Insights on biodiversity for National Parks

By: Cheuk Kwan Wong (Andrea)

Instructions

- A section describing the data in **species_info.csv**. Be sure to include some (or all) of what you noticed while working through the notebook.
- A section describing the significance calculations that you did for endangered status between different categories of species.
- A recommendation for conservationists concerned about endangered species, based on your significance calculations
- A section describing the sample size determination that you did for the foot and mouth disease study
- All of the graphs that you created in the notebook

Overview

1. Data in Species_info.csv
2. Breakdown of species by conservation status
3. Endangered Species by category
4. Chi-Squared Test for Significance (Mammals & Birds)
5. Chi-Squared Test for Significance (Mammals & Reptiles)
6. Chi-Squared Test for Significance (Mammals & Amphibians)
7. Chi-Squared Test for Significance Conclusion
8. Foot and Mouth Disease Reduction Effort at National Parks
9. No. of weeks observing the sheeps

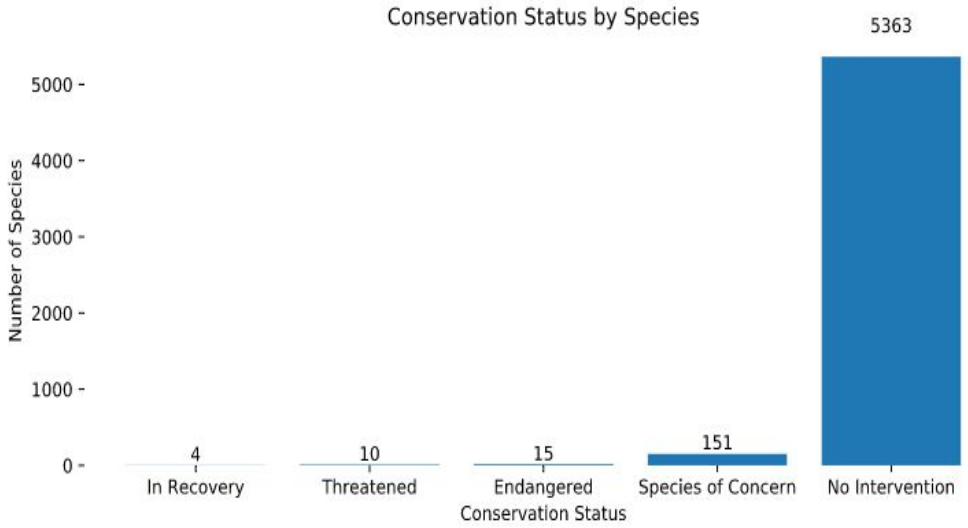
Data in Species_info.csv

The data in Species_info.csv includes :

Column	Data Type
Category (Amphibian, Bird, Fish, Mammal, Nonvascular Plant, Reptile, Vascular Plant)	String
Scientific Names	String
Common Names	String
Conservation Status (Endangered, In recovery, Species of Concern, Threatened, NaN (No Intervention))	String

Breakdown of species by conservation status

	Conservation Status	Species	Percentage (%)
0	In Recovery	4	0.07
1	Threatened	10	0.18
2	Endangered	15	0.27
3	Species of Concern	151	2.72
4	No Intervention	5363	96.75



Species that fall under ‘No Intervention’ takes up 96.75% followed by ‘Species of Concern that take up 2.72%

Endangered Species by category

Next, looking at the types of species that are listed under 'Endangered', I wanted to explore whether certain categories of species are more likely to be endangered.

	Category	Not Protected	Protected	Percent Protected (%)
0	Amphibian	72	7	0.089
1	Bird	413	75	0.154
2	Fish	115	11	0.087
3	Mammal	146	30	0.170
4	Nonvascular Plant	328	5	0.015
5	Reptile	73	5	0.064
6	Vascular Plant	4216	46	0.011

Mammals have a higher percentage of protected species as compared to the other categories.

Significance tests will still need to be conducted to determine whether certain types of species are more likely to be endangered than others.

Chi-Squared Test for Significance (Mammals & Birds)

A chi square test was used to determine whether there is a significant association between two variables. In this case, between two categories .

