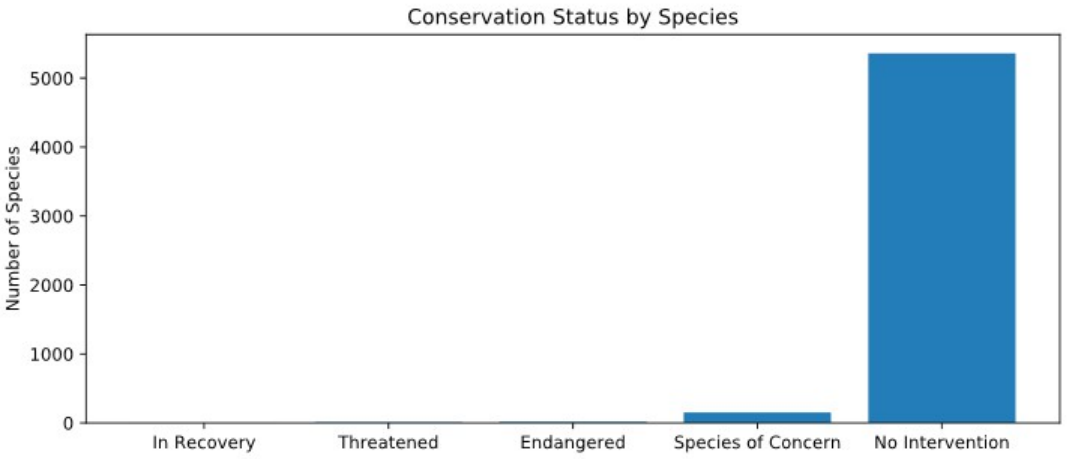
Capstone Project

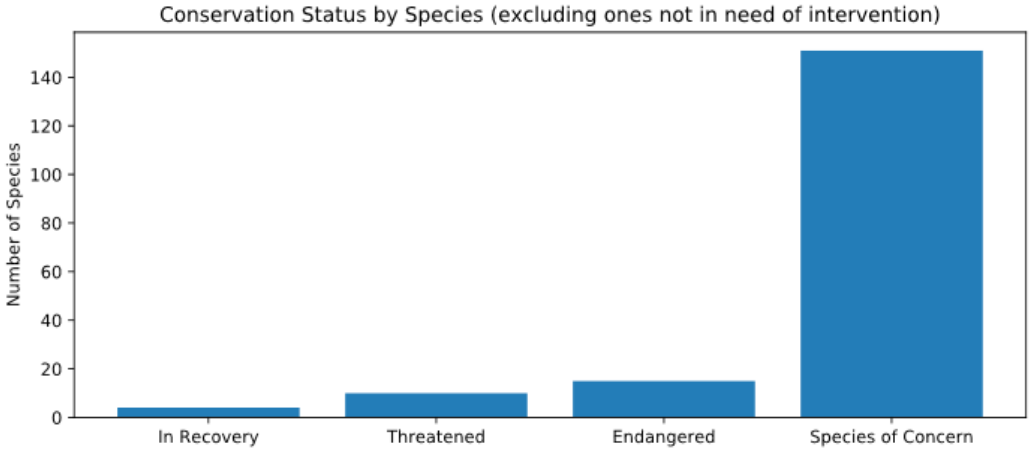
Welcome to my Introduction to Data Analysis Capstone Project

Biodiversity study

According to obtained data, we have tracked 5543 unique species across 5 different vulnerability categories ('Species of Concern', 'Endangered', 'Threatened', 'In Recovery' and those that require no intervention):

Conservation Status	Number of unique species:
In Recovery	4
Threatened	10
Endangered	15
Species of Concern	151
No Intervention	5363





Species description

All species in the database were assigned to one of the following 7 categories, based on type:

'Mammal', 'Bird', 'Reptile', 'Amphibian', 'Fish', 'Vascular Plant', or 'Nonvascular Plant'.

It can be clearly seen that different number of organisms in each category reqieres protection. We have been asked to see whether this is a statistically significant difference.

Category:	Not protected:	Protected:	Percent of protected:
Amphibian	72	7	8.86%
Bird	413	75	15.37%
Fish	115	11	8.73%
Mammal	146	30	17.05%
Nonvascular Plant	328	5	15.02%
Reptile	73	5	6.41%
Vascular Plant	4216	46	1.08%

Test results

According to chi-square test and assuming 95% certainty, only the difference in number of protected species between mammals and reptiles is statistically significant:

Pvalue for mammals vs reptiles: 0.0383555902297 < 0.05

However, the differences between birds and reptiles and between mammals and fish should also be noted as they are on the verge of significance:

Pvalue for birds vs reptiles: 0.0531354223215 > 0.05

Pvalue for mammals vs fish: 0.0561483484489 > 0.05

All other tested results were not statistically significant:

Pvalue for mammals vs birds: 0.687594809666 > 0.05

Pvalue for mammals vs amphibians: 0.127576696932 > 0.05

Pvalue for birds vs amphibians: 0.175936132496 > 0.05

Pvalue for vascular vs nonvascular plantss: 0.662341949138 > 0.05

Recommendations

Not all species are endangered to the same extent: extra care should be taken when dealing with mammals.

