Biodiversity for the National Parks

A Study on Endangered Species

By Olivia Hoewing

Data Analyst @ National Park Service

In collaboration with

codecademy

Introduction

This presentation examines data on endangered species.

The analysis of two datasets aims to answer the following questions:

1. **Are there certain types of species that are more endangered than others?**
   - Based on Dataset 1 – 'Species'
2. **How long will it take to test if a disease improvement program was effective?**
   - Based on Dataset 2 – 'Observations'

1) DATASET *SPECIES*

Exploring the Dataset ‘Species’ will answer questions on the most endangered species.

codecademy

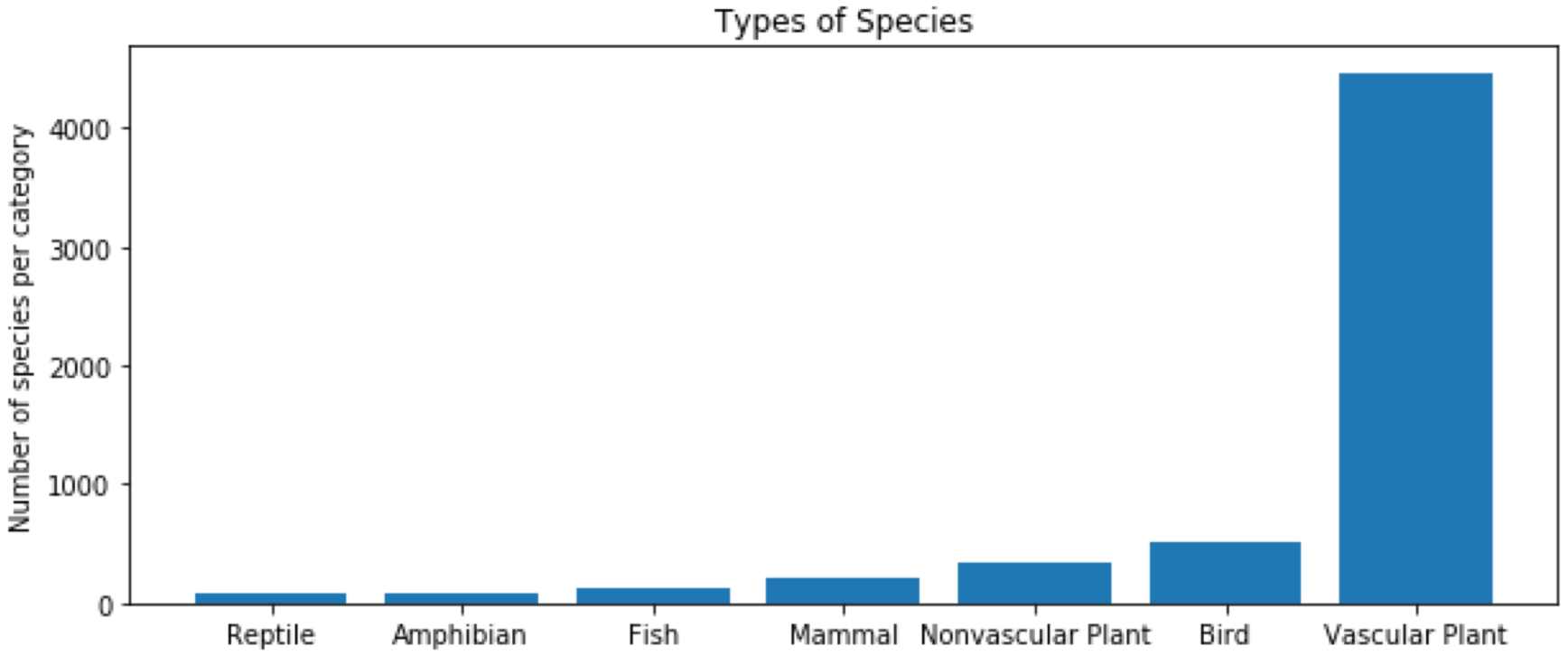
Dataset ‘Species’

This is its data schema.

