

Biodiversity Capstone Project Codecademy Data Analysis Course

Tyler Redman

June 27, 2018

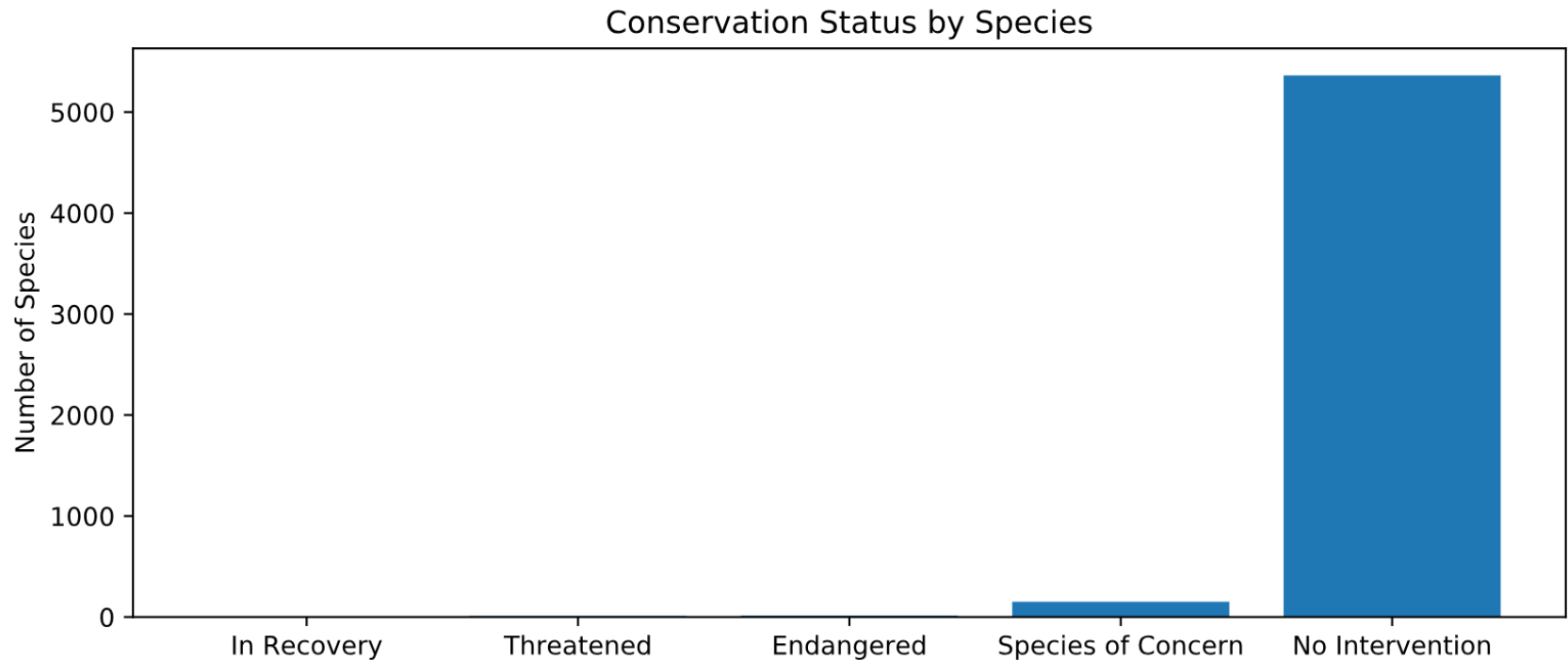
Data Set Intro

- The species_info.csv file has a collection of data for different animal species and their conservation status.
 - Comprehensive list covering various mammals, birds, reptiles, amphibians, fish, and plants.
- A quick look at this data shows that majority of species being monitored are not currently endangered
 - Only 180 out of 5541 requiring intervention as shown by Figure 1 on the following slide.
- Mammals and birds have the highest rate of protected species
- Nonvascular and vascular plants have the lowest rate of requiring protection.

Table 1: Protection Rate By Species Type

Category	Not Protected	Protected	% Protected
Amphibian	73	7	8.75%
Bird	442	79	15.16%
Fish	116	11	8.66%
Mammal	176	38	17.76%
Nonvascular Plant	328	5	1.50%
Reptile	74	5	6.33%
Vascular Plant	4424	46	1.03%

Figure 1 – Conservation Status



Significance of Protection Rates

- While both protection rates are high, the data in Table 1 suggests that mammals could be more endangered than birds.
- To determine if this result is significant, a Chi Squared test was performed using the following inputs:
 - Mammal: 38 protected, 176 not protected
 - Birds: 79 protected, 442 not protected
- The Chi Squared test returned a p-value of 0.4459
 - We can **not** reject the null hypothesis in this case.
 - **It is possible birds and mammals are equally endangered and this difference was entirely due to chance in the same of data.**
- The test was performed again to check against reptiles, which have a lower rate of protection required.
 - Reptiles: 5 protected, 74 not protected
- The second test returned a p-value of 0.0233
 - We **can** reject the null hypothesis and be certain that mammals are more endangered than reptiles!

Recommendation

- The data shows us that mammals and birds are definitely the most endangered species in this study.
 - However, it can not tell us which of two are the most endangered.
- Recommend focusing heavily on conservation efforts for both birds and mammals, and treating both groups as if they are equally as likely to require protection.

Foot and Mouth Reduction Effort – Sample Size Determination

- From past data, we know that there was a 15% occurrence rate of foot and mouth disease at Bryce Canyon.
 - This will be used as the baseline conversion rate.
- The scientists would like to be able to detect reductions of 5 percentage points.
 - This results in a minimum detectable effect of 33.33%
 - $(5 / 15) * 100 = 33.33$
- Plugging the above values into a sample size calculator with a 90% level of significance, we determine that we need to observe 870 sheep.
- At Yellowstone, where 507 sheep are observed each week, this sampling would take 1.72 weeks.
- At Bryce Canyon, where only 250 sheep are observed each week, this sampling would take 3.48 weeks.

Figure 2 – Sheep Observations

