Biodiversity Analysis

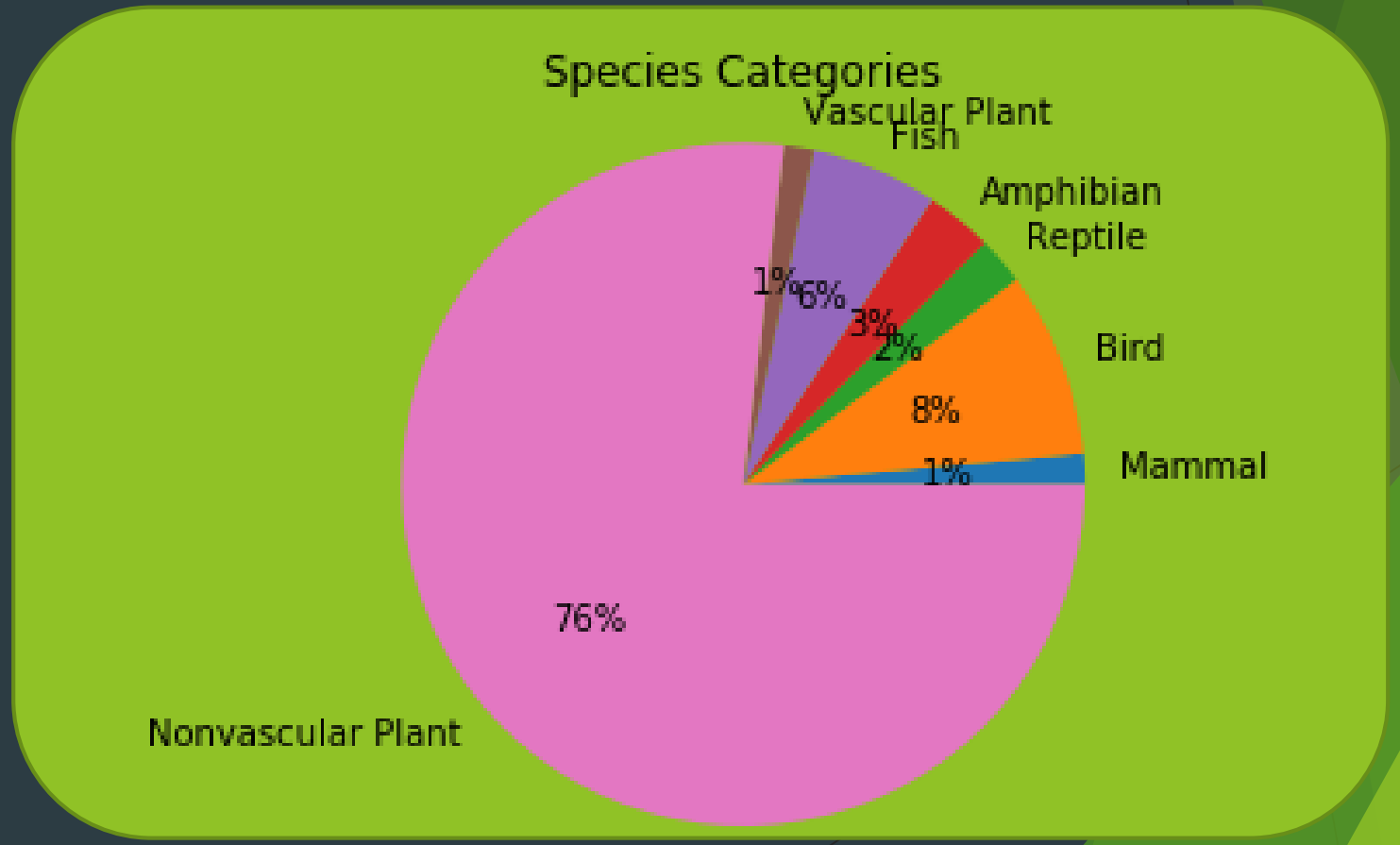
Scott Perry
Introduction to Data Analysis