Category	Scientific Name Common Name	Conservation Status
1. Mammal 2. Bird 3. Reptile Amphibian 4. Fish 5. Vascular Plant 6. Nonvascular Plant	**5541** unique species	1. No Intervention 2. Species of Concern 3. Endangered 4. Threatened 5. In Recovery

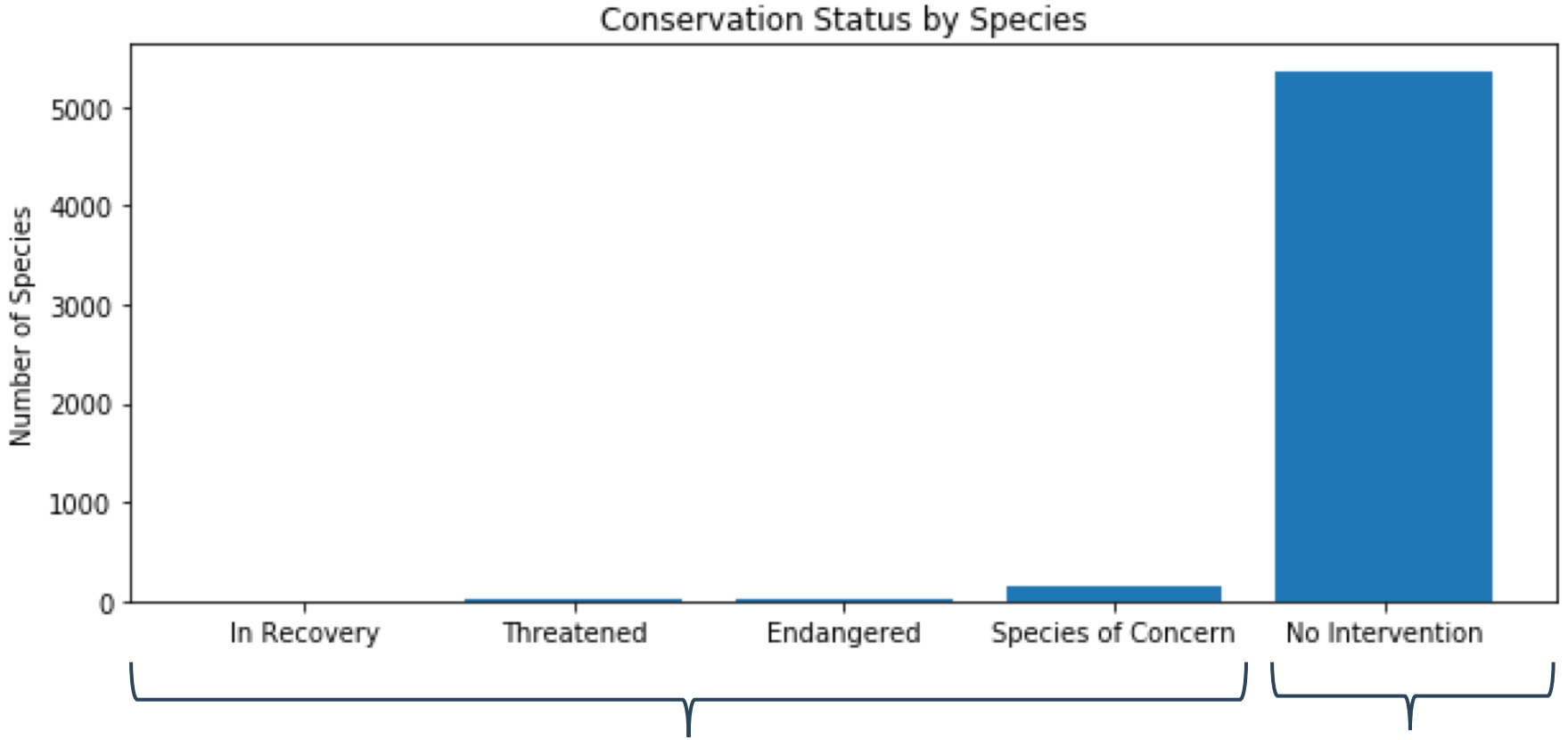
Data Overview I

How many species are there for each type?



