Data Analysis Report

Biodiversity in National Parks

By Ranchana Kiriyapong

1. Introduction

This is the study of biodiversity in four national parks, Yellowstone, Yosemite, Bryce and Great Smoky Mountains national parks in the U.S. The dataset was from national park service. Given observations records of species in each national park and information associated with that species, this analysis report aims at analyzing the biodiversity of lives by categories and across four national parks, listing the lives in specific conservation status and examining the distributions of observations in aspects of categories, national parks, and conservation status.

However, these were analyzed regarding their total number of observations or number remaining. Other aspects of evaluations of their conservation status such as their increase / decrease in populations over time or known threats were not included.

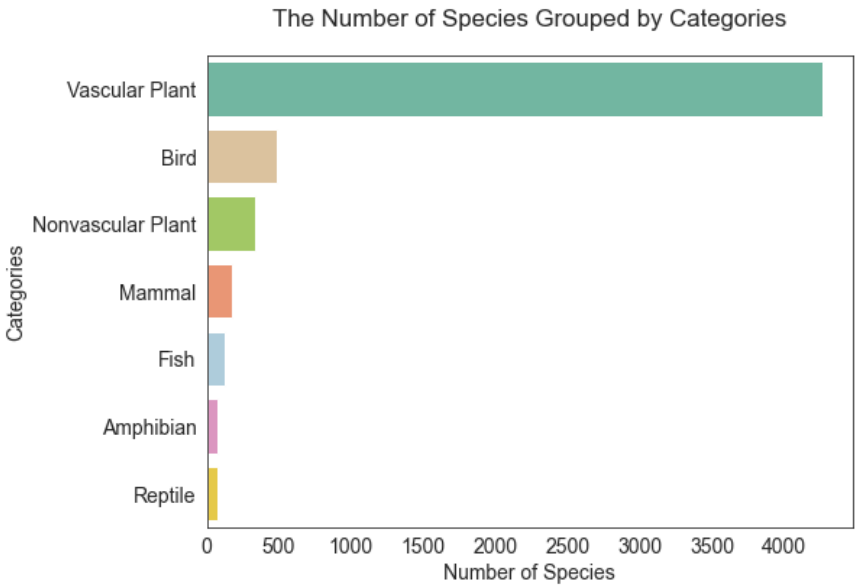
2. Analysis

Getting data from four national parks service, the dataset includes observations of each species found in each park.

2.1 Biodiversity by Categories

Method: Counting the number of species (distinct scientific name) found in every park, grouped by categories and plotted a bar chart to see the proportions of each category.

category	species_num	percentage
Vascular Plant	4262	76.92%
Bird	488	8.81%
Nonvascular Plant	333	6.01%
Mammal	176	3.18%
Fish	125	2.26%
Amphibian	79	1.43%
Reptile	78	1.41%
Total	5541	100.00%



Analysis:

From total 5,541 species found in all parks, vascular plants dominated the proportions of species since there were 4,262 species of them (76.92%). Meanwhile, 488 species (8.81%) were birds. There were 333 species of nonvascular plants (6.01%), 176 species of mammals (3.18%), 125 species of fish (2.26%), 79 species of amphibians (1.43%) and 78 species of reptiles (1.41%).

Conclusions:

The majorities of all species were vascular plants which had 4,262 of total 5,541 species. Birds were the second category that had a lot of variations (488 species). Amphibian and Reptile had the least diversity as they had got only 79 and 78 species.

2.2 Biodiversity by Parks

Method: Counting the number of species (distinct scientific name) found in each park as shown in the following table.

park_name	species_num
Bryce National Park	5541
Great Smoky Mountains National Park	5541
Yellowstone National Park	5541
Yosemite National Park	5541

Analysis:

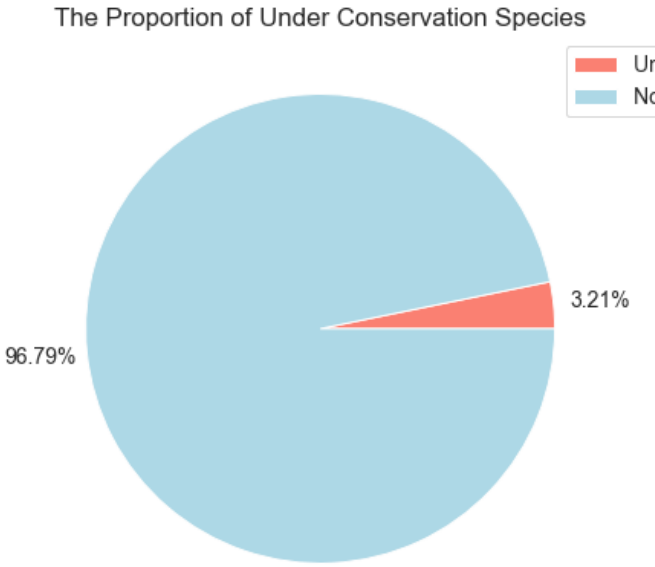
From above table, there was no difference between the numbers of species found in each park. Every national parks had 5,541 species of lives. Also, there were the same number of species by categories. (Appendix 1)

Conclusions:

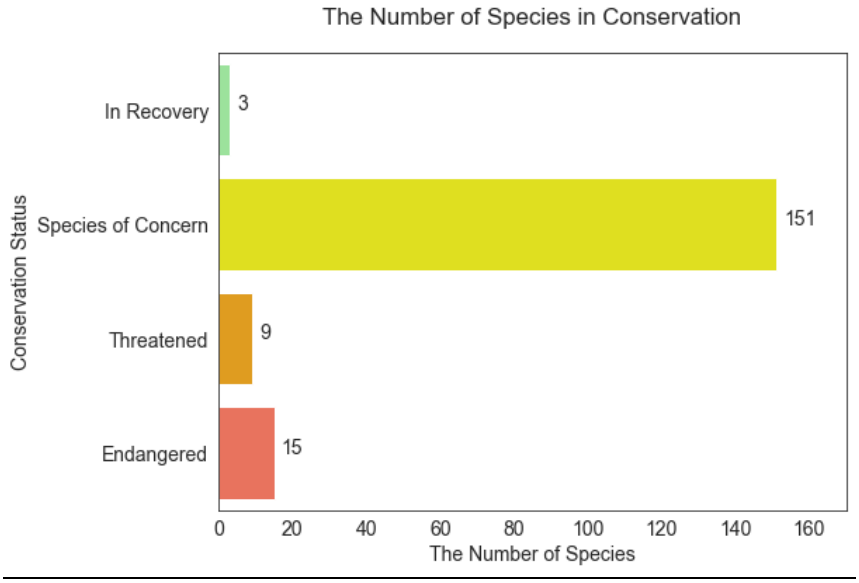
All national parks had the same number of species observed in the area (5,541 species). The cause of this issue might arise from collecting or creating the dataset.

2.3 Conservation status

Method: Counting the number of species (distinct scientific name) grouped by conservations status. Creating a pie chart to compare the percentage of species under conservation versus species that were not under conservations.



conservation_status	species_num
Endangered	15
In Recovery	3
Species of Concern	151
Threatened	9
NaN	5363
	5541



Analysis:

From the dataset, there were five conservations status including null values. Assuming the conservation status with null values meant that those species were not at risk. Only 3.21% of all species were under conservations. There were 151 species of concern, 9 species threatened, 15 species endangered and 3 species in recovery. To discover more about the categories of each conservation status, please see the appendix. (Appendix 2)

Conclusions:

From total 5,541 species, 178 species were under conservation status (3.21%). Most of them were the species of concern (151 species). The others were endangered (15 species), threatened (9 species), and in recovery (3 species).

2.4 Endangered Species

Method: Selecting the species that had "Endangered" conservations status, sorting by species' observations and then reporting the five least numbers of observations using a table form to show the number of observations by species and by national parks.

scientific_name	common_names	category	Bryce National Park	Great Smoky Mountains National Park	Yellowstone National Park	Yosemite National Park	Total
Canis rufus	Red Wolf	Mammal	30	13	60	34	137
Grus americana	Whooping Crane	Bird	24	13	57	37	131
Noturus baileyi	Smoky Madtom	Fish	22	23	67	31	143
Rana sierrae	Sierra Nevada Yellow-Legged Frog	Amphibian	31	11	60	42	144
Vermivora bachmanii	Bachman's Warbler	Bird	20	18	58	45	141
		Total	127	78	302	189	696

Analysis:

Top five endangered lives were animals: Whooping Crane, Red Wolf, Bachman's Warbler, Smoky Madtom, and Yellow-legged Frog. They had very low numbers in each park. However, the numbers in Great Smoky Mountains National Park were the smallest compared to other national parks. The range of endangered species were between 11 – 23 observations which were critical. On the other hand, Yellowstone National Parks had three times observations than Great Smoky Mountains National Park.

