Biodiversity for the National Parks

Endangered Species Analysis Report

Noriko Nakayama

10/02/2018

Summary of Data in “species_info.csv”

4 columns

5824 rows

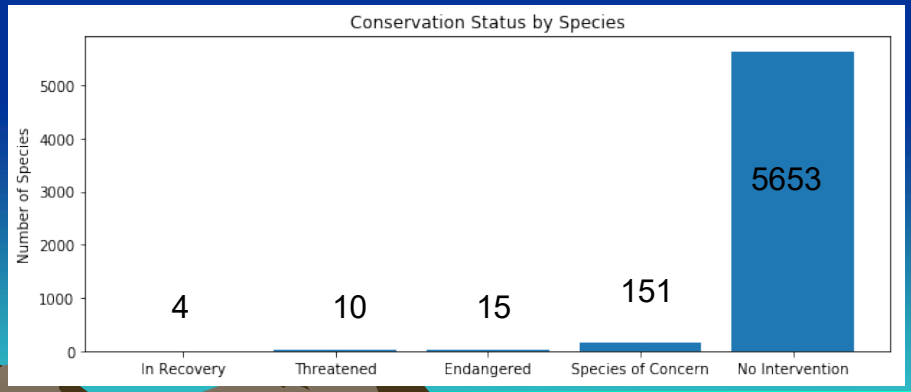
category	scientific_name	common_names	conservation_status

7kinds of values : 'Mammal'
'Bird‘
'Reptile'
'Amphibian'
'Fish'
'Vascular Plant'
'Nonvascular Plant'

The number of unique names is 5541

4 kinds of values : 'Species of Concern'
‘Endangered'
'Threatened'
'In Recovery‘
NaN case means “No Intervention”.

How many species are in each categories ?



Which category is more likely to be endangered?

Using the dataset, the percentage of protected species for each category can be calculated.

category	not_protected	protected	percent_protected
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015
Reptile	73	5	0.064103
Vascular Plant	4216	46	0.010793

The above data says mammal has the highest percentage. Let us check whether the difference is significant or not.

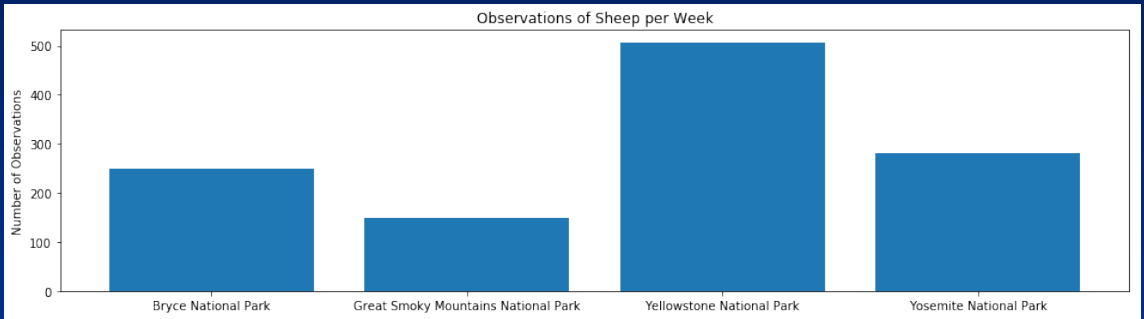
The data is numerical or categorical? -> categorical
How many pieces of data are you comparing? -> two
Select *"chi squared test"*.

Results are follows:
Reptile vs mammal-> Pval is 0.0383555902297. Significant!
Amphibian vs mammal-> Pval is 0.127576696932. Not Significant!
Bird vs mammal-> Pval is0.687594809666. Not Significant!

Mammals are in more serious condition than reptile.
More efforts should be made for mammal protection.
Amphibian and bird may be in the same danger level as mammal.

Sample size needed for foot and mouth disease program

Combining the data in “species_info.csv” with the data in “observations.csv”, we can get the number of sheep sightings at different national parks for the past 7 days



The 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 detect reductions of at least 5 percentage points.
How many weeks would you need to observe sheep to know whether or not this program is working with confidence level 90%?

For Bryce National Park, needed sample size is calculated as 890 from sample size calculator, that is approximately 2.5 weeks.

For Yellowstone National Park, 10% of sheep have foot and mouth disease. needed sample size is calculated as 890 from sample size calculator, that is approximately 1.8 weeks.