Biodiversity of the National Parks

Presentation by Aaron Bauer
Codecademy Data Analysis Pro-Intensive

Inspecting the Data

National Parks are host to thousands of animal species

The data shows there are at least **5541 different species** within our national parks

Data recorded the following characteristics

- The species type
- The scientific name of each species
- The common names of each species
- The species conservation status

Conservation Status Breakdown

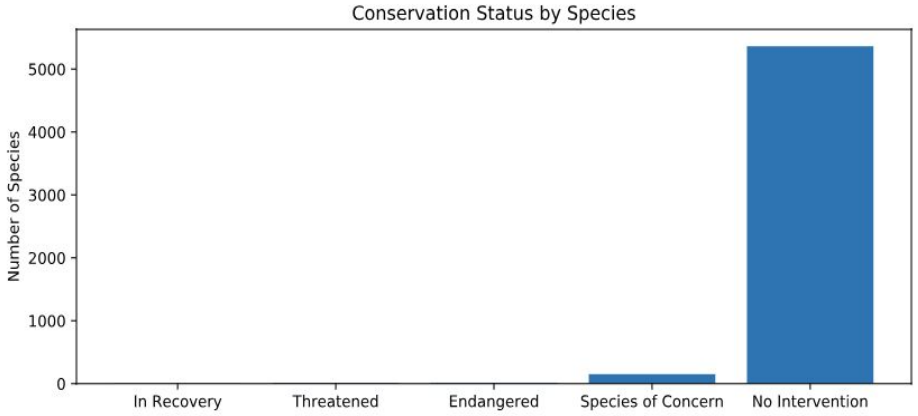
If we break down the **5541 different species by scientific name,** we see the following conservation status grouping

Conservation Status	Count
Species of Concern	151
Endangered	15
Threatened	10
In Recovery	4
Not Threatened	5363

What have we learned so far?

Good news: the majority of species are not threatened

Bad news: 176 species are at risk (about 3%)



Species Type Breakdown

We can further break down the data based on species category to gain a better sense of the groups present in our parks.

- Mammal
- Bird
- Reptile
- Amphibian
- Fish
- Vascular Plant
- Nonvascular Plant

Are certain types of species more likely to be threatened?

Endanger Risk by Species

Type	Not Protected	Protected	Percent 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	1.50%
Reptile	73	5	6.41%
Vascular Plant	4216	46	1.08%

Testing for Risk Significance

Mammals and birds have the highest protection rate overall, but are mammals more at risk?

- Mammals have a protection rate of **17.05%**
- Birds have a protection rate at **15.37%**

P value testing between mammal and birds returns a result of **0.688.**

Our null hypothesis is false; ***mammals are not more likely to be endangered than birds and is a result of chance***.

What about other species comparisons?

Testing for Risk Significance

Mammals have a higher protection rate over reptiles. Is this significant?

- Mammals have a protection rate of **17.05%**
- Reptiles have a protection rate at **6.41%**

P value testing between mammal and reptiles returns a result of **0.038**.

Our null hypothesis is true; ***certain species are more likely to be endangered than others and is not a result of chance.***

How do mammals compare to the remaining species?

Mammals vs. All Other Species

Species	P-value vs. Mammals	Significant? (<.05)
Bird	0.688	No
Reptile	0.038	Yes
Amphibian	0.128	No
Fish	0.056	No*
Vascular Plant	1.44	No
Nonvascular Plant	1.44	No

**Technically not significant, but very close to the threshold*

Recommendations for Conservation

Prioritize protection for reptiles species before mammals

Fish are very close to the significance threshold vs. mammals and should be closely observed

Plant and amphibian species do not appear to be at risk compared to mammals, so effort should be more focused on other species at risk

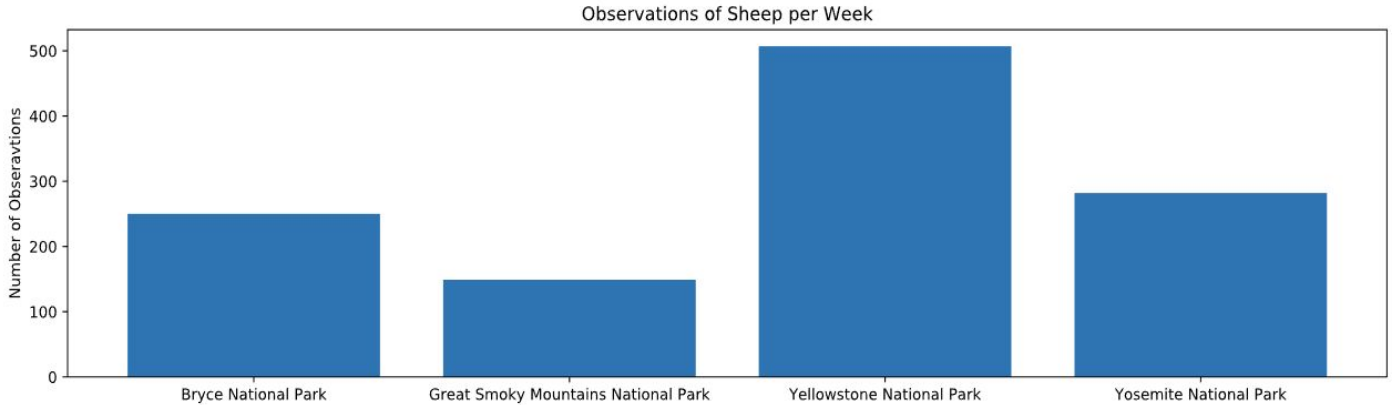
Mammals should not be neglected, but some effort could be re-focused to other species

Foot and Mouth Disease: In Search of Sheep

Seven days worth of sheep observation data within our parks has been collected to help reduction efforts for Foot and Mouth disease.

Park	Observations
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite National Park	282
Total Observations	1188

Visualizing the Observations



Determining Observation Sample Size

Park Rangers are targeting a %5 decrease in Foot and Mouth across all national parks

15% of sheep at Bryce National Park have foot and mouth disease last year. This will be our base line.

- Baseline: 15%
- Statistical significance: 90% (default)
- Minimum detectable effect: 33.33%

Sample size target is ***870 sheep observations per park***

Foot and Mouth Reduction Effort

We've determined our target sample size, but how long will it take scientists to collect this data for each national park to determine if targeted 5% reduction is working?

Park	Weeks
Bryce National Park	3.48
Great Smoky Mountains National Park	5
Yellowstone National Park	1.72
Yosemite National Park	3

Thank you.

Questions?