Biodiversity Capstone Project

Andrew C. Serino

The State of the National Parks

Exploring the Well-Being of Present Species

Species Data: Overview

- **Information Contained in species dataset:**

category	scientific_name	common_names	conservation_status
0	Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole
1	Mammal	Bos bison	American Bison, Bison
2	Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Domesticated Cattle
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)
4	Mammal	Cervus elaphus	Wapiti Or Elk

- **Entire dataset contains 5541 unique species of:**
  - **Mammal**
  - **Bird**
  - **Reptile**
  - **Amphibian**
  - **Fish**
  - **Vascular Plant**
  - **Nonvascular Plant**

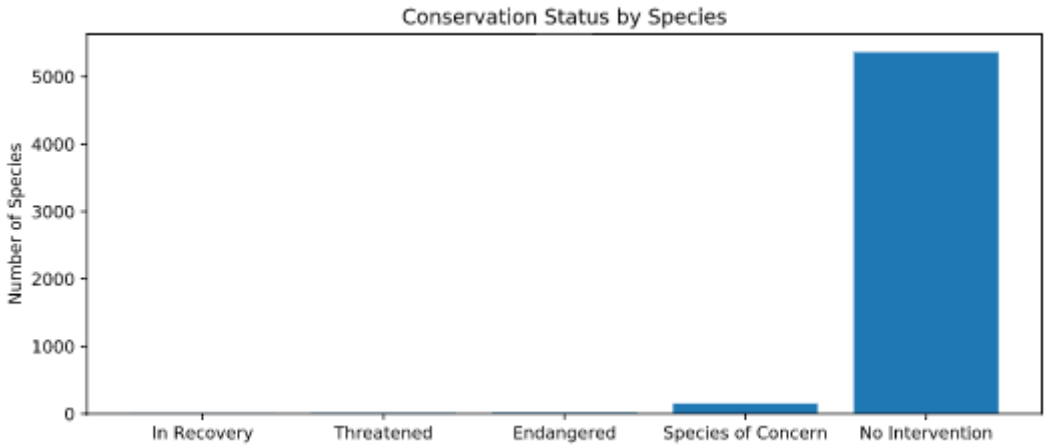
- **With conservations status of**
  - **Not provided**
  - **Species of Concern**
  - **Endangered**
  - **Threatened**
  - **In Recovery**

Analyzing Conservation Status

- **Observations:**
  - **At first glance, data indicates the population of a large number of species is in decline.**
  - **However, inclusion of the entire dataset reveals a large majority of species requires no intervention.**

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

- **Conclusions and recommendations:**
  - **Several species have been identified that require help, recommend immediate action to ensure survival**
  - **Observations provide useful for tracking the state of at-risk species, recommend continuing current tracking efforts**



Investigating Endangered Species

- **Are certain types of species more likely to be endangered?**
  - **To explore this question, I analyzed whether or not a species required 'No intervention'.**

is_protected	category	not_protected	protected	percent_protected
0	Amphibian	73	7	0.0875
1	Bird	442	79	0.1516
2	Fish	116	11	0.0866
3	Mammal	176	38	0.1776
4	Nonvascular Plant	328	5	0.0150
5	Reptile	74	5	0.0633
6	Vascular Plant	4424	46	0.0103

- **From these percentages, mammals appear to be most likely to be endangered, with birds close behind. But is it statistically significant?**

Significance of Difference Between Mammals and Birds

- **Null hypothesis is that the difference in endangered rates between mammal and bird is due to chance**

category	protected	not-protected
Mammal	176	38
Bird	442	79
Reptile	74	5

  - **Chi-squared test for significance of difference between mammals and birds:**
    - **P-value ≈ 0.45 > 0.05 , insignificant!**

- **Difference between mammals and birds is *not* significant.**

  - **Chi-squared test for significance of difference between mammals and reptiles:**
    - **P-value ≈ 0.02 < 0.05 , significant!**

- **The difference between mammals and reptiles *is* significant.**

Counting Sheep

Observations Data: Overview

- **Information contained in observations dataset:**

	scientific_name	park_name	observations
0	Vicia benghalensis	Great Smokey Mountains National Park	68
1	Neovison vison	Great Smokey Mountains National Park	77
2	Prunus subcordata	Yosemite National Park	138
3	Abutilon theophrasti	Bryce National Park	84
4	Githopsis specularioides	Great Smokey Mountains National Park	85

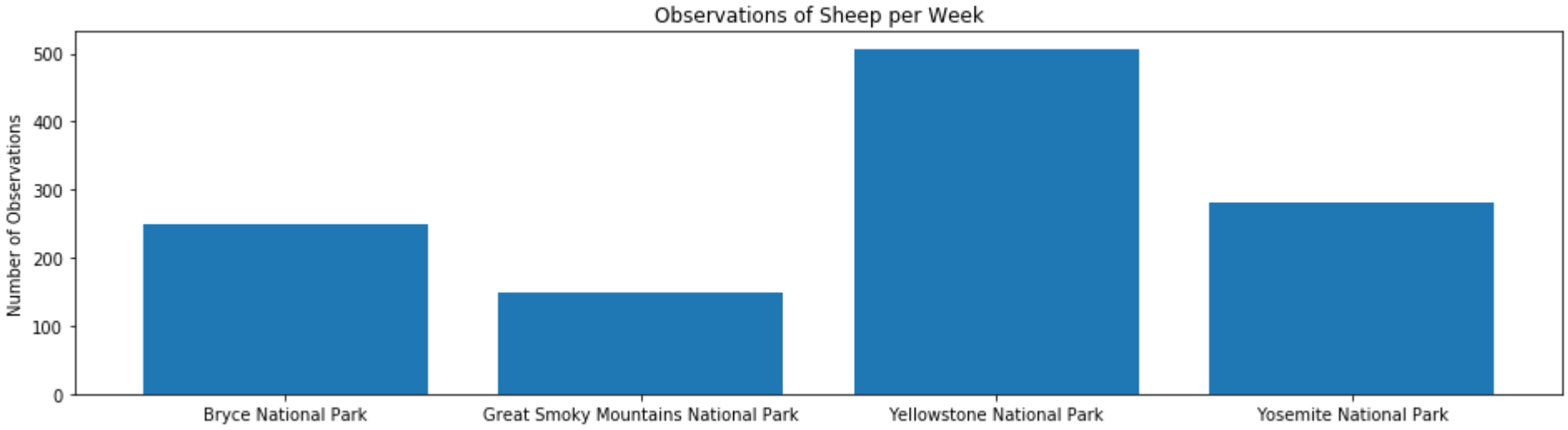
- **Dataset catalogs how many each species was sighted at different national parks.**

- **Using this dataset and the species dataset, we can identify all the species of 'sheep', how many times they were seen in each park, and whether the sheep falls under the 'is protected' category.**

Counting Sheep

- **Total number of sheep across all parks over the past 7 days**

	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountaains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282



Foot and Mouth Disease Reduction Program

- 15% of sheep at the Bryce National park have foot and mouth disease
- Yellowstone National Park rangers have been running a program to reduce the occurrence of foot and mouth disease among sheep
- We were tasked with identifying whether it is possible to track the success of this program and how long it would take to implement observations

- Tracking criteria:
  - 5 percentage point reduction from baseline of 15%
  - 90% statistical significant
  - Minimum detectable effect is 33.33%
    - 100*5/15
  - Sample size per variant is 870*

- From this information, it would take :
  - 1.7 Weeks to observe enough sheep at Yellowstone National Park
  - 3.48 weeks to observe enough sheep at Bryce National Park

*source: *Optimizely.com/sample-size-calculator/*