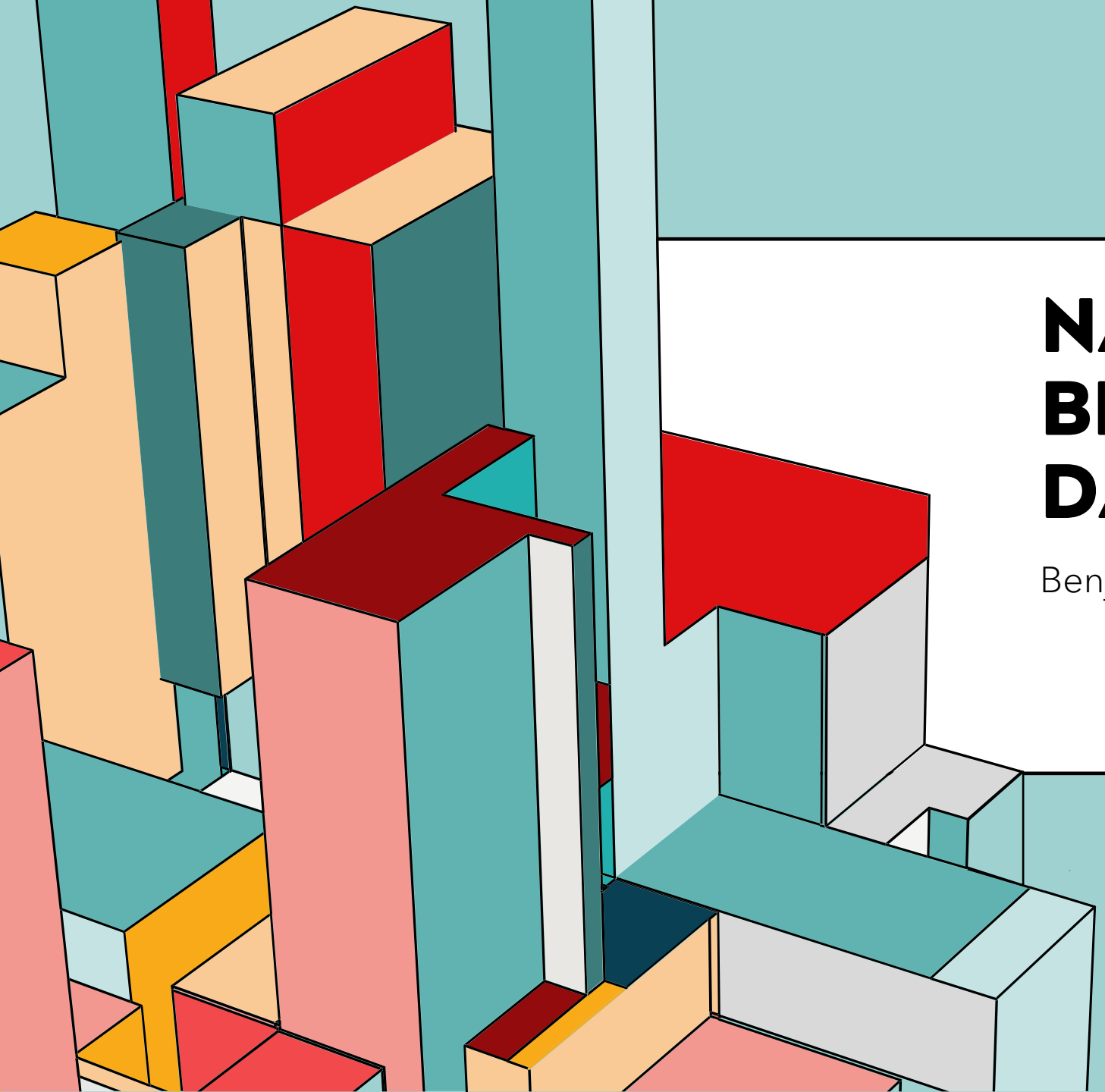
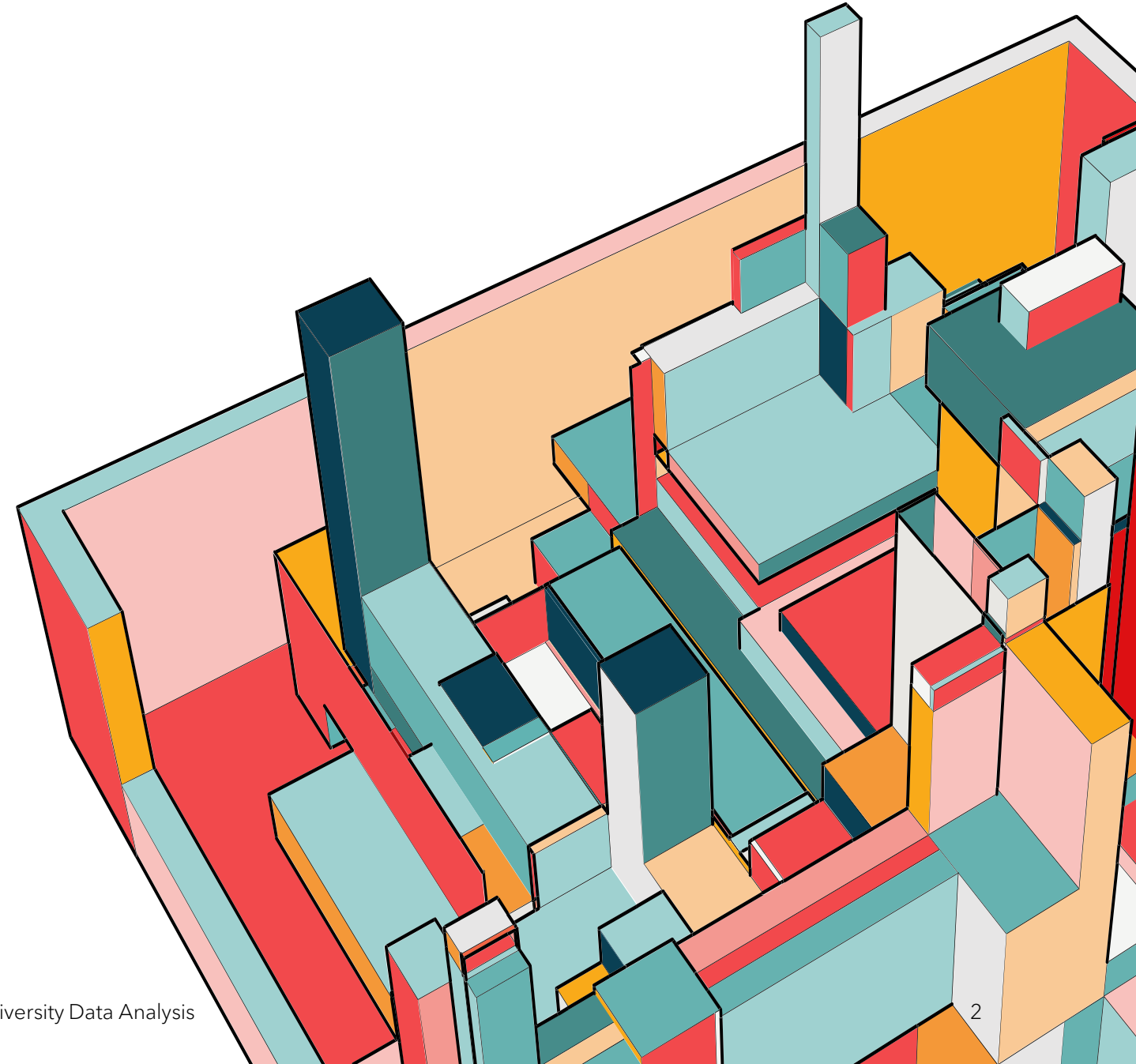


NATIONAL PARK BIODIVERSITY DATA ANALYSIS

Benjamin Sandmann

ABOUT THE DATA

- We have two datasets from Codecademy.com named 'observations.csv' and 'species_info.csv'.
- The data is fictional but inspired by real data from the National Parks Service.
- The 'observations.csv' dataset contains information on four United States National Parks.
- The 'species_info.csv' dataset contains information on different species found in United States National Parks.



DATASET BREAKDOWN

observations.csv

23,296 Rows 3 Columns

scientific_name – The scientific name of each species

park_name – The name of the National park

observations – The number of observations in the past 7 days

species.csv

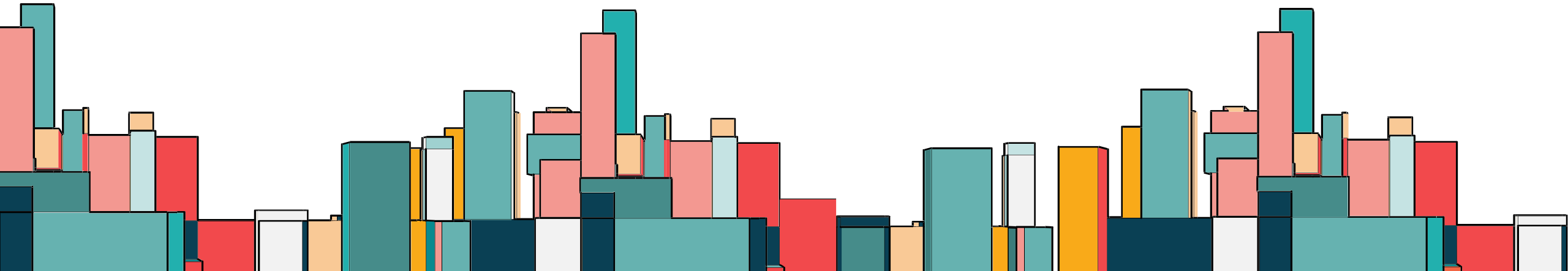
5824 Rows 4 Columns

category – The category that each species falls under

scientific_name – The scientific name of each species

common_names – The common names for each species

conservation_status – The species conservation status

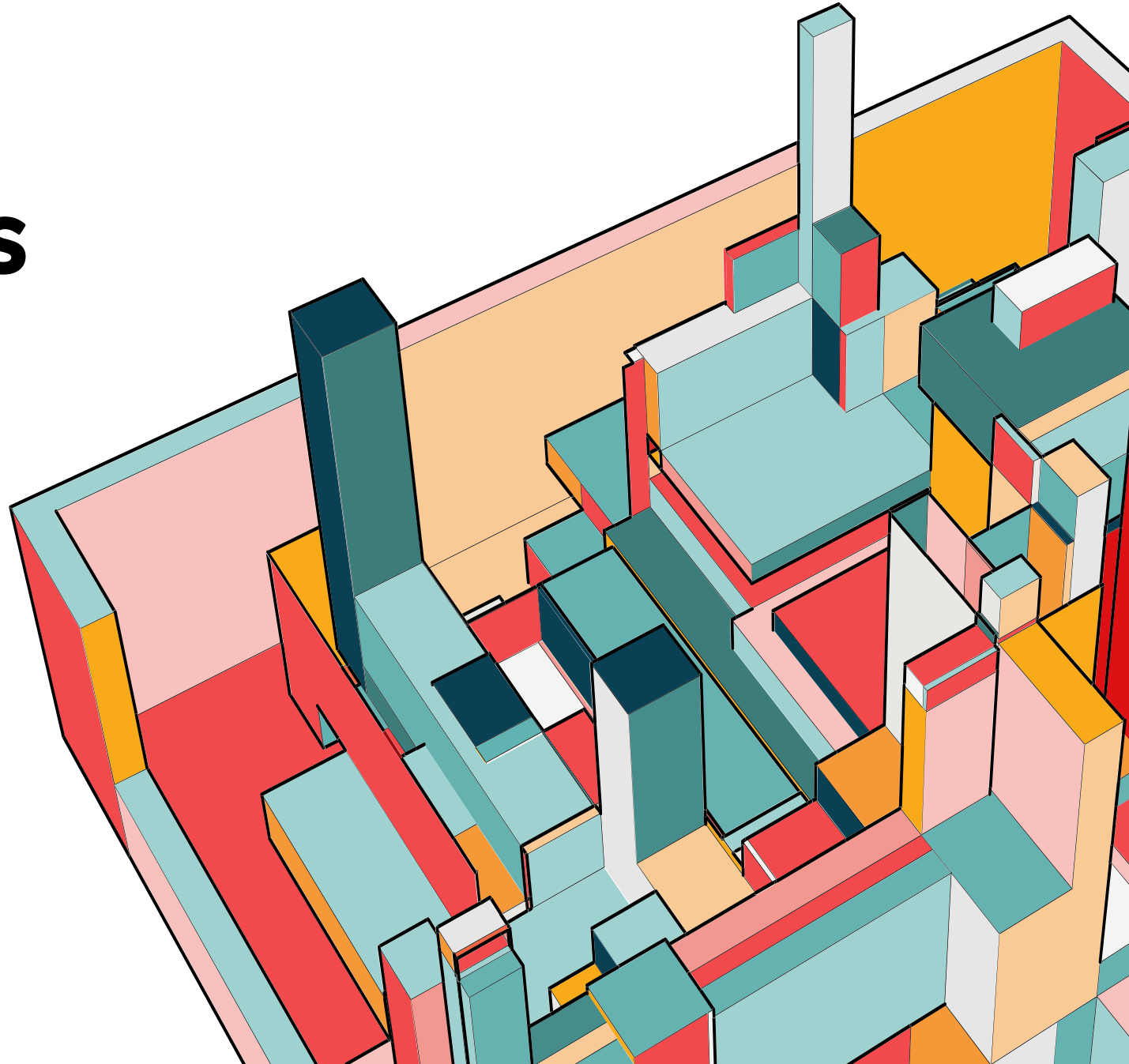


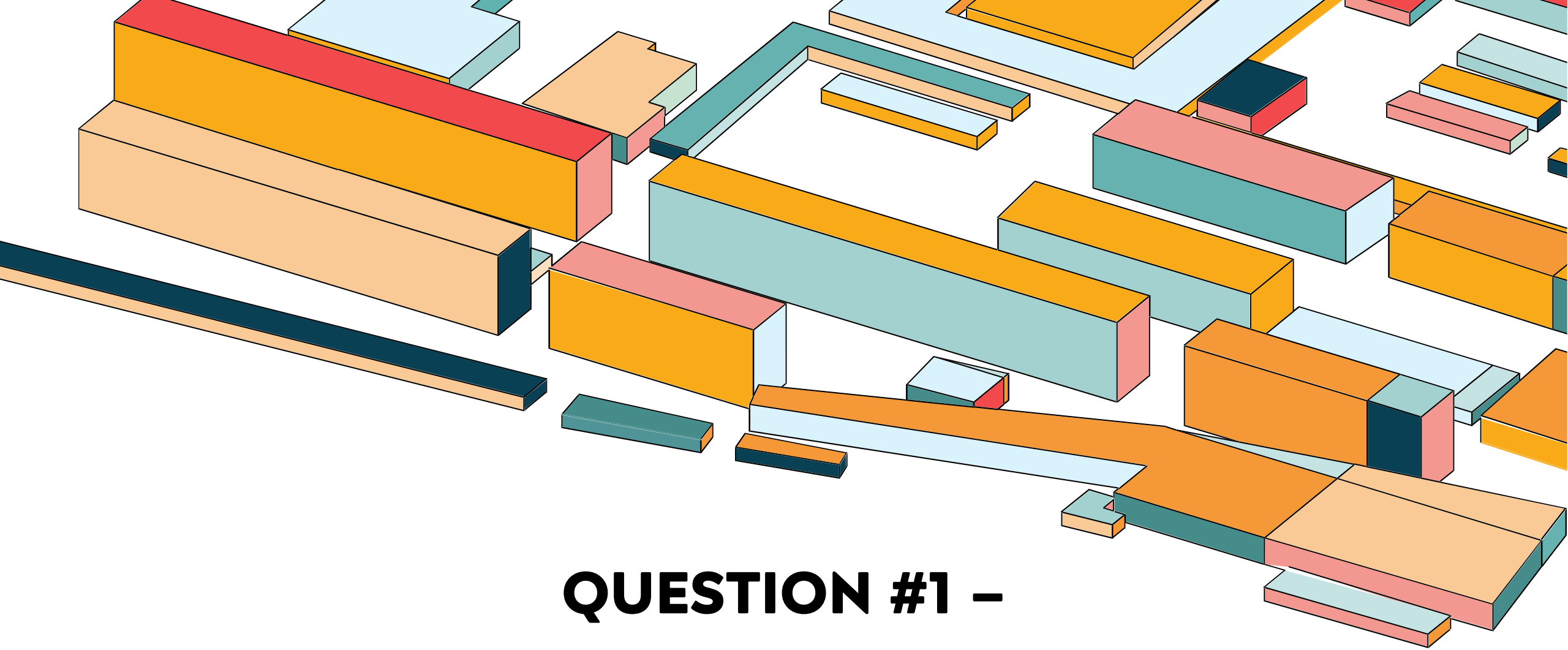
ABOUT THE ANALYSIS

By analyzing the datasets, we hope to gain some insights into the species that live within the parks and discover relationships within the data that may affect a species respective conservation status.

Conservation Statuses (Greatest risk to Least risk):

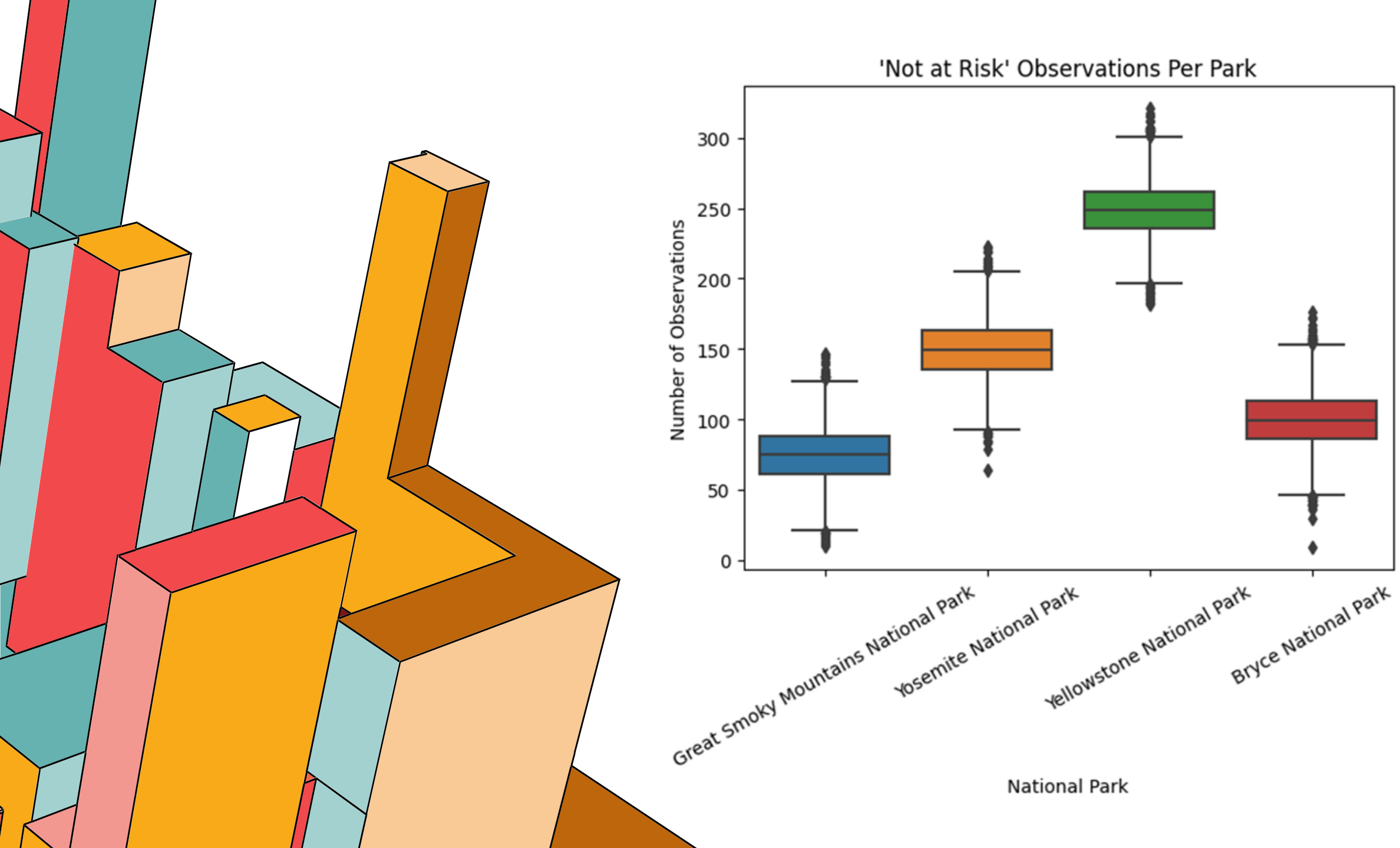
- Endangered
- Threatened
- In Recovery
- Species of Concern
- Not at Risk (Added to the data for analysis purposes)



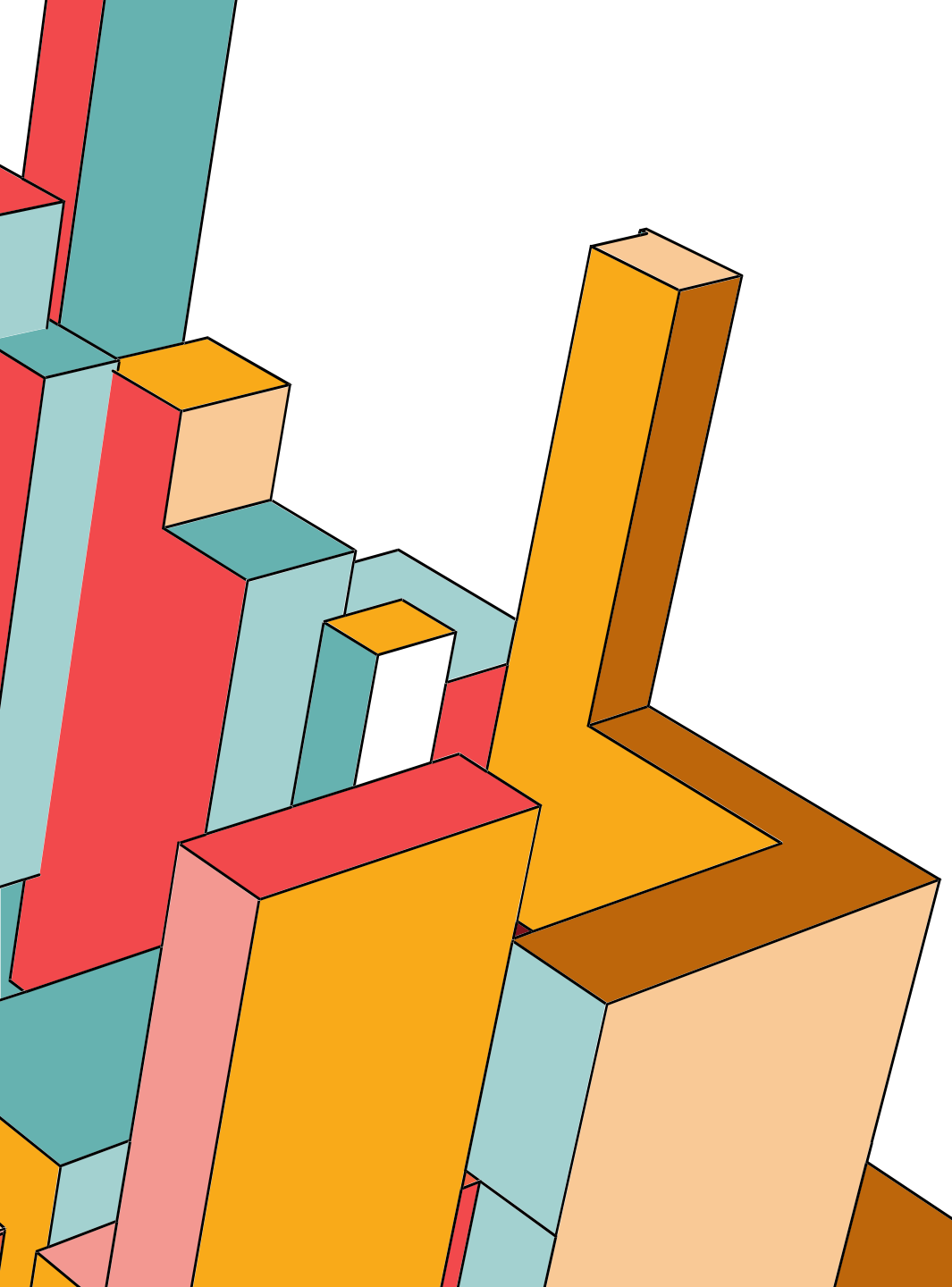


QUESTION #1 –

HOW MANY OBSERVATIONS OF ‘NOT AT RISK’ SPECIES ARE THERE IN EACH PARK?



