

# A Study of Biodiversity in Four National Parks

...

Vince Gregoric  
2/16/23

# Project Goals

- Explore the variety of species found in national parks in the US
- Determine what factors affect how often a species is observed in the wild
- Evaluate the accuracy of conservation status assignments
- Practice data cleanup, analysis, and visualization

# Overview of Data

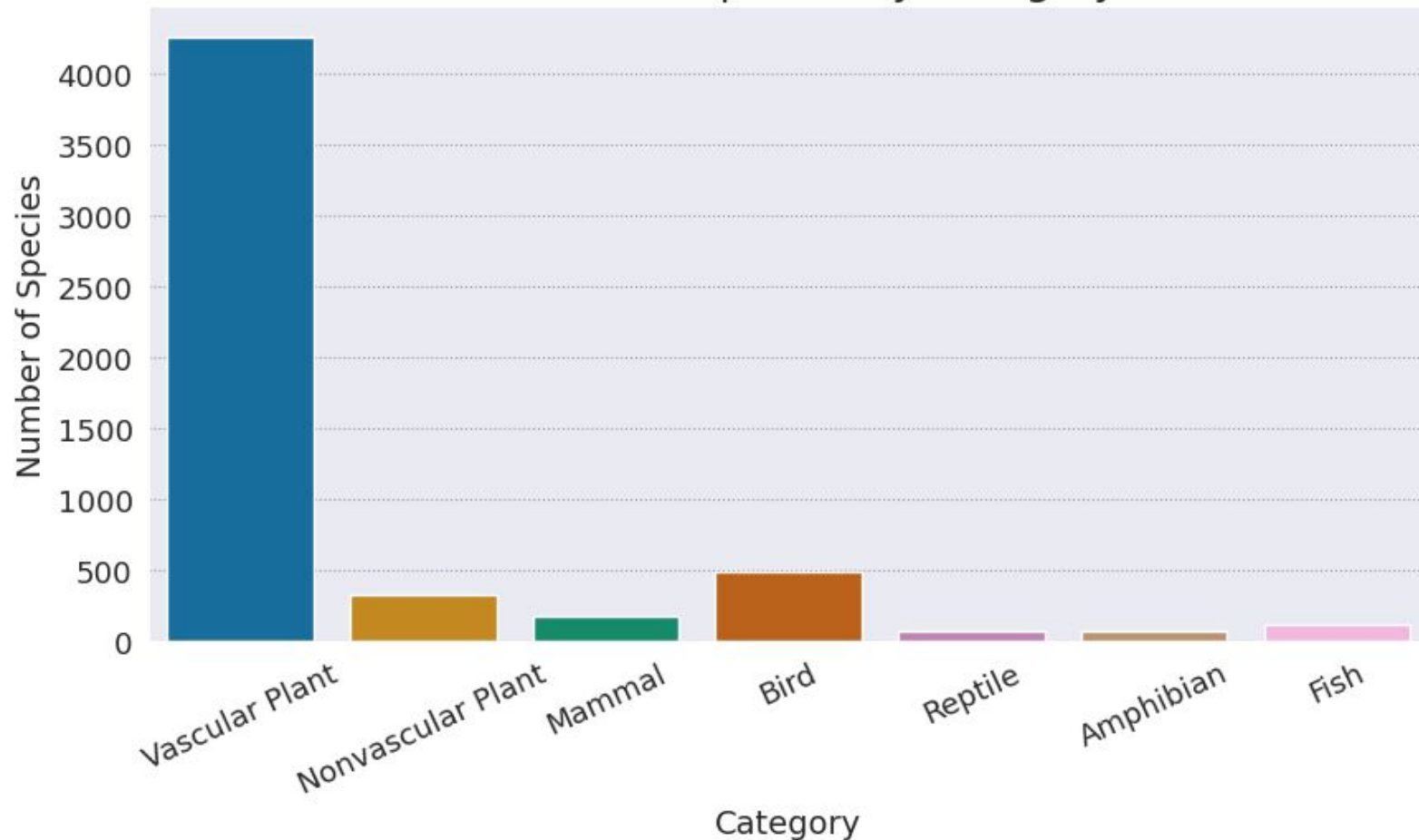
- Obtained from Codecademy.com
- 5541 species and 4 parks represented
- For each species:
  - Scientific name
  - Common names
  - Species category
  - Conservation status
  - Number of observations at each park over 7 days

# Breakdown by Category

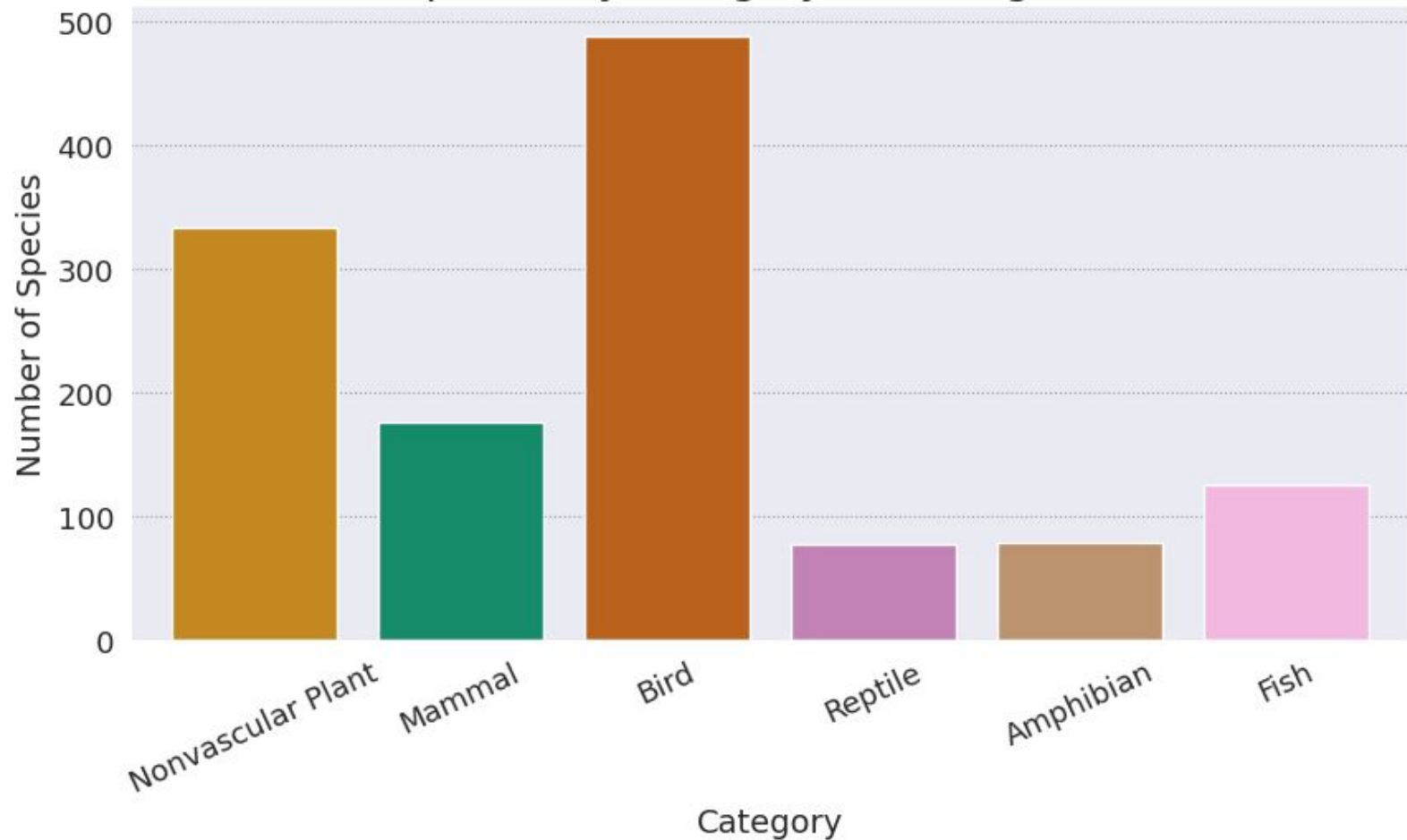
- Most species are vascular plants:

