Biodiversity for the National Parks

By Santiago Moscatelli

Codeacademy: Introduction to Data Analysis

Endangered Species

Introduction: Population Description

According to the data provided by the National Park Service, there are

5,541

different species, across

7

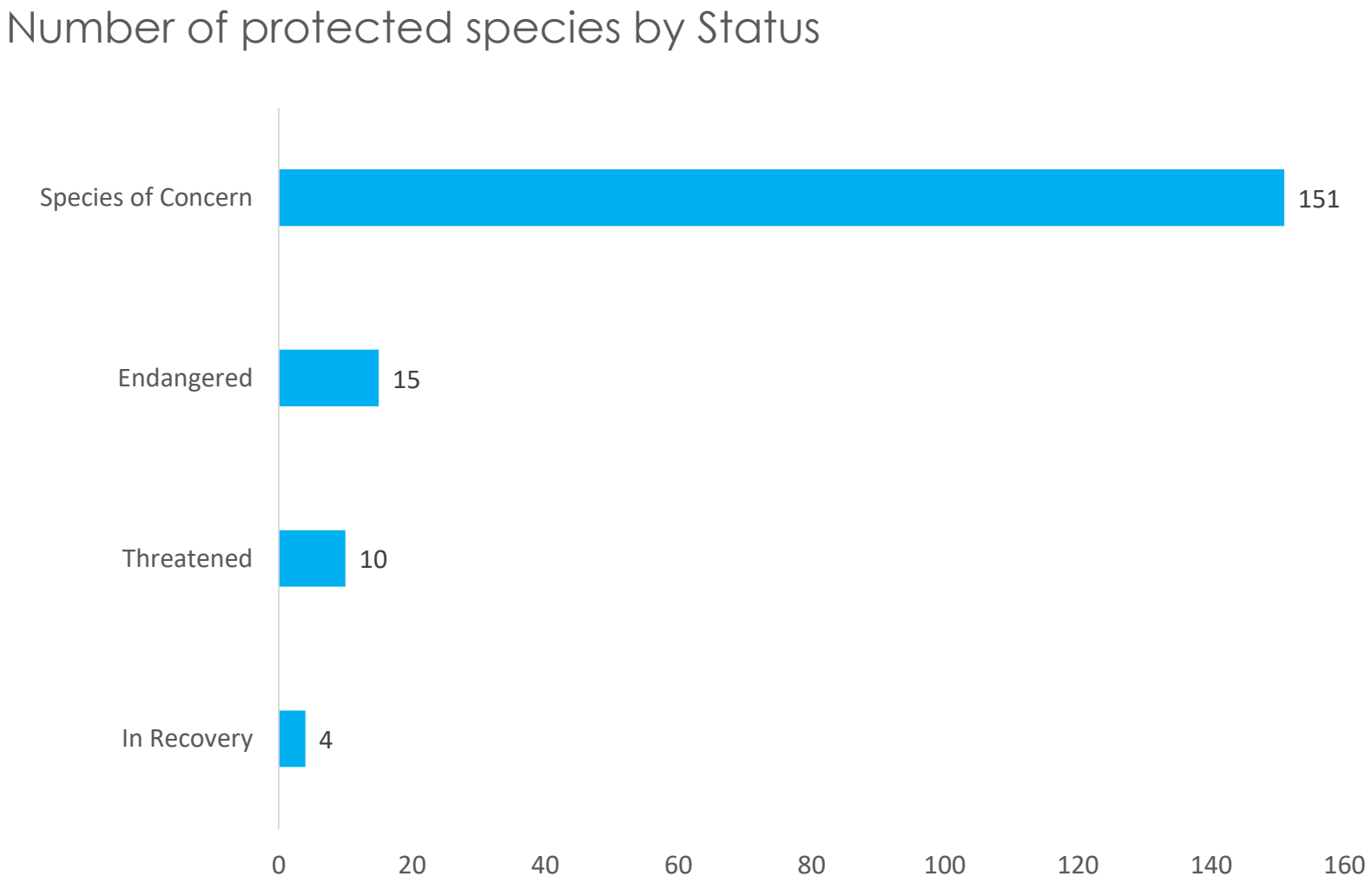
different categories.
The amount of protected species is

