

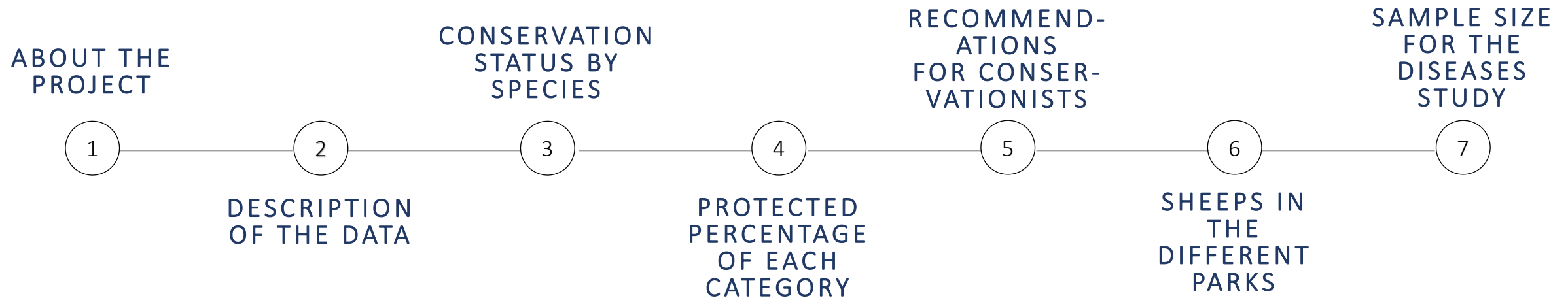
Capstone: Biodiversity for the National Parks

- Introduction to Data Analysis -



Margaux Breda – 18th of June 2018

Outline



About the Project

The **National Parks Service** would like to perform some data analysis on the conservation statuses of these species and to investigate if there are any patterns or themes to the types of species that become endangered.

Description of the Data

Species_info.csv file contains many data describing the different species in the National Parks.

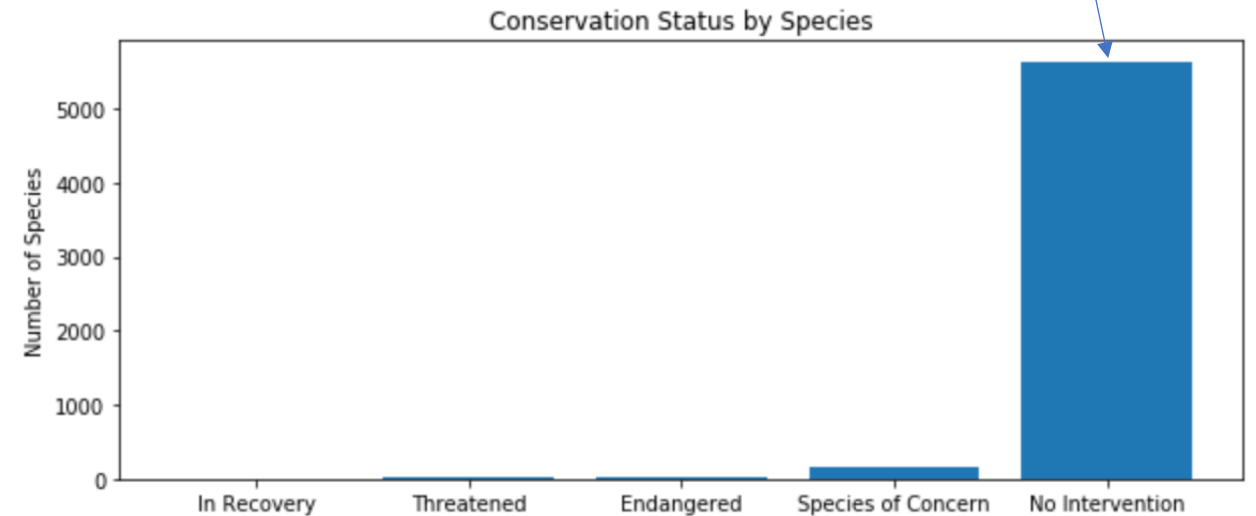
Those data include:

- The scientific name of each species (5541 different species)
- The common name of each species (common names for each scientific name)
- The species conservation status (5 different status: "Species in Concern", "Endangered", "Threatened", "In Recovery" & "NaN")
- The category of the species (7 different category: "Mammal", "Bird", "Reptile", "Amphibian", "Fish", "Vascular Plant" & "Nonvascular Plant")

Conservation Status by Species

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

Significant Number of
"No Intervention"



Significant Number of
"No Intervention"

Protected Percentage of each category

	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

Bird (15,3%) and **Mammal (17%)** are the most protected category. The difference between those two category is not significant and has been tested with a **Chi-Squared Test**.

Protected Percentage of each category

	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

Nonvascular Plant (1,5%) and **Vascular Plant (1,1%)** are the least protected category.

→ National Parks Services should invest in new conservation projects for these categories ←

Recommendations for Conservationists

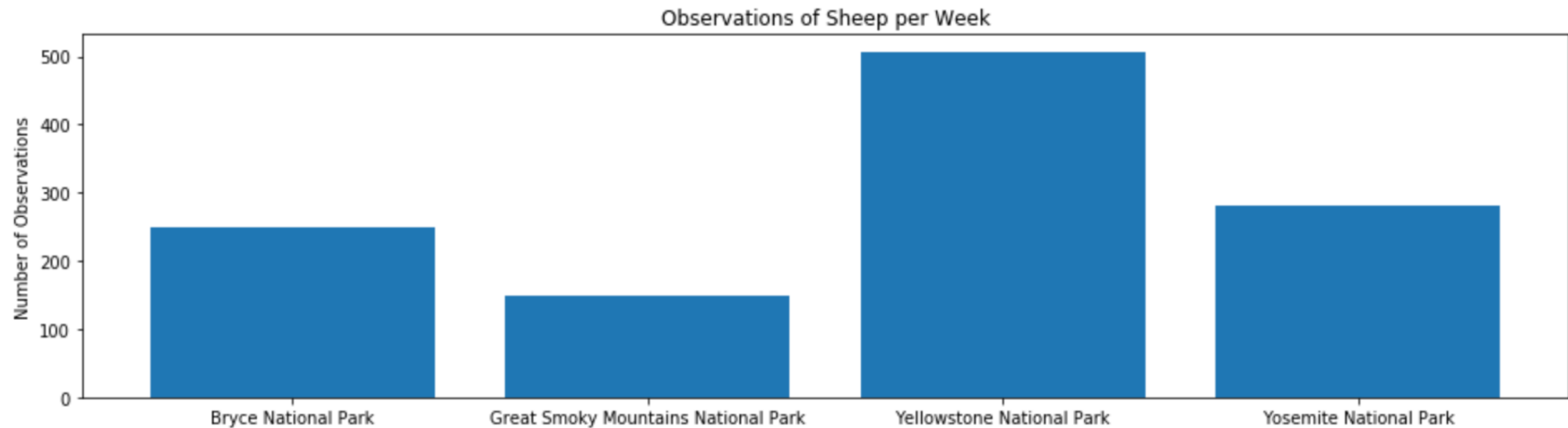
————→ **A need to better protect all the species** ←————

It is important to be conscient that all different species form a **chain** together.
Therefore, one species disappearing will also affect all the other species in the chain.

————→ **Conservationists should take a step further and investigate how to better** ←————
protect the species in the Parks

Sheep in the Different Parks

There are many different types of sheep across the parks



This **bar chart** depicts the number of sheep we observe in each park. Indeed, Yellowstone National Park has the highest number of observation of sheep per week.

Sample Size for the Diseases Study

Given a baseline of **15% occurrence** of foot and mouth disease in sheep at **Bryce National Park**, you found 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 **870 sheep**.

This would take approximately **one week** of observing in **Yellowstone** to see that many sheep, or approximately **two weeks** in **Bryce** to see that many sheep.