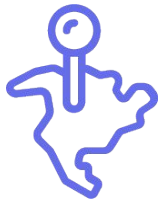
Biodiversity for the National Parks Data Story

Alejandro Rodriguez Perez

INTRO

What is this data story about?



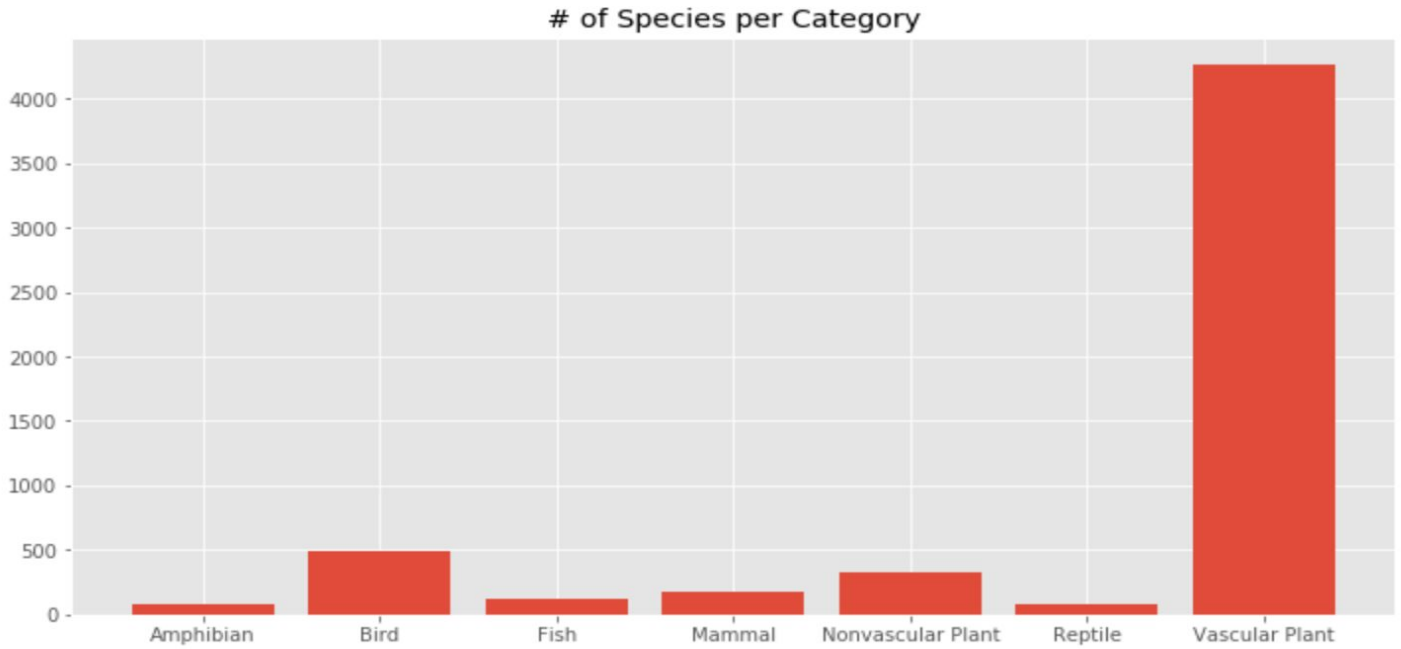
Overview of endangered species across 5 North American National Parks.

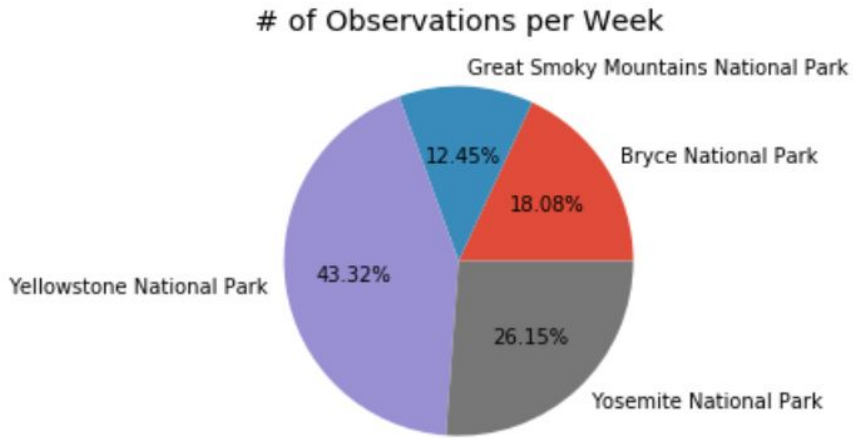
Species = ['Mammal', 'Bird', 'Reptile', 'Amphibian', 'Fish', 'Vascular Plant', 'Nonvascular Plant]

