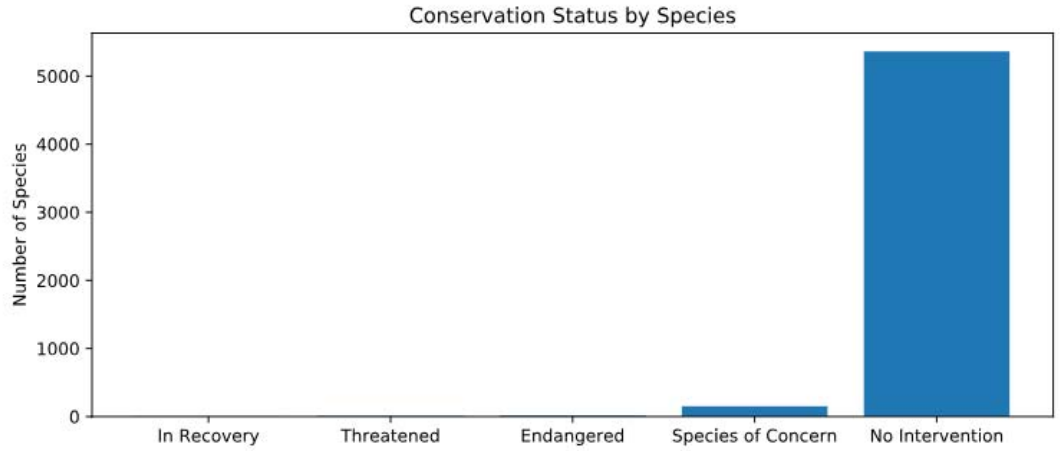
Biodiversity for the National Parks

Bryan Dominguez

Data in Species_info.csv

- When I first loaded the data I realized that there was a total of 5,541 different species.
- Looking at the conservation count I began to realize that there was a high number of species of concern for 151. This was followed by 15 endangered species.
- There are a good number of species that are not needing some sort of protection.
- However, that doesn't mean there isn't any work that needs to be done!



Significance Calculation: Endangered status between different categories of species

- During this phase we analyzed the percent protected of various endangered species to see if mammals are more likely to be endangered than birds.

	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

Significance Calculation (Continued)

- The null hypothesis was that the difference is due to chance
- With the dataset provided we knew that we have two categorical variables that come from the same population. In order to find out whether the cause of their endangerment is independent I(we) decided to use a Chi-Square test.
- Pval = 0.687594809666 which has no significant difference between the percentages of protected birds and mammals.
- Pval_reptile_mamal = 0.0383555902297 did have a significant difference because the pval was less than 0.05.
- Therefore, I concluded that depending on the types of species there are some that are more likely to be endangered than others.

My recommendation for conservationists

- Since we were able to tell that some species are more likely to be endangered then others I would suggest conservationists to further study what variables or attributes are causing some species to have a greater chance of being endangered.
- Being able to know the variables that may be causing this disadvantage can lead to programs with greater ability to help species that endangered and those species with the chance of being endangered.
- The ability to refocus their energy and resources on those that may need the support more can hopefully improve program efficiencies. I would caution though that analyzing what is done well that is keeping those certain species not endangered is important so that they don't lose sight of the whole groups needs.

Sample size determination for the foot and mouth disease study

- The question that was posed to us was what would be the appropriate sample size for each location to significantly see a 5% drop in observed cases of foot and mouth disease in sheep.
- I used the baseline of 15 % provided by the historical data given by the scientists. Then calculated the Minimum Detectable Effect by doing a small calculation:

Minimum_detectable_effect = (100 * 5)/15 – since the scientist want to see a reduction of 5 %.

- Using the following variables I was able to come up with the sample size per variant of 870.

Sample size determination (Continued)

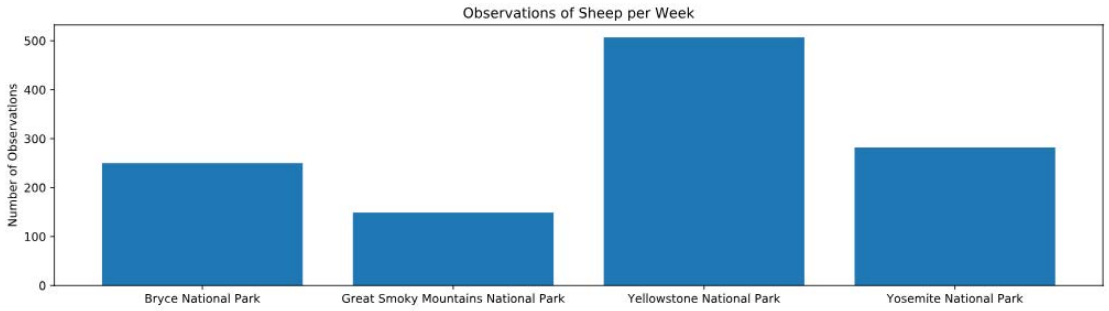
- Knowing that the sample size per variant in order to be significant would have to be 870. I would now need to calculate the amount of time it would take to see those 870 for each location. Therefore, I took the data recorded of sightings by location for the past 7 days to calculate how long each location would take to see 870.

- Yellowstone would take about one week.
- Bryce would take about two weeks.

Graphs for Foot and Mouth Reduction Effort

	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

	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



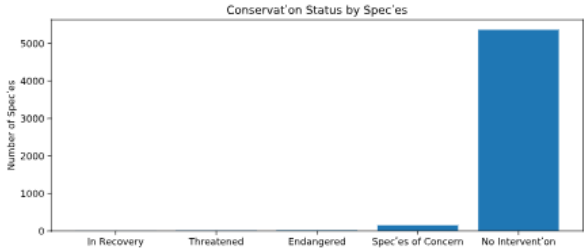

```
baseline = 15

minimum_detectable_effect = 100*5./15

sample_size_per_variant = 870

yellowstone_weeks_observing = sample_size_per_variant/507.

bryce_weeks_observing = sample_size_per_variant/250.
```

Graphs for Biodiversity Project

	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

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	Species of Concern	151
3	Threatened	10

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	No Intervention	5363
3	Species of Concern	151
4	Threatened	10



	category	is_protected	scientific_name
0	Amphibian	False	72
1	Amphibian	True	7
2	Bird	False	413
3	Bird	True	75
4	Fish	False	115

is_protected	category	False	True
0	Amphibian	72	7
1	Bird	413	75
2	Fish	115	11
3	Mammal	146	30
4	Nonvascular Plant	328	5
5	Reptile	73	5
6	Vascular Plant	4216	46

	category	not_protected	protected	percent_protected
0	Amphibian	73	7	0.087500
1	Bird	442	79	0.151631
2	Fish	116	11	0.086614
3	Mammal	176	38	0.177570
4	Nonvascular Plant	328	5	0.015015