Contents

Data Breakdown

Species_info.csv

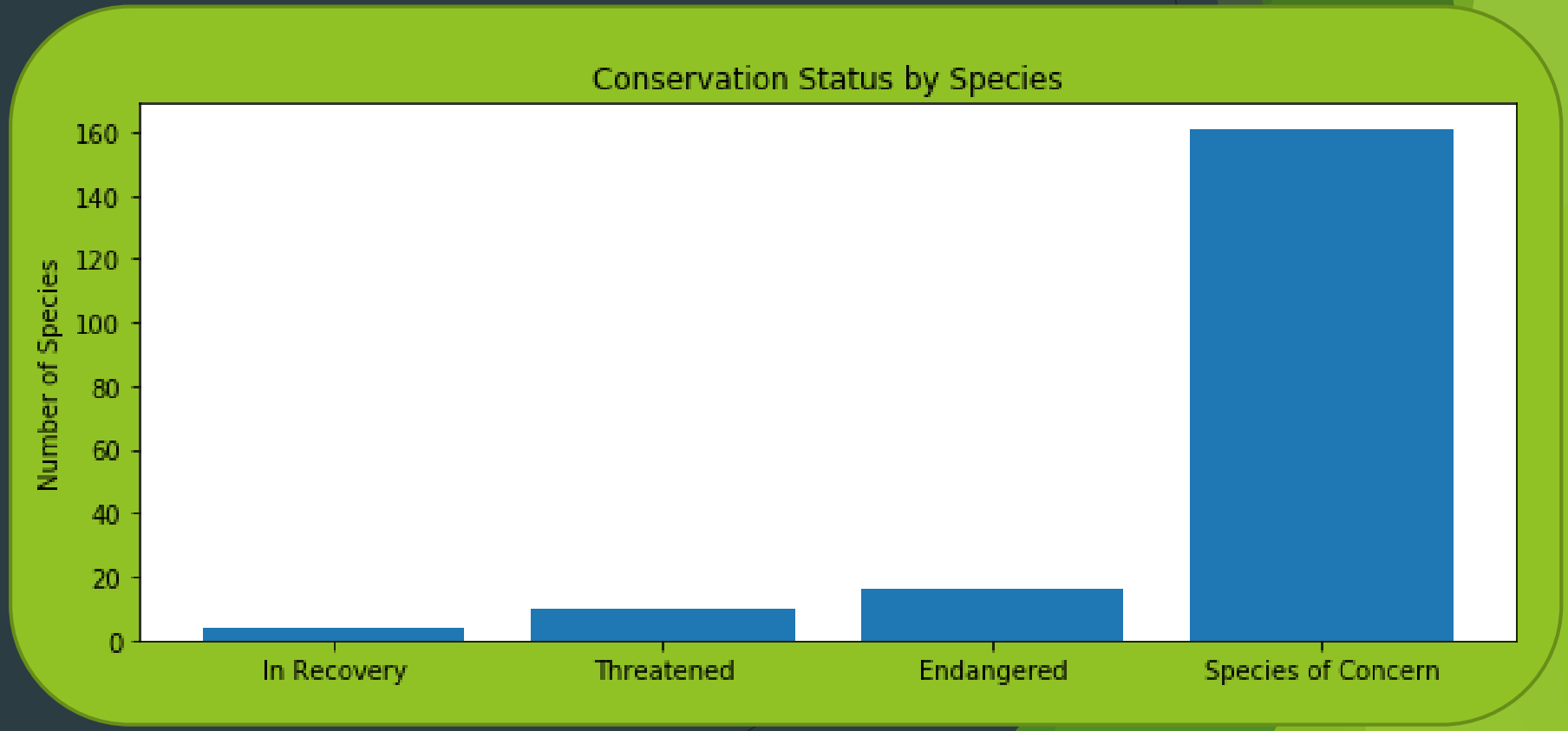
- A table of 5,541 unique species classified by 7 categories:
  - Mammal
  - Bird
  - Reptile
  - Amphibian
  - Fish
  - Vascular Plant