Null hypothesis Ho = The difference in the percentages of mammals and birds is a result of chance.

Alternative hypothesis H1 = The difference in the percentages of mammals and birds is not a result of chance.

Contingency Table

Category	Protected	Not Protected
Mammals	30	146
Bird	75	413

P value = 0.68759

There is no significant difference as the P value is > 0.05. The difference between the percentages of protected birds and mammals is not significant and is a result of chance.

Chi-Squared Test for Significance (Mammals & Reptiles)

Contingency Table

Category	Protected	Not Protected
Mammals	30	146
Reptiles	5	73

P value = 0.03836

There is a significant difference as the P value is < 0.05. The difference between the percentages of protected birds and mammals is significant and not a result of chance.

Chi-Squared Test for Significance (Mammals & Amphibians)

Contingency Table

Category	Protected	Not Protected
Mammals	30	146
Amphibians	7	72

P value = 0.12758

There is no significant difference as the P value is < 0.05. The difference between the percentages of protected birds and mammals is not significant and a result of chance.

Chi-Squared Test for Significance Conclusion

Category	Category	P Value
Mammals	Amphibians	0.12758
Mammals	Reptiles	0.03836
Mammals	Birds	0.68759
Mammals	Fish	0.68759

For conservationists concerned about endangered species, based on the significance calculations, we can conclude that certain types of species are more likely to be endangered than others.

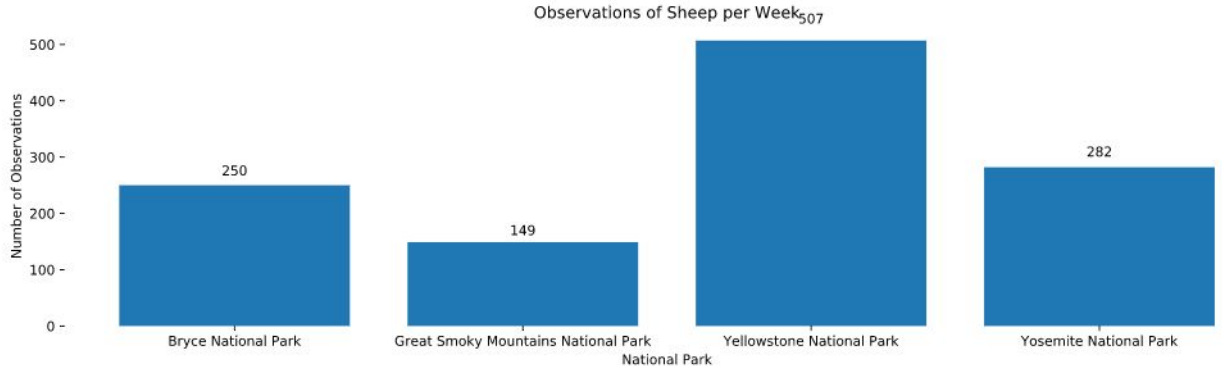
Foot and Mouth Disease Program at Yellowstone National Park

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. In order to measure the the reduction effort of the park, we will need:

1) Details on sheep sightings at different national parks
2) Determine the sample size for the calculations

Foot and Mouth Disease Reduction Effort at National Parks

Sheep Sightings by National Park



National Park	Bryce National Park	Great Smoky Mountains National Park	Yellowstone National Park	Yosemite National Park
Number of sheep observations per week	250	149	507	282

Sample Size Determination

1) Baseline percentage for sample size determination
    a) 15% (last year it was recorded that 15% of sheep at Bryce National Park have foot and mouth disease.)
2) Minimum Detectable Effect (Scientists want to detect reductions of at least 5 percentage points.)
    a) 100 * X / Baseline = (100*5)/15 = 33.333333
3) Statistical Significance = 90%
4) Sample Size per Variant = 870

No. of weeks observing the sheeps

After determining the sample size per variant, we can now estimate how many weeks the scientists will need to spend at the parkes to gather the number of observations.

Yellowstone National Park = 870 / 507 = 1.7 Weeks

Bryce National Park = 870 / 250 = 3.48 Weeks

End!

Thank You