Category	Number of Species	Percent
Vascular Plant	4262	76.92
Nonvascular Plant	333	6.01
Mammal	176	3.18
Bird	488	8.81
Reptile	78	1.41
Amphibian	79	1.43
Fish	125	2.26

# Number of Species by Category



Number of Species by Category, Omitting "Vascular Plant"

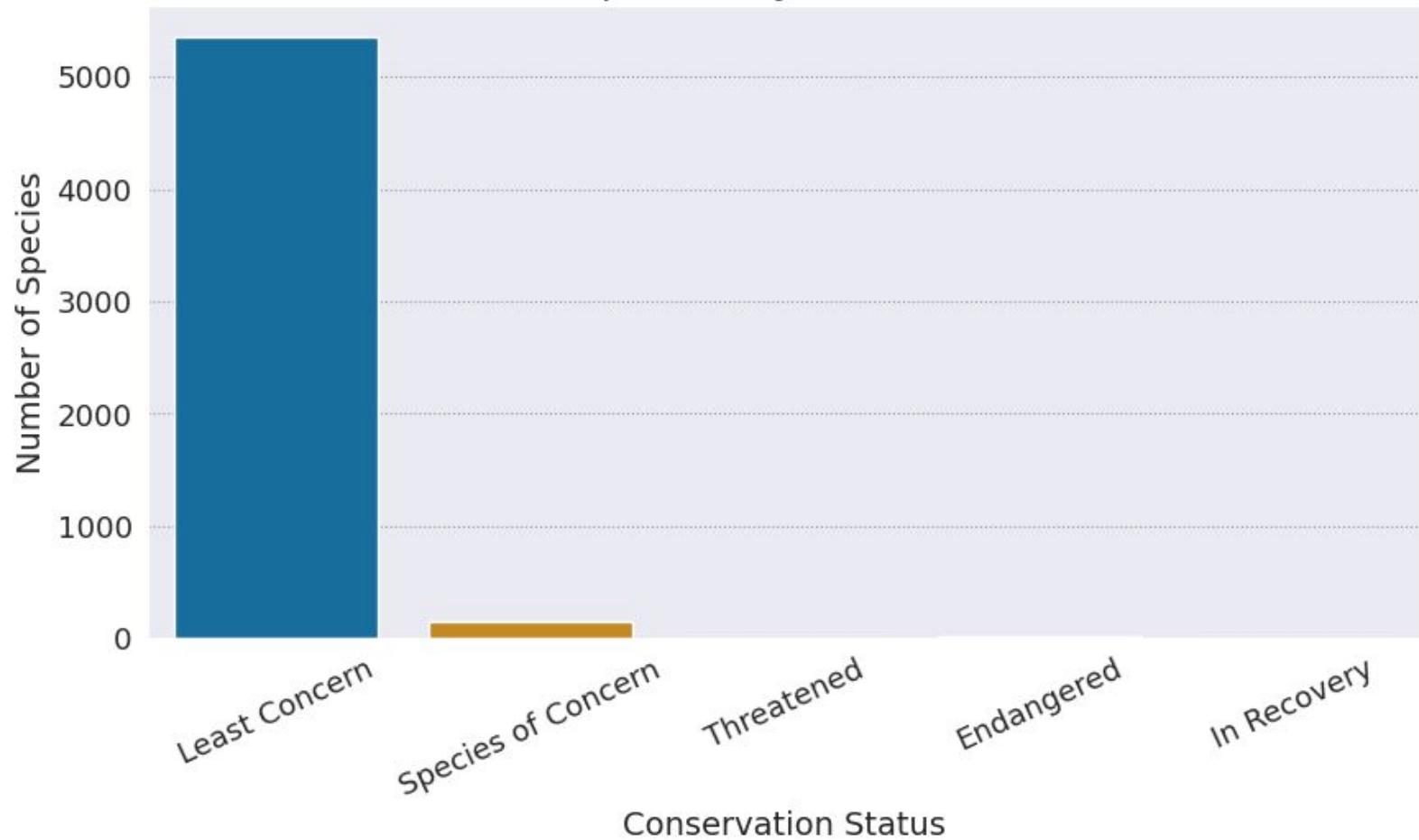


# Breakdown by Conservation Status

- Most species fall under “Least Concern” status:

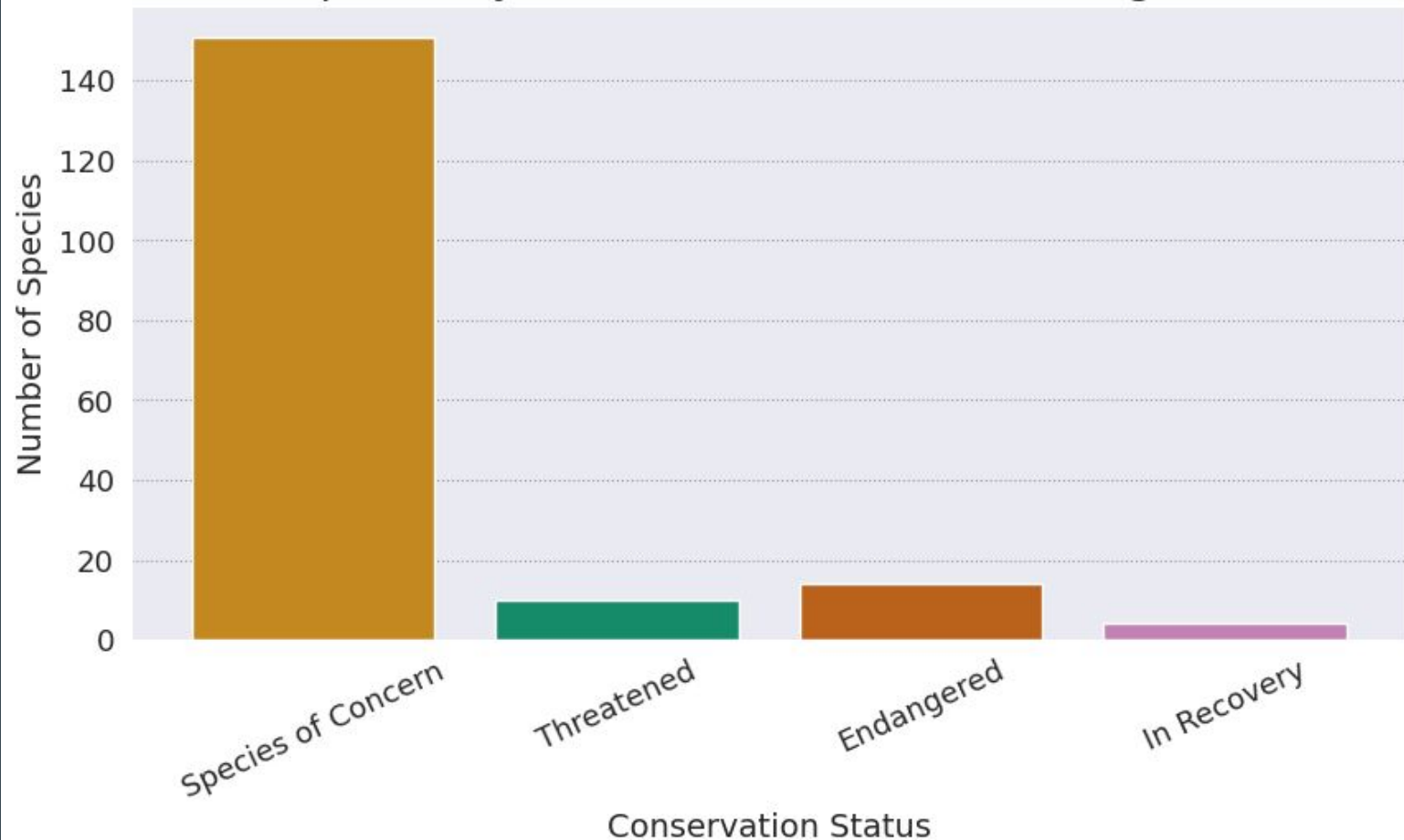
Conservation Status	Number of Species	Percent
Least Concern	5362	96.77
Species of Concern	151	2.73
Threatened	10	0.18
Endangered	14	0.25
In Recovery	4	0.07

