Biodiversity Analysis and Conservation Insights

Prepared by: Timur Rakhimov

Overview of the Data

Species Data (species_info.csv)

Species Data (species_info.csv):

- **Total Records**: 5,824 species entries.
- **Categories of Species**:
    - **Vascular Plants** dominate with 4,470 entries.
    - Other categories include Birds (521 entries), Nonvascular Plants (333), Mammals (214), Fish (127), Amphibians (80), and Reptiles (79).
- **Conservation Status**:
    - 161 species are listed as "Species of Concern."
    - 16 species are "Endangered."
    - 10 species are "Threatened."
    - 4 species are "In Recovery."
    - The majority of species (5,633 entries) do not have a specified conservation status.
- **Unique Species**:
    - There are 5,541 unique scientific names, indicating significant biodiversity.

Observations Data (observations.csv)

- **Number of Entries**: 23,296
- **Observation Statistics**:
    - Average observations per species: ~142
    - Maximum observations: 321
    - Minimum observations: 9
- **Parks**: Data spans multiple national parks, including Great Smoky Mountains, Yosemite, and Bryce.

Significance Calculations

Endangered Status Across Categories

- **Mammals** have the highest count of endangered species, with 7 out of 16 classified as such.
- **Birds** dominate the "Species of Concern" category, with 72 entries.
- **Fish** stand out in the "Threatened" category, accounting for 4 out of 10 species.
- **Vascular Plants** have the most representation overall but relatively fewer in critical conservation statuses.

Statistical Insights

- Mammals and Birds show statistically significant representation in conservation-related statuses.
- The data highlights the urgency of addressing specific categories with high endangered counts, such as Mammals and Birds.

Recommendations for Conservationists

Based on the analysis, the following recommendations are made for conservation efforts:

- **Focus on Mammals**:
    - With the highest count of endangered species, immediate conservation programs should prioritize Mammals.
- **Bird Monitoring and Protection**:
    - Birds have the highest number of "Species of Concern." Monitoring programs and habitat preservation should be intensified.
- **Fish Conservation**:

    - Given their prominence in the "Threatened" category, efforts should focus on water ecosystem health and pollution control.
- **Vascular Plants**:
    - As the dominant category, preserving biodiversity and preventing habitat loss are key strategies for maintaining ecosystem balance.

Graphical Analysis

Included Visualizations

1. **Distribution of Species Categories**:
    - Bar chart showing the dominance of vascular plants and birds in the dataset.
2. **Conservation Status by Category**:
    - A stacked bar chart detailing the distribution of conservation statuses across species categories.
3. **Top 10 Most Observed Species**:
    - Highlighting the most frequently observed species across parks.
4. **Significance of Endangered Species**:
    - Heatmap visualizing the differences in endangered status across categories.
5. **Observation Trends by Park**:
    - Bar chart showing species observations broken down by park location.
6. **Sample Size Justification**:
    - Visualization supporting sample size recommendations for detecting disease presence.

Conclusion

This analysis provides a comprehensive understanding of biodiversity trends and conservation priorities. The data-driven insights can guide efforts to protect vulnerable species and maintain ecological balance. Conservationists are encouraged to leverage these findings for targeted interventions and policy planning.