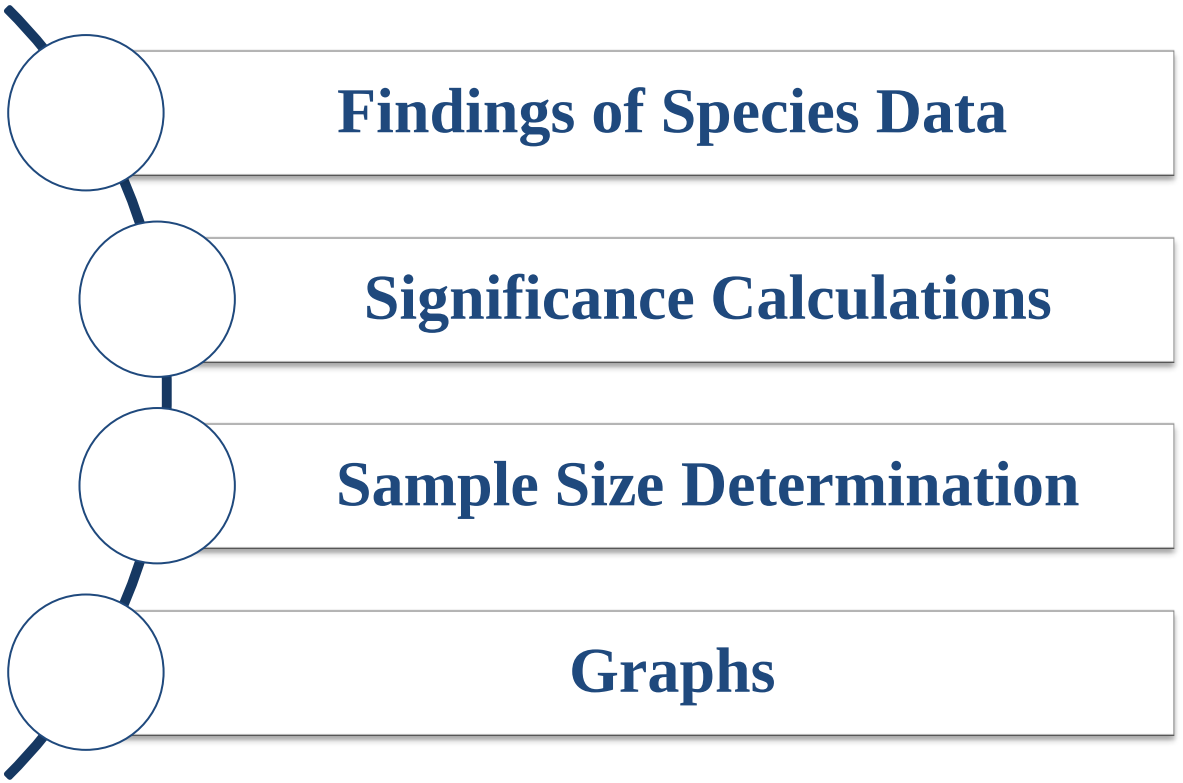
Biodiversity Analysis for the National Parks

- Fan Ye

Capstone Project of Intro to Data Analysis

Content


Findings of Species Data
Significance Calculations
Sample Size Determination
Graphs

Findings of Species Data

- A total of 5824 species are included in the dataset
- Species are divided into seven categories, Mammal, Bird, Reptile, Amphibian, Fish, Vascular Plant, and Nonvascular Plant
- Their conservation status include 'Species of Concern‘, 'Endangered‘, 'Threatened' and 'In Recovery’.

Findings of Species Data

Conservation Status	Number of Species
In Recovery	4
Threatened	10
Endangered	16
Species of Concern	161
No Intervention	5633
Total	5824

5633 Species are in the status of ‘No Intervetion’, leaving 191 species (3.3%) in need of protection to a certain degree.

Significance Calculations for Endangered Status of Different Species

Category	Not Protected	Protected	Percent Protected(%)
Amphbian	73	7	8.8
Bird	**442**	**79**	**15.2**
Fish	116	11	8.7
Mammal	**176**	**38**	**17.8**
Nonvascular Plant	328	5	1.5
Reptile	74	5	6.3
Vascular Plant	4424	46	1.0

p value of chi square test between categories of bird and mammal is 0.446 (>0.05), which means the difference is not significant.

