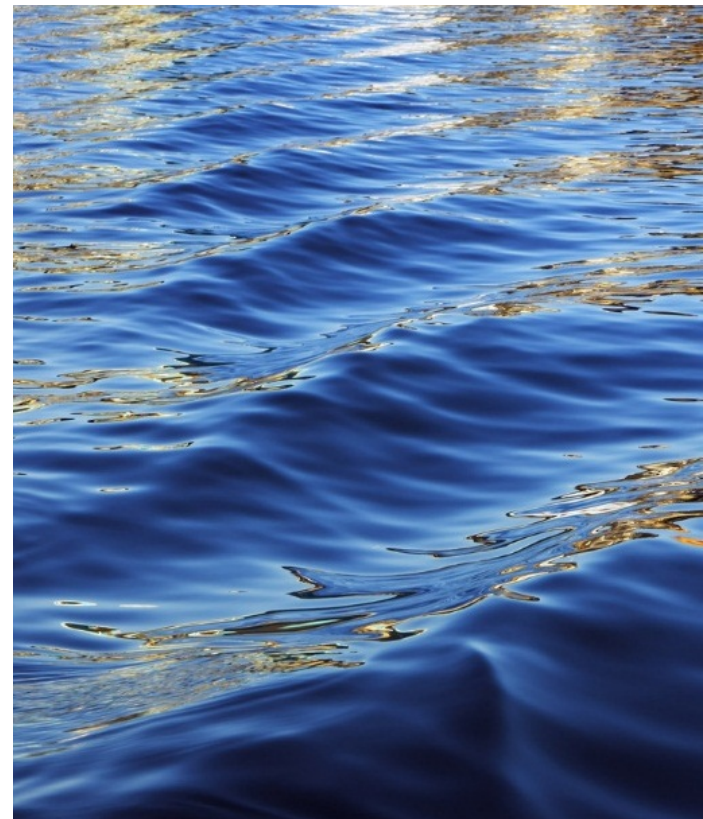




Biodiversity for the national parks

Sooyun (Christine) Lee
sooyun.career@gmail.com



Agenda

- Data description
 - Species in national park
 - 7 day observation from four national parks
- Data analysis summary
- Endangered species
 - The significance calculations for endangered status between different categories of species
 - A recommendation for conservationists concerned about endangered species
- Foot and mouth disease study
 - Sample size determination
 - A recommendation for conservationists

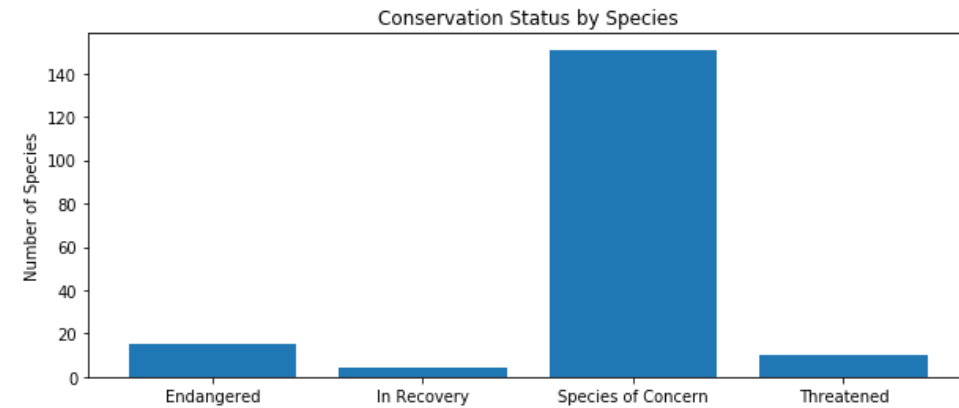
Data description

The given data shows;

