Analysis of Biodiversity in National Parks

Jonathan Surette

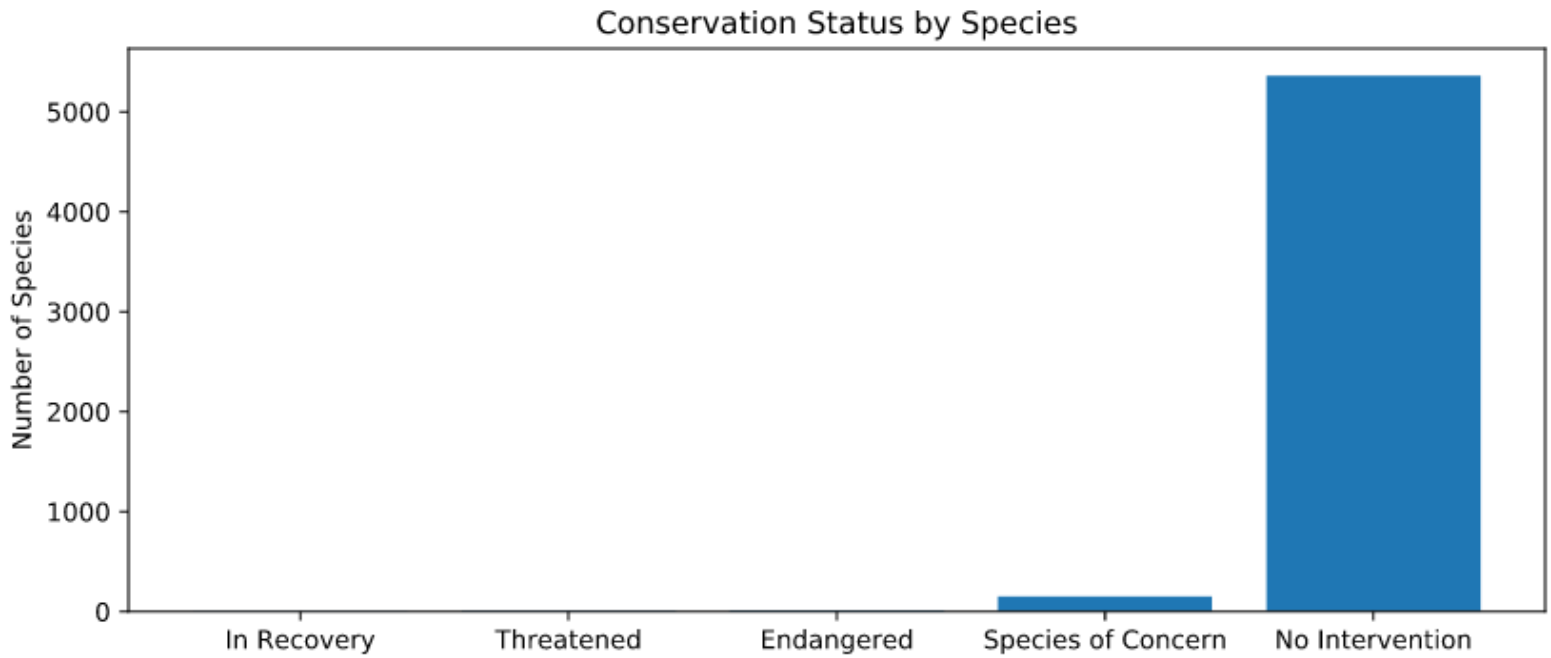
Species_info data

- Species_info.csv oringally contained fields such as the evolutionary class of the animal (mammal, reptile, etc.)
- Data also contained information on the scientific and common names for each entry