Species_info.csv

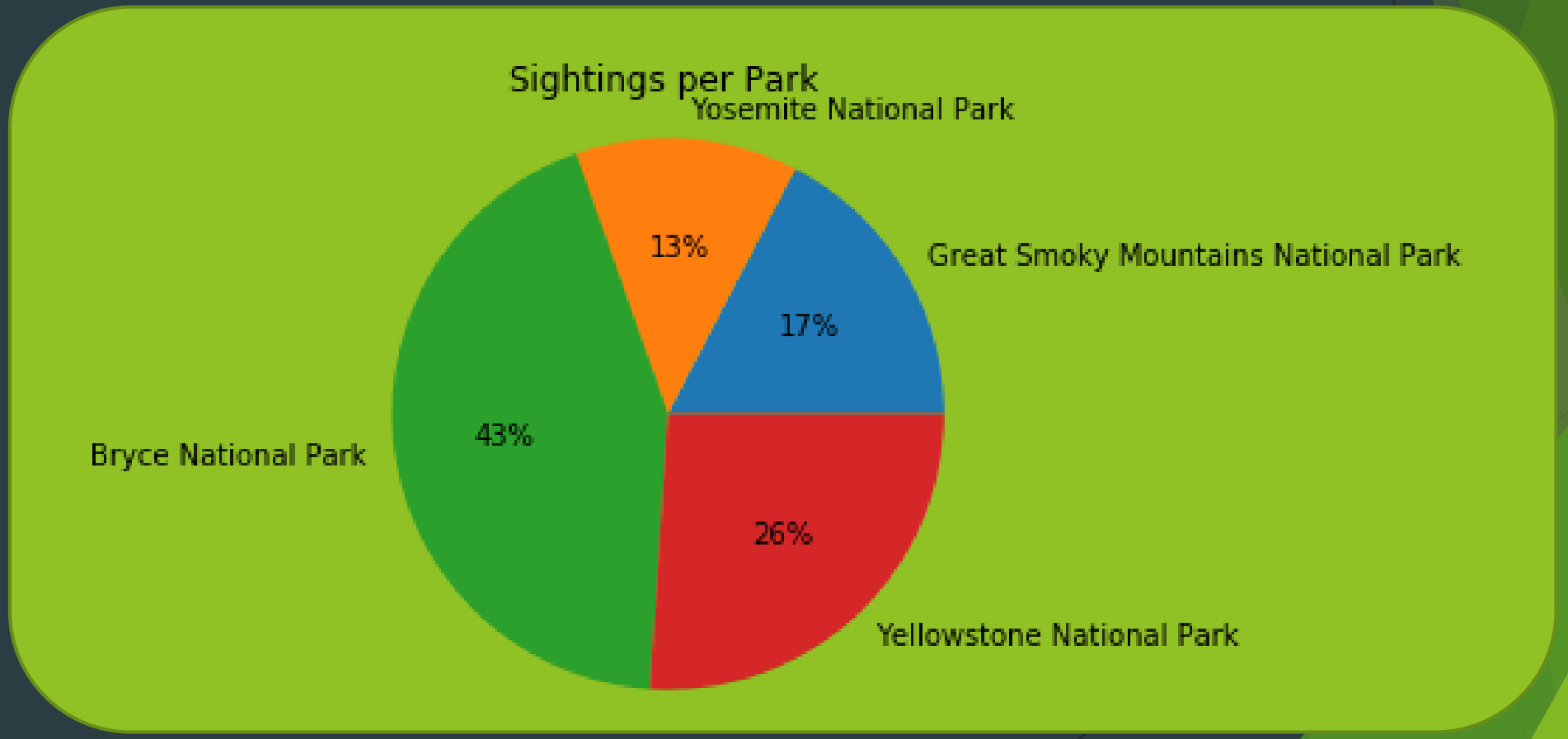
- Additionally, there is a classification of the conservation status of each species:
  - Species of Concern
  - Endangered
  - Threatened
  - In Recovery

- 96.7% of the species require no intervention and are not listed with a conservation status



observations.csv

- A table of the 5,541 species of species.csv with an entry for the number of observations at each park



Endangered Status Analysis

Definition of Endangered

- ‘Endangered’ generally refers to any of the conservation intervention statuses
- We will use that term going forward
- 191 of the 5,541 species in the data (3.3%) are classified as endangered

Endangerment Rates

- By grouping the not protected and protected status of each species, a percent protected statistic is calculated:

Category	Not Protected	Protected	Percent Protected
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015
Reptile	73	5	0.064103
Vascular Plant	4216	46	0.010793

Relative Endangerment of Species

- A chi squared test was performed on all of the species categories compared to mammals

Category	P Value (to Mammals)	Significant Difference?
Mammal	-	-
Bird	0.688	No
Reptile	0.038	**Yes**
Amphibian	0.128	No
Fish	0.056	No
Vascular Plant	1.441e-55	**Yes**
Nonvascular Plant	1.480e-10	**Yes**

Relative Endangerment of Species

- Unsurprisingly, the endangerment rates of both plant categories significantly differ from mammals
- Of the other animal categories, only reptiles differed significantly from the endangerment rate of mammals
- 17% of mammals are endangered. Reptiles manage to have a lower rate of 6.4%

Recommendations to Conservationists

Recommendations - Significance

- The relationship between human impact on the environment and reptiles warrants detailed study to learn of better ways to insulate other species from human activities.

- Better methods may also be discovered by comparing current conservation techniques of reptiles to the techniques used with other species.

Recommendations - General

- For a comprehensive comparison of all the categories without mammals as a basis, a Tukey's Range Test is recommended.