180

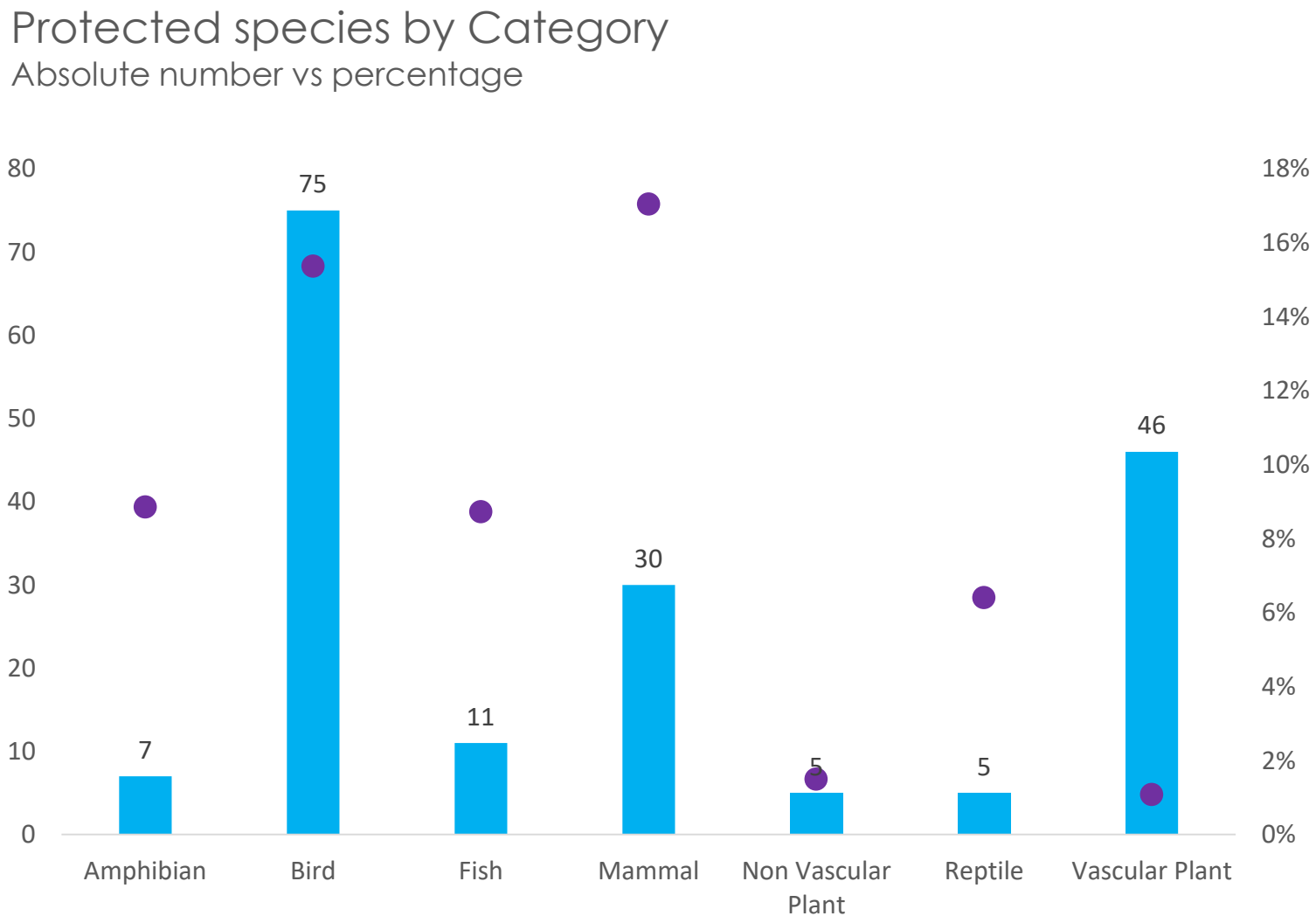
which represents a

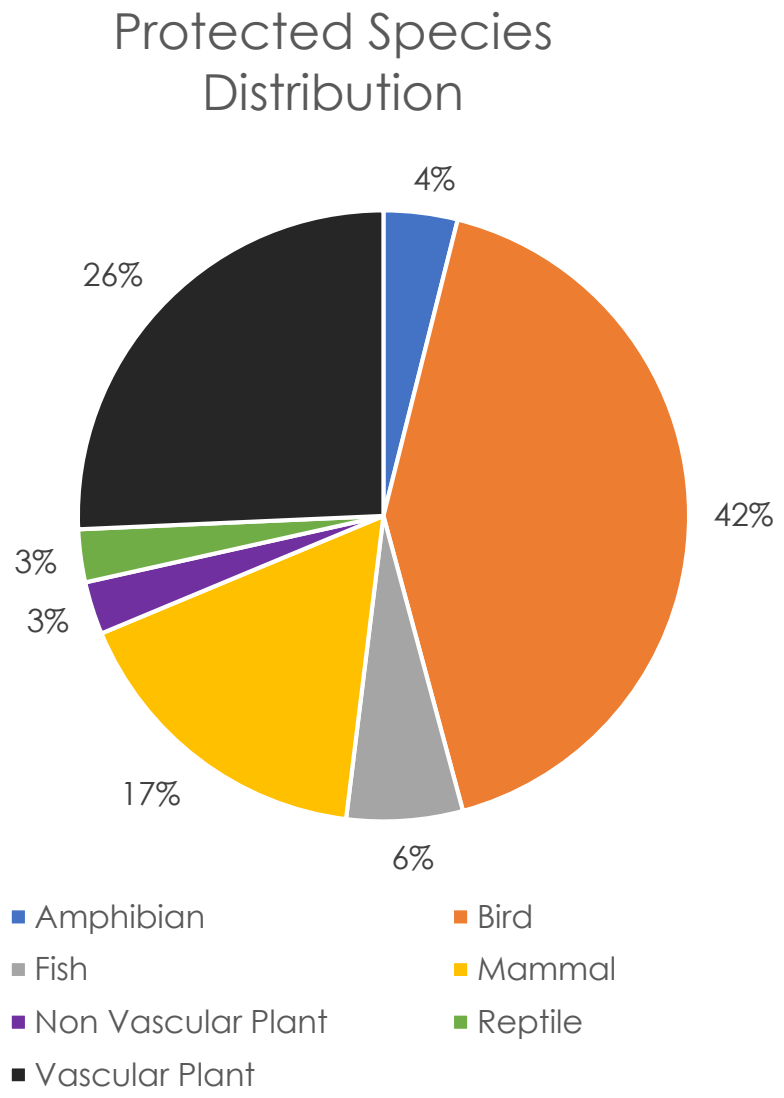
3.2%

of the total amount of different species.



Endangered Categories





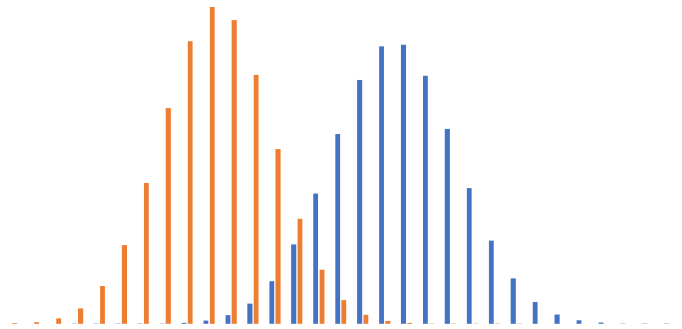
Differences Between Categories

To determine differences in categories' likeliness to be protected, multiple Chi Square tests were run with the following hypothesis:

