

BIODIVERSITY IN NATIONAL PARKS

CAPSTONE PROJECT
CODECADEMY SKILL PATH:
ANALYZE DATA WITH PYTHON



PROJECT DESCRIPTION

- THIS PROJECT EXPLORES NATIONAL PARKS SERVICE DATA TO GAIN INSIGHTS INTO SPECIES AT ITS NATIONAL PARKS.
- THE DATA USED WAS INSPIRED BY REAL DATA BUT IS MOSTLY FICTIONAL.

DATASET SUMMARY

- DATAFILE: "SPECIES_INFO.CSV"
 - 5,541 DIFFERENT SPECIES IN THE DATASET.
 - COLUMNS: CATEGORY, SCIENTIFIC NAME, COMMON NAMES, CONSERVATION STATUS
 - CATEGORIES:
 - MAMMAL, BIRD, REPTILE, AMPHIBIAN, FISH, VASCULAR PLANT, NONVASCULAR PLANT
 - CONSERVATION STATUSES:
 - SPECIES OF CONCERN, ENDANGERED, THREATENED, IN RECOVERY

DATASET SUMMARY

- DATAFILE: “OBSERVATIONS.CSV”
 - DATA PROVIDED BY CONSERVATIONISTS WHO RECORDED SIGHTINGS OF DIFFERENT SHEEP SPECIES AT SEVERAL NATIONAL PARKS FOR THE PAST 7 DAYS.
 - COLUMNS: SCIENTIFIC NAME, PARK NAME, OBSERVATIONS
 - PARK NAMES:
 - GREAT SMOKY MOUNTAINS NATIONAL PARK, BRYCE NATIONAL PARK, YOSEMITE NATIONAL PARK, YELLOWSTONE NATIONAL PARK

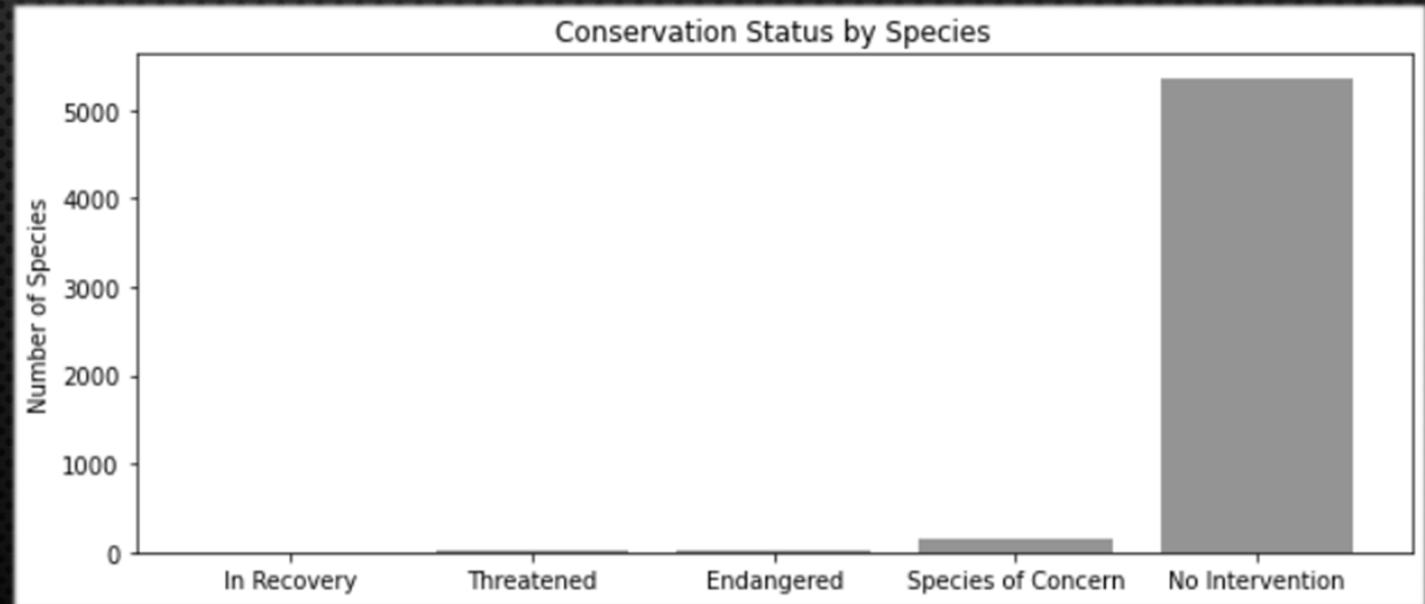
SPECIES ANALYSIS

HOW MANY SPECIES MEET EACH OF THE CONSERVATION STATUS CRITERIA?