# Number of Species by Conservation Status





Number of Species by Conservation Status, Omitting "Least Concern"

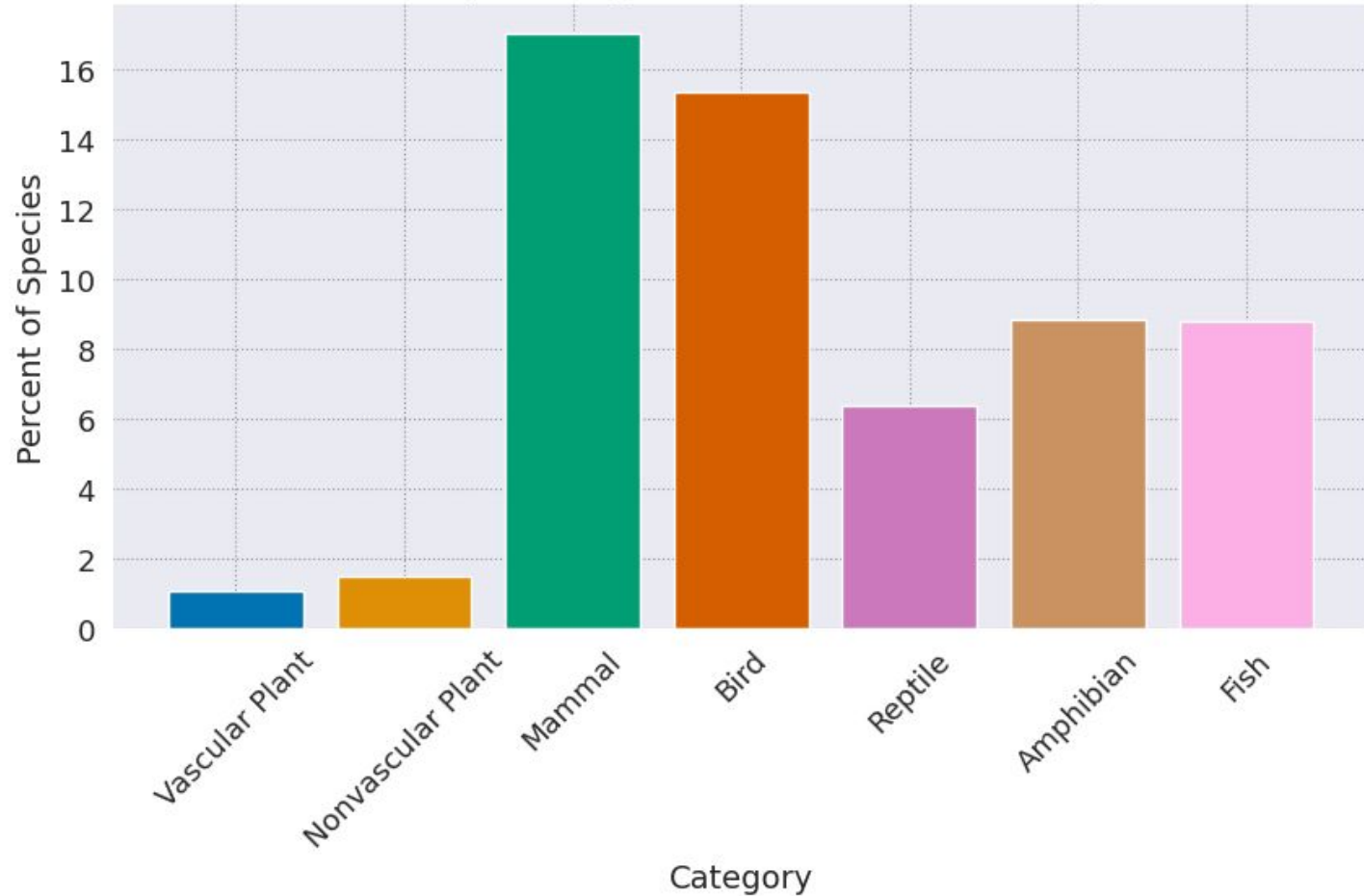


# Species Under “Protected” Status

- Protected status encompasses all but Least Concern. This includes:
  - Species of Concern
  - Threatened
  - Endangered
  - In Recovery
- Mammals and birds are more likely to be under protected status, plants are less likely to be under protected status

Category	Percent of Species Protected
Vascular Plant	1.08
Nonvascular Plant	1.50
Mammal	17.05
Bird	15.37
Reptile	6.41
Amphibian	8.86
Fish	8.80

Percent of Protected Species (all but "Least Concern") in each Category

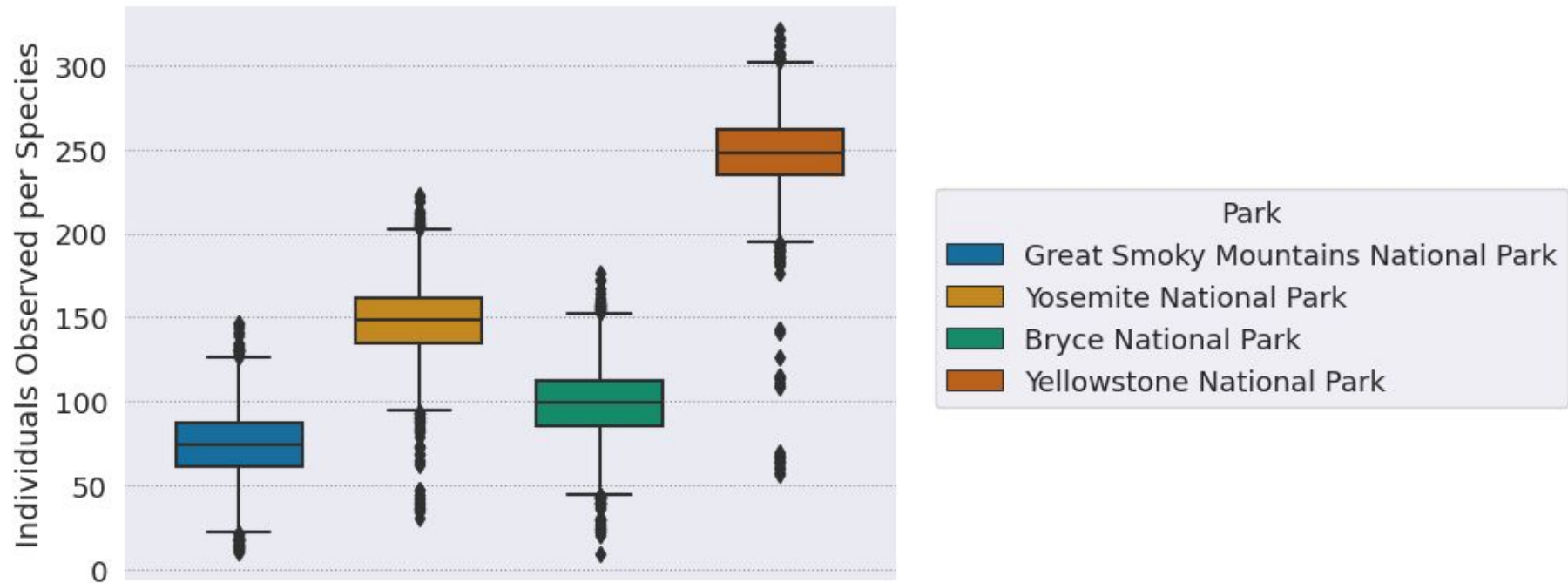


# Breakdown by Park

- Yellowstone had the largest number of total observations:

Park	Total Observations	Percent
Great Smoky Mountains	410825	13.03
Yosemite	821472	26.05
Bryce	548159	17.38
Yellowstone	1373535	43.55

## Distribution of Observations For Different Parks

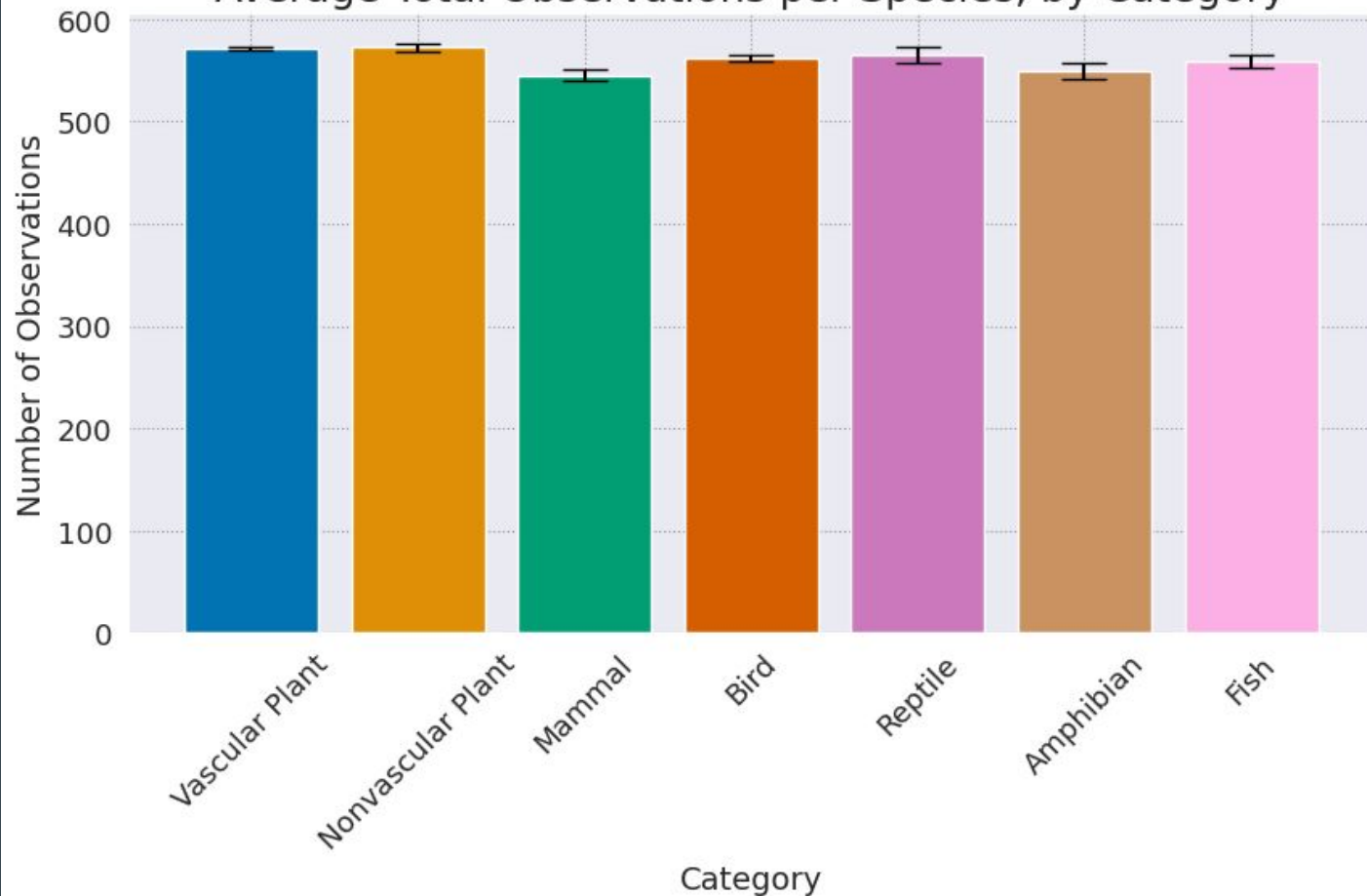


# Total Observations by Category

- The average of the total observations per species (sum of observations in all four parks) did not vary much across category:

Category	Average Observations per Species
Vascular Plant	571.5
Nonvascular Plant	572.5
Mammal	545.0
Bird	561.8
Reptile	565.4
Amphibian	549.6
Fish	558.2

Average Total Observations per Species, by Category



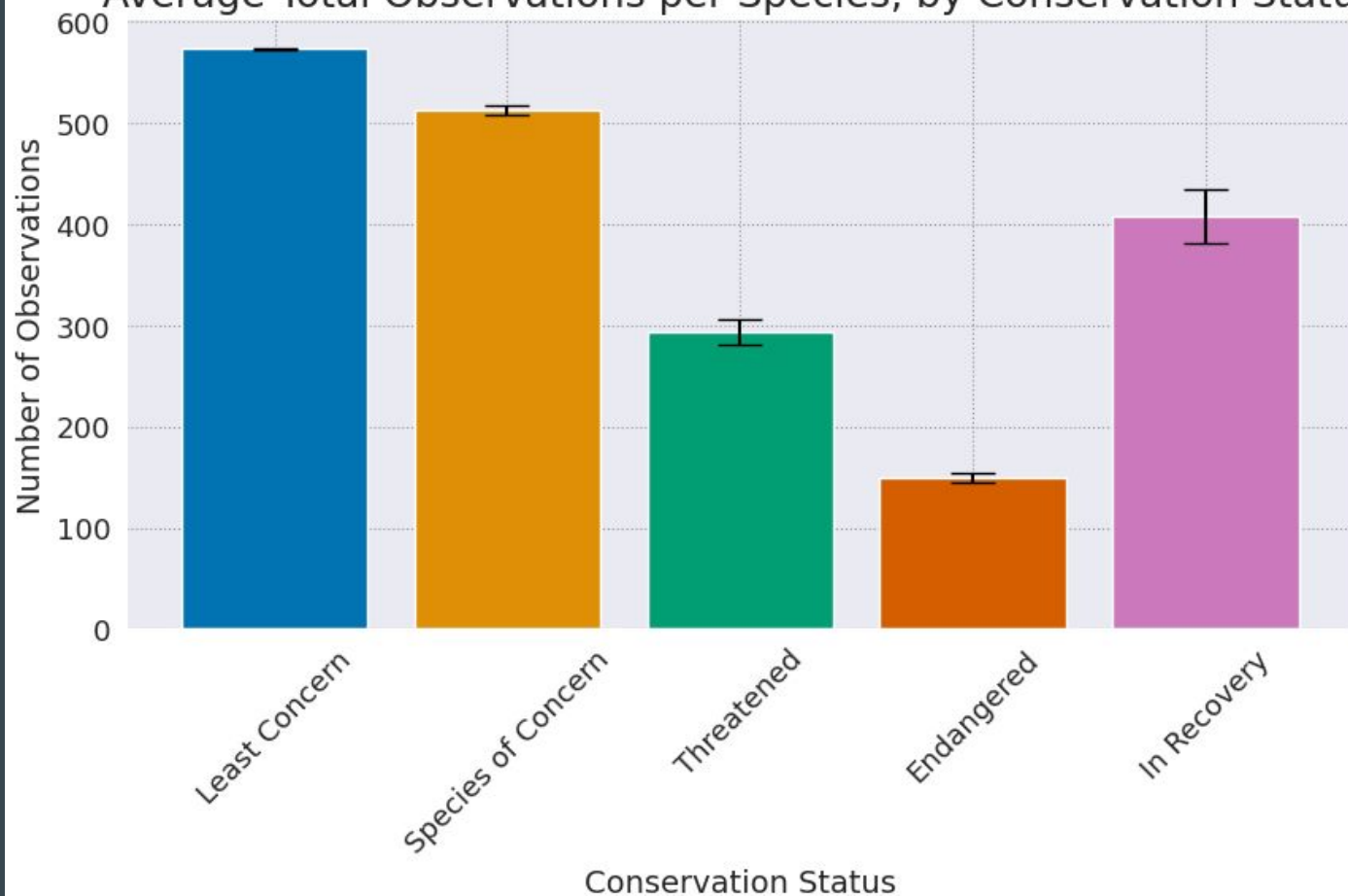
# Total Observations by Conservation Status