Sheep data

There are 3 species of sheep observable across all our National Parks:

- A common *Ovis aries* (feral domestic sheep) that is neither protected nor endangered but can still transmit disease
- *Ovis canadensis* (Bighorn Sheep) that is a species of concern
- *Ovis canadensis sierrae* (Sierra Nevada Bighorn Sheep), an endangered species

Their total sighting per National Park looks as follows:

Species:	Park name:	Observations:	Common Name:	Conservation status:
Ovis canadensis	Yellowstone	219	Bighorn Sheep, Bighorn Sheep	Species of Concern
Ovis canadensis	Bryce	109	Bighorn Sheep, Bighorn Sheep	Species of Concern
Ovis canadensis	Yosemite	117	Bighorn Sheep, Bighorn Sheep	Species of Concern
Ovis canadensis	Great Smoky	48	Bighorn Sheep, Bighorn Sheep	Species of Concern
Ovis canadensis sierrae	Yellowstone	67	Sierra Nevada Bighorn Sheep	Endangered
Ovis canadensis sierrae	Yosemite	39	Sierra Nevada Bighorn Sheep	Endangered
Ovis canadensis sierrae	Bryce	22	Sierra Nevada Bighorn Sheep	Endangered
Ovis canadensis sierrae	Great Smoky	25	Sierra Nevada Bighorn Sheep	Endangered
Ovis aries	Yosemite	126	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention
Ovis aries	Great Smoky	76	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention
Ovis aries	Bryce	119	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention
Ovis aries	Yellowstone	221	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention

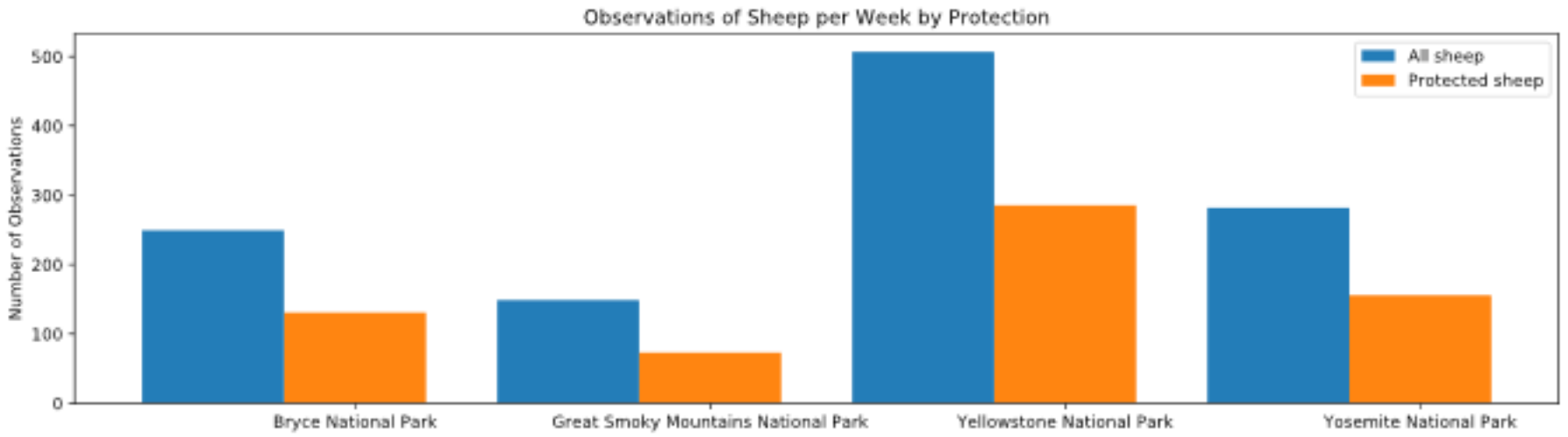
Sheep sightings

Total sheep observations per week per Park:

Park name:	Number of observations:
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite National Park	282

Sheep observation per park accounting for vulnerability:

Park name:	Is vulnerable?	Number of observations:
Bryce National Park	No	119
Bryce National Park	Yes	131
Great Smoky Mountains	No	76
Great Smoky Mountains	Yes	73
Yellowstone National Park	No	221
Yellowstone National Park	Yes	286
Yosemite National Park	No	126
Yosemite National Park	Yes	156



Evaluation of disease control attempts

We currently estimate that 15% of all sheep are affected by foot and mouth disease. If we wish to check whether the results of disease control are significant (i.e. whether we have managed to reduce the number of affected animals by 33%, to 10% of total population), then assuming we aim for 90% certainty, we would require a sample size of 890 animals to be tested.

Since the typical sheep sightings equal to about 507 in Yellowstone National Park and about 250 in Bryce National Park, we would expect the tests to take roughly 3.5 weeks for Bryce and 2 weeks for Yellowstone (or 3.6 and 1.8 weeks respectively if we want to be precise).