Biodiversity and Conservation Analysis in National Parks

Biodiversity and Conservation Status Analysis

A data-driven study of species protection and conservation significance across U.S. National Parks.

Description of the Species Data (species_info.csv)

The species_info.csv file contains comprehensive information about various species found in U.S. National Parks and their conservation status, forming the bedrock of our analysis.

Dataset Overview

- **Total rows:** 5,824
- **Unique species:** 5,541
- **Main columns used:**

- scientific_name
- common_names
- category (type of species)
- conservation_status

Species Categories

The dataset encompasses a broad range of life, categorized as:

- Vascular Plants
- Birds
- Mammals
- Fish
- Reptiles
- Amphibians
- Nonvascular Plants

Key observation: Vascular plants constitute the largest portion, while mammals and birds represent a much smaller share.

Conservation Status Breakdown

Species are classified into four critical conservation groups:

- Not Protected
- Species of Concern
- Threatened
- Endangered

*Note: Missing conservation values were uniformly treated as **"Not Protected"** for consistency.*

Protected Species

191 species are under protection.

Not Protected

5,633 species lack formal protection.

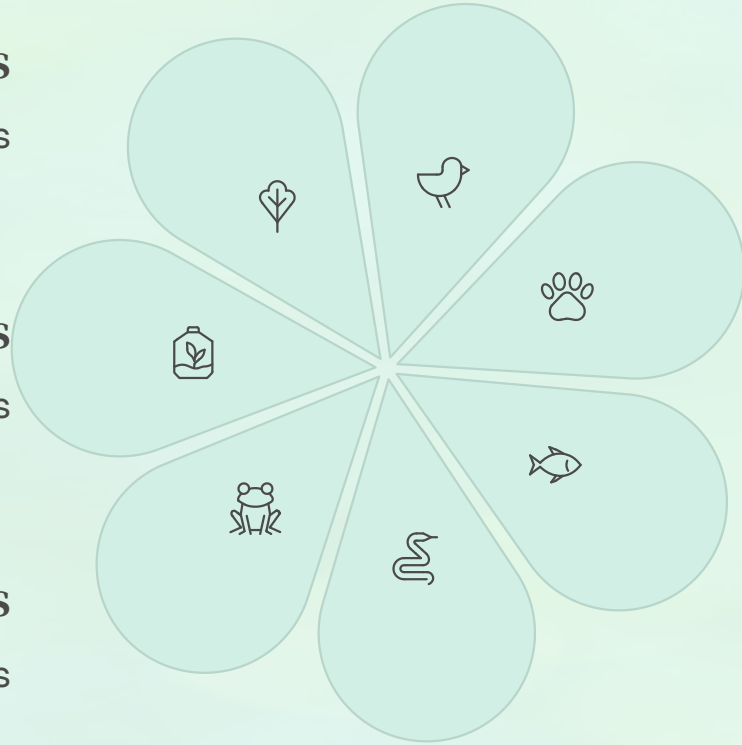
Important Finding: Over 96% of species are not currently under conservation protection, highlighting a significant gap in current efforts.

Species Categories: A Snapshot of Park Biodiversity

Vascular Plants
4,470 Species

Nonvascular Plants
135 Species

Amphibians
80 Species



Birds
521 Species

Mammals
214 Species

Fish
127 Species

Reptiles
79 Species

Key takeaway: Vascular plants overwhelmingly dominate the dataset, highlighting their foundational role in these ecosystems. While mammals and birds represent a smaller proportion, their ecological importance often warrants specific conservation focus.

Significance Calculations for Endangered Status

To assess whether certain species categories face a disproportionately higher risk of endangerment, statistical significance tests were performed.

Categories Compared

- Mammals vs. Birds
- Mammals vs. Reptiles

Method Used

A **Chi-Square test** was employed to compare the proportion of protected versus non-protected species between categories. This method helped determine if observed differences were statistically meaningful or merely due to chance.

