

# Biodiversity Project

By Ross Taylor

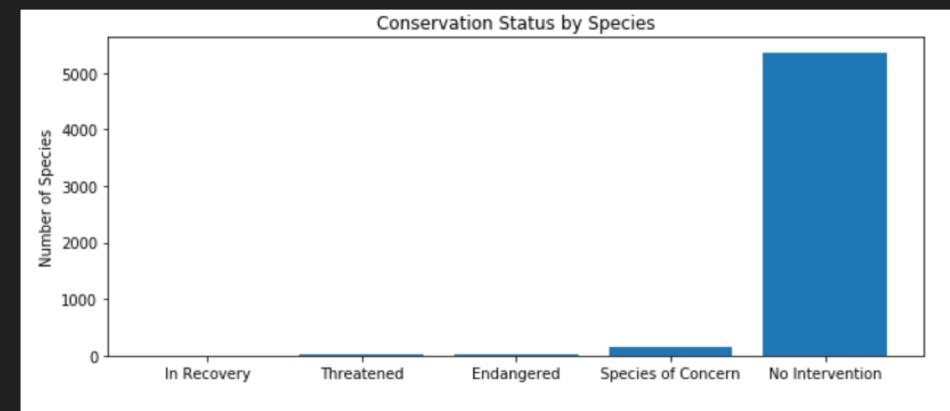
# The Species Dataframe

category	scientific_name	common_names	conservation_status
0 Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole	nan
1 Mammal	Bos bison	American Bison, Bison	nan
2 Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Domesticated Cattle	nan
3 Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	nan
4 Mammal	Cervus elaphus	Wapiti Or Elk	nan

- The species dataframe contained 5,541 different species. For each species the table provided us with their category (e.g. mammal), scientific name, common names (where applicable) and conservation status.
- There were 7 different categories which were mammal, bird, reptile, amphibian, fish, vascular plant and non-vascular plant.
- There were also 5 different conservation statuses including Endangered, in recovery, no intervention (aka. nan), species of concern and threatened

# The Species Dataframe (part 2)

- A large number of species required no intervention (see bar chart bottom right), which was positive, so next we thought to compare the percentage of species who needed intervention from different categories.



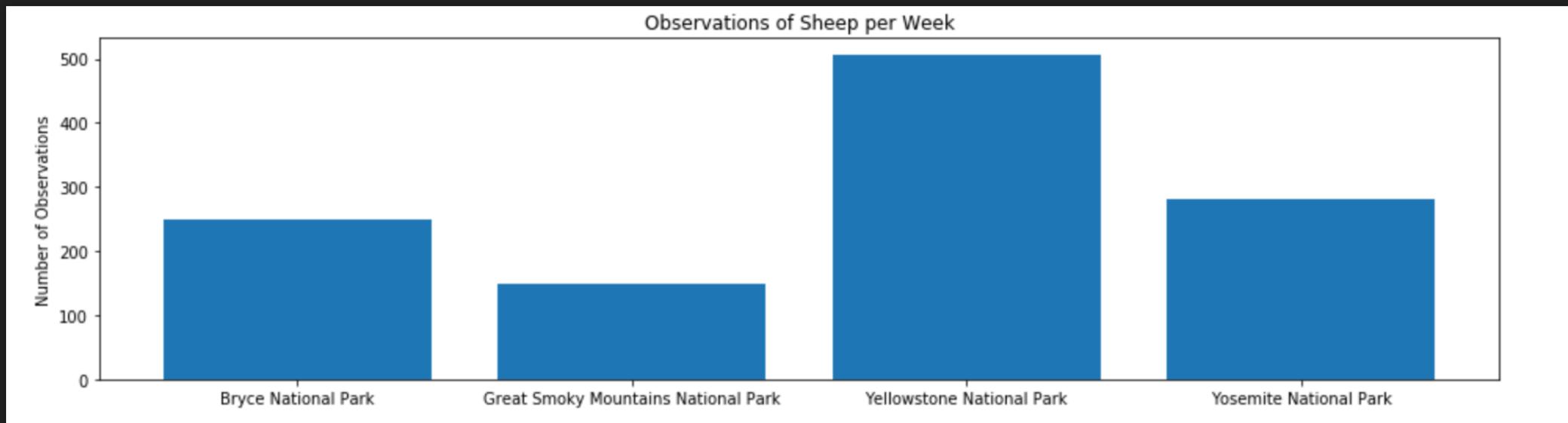
# The Species Dataframe (part 3)

- We first used the chi-squared test to compare the percent of mammals to birds which required protection, but the results were insignificant.
- Then we compared mammals to Reptiles, and this time the results were significant, and using this along with the graph (bottom right) allows us to conclude that mammals are significantly more likely to need protecting than reptiles.
- Which leads us to the conclusion that conservationists concerned about endangered species should focus on mammals over reptiles, however more tests would have to be done to get a more detailed view of the other categories.

	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

# Foot and mouth disease – sample size

- We determined that with a baseline conversion rate of 15%, statistical significance of 90%, and minimum detectable effect of 33.3% reoccurring we would need a sample size of 870 sheep.
- Meaning that at Yellowstone the scientists would need at least 2 weeks to observe enough sheep and at least 4 weeks at Bryce.



**Thank you for  
listening, any  
questions?**

Slide by Ross Taylor

Website: [fictional\\_datascientist.com](http://fictional_datascientist.com)

Number: 0800- fictional-datascientist-1234