Working with species_info.csv

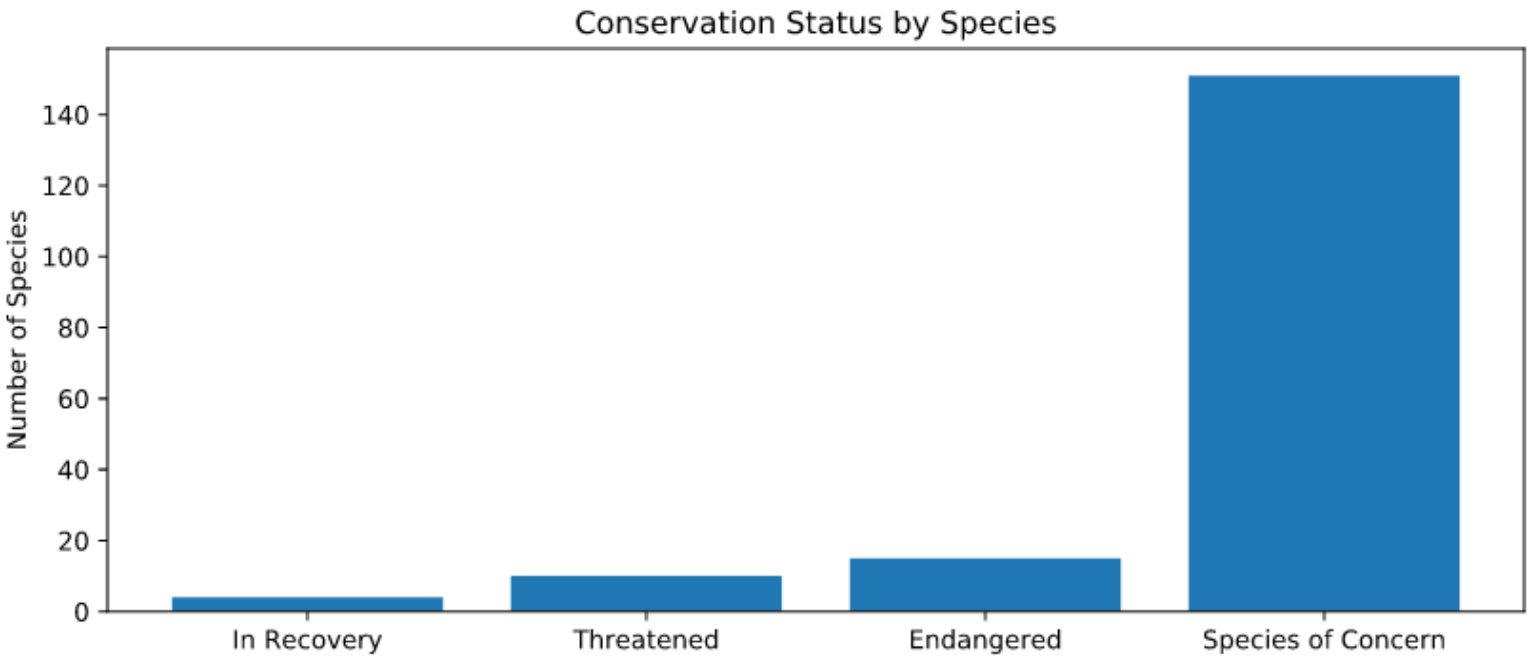
- Step-wise process used to create off-shoot dataframes, where additional fields could be added to the data and then analyzed
  - E.g. counting animals by conservation status, then comparing class of animals with conservation status to determine whether certain types of animals were more likely to be endangered (see next slide)

Protection status by volume overall



- Graph shows the vast majority of species are not considered to be in one of the at-risk conservation statuses

Range of protected statuses



- Graph shows a small volume of species in recovery
- Majority of species fall into “of concern” status

Protected Status by Category

Category	Percent Protected
Amphibian	8.86%
Bird	15.37%
Fish	8.73%
Mammal	17.04%
Nonvascular Plant	15.02%
Reptile	6.41%
Vascular Plant	1.08%

