Biodiversity Capstone Project Slide Deck - Codeacademy Data Analysis

An analysis written by Luciano Leone

Investigating Protected Species 1/3

Calculations' observations:

Figure 1:

	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

Looking at the following data frame, we assumed that Mammals were more inclined to be enderanged than birds. During the project, we ran a significance test to evaluate whether or not it was true. We ran a chi squared test - the null hypothesis is that this difference is due to chance.

We ran 2 tests comparing the following species:

- Bird and Mammal
- Reptile and Mammal

Investigating Protected Species 2/3

Results

Here below, you will find the results related to the two Chi-squared tests:

➔ **Bird and Mammal**

P-value after the chi-squared test: **~0.688**
We can conclude that the difference between the percentages of protected birds and mammals is not significant and is a result of chance.

➔ **Reptile and Mammal**

P-value after the chi-squared test: **~0.038**
The difference is significant and can conclude that certain types of species are more likely to be endangered than others.

Investigating Protected Species 3/3

A Recommendation for conservationists:

We only ran Chi-squared tests for a few species, I will advise them to run Chi-squared tests for all the other species and test whether or not there is a difference.

After the results, they will have a clearer picture which species are more inclined to be endangered.

Foot and Mouth Reduction Effort 1/2

Description:

Park Rangers at Yellowstone National park have been running a program to reduce the rate of foot and mouth disease at the park. The scientists want to know whether or not the program is working.
They want to be able to detect reductions of at least 5 percentage point. For instance, if 10% of sheep in Yellowstone have the disease, they would like to know it with confidence.

The only information in our possession is last year 15% of sheep at Bryce had the disease. Using an A/B Test Sample Size Calculator, we need to calculate the number of sheeps they would need to observe from each park to make sure their foot and mouth percentages are significant and using a level of significance of 90%

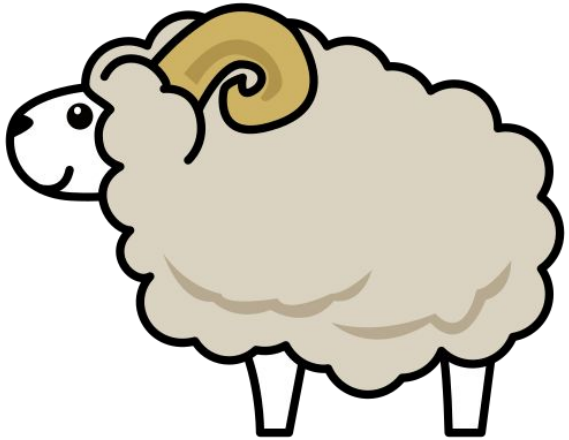
Foot and Mouth Reduction Effort 2/2

Calculation:

First, we needed the baseline which was given to us: 15%

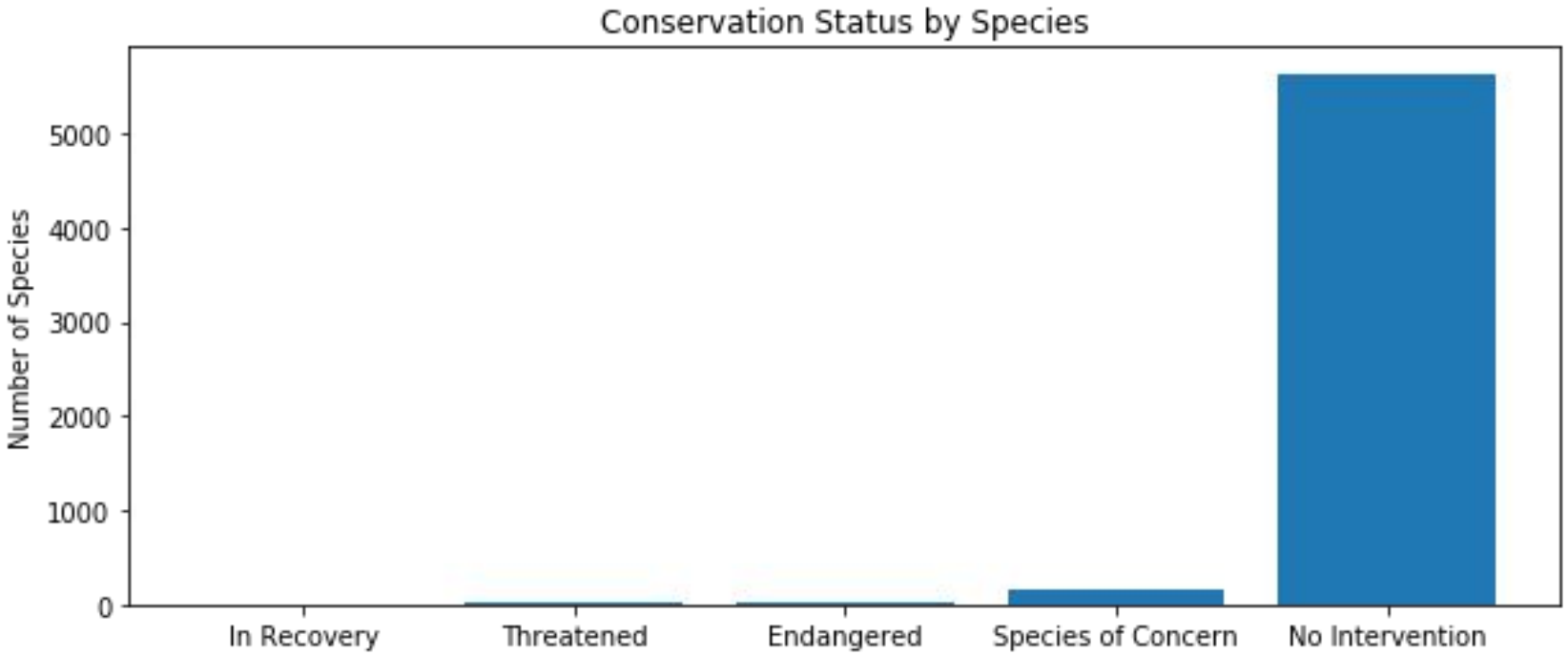
Then, we had to calculate the minimum detectable effect, remember the scientists want to be able to detect reductions of at least 5 percentage point.

Minimum Detectable Effect is a percent of the baseline, so if we wants to observe 5% change , our minimum detectable would be 100 * (0.05/0.15), which give us: 33.333%



Then we can implement all the data into the A/B Test Sample Size Calculator which gives us the sample size of 510. Then we divide this number by the number of observations seen for the past 7 days (i.e. for Yellowstone 507). We can say that scientists will need to spend 1 week at Yellowstone in order to observe enough sheep. We can obviously use the sample size of 510 to calculate for the other parks.

Graphs created in the notebook 1/2



Graphs created in the notebook 2/2