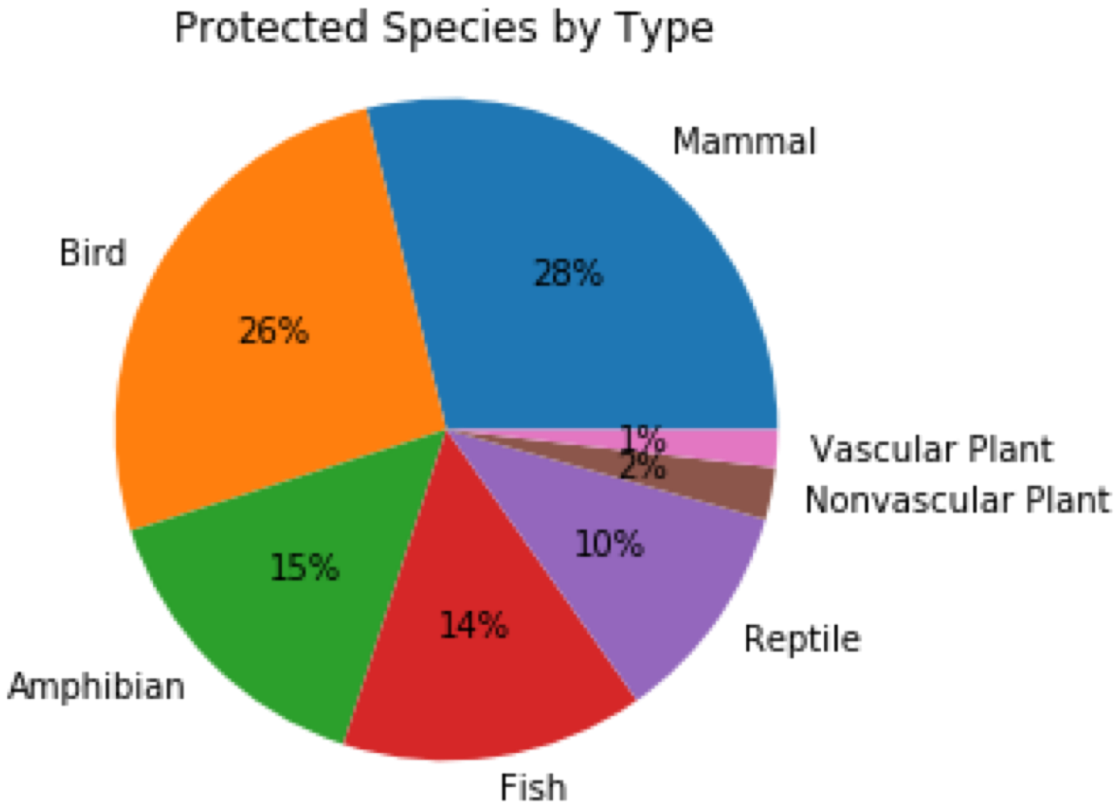
Data Overview II

How many species are protected vs not protected?



Focus on Protected Species

Which types of species are most endangered?



Significance Analysis

Are some types of species more endangered than others?

Significant Differences were found between the following pairs of types of species:

- Amphibian - Non-vascular Plant
- Amphibian - Vascular Plant
- Bird - Non-vascular Plant
- Bird - Vascular Plant
- Fish - Non-vascular Plant
- Fish - Vascular Plant
- Mammal - Non-vascular Plant
- Mammal - Reptile
- Mammal - Vascular Plant
- Non-v. Plant - Reptile
- Reptile - Vascular Plant

Key Conclusions

- Species that belong to **Amphibians, Birds, Fish, Mammals and Reptiles** are significantly more endangered than species that are vascular and non-vascular plants.
- Mammal Species are also more endangered than Reptile Species.

Recommendation

Based on the Significance Analysis.

- **Amphibians, Birds, Fish, Mammals and Reptiles** are significantly more endangered than vascular and non-vascular plants and will therefore require extra care.
- It is possible that **the conservation status of some species require reassignments** in order to focus protection efforts on the most endangered species.