- The average of the total observations per species varies significantly across conservation status:

Conservation Status	Average Observations per Species
Least Concern	572.5
Species of Concern	512.5
Threatened	293.2
Endangered	148.8
In Recovery	407.2



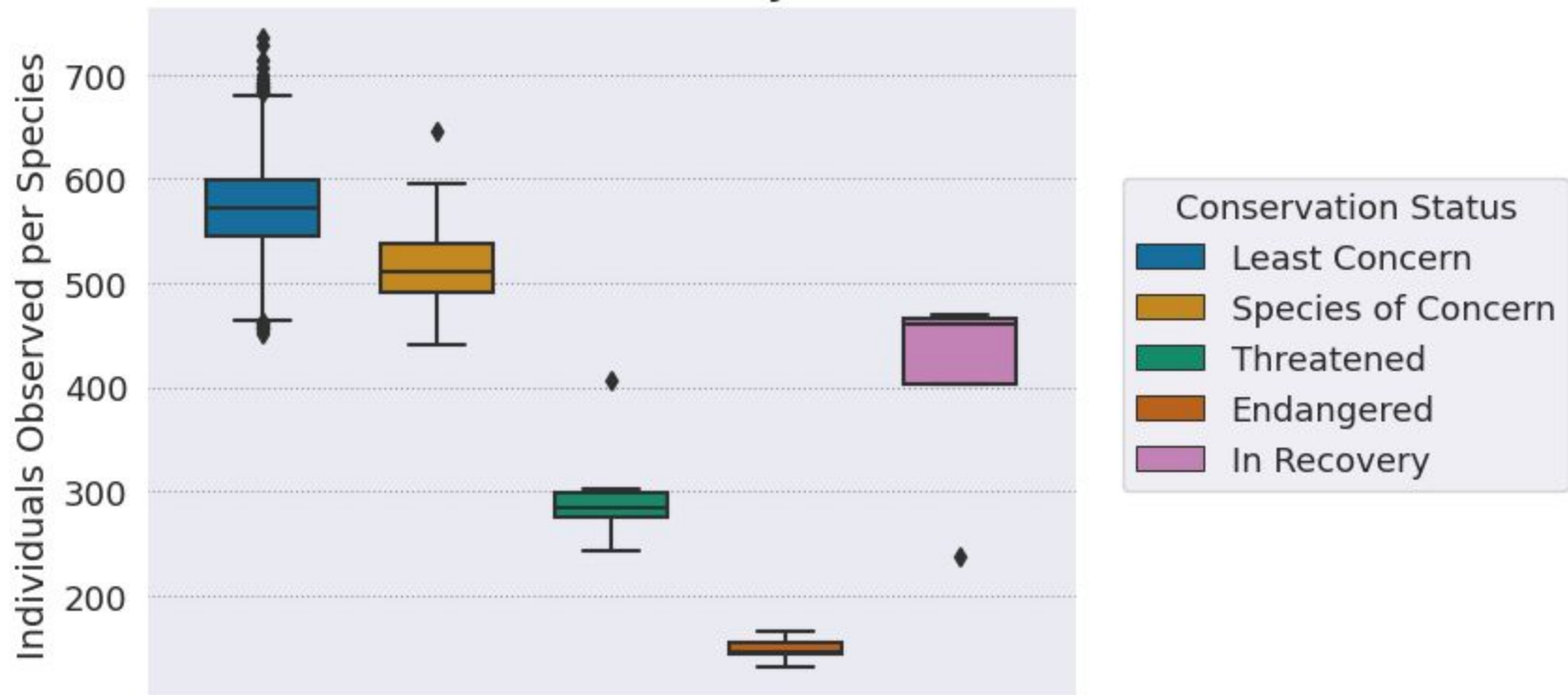
Average Total Observations per Species, by Conservation Status



