Biodiversity for the National Parks

Victor R. Naive Jr.

What’s in the species_info.csv?

The csv contains the different species in the National Park. This has the following columns:

-category

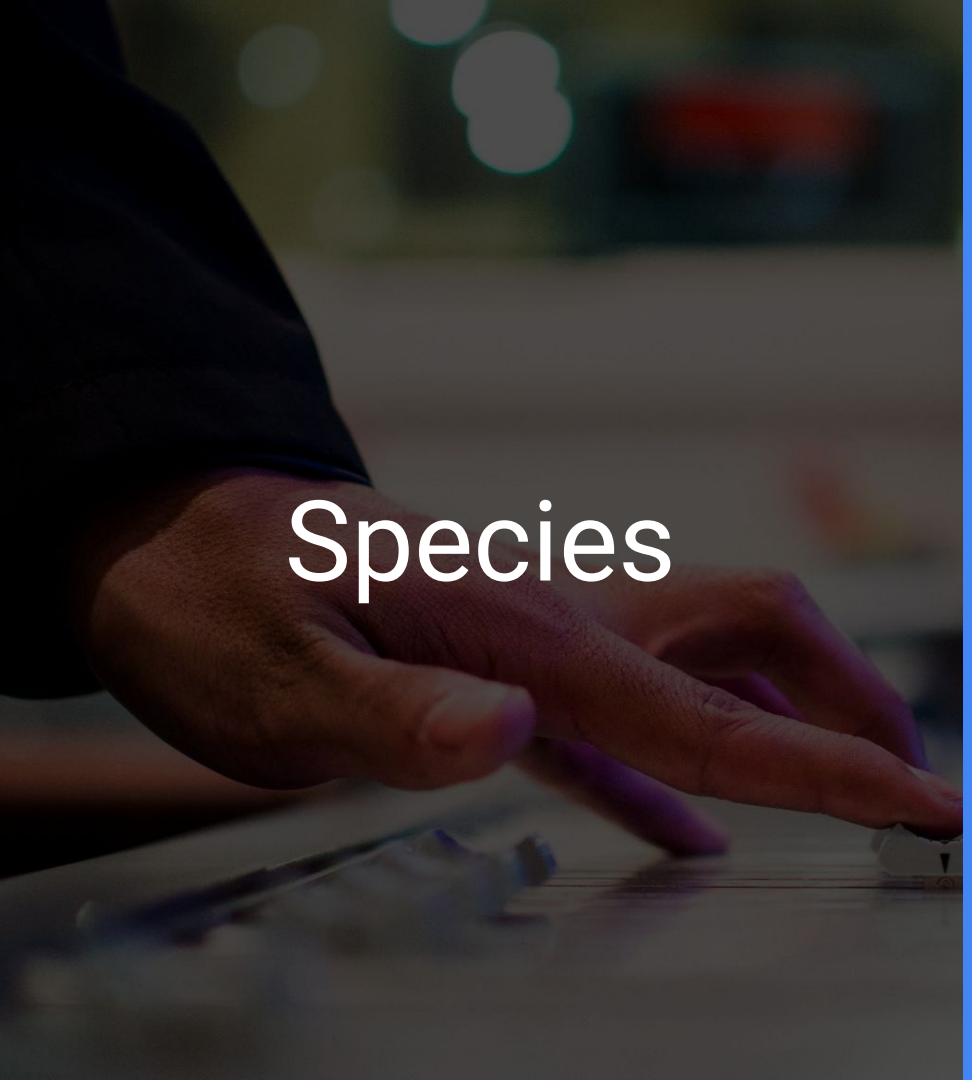
-scientific_name

-common_names

-conservation_status

Out[2]:

	category	scientific_name	common_names	conservation_status
0	Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole	NaN
1	Mammal	Bos bison	American Bison, Bison	NaN
2	Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Dom...	NaN
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	NaN
4	Mammal	Cervus elaphus	Wapiti Or Elk	NaN
5	Mammal	Odocoileus virginianus	White-Tailed Deer	NaN
6	Mammal	Sus scrofa	Feral Hog, Wild Pig	NaN
7	Mammal	Canis latrans	Coyote	Species of Concern
8	Mammal	Canis lupus	Gray Wolf	Endangered
9	Mammal	Canis rufus	Red Wolf	Endangered
10	Mammal	Urocyon cinereoargenteus	Common Gray Fox, Gray Fox	NaN
11	Mammal	Vulpes fulva	Black Fox, Cross Fox, Red Fox, Silver Fox	NaN
12	Mammal	Vulpes vulpes	Red Fox	NaN
13	Mammal	Felis concolor	Mountain Lion	NaN
14	Mammal	Felis silvestris	Wild Cat, Wildcat	NaN
15	Mammal	Lynx rufus	Bobcat	NaN
16	Mammal	Puma concolor	Panther (Mountain Lion)	NaN
17	Mammal	Mephitis mephitis	Striped Skunk	NaN
18	Mammal	Spilogale putorius	Eastern Spotted Skunk	NaN
19	Mammal	Lontra canadensis	River Otter	NaN
20	Mammal	Lutra canadensis	Northern River Otter	NaN



Species

There are **5541** different number of species

Categories of Species

There are (7) seven different values of categories in species.

Mammal

nourish their young with milk secreted by mammary glands, have the skin usually more or less covered with hair, and include humans.

Bird

also known as Aves, are a group of endothermic vertebrates, characterised by feathers.

Reptile

are tetrapod animals in the class Reptilia, comprising today's turtles, crocodilians, snakes, amphisbaenians, lizards, tuatara, and their extinct relatives.

Amphibians

are ectothermic, tetrapod vertebrates of the class Amphibia.

Categories of Species

There are (7) seven different values of categories in species.

Fish

are gill-bearing aquatic craniate animals that lack limbs with digits.

Non Vascular Plant

also known as Aves, are a group of endothermic vertebrates, characterised by feathers.

Vascular Plant

also known as tracheophytes and also as higher plants

```
Out[6]: array([nan, 'Species of Concern', 'Endangered', 'Threatened', 'In Recovery'], dtype=object)
```

Conservation

Different values of conservation status

NaN

None

Species of Concern

declining or appear 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 neither in danger of extinction throughout all or a significant portion of its range

5363

No Intervention are being made to 5633 species.

The data shows that most of the species has no intervention yet. We still don’t know the conservation status of the species. Thus, it is necessary to conduct interventions the soonest.

Significance Calculations

Null Hypothesis: It looks like species in category Mammal are more likely to be endangered than species in Bird.

I did a significance test calculations for Mammal and Bird species using Chi Square test. Have created a contingency table using chi2_contingency function in Python.

	protected	not protected
Mammal	?	?
Bird	?	?

```
In [63]: from scipy.stats import chi2_contingency
```

Significance Calculations

After the test, it showed that there is no significant difference between Mammal and Bird.Therefore, we reject the null hypothesis.

In [58]:
```
pval = chi2_contingency(contingency)[1]
print (pval)

if pval < 0.05:
    print ("There is significant difference between mammals and birds")
else:
    print ("There is no significant difference between mammals and birds")
```

```
0.6875948096661336
There is no significant difference between mammals and birds
```

It looks like this difference isn't significant!

Difference between Reptile & Mammal?

Have took another test for Reptile and Mammal.

Null Hypothesis:

There is a significant difference between Reptile & Mammal.

Difference between Reptile & Mammal?

```
contingency_rep_mammal = [30, 146], [5, 73]
pval_rep_mammal = chi2_contingency(contingency_rep_mammal)[1]
pval_rep_mammal

if pval_rep_mammal  < 0.05:
    print ("There is significant difference between mammals and reptile")
else:
    print ("There is no significant difference between mammals and reptile")
```

```
There is significant difference between mammals and reptile
```

Based on the result, we are going to accept the null hypothesis.