RESULTS:

- LARGE MAJORITY REQUIRE NO INTERVENTION.
- 4 SPECIES IN RECOVERY.

conservation_status	scientific_name	
1	In Recovery	4
4	Threatened	10
0	Endangered	15
3	Species of Concern	151
2	No Intervention	5363

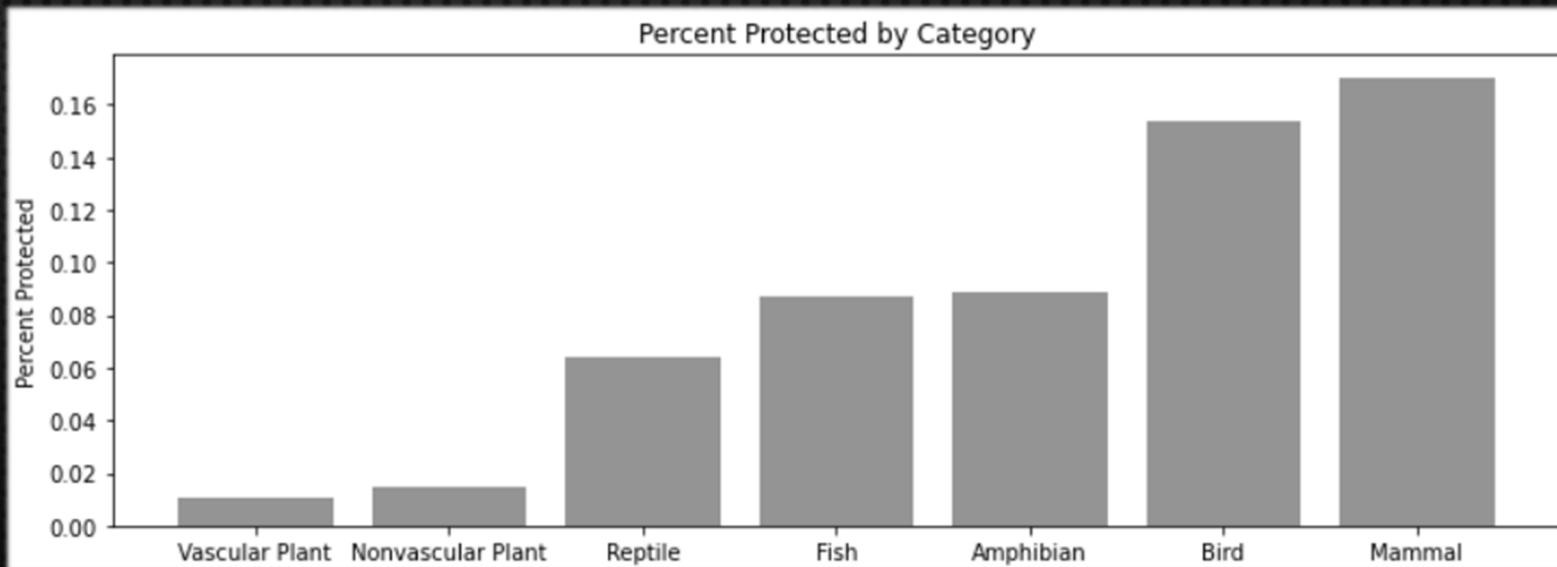


ARE CERTAIN TYPES OF SPECIES MORE LIKELY TO BE ENDANGERED?

- MAMMAL SPECIES ARE MOST ENDANGERED AT 17%.
- BIRD SPECIES ARE SECOND MOST ENDANGERED AT 15.4%.
- 6.4% OF REPTILE SPECIES ARE ENDANGERED.