- For an increased understanding of a family of species, conservationists can compare the endangerment rates of each individual genus using the same analysis performed at the category level.

Foot and Mouth Disease Study

Grouping Data

- Foot and Mouth Disease is an affliction that affects all of the Ovis genus commonly referred to as sheep

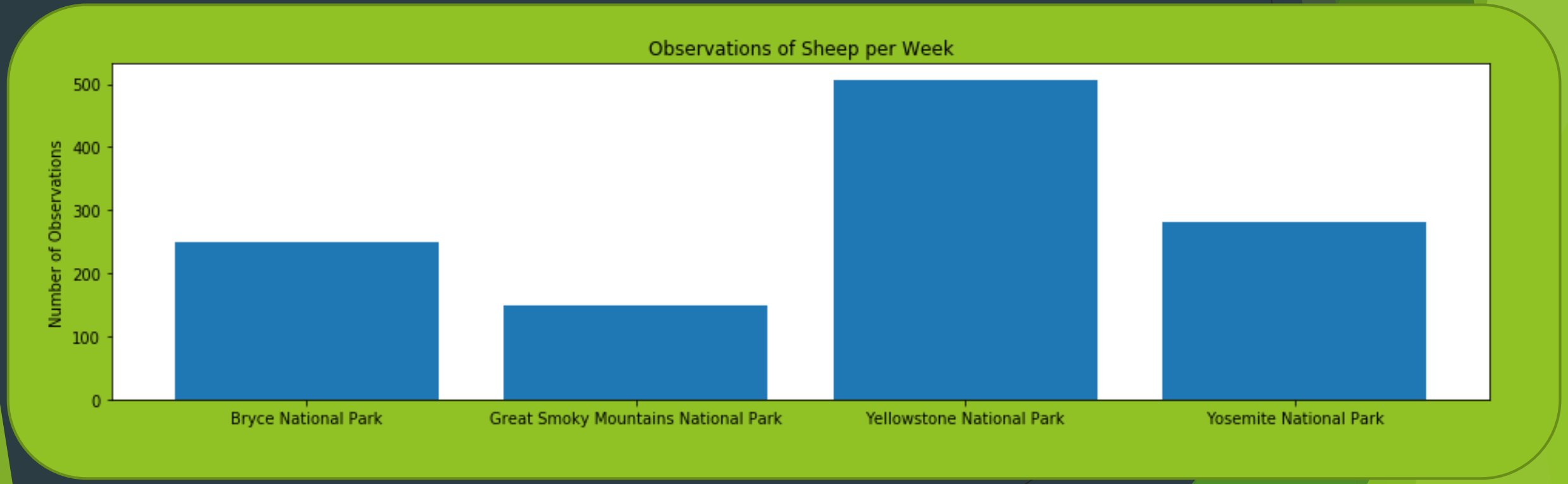
- A study would require knowledge of how many sheep were observed in total over the parks included in observations.csv

- The data was grouped for any species with a common name sheep:

Scientific Name	Common Names	Conservation Status
Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention
Ovis canadensis	Bighorn Sheep, Bighorn Sheep	Species of Concern
Ovis canadensis sierrae	Sierra Nevada Bighorn Sheep	Endangered

Ovis by Park

- This data was merged with the observations registry, then aggregated by species



Evaluating Reduction Effectiveness

- The baseline rate of Foot and Mouth Disease in sheep is 15%
- Our scientists wish to detect at least a 5% reduction rate from their trial program
- To find the required sample size for this study, we use the following parameters with a sample size calculator:
  - afflicated_pop = 0.15
  - target_effect = 0.05
  - minimum_detectable_effect = 100 * (target_effect / afflicated_pop) #33.33%
  - baseline = 15
  - statistical_sig = 0.90
  - sample_size_per_variation = 510