2) DATASET *OBSERVSTIONS*

Exploring the Dataset ‘observations ‘ will help answer questions on the Foot and Mouth Disease Improvement Program.

codecademy

Dataset ‘Observations’

This is a snapshot of the data.

Scientific Name	Park Name	Observations
Vicia benghalensis	Great Smoky Mountains National Park	68
Neovison vison	Great Smoky Mountains National Park	77
Prunus subcordata	Yosemite National Park	138
Abutilon theophrasti	Bryce National Park	84
Githopsis specularioides	Great Smoky Mountains National Park	85

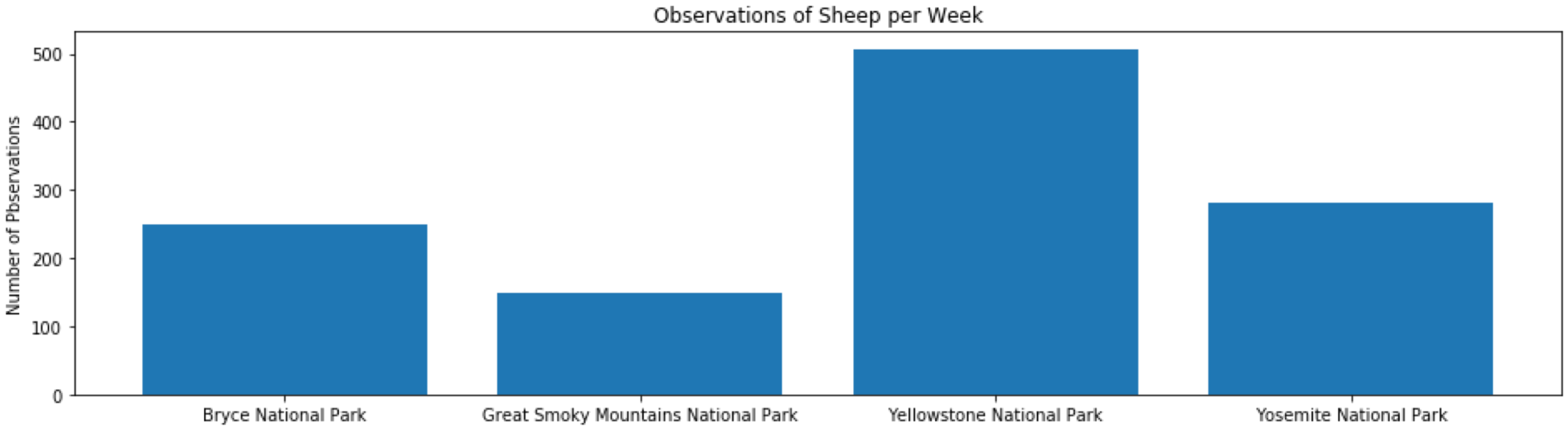
Foot and Mouth Disease

The effectiveness of a new disease improvement program will be tested among Sheep. How large will the test samples need to be?

1. First, we are going to filter the dataset for all species that are Sheep.
2. A new program, that helps to cure foot and mouth disease, will be tested among sheep in two out of four national parks.
3. In order to find out if the improvement program was truly effective - rather than influenced by chance -, we will need to extract samples of sheep from two national parks.
4. We will determine how large each sample needs to be so that we can be 90% sure that the differences between the two samples are not due to chance.

Dataset ‘Observations’

Below you will find the number of sheep observed during one week at four National Parks.



Improvement Program

How large will each sample size need to be?

- Currently 15% of all sheep have Foot and Mouth disease.
- The goal is to improve this rate by 33%.
- The improvement program will first be launched at Yellowstone and Bryce National Parks.

➢ **In order to test the impact of the program, we will need to:**

- **Observe 510 sheep at each park (1020 in total)**
  - **50% of sheep will be part of the improvement program**
  - **50% of sheep will be part of the control group**
- **Observing 510 sheep will take 2 weeks at Bryce National Park …**
- **… and 1 week at Yellowstone National Park.**

Thank You

Please reach out for more information.

codecademy