph



Based on the previous data, we can say:

- According to our calculations the baseline was **15%**, and the minimum detectable effect was **33.33%**.
- Based on this data using the calculator's results we can safely say that the scientists need to sample at least **870** animals to have a significant result for their test.

- **How many weeks they need for sampling enough animals?**
  In case of Yellowstone National Park:
  ~ 1,72 weeks
  In case of Bryce National Park:
  ~ 3,48 weeks

Thank you for your attention!