Results

Mammals vs. Birds

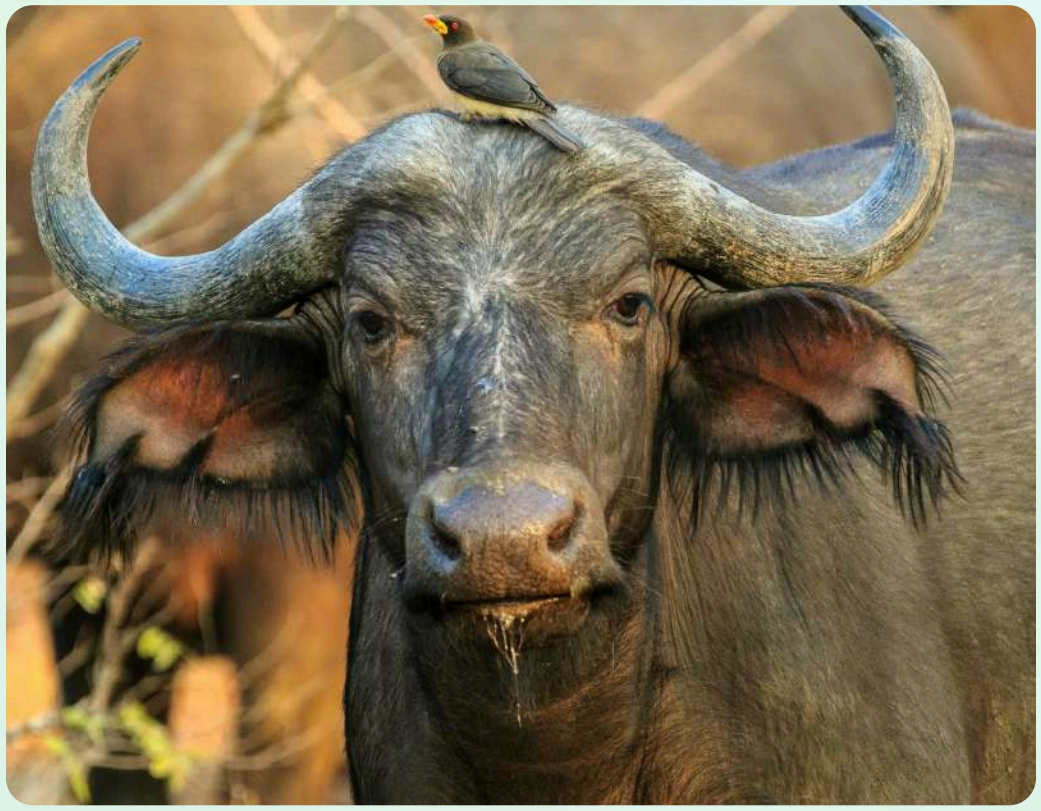
No statistically significant difference was found, indicating similar protection rates for both categories.