- The percentage of protected species is the same across two different categories – H_0: $p_{categoryA} = p_{categoryB}$

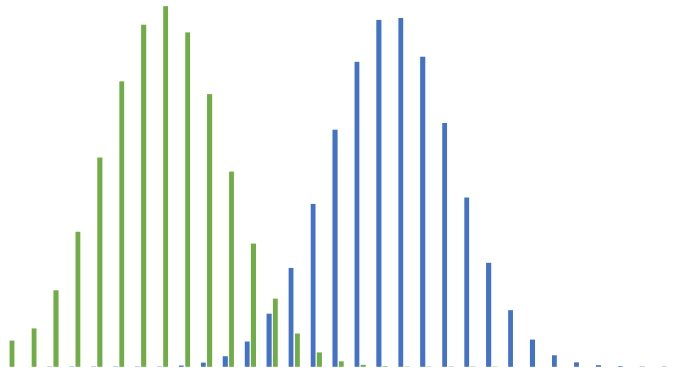
Mammal vs Fish

We **can not conclude** that Mammal's and Fish's percentage of protected species differ significantly at a 95% confidence level.
The Chi Square test returned a p value of 0,6876 which is larger than 0,05.



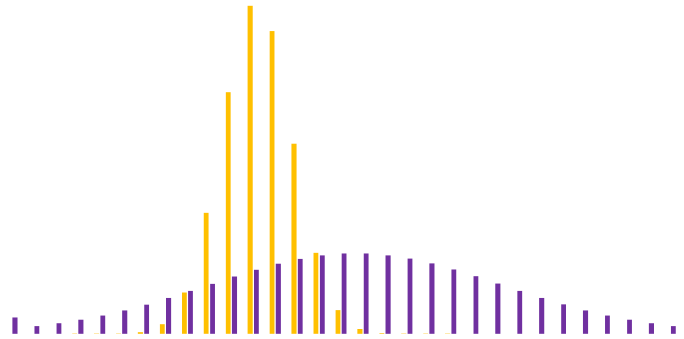
Mammal vs Reptile

We **can conclude** that Mammal's and Reptile's percentage of protected species differ significantly at a 95% confidence level.
The Chi Square test returned a p value of 0,0,0384 which is smaller than 0,05.



Vascular vs Non Vascular Plants

We **can not conclude** that both types of plants' percentages of protected species differ significantly at a 95% confidence level.
The Chi Square test returned a p value of 0,6623 which is larger than 0,05.

