BIODIVERSITY CAPSTONE PROJECT

-By Sharvari Sawant

Overview

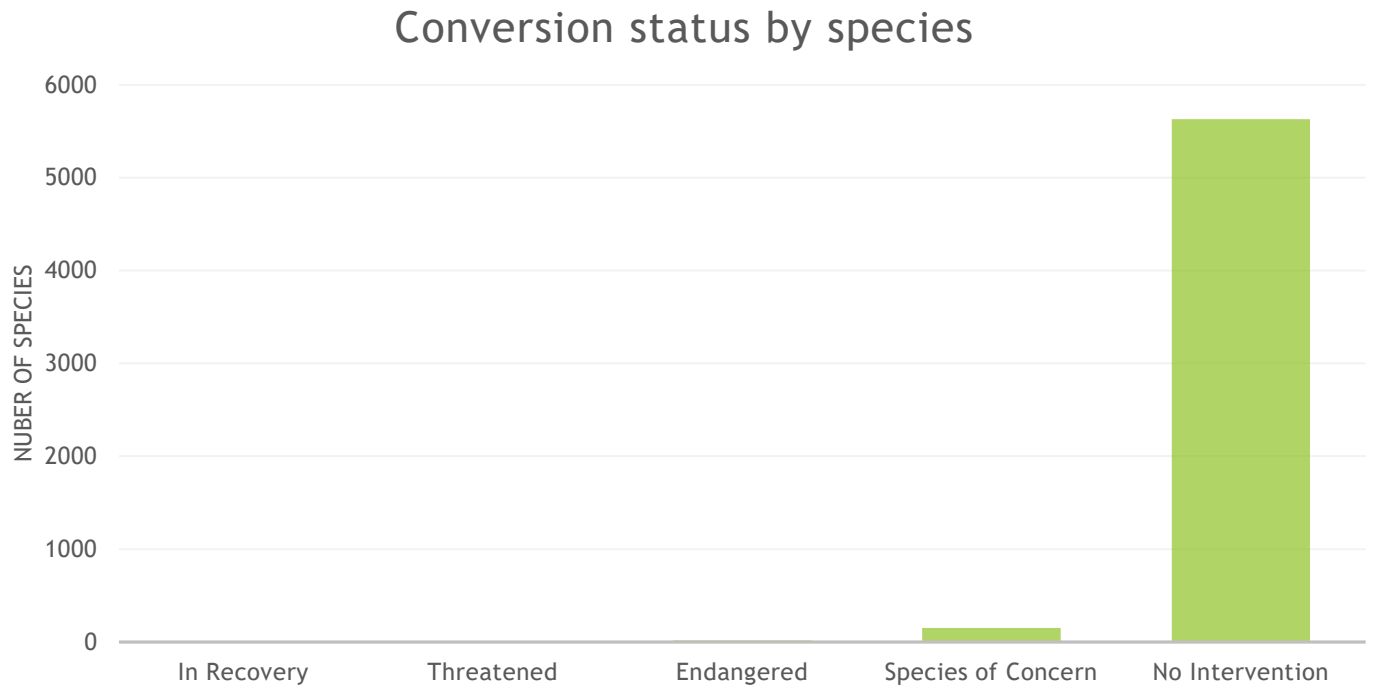
Analysis was performed on 2 sets of data

Species_info file

- It consisted of 5424 number of records
- This file provided information like the category, scientific name, common names and conservation status of the species

Observation file

- Data set consisted of 23296 records
- It provided data containing the park name scientific name and observations



- The graph shows that around 5633 species fall into “No intervention Category”
- The “In Recovery” contains the least number of species which is around 4

Types of species more likely to be endangered

- It is observed that Vascular Plant are most highly endangered, followed by birds and Nonvascular Plant

is_protected	category	not protected	protected
0	Amphibian	72	7
1	Bird	413	75
2	Fish	115	11
3	Mammal	146	30
4	Nonvascular Plant	328	5
5	Reptile	73	5
6	Vascular Plant	4216	46

Chi-Squared Test for Significance

- To prove that some species are more endangered than the other Chi- Squared Test was performed
- The result of Chi- Squared Test carried out between Birds and Mammals gave a p-value 0.68 which is greater than 0.05, hence we can say there is no significant difference
- Also, Chi- Squared Test carried out between Reptiles and Mammals, this generated a p-value of 0.038. This proves that there is a significant difference.