Mammals vs. Reptiles

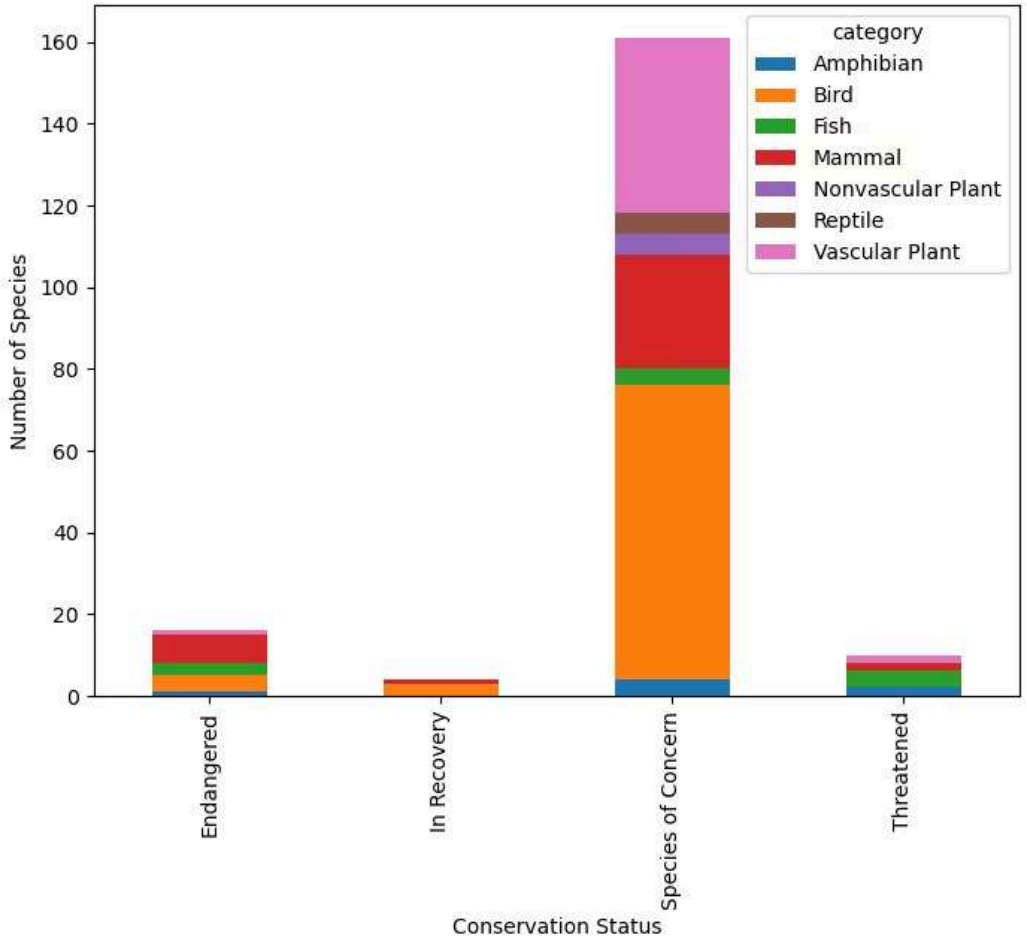
A statistically significant difference was observed, with mammals demonstrating a higher likelihood of being under protection than reptiles.

Interpretation

This disparity suggests that conservation efforts are not evenly distributed across different species categories, pointing to potential biases in current protection strategies.



Conservation Status Distribution: A Stark Reality

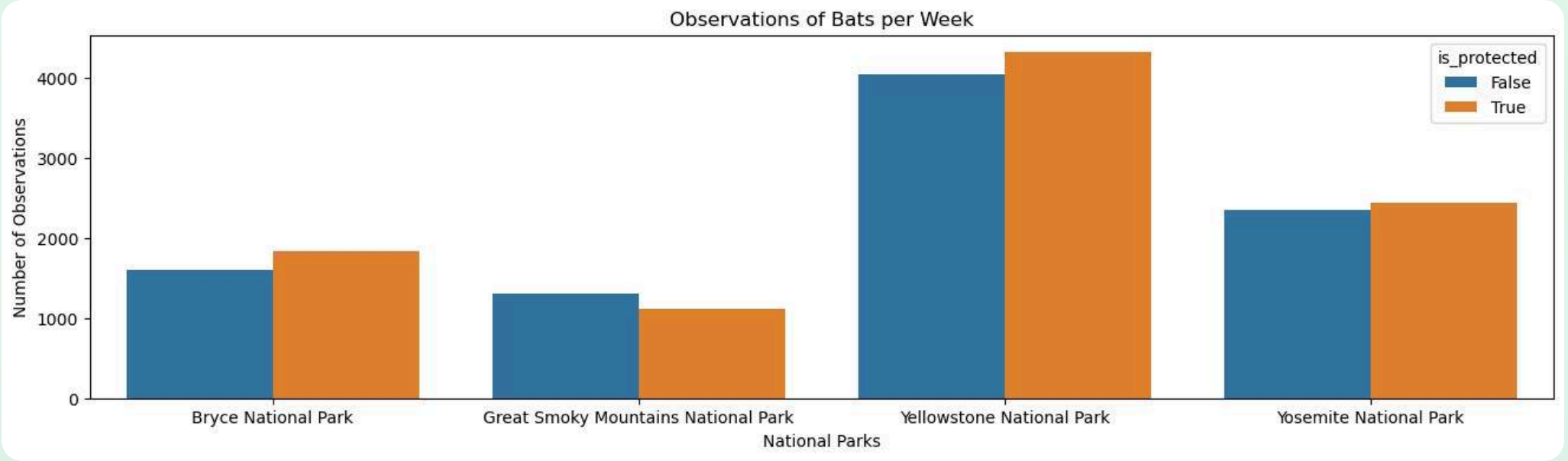


Species are categorized into four conservation statuses, revealing a critical imbalance:

- **Not Protected (NaN):** 5,633
- **Species of Concern:** 161
- **Endangered:** 16
- **Threatened:** 10

Key insight: A staggering 96% of species in our dataset are not currently under active conservation status. Only a small fraction (191 species total) receive formal protection, underscoring the vast scope of biodiversity that remains unmonitored or unprotected.