Recommendation

Based on the results, the category that has least protection are plants namely Vascular and Nonvascular plants. Next to these is the Reptiles.

```
In [18]: category_pivot
```

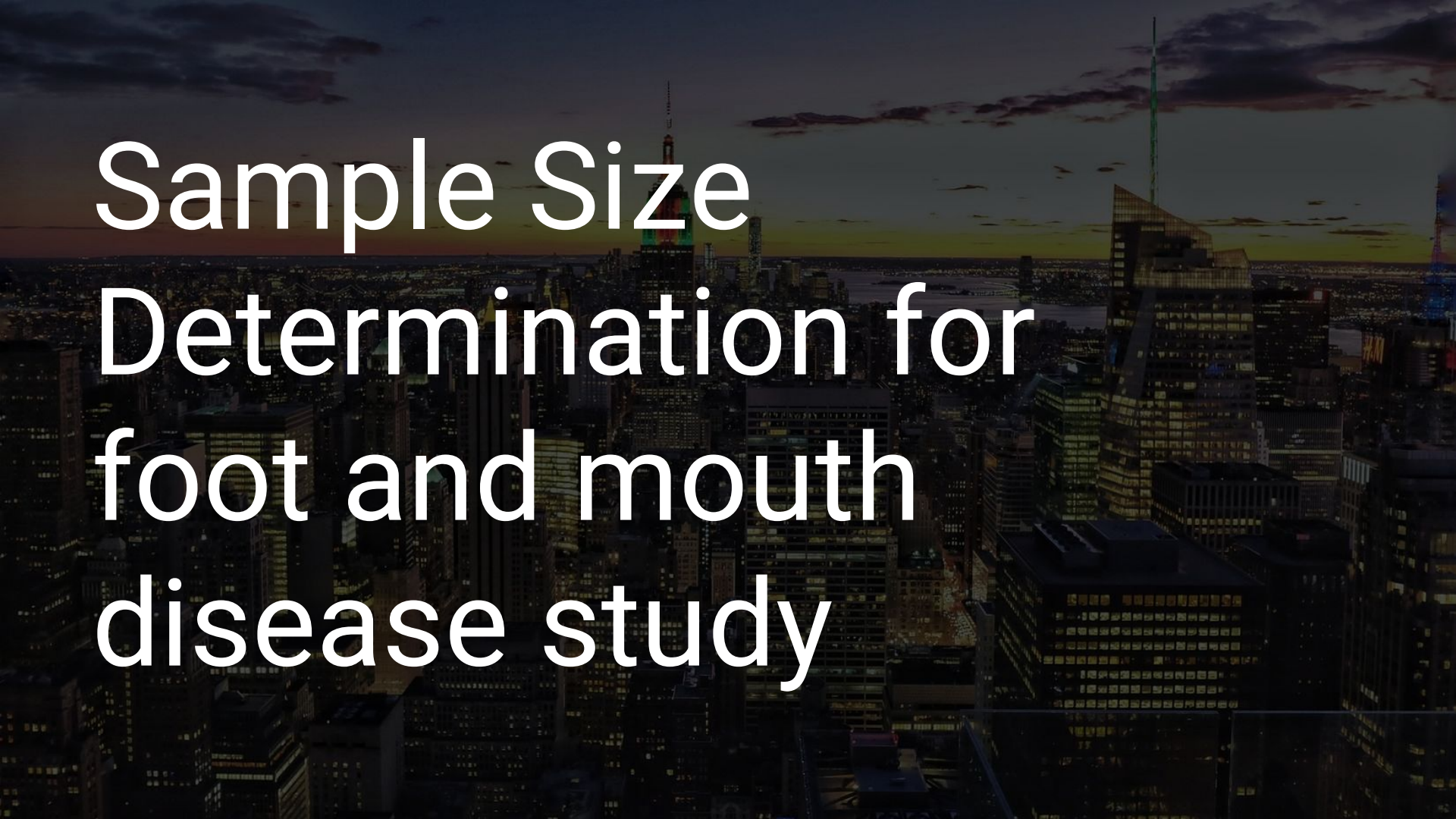
Out[18]:

	category	not_protected	protected	percent_protected
0	Amphibian	73	7	0.087500
1	Bird	442	79	0.151631
2	Fish	116	11	0.086614
3	Mammal	176	38	0.177570
4	Nonvascular Plant	328	5	0.015015
5	Reptile	74	5	0.063291
6	Vascular Plant	4424	46	0.010291

Recommendation

I highly recommend to do intervention and protection to our plants both Vascular and Nonvascular plants. Aside from the fact that they have given least attention, they are also the food of the aforementioned animals.