Conclusions:

Five species which had been endangered were Grus Americana, Canis rufus, Vermivora bachmannii, Noturus baileyi, and Rana sierrae. Their remaining numbers were the least especially, in Great Smoky Mountains National Park.

2.5 Threatened Species

Method: Selecting the species that had "Threatened" conservations status, sorting by species' observations and then reporting the five least numbers of observations using a table form to show the number of observations by species and by national parks.

scientific_name	common_names	category	Bryce National Park	Great Smoky Mountains National Park	Yellowstone National Park	Yosemite National Park	Total
Anaxyrus canorus	Yosemite Toad	Amphibian	43	30	114	87	274
Erimonax monachus	Spotfin Chub	Fish	46	18	109	69	242
Isotria medeoloides	Small Whorled Pogonia	Vascular Plant	42	36	114	82	274
Spiraea virginiana	Virginia Spiraea	Vascular Plant	57	45	114	62	278
Ursus arctos horribilis	Grizzly Bear	Mammal	38	50	115	73	276
		Total	226	179	566	373	1344

Analysis:

Top five threatened lives included both plants and animals. Threatened animals were Spotifn Chub, Yosemite Toad, and Grizzly Bear. Threatened plants were Small Whorled Pogonia and Virginia Spiraea. The number remaining of these lives were low but still higher than the endangered. As said earlier, Great Smoky Mountains National Park had least number of observations.

Conclusions:

Five species which had been threatened were Erimonax monachus, Anaxyrus canorus, Isotria medeoloides, Ursusartos horribilis, and Spiraea virginiana. Their remaining numbers were higher than the endangered. However, in Great Smoky Mountains National Park these lives were the most risky.

2.6 Species of Concern

Method: Selecting the species that had “Species of Concern” conservations status, sorting by species’ observations and then reporting the five least numbers of observations using a table form to show the number of observations by species and by national parks.

scientific_name	common_names	category	Bryce National Park	Great Smoky Mountains National Park	Yellowstone National Park	Yosemite National Park	Total
Accipiter striatus	Sharp-Shinned Hawk	Bird	62	69	200	115	446
Buteo lineatus	Red-Shouldered Hawk	Bird	85	31	224	109	449
Cathartes aura	Turkey Vulture	Bird	49	47	237	115	448
Certhia americana	Brown Creeper	Bird	79	63	189	115	446
Phalacrocorax auritus	Double-Crested Cormorant	Bird	61	45	218	117	441
		Total	336	255	1068	571	2230

Analysis:

Top five species of concern were from bird category. They consisted of Sharp-Shinned Hawk, Red-Shouldered Hawk, Turkey Vulture, Brown Creeper, and Double-Crested Cormorant. Total number of these species across all national parks were around 450 observations.

Conclusions:

Five species which were considered in concern were Accipiter striatus, Buteo lineatus, Cathates aura, Certhia Americana, and Phalacrocorax auritus. Their remaining numbers were higher than the threatened as well.

2.7 Distributions of Observations

Method: Using descriptive statistics

count	22164
mean	142.31465
std	69.805762
min	9
25%	86
50%	124
75%	195
max	321

Analysis:

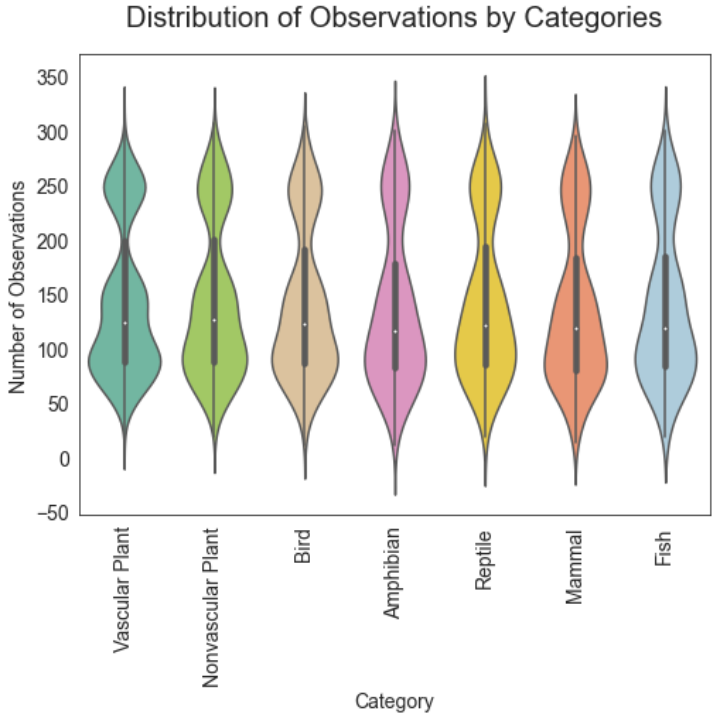
As one row of dataset was the record of a species in one national park, all species' observations varied from 9 to 321 observations each park. The observations' mean was 142. The interquartile range was 86-195 observations.