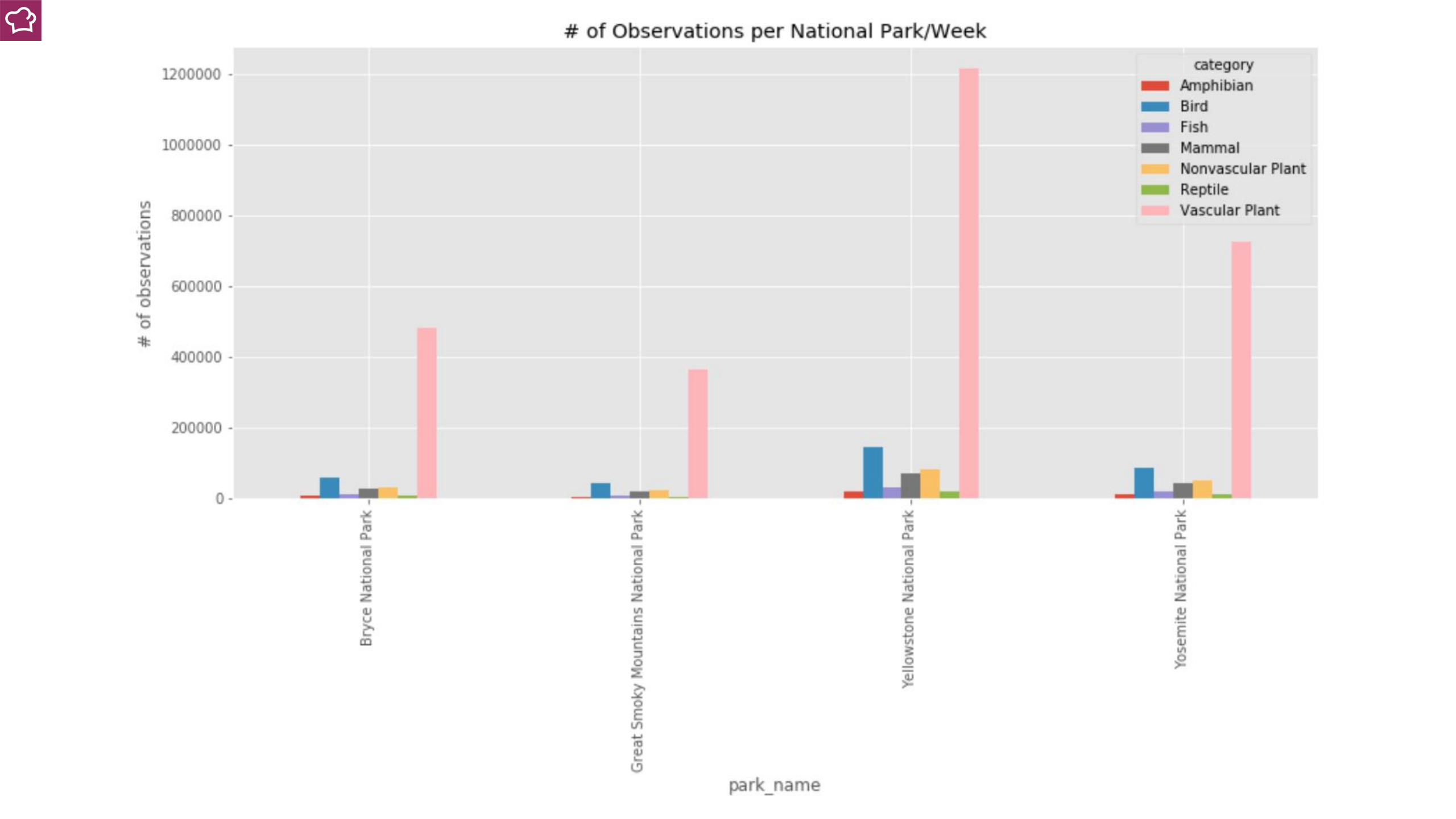
National Parks = ['Great Smoky Mountains National Park', 'Yosemite National Park', 'Bryce National Park', 'Yellowstone National Park']

Highlights

- **Category** Ranking: #1 Vascular Plants, #2 Birds, #3 Nonvascular Plants
- **Park** Ranking: #1 Yellowstone, #2 Yosemite , #3 Brice



