(Biodiversity for the National Parks)

8 July 2018

Conservation: where should efforts be focused?

Overview of underlying data

- The dataset analyzed in this part of the project contains 5824 records
- Each record describes an individual species in terms of category, scientific name, common name(s), and conservation status
  - Categories consist of: Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, and Nonvascular Plant
  - Conservation status may be: Species of Concern, Endangered, Threatened, In Recovery, or none (No Intervention)
- 5541 different species (by unique scientific name) are represented
- Most species (> 4200, or ~75%) are vascular plants
- Most species (> 5300, or ~95%) have no conservation status (see *Supporting graphs (1)*)

Conservation by category

- Birds have the most protected species (75) while mammals have the highest percentage of protected species (~17%)
- The two are not statistically different in terms of the likelihood of requiring a conservation status:
- A chi-squared test that takes into account the numbers of protected and total species for mammals & birds cannot reject the null hypothesis (assuming standard p-value threshold of 0.05) that mammals and birds are equally likely to be protected
- However, there ARE categories that are statistically different; for example, mammals are statistically more likely to be protected than reptiles

Recommendation

- The populations that show statistical difference (in terms of likelihood of requiring protection) from mammals & birds are shown in the table below

Significant difference between mammal or bird and specified other category?

	Amphibian	Bird	Fish	Mammal	Nonvascular plant	Reptile	Vascular plant
Mammal	N	N	N	N/A	Y	Y	Y
Bird	N	N/A	N	N	Y	N	Y

- Based on this information, conservation efforts would do well to focus on animals over plants, but should avoid heavily prioritizing efforts for one animal type over another

Foot & mouth disease: how long will it take to assess incidence?

Initial conditions

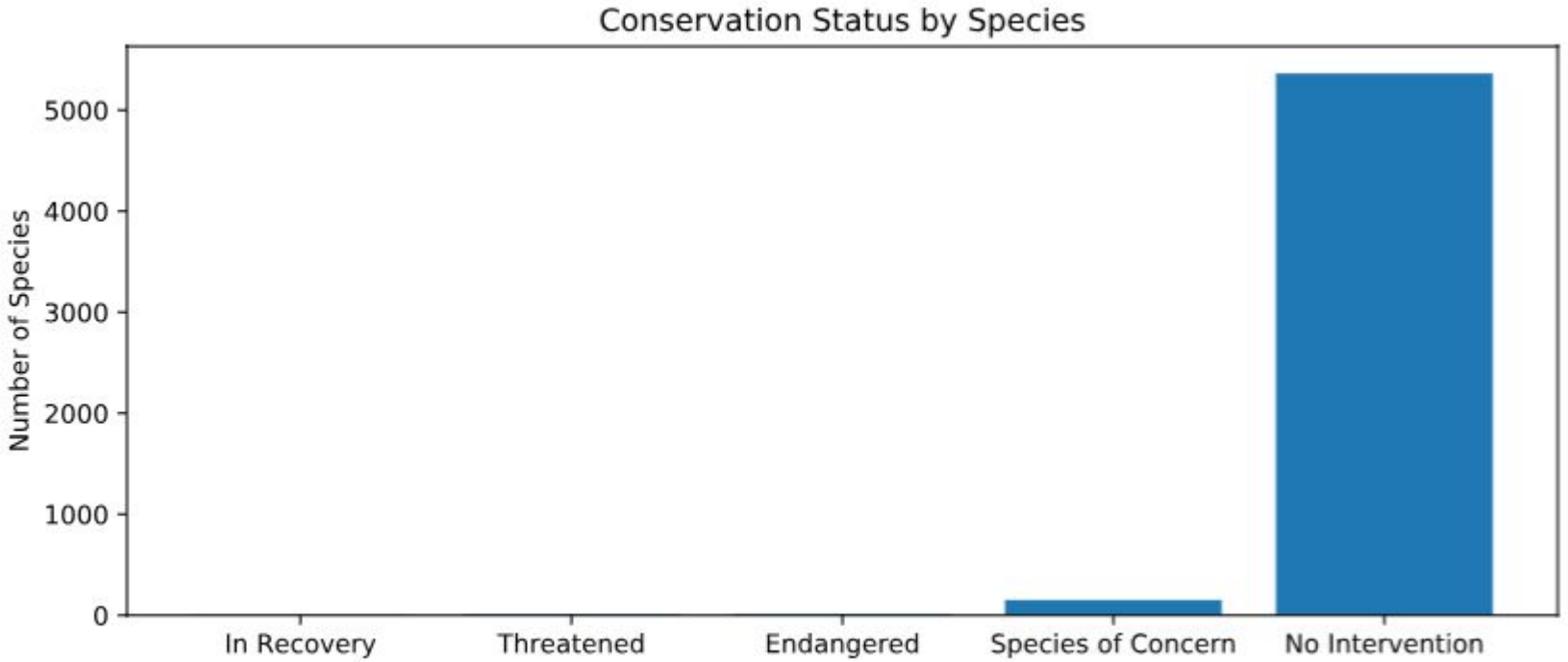
- Park Rangers at Yellowstone want to determine how many data points are necessary to ensure they are generating good data describing the incidence of foot & mouth disease in sheep
- The baseline expectation of incidence, from a previous year's results, is that 15% of sheep suffer from the disease
- The minimum effect that they would like to be able to detect is +/- 5% from the baseline, also described as 33% of the baseline (5/15 = 0.333)
- The Park Rangers want to use a standard statistical significance level of 90%

Results

- Given these initial conditions, the Park Rangers need to gather 890 sheep samples to understand incidence of disease
- Based on observation data (see *Supporting graphs (2)*) showing that Park Rangers have seen 507 sheep in the previous week at Yellowstone, this implies that ~1.75 weeks will need to be spent to collect the required samples
- If the experiment is repeated at Bryce National Park, Park Rangers will need to collect data for ~3.5 weeks, as only 250 sheep are expected to be seen on a weekly basis

Supporting graphs

Supporting graphs (1)



Supporting graphs (2)