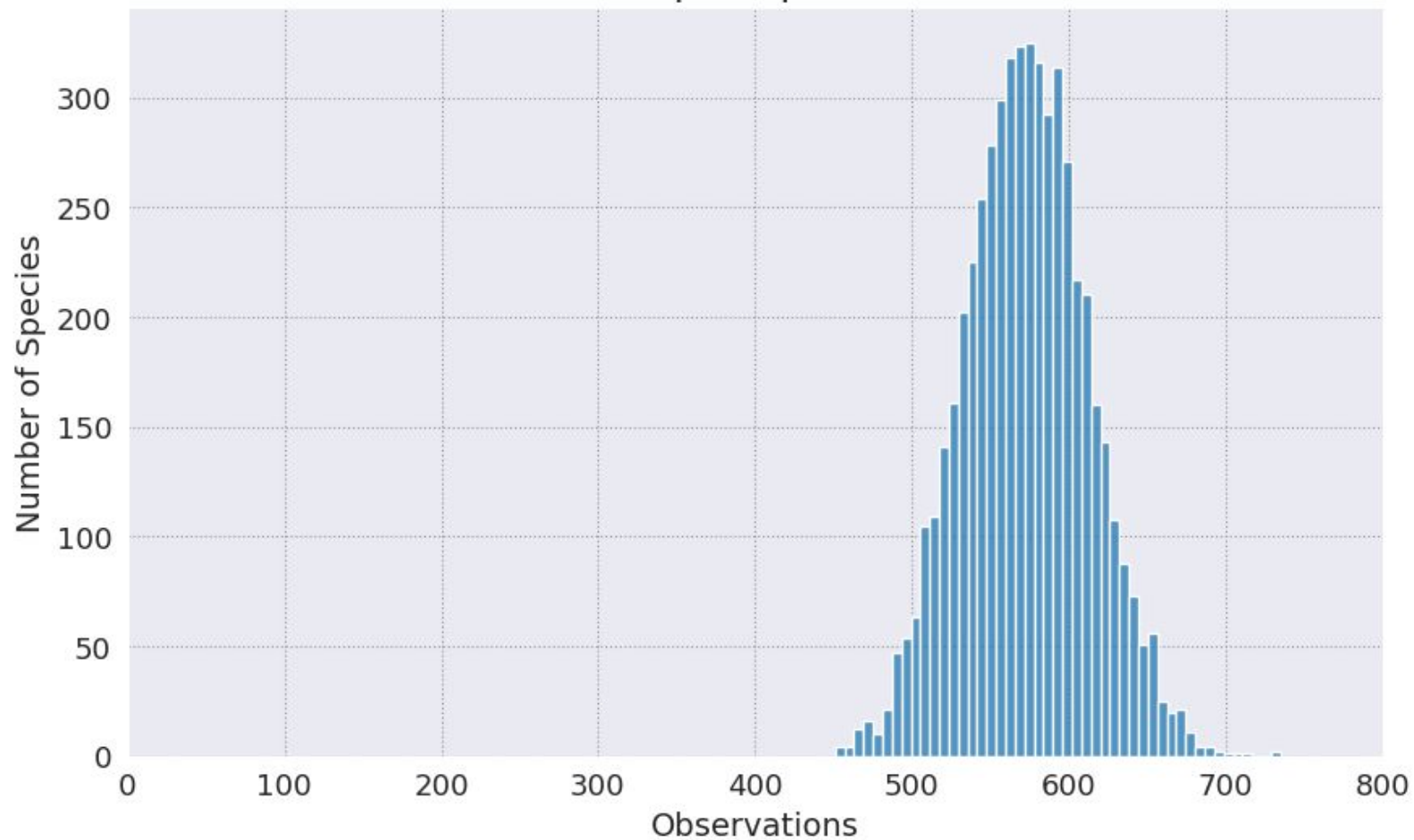
# Total Observations by Conservation Status

- The trend of decreasing observations for increasing risk of extinction can also be clearly seen in the full distributions:

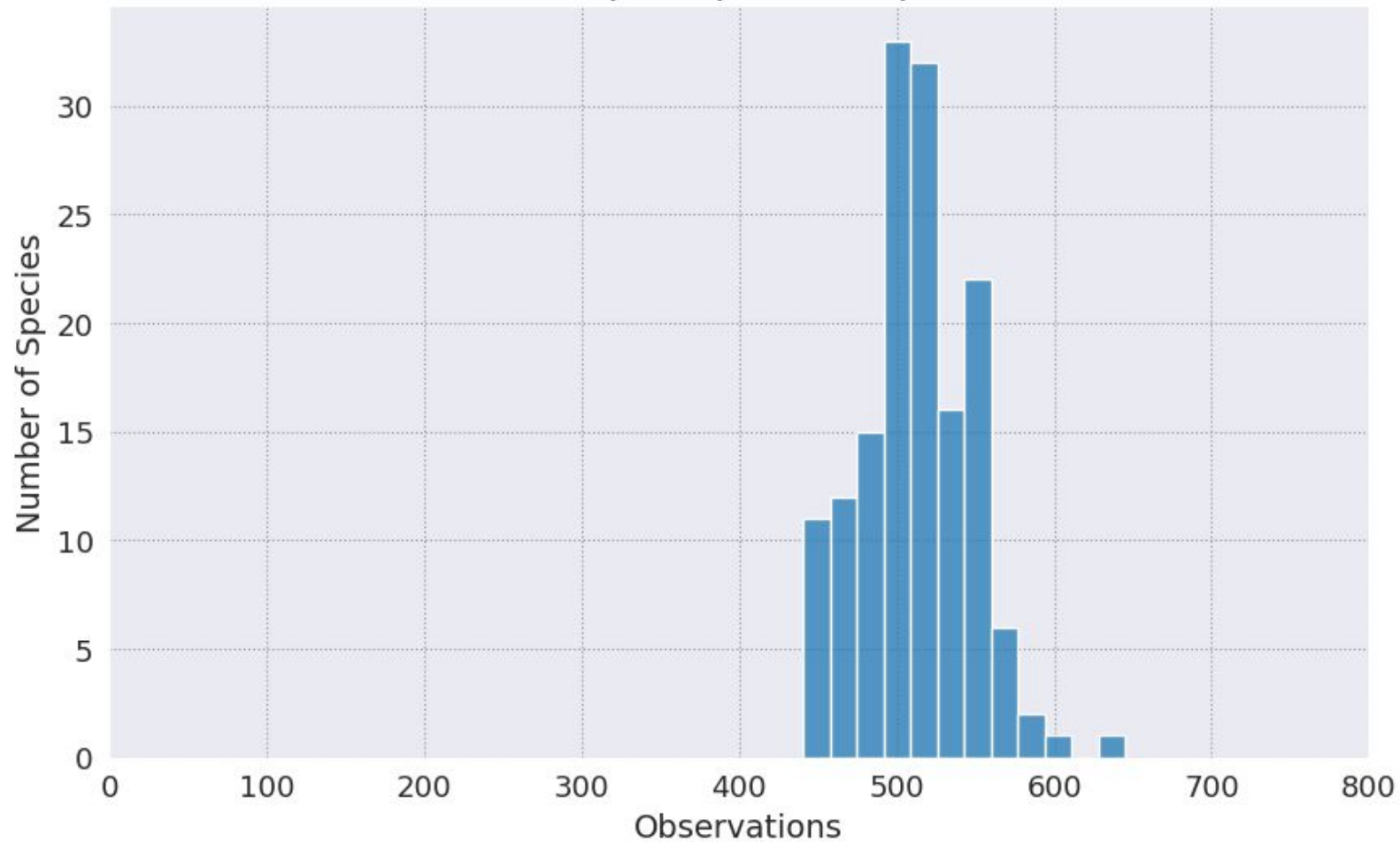
## Distribution of Total Observations by Conservation Status