Conclusion:

50% of species in each area were in range of 86 -195 observations. The average observations were 142. The minimum and the maximum numbers were 9 and 321, in turn.

2.7.1 Distributions of Observations by Categories

Method: Plotting violin plots to illustrate overall distributions of observations and distributions of observations in aspects of categories.



Analysis:

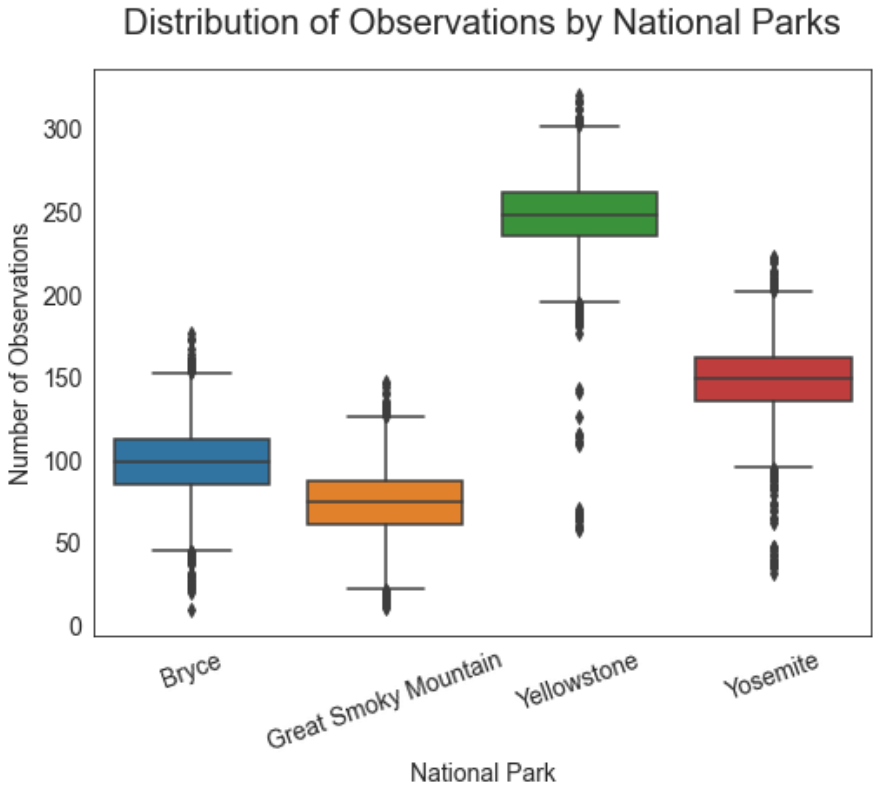
Even though the vascular plant category had much more varieties than other categories, number of observations was indifferent to the rest. All groups had almost similar distributions. The distributions were bimodal with the lower peak at about 250 observations and the high peak at about 100 observations.

Conclusion:

There were similar distributions of observations by categories which ranged from 9 – 321 observations referring to descriptive statistics. However, the distributions were bimodal with a low peak at around 250 observations and a high peak at around 100 observations.

2.7.2 Distributions of Observations by National Parks

Method: Plotting boxplots to illustrate overall distributions of observations and distributions of observations in aspects of national parks.



