Species Conservation in the National Parks

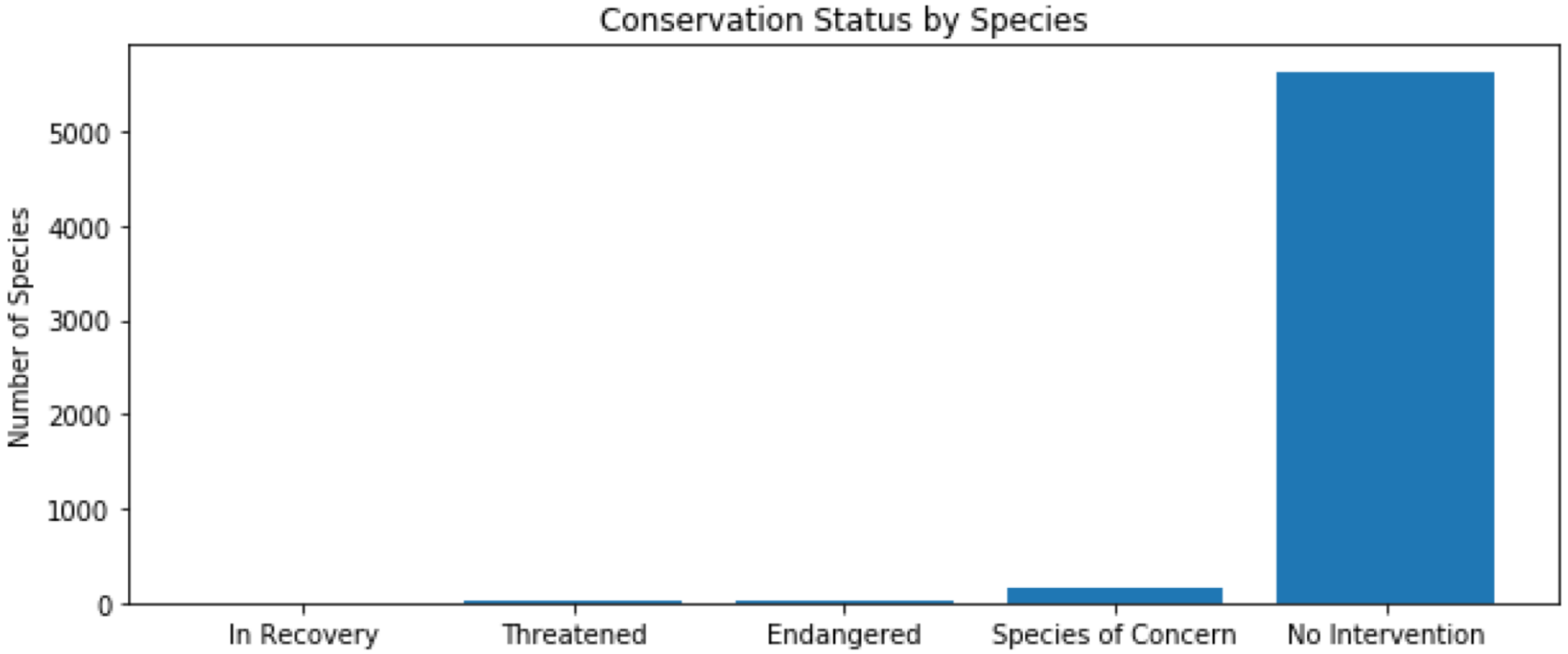
CD Jones

Codecademy Introduction to Data Analysis

03/02/2018

Species Data

- species_info.csv is a dataset listing species:
  - category: 'Mammal', 'Bird', 'Reptile', 'Amphibian', 'Fish', 'Vascular Plant', 'Nonvascular Plant’
  - scientific_name: actual scientific name of species
  - common_names: list of common names associated with the scientific name
  - conservation_status: nan, 'Species of Concern', 'Endangered', 'Threatened', 'In Recovery’
- There are over 5,000 species listed, and close to 200 of them have a protection status and need intervention
- Plants are the least likely categories to require intervention (~1% of species protected), while mammals and birds are the most likely (>15% protected)