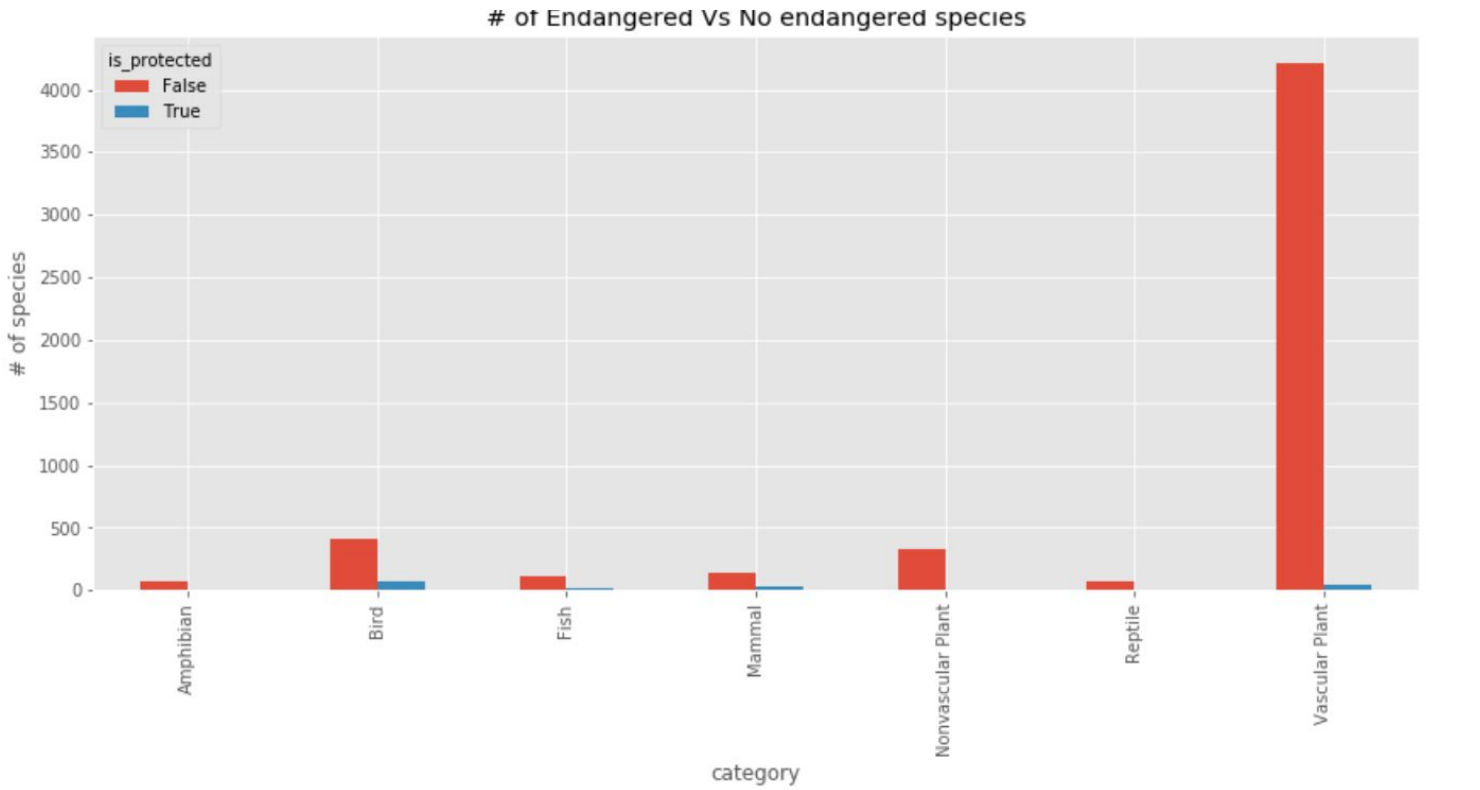



# of Observations per National Park/Week
# of observations
1200000
1000000
800000
600000
400000
200000
0
category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Bryce National Park
Great Smoky Mountains National Park
Yellowstone National Park
Yosemite National Park
park_name

Protected Species Per Animal Category

- There is a total of 179 protected species (3.23%)

- The largest number of Protected species belongs to the Bird Category, followed by Vascular plants and with mammals in the third position



is_protected category	not_protected	protected	percent_protected
Amphibian	72	7	8.86%
Bird	413	75	15.37%
Fish	115	11	8.73%
Mammal	146	30	17.05%
Nonvascular Plant	328	5	1.50%
Reptile	73	5	6.41%
Vascular Plant	4216	46	1.08%

Significance Analysis

Is there a significant difference in between the # of endangered species among Mammals, Reptiles & Birds?

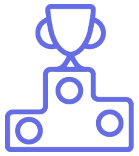
Mammals VS Birds

```
hi2_stat, pvalue, dof, t = chi2_contingency(contingency)
print('There is no significant difference as the pvalue is > 0.05: {:.5f}'.format(pvalue))
```

```
There is no significant difference as the pvalue is > 0.05: 0.68759
```

There is no difference

Mammals VS Reptiles



```
contingency = [[73, 5],
               [146, 30]]

hi2_stat, pvalue, dof, t = chi2_contingency(contingency)
print('There is significant difference as the pvalue is < 0.05: {:.5f}'.format(pvalue))
```

```
There is significant difference as the pvalue is < 0.05: 0.03836
```

There is a significant difference, a larger # of mammals species are endangered :(

Reptiles VS Birds

```
contingency = [[73, 5],
               [413, 75]]

hi2_stat, pvalue, dof, t = chi2_contingency(contingency)
print('There is no significant difference as the pvalue is > 0.05: {:.5f}'.format(pvalue))
```

```
There is no significant difference as the pvalue is > 0.05: 0.05314
```

There is no difference

Sampling

- Scientist will like to reduce foot and mouth disease by 5%
- **Baseline** = 15%
- **Confidence Level** = 90%
- **Minimum Detectable Effect** = 33.3%