Analysis:

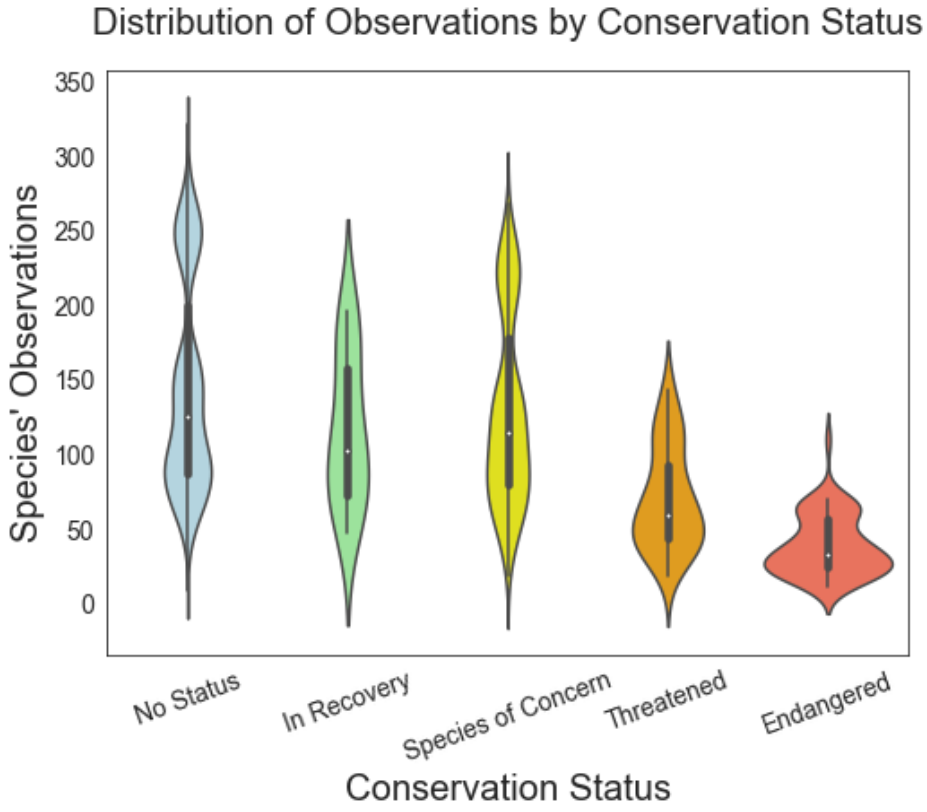
As seen before, all national parks had the same number of species. However, the number of observations were distinct. First, most of the species in Yellowstone National Park had the highest numbers of observations, 50% of species had about 250 observations. Second, Yosemite National park had the interquartile range at about 150 observations. Third, 50% of species in Bryce National Park had around 100 observations found. Great Smoky Mountains had the least number of observations at approximately 75 observations.

Conclusion:

The distributions of observations in each park were totally different. Half of species in Yellowstone, Yosemite, Bryce and Great Smoky Mountains were around 250, 150, 100 and 75 observations, respectively. This means the Great Smoky Mountains might be considered at risk.

2.7.3 Distributions of Observations by Conservation Status

Method: Plotting violin plots to illustrate overall distributions of observations and distributions of observations in aspects of conservation statuses. This visualization also included species without status.



Analysis:

There was not much difference between non-status species and species of concern. Species in all statuses of conservations had the range of number of observations that were overlapped. This indicated that the number remaining is not the only factor of conservation status labeling. However, most species in the threatened and the endangered groups had small numbers of observations. The interquartile ranges of both groups started from approximately 50 – 100 observations and 25-50 observations, in turn. Moreover, the peak of distribution in endangered group was at 25 observation and the mean was very low at around 30 observations.

Conclusion:

Species in each conservation status had different ranges of observations. Even though there were some overlapped ranges. Species threatened had shorter ranges and species endangered had the shortest ranges. Half of threatened group had the observations from 50 up to 100 observations. Meanwhile, half of endangered group had the observations from 25 to 50 observations. There was also high density of observations of 25 among endangered species.

3. Conclusions

There were 5,541 species of lives each national park. All of them can be divided into seven categories: vascular plant, non-vascular plant, bird, fish, mammal, amphibian, and reptile. Vascular plants had exceedingly high number of species (4,262 species). Meanwhile, birds ranked the second place regarding their biodiversity (488 species). 3.21% of species needed conservations. There were 151 species of concern, 9 species threatened, 15 species endangered and only 3 species in recovery. The endangered animals were as follows:

1) Grus Americana (Whooping Crane - bird)
2) Canis Rufus (Red Wolf - mammal)
3) Vermivora Bachmanii (Bachman's Warbler - bird)
4) Noturus Baileyi (Smoky Madtom - fish)
5) Rana Sierrae (Yellow-Legged Frog – amphibian)

However, the levels of severity at each national parks were different. In Great Smoky Mountains national park, the number of these animals were at risk (about 10 observations per species). Similarly, Bryce national park had small numbers of them (20-30 observations). These were along with the total number of observations in the parks as Great Smoky Mountains and Bryce national parks had the least number of observations compared to other parks.