'Not at Risk' Observations Per Park
Number of Observations
0
50
100
150
200
250
300
Great Smoky Mountains National Park
Yosemite National Park
Yellowstone National Park
Bryce National Park
National Park



Question #1 Findings –

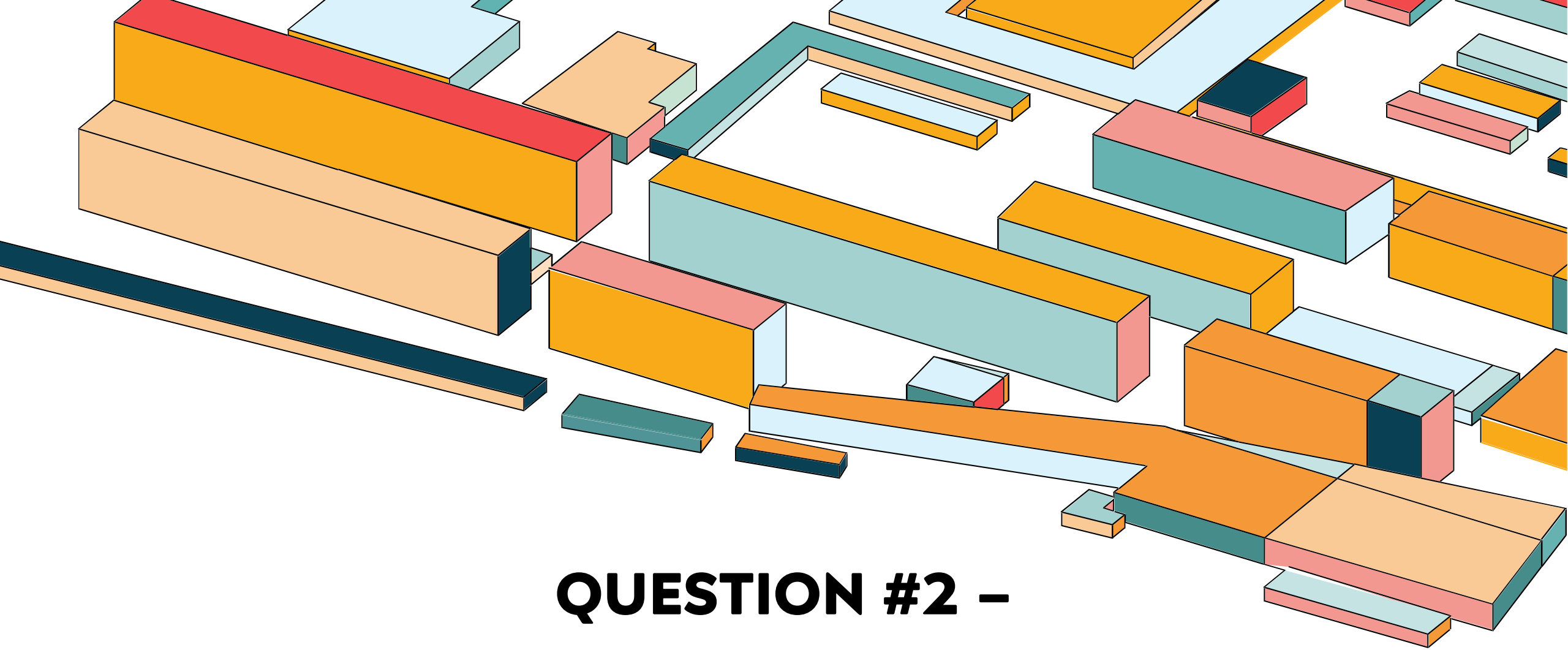
At first glance, it looks like Yellowstone National Park has significantly more species who are 'Not at Risk' compared to the other parks, but this isn't entirely true.

This poses more questions:

- Why are there more 'Not at Risk' species in certain parks and not in others?

- Are some parks more sustainable for wildlife than others?

- Is this only because certain parks have larger populations than others?



QUESTION #2 –

WHAT IS THE TOTAL NUMBER OF OBSERVATIONS IN EACH PARK?

	Conservation Status	Endangered	In Recovery	Not at Risk	Species of Concern	Threatened	Total Obsv's.
Park Name							
Bryce National Park		402	258	533093	13979	427	548,159
Great Smoky Mountains National Park		294	189	400343	9659	340	410,825
Yellowstone National Park		1008	559	1337313	33569	1087	1,373,536
Yosemite National Park		616	386	799611	20187	672	821,472

Question #2 Findings –

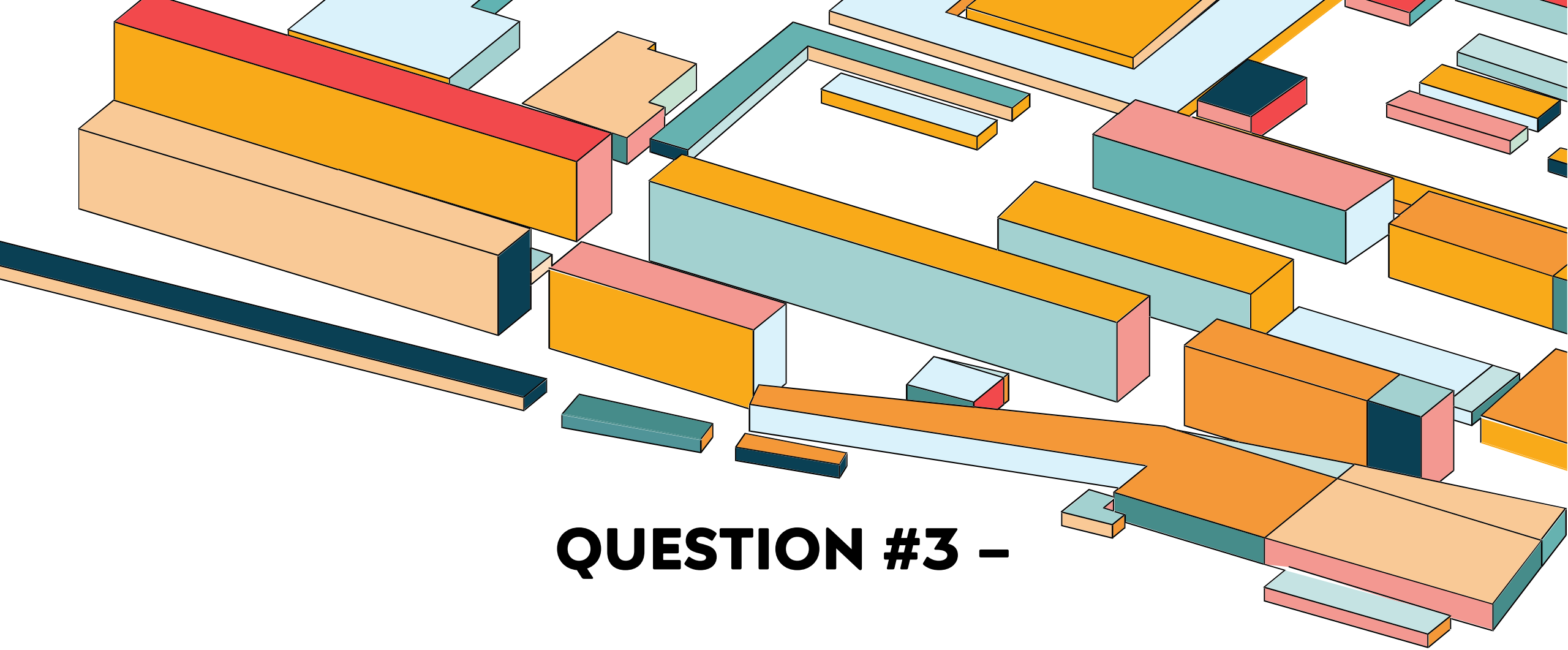
When looking at the table in the previous slide, we can see the total observations in each park have a lot of variation.

Each park's total observations from greatest to least observations:

- Yellowstone National Park - - - 1,373,536 Observations
- Yosemite National Park - - - 821,472 Observations
- Bryce National Park - - - 548,159 Observations
- Great Smoky Mountains National Park - - - 410,825 Observations

Look back at the graph and you'll see this same order is seen when looking at the amount of 'Not at Risk' observations in each park.

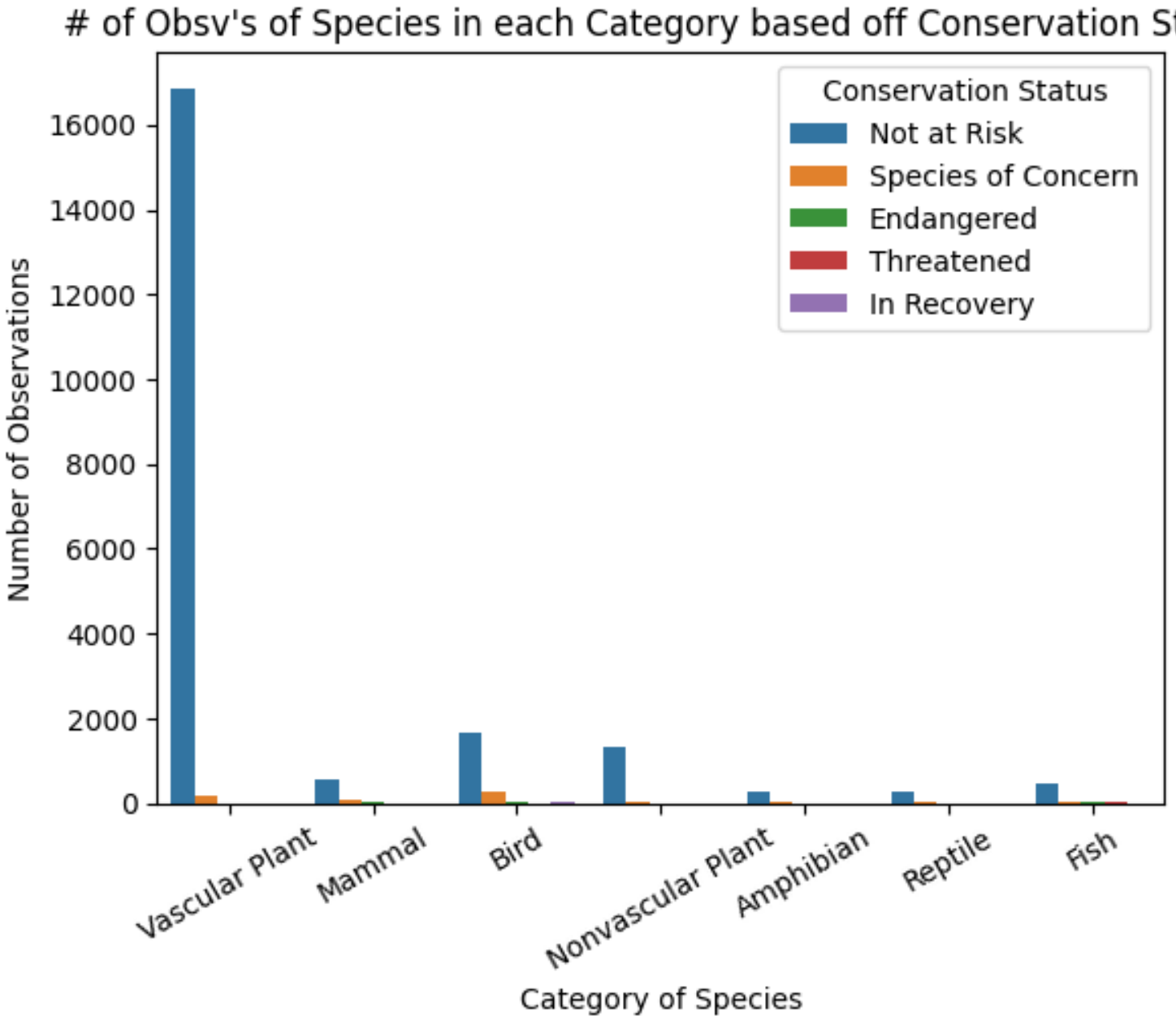
This means that Yellowstone National Park has more sightings of 'Not at Risk' species, not because it is a safer environment for wildlife, but because it has a larger total wildlife population.



QUESTION #3 –

WHAT PERCENT DO THE OBSERVATIONS OF EACH CONSERVATION STATUS MAKE UP WITHIN EACH PARK?

	Conservation Status	Endangered	In Recovery	Not at Risk	Species of Concern	Threatened
Park Name						
Bryce National Park		0.073%	0.047%	97.252%	2.55%	0.078%
Great Smoky Mountains National Park		0.072%	0.046%	97.449%	2.351%	0.083%
Yellowstone National Park		0.073%	0.041%	97.363%	2.444%	0.079%
Yosemite National Park		0.075%	0.047%	97.339%	2.457%	0.082%



What's the first thing you noticed when looking at the table on the last slide?

Every park looked nearly identical in terms of the conservation statuses and what percent they made up within each park.

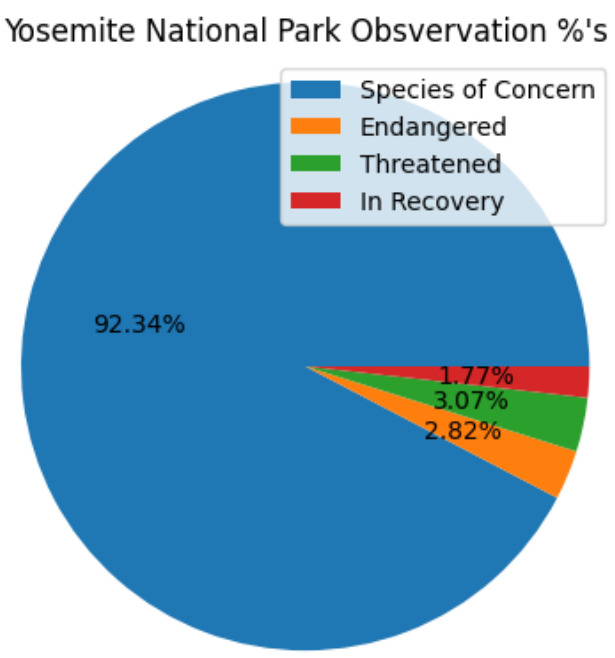
For example, no matter which park you go to, roughly 97% of the parks population will consist of species that are 'Not at Risk'.

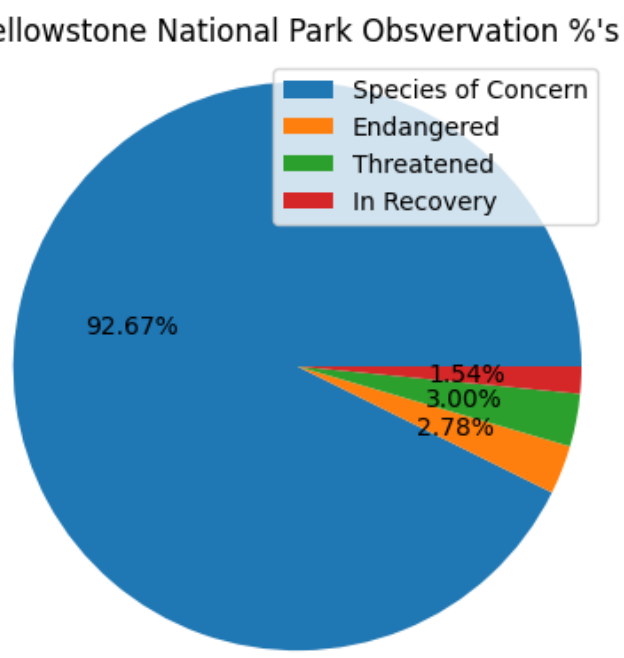
Check out the graph on the left to see a visual comparison of how many sightings of 'Not at Risk' species there were compared to species of other conservation statuses.

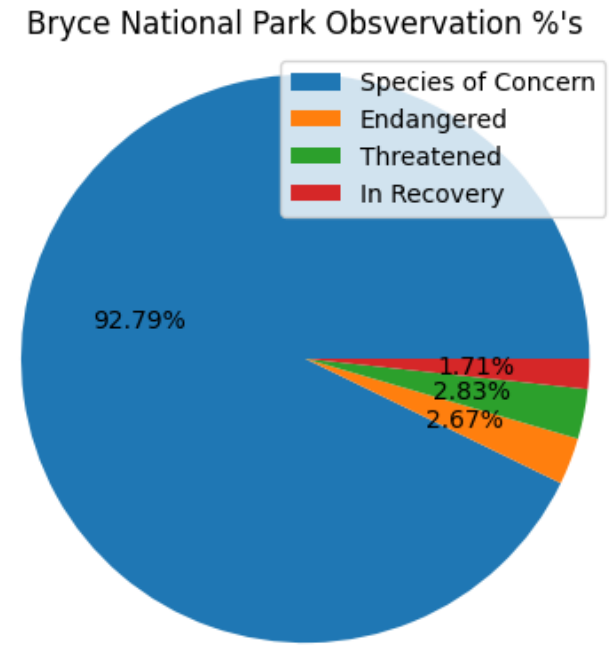
Wow, that's a lot, and it's hard to understand much about the other conservation statuses like this.

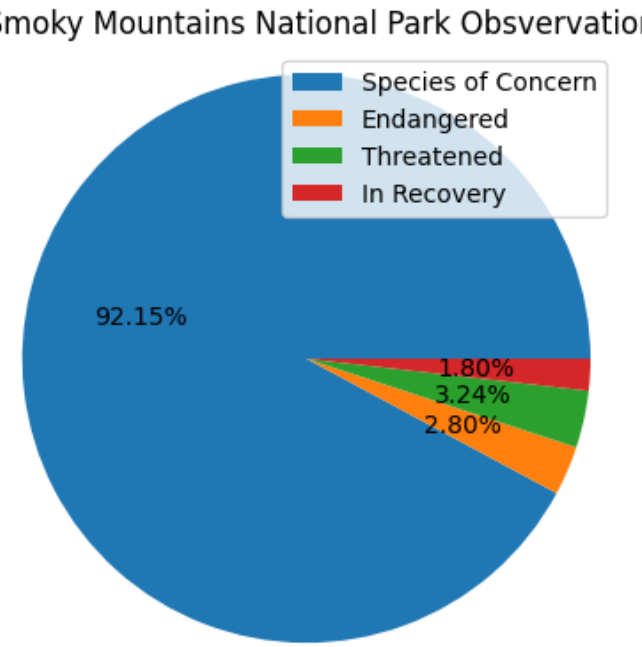
For this reason, a lot of the graphs will exclude 'Not at Risk' species.

With the ‘Not at Risk’ species removed, we can actually see the distribution of other conservation statuses in each park









Again, you can see that, among these ‘At Risk’ species, each park has an almost identical distribution

Question #3 Findings –

The greatest takeaway is that every park has a near equal percentage of sightings for each conservation status.

