APRIL 2018   CHRISTINA LEE

BIODIVERSITY FOR THE NATIONAL PARKS

CAPSTONE PROJECT 2

OBJECTIVE

▶ **PART I**
Analyze the data on the conservation statuses on endangered species from several National Parks

　　▶ Examine data to find how many species are endangered and how many species are protected by their species categories

　　▶ Can we find that certain types of species are more or less likely to be endangered than others?

▶ **PART II**
Help the National Park Service track the movements of all sheep species across these parks

　　▶ Find where all species of sheep were accounted for and combine these numbers by the various National Parks.

　　▶ Use this information to run a sample size determination to help reduce the rate of foot and mouth disease in these sheep.

PART I: INVESTIGATING PROTECTED SPECIES

Analyze the data on the conservation statuses on Endangered species from National Parks in the US

▶ In the Species Info data, we are provided with data that includes species' scientific names, common names and their conservation statuses.

▶ After sorting all species by their conservation statuses, we found 180 species identified as in recovery, threatened, endangered or species of concern.
No Intervention replaced the species that had N/A in its row.

▶ In order to see whether certain types of species were more likely to be endangered, we organized the data into a pivot to show species with the number of those protected / not protected with their percentages to run a chi squared test to see whether or not some species were more or less likely to become endangered
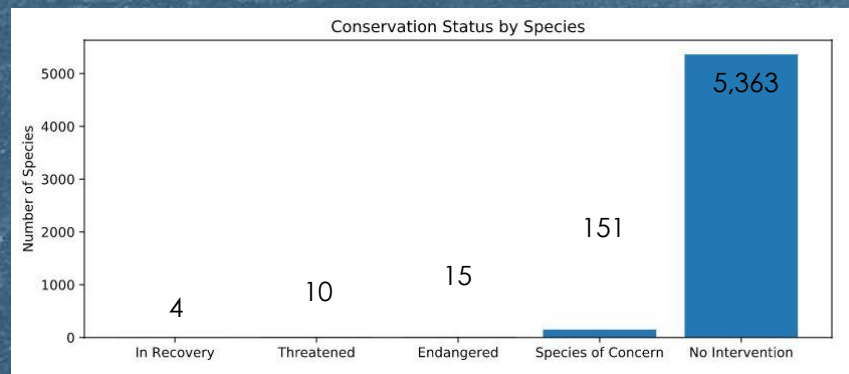
CATEGORY	NOT PROTECTED	PROTECTED	% PROTECTED
AMPHIBIAN	72	7	8.86%
BIRD	413	75	15.37%
FISH	115	11	8.73%
MAMMAL	146	30	17.05%
NONVASCULAR PLANT	325	5	1.50%
REPTILE	73	5	6.41%
VASCULAR PLANT	4,216	46	1.08%

PART I: RESULTS – INVESTIGATING PROTECTED SPECIES

▶ CHI SQUARED TEST I
  ▶ The first contingency created was to compare if Mammals are more likely to be endangered than Bird species
  ▶ The Chi Square test run with this data proved that the difference between the 2 categories were not significant to prove with the answer from the test being 0.6876

CATEGORY	PROTECTED	NOT PROTECTED
MAMMAL	30	146
BIRD	75	413

▶ CHI SQUARED TEST II
  ▶ We then compared if Mammals are more likely to be endangered than Reptile species
  ▶ The results of the Chi Square test here proved that the difference between the two categories were in fact significant with the answer being 0.0384

CATEGORY	PROTECTED	NOT PROTECTED
MAMMAL	30	146
REPTILE	5	73

▶ The results from the second test proves that there are certain types of species more likely to be endangered – i.e. Mammals are more likely to be endangered than Reptiles

PART II: SHEEP PLOTTING + DISEASE REDUCTION

FINDING SHEEP IN THE DATA FIRST

Observations of Sheep per Week

▶ Using the both the Species Info from the previous test and a new Observations dataset where conservationists have recorded their sightings of various species at these national parks over the past week, we are able to focus in on sheep sightings with the following method:

▶ Sort the Species data to include sightings where "sheep" is present in the common names

▶ Because some vascular plants have species with "sheep" in its common name, filter the data once more to include "sheep" in its common name and category set to equal "Mammal" in order to get the sheep we are looking for in this data

▶ Merge the data found filtered from the Species data with Observations by connecting it by scientific name to see the number of observations by scientific name and park name

▶ We then grouped the data to see the number of sheet sightings by National Parks as shown in the graph to the left

PART II: RESULTS – SHEEP DISEASE REDUCTION

FOOT + MOUTH DISEASE IN SHEEP

▶ At Bryce National Park, the scientists found that 15% of all sheep there have foot and mouth disease

▶ Yellowstone National Park has an ongoing program to reduce this disease rate and want to know whether this program has been working and to detect reductions of at least 5 percentage points

SAMPLE SIZE DETERMINATION

▶ With the 15% of sheep having the disease at Bryce, we know that the minimum detectable effect is 33% - the number to observe percent change with confidence

▶ The sample size needed to observe change with confidence is 890 based on the data below:

Baseline = 15 | Statistical Significance = 90% | Minimum detectable effect = 33

▶ By dividing 520 by the number of sightings shown at Bryce and Yellowstone National Parks, it will take the following number of weeks below to observe enough sheep to determine the program's efforts in reducing foot and mouth disease in the sheep population:

NATIONAL PARK	SAMPLE SIZE / OBSERVED SIZE	TOTAL NUMBER OF WEEKS
BRYCE NATIONAL PARK	520 / 250	2.08 WEEKS
YELLOWSTONE NATIONAL PARK	520 / 507	1.03 WEEKS

THANK YOU CODECADEMY!