Conservation Status by Species
Number of Species
5000
4000
3000
2000
1000
0
In Recovery
Threatened
Endangered
Species of Concern
No Intervention

Significance Calculations

- Mammals have a higher percentage of species that are protected (17.05%) than birds (15.37%)
  - These are categorical observations (averages and standard deviations cannot be calculated)
  - We could potentially compare all categories, which is a value 2 or greater
    - We will use a chi squared test
  - P-value that results from comparing mammals and birds is 0.7! **It is NOT actually more likely that a mammal need protection than a bird!**
- Reptiles need protection 6.4% of the time
  - P-value here is 0.038, which is less than 0.05. **Reptiles are actually less likely to need protection than mammals!**

Recommendation

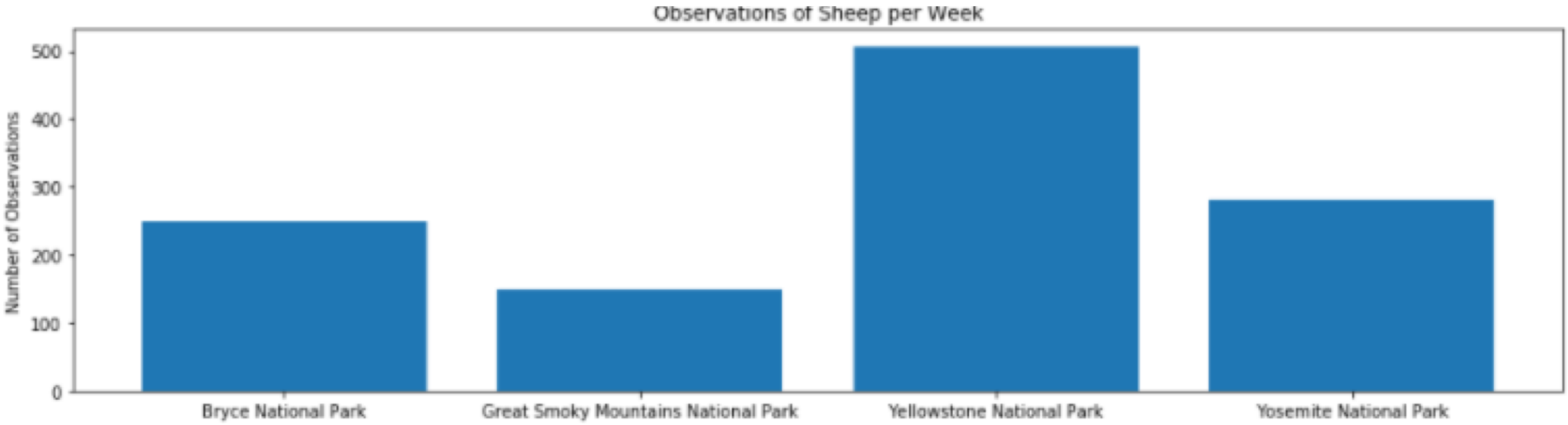
- To have the greatest impact, the categories of species that have the greatest protection needs should be focused on FIRST
- Birds and Mammals are equally at risk, so no prioritization between these categories is justified
- Plants can be ignored

- **Focus observations on mammals and birds**

Foot and Mouth in Sheep

How many sheep do we need to observe?

- We have a known value of 15% of sheep in Bryce National Park infected with foot and mouth; this is our baseline
- An effort has been made in Yellowstone to reduce the disease, but a 5% change is needed to know if there has been an impact
- The minimum detectable effect is the change of 5% from 15%, or a 33.3% change
- We can accept 90% confidence, so an online calculator shows **we need 510 observations of sheep** to know if our reduction effort has been significant
  - In Bryce we need 2 weeks of observation time
  - In Yellowstone we only need one week


Observations of Sheep per Week
Number of Observations
0
100
200
300
400
500
Bryce National Park
Great Smoky Mountains National Park
Yellowstone National Park
Yosemite National Park