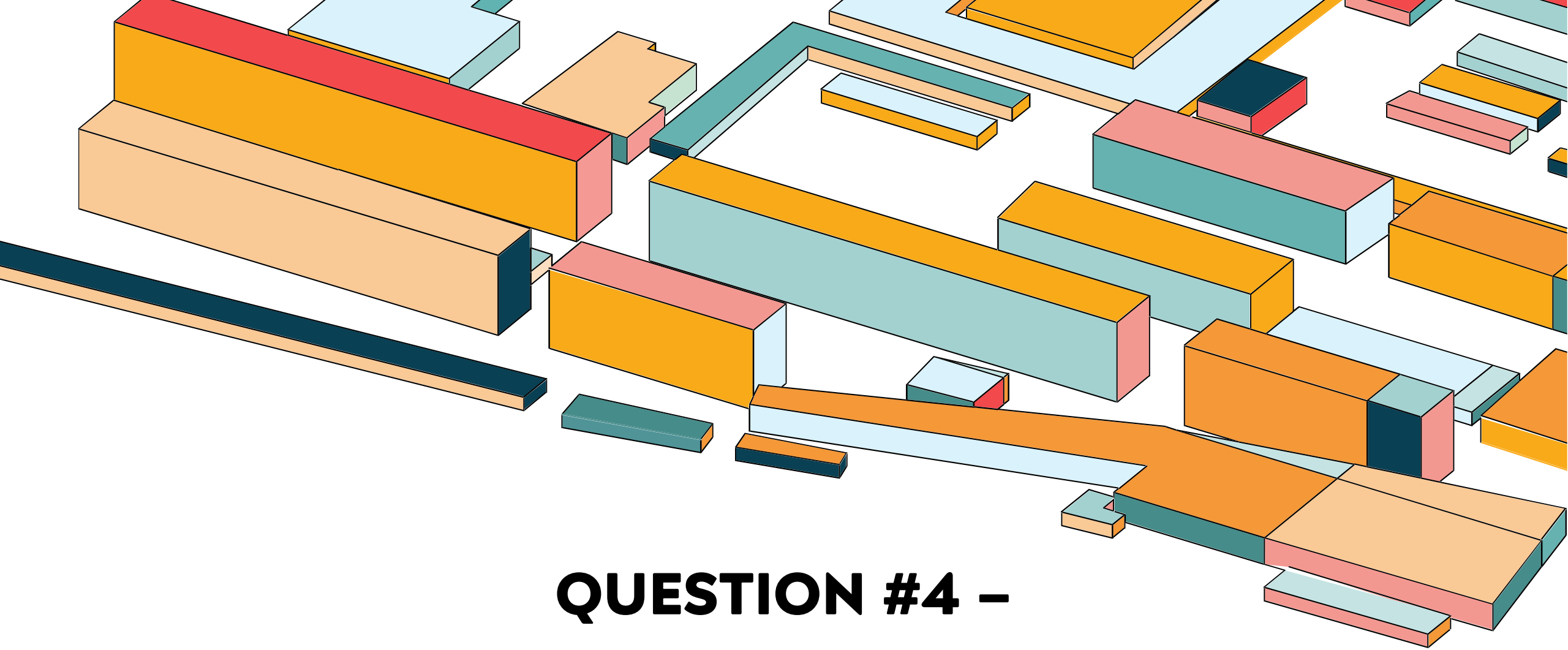
We know that the 'Not at Risk' species make up most of the sightings of species within each park at roughly 97%.

When you exclude the 'Not at Risk' species from the data, the species who are 'Species of Concern' (meaning they are at risk, but not significant risk), make up a large portion of sightings in each park at roughly 92%.

Amongst the 'At Risk' species, only 2% – 3% of those sightings are of 'Endangered' species.

Which leads to my next questions:

- What do the observations of 'Endangered' species look like?
- Less or more sightings on average? No correlation at all?



QUESTION #4 –

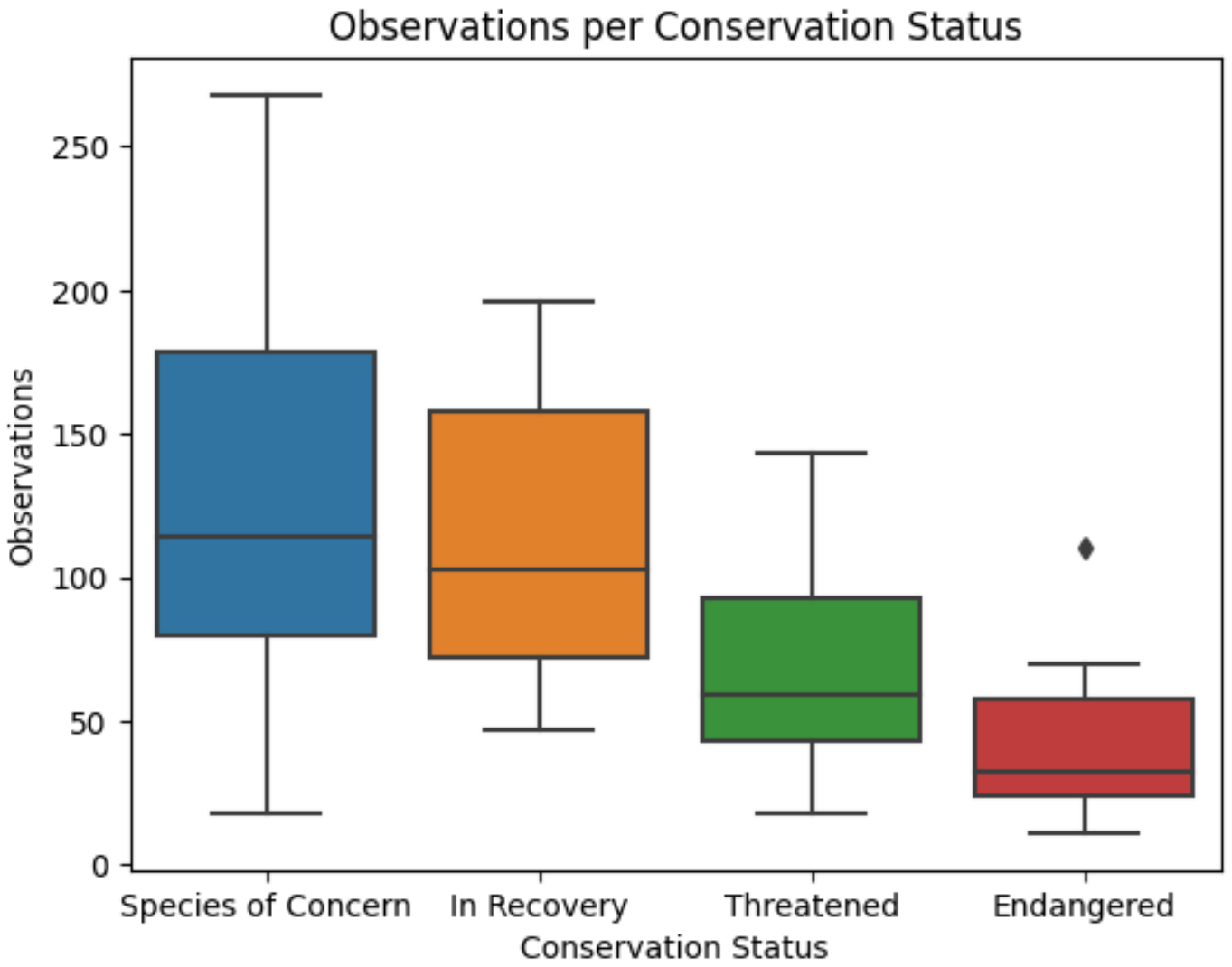
DO ‘ENDANGERED’ SPECIES HAVE LESS OBSERVATIONS ON AVERAGE?

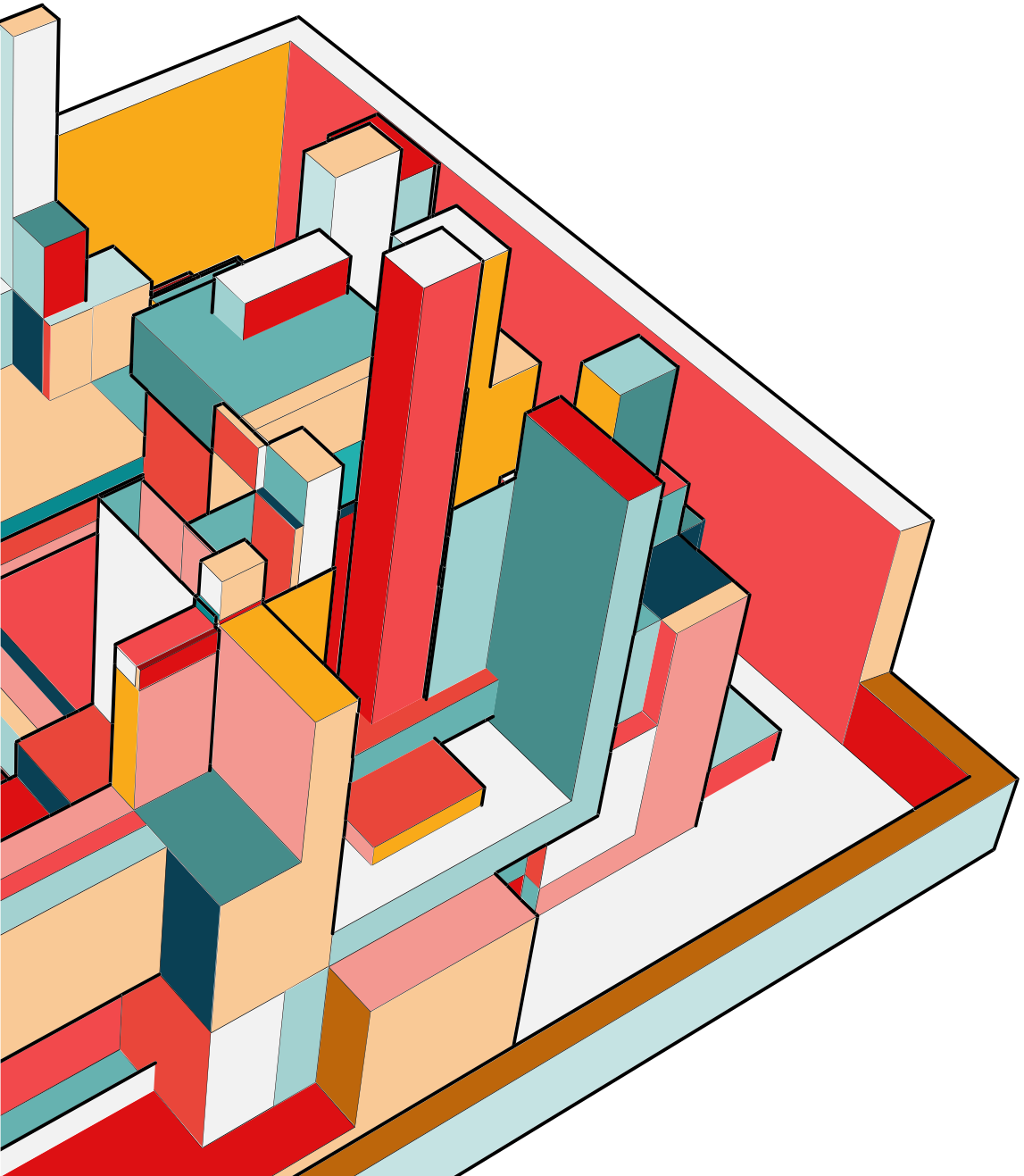
The graph shows all the observations for each species in each conservation status.

In each colored box there is a horizontal line, that line is the average number of observations for that conservation status.

We're looking at the line in each box to determine if there is a lower amount of observations on average for endangered species.

According to this graph, does this seem to be true?





Question #4 Findings –

According to the graph, it is clear that, on average, 'Endangered' species are observed less than any other conservation status.

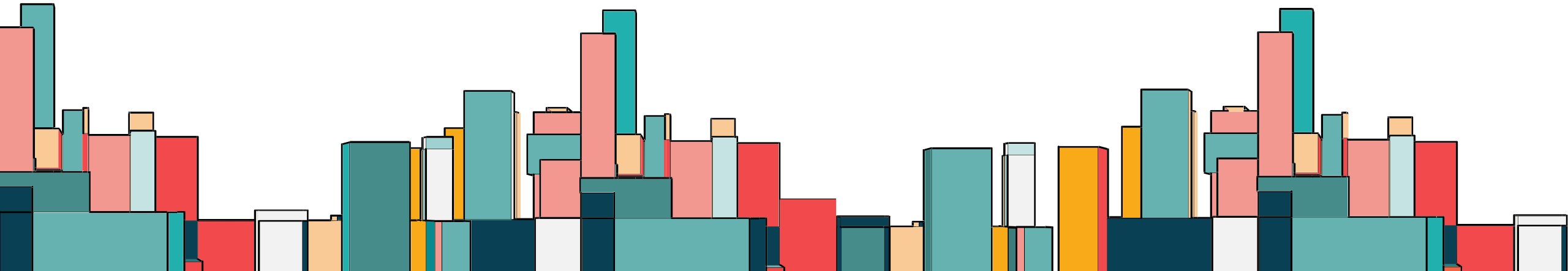
This makes sense because if the species is 'Endangered' then they should have a decreasing population when compared to other species who are not 'Endangered'.

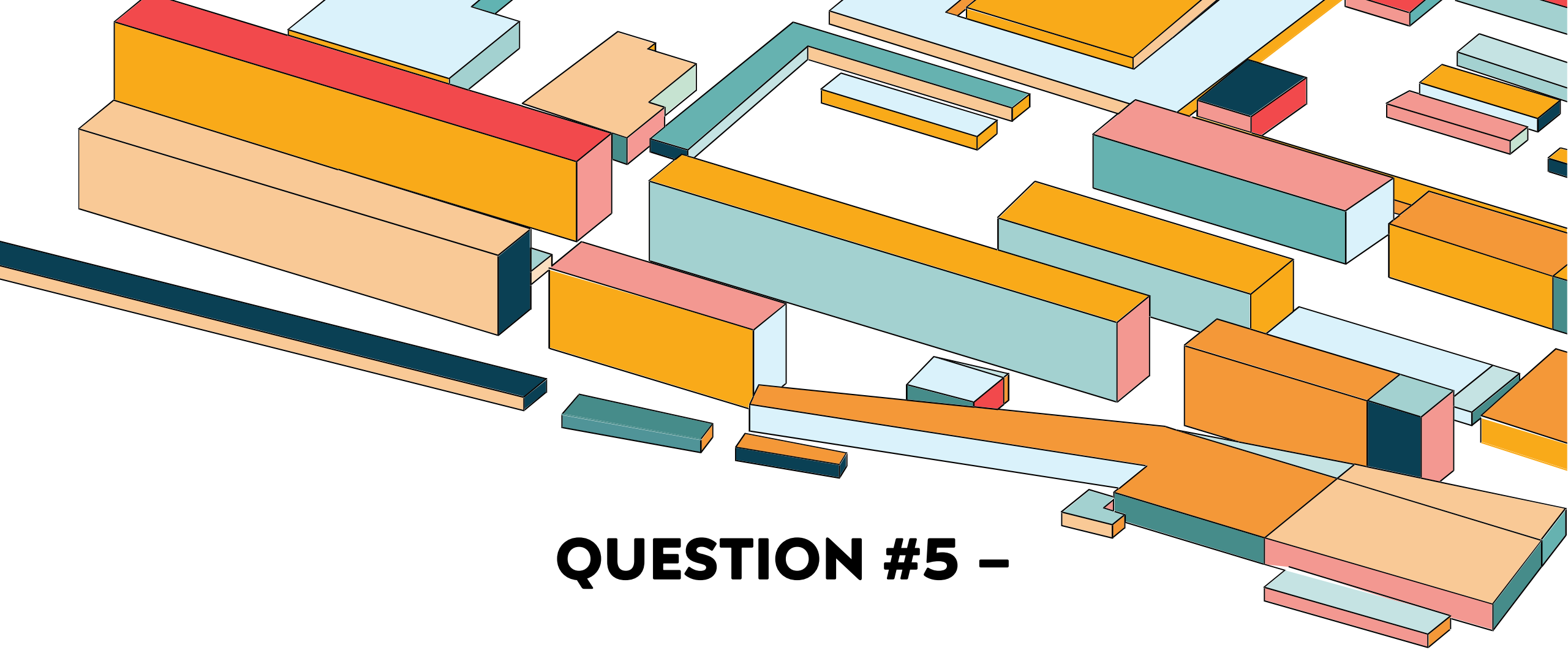
On average, as the severity of the conservation status increases for a specific species, the sightings of that specific species decreases

For example, if a species goes from being 'Threatened' to 'Endangered', then the severity of their conservation status has increased.

Moving on from comparing observations and parks, we need to investigate other areas of the data as well.

The focus of the following questions will be directed towards the different ***categories*** of species and the ***conservation statuses*** of those species.

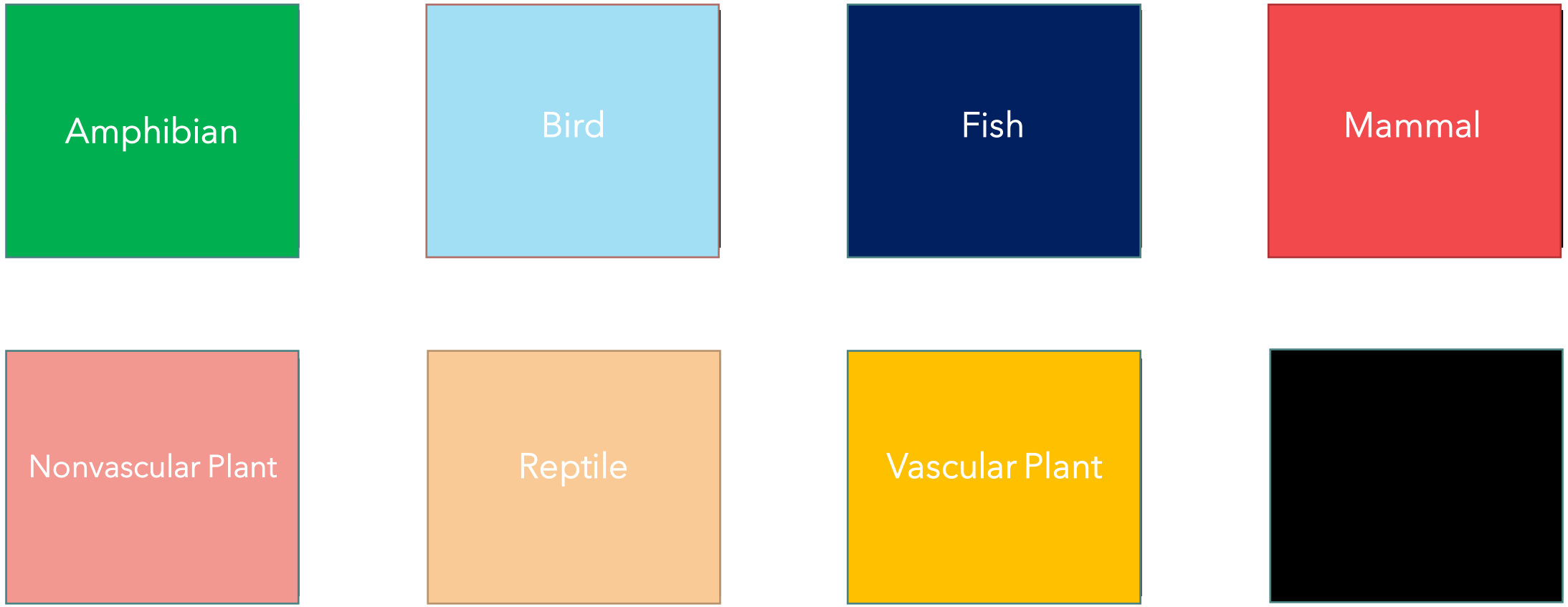


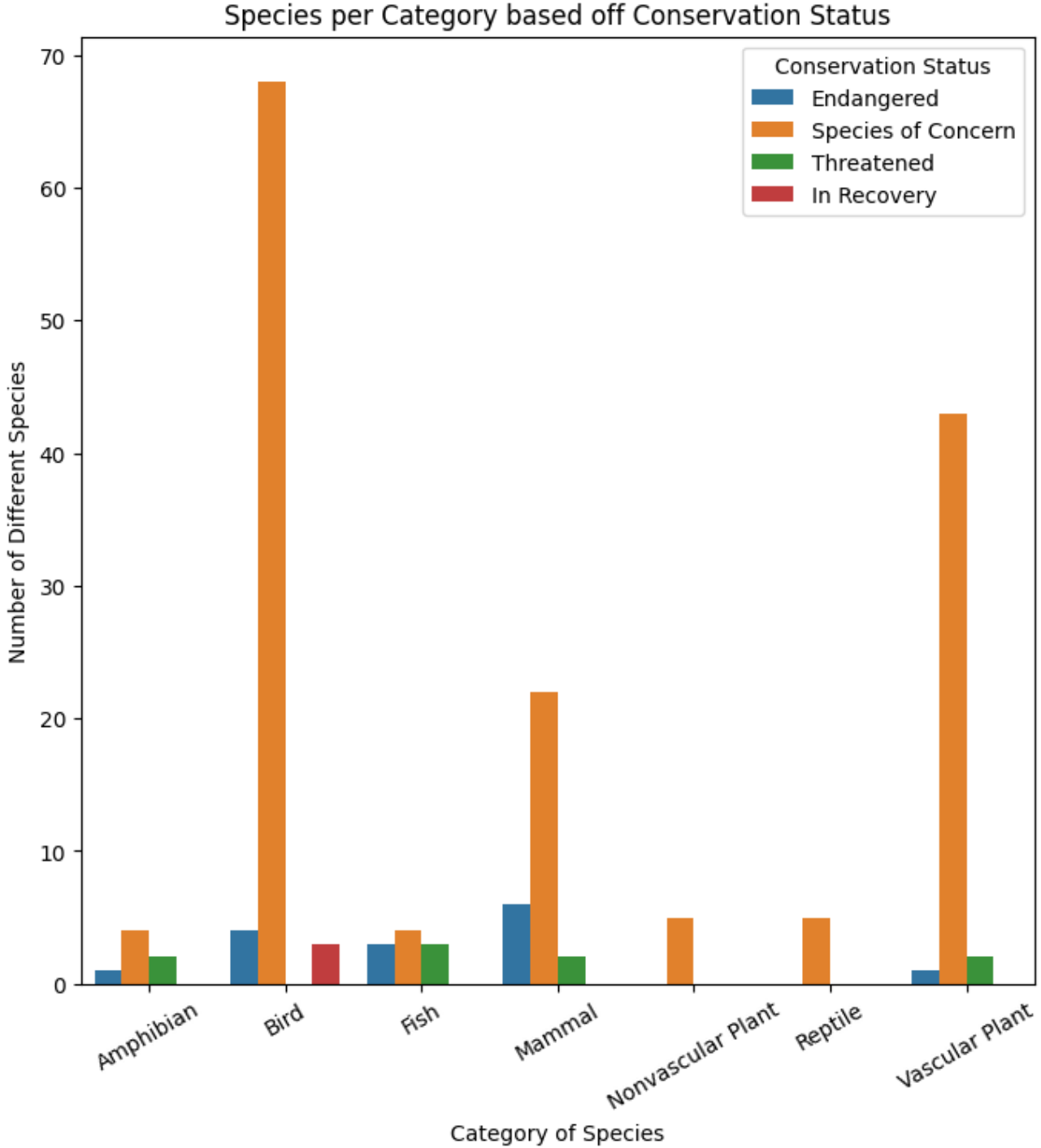


QUESTION #5 –

WHAT CAN WE LEARN ABOUT CATEGORIES BY LOOKING INTO THE NUMBER OF SPECIES IN EACH CONSERVATION STATUS?

