3/25/2018

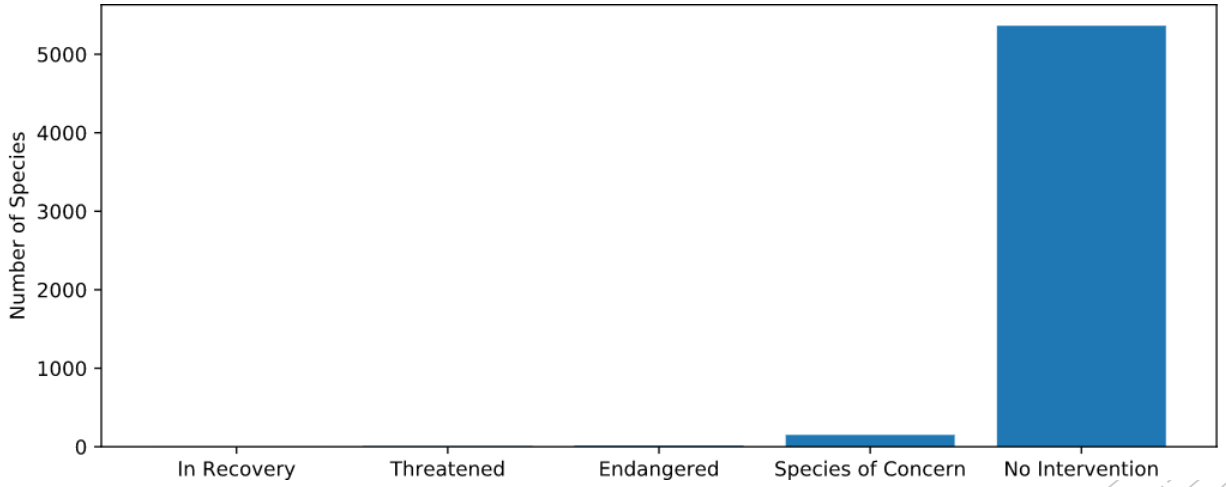
Capstone Option 2:

Biodiversity for the National Parks

Philip Skogsberg

Species Data Set

- The data is comprised of observations of species made at various national parks.
- The majority of species observed in the parks are currently not protected, or more specifically, lack an assigned conservation status (No intervention).
- Having the property of **not** being protected doesn't necessarily imply that a species is endangered, but it could imply that they are at risk of falling into that category if we can't say for sure that they aren't protected. If they are protected we know their status and can take the necessary actions to conserve them.
- Unsurprisingly, the most numerous category of observed species in the dataset are plants. (Of course, this may differ slightly from the total population in the parks, as it's a sample.)
- It seems like the smaller, more numerous, and less "animal-like" kinds of species like amphibians, plants and reptiles are more likely to be unprotected (or rather to lack a conservation status). This could be because the bigger and more prominent animals like mammals are much easier to spot and track for the The National Parks Service, and/or because non-mammalian species seem less important to us as humans (a bias).



Significance calculations

- Question: **are certain types of species more likely to be endangered?**
- As outlined in the previous slide, it seems that that some are! More specifically:
  - There's no significant difference (that we can establish with confidence) between Mammals and Birds in their protection statuses.
  - There is however a significant difference between Reptiles and Mammals (and thus with a high likelihood, a difference between Amphibians and Mammals, and Fish and Mammals).
  - I also wanted to know whether there are any significant differences among amphibians, fish and reptiles, but there turned out not to be.
  - The next step "down" in the rate of protection were the plant species, and as explained in the previous slide, there seems to be a stark difference in the rate of protection between non-plant species and plants, as the significance test between reptiles and nonvascular plants showed.
  - There was no significant difference in protection status between vascular and nonvascular plants, however.

Recommendations

- As a result of these calculations, we can observe that there are statistically significant differences of conservation status at three different "levels":
  - First around larger "animal-like" species like Mammals and Birds,
  - Then around smaller species like fish, amphibians, and reptiles,
  - and finally among plants, which are the least likely to be protected.
- My recommendations would be to focus more on plant life discovery and protection since this category of species are the least protected. However, bigger fauna are also more likely to be noticed and hunted, and can be more sensitive to ecological changes since there are a fewer number of individuals in a population, compared to smaller species like plants (in general). Thus, it can make sense to continue the outsized focus on larger animals as well.
- Finally, the conservationists must of course approach their protection efforts on a systematic level taking into account all the species in the ecosystem as a whole. As many species rely on each others existence or non-existence in very intricate and complicated ways. The data analysis performed in this exercise can serve as a starting point for them.

Sample Size Determination

- Some sheep in the various national parks have come down with a foot and mouth disease. Scientists have taken certain measures to reduce the incidence of this disease at Yellowstone National Park, but they only new the current/previous rate of occurrence, of the disease, at Bryce National Park (15%).
- To be sure that there has been at least a 5% drop in occurrence, we use the baseline data from Bryce to figure out how many Sheep they have to test in Yellowstone to be confident that they actually discover a difference that is significant. It turns out it will take about a week to observe a minimum of 510 sheep.
- With this information, the scientists will be able to determine if their efforts to reduce the prevalence of the disease among the sheep is working or not.