Biodiversity Project

Yao Chen

How the Species_Info.Csv File Looks Like?

	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

- scientific_name: the scientific name of each species
- common_name: the common names of each species
- conservation _statues: the species conservation status

Some Information of species_info.csv

- 5541 different species in the DataFrame

- The species types include: Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, Nonvascular Plant

- The conservation status and fixed conservation status:

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

the Conservation Status

- 5363 species are without any intervention
- 151 species have the possibility to distinct or need conservation in the future
- 15 species are in seriously at risk of extinction
- 10 species are vulnerable to endangerment in the near future
- 4 species are in recovery(formerly Endangered, but wold not distinct in the near future)

Conservation Status by Species
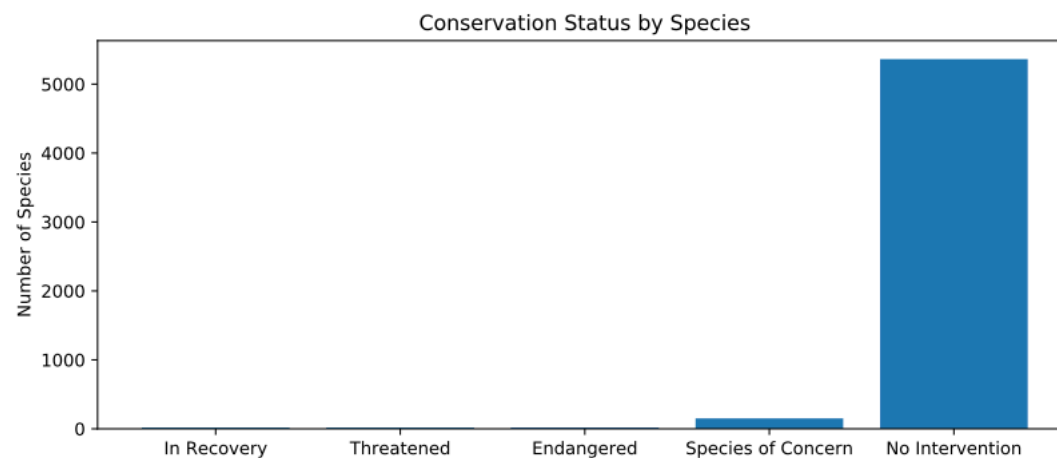
Most

Info of Endangered Species

- Data of every species

	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
	SUM	5363	179	0.033377

- In average, 3.34% of the species are protected
- In Chi-Squared Test:
  - On mammal and bird, the p value is 0.687594809666, no significant difference
  - On reptile and mammal, the p value is 0.0383555902297, significant difference

Observations.Csv File & Sheep

- Merge Species_info.csv & Observations.csv File

	category	scientific_name	common_names	conservation_status	is_protected	is_sheep	park_name	observations
0	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True	Yosemite National Park	126
1	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True	Great Smoky Mountains National Park	76
2	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True	Bryce National Park	119
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True	Yellowstone National Park	221
4	Mammal	Ovis canadensis	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True	Yellowstone National Park	219

- Observe sheeps by park in last 7 days

	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
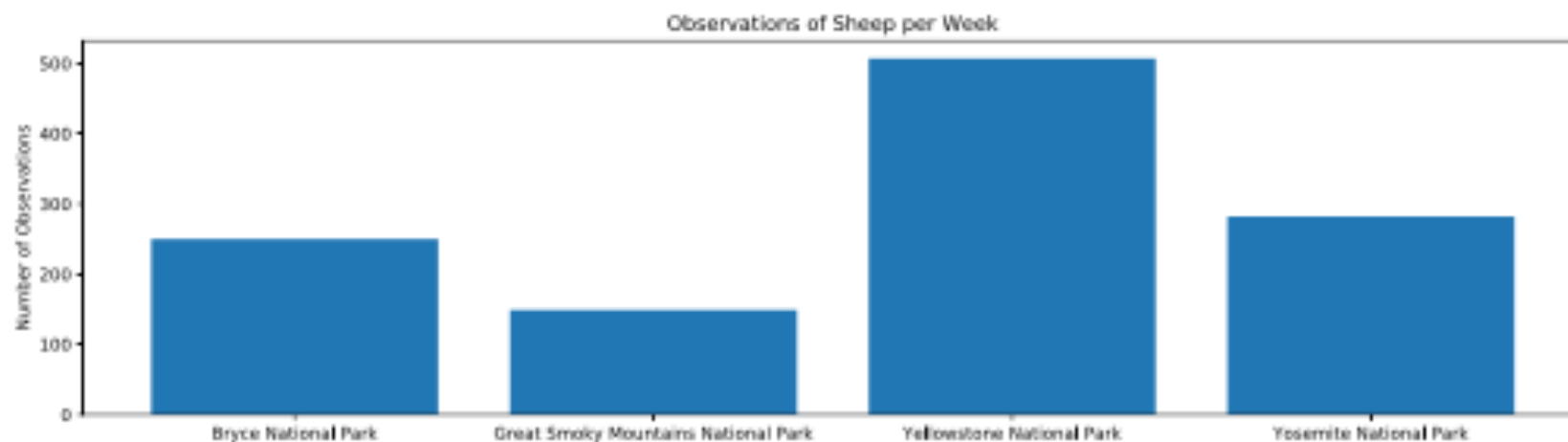
Foot and Mouth Reduction Effort

- baseline = 15

- minimum detectable effect is 33.3333333333

- sample size per variant is 870(significance is 90%)

- To observe enough sheeps, 1.71597633136 weeks are needed at Yellowstone National Park

- To repeat the experiment and observe enough sheeps, 3.48 weeks are needed at Bryce National Park