CATEGORIES OF SPECIES



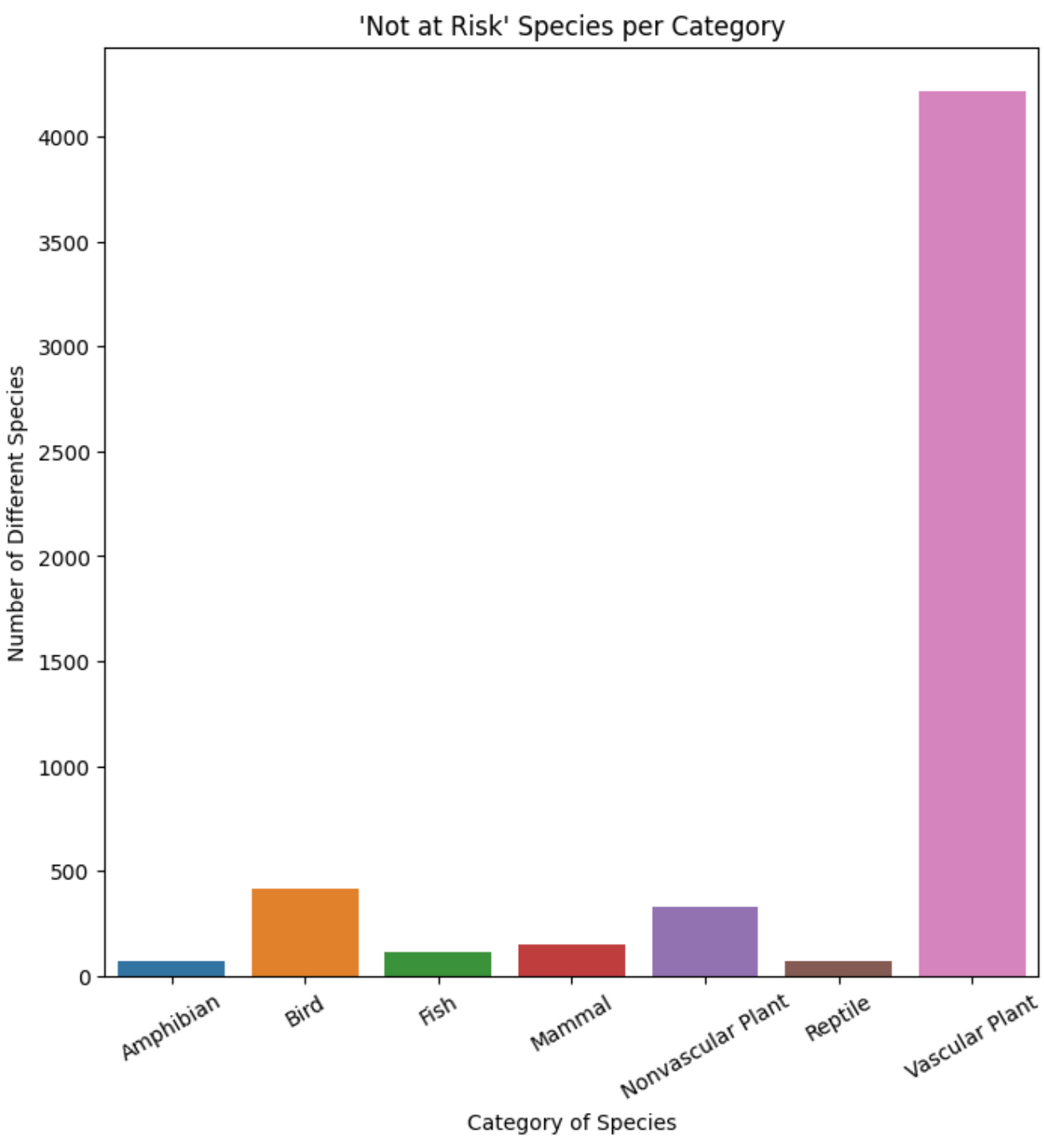

Species per Category based off Conservation Status
Conservation Status
Endangered
Species of Concern
Threatened
In Recovery
Number of Different Species
0
10
20
30
40
50
60
70
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category of Species

You can clearly see some major differences between the categories of species who are ‘At Risk’!

If you were to rank the categories from greatest to least by how much attention each category needs, then the ranking would be as follows...

1. *Bird (Huge concern for nearly the entire category of species)*
2. *Vascular Plant (Huge concern for nearly the entire category of species)*
3. *Mammal (Has the most endangered species but far less species who are of concern)*
4. *Fish (Moderate concern)*
5. *Amphibian (Moderate concern)*
6. *Nonvascular Plant & Reptile (No Reptiles or Nonvascular Plants are endangered and very few are of concern)*

Let's see if there is an inverse relationship when you plot the number of species which are not at risk.


'Not at Risk' Species per Category
Number of Different Species
0
500
1000
1500
2000
2500
3000
3500
4000
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category of Species

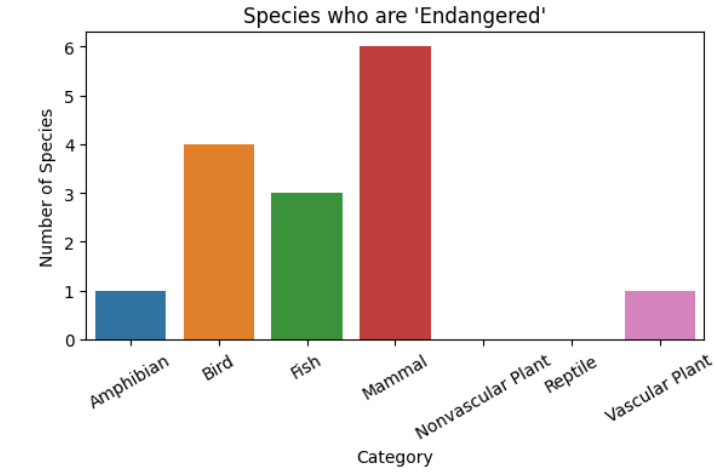
The plot doesn't show us too much, but it does show us that there is a very large amount of Vascular Plant species.

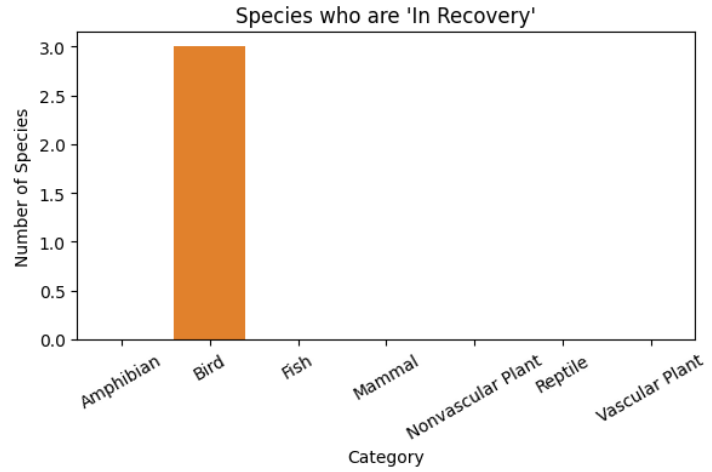
You could spend some time making assumptions and thinking about totals and percentages based off of these two graphs, but a better idea would be to make some more tables!

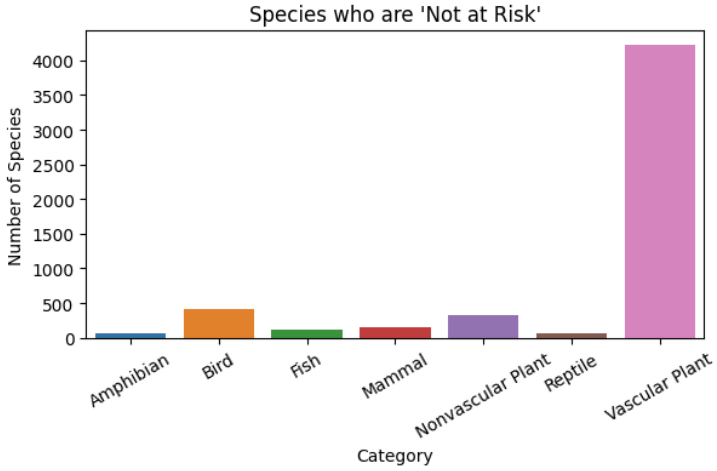
We will learn more about the data if we can visualize the percentages of the total number of species within each category.

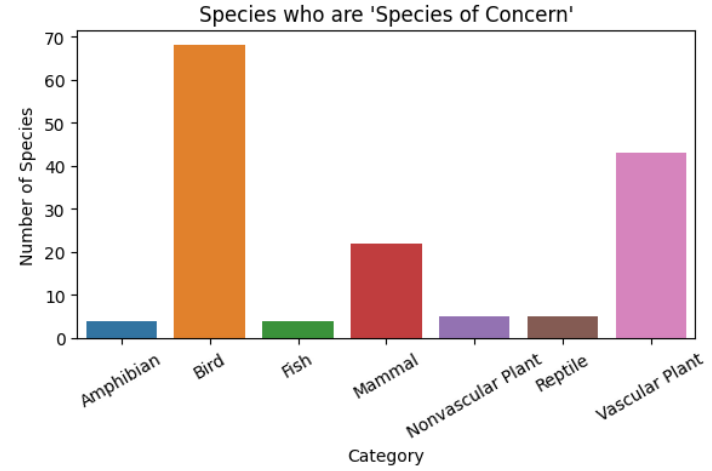
	Conservation Status	Endangered	In Recovery	Not at Risk	Species of Concern	Threatened	Total Species
Category							
Amphibian		1.0	0.0	72.0	4.0	2.0	79.0
Bird		4.0	3.0	413.0	68.0	0.0	488.0
Fish		3.0	0.0	115.0	4.0	3.0	125.0
Mammal		6.0	0.0	146.0	22.0	2.0	176.0
Nonvascular Plant		0.0	0.0	328.0	5.0	0.0	333.0
Reptile		0.0	0.0	73.0	5.0	0.0	78.0
Vascular Plant		1.0	0.0	4216.0	43.0	2.0	4262.0

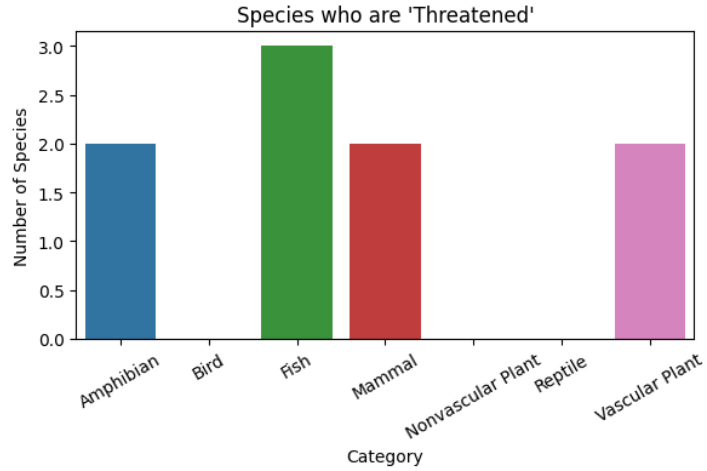
	Conservation Status	Endangered	In Recovery	Not at Risk	Species of Concern	Threatened
Category						
Amphibian		1.27%	0.0%	91.14%	5.06%	2.53%
Bird		0.82%	0.61%	84.63%	13.93%	0.0%
Fish		2.4%	0.0%	92.0%	3.2%	2.4%
Mammal		3.41%	0.0%	82.95%	12.5%	1.14%
Nonvascular Plant		0.0%	0.0%	98.5%	1.5%	0.0%
Reptile		0.0%	0.0%	93.59%	6.41%	0.0%
Vascular Plant		0.02%	0.0%	98.92%	1.01%	0.05%

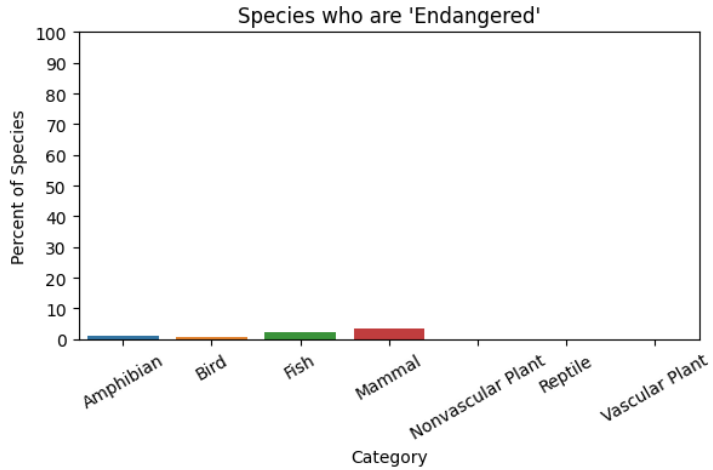

Species who are 'Endangered'
Number of Species
0
1
2
3
4
5
6
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

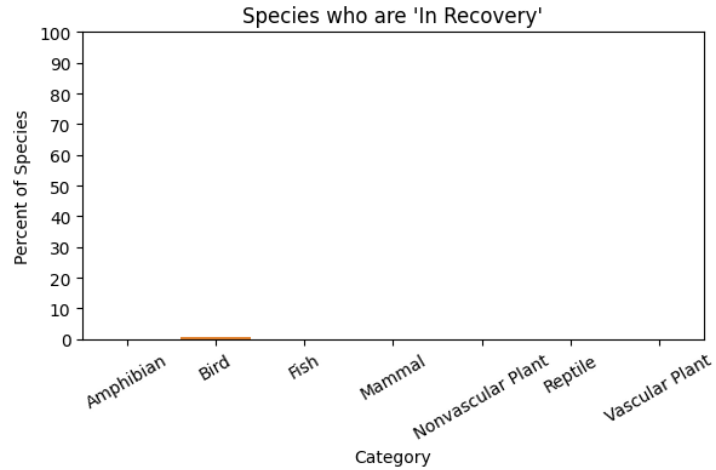

Species who are 'In Recovery'
Number of Species
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

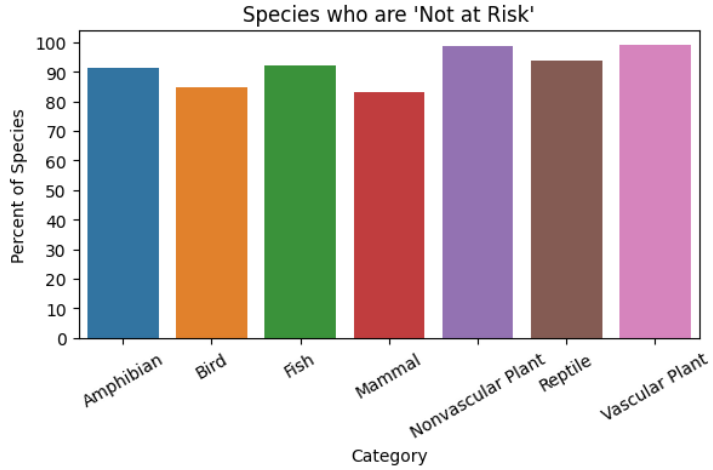

Species who are 'Not at Risk'
Number of Species
0
500
1000
1500
2000
2500
3000
3500
4000
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

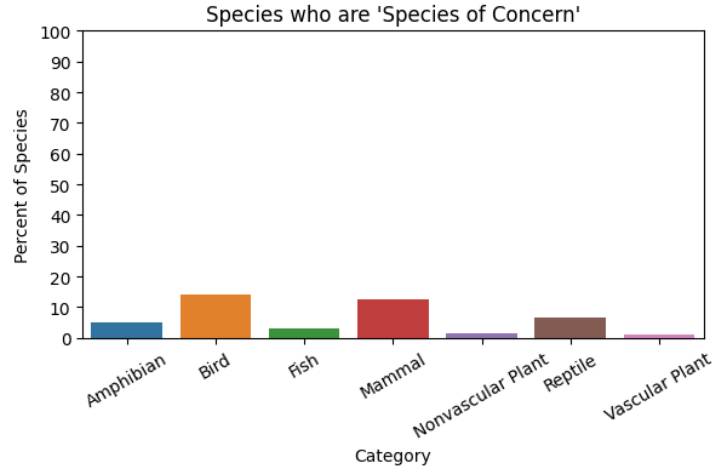

Species who are 'Species of Concern'
Number of Species
0
10
20
30
40
50
60
70
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

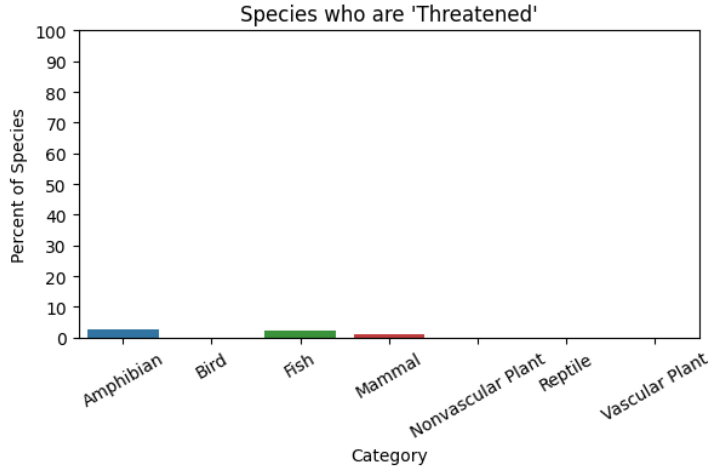

Species who are 'Threatened'
Number of Species
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

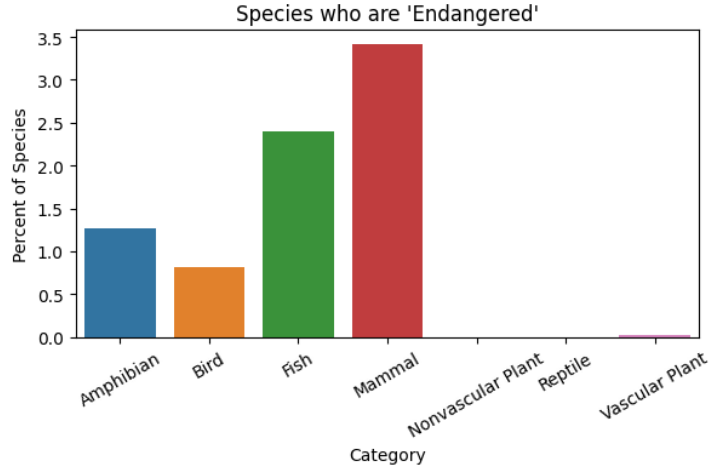

Species who are 'Endangered'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

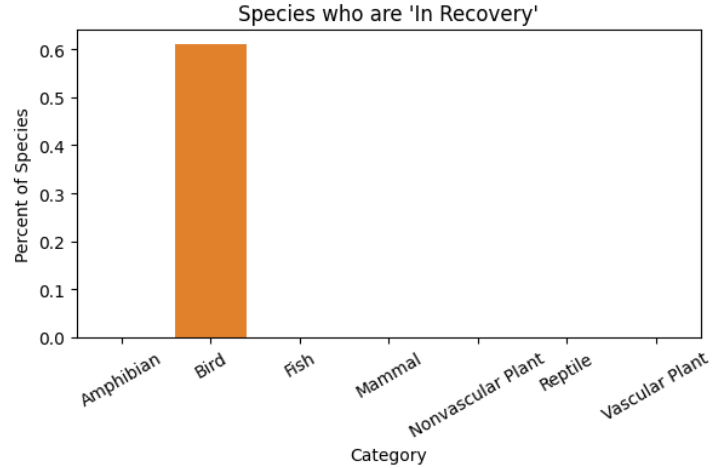

Species who are 'In Recovery'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

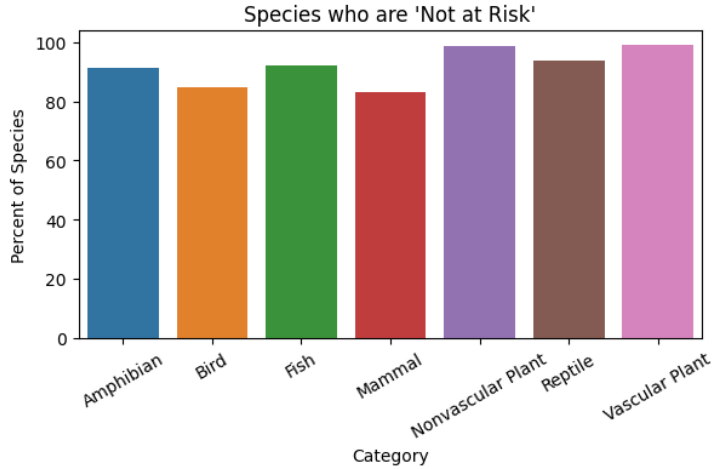

Species who are 'Not at Risk'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

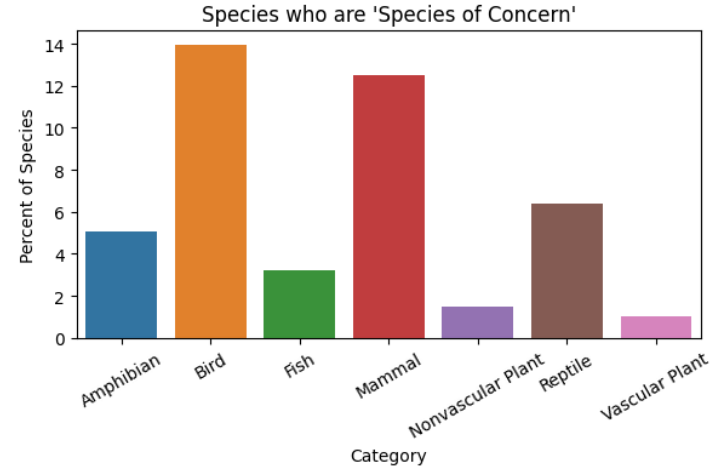

Species who are 'Species of Concern'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

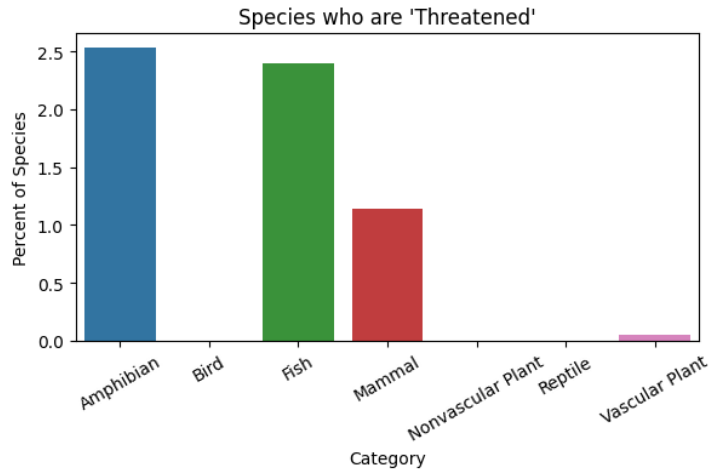

Species who are 'Threatened'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

