Biodiversity for the National Parks

By Adam Kinnane

CONTENTS

1. Data Source: Species_info.csv.
2. Calculations of the status of various endangered species.
3. Recommendations for conservationists concerned about endangered species.
4. Sample size determination of foot and mouth disease study
5. Graphs

1. Data Source: Species_info.csv.

The data containd in 'Species_info.csv' included the following information:
- Category (Categories of species)
  - There are 7 categories in the dataset.
    - ♣ 'Mammal', 'Bird', 'Reptile', 'Amphibian', 'Fish', 'Vascular Plant', and 'Nonvascular Plant'
- Scientific names (Scientific names of species)
  - There are 5,541 unique scientific names in the dataset.
- Common names (Common names of species)
  - There are various common names for the species. Some examples are below:
    - ♣ Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral), Wapiti Or Elk
- Conservation Status (Conservation Status of species)
  - Conservation status is broken down into 'nan' 'Species of Concern' 'Endangered' 'Threatened' 'In Recovery'
  - The below outlines the amount of species by scientific name are categorized by Conservation Status.
    - ♣ Endangered = 15
    - ♣ In Recovery = 4
    - ♣ Species of Concern = 151
    - ♣ Threatened = 10
    - ♣ No intervention (Nan) = 5,363

2. Calculations of the status of various endangered species.

Conservation Status	Species (Unique Scientific Name)
Endangered	15
In Recovery	4
Species of Concern	151
Threatened	10
No intervention	5,363

This table represents the Conservation Status types outlined in section 1 and the number of species that fall under each category. These species are group with each Conservation Status by unique scientific name. For example there are 15 unique endangered species in the dataset.

3. Recommendations for conservationists concerned about endangered species.

Category	Not Protected	Protected	Percent Protected *
Amphibian	72	7	0.09%
Bird	413	75	0.15%
Fish	115	11	0.09%
Mammal	146	30	0.17%
Nonvascular Plant	328	5	0.02%

The table above outlines the amount of species per category that are protected, or endangered. From the above you can see that the categories with the highest percent of protected species are as follows.

- Mammals with 0.17% of their species being protected, and
- Birds with 0.15% of their species being protected.

By observing this data we investigated if Mammals are more likely to be endangered than Birds or if the data was similar due to chance.

3. Recommendations for conservationists concerned about endangered species - Continued.

After running a Chi-Squared significance test to see if there were any significant differences between the Mammal and Birds data. Our results (see below) confirmed that there is no significant difference between the 2 data sets. Our recommendation to conservationists is to treat these 2 types of species equally for now.

- Null hypothesis is that the difference in Mammal & Bird is due to chance
    - If p-value of Chi-Squared significance test is < 0.05 then reject Null Hypothesis
- Endangered likeliness: p-valus of Mammal vs Bird 0.687594809666

However, it should be noted that we also ran a Chi-Squared significance test against Mammals and Reptiles (see below) and the data confirmed that reptiles are more likely to become endangered than Mammals. Conservationists should apply this data to their efforts.

- Null hypothesis is that the difference in Mammal & Reptile is due to chance
    - If p-value of Chi-Squared significance test is < 0.05 then reject Null Hypothesis
- Endangered likeliness: p-value of Mammal vs Reptile 0.0383555902297

4. Sample size determination of foot and mouth disease study.

The sample size determination study used the following inputs to determine how many sheep would need to be observed to acuratly determine if Foot & Mouth disease was being reduced in sheep in Yellowstone National Park.

- Baseline = 15%
- Statistical Significance = 90%
- Minimum Detectable effect = 33%

The baseline was set at 15% as this was the only observed data available for the study. This data was taken from a study recorded at Bryce National Park.

The Statistical Significance was set at 90% as a default.

The Minimum Detectable Effect was set at 33% as Yellow Stone National Park wanted to see a minimum of a 5 percent points decrease. To find this value we used the following formulae:

- Minimum Detectable Effect = 100 * (X% Change / Basline)
    - 33% = 100 * (5 / 15)

# 4. Sample size determination of foot and mouth disease study. - Continued.

With this information we were able to learn the following results by plugging the above numbers into a Sample Size Calculator.

- Sample Size = 890

When analysing the observed sheep at the national parks against the sample size required to identify any difference in Foot & Mouth disease we discovered the following.
- Yellowstone National Park would require 1.76 weeks to observe a difference.
- Bryce National Park would require 3.56 weeks to observe a difference.