Significance Calculations for Endangered Status of Different Species

Category	Not Protected	Protected	Percent Protected(%)
Amphbian	73	7	8.8
Bird	442	79	15.2
Fish	116	11	8.7
Mammal	**176**	**38**	**17.8**
Nonvascular Plant	328	5	1.5
Reptile	**74**	**5**	**6.3**
Vascular Plant	4424	46	1.0

p value of chi square test between categories of reptile and mammal is 0.023 (<0.05), which indicates a significant difference between their endangered status.

Recommendations for Conservationists

Based on the significance calculations, the percentage of birds and mammals that need protection are significantly higher than the other categories. Conservationists should pay special attention to these two categories and give them more specific care.

Sample Size Determination

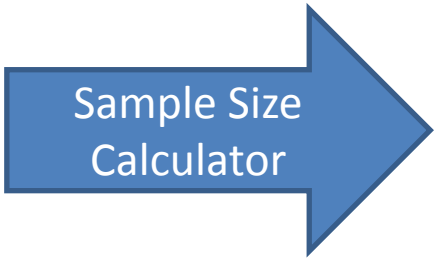
Baseline conversion rate: 15%

15% of sheep at Bryce National Park have foot and mouth disease

Minimum effect of detection: 5% / 15% = 33%

should be able to detect at least 5 percent points of reduction

Level of significance: 90% by default



Sample Size: 890

Sample Size Determination

Number of sheep observed per week in **Bryce National Park**: 250

Weeks needed to observe enough sheep in Bryce National Park

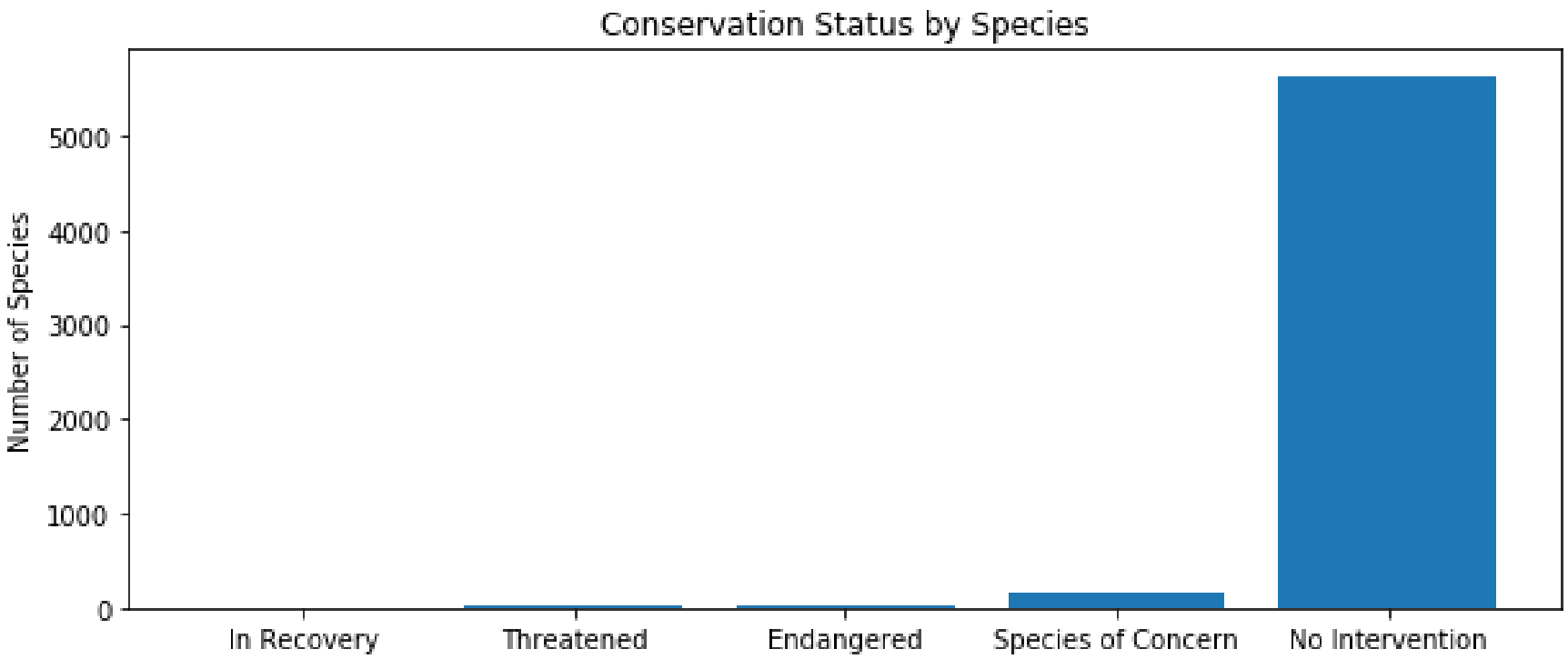
= math. ceil(890 / 250) = 4

Number of sheep observed per week in **Yellowstone National Park**: 507

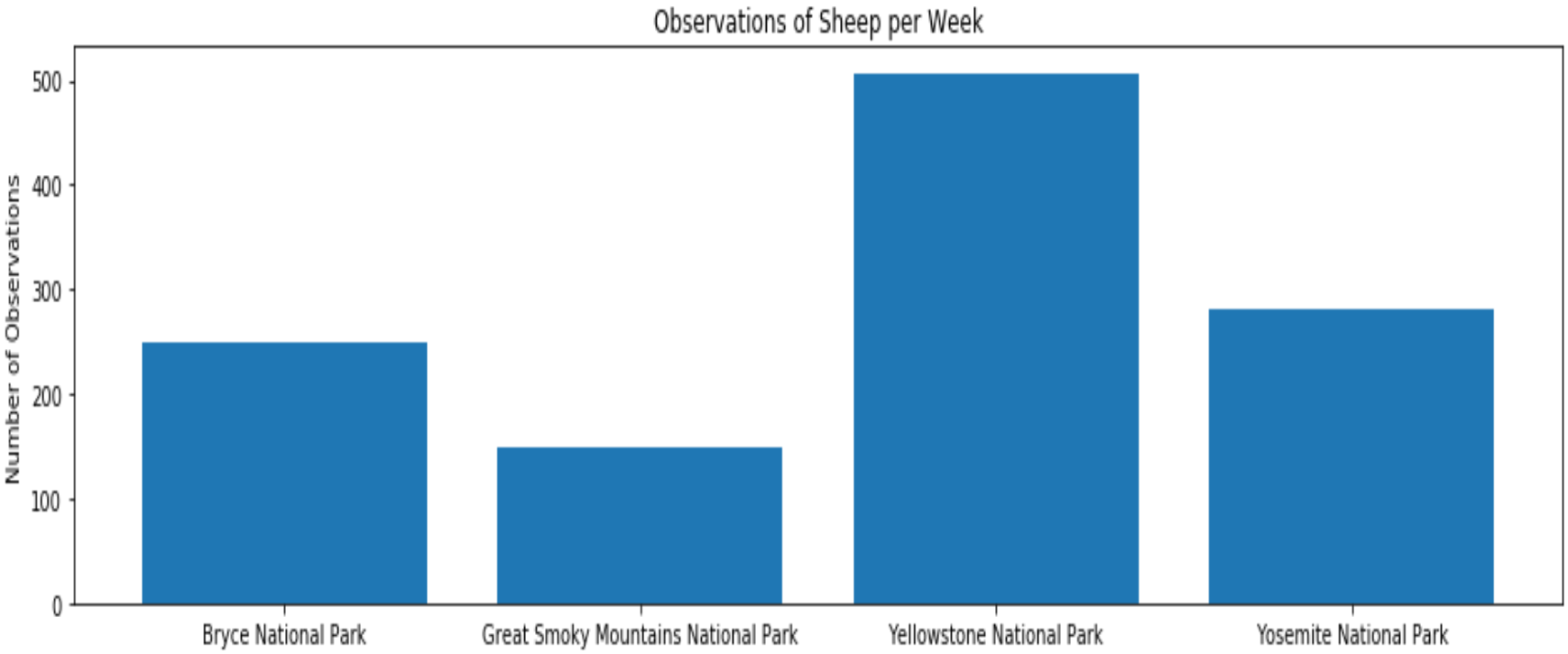
Weeks needed to observe enough sheep in Yellowstone National Park

= math. ceil(890 / 507) = 2

Graphs


Conservation Status by Species
Number of Species
5000
4000
3000
2000
1000
0
In Recovery
Threatened
Endangered
Species of Concern
No Intervention

Graphs


Observations of Sheep per Week
Number of Observations
0
100
200
300
400
500
Bryce National Park
Great Smoky Mountains National Park
Yellowstone National Park
Yosemite National Park