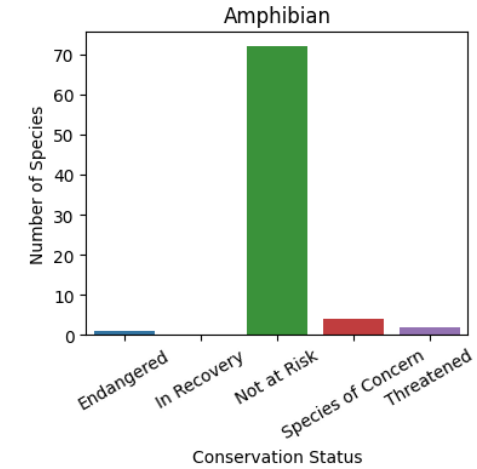

Species who are 'Endangered'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

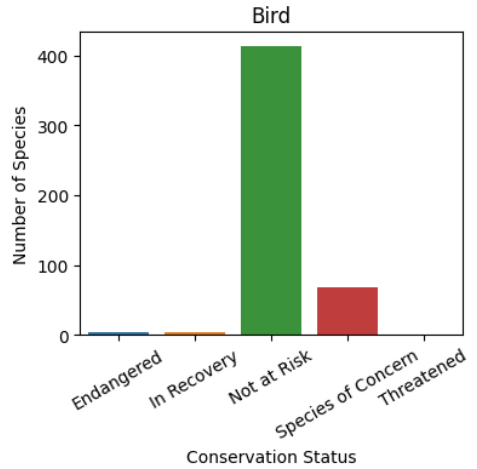

Species who are 'In Recovery'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

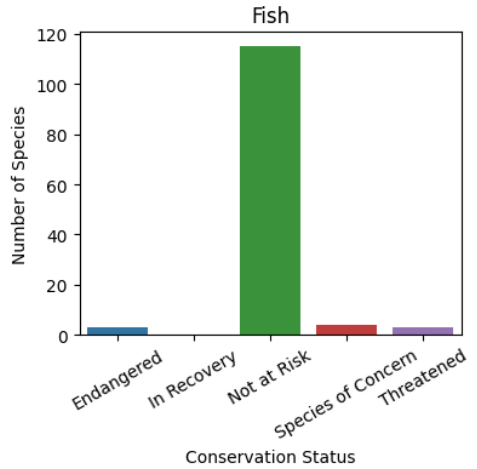

Species who are 'Not at Risk'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

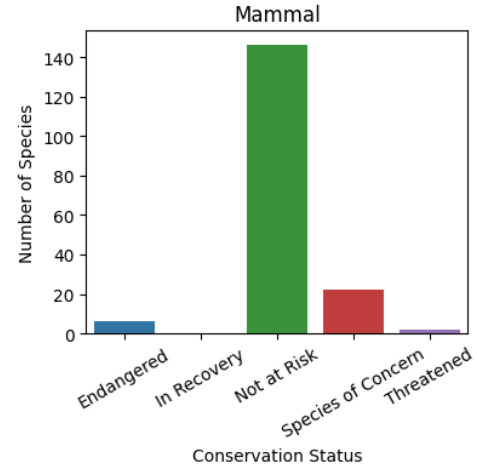

Species who are 'Species of Concern'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

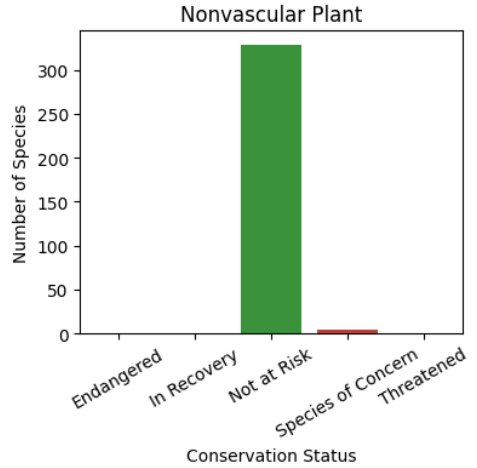

Species who are 'Threatened'
Percent of Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

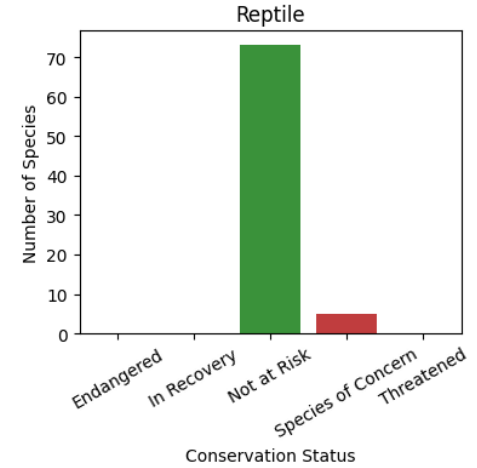

Amphibian
Number of Species
0
10
20
30
40
50
60
70
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

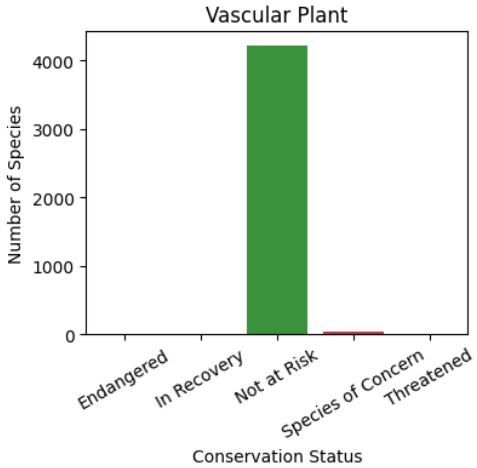

Bird
Number of Species
0
100
200
300
400
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

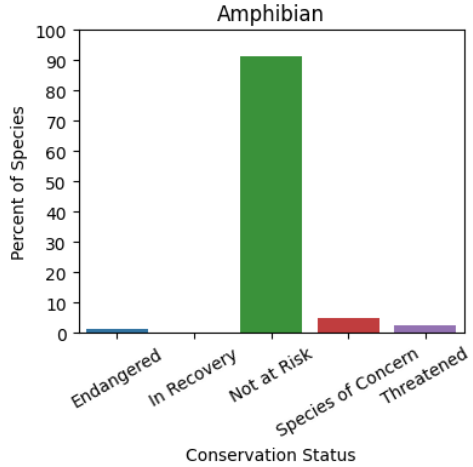

Fish
Number of Species
0
20
40
60
80
100
120
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

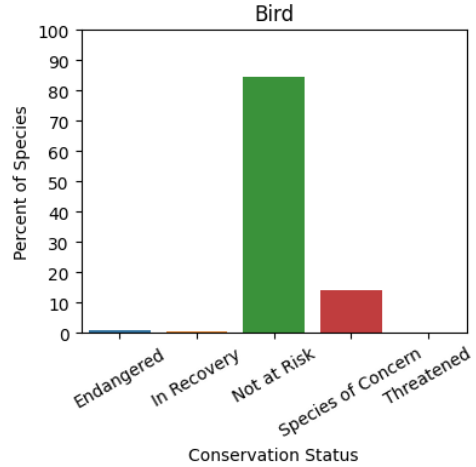

Mammal
Number of Species
0
20
40
60
80
100
120
140
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

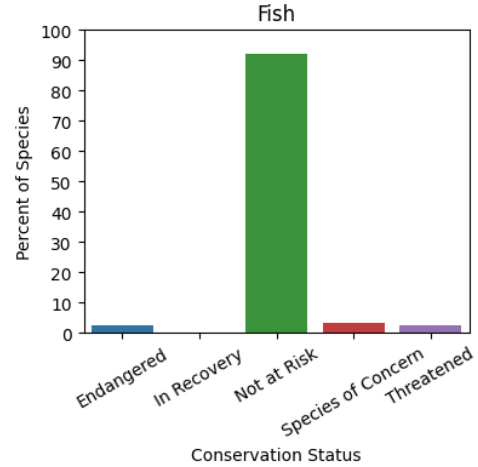

Nonvascular Plant
Number of Species
0
50
100
150
200
250
300
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

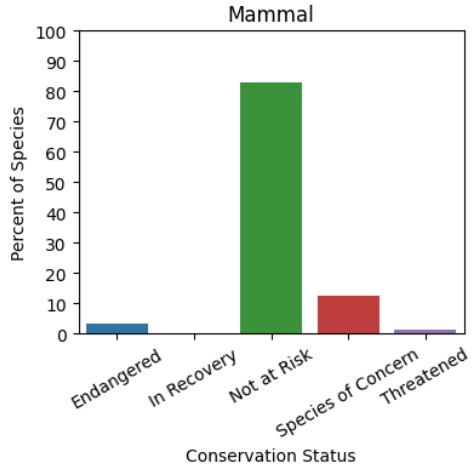

Reptile
Number of Species
0
10
20
30
40
50
60
70
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

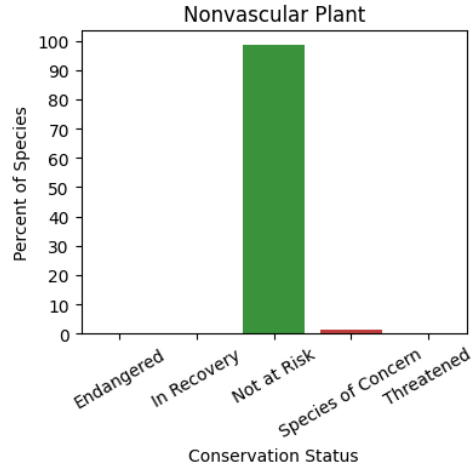

Vascular Plant
Number of Species
0
1000
2000
3000
4000
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

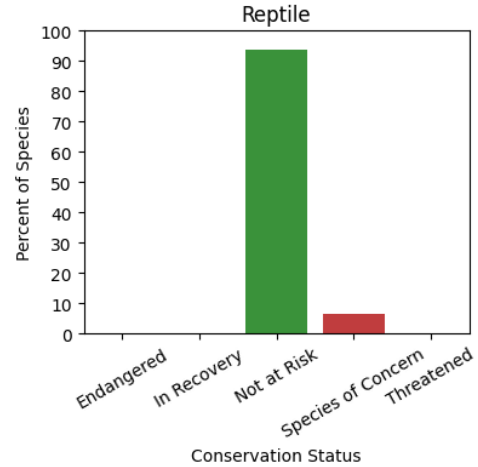

Amphibian
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

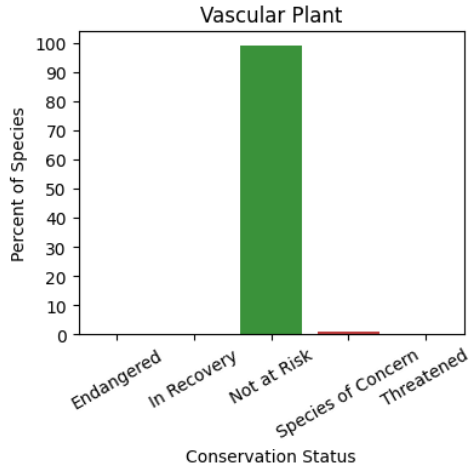

Bird
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status


Fish
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status


Mammal
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status


Nonvascular Plant
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status


Reptile
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status


Vascular Plant
Percent of Species
Endangered
In Recovery
Not at Risk
Species of Concern
Threatened
Conservation Status

Question #5 Findings –

When comparing all conservation statuses (Includes 'Not at Risk')…

- *Mammals are the most 'At Risk' category of species overall*
- *Both Plant categories are the least 'At Risk' categories of species overall*

- *Mammals and Fish are the most 'Endangered' categories with mammals being the most 'Endangered'*
- *Birds and Mammals are the categories of species with the most concern with birds having the highest concern*

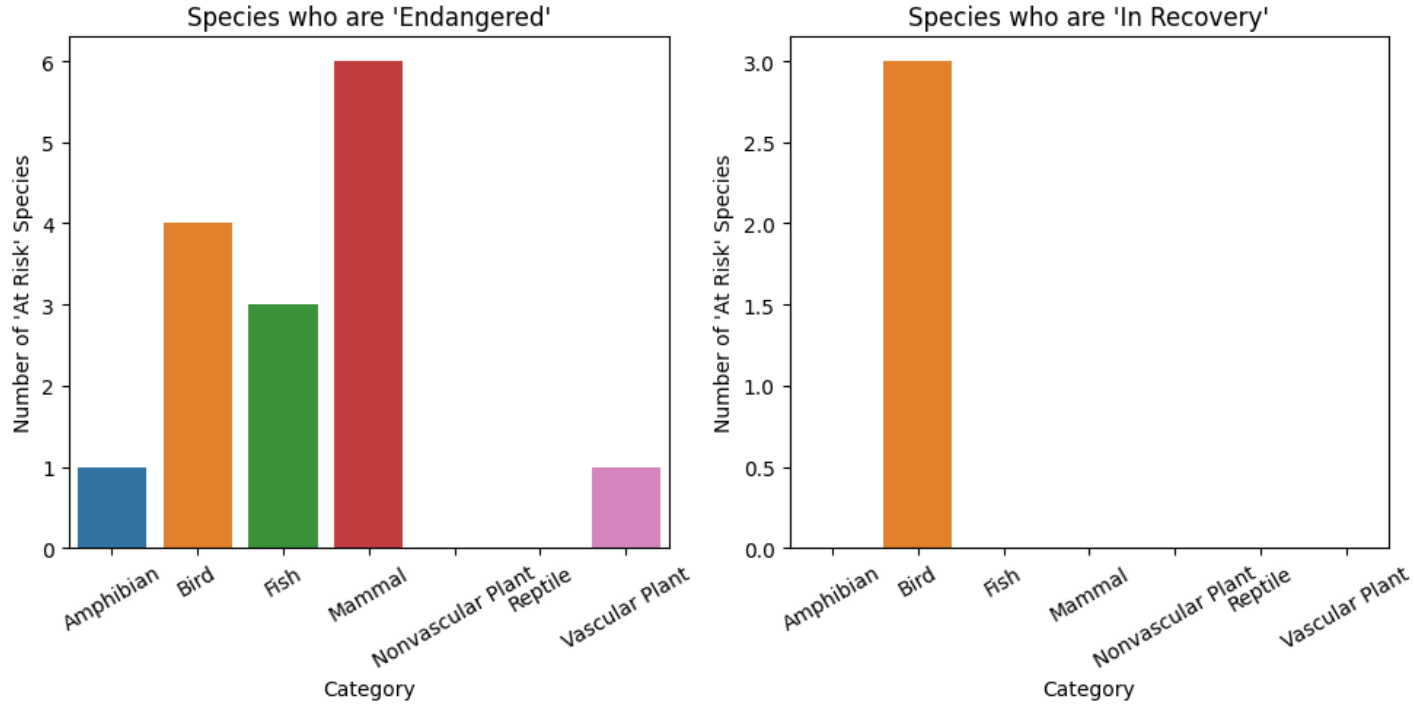
- *Reptiles have no 'Endangered' species but they do have the third highest 'Species of Concern' status*
- *Nonvascular Plants have no 'Endangered' species*

QUESTION #6 –

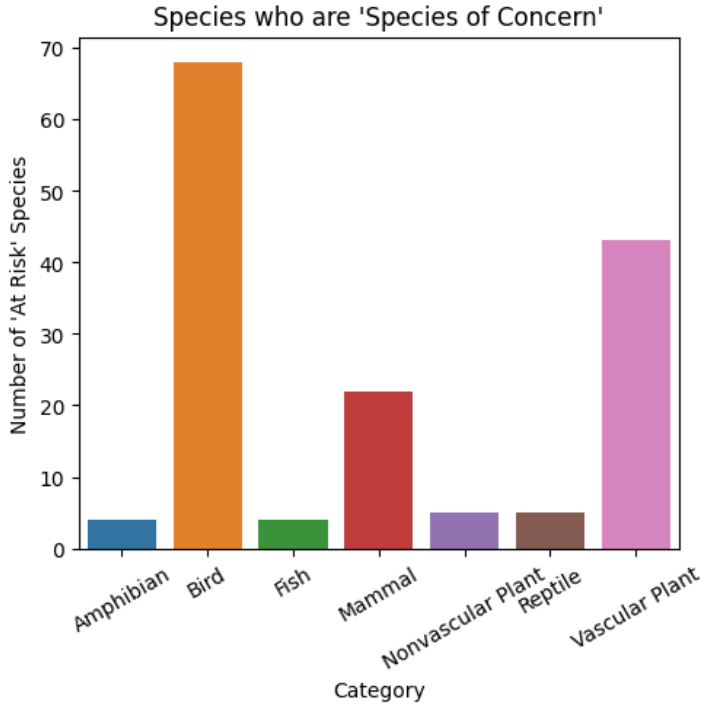
WHAT HAPPENS WHEN WE EXCLUDE THE ‘NOT AT RISK’ SPECIES WHILE ANALYZING CATEGORIES AND CONSERVATION STATUSES?

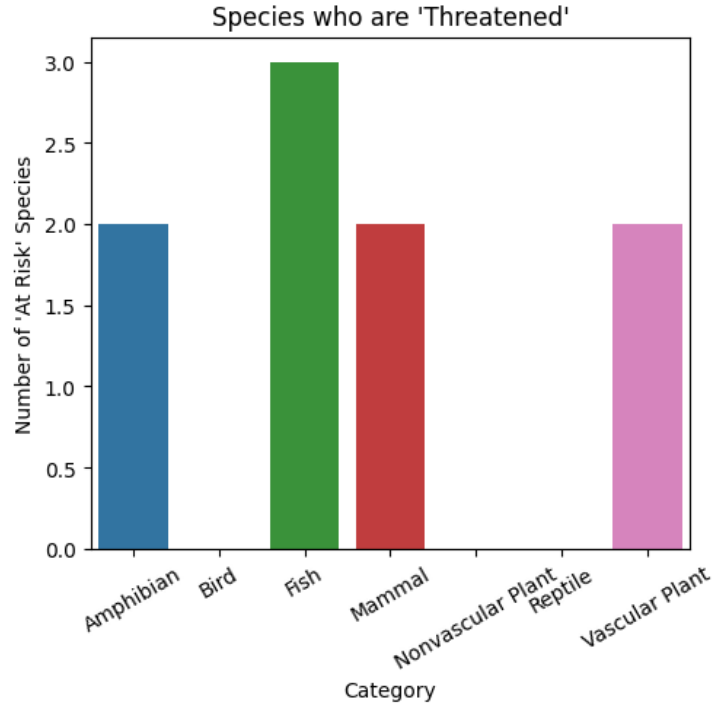
	Conservation Status	Endangered	In Recovery	Species of Concern	Threatened	Total Species
Category						
Amphibian		1.0	0.0	4.0	2.0	7.0
Bird		4.0	3.0	68.0	0.0	75.0
Fish		3.0	0.0	4.0	3.0	10.0
Mammal		6.0	0.0	22.0	2.0	30.0
Nonvascular Plant		0.0	0.0	5.0	0.0	5.0
Reptile		0.0	0.0	5.0	0.0	5.0
Vascular Plant		1.0	0.0	43.0	2.0	46.0

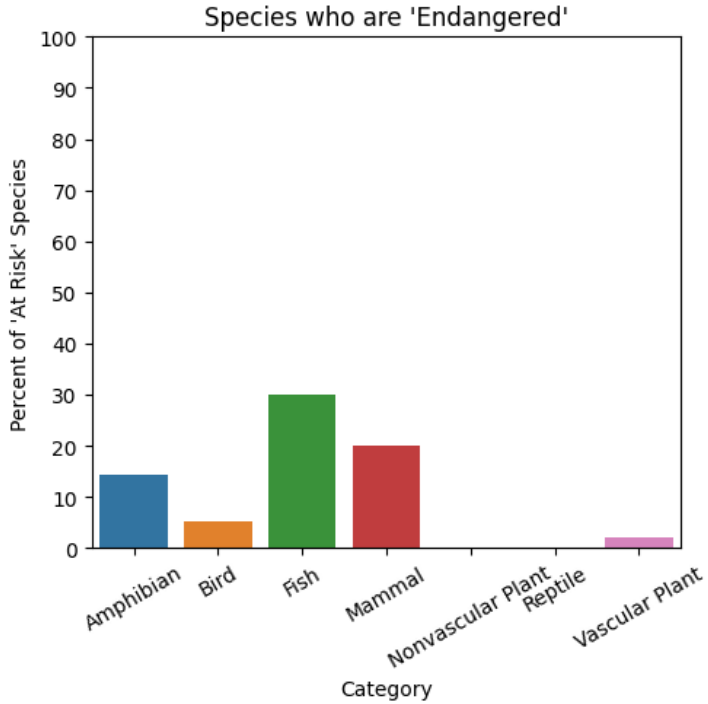
	Conservation Status	Endangered	In Recovery	Species of Concern	Threatened
Category					
Amphibian		14.29%	0.0%	57.14%	28.57%
Bird		5.33%	4.0%	90.67%	0.0%
Fish		30.0%	0.0%	40.0%	30.0%
Mammal		20.0%	0.0%	73.33%	6.67%
Nonvascular Plant		0.0%	0.0%	100.0%	0.0%
Reptile		0.0%	0.0%	100.0%	0.0%
Vascular Plant		2.17%	0.0%	93.48%	4.35%

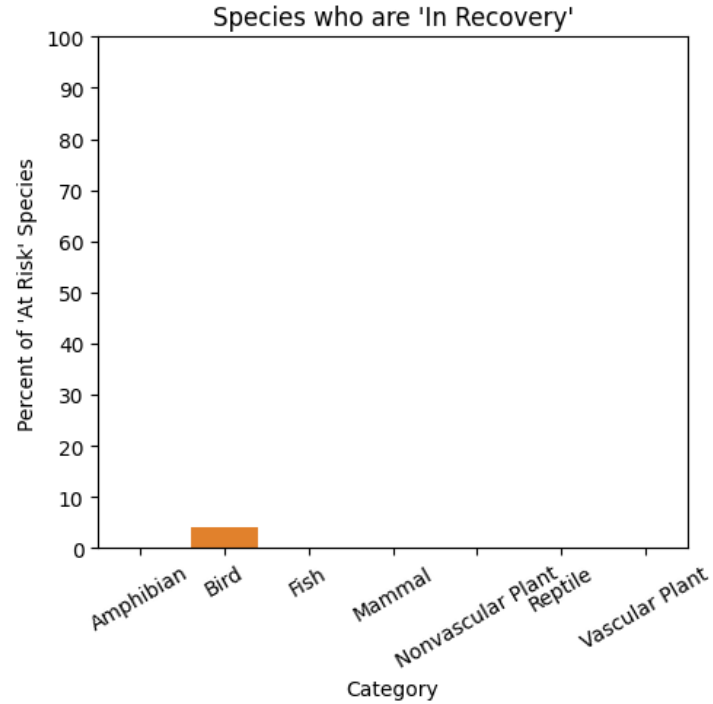

Species who are 'Endangered'
Number of 'At Risk' Species
0
1
2
3
4
5
6
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