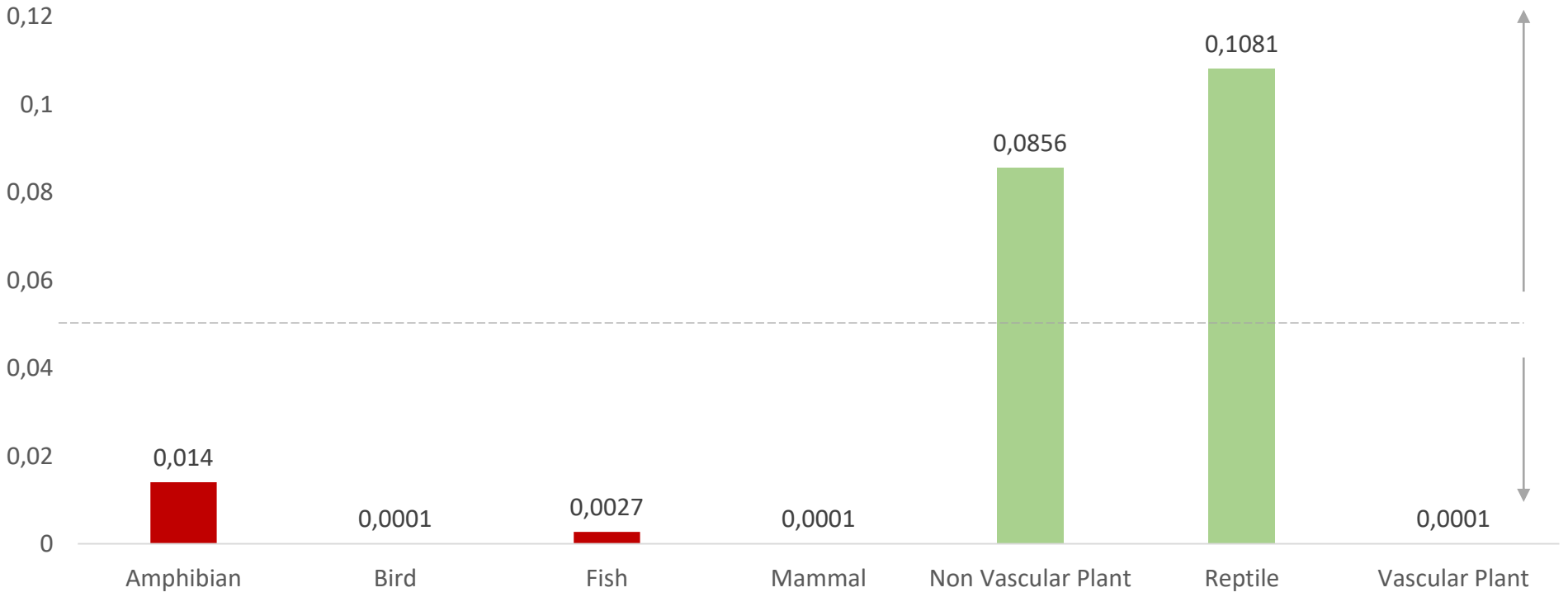


Most Susceptible Categories

To determine which categories are more likely to be protected than the average, multiple Binomial tests were run with the following hypothesis:

- Total population percentage of protected species is the same as each category percentage of protected species – H_0: $p_i = p_{opulation}$ – at a confidence level of 95%

p Value for every category against the population



The hypothesis that likeliness of being endangered is the same between the total population and these categories can not be rejected.

Categories are more likely to be endangered than the average, with the exception of Vascular Plants, which are less likely than the average to be endangered.

Recommendations

Due to the statistical analysis results, the following insights could improve the species protection program:

- As Amphibians, Birds, Mammals and Fish record a significantly higher percentage of protected species than the average, more efforts should be allocated in to programs that protect these categories.

- While Mammals and Fish should be treated similarly, Reptiles are currently in a safer situation than Mammals and, therefore, some of the efforts allocated to Reptiles could be shifted in to Mammals, as they present a higher likeliness to be endangered.

- Although the absolute number of protected Vascular Plants is much higher than the number of protected non Vascular Plants. The percentage of protected species on both groups does not differ significantly. Efforts allocated to both programs should be proportionally similar.

- Categories could be classified into 3 different group, taking in to account if their percentage of endangered species is significantly larger, smaller or similar to the average:
    1. Categories more endangered than average
    2. Average endangered categories
    3. Categories less endangered than average

Sheep Observations

Introduction: Population Description

According to the information provided by the National Parks Services data, there are 3 different types of what is commonly named as 'sheep'.

