Biodiversity Project

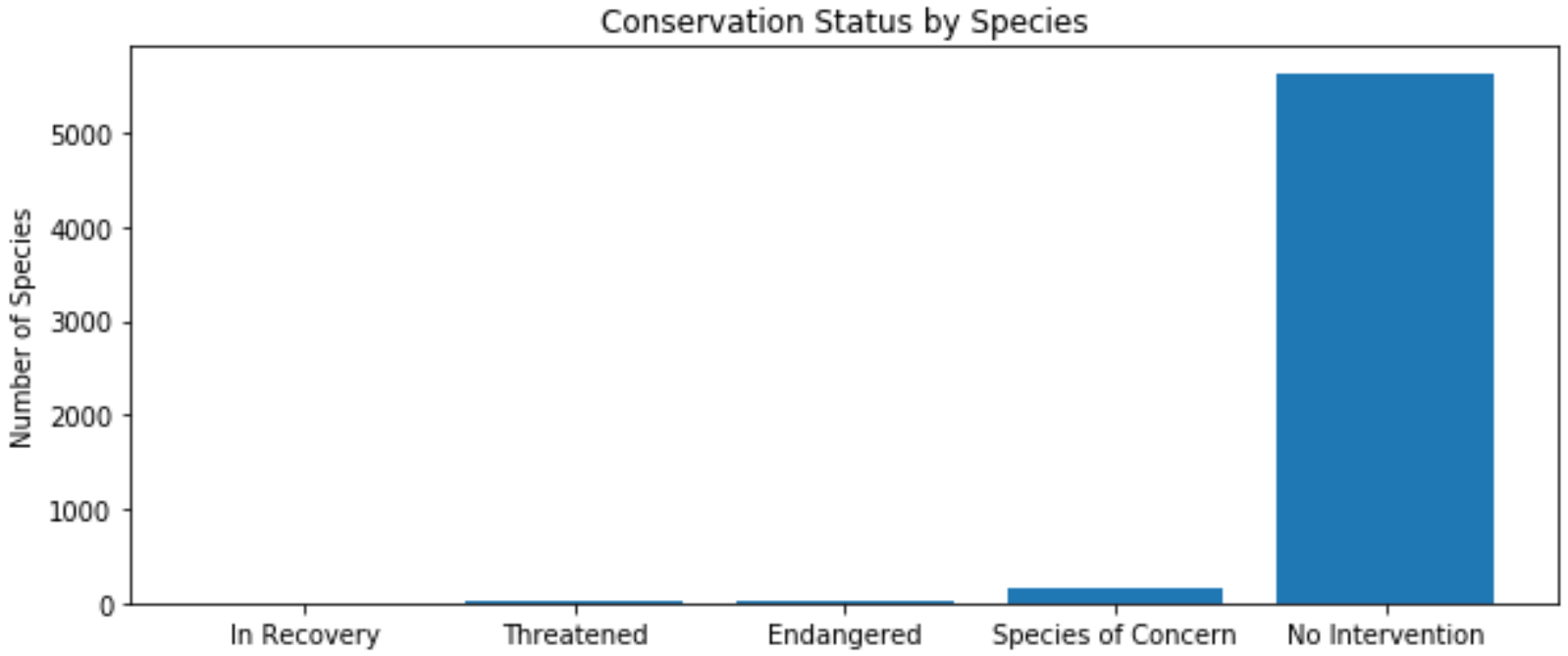
By: Matt Byrne

Species Info Data

- The CSV species_info.csv includes data about different species in various National Parks.
- The data is sorted by the scientific name, common name and conservation status of each species.
  - 5541 different species
  - 7 categories
  - 5 categories of conservation status (many of the species initially listed as ‘nan’)

Analysis Steps

- Previously mentioned: many conservation statuses listed as ‘none’
  - Corrected this so “none” is replaced by a more accurate “No Intervention”.
- The results of grouping our species by conservation status can be seen in the figure below.



Analysis Steps

- We arranged data to see what species required protection (had their numbers threatened) and rearranged categories to create a more useful pivot table.
- We found the percent protected by dividing our protected species by our protected species plus our non protected species. The data from this calculation is available in the Table 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

Analysis Steps

- We suspected that Mammals were more likely to be endangered than Birds.
- We conducted a chi squared test, first creating a contingency table for mammals and birds, with categories “protected” and “not protected”.
- We found that our results were not significant for Mammals and Birds but a second test between Reptiles and Mammals yielded significant differences.

Analysis Steps - Observations

- We analyzed data concerning sightings of species in National Parks.
- We sorted our data to extract which species were sheep (Figure A) and further clarified out sorting to extract plant species that had made our initial cut (Figure B).