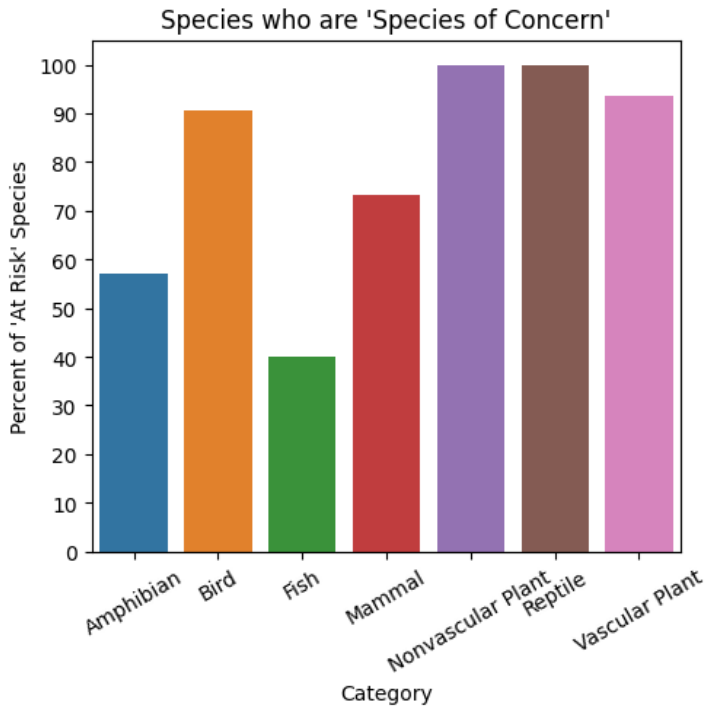
Species who are 'In Recovery'
Number of 'At Risk' Species
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

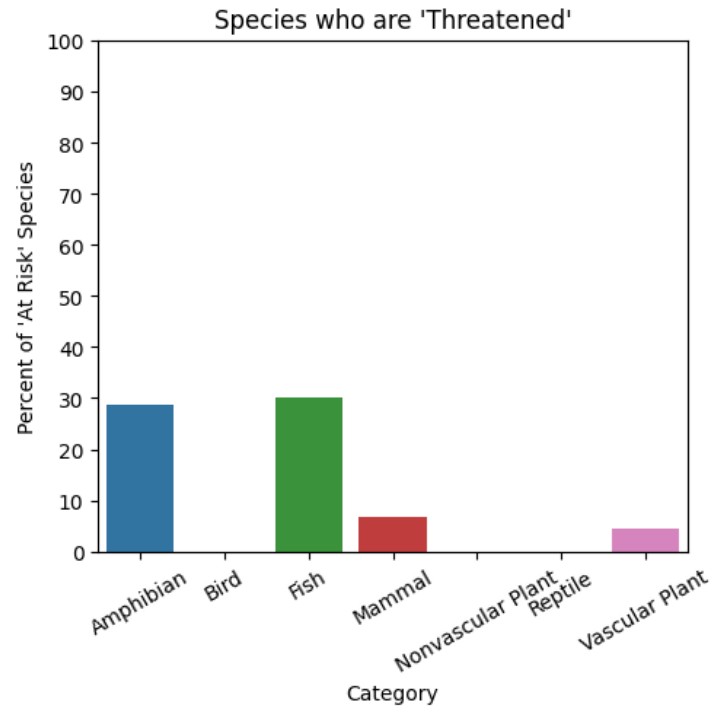

Species who are 'Species of Concern'
Number of 'At Risk' Species
0
10
20
30
40
50
60
70
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

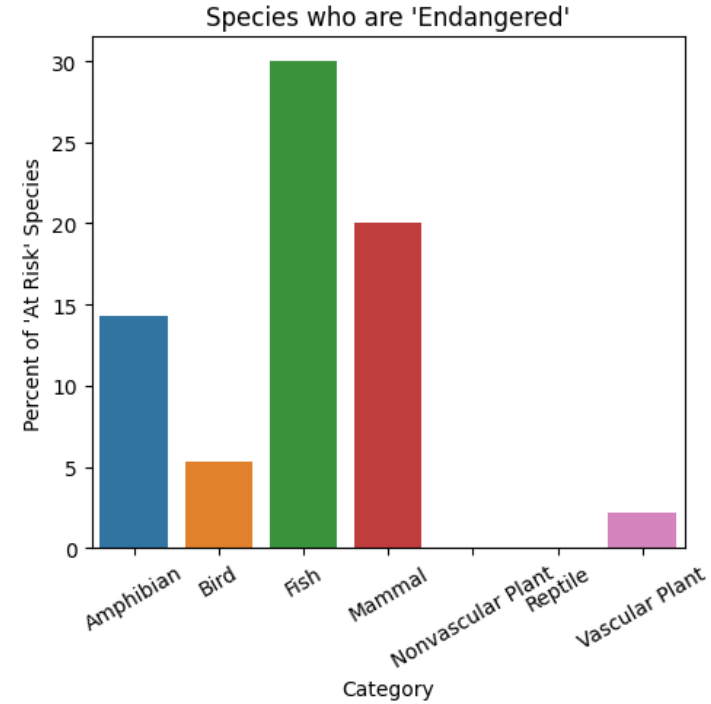

Species who are 'Threatened'
Number of 'At Risk' Species
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

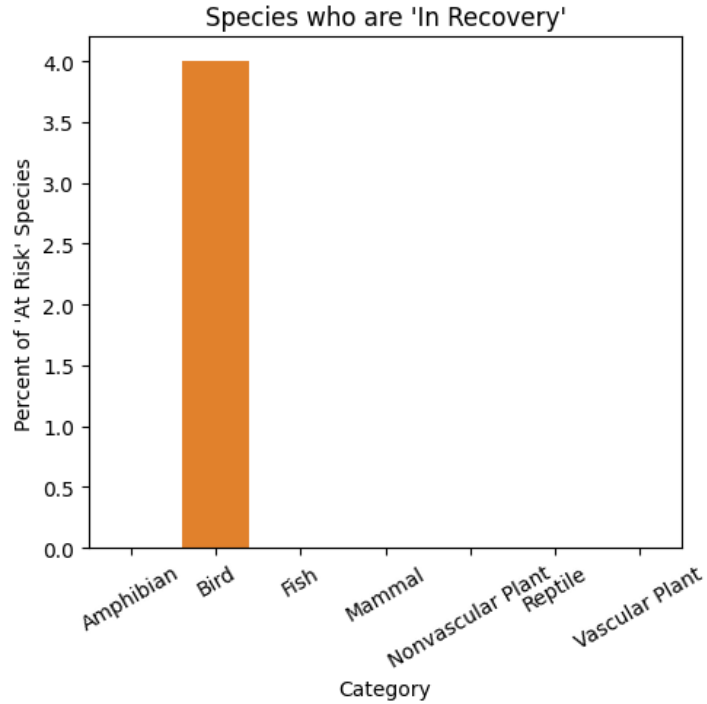

Species who are 'Endangered'
Percent of 'At Risk' Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

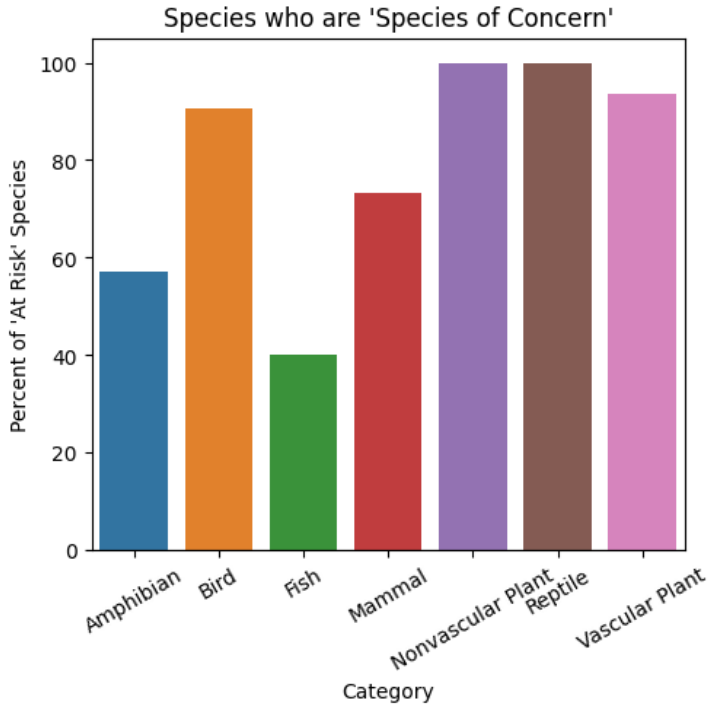

Species who are 'In Recovery'
Percent of 'At Risk' Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

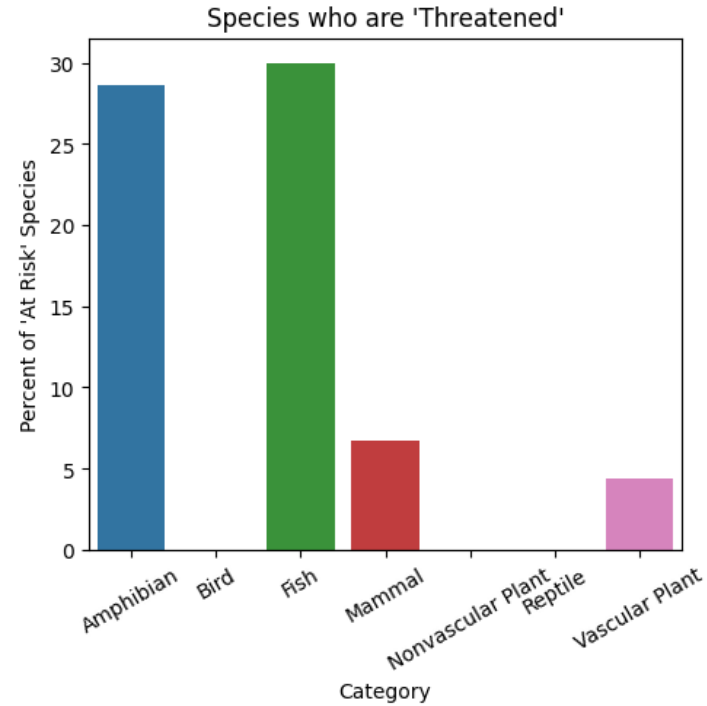

Species who are 'Species of Concern'
Percent of 'At Risk' Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

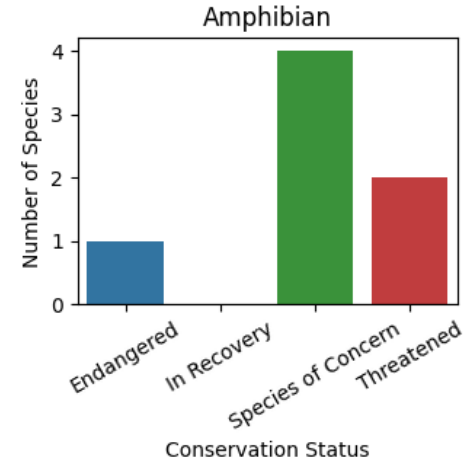

Species who are 'Threatened'
Percent of 'At Risk' Species
Category
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant

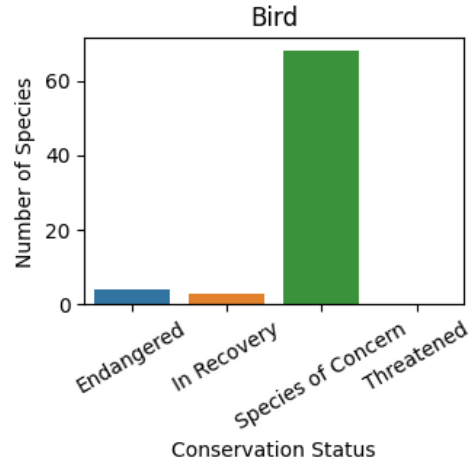

Species who are 'Endangered'
Percent of 'At Risk' Species
0
5
10
15
20
25
30
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

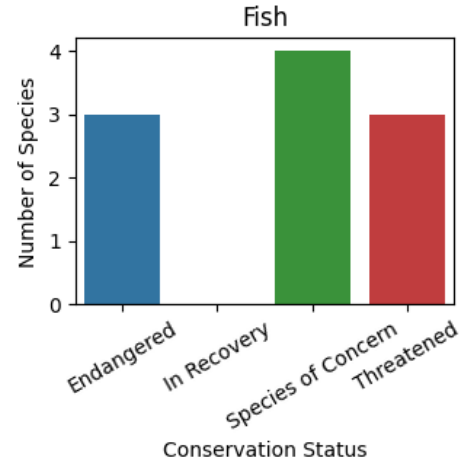

Species who are 'In Recovery'
Percent of 'At Risk' Species
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

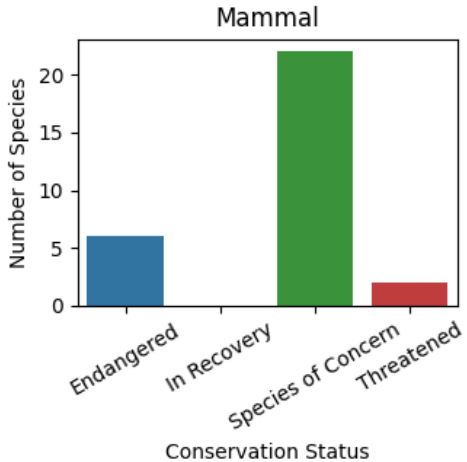

Species who are 'Species of Concern'
Percent of 'At Risk' Species
0
20
40
60
80
100
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

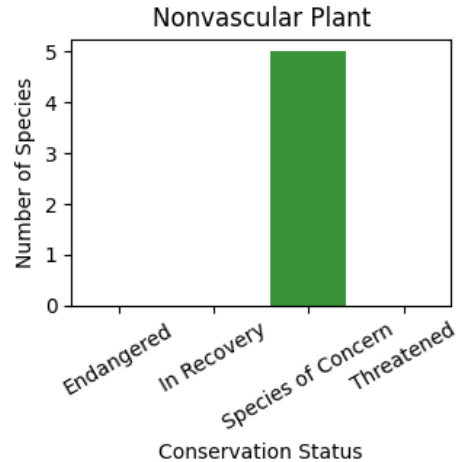

Species who are 'Threatened'
Percent of 'At Risk' Species
0
5
10
15
20
25
30
Amphibian
Bird
Fish
Mammal
Nonvascular Plant
Reptile
Vascular Plant
Category

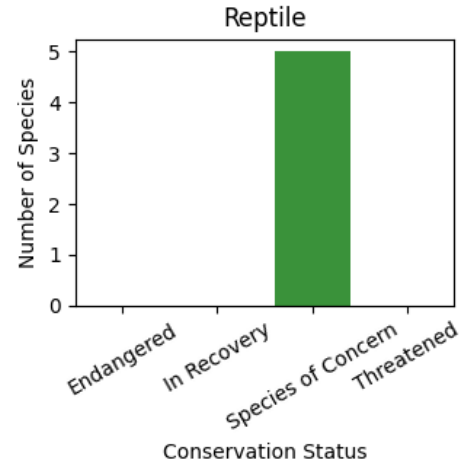

Amphibian
Number of Species
0
1
2
3
4
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

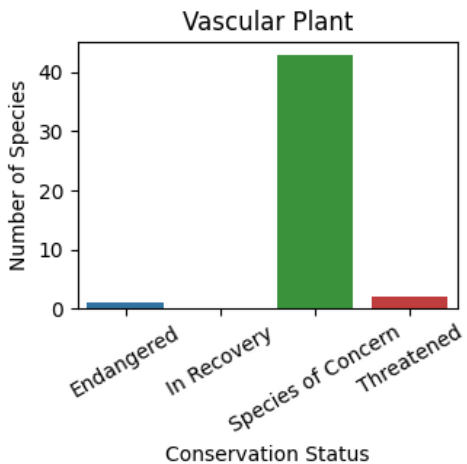

Bird
Number of Species
0
20
40
60
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

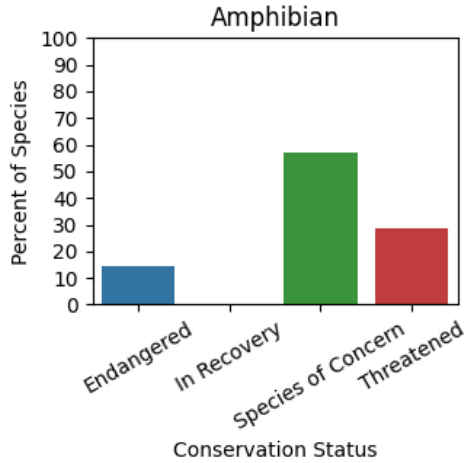

Fish
Number of Species
0
1
2
3
4
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

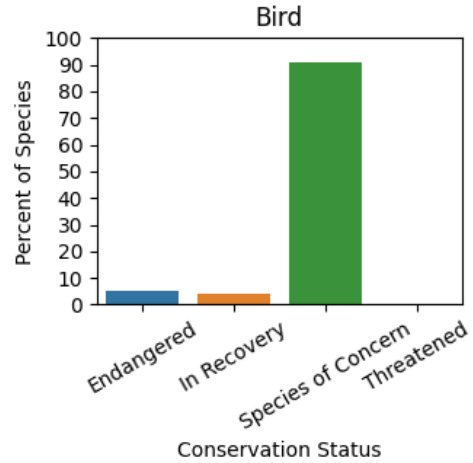

Mammal
Number of Species
0
5
10
15
20
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

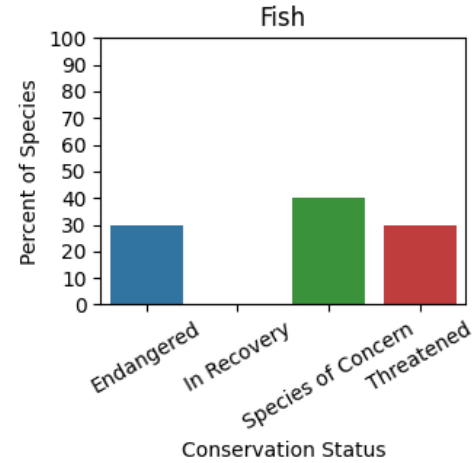

Nonvascular Plant
Number of Species
0
1
2
3
4
5
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

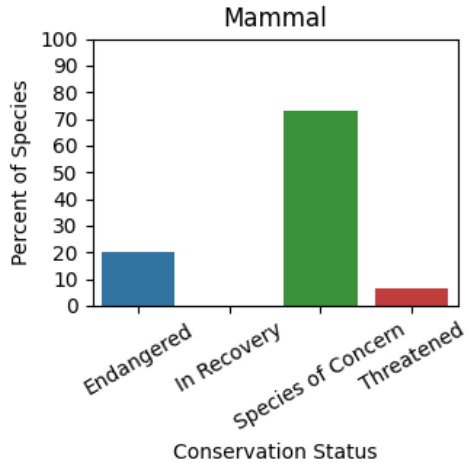

Reptile
Number of Species
0
1
2
3
4
5
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

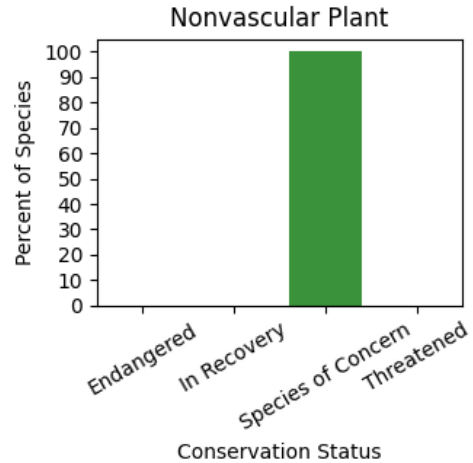

Vascular Plant
Number of Species
0
10
20
30
40
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

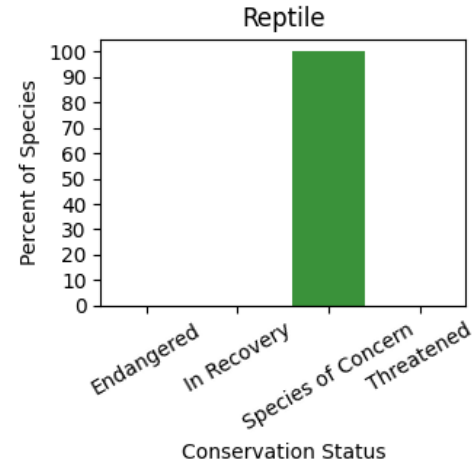

Amphibian
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

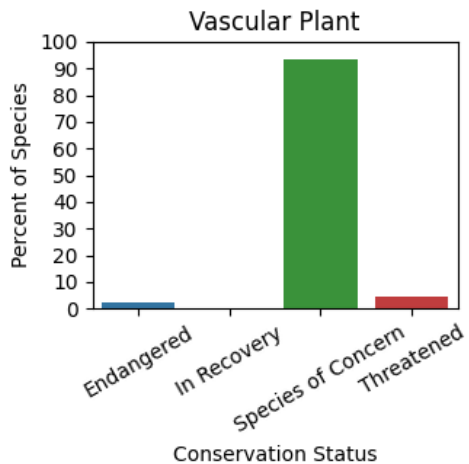

Bird
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status


Fish
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status


Mammal
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status


Nonvascular Plant
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status


Reptile
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status


Vascular Plant
Percent of Species
Endangered
In Recovery
Species of Concern
Threatened
Conservation Status