SPECIES ANALYSIS

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



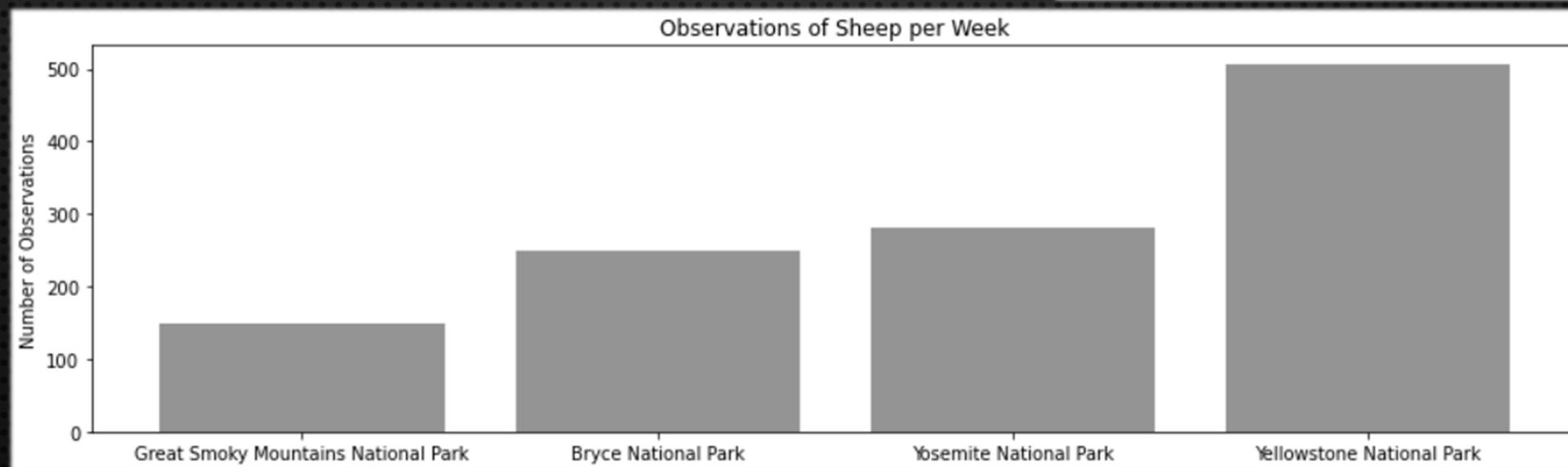
HOW SIGNIFICANT?

- SIGNIFICANCE TEST: CHI SQUARED TEST
 - WHY:
 - CATEGORICAL DATA
 - NEED TO UNDERSTAND ASSOCIATION BETWEEN PAIRS OF CATEGORICAL VARIABLES.
 - P-VALUE = 0.688 → MAMMALS ARE NOT SIGNIFICANTLY MORE ENDANGERED THAN BIRDS AT THE NATIONAL PARKS.
 - P-VALUE = 0.038 → MAMMALS ARE SIGNIFICANTLY MORE ENDANGERED THAN REPTILES AT THE NATIONAL PARKS.

OBSERVATIONS ANALYSIS

HOW MANY TOTAL SHEEP OBSERVATIONS
(ACROSS ALL THREE SPECIES) WERE MADE AT
EACH NATIONAL PARK?

	park_name	observations
1	Great Smoky Mountains National Park	149
0	Bryce National Park	250
3	Yosemite National Park	282
2	Yellowstone National Park	507



FOOT AND MOUTH DISEASE IMPROVEMENT STUDY

- KNOWN: 15% OF SHEEP AT BRYCE NATIONAL PARK HAVE FOOT AND MOUTH DISEASE.
- PARK RANGERS AT YELLOWSTONE HAVE BEEN TRYING TO REDUCE THE DISEASE RATE AND WANT TO KNOW IF REDUCTION EFFORTS HAVE BEEN EFFECTIVE.
- HOW MANY SHEEP MUST BE OBSERVED AT EACH PARK? (SAMPLE SIZE CALCULATION)
- HOW LONG WILL OBSERVATIONS TAKE? (STUDY DURATION)

SAMPLE SIZE CALCULATION, STUDY DURATION

- HOW MANY SHEEP MUST BE OBSERVED AT EACH PARK?
 - KNOWN (BASELINE): 15% WITH FOOT AND MOUTH DISEASE
 - MINIMUM DETECTABLE EFFECT: 33.3%
 - LEVEL OF SIGNIFICANCE: 90%
 - SAMPLE SIZE: 870
- HOW LONG WILL OBSERVATIONS TAKE AT BRYCE AND YELLOWSTONE NATIONAL PARKS?
 - 3.5 WEEKS TO COLLECT DATA AT BRYCE NATIONAL PARK
 - 1.5 WEEKS TO COLLECT DATA AT YELLOWSTONE NATIONAL PARK