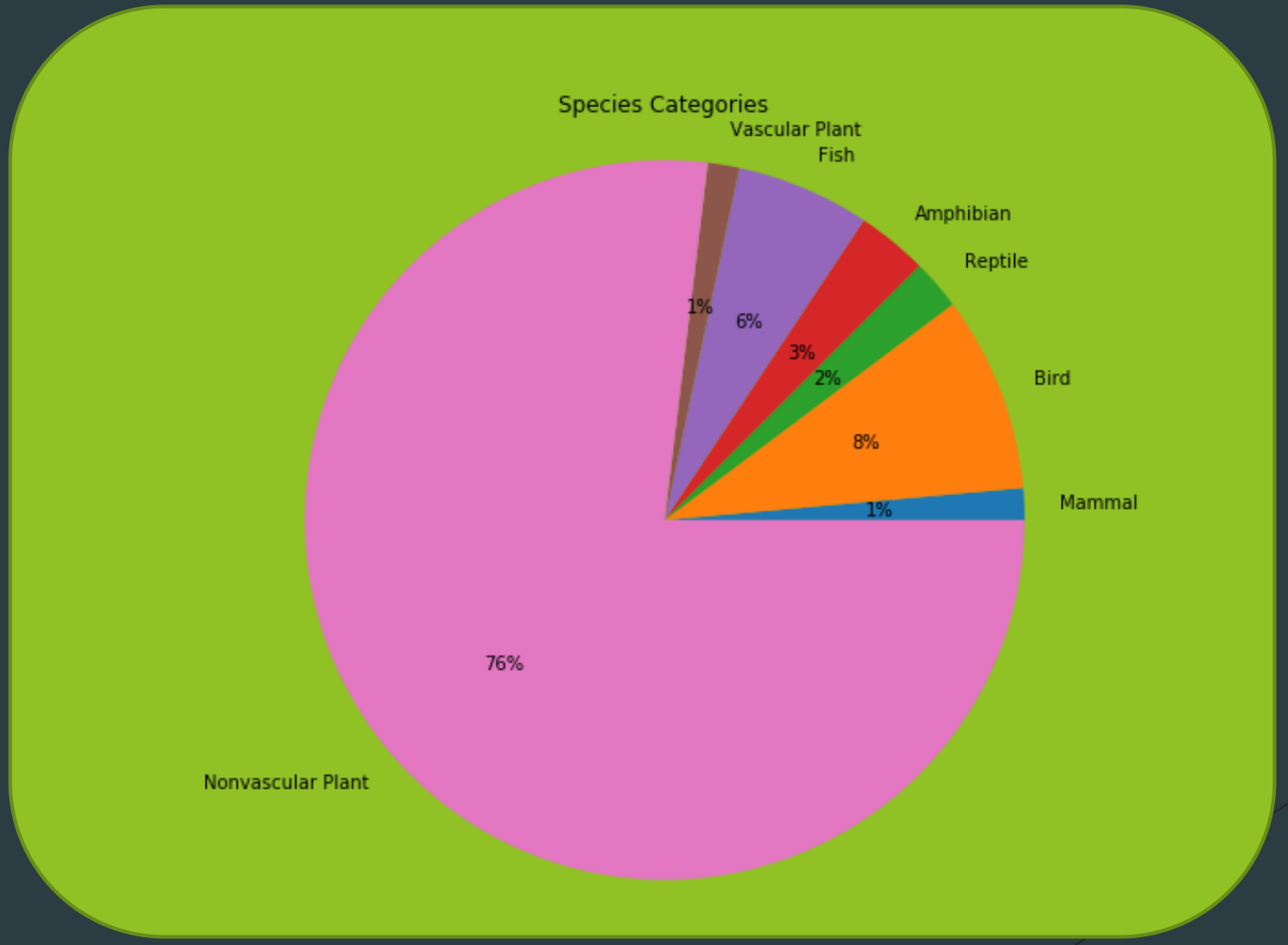
Minimum Study Length

- Since the observations over seven days is known, and assuming that week to be representative of the average sighting rate, the following chart presents the required length of study for the populations of each park:

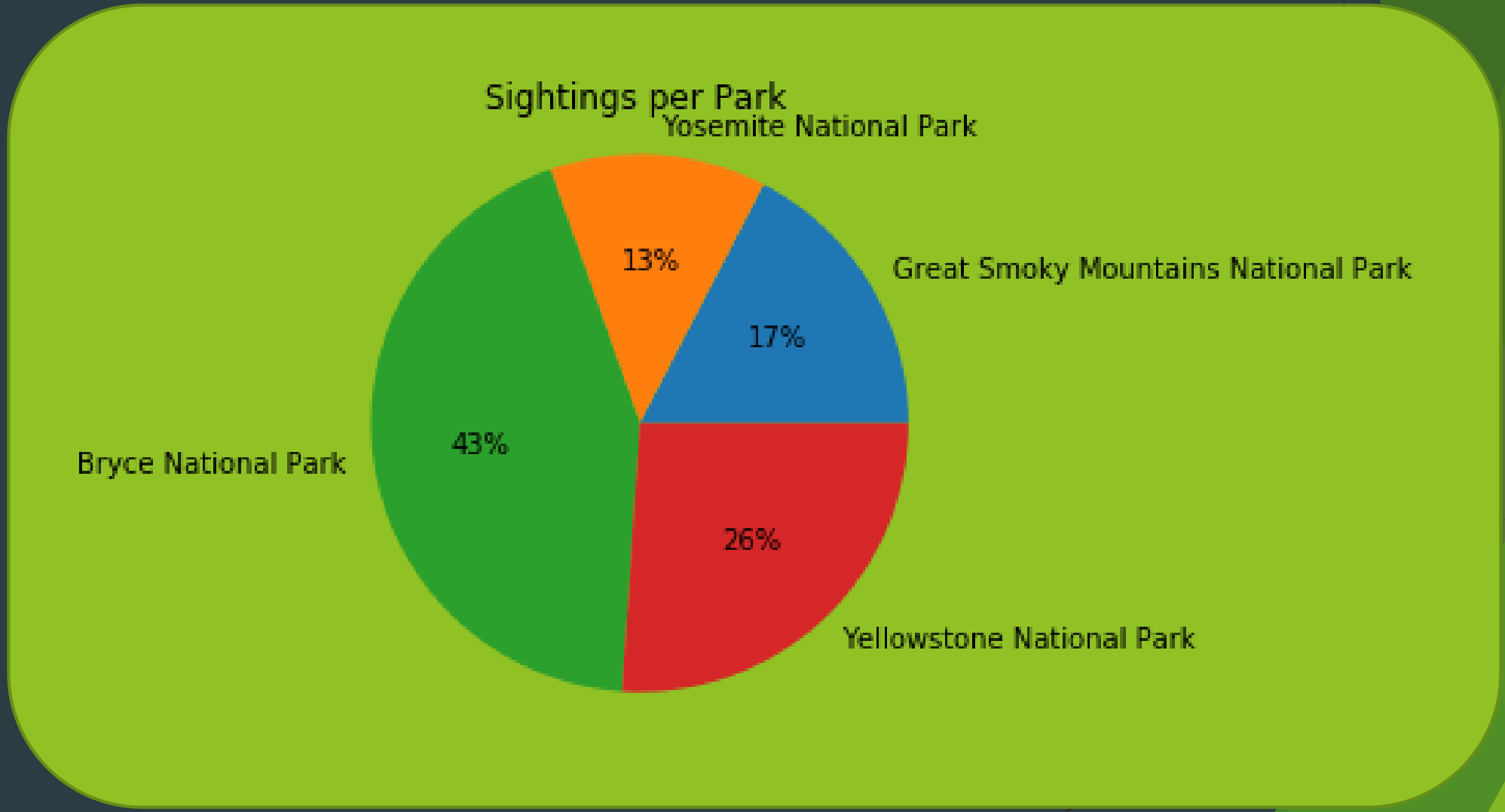
Park	Observations per Week	Weeks of Study
Bryce National Park	250	2.04
Great Smoky Mountains National Park	149	3.42
Yellowstone National Park	507	1.01
Yosemite National Park	282	1.81