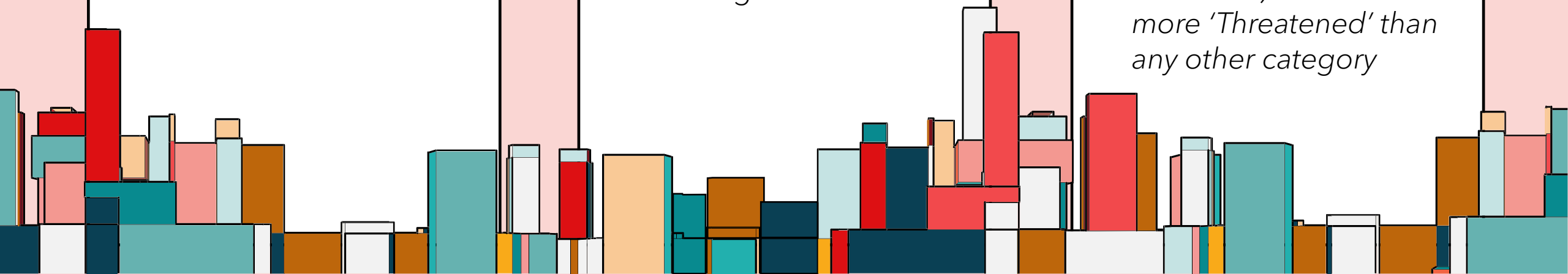
Question #6 Findings –

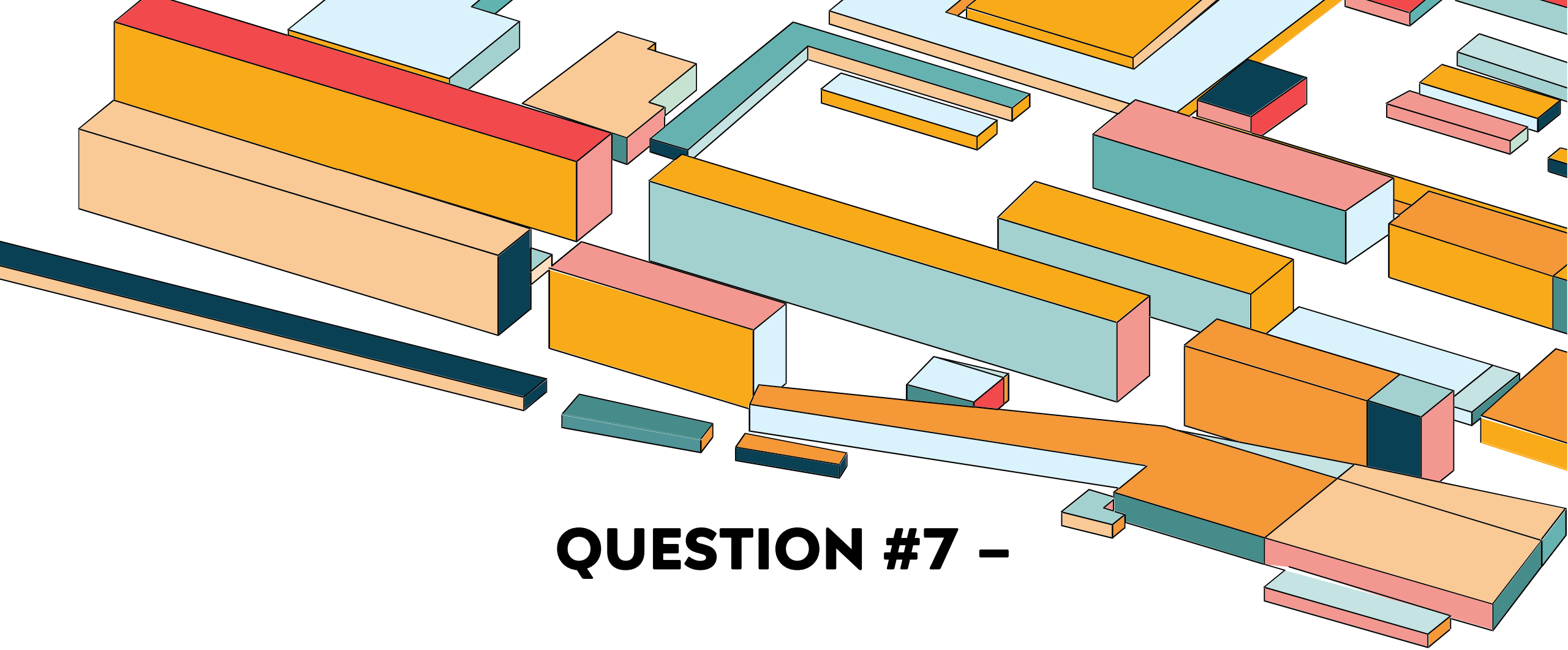
When comparing 'At Risk' conservation statuses (Excludes 'Not at Risk')…

- *Fish have the highest proportion of 'Endangered' species (30%) with Mammals being the second highest (20%)*

- *If preventative actions are taken, Birds, Reptiles, Vascular Plants and Nonvascular Plants have a much lower chance of moving towards 'Endangerment'*

- *Species that live primarily in the water (Amphibians and Fish) have high proportions of 'Endangerment' (Amphibians are 3rd highest behind Mammals) and are more 'Threatened' than any other category*



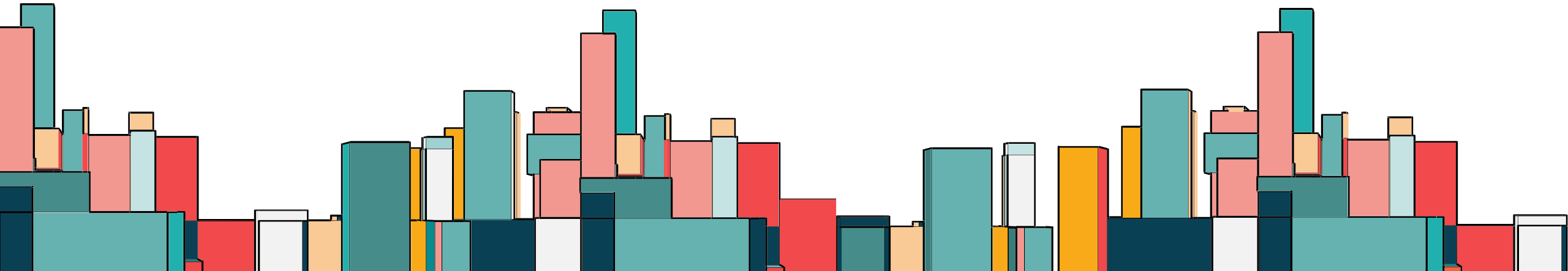


QUESTION #7 –

WHAT PERCENT DO THE OBSERVATIONS OF EACH CATEGORY MAKE UP WITHIN EACH PARK?

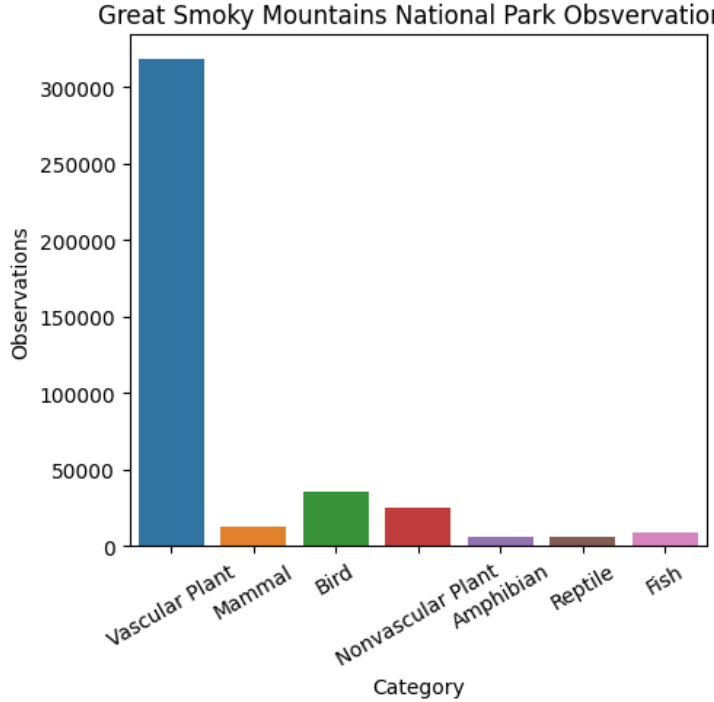
The focus of this section will be on the number of *observations* for each *category* of species in each *park*.

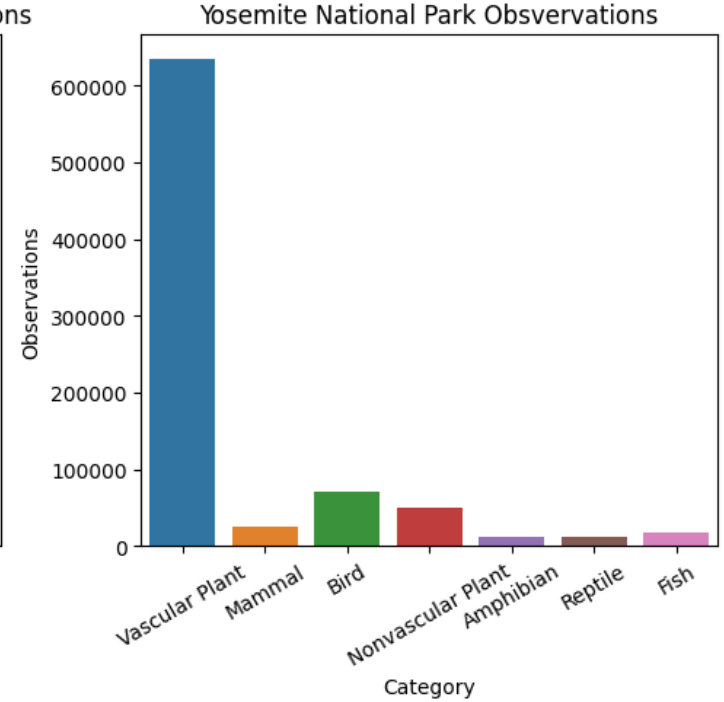
Very similar to what was analyzed earlier but now instead of looking at the observations for each conservation status, we'll examine them for each category of species.

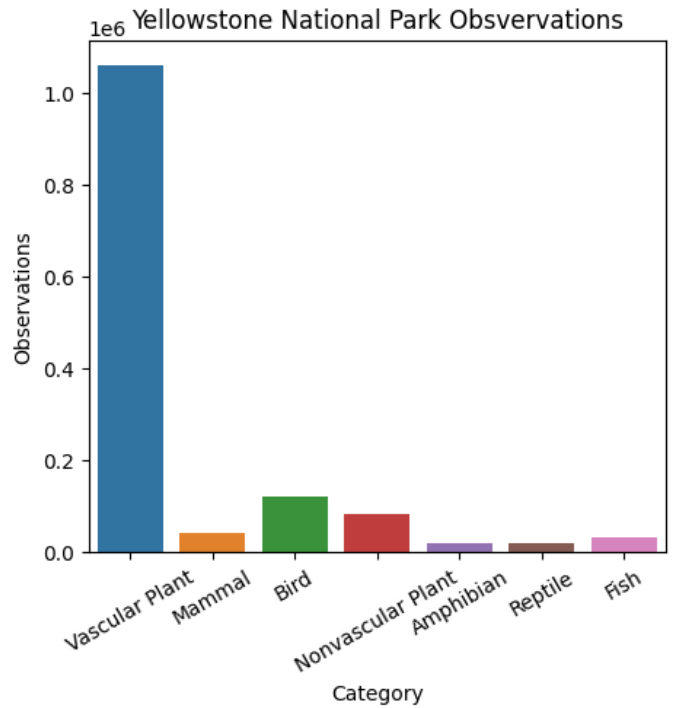


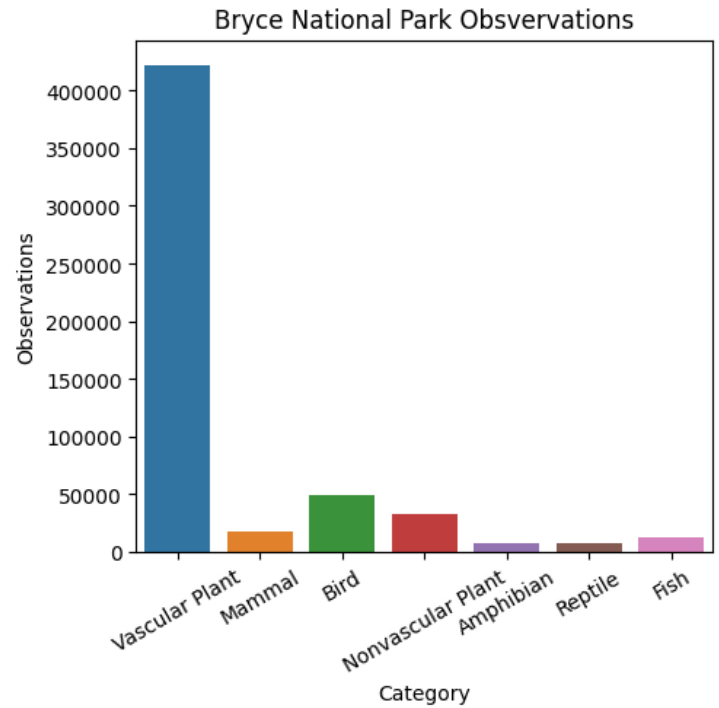
	Category	Amphibian	Bird	Fish	Mammal	Nonvascular Plant	Reptile	Vascular Plant	Total Obsv's
Park Name									
Bryce National Park		7,299	48,383	12,223	16,823	32,992	7,854	422,585	548,159
Great Smoky Mountains National Park		5,622	35,290	9,068	12,301	24,857	5,616	318,071	410,825
Yellowstone National Park		19,191	119,219	30,131	41,905	83,021	19,300	1,060,769	1,373,536
Yosemite National Park		11,309	71,290	18,353	24,887	49,783	11,335	634,515	821,472

	Category	Amphibian	Bird	Fish	Mammal	Nonvascular Plant	Reptile	Vascular Plant
Park Name								
Bryce National Park		1.33%	8.83%	2.23%	3.07%	6.02%	1.43%	77.09%
Great Smoky Mountains National Park		1.37%	8.59%	2.21%	2.99%	6.05%	1.37%	77.42%
Yellowstone National Park		1.4%	8.68%	2.19%	3.05%	6.04%	1.41%	77.23%
Yosemite National Park		1.38%	8.68%	2.23%	3.03%	6.06%	1.38%	77.24%

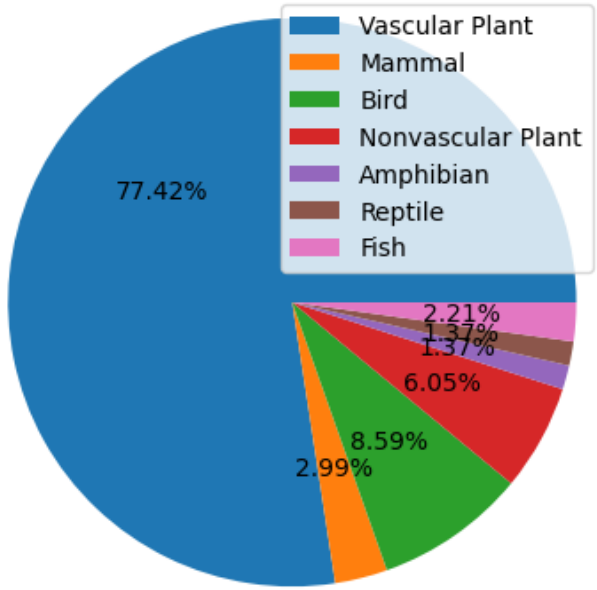

Great Smoky Mountains National Park Obsvervations
Observations
300000
250000
200000
150000
100000
50000
0
Vascular Plant
Mammal
Bird
Nonvascular Plant
Amphibian
Reptile
Fish
Category


Yosemite National Park Obsvervations
Observations
600000
500000
400000
300000
200000
100000
0
Vascular Plant
Mammal
Bird
Nonvascular Plant
Amphibian
Reptile
Fish
Category

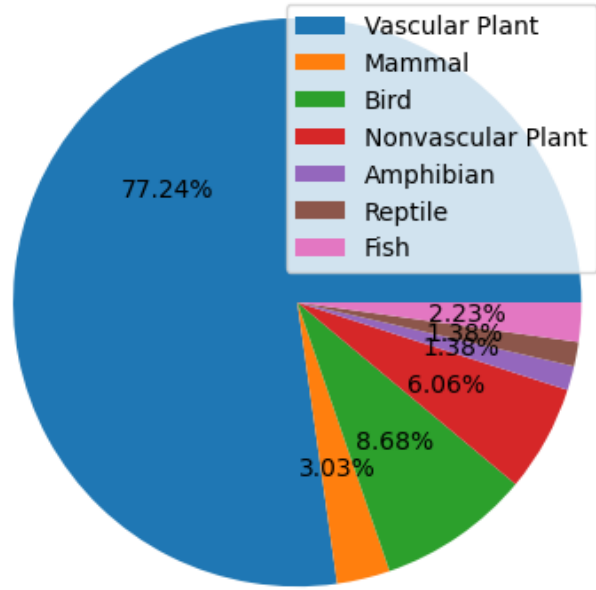

Yellowstone National Park Obsvervations
1e6
Observations
1.0
0.8
0.6
0.4
0.2
0.0
Vascular Plant
Mammal
Bird
Nonvascular Plant
Amphibian
Reptile
Fish
Category


Bryce National Park Obsvervations
Observations
400000
350000
300000
250000
200000
150000
100000
50000
0
Vascular Plant
Mammal
Bird
Nonvascular Plant
Amphibian
Reptile
Fish
Category

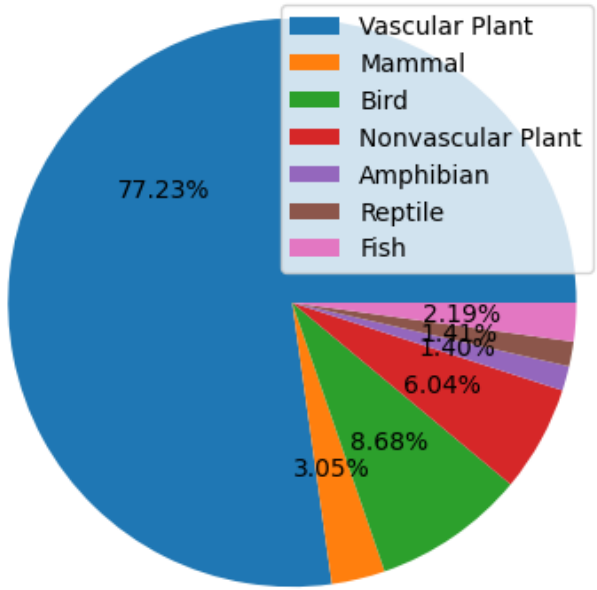
Great Smoky Mountains National Park Obsvervation %'s