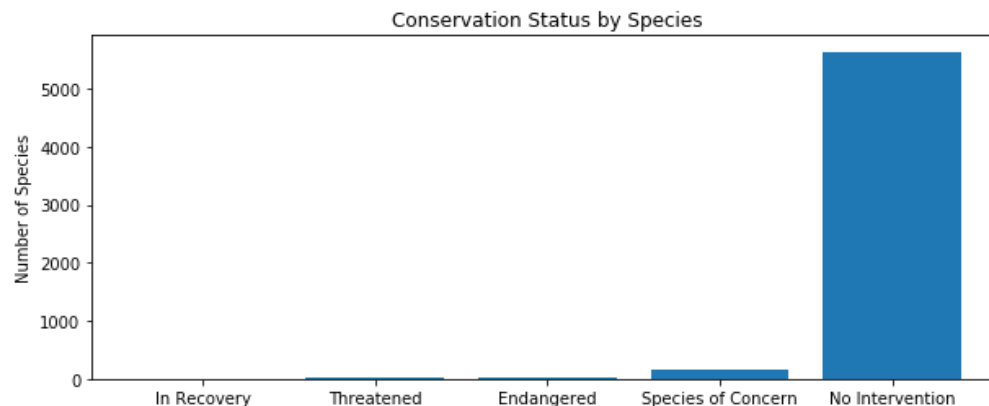
- total **5824** species (raw data) in the national parks
- **5504** unique species with their **common name**
- **5541** unique scientific name of species
- **7 species categories** — Amphibian, Bird, Mammal, Reptile, Nonvascular Plant, Vascular Plant
- Current conversation status of species in the national parks as
 - **Species of Concern** - declining population or appears to be in need of conservation.
 - **Endangered** - seriously at risk of extinction.
 - **Threatened** - vulnerable to endangerment in the near future.
 - **In Recovery** - formerly Endangered, but currently not in danger of extinction throughout all
 - **No intervention** – not categorized, no data

Data analysis I – current conversation status

Current conversation status of approximately 200 (3.5%) of species in national parks show in the given data (graph on the right).



Categorised species



It indicates the rest of species (5633) aren't categorised and marked as "no intervention". It means those species aren't being threatened and not require a protection

Data analysis II - Investigating Endangered Species

The table below shows the proportion of 'not protected' and 'protected' species by its category. By looking at the protected percentage (%) column, Mammal has the highest percentage of protected species.

Is mammal more likely to be endangered?

Category	Not protected	Protected	Protected(%)
Amphibian	72	7	8.86%
Bird	413	75	15.36%
Fish	115	11	8.73%
Mammal	146	30	17.03%
Nonvascular plant	328	5	1.5%
Reptile	73	5	6.41%
Vascular plant	4216	46	1.07%

category	not_protected	protected	percent_protected
Amphibian	72	7	8.86
Bird	413	75	15.36
Fish	115	11	8.73
Mammal	146	30	17.04
Nonvascular Plant	328	5	1.50
Reptile	73	5	6.41
Vascular Plant	4216	46	1.07

Data analysis III – Investigating Endangered Species

using Chi-Squared Test for Significance

The data shows Mammals are more likely to be endangered than Birds, but is it true? We can do a **significance test** to validate this statement. In this test, our **null hypothesis** is that this difference is due to chance

Mammal vs. Bird

	Protected	Not protected
Mammal	30	146
Bird	75	413

p-value = 0.688

Mammals are not more endangered species than Birds.

Mammal vs. Reptile

	Protected	Not protected
Mammal	30	146
Reptile	5	73

p-value = 0.038 SIGNIFICANT!

