Biodiversity Project analysis

Mor Rubinstein

The dataset

The dataset has 5541 unique records. I has 4 columns that has string values:

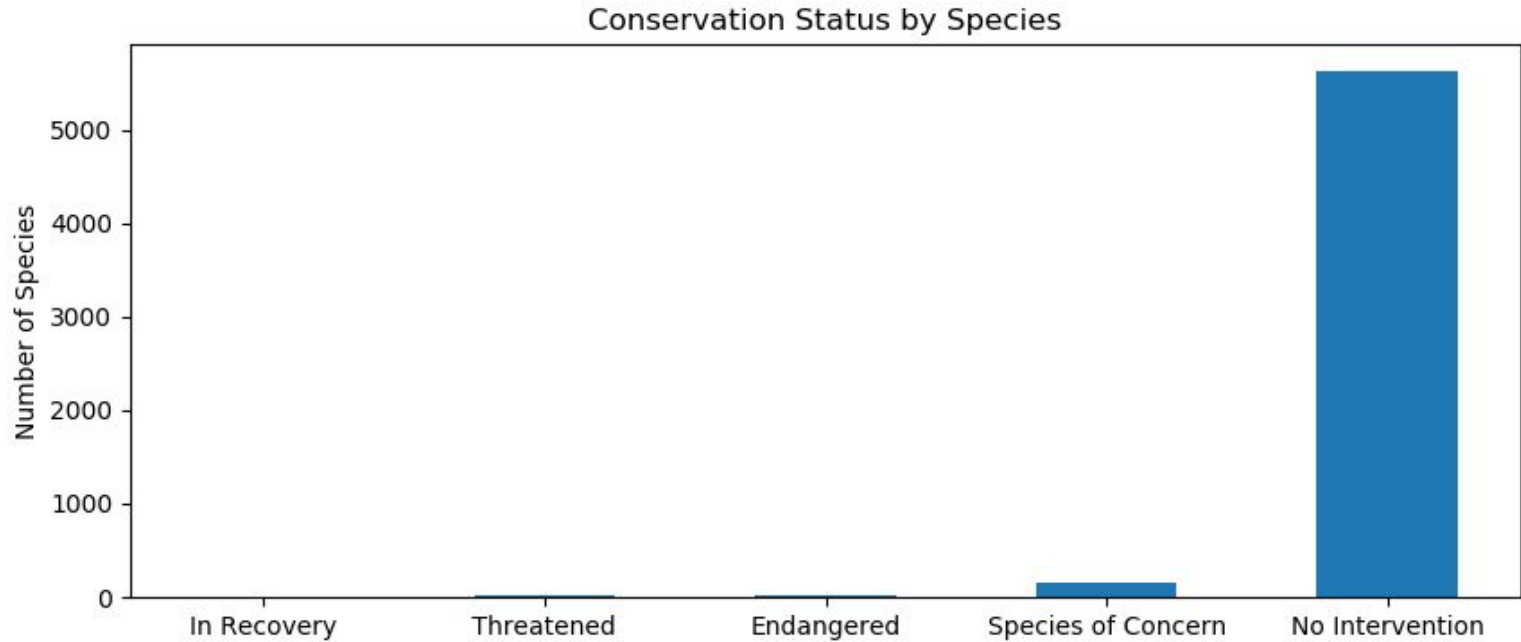Category - with the 7 values 'Mammal' 'Bird' 'Reptile' 'Amphibian' 'Fish' 'Vascular Plant'
 'Nonvascular Plant'
Scientific_name - free text
 common_names - free text
 conservation_status - with the values nan, 'Species of Concern' , 'Endangered' ',Threatened' ,'In Recovery'

Conservation Status

Conservation Status by Species

Conservation Status Analysis

We can see that most species in the park do not need intervention - 96% of species actually!

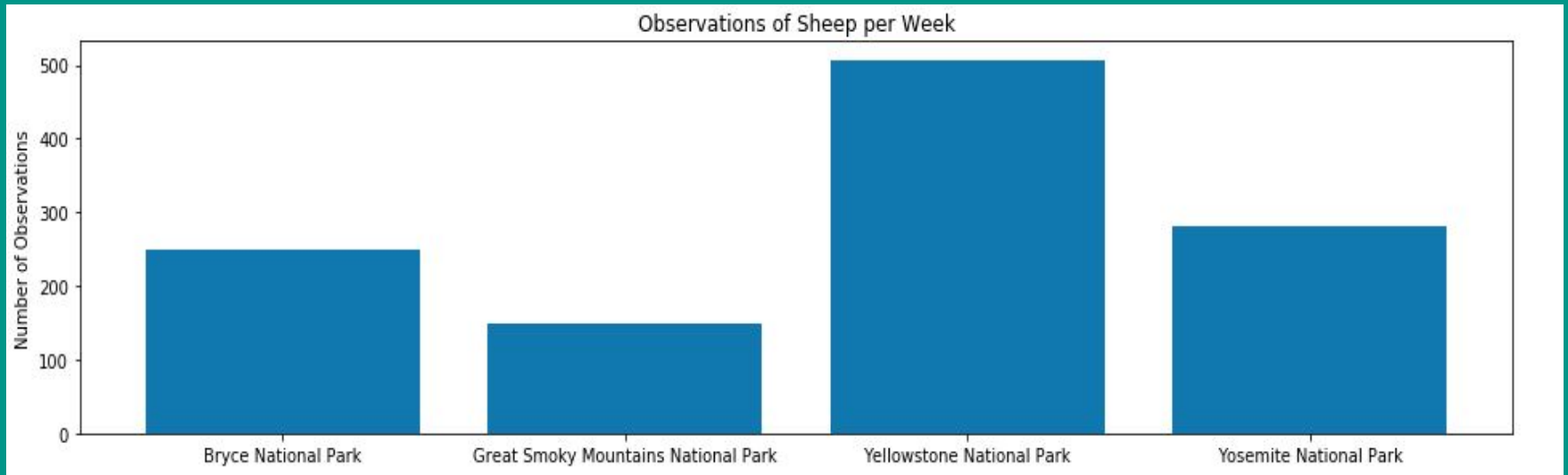Only 0.001 species are threatened, very rare, but we need to do all in our power to conserve them!

Protected species

Birds are the most protected with 75 species, 2.5 times more species than the second most protected group, mammals with 30 species.

However, in percentage, 17% of mammals are protected, compared to 15% of birds. However, according to our hypothesis testing, mammals are not inclined to be more protected than birds, since the p values was bigger than 0.05.

Nonvascular plants are the least protected with 1.5% protected.

Sheep

Sheep observations

Yellowstone park has the most sheep observations, almost twice than Yosemite and more than twice Bryce park