A

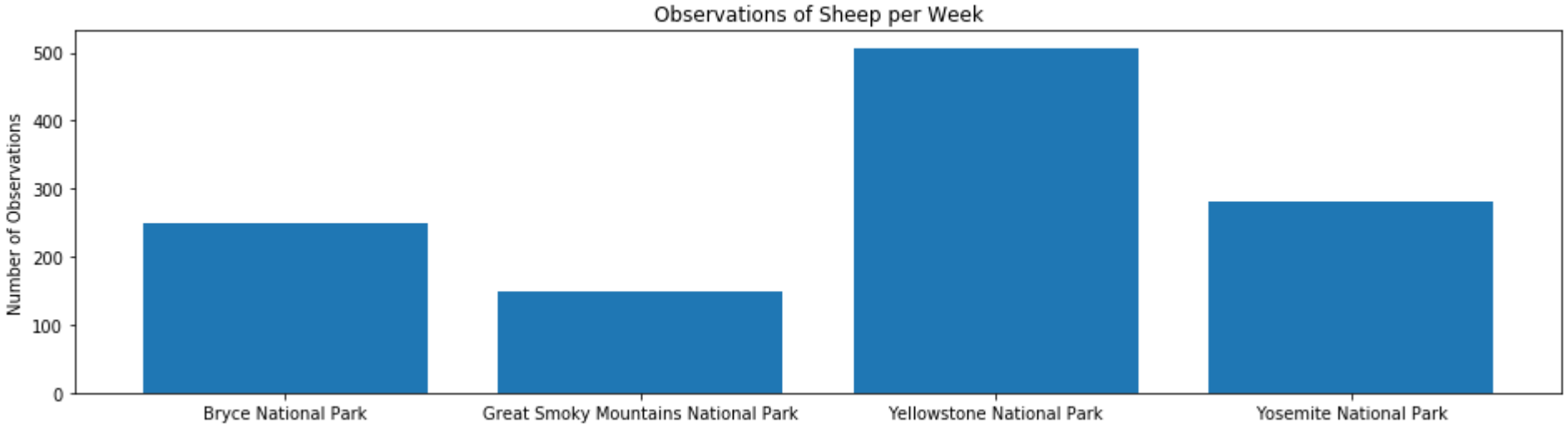
	category	scientific_name	common_names	conservation_status	is_protected	is_sheep
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
1139	Vascular Plant	Rumex acetosella	Sheep Sorrel, Sheep Sorrell	No Intervention	False	True
2233	Vascular Plant	Festuca filiformis	Fineleaf Sheep Fescue	No Intervention	False	True
3014	Mammal	Ovis canadensis	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
3758	Vascular Plant	Rumex acetosella	Common Sheep Sorrel, Field Sorrel, Red Sorrel,...	No Intervention	False	True
3761	Vascular Plant	Rumex paucifolius	Alpine Sheep Sorrel, Fewleaved Dock, Meadow Dock	No Intervention	False	True
4091	Vascular Plant	Carex illota	Sheep Sedge, Smallhead Sedge	No Intervention	False	True
4383	Vascular Plant	Potentilla ovina var. ovina	Sheep Cinquefoil	No Intervention	False	True
4446	Mammal	Ovis canadensis sierrae	Sierra Nevada Bighorn Sheep	Endangered	True	True

B

	scientific_name	park_name	observations	category	common_names	conservation_status	is_protected	is_sheep
0	Ovis canadensis	Yellowstone National Park	219	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
1	Ovis canadensis	Bryce National Park	109	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
2	Ovis canadensis	Yosemite National Park	117	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
3	Ovis canadensis	Great Smoky Mountains National Park	48	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
4	Ovis canadensis sierrae	Yellowstone National Park	67	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
5	Ovis canadensis sierrae	Yosemite National Park	39	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
6	Ovis canadensis sierrae	Bryce National Park	22	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
7	Ovis canadensis sierrae	Great Smoky Mountains National Park	25	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True
8	Ovis aries	Yosemite National Park	126	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
9	Ovis aries	Great Smoky Mountains National Park	76	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
10	Ovis aries	Bryce National Park	119	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
11	Ovis aries	Yellowstone National Park	221	Mammal	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True

Analysis Steps - Observations

- We could then find how many observations were made, across all 3 sheep species by park. These results are presented in a Table and Figure 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

Analysis Steps – Recommendations

- The scientists knew a 15% rate of disease exists in the sheep of Bryce NP. Yellowstone scientists want to know whether their program is reducing their rate of diseased sheep.
- Using a significance level of 90%, we found a sample size of 510 needed to be observed for our minimum detectable effect of 33.33%
- We found that scientists needed to observe sheep at Bryce a little over 2 weeks (2.04) in order to get enough sheep.
  - Scientists at Yellowstone would need to observe a little over 1 week (1.006).