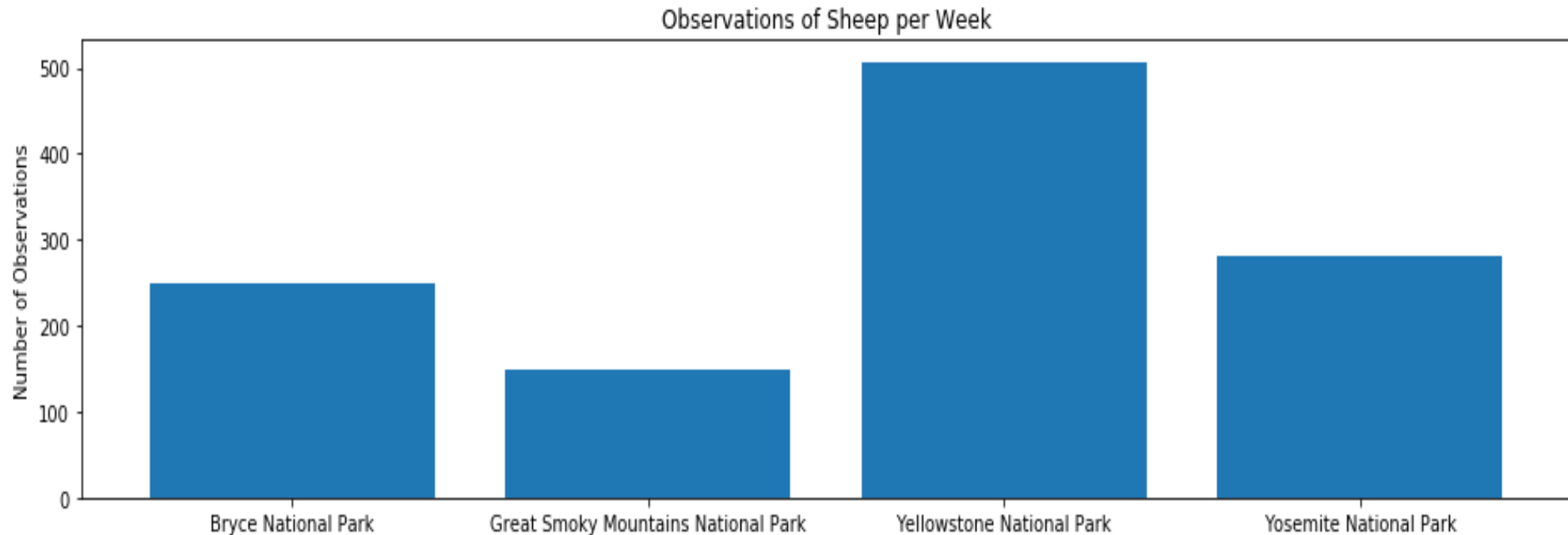
Mammals are more endangered than reptiles

Recommendation - Investigating Endangered Species

- The given data suggests, Mammals may be the most endangered species in the national parks
- However, national parks should proceed further testings using the method we used to identify the most endangered species. This way, national parks can prioritise which species need an immediate attention.
- We also recommend, national parks should have a close look at species categorised in no intervention if any of them is on the line between 'no intervention' and 'species of concerns'. Prevention is always better!

Data description - Foot and mouth disease study

The bar chart below shows number of sheep in each national parks. The chart has produced from two data sets – original species data and 7 day park observation data from each national park.



Data analysis - Foot and mouth disease study

The only information the scientists currently have about the foot and mouth disease is that last year it was recorded that **15% of sheep at Bryce National Park** have foot and mouth disease.

Park Rangers at Yellowstone National Park have been running a program to reduce the rate of foot and mouth disease at that park. The scientist wants to validate if the reduction program works.

We will determine the sample size and make a recommendation on how many weeks for selected parks need to observe enough sheep. **Bryce National Park** and **Yellowstone National Park** are selected to make a recommendation.

Foot and mouth disease study – sample size determination

- **15%** of sheep at Yellowstone national park has foot and mouth diseases
- Scientists want to know when the disease has reduced by **5% confidently**
- The recommended sample size is ***510**

*The sample size is being calculated as 870 when you do this exercise is completed from the web version

Baseline conversion rate:	<input type="text" value="15"/>
	%
Statistical significance:	<input type="button" value="85%"/> <input checked="" type="button" value="90%"/>
	<input type="button" value="95%"/>
Minimum detectable effect:	<input type="text" value="33.33"/>
	%
Sample size:	870

Recommendation - Foot and mouth disease study

- Two selected national parks for recommendation to be observed for duration as stated in the table below;

Sample size (The number of sheep must observed from each park)		510
Yellowstone national park observation duration		1 week
Bryce national park observation duration		2 weeks

- All other national parks should use this method to determine how long each of them will take to ensure enough sheep have been observed.