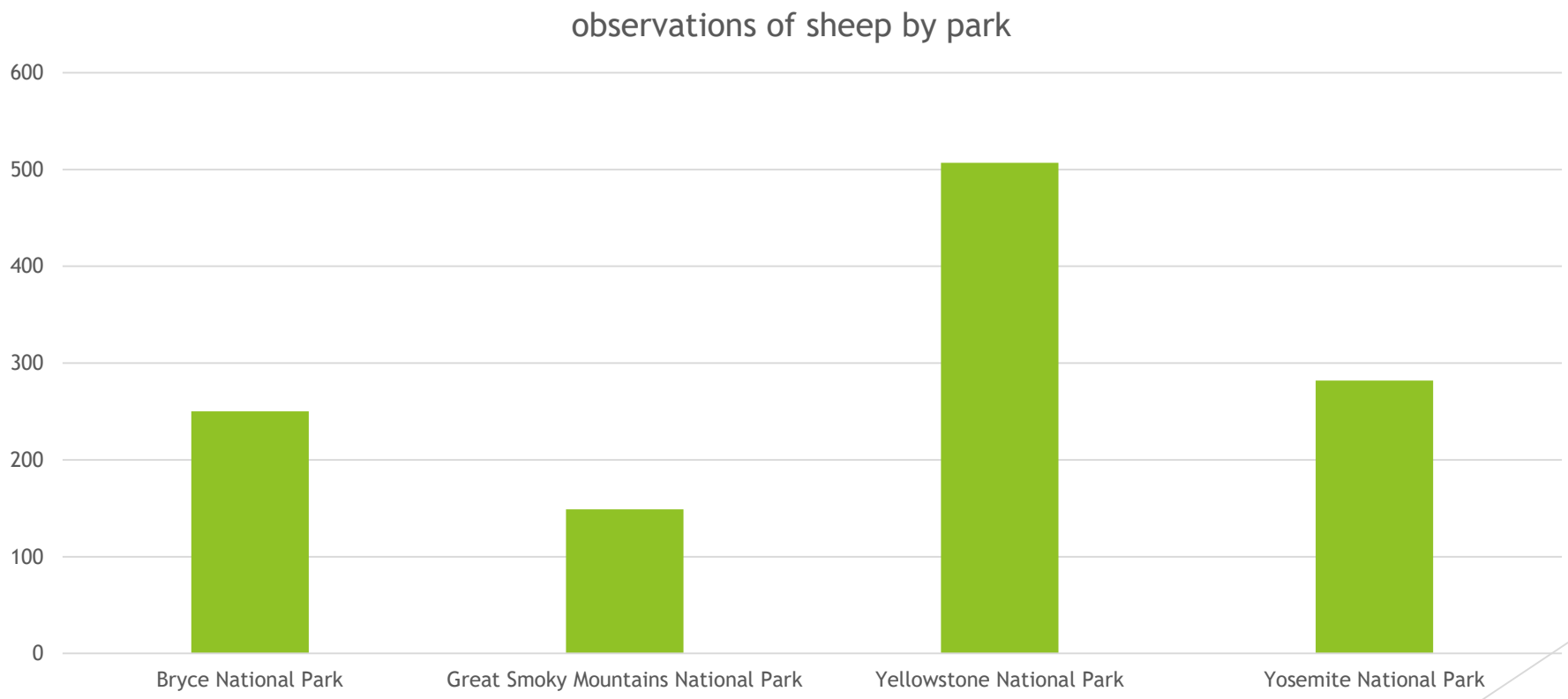
Observations DataFrame(In Search of Sheep)

- The given table describes the total number of sheep observed in each park over the past 7 days

	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

- It can be observed that the number of sheep is lowest in Great Smoky Mountains National Park

Observation of sheep by parks



Foot and Mouth Reduction Effort

- Observation:
  - If scientists want to be sure that a >5% drop in foot and mouth disease in sheep at Yellowstone is significant, then they would have to observe at least 870 sheep for the given baseline of 15%