A photograph of a dense, moss-covered forest floor. Large, gnarled tree roots are covered in thick green moss and hang down from the canopy. The ground is covered with various types of ferns and other green plants. The lighting is soft and dappled, creating a sense of a deep, ancient forest.

Codecademy Unit 6: Capstone Project

Biodiversity for the National Parks

Investigating species_info.csv

- This file contained 4 distinct comma-separated vales: category, scientific_name, common_names, and conservation_status, with 5,825 records total
- “category” had 7 options: Amphibian, Bird, Fish, Mammal, Nonvascular Plant, Reptile and Vascular Plant
- “scientific_name” had 5,539 different options corresponding to the specific scientific name for a species
- “common_name” had 5,502 different options corresponding to the given common name for a species
- “conservation_status” had 5 options: Endangered, In Recovery, No Intervention, Species of Concern and Threatened

Endangered Status Significance Calculations

- By manipulating the data into a grouped table, with the number of protected, not protected, and percent protected species per “category”
- Using this table, two categories with similar percent protection values could be analyzed by using a chi-squared test



- When asked to compare Bird & Mammal, as well as Reptile & Mammal, results showed that there was no significant difference between the first pair (p-value of 0.688), but a significant difference between the second pair (p-value of 0.038)
- THEREFORE, some types of species are more likely to be endangered, with percent protection difference not due to chance (ex. Reptile & Mammal)

Recommendations for Conservationists



- Using the Mammal category (highest percent protected) as a comparison point for the chi-squared contingency tests against different categories allows for investigation into whether a species is more likely to be endangered than mammals
 - Bird, Nonvascular Plant, Vascular Plant are not; just chance
 - Amphibian, Fish (on the cusp), Reptile are; not due to chance
- Given this information, conservationists concerned about how best to dedicate themselves to the protection of species could allocate additional time and resources to the protection and care of amphibians, fish, and reptiles, as they are more likely to be endangered

Foot and Mouth Disease Study in Sheep: Sample Size Determination

After manipulating and combining data to obtain about observations of sheep at specific parks per week.



- Through finding the baseline conversion rate, statistical significance, and minimum detectable effect corresponding to a program the Yellowstone National Park has undertaken to reduce the rate of foot and mouth disease, the sample size needed for significance of the program working could be determined.
- **RESULTS:** it would take scientists at Yellowstone ~1.76 weeks to observe enough sheep for a significant conclusion about the program's reduction in the disease.
- For the same experiment to be ran at Bryce National Park, scientists would need to observe sheep for the disease over ~3.56 weeks.

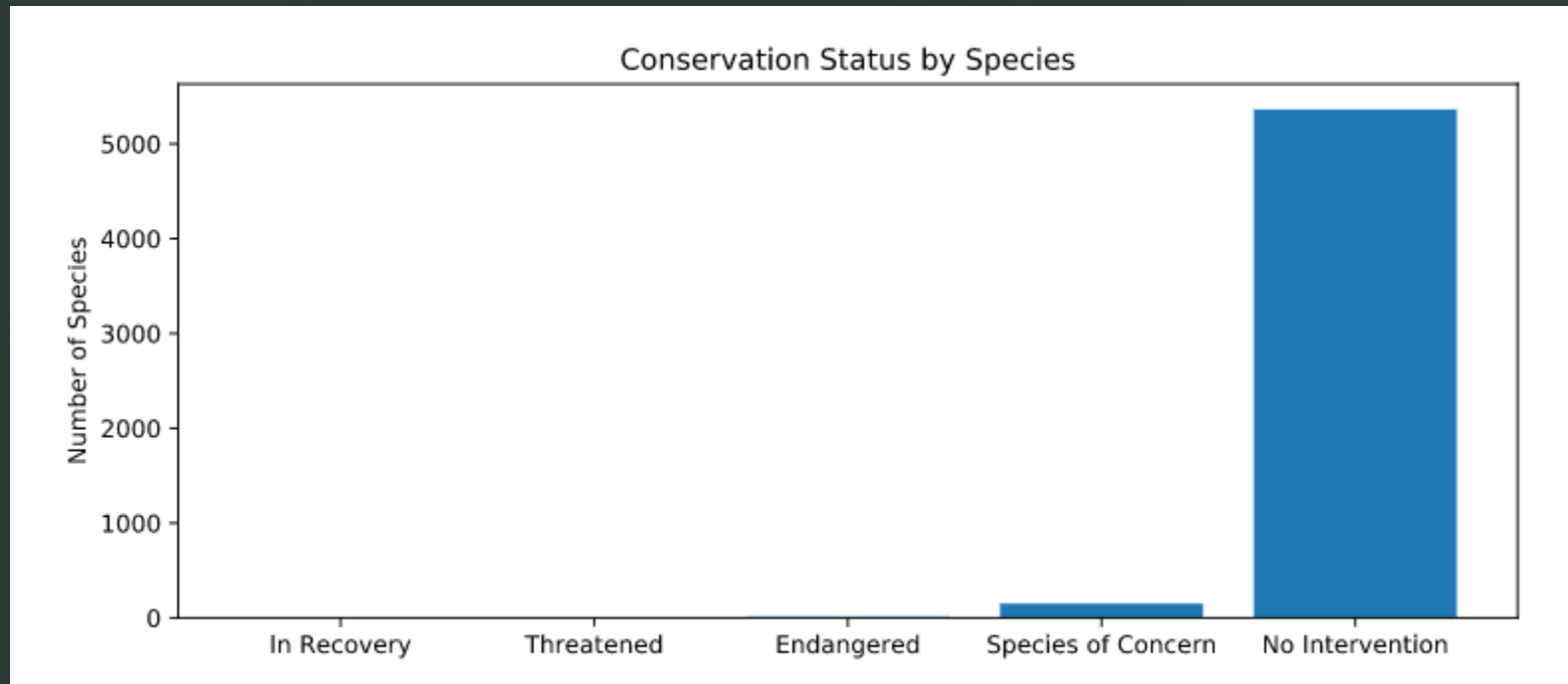
Graphs (from Codecademy)

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	Species of Concern	151
3	Threatened	10

Graphs (from Codecademy)

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	Species of Concern	151
3	Threatened	10

Graphs (from Codecademy)



Graphs (from Codecademy)

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	8.860759
1	Bird	413	75	15.368852
2	Fish	115	11	8.730159
3	Mammal	146	30	17.045455
4	Nonvascular Plant	328	5	1.501502
5	Reptile	73	5	6.410256
6	Vascular Plant	4216	46	1.079305

Graphs (from Codecademy)

	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282

Graphs (from Codecademy)