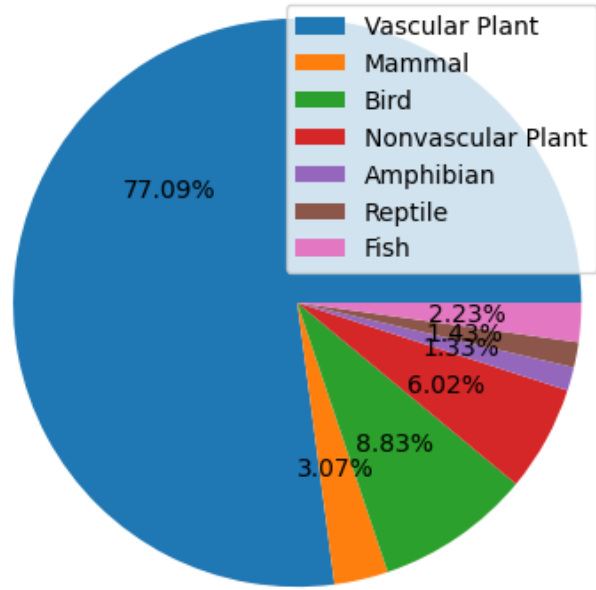
Yosemite National Park Obsvervation %'s



Yellowstone National Park Obsvervation %'s



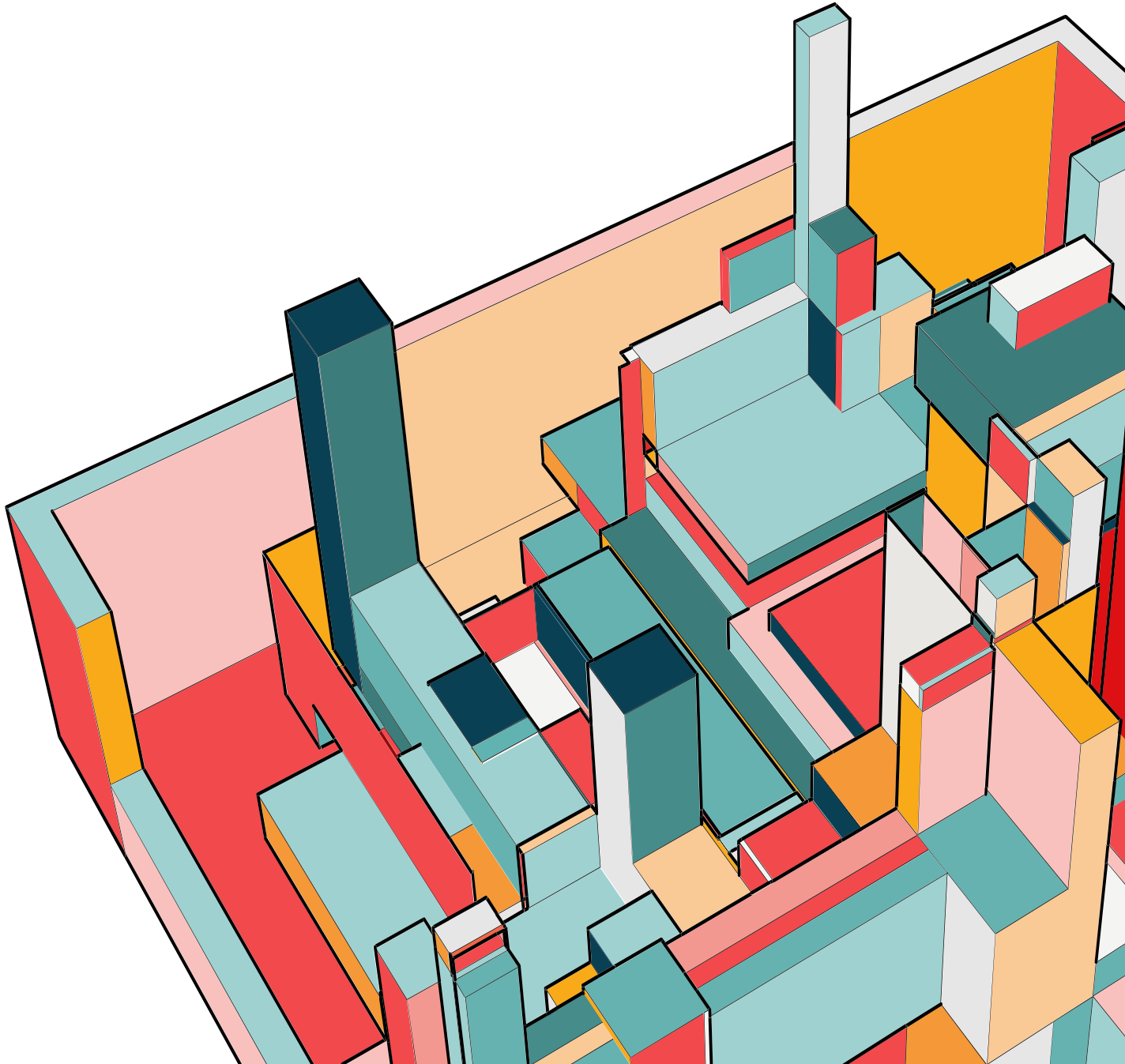
Bryce National Park Obsvervation %'s

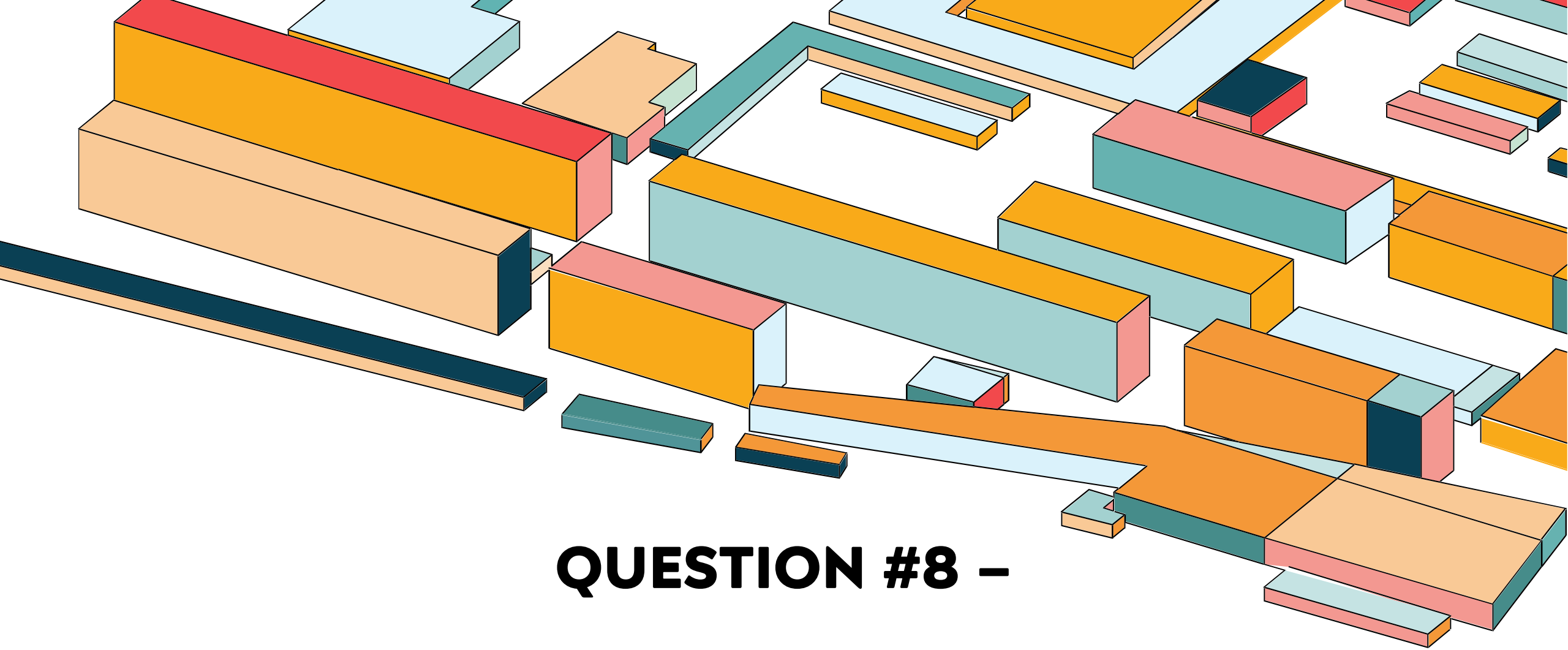


Question #7 Findings –

Just as you discovered earlier with the observations of conservation statuses per park, the percentage of observations of categories from park to park is almost identical.

For example, no matter which National Park you are in, Vascular Plants make up roughly 77% of the species seen within the park.





QUESTION #8 –

WHAT IS THE MOST ENDANGERED CATEGORY OF SPECIES AND IS THIS STATISTICALLY SIGNIFICANT?

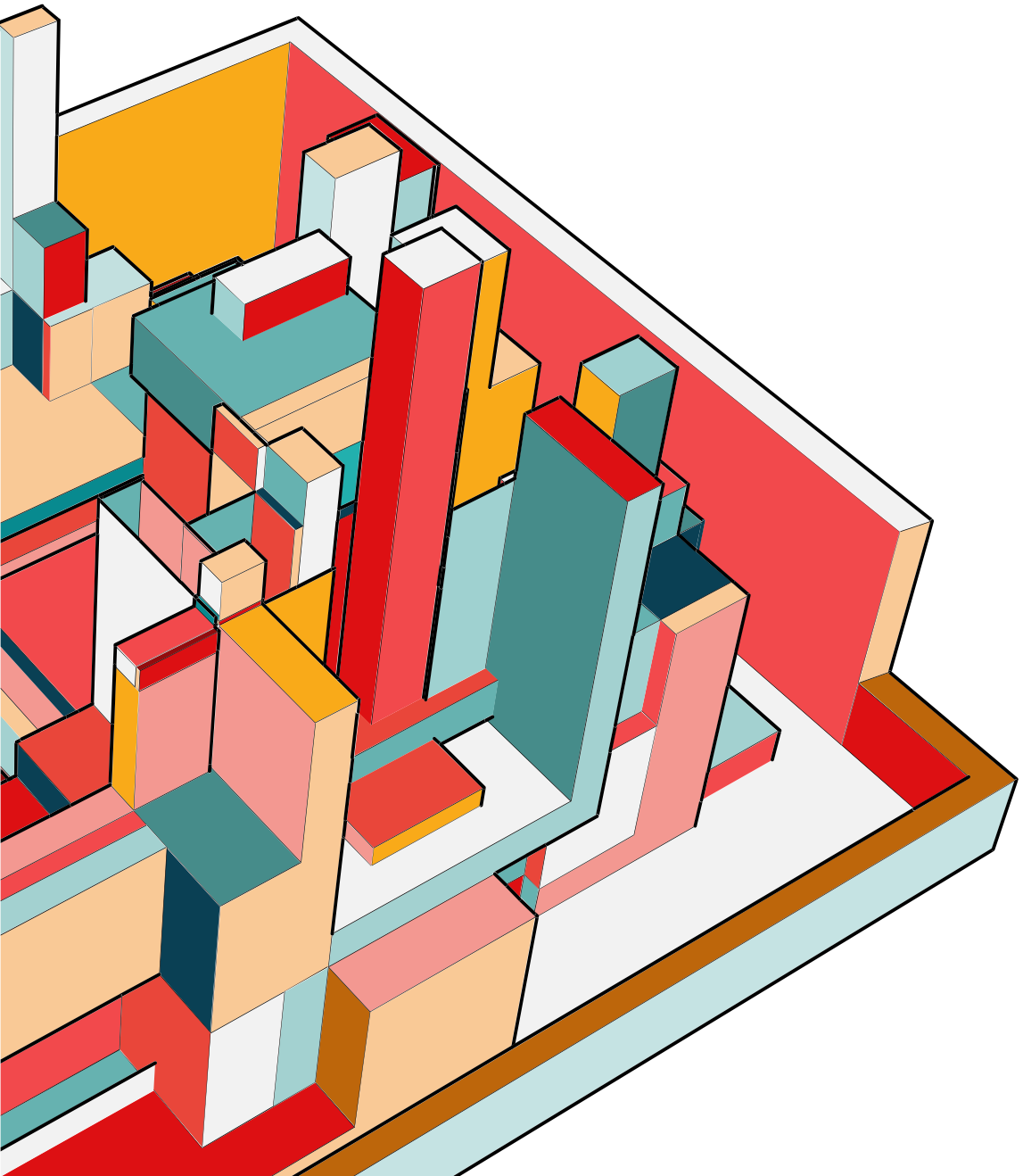
Referring to a table from earlier…

This table shows the percentage that each conservation status makes up within each category of species

	Conservation Status	Endangered	In Recovery	Not at Risk	Species of Concern	Threatened
Category						
Amphibian		1.27%	0.0%	91.14%	5.06%	2.53%
Bird		0.82%	0.61%	84.63%	13.93%	0.0%
Fish		2.4%	0.0%	92.0%	3.2%	2.4%
Mammal		3.41%	0.0%	82.95%	12.5%	1.14%
Nonvascular Plant		0.0%	0.0%	98.5%	1.5%	0.0%
Reptile		0.0%	0.0%	93.59%	6.41%	0.0%
Vascular Plant		0.02%	0.0%	98.92%	1.01%	0.05%

This table includes all the species (‘Not at Risk’ species as well) and the Mammal category has the highest percentage of ‘Endangered’ species.

For example, 3.41% of all Mammals are ‘Endangered’.



We know that Mammals have the highest rate of ‘Endangerment’, but to test the significance of this finding (the accuracy of the finding) we will need to do some statistical testing.

The goal is to find out if Mammals have higher rates of ‘Endangerment’ when compared to other categories all the time, or just in this circumstance.

Basically, it’s a test that tests the significance of our finding that Mammals have the highest rate of ‘Endangerment’.

For this test, we’ll be using a Chi-Squared test, which will compare another category to the Mammal category.

For the Chi-Squared test, we’ll need to know the total of species for each category, minus the number of ‘Endangered’ species

The test will be run 6 times, comparing Mammals to another category each time.

	Conservation Status	Endangered	In Recovery	Not at Risk	Species of Concern	Threatened	Total (Excluding 'Endangered')
Category							
Amphibian		1.0	0.0	72.0	4.0	2.0	78.0
Bird		4.0	3.0	413.0	68.0	0.0	484.0
Fish		3.0	0.0	115.0	4.0	3.0	122.0
Mammal		6.0	0.0	146.0	22.0	2.0	170.0
Nonvascular Plant		0.0	0.0	328.0	5.0	0.0	333.0
Reptile		0.0	0.0	73.0	5.0	0.0	78.0
Vascular Plant		1.0	0.0	4216.0	43.0	2.0	4261.0

The actual tests were conducted using Python and a Python library, so the tests won't be shown, but here is an example of what information the test would need to make a comparison.

For example, if you were comparing Mammals to Fish…

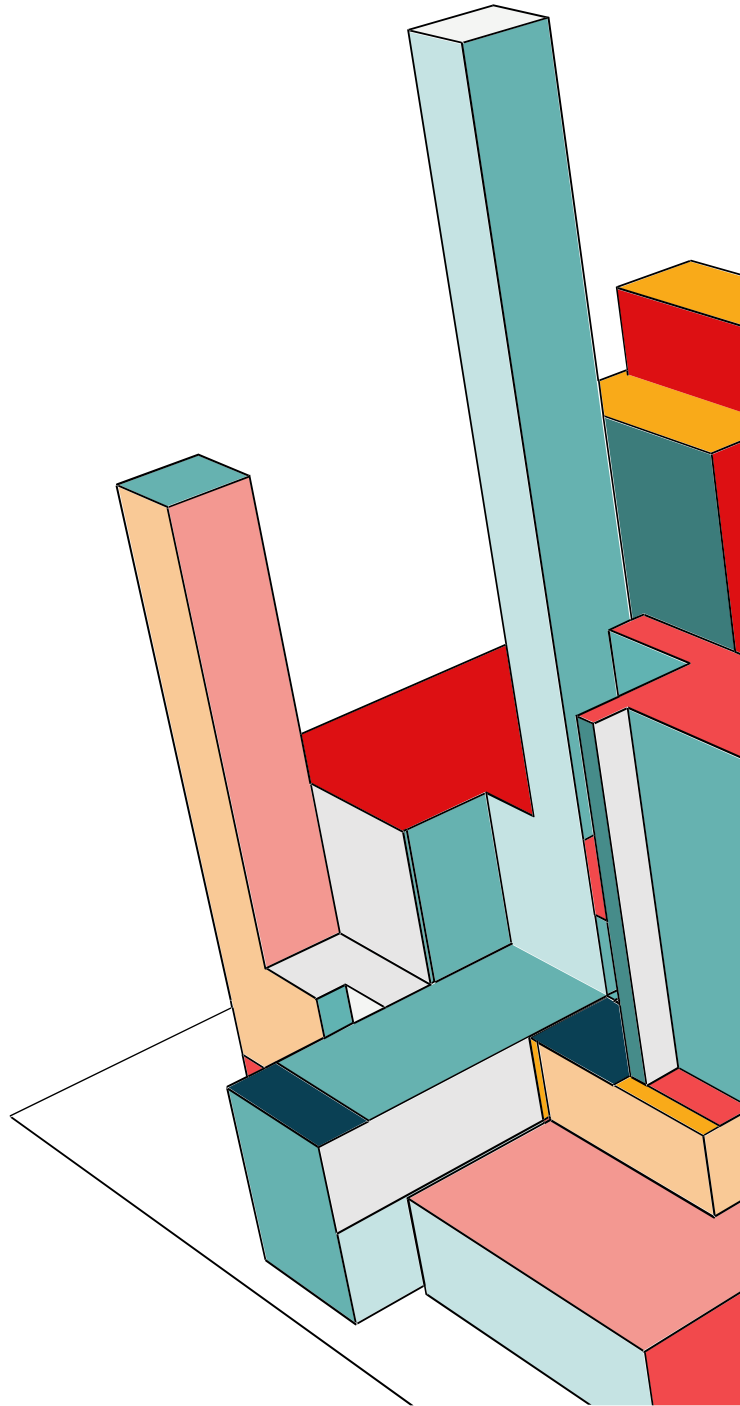
	Endangered	Total (Excluding 'Endangered')
Mammal	6.0	170.0
Fish	3.0	122.0

Before showing test results, we need to first explain what the results are showing.

The results will show a *P-value* for each category that Mammals get compared to. A *P-value* is essentially a probability value that describes the probability that an event will occur.

In our case, the event that is occurring is Mammals having a higher percentage of 'Endangered' species than another category of species.

The *P-value* is a confidence measurement, we need a *P-value* that is less than 0.05 (5%) to prove that the finding is statistically significant.



Chi-Squared Test Results –

Amphibian P-Value: 0.5795
The difference in the rate of endangerment between Mammals and Amphibian(s) is NOT statistically significant

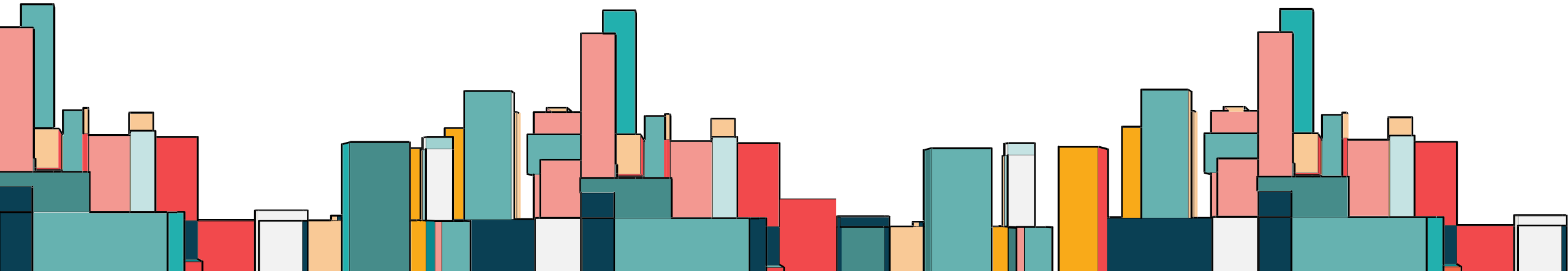
Bird P-Value: 0.0397
The difference in the rate of endangerment between Mammals and Bird(s) *is statistically significant*

Fish P-Value: 0.8704
The difference in the rate of endangerment between Mammals and Fish(s) is NOT statistically significant

Nonvascular Plant P-Value: 0.0031
The difference in the rate of endangerment between Mammals and Nonvascular Plant(s) *is statistically significant*

Reptile P-Value: 0.2292
The difference in the rate of endangerment between Mammals and Reptile(s) is NOT statistically significant

Vascular Plant P-Value: 0.0
The difference in the rate of endangerment between Mammals and Vascular Plant(s) *is statistically significant*



Question #8 Findings –

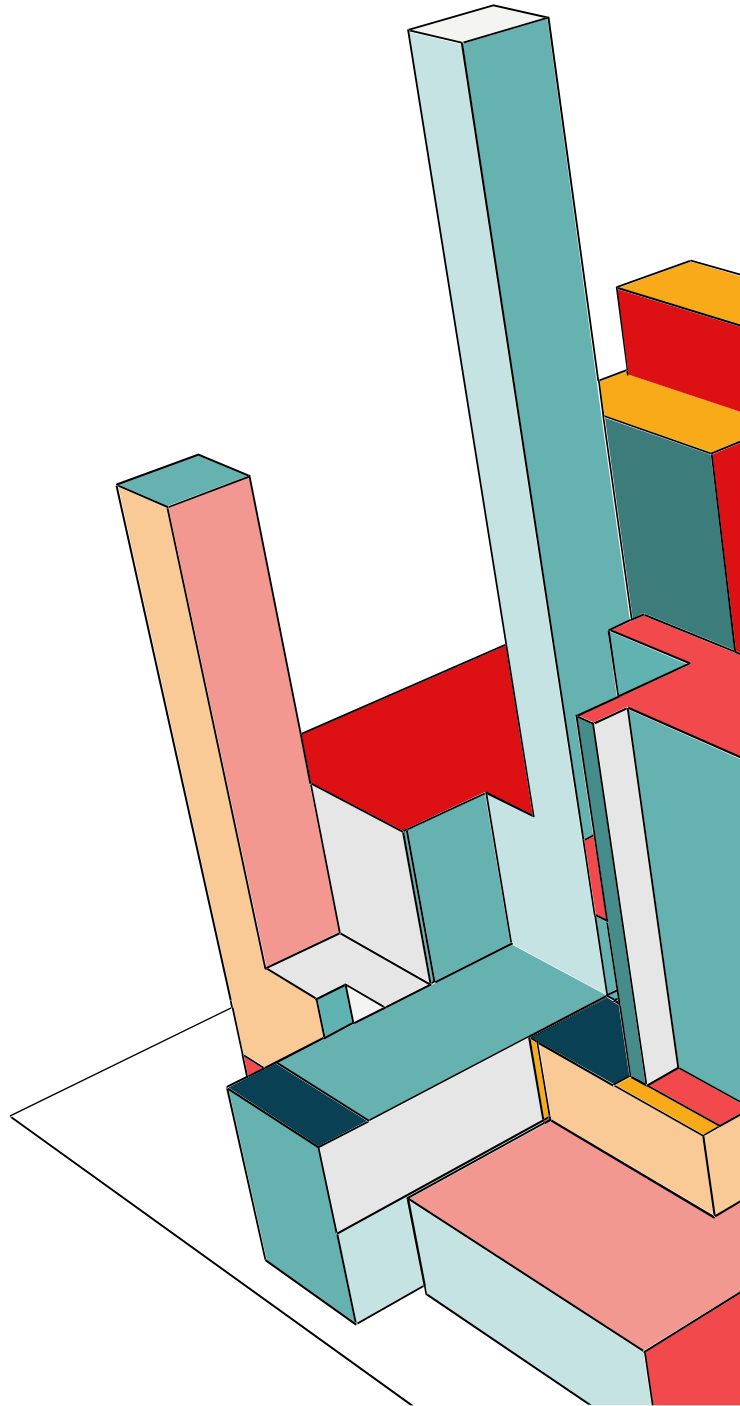
Mammals have statistically significant higher rates of endangerment when compared to Birds, Nonvascular Plants, and Vascular Plants (More than 95% of the time)

Being able to explore data come to relevant conclusions is crucial for discovering powerful solutions to problems.

In this instance, we understand that Mammals require more attention than Birds, Nonvascular Plants, and Vascular Plants when it comes to endangerment prevention.

This doesn't mean that all the other categories don't need attention as well, just that Mammals most likely need more attention than Birds, Nonvascular Plants, and Vascular Plants, specifically.

For more information on the other categories, we would need to run separate tests for them as well.



Questions/Findings Overview –

Question 1 Finding – At first glance, it looks like Yellowstone National Park has significantly more species who are 'Not at Risk' compared to the other parks.

Question 2 Finding – Yellowstone National Park has more sightings of 'Not at Risk' species, not because it is a safer environment for wildlife, but because it has a larger total wildlife population.

Question 3 Finding – Every National Park has a near equal percentage of sightings for each conservation status.

Question 4 Findings – 1. On average, 'Endangered' species are observed less than any other conservation status.
2. On average, as the severity of the conservation status increases for a specific species, the sightings of that specific species decreases.

Question 5 Findings – When comparing all conservation statuses (Includes 'Not at Risk')…
1. Mammals are the most 'At Risk' category of species overall.
2. Both Plant categories are the least 'At Risk' categories of species overall.
3. Mammals and Fish are the most 'Endangered' categories with mammals being the most 'Endangered'.
4. Birds and Mammals are the categories of species with the most concern with birds having the highest concern.
5. Reptiles have no 'Endangered' species but they do have the third highest 'Species of Concern' status.
6. Nonvascular Plants have no 'Endangered' species.

Question 6 Findings – When comparing 'At Risk' conservation statuses (Excludes 'Not at Risk')…
1. Fish have the highest proportion of 'Endangered' species (30%) with Mammals being the second highest (20%).
2. If preventative actions are taken, Birds, Reptiles, Vascular Plants and Nonvascular Plants have a much lower chance of moving towards 'Endangerment'.
3. Species that live primarily in the water (Amphibians and Fish) have high proportions of 'Endangerment' (Amphibians are 3rd highest behind Mammals) and are more 'Threatened' than any other category.

Question 7 Finding – Just as you discovered earlier with the observations of conservation statuses per park, the percentage of observations of categories from park to park is almost identical.

Question 8 Finding - Mammals have statistically significant higher rates of endangerment when compared to Birds, Nonvascular Plants, and Vascular Plants (More than 95% of the time).

THANK YOU

Benjamin Sandmann

Bsandma1@gmail.com