Appendix: All Graphs Generated

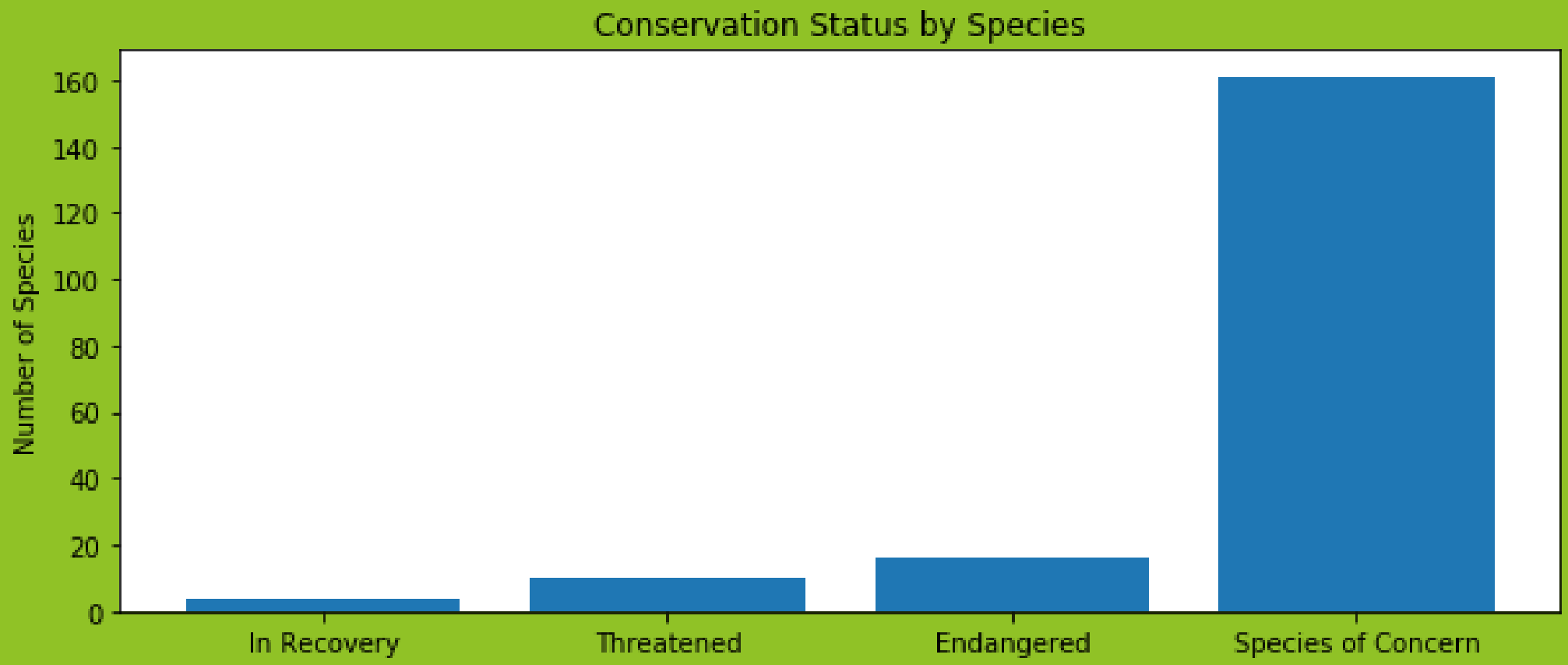
Total Species By Category



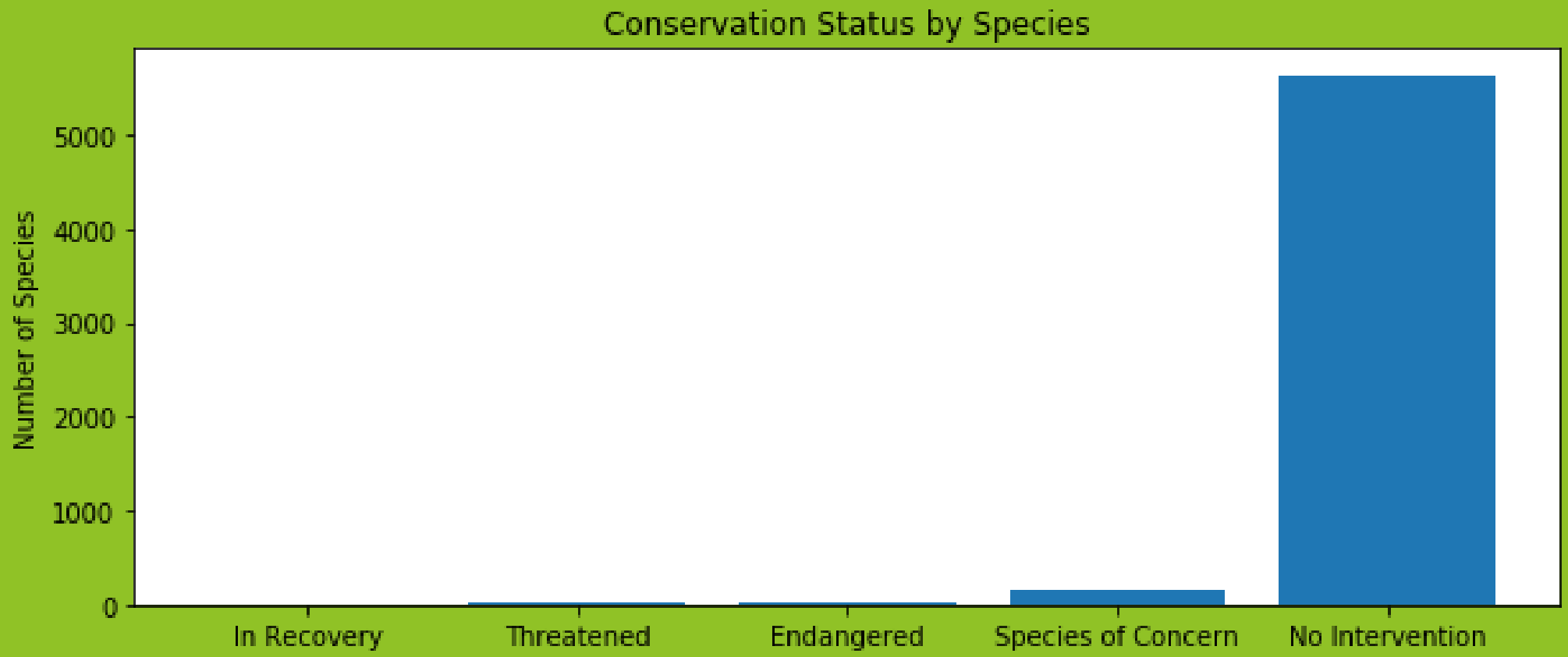
Observations by Park



