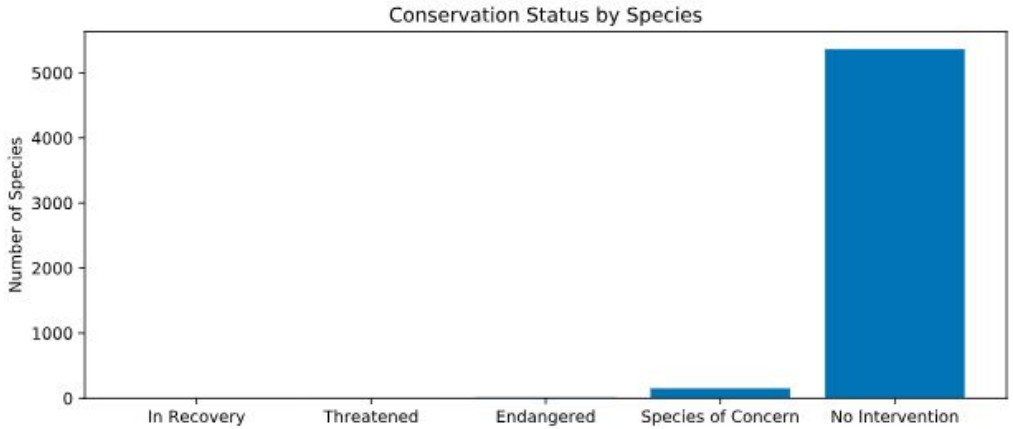
Biodiversity for the National Parks

Capstone Project

by Nic C

Species Data

- In our first data set, ‘species.csv’, we can observe information on 5,543 species living in the national parks system including their scientific and common names
- These species span 7 categories: 'Mammal', 'Bird', 'Reptile', 'Amphibian', 'Fish', 'Vascular Plant', and ‘Nonvascular Plant’
- These species are also classified by their conservation status:
    - Endangered 15
    - In Recovery 4
    - No Intervention 5363
    - Species of Concern 151
    - Threatened 10



Significance Calculations

- Chi-squared tests for significance were performed to better understand the differences between the protected and not protected statuses of various species:
  - Mammal vs Bird = 0.687594809666, **not** statistically significant at the 0.05 p-value
  - Mammal vs Reptile = 0.0383555902297, statistically significant at the 0.05 p-value

Recommendations for Conservation

Based upon the data provided, birds and mammals should be the main focus of conservation efforts because they have the largest overall percentage of protected species

The calculations found that the mammals are significantly more likely to be protected than reptiles

Plants (vascular and non) are safest

category	not_protected	protected	percent_protected
Amphibian	72	7	8.860759
Bird	413	75	15.368852
Fish	115	11	8.730159
Mammal	146	30	17.045455
Nonvascular Plant	328	5	1.501502
Reptile	73	5	6.410256
Vascular Plant	4216	46	1.079305

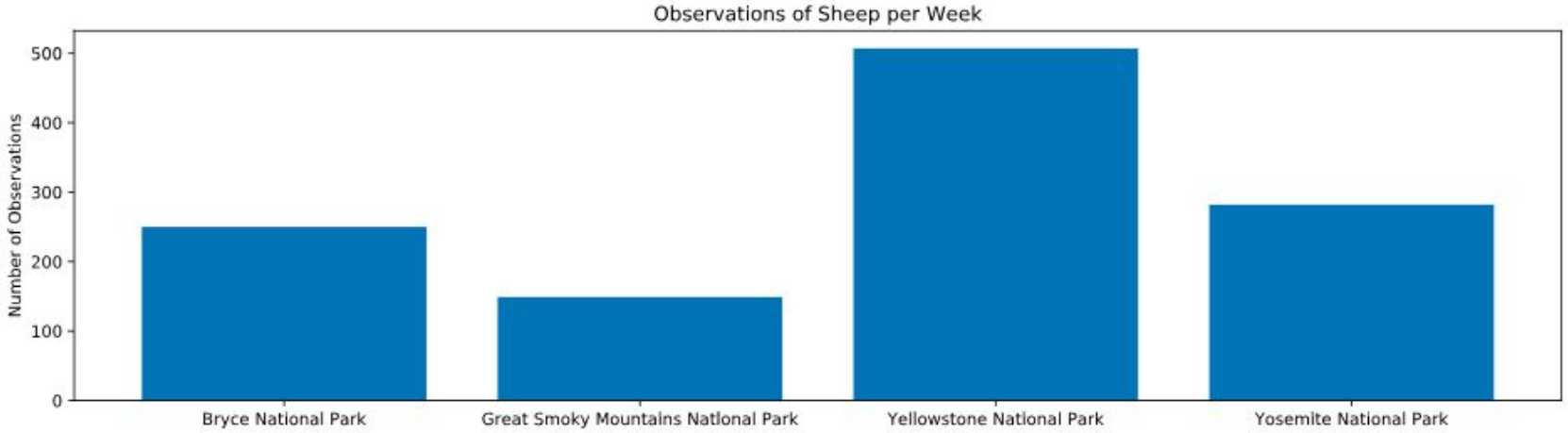
Observations Data

Conservationists have been recording sightings of different species at several national parks for the past 7 days. Their observations have been stored in our second data set called ‘observations.csv’.

By cross-referencing the observations data set with the species data, we can narrow our sheep species to three:

1. Domestic Sheep
2. Bighorn Sheep
3. Sierra Nevada Bighorn Sheep

Sheep Sighting by Park



Foot and Mouth Program

- Yellowstone National Park is piloting a program to reduce foot and mouth disease and want to see if it is is reducing cases by 5%
- The only baseline information that is present is from Bryce National Park, where 15% of the sheep have the disease
- Based upon the available data:
  - The baseline is 15%
  - The minimum detectable difference is 33.333%
  - The level of statistical significance is 90%
- Therefore, our sample size is 870 observations
- In Yellowstone, this will need 1.72 weeks (~12 days) and 3.48 weeks (~24 days) in Bryce

Thanks!