Species Observations Across Parks: Bat Activity



Most Frequently Observed Species

Our data indicates that bats are the most commonly observed species group across the national parks included in this study. This prevalence warrants further investigation into their ecological health and specific conservation needs.

Distribution of Bat Observations

- **Yellowstone National Park** recorded the highest number of bat sightings, suggesting a robust population or concentrated monitoring efforts.
- Bryce Canyon and Yosemite National Parks showed moderate levels of bat observations.
- Great Smoky Mountains National Park reported comparatively fewer protected bat observations, which could signify varying populations or differing levels of survey intensity.

Important note: Higher observation counts do not automatically equate to healthier populations. These figures can also be influenced by factors such as park size, habitat diversity, and the intensity or methodology of sampling.

Recommendation for Conservationists

Based on the compelling findings from our significance calculations, targeted adjustments to conservation priorities are warranted.

Current Focus

Conservation efforts predominantly prioritize **mammals and birds**, which receive similar levels of attention and protection.

Underrepresented Groups

Reptiles exhibit significantly lower protection rates, despite their crucial ecological roles and increasing threats.

Recommendation

Re-evaluate Strategies

Conservationists must re-evaluate existing protection strategies, particularly for underrepresented species like reptiles.

Embrace Data-Driven Methods

Utilize data-driven approaches to mitigate biases that favor more visible or charismatic species over others.

Prioritize Ecological Risk

Allocate conservation resources based on objective ecological risk assessments, rather than solely on public attention or appeal.

This strategic shift can lead to more balanced, effective, and equitable biodiversity protection across all species.

Sample Size Determination: Foot and Mouth Disease Study

In a separate, yet equally critical analysis, we focused on determining the optimal sample size needed to detect Foot and Mouth Disease (FMD) with robust statistical confidence.

Goal

The primary objective was to estimate the minimum number of animals that need to be observed to reliably detect the presence of FMD within a population, ensuring accuracy and efficiency in disease surveillance.

Key Assumptions

- The baseline disease rate was carefully estimated from prior epidemiological data.
- A desired confidence level of **95%** was established to ensure high reliability.
- The minimum detectable effect (the smallest difference or effect worth detecting) was clearly defined before the analysis began.

Outcome

1

Larger Samples for Rarity

The calculations revealed that a **larger sample size** is imperative when the occurrence of the disease is rare, to avoid missing critical cases.

2

Risk of Small Samples

Conversely, employing small samples significantly **increases the risk of false negatives**, potentially overlooking true disease outbreaks.

Why This Matters

Proper sample size determination is fundamental for:

- Ensuring **reliable conclusions** from surveillance data.
- Reducing the incidence of **false negatives**, which can have severe implications.
- Facilitating **better decision-making** in disease control and prevention strategies, safeguarding animal health and economic stability.

Final Summary: Key Insights for Conservation

Our comprehensive analysis across species data and epidemiological studies yields several critical insights for advancing conservation efforts.

1

Species Protection Gap

The vast majority of species within U.S. National Parks are **not currently under formal conservation protection**.

2

Conservation Focus

Mammals and birds receive comparable levels of conservation attention, highlighting a traditional focus.

3

Reptile Disparity

Mammals are **significantly more protected than reptiles**, indicating an imbalance in current conservation strategies.

4

Data-Driven Conservation

Adopting more data-driven approaches can significantly enhance the effectiveness and fairness of conservation strategies.

5

Sample Size Imperative

Accurate sample size calculations are **critical for reliable disease studies**, minimizing false negatives and ensuring robust public health decisions.