CAPSTONE PROJECT

BIODIVERSITY

Ladies and gentlemen! Welcome to Biodiversity Capstone Project!

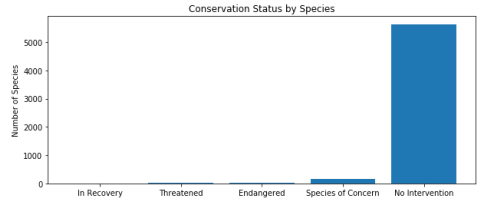
- Let me walk you through the initials of all the data gathered while performing the analysis on National Parks Service info.
- In the first part of the project I've learned that there are 5541 different species represented across National Parks and the variety of categories of those species range from 'Mammal' and 'Bird', 'Reptile' and 'Amphibian', 'Fish' and 'Vascular Plant' to 'Nonvascular Plant'.

```
  category                scientific_name  \
0   Mammal  Clethrionomys gapperi gapperi
1   Mammal                      Bos bison
2   Mammal                     Bos taurus
3   Mammal                     Ovis aries
4   Mammal                  Cervus elaphus

                                        common_names
conservation_status
0                           Gapper's Red-Backed Vole
NaN
1                              American Bison, Bison
NaN
2  Aurochs, Aurochs, Domestic Cattle (Feral), Dom...
NaN
3  Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)
NaN
4                                    Wapiti Or Elk
NaN
5541
['Mammal' 'Bird' 'Reptile' 'Amphibian' 'Fish' 'Vascular Plant'
 'Nonvascular Plant']
[nan 'Species of Concern' 'Endangered' 'Threatened' 'In
Recovery']
```

Conservation Status across different categories of species

- I’ve taken into consideration the conservation status of each species and got to know that Species of Concern are revealing the biggest number of certain species across National Parks to where endangered as well as threatened ones were not so indicative in terms of figures.
- The good news was to also learn that 5363 species proved to be the ones needing no intervention.



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

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	No Intervention	5363
3	Species of Concern	151
4	Threatened	10

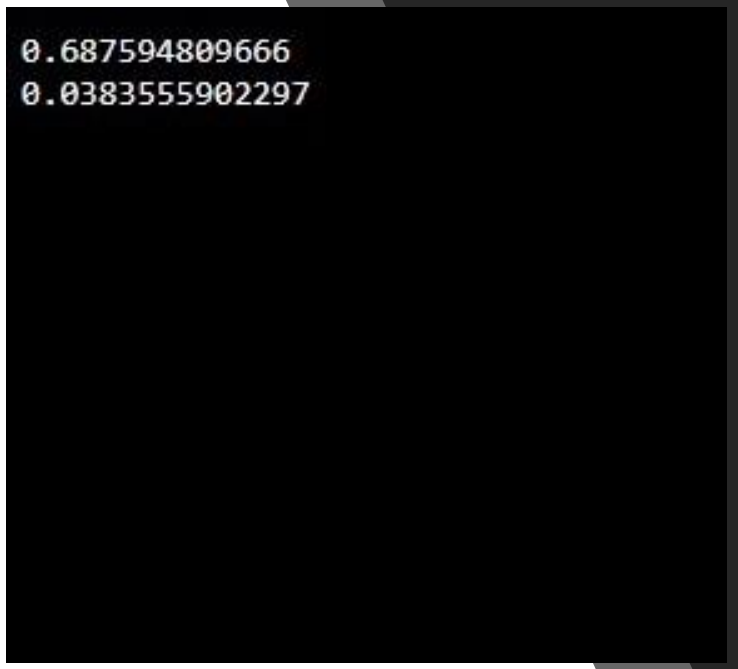
The most endangered species

- Further analysis of species has shown that certain types like ‘Nonvascular or Vascular Plants’ ‘Bird’, or ‘Mammal’ are more endangered unlike ‘Reptiles’, ‘Fish’ or ‘Amphibian’ whose figures reveal those species as more protected.

is_protected	category	False	True
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

	category	not_protected	protected	percent_protected
0	Amphibian	72	7	0.088608
1	Bird	413	75	0.153689
2	Fish	115	11	0.087302
3	Mammal	146	30	0.170455
4	Nonvascular Plant	328	5	0.015015
5	Reptile	73	5	0.064103
6	Vascular Plant	4216	46	0.010793

Significance of acquired results



- To make sure the acquired data are significant and considering the fact I was dealing with categorical data sets I decided to carry out Chi Square Test and see if the numbers I got earlier with regards to two species like ‘Mammal’ and ‘Bird’ were significant. The results of the hypothesis test revealed there was no significant difference.
- To be absolutely confident I compared two other data sets relating to ‘Mammal’ vs. ‘Reptiles’ this time. And the results of the test proved to be significant because P-value was less than 0.5

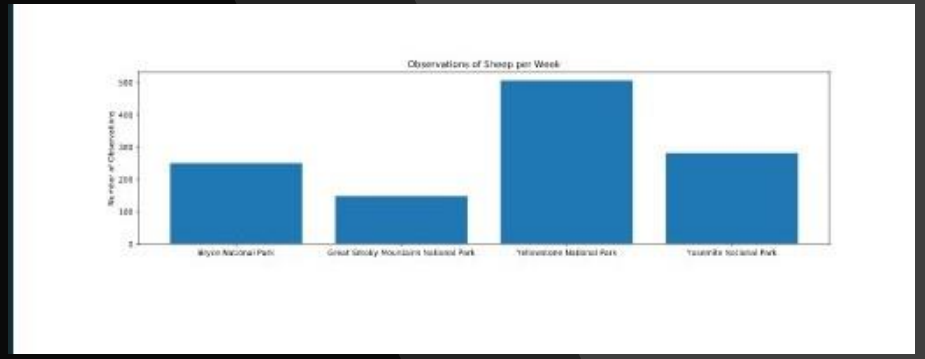
Recommendation to Conservationists

- Based on the results acquired while performing tests using data relating to different species I would recommend to the conservationists to pay more attention to such species as ‘Vascular Plants’ which occurs to be the most endangered one, just like ‘Nonvascular Plants’ and ‘Reptiles’ compared to other species analyzed where a slight difference occurring in the percentages of certain species falling into a protected category was a result of chance.

Foot and Mouth Reduction Study

- Second part of my project was devoted to analyzing data received from the conservationists who have been recording sightings of different species at several national parks for the past 7 days. A team of ruminant-enthused scientists has been tracking the movements of various species of sheep across different national parks.
- Having analyzed provided information I've discovered the total number of sheep observed in each park as shown in a bar chart and a table.

	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



Sample Size Determination

- Park Rangers at Yellowstone National Park have been running a program to reduce the rate of foot and mouth disease at that park. The scientists wanted me to test whether or not this program was working. They wanted to be able to detect reductions of at least 5 percentage point.
- Given a baseline of 15% occurrence of foot and mouth disease in sheep at Bryce National Park, I found out that if the scientists wanted to be sure that a >5% drop in observed cases of foot and mouth disease in the sheep at Yellowstone was significant they would have to observe at least 520 sheep.
- After that, using the observation data I analyzed earlier, I discovered that this would take approximately one week of observing in Yellowstone to see that amount of sheep, or approximately two weeks in Bryce park.