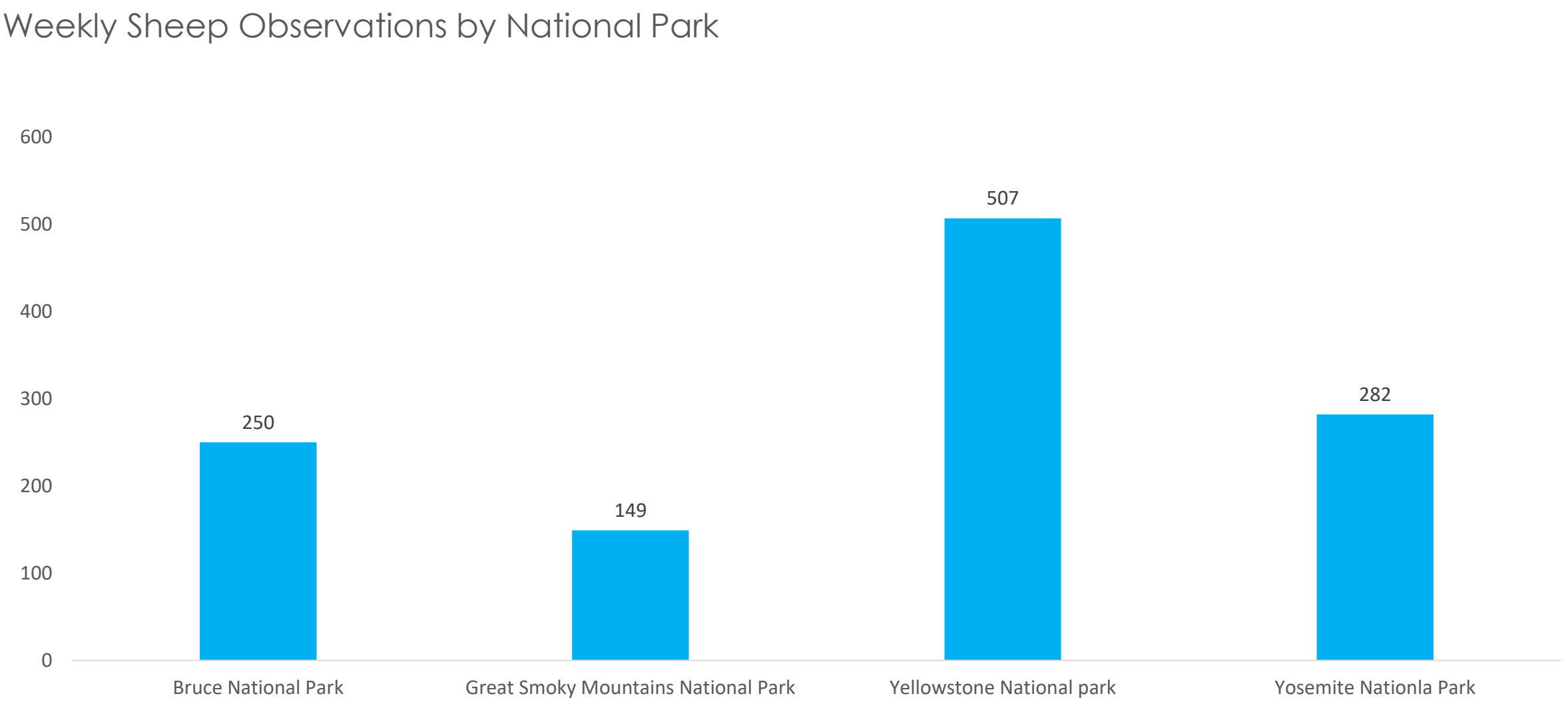
- Ovis Aries
- Ovis Canadensis
- Ovis Canadensis Sierral

A foot and mouth disease study will be conducted with observations of these sheep species in the following National Parks:

- Bryce National Park
- Great Smoky Mountains National park
- Yellowstone National Park
- Yosemite National Park

Specie	Endangered Status
Ovis Aries	Not Protected
Ovis Canadensis	Specie of Concern
Ovis Canadensis Sierral	Endangered

National Parks Observations



Foot and Mouth Disease Study: Sample Size

Input	Value
Last Year % of Diseased Sheep	15%
Minimum Detectable Effect	33.33%
Statistical Significance	90%

Minimum sample size to get results with statistical significance:

870

sheep

Foot and Mouth Disease Study: Weekly Observations

According to the weekly sheep observations in the different National Parks and the minimum sample size necessary to reach results with statistical significance, the numbers of weeks that the study should be run in each National Park is as follows:

Bryce National Park

3.48

driven from 870 divided by 250

Great Smoky Mountains National Park

5.84

driven from 870 divided by 149

Yellowstone National Park

1.72

driven from 870 divided by 507

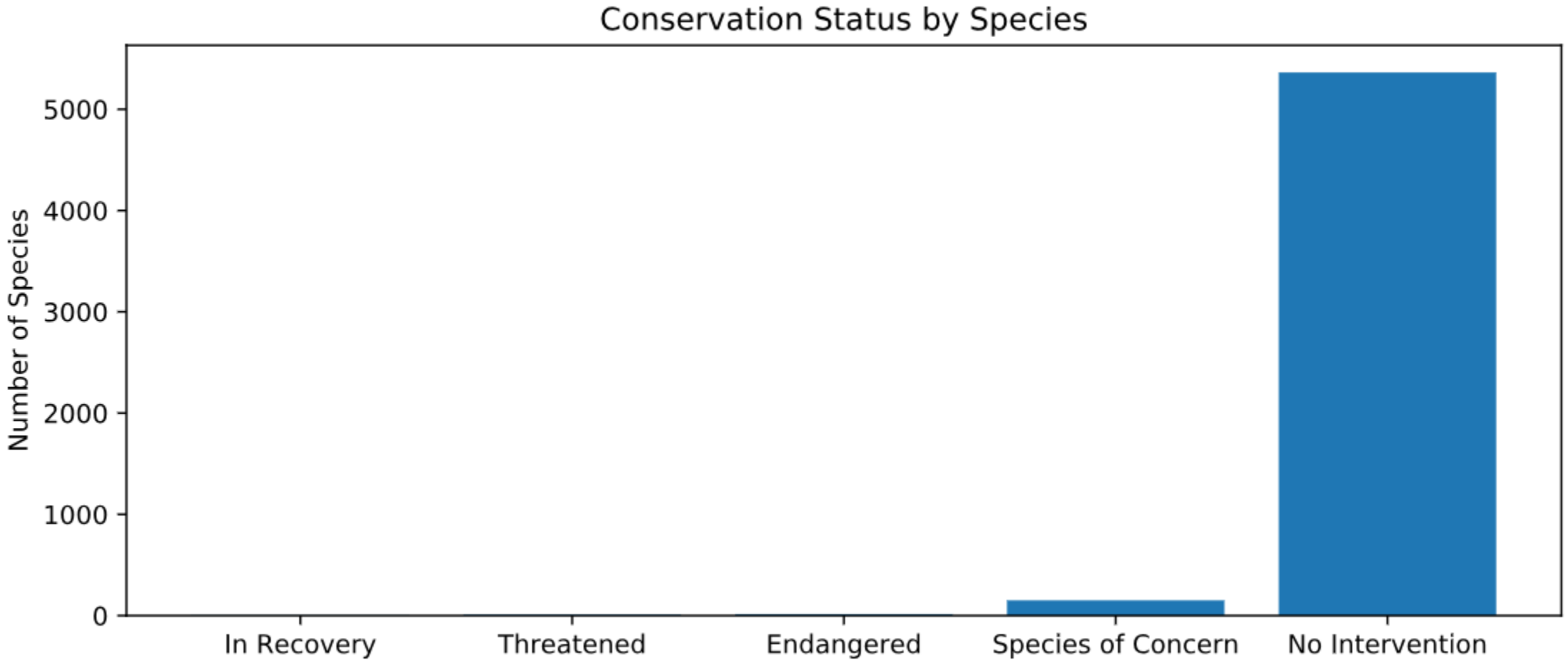
Yosemite National Park

3.09

driven from 870 divided by 282

Appendix: Graphs

Conservation Status by Species Graph



Sheep Sightings Graph