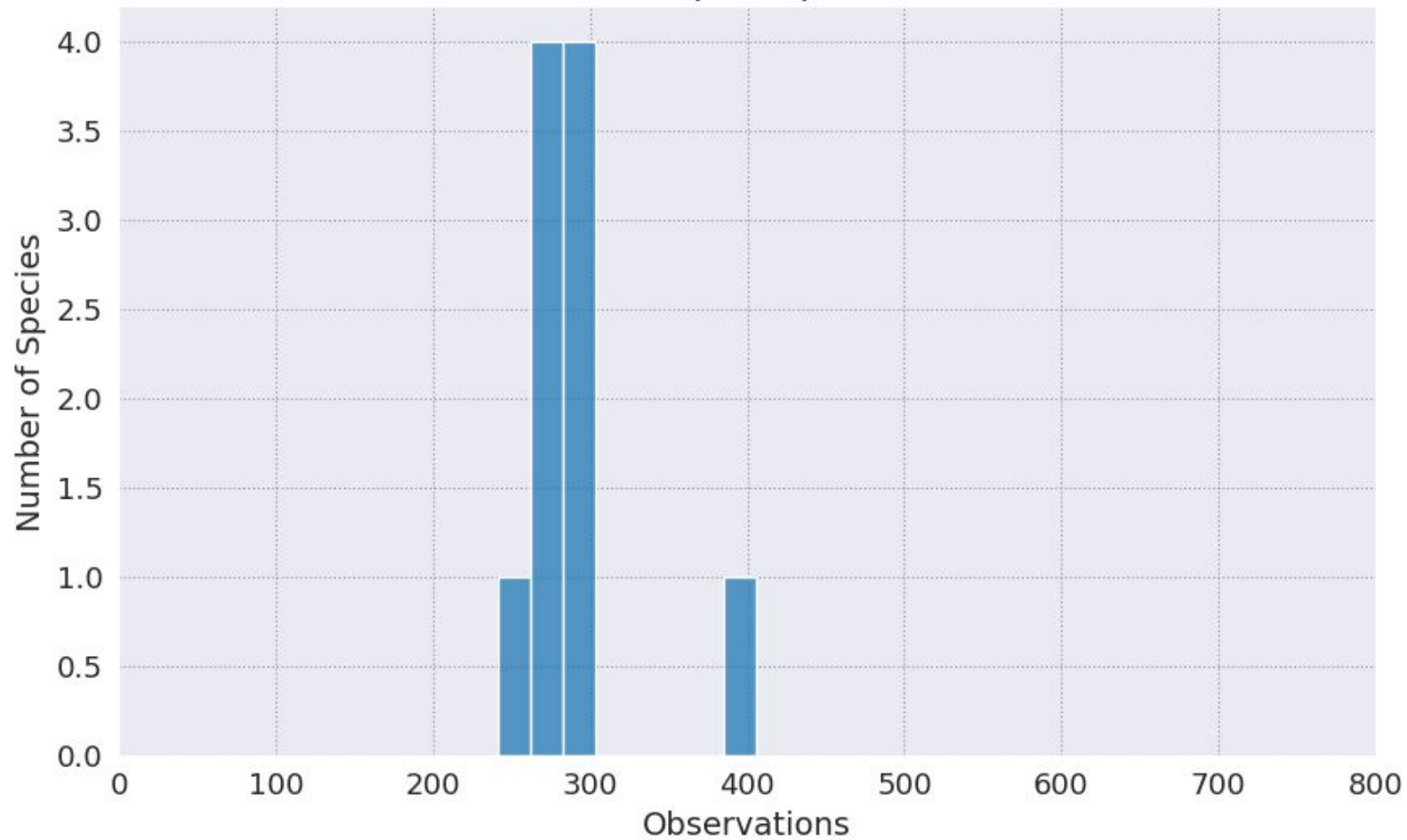
Total Observations per Species, Least Concern



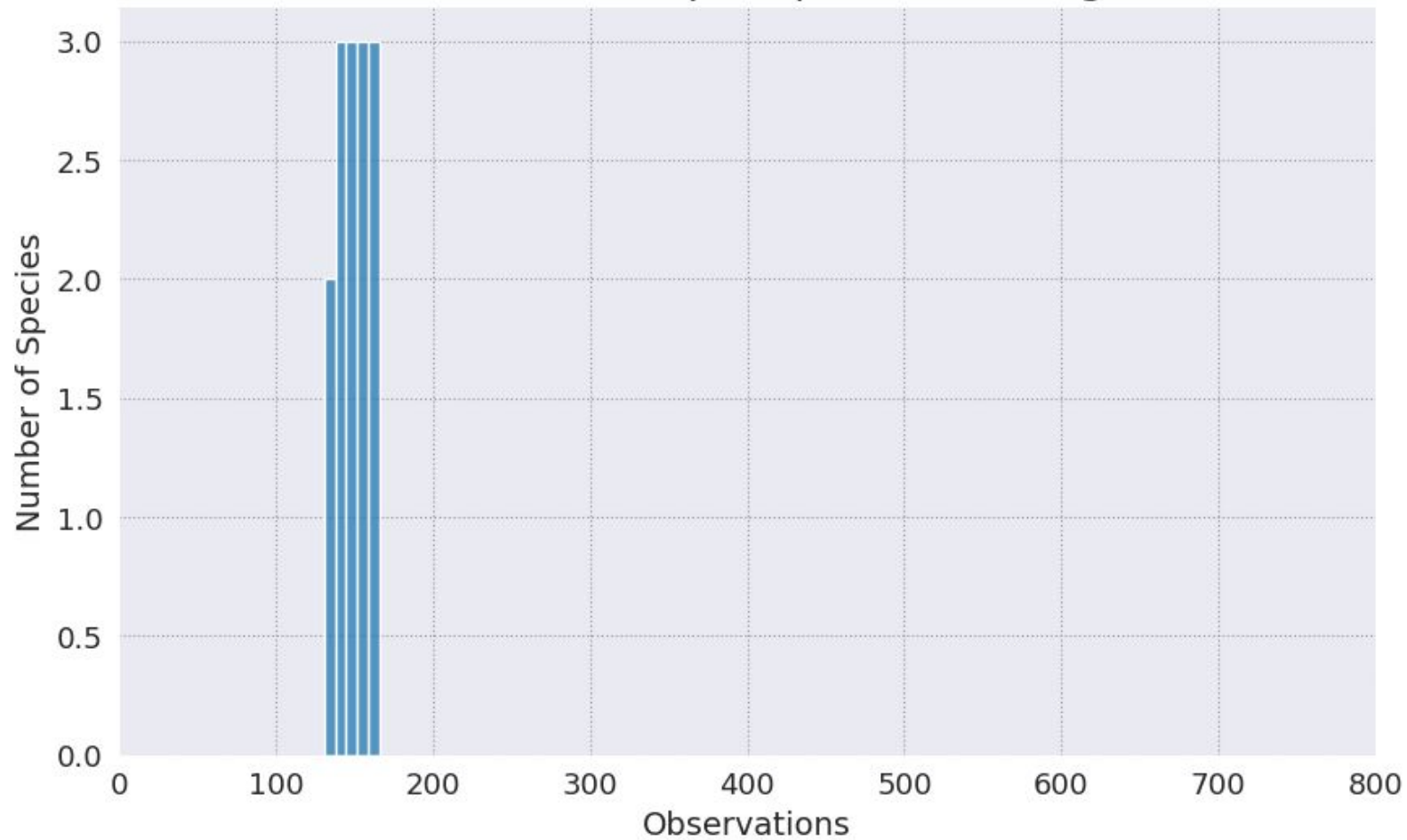
Total Observations per Species, Species of Concern