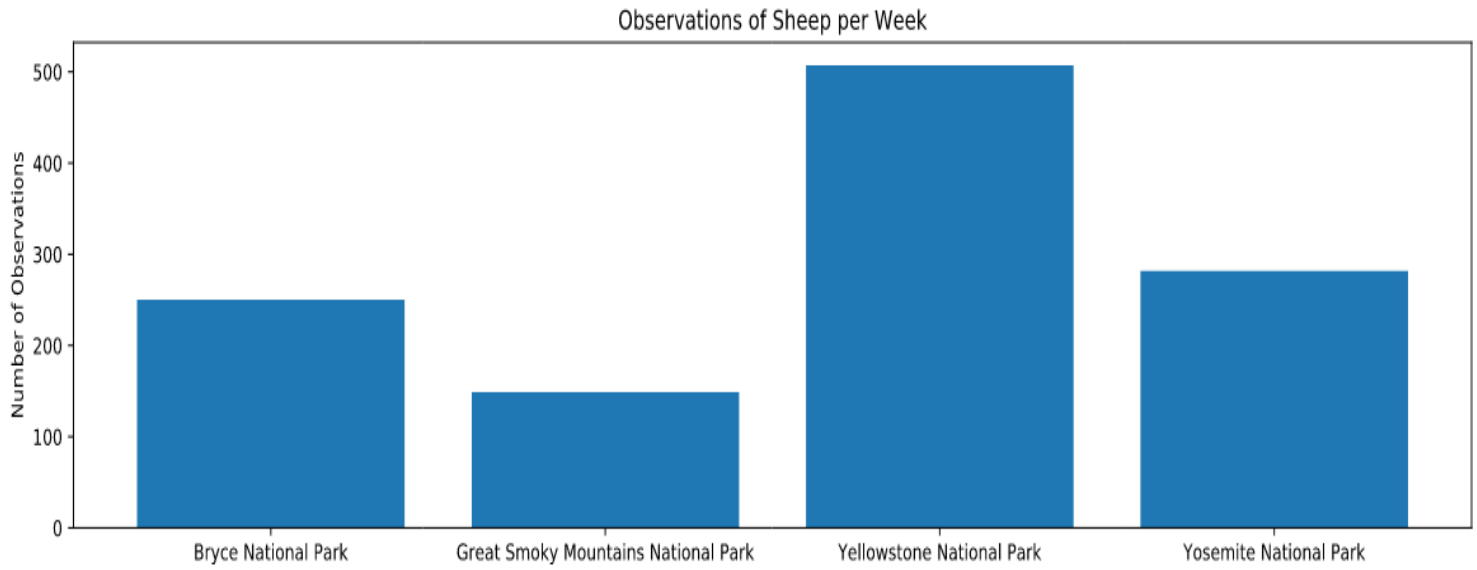
Protected Status by Category

- Subsequent Chi-square analyses comparing rates of protected status of mammals compared to those of birds showed no significant difference between the two
- The same analysis DID show that mammals fall into protected statuses significantly more than reptiles

Conservation Recommendations

- Analysis determined a significant difference in the rates of at-risk conservation statuses between mammals and reptiles
- Major recommendation to conservationists would be to examine whether categories with higher conservation rates (mammals, birds, nonvascular plants) are preferentially offered protected statuses
- Either those categories suffer particular causes of population loss disproportionally and need environmental intervention-
  - Or certain categories (eg. Vascular plants, reptiles) have less popular appeal and so are not offered protected statuses as readily

Foot & Mouth Disease Study



- Study of foot & mouth disease in sheep populations of the above national parks

Sample Size Determination

- Based on observations from Bryce Nat'l Park last year, rate of foot & mouth among sheep populations was estimated to be 15%
- Scientists wanted the ability to determine reductions of the disease by at least 5% increments (~33% of the baseline = MDE)
- Given that minimum detectable effect and baseline, it was determined 520 observations would be needed to measure changes at a 90% confidence level

Sample Size Determination (continued)

- Based on the requirement of 520 observations, each park would require different windows of time to accurately sample populations
  - Yellowstone observations would last slightly over a week
  - Compared to Bryce National Park, where observations would most likely take over two weeks

Thank you!

- This class has been a great introduction to Pandas, numPy, sciPy, and SQL
- CodeAcademy staff have been courteous and helpful
- Thanks again – Jonathan Surette