5. Biodiversity for the National Parks Graphs.
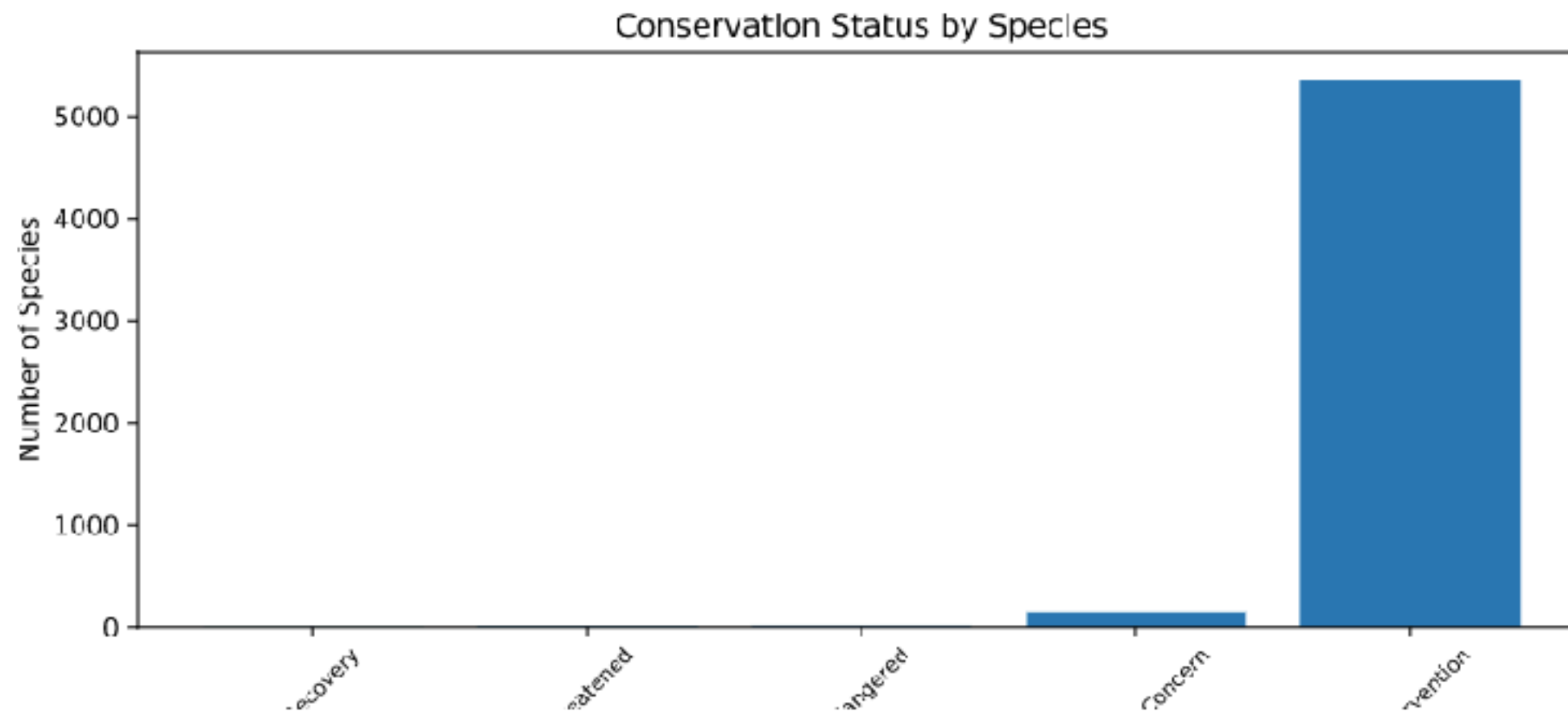
This table represents the number of species that fall into each level of Conservation Status. As you can see the vast majority fall into the last segment of 'No Concern' followed by 'Concern'. For more information see slide 4.

5. Biodiversity for the National Parks Graphs - Continued.
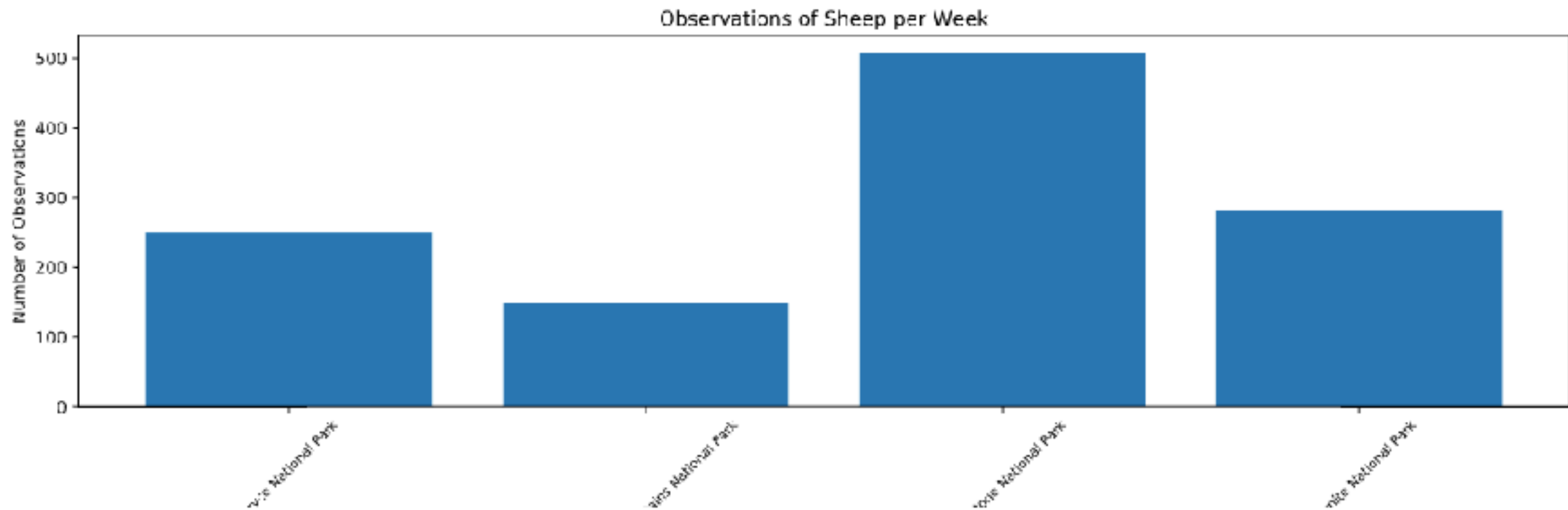
This table represents the number of sheep observed in the various national parks. Yellow Stone National Park observed the most sheep followed by Bryce National Park. For more information see slides 7 and 8.