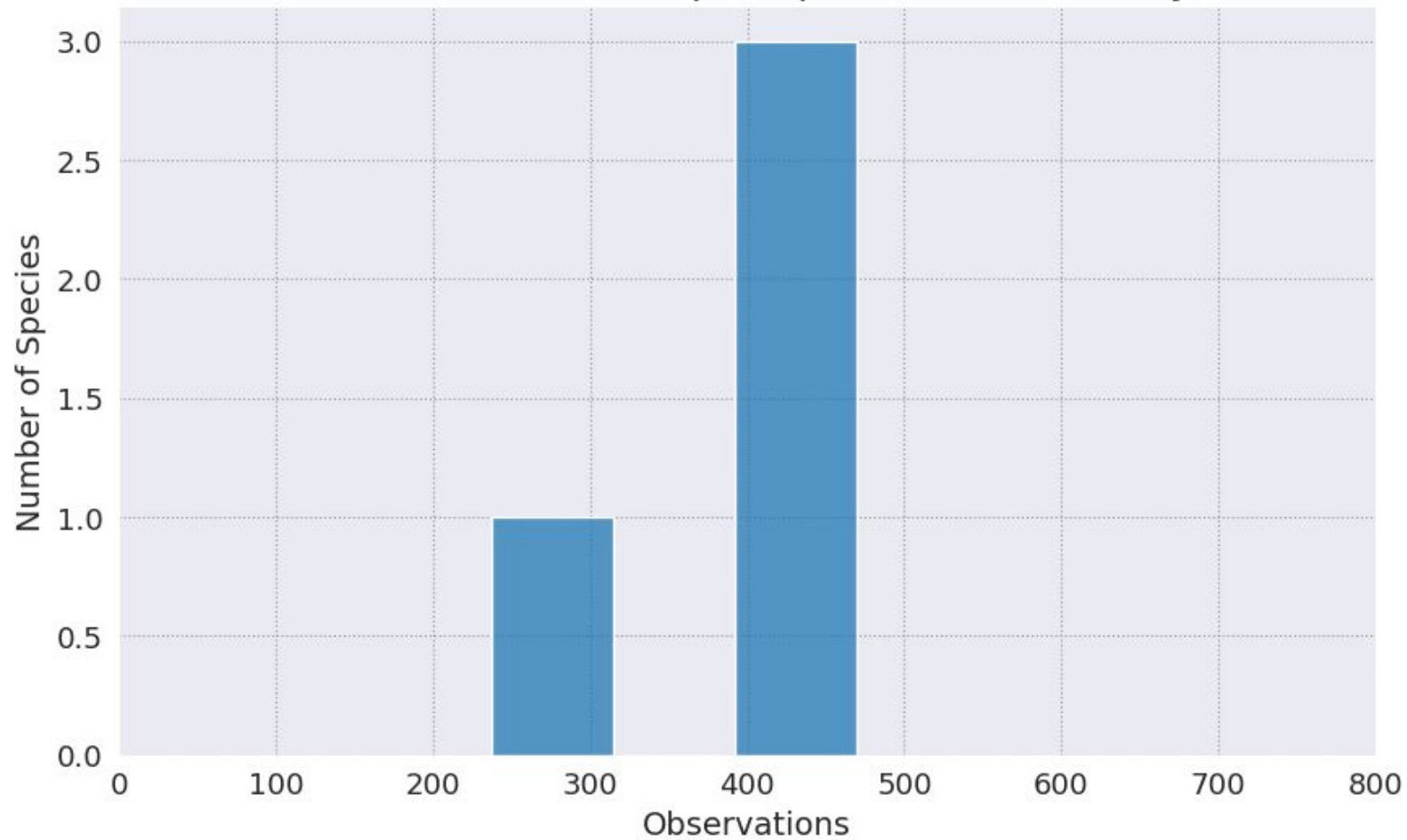
Total Observations per Species, Threatened



Total Observations per Species, Endangered

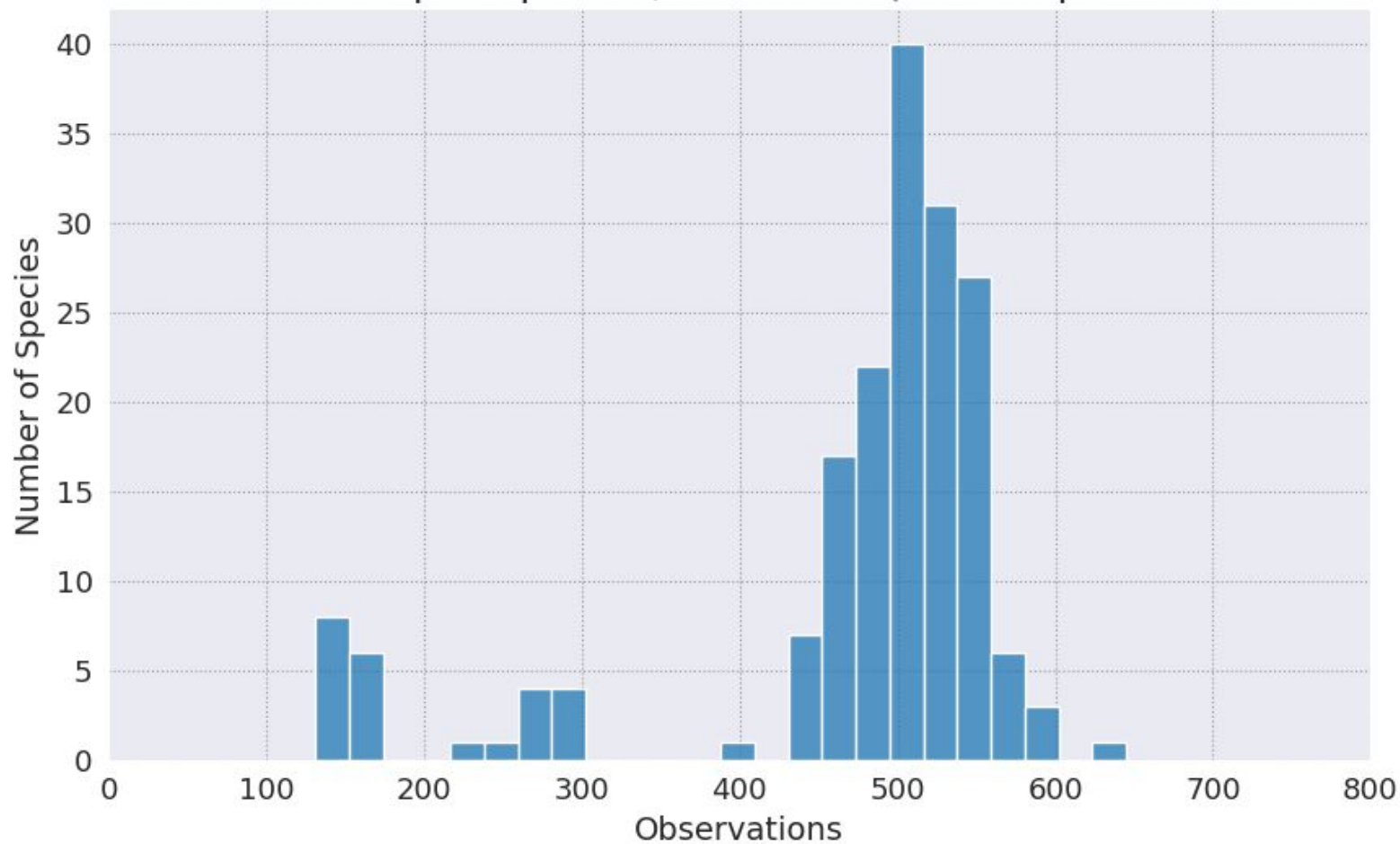


Total Observations per Species, In Recovery





Total Observations per Species, Protected (all except "Least Concern")



# Summary of Results

- Most species are vascular plants (77%)
- Most species are under Least Concern status (97%)
- Mammals and birds are more likely to be under protected status, plants are less likely to be under protected status
- The number of times a species was observed did not vary much by category, but was highly dependent on conservation status (fewer observations for species in greater danger of extinction)
  - This is evidence that the conservation statuses are meaningful, and accurately assigned

# Future Work

- Expand study to include more parks
- Gather data for observations over a wider range of time
  - Look for seasonal variation
  - Determine long-term trends (e.g., can any endangered species be shifted to “in recovery” status?)
- Determine why some parks had more observations than others

# Appendix: Data Cleanup

- NaN for conservation status assumed to be “Least Concern”
- Dropped duplicate values for scientific names
- Replaced duplicate values for observations of a given species in a given park with the average value
- Two species had multiple entries with conflicting information for conservation status:
  - Gray Wolf: kept “In Recovery,” dropped “Endangered”  
[https://en.wikipedia.org/wiki/Wolf#Status\\_and\\_conservation](https://en.wikipedia.org/wiki/Wolf#Status_and_conservation)
  - Rainbow Trout: kept “Threatened,” dropped “Least Concern”  
[https://en.wikipedia.org/wiki/Rainbow\\_trout#Conservation](https://en.wikipedia.org/wiki/Rainbow_trout#Conservation)