Capstone Option 2: Biodiversity for the National Parks

By Tara Kamal Ahmadi

Course Introduction to Data Analysis

Jul 10, 2018 – Oct 2, 2018

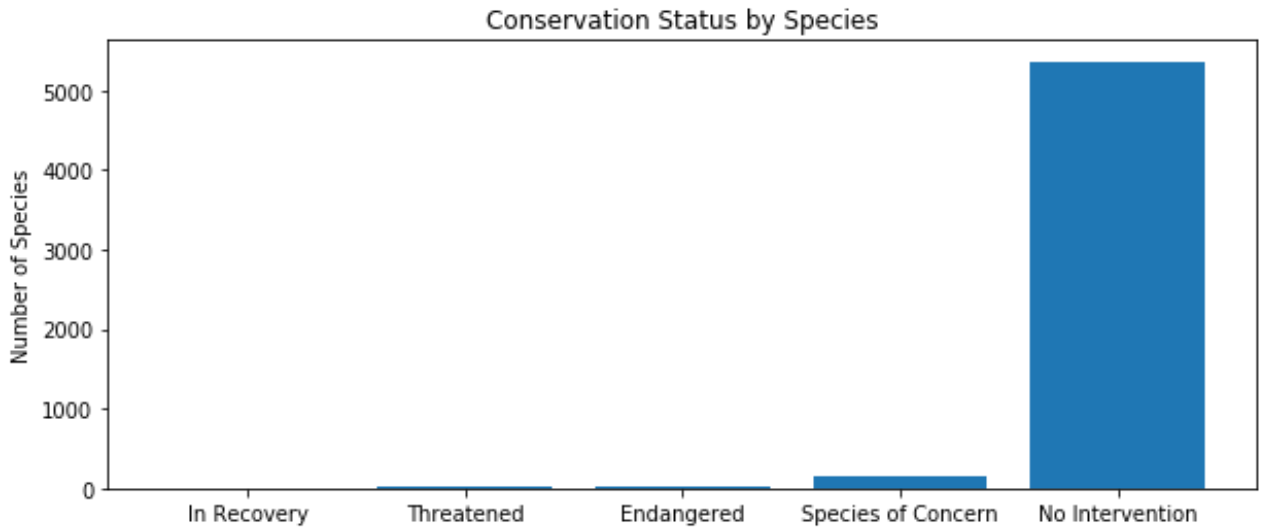
Data Description

- Dataset (species_info.csv) includes :
  - category
  - scientific_name
  - common_names
  - conservation_status
- The park has 5541 different types of species.
- The park has 'Mammal‘, 'Bird‘, 'Reptile‘, 'Amphibian‘, 'Fish‘, 'Vascular Plant‘, and 'Nonvascular Plant‘
- The conservations statuses labeled as 'Species of Concern' 'Endangered‘, 'Threatened‘, 'In Recovery‘.

Conservation Status by Species

- Only (n=151) are in need of conservation. There are only 4 species that are not in danger of extinction.

conservation_status	Total Number
Endangered	15
In Recovery	4
No Intervention	5363
Species of Concern	151
Threatened	10



Chi square test

null hypothesis here is that this difference was a result of chance.

1. Are mammals more likely to be endangered than birds? No the difference is not significant (p= 0.68)

2. Are mammals more likely to be endangered than Reptile? Yes, the difference is significant (p =0.03)

category	not_protected	protected	percent_protected
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015
Reptile	73	5	0.064103
Vascular Plant	4216	46	0.010793

CONCLUSION = certain types of species *are* more likely to be endangered than others.

Data Description

- Data Frame Observations.csv contains:
  - scientific_name
  - park_name
  - number of observations

In search of sheep:

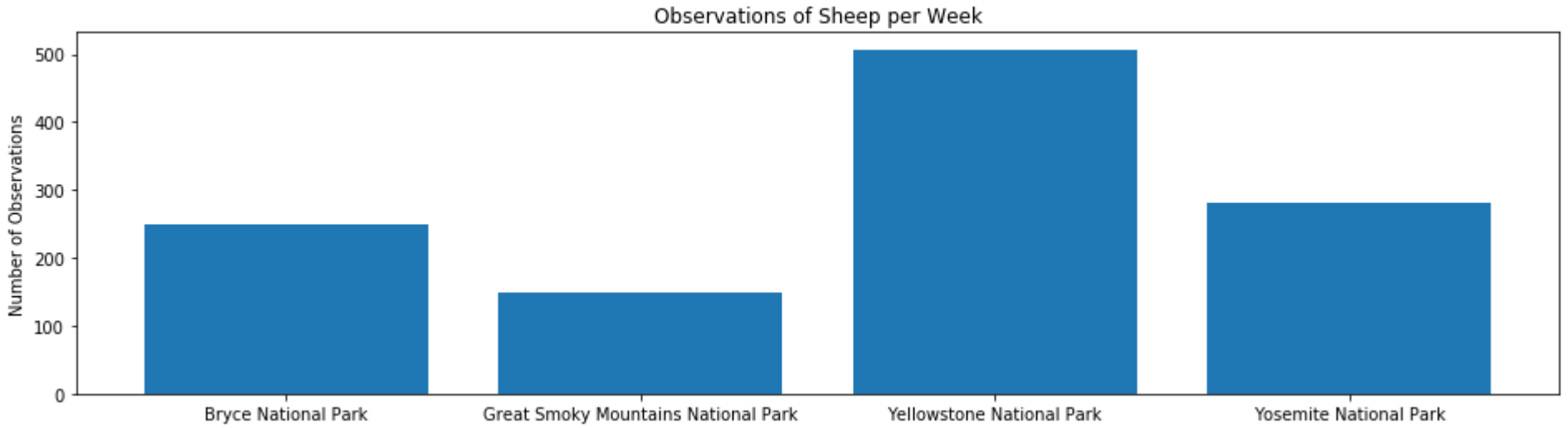
- Designate species which are sheep and mammals:

category	scientific_name	common_names	conservation_status	is_protected	is_sheep
Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention	False	True
Mammal	Ovis canadensis	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
Mammal	Ovis canadensis sierrae	Sierra Nevada Bighorn Sheep	Endangered	True	True

Where these sheep are locating?

park_name	Total number of sheep observations in last 7 days
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite National Park	282

- The most common location for sheep are Yellowstone National Park location.
- The least common location for sheep is Great Smoky Mountains National Park.



Foot and Mouth Reduction Effort - Sample Size Determination

- baseline = 15%
- minimum_detectable_effect= 33.33
- Statistical significance = 90 %
- **Sample size = 870**

- Thus they need to observe sheep for about 1.7 week for Yellowstone National Park and about of 3.5 weeks for Bryce National Park to assure that 5% drop was significant.