The distributions of observations were indifferent across all seven categories. There were distinct distribution of observations in regard to national parks. Most species in Yellowstone National Park had the greatest number of observations. Half of species in Yellowstone, Yosemite, Bryce and Great Smoky Mountains were around 250, 150, 100 and 75 observations, respectively. This means the Great Smoky Mountains might be considered at risk.

4. Appendices

Appendix 1 Biodiversity by Categories in Each National Park

Method: Counting the number of species found grouped by categories as regards national parks. Formatting the results in the tabular format.

Category	Bryce National Park	Great Smoky Mountains National Park	Yellowstone National Park	Yosemite National Park	Total
Amphibian	79	79	79	79	316
Bird	488	488	488	488	1952
Fish	125	125	125	125	500
Mammal	176	176	176	176	704
Nonvascular Plant	333	333	333	333	1332
Reptile	78	78	78	78	312
Vascular Plant	4262	4262	4262	4262	17048
Total	5541	5541	5541	5541	22164

Analysis:

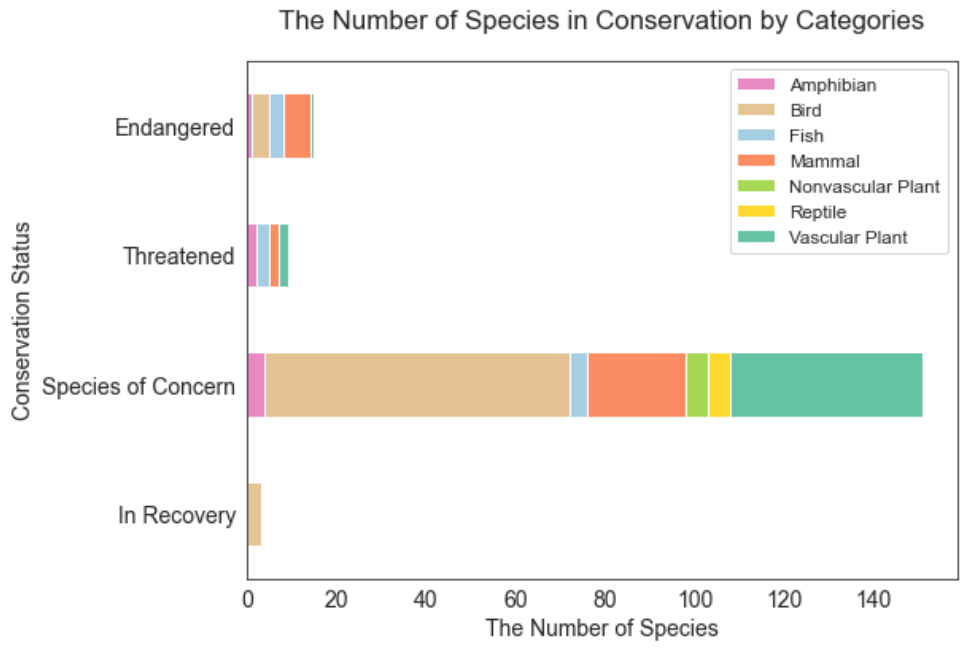
The above table shows that every national park had the same total number of species found and also the same number of species in each category.

Conclusion:

All four national parks had the same exact number of species by categories. This might be caused by data collection.

Appendix 2 Categories under conservation status

Method: Counting the number of species found grouped by categories as regards their conservation status. Creating a stacked bar chart to demonstrate the categories which were in each conservation status.



Analysis:

This stacked bar graph shows that top three categories which were species of concern were bird, vascular plant and mammal. Meanwhile, the categories of threatened species were amphibian, fish, mammal and vascular plant. Most species endangered were from mammals and birds and fish.

Conclusion:

Main categories which were endangered were mammal, bird and fish. For threatened species, they were from amphibian, fish, mammal and vascular plant categories.