Also, give attention to Reptiles. Extinction of these species may lead to imbalance of our ecosystem.


Sample Size
Determination for
foot and mouth
disease study

Minimum Detectable Effect

Our scientists know 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 scientists want to test whether or not this program is working. They want to be able to detect reductions of at least 5 percentage points. For instance, if 10% of sheep in Yellowstone have foot and mouth disease, they'd like to be able to know this, with confidence.

Use Codecademy's sample size calculator to calculate the number of sheep that they would need to observe from each park. Use the default level of significance (90%).

33.33 % Minimum Detectable Effect

We used the 15% baseline in calculating the Minimum Detectable Effect

Minimum Detectable Effect = 100 * (old - new) / old

Minimum Detectable Effect = 100 * (15-10) / 10

33.33 % Minimum Detectable Effect

```
In [2]: baseline = 15

        minimumDetectableEffect = 100 * (15-10) / 15 #old-new/old
        minimumDetectableEffect


        minimumDetectableEffect

Out[2]: 33.333333333333336
```

Number of Weeks

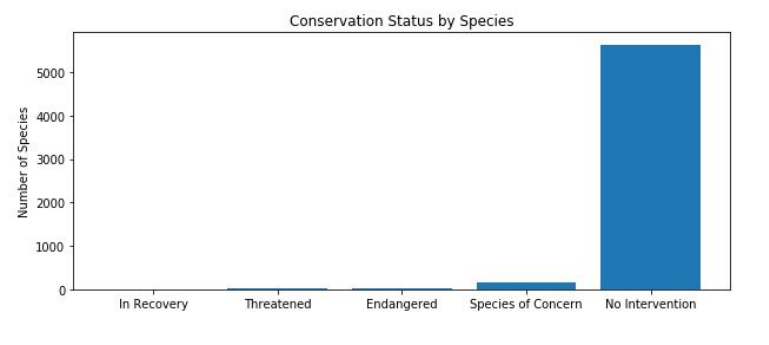
How many weeks would you need to observe sheep at Bryce National Park in order to observe enough sheep? How many weeks would you need to observe at Yellowstone National Park to observe enough sheep?

```
In [5]: samplesize = 870
        bryce = 870 / 250.
        yellowstone = 810 / 507

In [6]: bryce
Out[6]: 3.48

In [8]: yellowstone
Out[8]: 1.5976331360946745
```

Number of Weeks

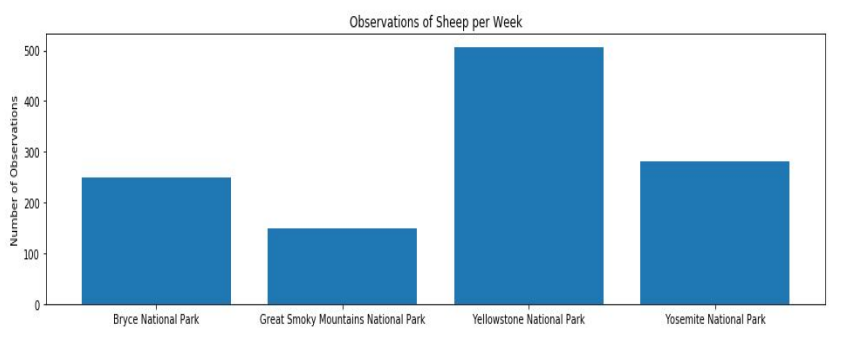
It would take about 3.5 weeks for bryce and 1.5 weeks for yellowstone.

BIODIVERSITY GRAPHS



Conservation Status by Species

The graph shows that that most of the species has No Intervention.



Conservation of Sheep per Week

The graph shows that Yellowstone has the greatest number of observations.

END