

# **Biodiversity for the National Parks**

[ Capstone project ]

Raphaël Grolimund

# Dataset

`species_info.csv`

includes the following information:

- category
- scientific name
- common name
- conservation status

# Categories

Possible values for category are Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, Nonvascular Plant.

# Names

Scientific names are provided for scientific soundness.

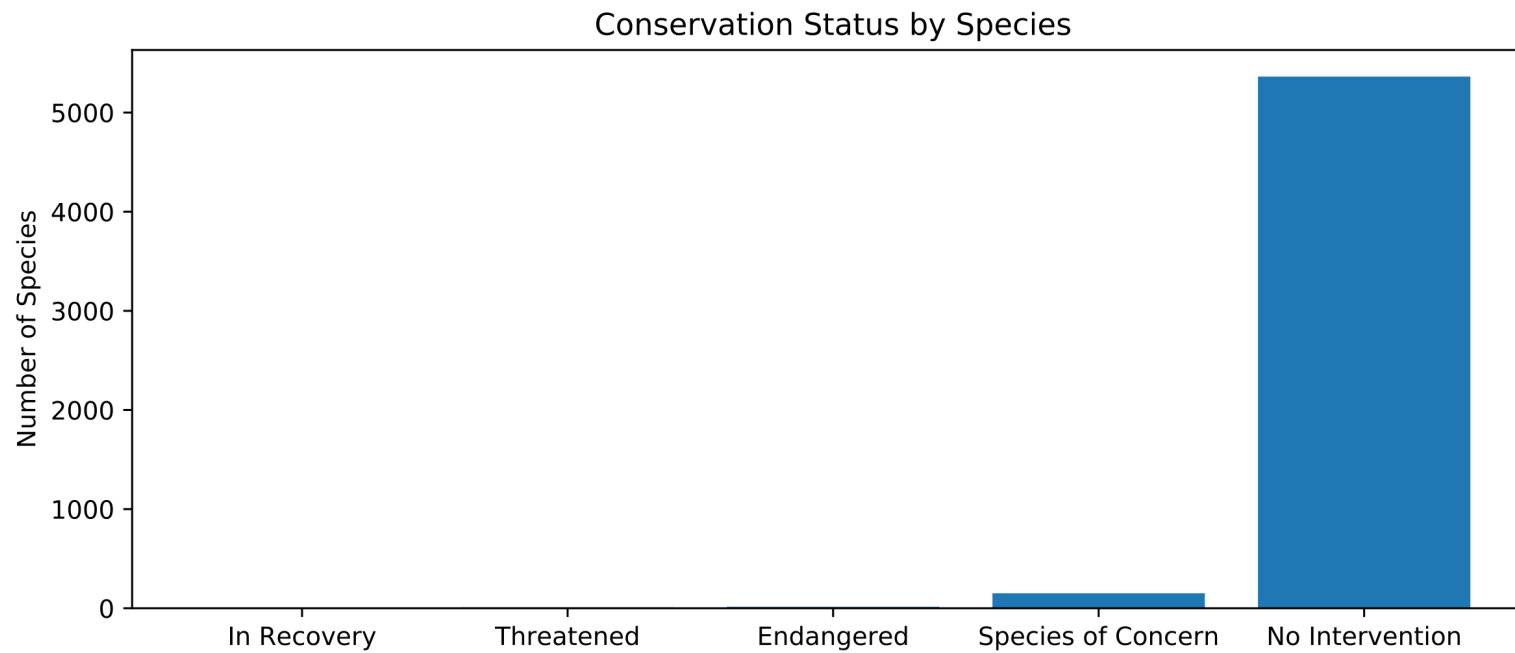
Common names are provided for easier comprehension of data.

# Status

Possible values for conservation status are Species of Concern, Endangered, Threatened, and In Recovery.

Note that rows can have no conservation status (NaN).

# Endangered species



<b>category</b>	<b>not_protected</b>	<b>protected</b>
Amphibian	72	7
Bird	413	75
Fish	115	11
Mammal	146	30
Nonvascular Plant	328	5
Reptile	73	5
Vascular Plant	4216	46

<b>category</b>	<b>%</b>
Amphibian	0.088608
Bird	0.153689
Fish	0.087302
Mammal	0.170455
Nonvascular Plant	0.015015
Reptile	0.064103
Vascular Plant	0.010793



# Conclusions

Mammals are **NOT** significantly more endangered than birds ( $p=0.687$ ).

But mammals are significantly more endangered than reptiles ( $p=0.038$ ).

# **Recommendations for conservationists**

Take great care of mammals and birds!

A lot of species in these categories are endangered.

# Foot and mouth disease

sample size determination

baseline = 15%

minimum detectable effect = 33.33%

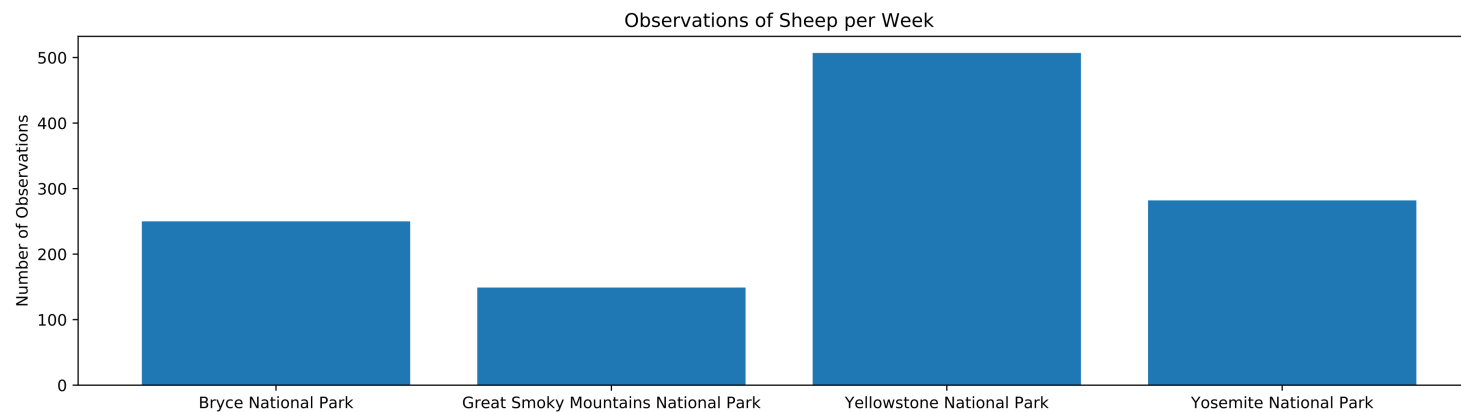
Statistical Significance = 90%

**sample size = 520**

# Observations

Bryce National Park: 250 per week

Yellowstone National Park: 507 per week



# **Recommended observation time**

2 weeks at Bryce National Park

1 week at Yellowstone National Park