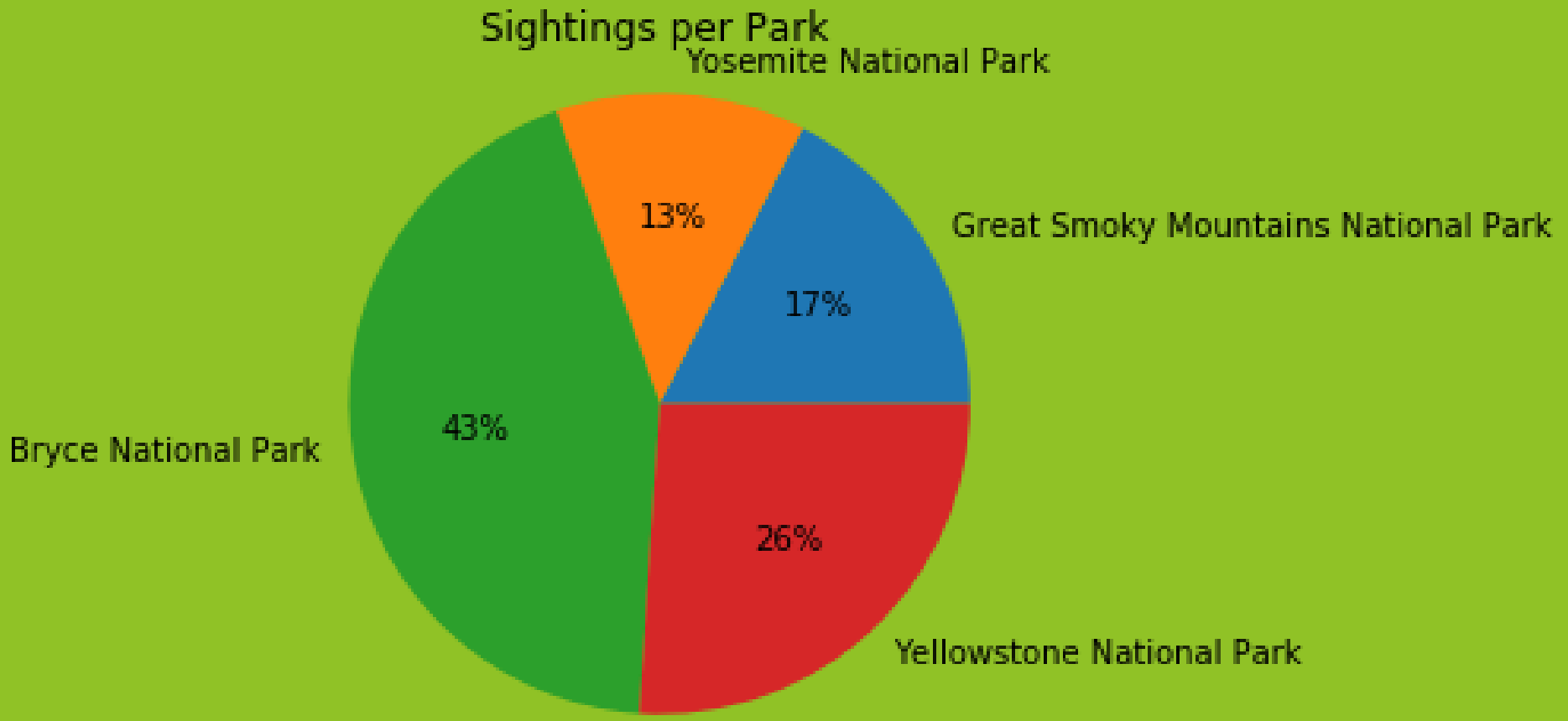
Conservation Status by Species



Conservation Status by Species – with Unendangered



Sightings per Park



Observations of Sheep per Week